

**TÜRK HAVA KURUMU ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**GRNN VE MLP METOTLARI KULLANILARAK GERİ DÖNEN MEME
KANSER TESBİTİ**

YÜKSEK LİSANS TEZİ

Omar Taha Ahmed ALKALİDİ

Elektrik ve Bilgisayar Anabilim Dalı

Elektrik ve Bilgisayar Mühendisliği Programı

KASIM 2017

**TÜRK HAVA KURUMU ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**GRNN VE MLP METOTLARI KULLANILARAK GERİ DÖNEN MEME
KANSER TESBİTİ**

YÜKSEK LİSANS TEZİ

Omar Taha Ahmed ALKALİDİ

1406010010

Elektrik ve Bilgisayar Anabilim Dalı

Elektrik ve Bilgisayar Mühendisliği Programı

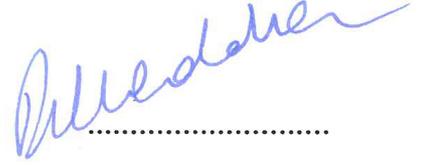
Tez Danışmanı: Yrd. Doç. Dr. Meltem İMAMOĞLU

Türk Hava Kurumu Üniversitesi Fen Bilimler, Enstitüsü'nün 1406010010 numaralı Yüksek Lisans öğrencisi, "Omar Taha Alkalidi", ilgili yönetmeliklerin belirlediği gerekli tüm şartları yerine getirdikten sonra hazırladığı "GRNN ve MLP Metotları Kullanılarak Geri Dönen Meme Kanseri Tesbiti" başlıklı tezini, aşağıda imzaları olan jüri önünde başarı ile sunmuştur.

Tez Danışmanı : Yrd. Doç. Dr. Meltem YILDIRIM İMAMOĞLU
Türk Hava Kurumu Üniversitesi



Jüri Üyeleri : Doç. Dr. Fırat HARDALAC
Gazi Üniversitesi



: Yrd. Doç. Dr. Javad RAHEBİ
Türk Hava Kurumu Üniversitesi



: Yrd. Doç. Dr. Meltem YILDIRIM İMAMOĞLU
Türk Hava Kurumu Üniversitesi



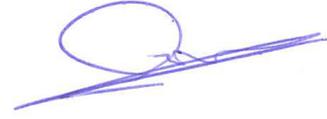
Tez Savunma Tarihi: 27 Kasım 2017

**TÜRK HAVA KURUMU ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜ'NE**

Yüksek Lisans Tezi olarak sunduğum, “GRNN ve MLP Metotları Kullanılarak Geri Dönen Meme Kanseri Tesbiti” adlı çalışmamın, tarafımdan akademik etik ve kurallara aykırı düşecek bir yardıma başvurmaksızın yazıldığını ve yararlandığım kaynakların kaynakçada gösterilenlerden oluştuğunu, bunlara atıf yapılarak yararlanılmış olduğunu belirtir ve bunu onurumla doğrularım.

14.11.2017

Omar Taha Ahmed ALKALİDİ



TEŐEKKÜR

Yüksek Lisans tez çalışma sürecinde beni yönlendiren, karşılaştığım zorlukları bilgi ve tecrübesi ile aşmamda yardımcı olan tez danışmanım değerli Yrd. Doç. Dr. Meltem İmamođlu'na teşekkürlerimi sunarım.

Kasım 2017

Omar Taha Alkalidi



İÇİNDEKİLER

TEŞEKKÜR.....	iv
İÇİNDEKİLER.....	v
TABLO LİSTESİ.....	vii
ŞEKİL LİSTESİ.....	viii
KISALTMALAR.....	ix
ÖZET.....	x
ABSTRACT.....	xi
BİRİNCİ BÖLÜM	1
1. GİRİŞ	1
1.1 Neden Meme Kanseri.....	1
1.2 MR Görüntüleme Sistemi.....	4
1.3 Araştırmanın Motivasyon Amacı.....	5
1.4 Meme Kanseri Riskini Arttıran Faktörler [4].....	6
1.5 Problem Durumu.....	6
1.6 Meme Kanseri Riskini Azaltan Faktörler.....	7
1.7 Tanı.....	7
1.8 Evreleri ve Tedavi.....	8
1.9 Problem Cümlesi.....	9
1.10 Araştırmanın Amacı ve Önemi.....	9
1.11 Kavramsal Çerçeve.....	10
1.12 Alt Problemler.....	10
1.13 Kısıtlılıklar.....	10
1.14 Raporun Ana Hatları.....	11
İKİNCİ BÖLÜM	12
2. TEMEL KAVRAMLAR VE LİTERATÜR ÖZETİ	12
2.1 Temel Kavramlar.....	12
2.2 Doku Yapısında Meme Kanserinin Yaygın Türleri.....	14
2.2.1 Yayılmacı Olmayan.....	14
2.2.2 Yayılmacı.....	15
ÜÇÜNCÜ BÖLÜM	25
3. YÖNTEM	25
3.1 Wisconsin Prognostic Breast Cancer (WPBC).....	25
3.2 Biyolojik Sinir Hücresi.....	27
3.2.1 Yapay Sinir Hücresi.....	28
3.2.2 Tek Katmanlı Algılayıcılar (TKA) [35].....	30
3.2.3 Çok Katmanlı Algılayıcılar (ÇKA) [35].....	30
3.2.4 Geriye Yayılım Algoritması.....	32
DÖRDÜNCÜ BÖLÜM	35
4 BULGULAR VE TARTIŞMA	35
4.1 Normalleştirme Yöntemi.....	35
4.2 Performans Ölçümü.....	36

4.3	Sınıflandırıcı Yöntemi 194 Kişi İçin	36
4.3.1	Genelleştirilmiş Regresyon Sinir Ağı.....	36
4.3.2	Destek Vektör Makinesi	39
4.3.3	Çok Katlı Perceptron	41
4.4	Sınıflandırıcı Yöntemi 569 Kişi İçin	46
4.4.1	SVM Sonuçları	46
4.4.2	GRNN Sonuçları	47
4.4.3	MLP Algoritması.....	48
BEŞİNCİ BÖLÜM		50
5. SONUÇLAR VE ÖNERİLER.....		50
5.1	Sonuç	50
5.2	İleri Çalışmalar	52
KAYNAKÇA		53
ÖZGEÇMİŞ.....		57



TABLO LİSTESİ

Tablo 1.1 : Türüne göre Türkiye’de kanser sıralaması.....	2
Tablo 1.2 : Türkiye’de en yüksek ölüm nedenleri.....	3
Tablo 1.3 : Türüne göre Irak’ta kanser sıralaması.....	3
Tablo 1.4 : Irak’ta en yüksek ölüm nedenleri.....	4
Tablo 2.1 : Meme kanser tesbiti hakkında çalışmalar	23
Tablo 3.1 : Katıştırma matrisi ve doğruluk değeri.....	27
Tablo 4.1 : Confüzyon matrisi.....	36
Tablo 4.2 : SVM, GRNN ve MLP karşılaştırılması.....	48
Tablo 5.1 : Bizim yöntemimiz ve diğer çalışmaların karşılaştırılması.....	51

ŞEKİL LİSTESİ

Şekil 1.1	: Meme kanseri için MRI görüntüleme sistemi.....	5
Şekil 1.2	: Bilgisayar tabanlı tanı sistemi işlem basamakları.....	9
Şekil 2.1	: Östrojen alıcı boyamasını gösterir örnek bir sayısal görüntü.....	13
Şekil 2.2	: Meme kanserinde tümör.....	13
Şekil 2.3	: HE boyamasında normal bir meme dokusunun görünümü.....	14
Şekil 2.4	: [21]'deki önerilen yöntem.....	22
Şekil 3.1	: Biyolojik sinir hücresi.....	28
Şekil 3.2	: Basit algılayıcı modeli.....	29
Şekil 3.3	: Tek katmanlı algılayıcı.....	30
Şekil 3.4	: Çok katmanlı algılayıcı.....	31
Şekil 3.5	: Sinyal akış gösterimi.....	31
Şekil 3.6	: Geri yayılım algoritması.....	33
Şekil 4.1	: Genelleştirilmiş regresyon sinir ağı mimarisi.....	37
Şekil 4.2	: Doğruluk yüzdesi çalışma zamanına göre.....	37
Şekil 4.3	: GRNN için karışıklık matrisi, a) en az doğruluk 16'ncı çalışmada, b) en yüksek değer 5. Çalışmada, c) test verileri sonucu düşük değer, d) test verileri sonucu yüksek değer.....	38
Şekil 4.4	: Önerilen generalized regression neural network.....	39
Şekil 4.5	: Verilerin sınıflandırılması için SVM.....	39
Şekil 4.6	: Ortalama 75.29 elde edildi.....	40
Şekil 4.7	: SVM için karışıklık matrisi, a) en az doğruluk 10'ncü çalışmada, b) en yüksek değer 3. Çalışmada, c) test verileri sonucu düşük değer, d) test verileri sonucu yüksek değer.....	41
Şekil 4.8	: YSA'nın yapısı.....	42
Şekil 4.9	: kullanıldığı YSA'ın modeli.....	43
Şekil 4.10	: Eğitim hatanın performansı.....	44
Şekil 4.11	: YSA için karışıklık matrisi, a) doğruluk oranı 56.9, b) test verileri sonucu.....	45
Şekil 4.12	: SVM için karışıklık matrisi, a) en az doğruluk 12'ncü çalışmada, b) en yüksek değer 4. Çalışmada, c) test verileri sonucu düşük değer, d) test verileri sonucu yüksek değer.....	46
Şekil 4.13	: GRNN için karışıklık matrisi, a) en az doğruluk 8'ncü çalışmada, b) en yüksek değer 1. Çalışmada, c) test verileri sonucu düşük değer, d) test verileri sonucu yüksek değe.	47
Şekil 4.14	: YSA için karışıklık matrisi, a) doğruluk oranı 95.139, b) test verileri sonucu.....	48
Şekil 4.15	: Önerilen MLP'nin mimarisi.....	49

KISALTMALAR

WPBC	: Wisconsin Prognostic Breast Cancer
YSA	: Yapay Sinir Ağları
HE	: Hemotoksin Eosin
YSKK	: Yayılmacı Süt Kanalı Kanseri
YSBK	: Yayılmacı Süt Bezi Kanseri
KSA	: Kıvrımlı sinirsel ağlardan
DVM	: Destek vektör makinesi
DKSA	: Derin Kıvrımsal Sinir Ağları
YİM	: Yerel İkili Modeller
YGH	: Yönlendirilmiş gradyanların histogramı
GFH	: Gradyan farklılıklarının histogramı
GOOESA	: Genetik Olarak Optimize Edilmiş Sinir Ağı
BDA	: Bilgisayar destekli algılama
TKA	: Tek katmanlı algılayıcılar
ÇKA	: Çok katmanlı algılayıcılar
GRNN	: Genelleştirilmiş regresyon sinir ağı
MLP	: Multilayer perceptron
MRI	: Support vector machines
WHO	: World health organization
RBF	: Radial basis function
SSL	: Secure sockets layer
PCR	: Polimeraze chain reaction
RNA	: Ribonükleik asit
DNA	: Deoksiribo nükleik asit
BDA	: Bilgisayar destekli algılama
ANFIS	: Adaptive neuro fuzzy inference sistem
TSLP	: Timik stromal lenfopoietin
TP	: Gerçek Pozitif
FN	: Hatalı Negatif
FP	: Hatalı Pozitif
TN	: Gerçek Negatif
TSLP	: Timik stromal lenfopoietin

ÖZET

GRNN VE MLP METOTLARI KULLANILARAK GERİ DÖNEN MEME KANSER TESBİTİ

Alkalidi, Omar Taha

Yüksek Lisans, Elektrik ve Bilgisayar Mühendisliği Bölümü

Tez Danışman: Yrd. Doç. Dr. Meltem İmamoğlu

Kasım 2017, 57 sayfa

Bu tezde, meme kanseri tahmini için Yapay Sinir Ağı ve Destek Vektör Makinesi ile GRNN uyguladık. Yapay Sinir Ağı için çok katmanlı perseptron metodunu kullandık ve her katman için farklı farklı neron sayıları deneyerek sonuçları elde ettik. Sonuçlarımızı SVM ile ve GRNN'le karşılaştırdık. Bu tezde, Wisconsin meme kanseri veri tabanını kullandık. Bu veri tabanında, gerçek klinik verileri içeren çok sayıda veri var. Önceki çalışmalarda yazarlar sinir ağı üzerinde çalıştı, ancak biz çalışmalarımızda sinir ağının hangi modelinin bize meme kanseri tespitinin daha yüksek doğruluğu verdiğini karşılaştıracacağız. Sonuçlarda GRNN yönteminin daha doğru sonuç verdiğini tespit edildi. Bu çalışmada 194 ve 569 örnek içeren Wisconsin Prognostik Meme Kanseri (WPBC) veri tabanı kullanılmıştır. Bu veri tabanını seçmenin nedeni; meme kanseri ile ilgili gerçek verileri içerdiği gerçeğidir. Bu tez çalışmasında, Çok Katmanlı Perceptron (MLP) ve Genelleştirilmiş Regresyon Sinir Ağı (GRNN) kullanılarak meme kanseri tahmini gerçekleştirilmiştir. Yapay sinir ağlarının sonuçları da, Destek Vektör Makinesi (SVM) ile elde edilen sonuçlar ile karşılaştırılmıştır. En iyi performans GRNN yöntemi kullanıldığında elde edilir. Sonuçlar ve gelecekteki çalışmalar tartışılmıştır.

Anahtar Kelimeler: GRNN, YSA, Destek vektör makinesi.

ABSTRACT

RECURRENCE BREAST CANCER DETECTION USING GENERAL REGRESSION NEURAL NETWORK AND MULTI LAYER PERCEPTRON

Omar Taha Alkalidi

Master, Department of Electrical and Computer Engineering

Thesis supervisor: Asst. Prof. Dr. Meltem İmamoğlu

November 2017, 57 pages

In this thesis study, we applied GRNN with an Artificial Neural Network and Support Vector Machine. We used multilayer perceptron method for the Artificial Neural Network and obtained the results by trying nerve cells in different numbers for individual layers. We compared our results with SVM and GRNN. In this thesis study, we made use of the Wisconsin breast cancer database. This database contains a substantial amount of data with actual clinical data. In the former studies, the authors researched the neural network; however, we will compare in our studies the neural network models to find out which model gives us the breast cancer diagnosis with more accuracy. The results suggest that GRNN method gives more accurate outcomes. In this study, we used Wisconsin Prognostic Breast Cancer (WPBC) database that contains 194 and 569 samples. The reason for selecting this database is that it contains actual breast cancer data. In this thesis study, we made a breast cancer prediction by using Multilayer Perceptron (MLP) and General Regression Neural Network (GRNN). Also, the outcomes of artificial neural networks were compared with the results obtained by means of the Support Vector Machine (SVM). The best performance is achieved by using the GRNN method. The results are future researches are discussed.

Keywords: GRNN, ANN, Support vector machine

BİRİNCİ BÖLÜM

GİRİŞ

Bu bölümde meme kanseri hakkında genel bilgiler anlatılmıştır ve Türkiye’de kanser sıralaması ve en yüksek ölüm nedenleri gösterilmiştir. Sonar Irak’ta en yüksek ölüm nedenleri ve Meme Kanseri Riskini Arttıran Faktörleri anlatılmıştır. En sonda tezin amacı ve alt problemler tartışılmıştır.

1.1 Neden Meme Kanseri

Meme kanseri meme hücrelerinde başlayan kanser türüdür. Akciğer kanserinden sonra, dünyada görülme sıklığı en yüksek olan kanser türüdür. Her 8 kadından birinin hayatının belirli bir zamanında meme kanserine yakalanacağı bildirilmektedir [1]. Erkeklerde de görülmeyle beraber, kadın vakaları erkek vakalarından 100 kat fazladır. 1970’lerden bu yana meme kanserinin görülme sıklığında artış yaşanmaktadır ve bu artışa modern, batılı yaşam tarzı sebep olarak gösterilmektedir. Kuzey Amerika ve Avrupa ülkelerinde görülme sıklığı, dünyanın diğer bölgelerinde görülme sıklığından daha fazladır. Meme kanseri, yayılmadan önce, erken tespit edilirse, hasta % 96 yaşam şansına sahiptir. Her yıl 44000’de bir kadın meme kanserinden ölmektedir. Meme kanserine karşı en iyi koruyucu yöntem erken teşhistir. Meme kanserinin birçok tipi vardır. En sık rastlanan duktal karsinoma, memenin süt kanallarında başlar. Meme kanseri memenin dışına yayıldığında koltuk altındaki lenfatik nodüller en sık görülen yayılım yerleridir. Kanser hücreleri memenin diğer Lenf Nodlarına, Kemiğe, Karaciğere ve Akciğere yayılabilir. Her kadın meme kanseri gelişme riskine sahiptir. Gerçekte meme kanseri gelişen kadınların çoğunda risk faktörleri belli değildir.

Türkiye’de ve dünyada, insanlar sınırlı sayıdaki tıbbi kaynaklardan ve medikal malzemelerden yararlanmak için uzun süre beklemekten acı çekiyorlar. Dünya Sağlık Örgütü’ne (WHO) göre, Türkiye sağlık sistemi alanında 190 ülkeden 70’inci olarak sıralanmıştır.

Türkiye'nin artan nüfusu, yaşlanan nüfus, modern yaşam tarzı, iklim değişikliği ve ortaya çıkan yeni hastalıklar. Türkiye sağlık kuruluşlarını ve eyalet hükümetlerini yeni prosedürler koymak ve planlamalara yöneltti. Bu kuruluşlar mevcut tıbbi kaynaklar, altyapı ve sağlık personelindeki eksikliklere rağmen halka iyi bir sağlık hizmeti sunmaktır. Buna ek olarak, tıbbi hizmetler tüm bireyler için vazgeçilmezdir ve tüm sakinler ve vatandaşlar için tıbbi altyapı ve hizmetleri geliştirmek ve korumak ulusun sorumluluğundadır. Tıbbi personeldeki ve teknolojiye ek olarak, reçetelendirme hatası olayları, hastalar için küçük sorunları daha büyük sorunlar haline getirmektedir.

Dünya Sağlık Örgütü'ne (WHO) göre, Meme kanseri şu anda dünyadaki tüm kadınlar arasında en yüksek kanser türüdür ve Meme kanseri kadınlarda ikinci en yüksek ölüm nedenidir. Türkiye'de Meme kanseri, kanser türleri arasında üçüncü yaygın kanser türüdür ve genel olarak dokuzuncu ölüm nedeni olarak bilinir (Bakınız Tablo 1.1, 1.2).

Tablo 1.1: Türüne göre Türkiye'de kanser sıralaması.

Türüne göre Türkiye'de kanser sıralaması					
Sıra	Kanser	Dünya Sıra.	Sıra	Kanser	Dünya Sıra.
1	Akciğer Kanseri	17	9	Pancreas Cancer	66
2	Prostat kanseri	38	10	Yumurtalık Kanseri	90
3	Meme kanseri	113	11	Yemek borusu kanseri	69
4	Mide kanseri	32	12	Karaciğer kanseri	149
5	Kolon Rectum Kanserleri	66	13	Rahim kanseri	63
6	Lenfomalar	14	14	Rahim ağzı kanseri	159
7	Mesane kanseri	4	15	Ağız kanseri	137
8	Lösemi	45	16	Cilt Kanseri	116

Tablo 1.2: Türkiye’de en yüksek ölüm nedenleri.

Türkiye’de en yüksek ölüm nedenleri					
Sıra	Ölüm nedeni	Dünya Sıra.	Sıra	Ölüm nedeni	Dünya Sıra.
1	Koroner kalp hastalığı	37	13	Astım	39
2	inme	82	14	Kolon Rectum Kanseri	66
3	Akciğer Kanseri	17	15	Doğuştan anomaliler	67
4	Akciğer hastalığı	42	16	Diğer Yaralanmalar	107
5	Prostat kanseri	38	17	İntihar	94
6	Romatizmal kalp rahatsızlığı	4	18	Lenfomalar	14
7	Hipertansiyon	93	19	Menenjit	47
8	Grip ve pnömoni	139	20	Mesane kanseri	4
9	Meme kanseri	113	21	Düşük doğum ağırlığı	99
10	Şeker hastalığı	132	22	Enflamatuvar / Kalp	124
11	Mide kanseri	32	23	Lösemi	45
12	Karayolu Trafik Kazaları	118	24	Karaciğer hastalığı	158

Irak’ta Meme kanseri, kanser türleri arasında birinci yaygın kanser türüdür ve genel olarak dokuzuncu ölüm nedeni olarak bilinir (Bakınız Tablo 1.3, 1.4).

Tablo 1.3: Türüne göre Irak’ta kanser sıralaması.

Türüne göre Irak’ta kanser sıralaması					
Sıra	Kanser	Dünya Sıra.	Sıra	Kanser	Dünya Sıra.
1	Meme kanseri	44	9	Karaciğer kanseri	116
2	Akciğer Kanseri	81	10	Yumurtalık Kanseri	105
3	Lenfomalar	13	11	Pankreas Kanseri	96
4	Mesane kanseri	3	12	Ağız kanseri	103
5	Lösemi	1	13	Rahim ağzı kanseri	162
6	Prostat kanseri	144	14	Yemek borusu kanseri	138
7	Kolon Rectum Kanseri	116	15	Cilt Kanseri	143
8	Mide kanseri	96	16	Rahim kanseri	164

Tablo 1.4: Irak'ta en yüksek ölüm nedenleri.

Irak'ta en yüksek ölüm nedenleri					
Sıra	Ölüm nedeni	Dünya Sıra.	Sıra	Ölüm nedeni	Dünya Sıra.
1	Koroner kalp hastalığı	22	13	Düşük doğum ağırlığı	58
2	inme	57	14	Doğuştan anomaliler	38
3	Karayolu Trafik Kazaları	2	15	Romatizmal kalp rahatsızlığı	17
4	Şeker hastalığı	47	16	Şiddet	52
5	Savaş	4	17	Enflamatuvar / Kalp	61
6	Böbrek hastalığı	5	18	Doğum Travması	60
7	Diğer Yaralanmalar	13	19	Astım	55
8	Hipertansiyon	28	20	Diğer Neoplaziler	4
9	Grip ve pnömoni	101	21	Lenfomalar	13
10	Meme kanseri	44	22	Mesane kanseri	3
11	Akciğer hastalığı	114	23	Endokrin Bozukluklar	73
12	Akciğer Kanseri	81	24	Lösemi	1

Etkili ve koruyucu bir önlem yaygınlaşmaya kadar, erken tanı ve ardından etkili bir tedavi, Meme kanserinden ölüm sonucunu azaltmanın tek yoludur. Meme kanseri, hasta tarafından Memete yumru kitleleri olarak algılanır. Bu kitleler genel olarak iyi huyludur (kanser değil), bu nedenle Meme kanserini teşhis etmek doktorun sorumluluğundadır. Teşhisin amacı malign (kanserli) ya da benign Meme kitlesini ayırt etmektir. Meme kanseri, hasta üzerinde yapılan çeşitli testlerin sonuçlarını analiz ederek hastalık teşhisi gerçekleştirilir. Testler yapıldıktan sonra, nihai teşhis tıbbi bir uzman için bile zor olabilir. Son yıllarda, doktorlar test sonuçlarının verilerini paylaşmaya başladılar ve bu paylaşım bilgisayarlı tanısal sistemlerin geliştirmesine yol açtı.

1.2 MR Görüntüleme Sistemi

Mastektomi uygulamaları sonrasında alınan memenin yerine, plastik cerrahi teknikler ile meme rekonstrüksiyonu yapılması ameliyatları söz konusu olabilmektedir. Günümüzde bazı deneysel tarama yöntemleri incelenmektedir. Bunlar; MR, duktal lavaj, ultrason, optik tomografi, PET taraması ve dijital mamografidir. Manyetik rezonans görüntüleme; birçok hastanın klinik tedavisinde önemli bir konudur. Tıbbi bilgi ayrıntılı manyetik rezonans, bilgisayarlı tomografi

(X-ışını CT), EİTÖ, mamografi gibi diğer görüntüleme yöntemleri. diğer yöntemlere göre daha fazla kullanılmaktadır. İki tür nükleer gevşeme süresi (T1 ve T2) ve hareket yoğunluğu; Manyetik rezonans tekniği kaynaklı olduğu için dört farklı bileşenden oluşmaktadır. Şekil 1.1 meme kanseri için MRI görüntüleme sisteminin örneklemini göstermektedir.



Şekil 1.1: Meme kanseri için MRI görüntüleme sistemi [2].

Meme MR çekimi için ideal dönem menstural siklusün 5. ve 12. günleri arasındır. Özellikle lüteal fazda yapılan MR çekimlerinde meme parankimi yoğun kontrast tutarak zemindeki olası kitle lezyonları gizlemekte ve yalancı negatifliğe neden olmaktadır. Oral kontraseptif veya hormon replasman terapisi kullanan olgularda meme MR çekimi ilaçların kesiminden 2-3 ay sonra yapılmalıdır. Hastalara damar yolu açılarak, MR cihazının bulunduğu odaya alınır ve meme MR için özel geliştirilmiş MR sargı- larına hasta pron pozisyonda yatırılmaktadır [3].

1.3 Araştırmanın Motivasyon Amacı

Bu araştırma motivasyonunun amacı, testlerden elde edilen sonuçları sınıflandırmak için WPBC veri tabanlı bir araç yaratmaktır. Bu çalışmada önerilen araç tamamen klinik destek sistemidir. Bu sistem teşhis için tıbbi veri analizini daha kısa sürede sağlamakta ve deneyimsizlik veya yorgunluk nedeniyle oluşan insan hatalarını engellemektedir. Yapay sinir ağlarının kullanımı, yöntemlerin çoğunun doğruluğunu artırır ve bu süreç içerisinde ihtiyaç duyulan uzman ihtiyacını azaltır.

1.4 Meme Kanseri Riskini Arttıran Faktörler [4]

- a) Kadın olmak
- b) 55 yaş ve üzerinde olunması
- c) Yakın akrabalarından biri meme kanseri ise, (anne veya kız kardeş meme kanseri ise, 2-3 kat daha fazla)
- d) Daha önceden diğer memede kanser tespit edilmiş olması
- e) Adet görmeye 12 yaşından önce başlamış olması
- f) Hiç gebe kalmamış olması
- g) Adet görmesi 50 yaşından sonra da devam ediyor olması
- h) Yoğun meme dokusuna sahip olmak
- i) Alkol ve sigara tüketmek
- j) Gece çalışmak
- k) Obez olmak
- l) Fiziksel aktiviteden yoksun olmak
- m) Deodorant kullanmak (Muhtemel)

1.5 Problem Durumu

Araştırmalar, meme hücreleri içerisinde, meme kanser riskini artıran bazı genler olduğunu göstermektedirler. Genetik değişiklikler, aileden (herediter) olabilir veya hayat boyu gelişebilirler. Meme kanseri genellikle tek bir hücrede başlar. Günümüzde meme kanserinin nedeni ve nasıl gelişim göstereceği tam olarak bilinmemektedir.

Meme kanseri kompleks bir hastalıktır. Her vaka birbirinin aynısı değildir. Meme kanserinin içinde bulunduğu evreye "stage" denir. Gerçek stage'in bilinmesi, doktorun tedavi planını yapmasını sağlayacaktır.

- a) İnsan, yaşamında meme kanserine sebep olacak herhangi bir yanlış yapmamış olsa da bu hastalığa yakalanabilir.
- b) Meme kanseri bulaşıcı değildir, başka bir hastadan bulaşmaz.
- c) Meme kanseri, stresle veya memeye travmayla (darbeyle) meydana gelmez.
- d) Meme kanseri gelişen çoğu kadının risk faktörü veya ailesinde hastalığa ait bir hikâye yoktur.

2015 yılında ABD'de Texas Üniversitesi Anderson Kanseri Merkezi'nde yapılan ve Kanseri Araştırma Dergisi'nde (Cancer Research) yayınlanan deneysel çalışmaya göre fruktoz ve sakaroz (glukoz ve fruktoz kombinasyonu) 12-lipoksijenaz (12-LOX) sinyalizasyonunu ve dolayısıyla da 12-hidroksi- 5Z, 8Z, 10E, 14z – eikosatetraenoik asit (12-HETE) seviyesini arttırarak meme kanseri riskini ve metastaz oranının yükselttiği tespit edilmiştir.

glukoz + fruktoz → 12-LOX ↑ → 12-HETE ↑ → meme kanseri ↑

1.6 Meme Kanseri Riskini Azaltan Faktörler

- a) Düzenli egzersiz ve fiziksel aktivite
- b) Emzirmek
- c) Yağlı besinleri kısıtlamak
- d) Alkol alımını kısıtlamak
- e) İdeal kiloyu korumak

1.7 Tanı

Meme kanserinde erken teşhis yöntemleri, hastanın taşıdığı risk faktörlerine göre değişkenlik gösterir. Bu faktörlerin arasında yaş ilk sırada gelmektedir. Genç yaşlarda görülebilmesine karşın, ileri yaş gruplarında bu risk artar. Bu nedenle ileri yaş gruplarında erken tanı konması için alınması gereken önlemler, erken yaş gruplarından daha farklıdır.

Yirmi yaş grubu, her ayın belirli bir döneminde kendi kendilerini muayene etmelidir. Bu kontrol sırasında meme dokusunda farklılık olup olmadığı araştırılır. Şişkinlik, yumru benzeri bir değişiklik saptanırsa derhal bir hekime baş vurulmalıdır. Bir değişiklik saptanmasa da, üç yılda bir kez hekim tarafından muayene edilmelidirler. Kırk yaş grubunun, kendi yaptıkları periyodik muayeneye ek olarak her yıl bir kez hekim tarafından muayene edilmeleri gereklidir. Ayrıca her yıl veya en az iki yıl arayla mamografi çektirmeleri gereklidir. Elli yaş grubu kadınlar, kendilerinin periyodik muayenelerine ve her yıl bir defa hekim muayenesine devam etmeli ve her yıl mamografi (meme filmi) çektirmelidir.

Kendi kendini kontrollerde onbeş günü aşkın sürede ele gelen sertlik veya kitle, deride kalınlaşma, şişme, renk değişikliği, meme başında kalınlaşma, kızarıklık veya

yara olması, memede veya meme başında içeri doğru çukurtu, meme şeklinde değişiklik, meme başlarının pozisyonlarında değişiklik ve meme başında akıntı gibi belirtiler derhal doktor kontrolü gerektirmektedir. Hekim muayenesi sonucu yapılacak mamografi taramasının ardından, ultrason, ince iğn aspirasyon biyopsisi ve normal biyopsi tetkikleriyle kesin tanı konulur.

1.8 Evreleri ve Tedavi

a) Evre 1: Tümör 20 mm. veya daha küçüktür. Bu durumda kanser lenf bezlerine sıçramamıştır. Tedavide meme koruyucu yöntemle lenf bezlerinin alınmasından sonra radyasyon tedavisi uygulanır. Desteklemek için kemoterapi ve/veya hormonoterapi eklenir. Bir diğer uygulama da mastektomi yöntemidir. Bu yöntemde kanserli meme alınarak koltuk altı lenf bezleri çıkarılır.

b) Evre 2-A: Tümörün 20 – 50 mm. arasında olup, lenf bezlerine sıçramamış halidir.

c) Evre 2-B: Tümörün 2-A evresindeki gibi bir boyut aralığında olup (50 mm. den büyük olabilir), koltuk altı lenf bezlerine sıçramamış halidir.

d) Evre 2: Evre 1 ile aynı tedavi yöntemi uygulanmakla birlikte, eğer tümör aşırı büyümüş ya da lenf bezlerine sıçramışsa kemoterapi, hormonoterapi ve radyasyon tedavisi tamamlayıcı olarak önerilir.

e) Evre 3-A: Tümör koltuk altı lenf bezlerine ve meme dışı dokulara sıçramış durumdadır ve bu halde mastektomi yöntemiyle tedavi yoluna gidilir. Cerrahi müdahaleden sonra kemoterapi ve hormon tedavisi uygulanır.

f) Evre 3-B: Bu aşamada tümörün boyutu dikkate alınmaz; tümör meme duvarına bağlıdır ve lenf bezlerine sıçramıştır. "Neoadjuvant" adı verilen tümörün boyunun küçültülmesi amaçlı kemoterapi uygulanmasının ardından, tümörün boyunun küçülmesinden sonra lampektomi veya mastektomi yapılır.

g) Evre 4: Bu aşamada kanser meme dışındaki vücut bölümlerine yayılmıştır. Bu evre tedavisinde hastanın yaşam süresini artırmak ve yaşam kalitesini yüksek düzeyde tutmak hedeflenir. Kemoterapi ve hormonoterapi yapılır. Hasta şikayetlerine bağlı olarak mastektomi de uygulanabilir.

1.9 Problem Cümlesi

Genelde meme kanseri teşhisi veritabanındaki dataların miktarlarına bağlı ve bu dataların iyi sınıflandırma edilememesinden dolayı çok zor ve hassas bir işlemdir. Bu tezin çalışmasında “GRNN” ve “YSA” ve “DVM” uygulayarak yeni bir metot ile probleme çözüm aranmıştır. Meme kanseri verileri, sağlanan çözümün meme kanserin iyi huylu veya kötü huylu olduğunarağmen daha hassas tanımlayabilmesi amacıyla MATLAB ile program geliştirilmiştir.

Bu tezde, otomatik meme kanseri teşhisi için bir yazılım önerilmiştir. Sistem üç aşama içermektedir. Bilgisayarlı sistemin ilk aşamasında, GRNN sınıflandırma yöntemini uyguladık. Sonar YSA ve akabinden DVM uygulandı.

1.10 Araştırmanın Amacı ve Önemi

Bu proje çalışması; yayılcı süt kanalı meme kanserinde östrojen alıcı durum tanısı amacıyla kullanılabilir çeşitli mühendislik araçlarının (yöntemlerinin), tıbbi tanı elde edilmesinde kullanımını ve bunların başarısının incelemesini amaçlar.

Bu amaç doğrultusunda, Şekil 1’de gösterildiği gibi giriş verisi olarak östrojen alıcı boyaması ile boyanmış sayısal bir resmi alan ve çeşitli işlem basamaklarını uyguladıktan sonra, çıkış olarak Allred puanlama sistemine göre östrojen tanısı koyan bilgisayar tabanlı bir sistem gerçekleştirilmiştir [5]. Bu sistemde giriş verisi, östrojen boyaması ile boyanmış biyopsi örneğinin sayısal görüntüsüdür. Bu görüntü, tümörü temsil eden bölgede patolog hekim tarafından kayıt edilmiştir.



Şekil 1.2: Bilgisayar tabanlı tanı sistemi işlem basamakları.

1.11 Kavramsal Çerçeve

Yapay Sinir Ağları, sağlık bilimleri alanında sağlık problemlerinin çözümünde ve çeşitli hastalıkların teşhisinde önemli bir rol oynamaktadır. Sinir ağları, sınıflandırma problemleri söz konusu olduğunda kullanılır. Bu çalışmanın amacı ameliyat sonrası rekürrens ve nüks etmeni öngörmeye iyi ve sağlam bir yöntem bulmaktır. Bu çalışmanın nihai amacı, riskleri azaltmak için çeşitli tıbbi prognoz sorunlarında klinik karar verme desteği için tahmini bilgisayar modelleri geliştirmektir. Şebekenin potansiyel aşırı eğitiminden dolayı doğruluk oranlarında yapay enflasyona karşı bir yineleme tahmin modeli üzerinde bir çapraz doğrulama veri örneği test etmek.

1.12 Alt Problemler

Meme kanseri veri tabanlarında meme kanseri teşhisi ve kaydedilmesi için bir grup bilgisayar algoritması olan meme kanseri sınıflandırma oldukça ilginçtir. Meme kanseri teşhisi ve analiz edilmesi için çok çeşitli yöntemler kullanılabilir. Bununla birlikte, meme kanseri analizinin oldukça araştırılan bir alanı olmasına rağmen, sınıflandırma uygulanmasında çok sayıda problem bulunmaktadır. Oluşturulan verisetlerindeki parazit ve yapay olgular, en basit algoritmaların meme kanseri teşhisi uygulanmasını zorlaştırmakta ve geliştirici hız ve sağlamlık arasında bir seçim yapmak zorunda kalmaktadır. Meme kanseri teşhisi uygulamalarda hem hız hem de sağlamlık hayati bir önem taşıdığından dolayı, tıbbi gerçekliğe uygun olmayan çözümler ortaya çıkmaktadır.

1.13 Kısıtlılıklar

Farklı sınıflandırma yöntemleri ve uygulamaları ve farklı iki veritabanı yüksek lisans tezi için eldeki zaman kısıtlı olduğundan, konu meme kanseri betimleme uygulaması ile sınırlandırılmıştır.

1.14 Raporun Ana Hatları

Tezin bu giriş bölümünden sonra, literatür taramasını, temel kavramlarını ve tanımlamalarını ele alan ikinci bölüm yer almaktadır. Arkasından, ana algoritmaların uygulanmasını, GRNN, YSA ve DVM algoritmalarını detaylı şekilde inceleyen üçüncü bölüm gelmektedir. Dördüncü bölümde ise uygulama, sonuçlar ve bu çalışmadan elde edilen çıkarımlar mevcuttur. Son bölüm olan beşinci bölümde sonuçlar tartışılmış ve daha sonraki çalışmalara ışık tutacak bulgular özet olarak anlatılmıştır.



İKİNCİ BÖLÜM

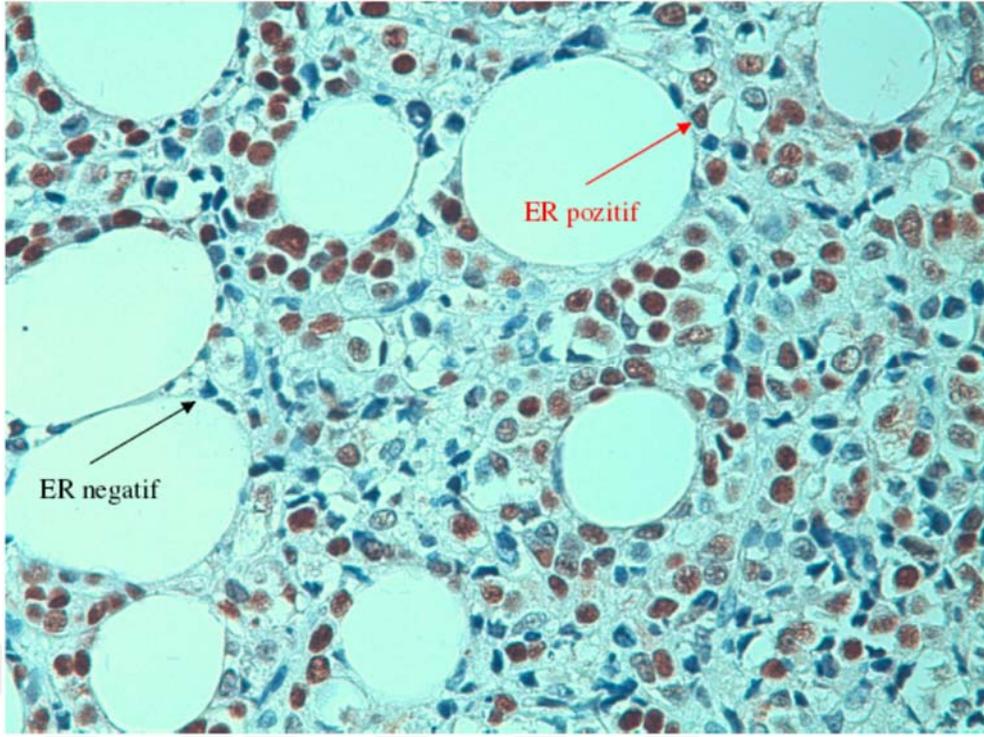
TEMEL KAVRAMLAR VE LİTERATÜR ÖZETİ

Bu bölüm mevcut çalışmaların daha ayrıntılı bir şekilde tanımlanmasına ve bu tezin temelini oluşturan ana kavramlar ve yöntemlerin açıklanmasına odaklanmaktadır. Meme kanseri, sınıflandırma yöntemleri literatür taraması bu bölümde anlatılmıştır.

2.1 Temel Kavramlar

Östrojen alıcı boyamasını gösterir örnek bir sayısal görüntü aşağıda verilmektedir (Şekil 2.1). Bu görüntü üzerinde proje çalışması kapsamında oluşturulan bilgisayar tabanlı tanı sistemi sırasıyla bölütleme, öznitelik çıkartma, öznitelik eleme, hücre çekirdeği sınıflama ve Allred puanı hesaplama işlemlerini uygulanır ve sonuçta östrojen alıcı varlığına dair tanı elde edilir. Bilgisayar tabanlı tanı sistemi, östrojen alıcı boyamasında östrojen alıcısı pozitif ve negatif çekirdekleri ayırt ederken kahve, mavi veya mor tonlarda boyanma özelliğinin yanı sıra hücre çekirdeklerinin desen özniteliklerini de dikkate alır.

Görünüşlerin karar vermeye yeterli olduğu alanlarda, patolojik incelemenin tanıya ve uygun tedavi yönteminin belirlenmesine katkısı da çok büyüktür. Günümüzde, tümörlerin tanısı başta olmak üzere, pek çok hastalığın kesin tanısı için patolojik inceleme gerekli ve zorunludur. Özellikle kanser tedavisinde kanserli dokunun değişiminin belirlenmesi, hastalığın seyri ve uygulanacak tedavi şeklinin belirlenmesinde büyük önem taşımaktadır. Elde edilen resimlerde çok sayıdaki organik yapı, şekil, renk ve desen bakımından kaynaşmış bir görüntü sergilemektedir. Mikrometre ölçeğinde çok küçük boyutlu resimlerde dahi çok sayıda hücrenin sayılması ve nicelik bakımından ayrı ayrı değerlendirilmesi zaman alıcı, yorucu ve hataya açık bir işlemdir [6].



Şekil 2.1: Östrojen alıcı boyamasını gösterir örnek bir sayısal görüntü [6].

Bir patolog aynı görüntüyü farklı zamanlarda farklı yorumlayabilmekte veya aynı görüntü için farklı hekimler farklı inceleme raporları sunabilmektedir. Kısacası değerlendirmeler soyuttur ve kişiye, zamana göre değişim gösterebilmektedir [6].



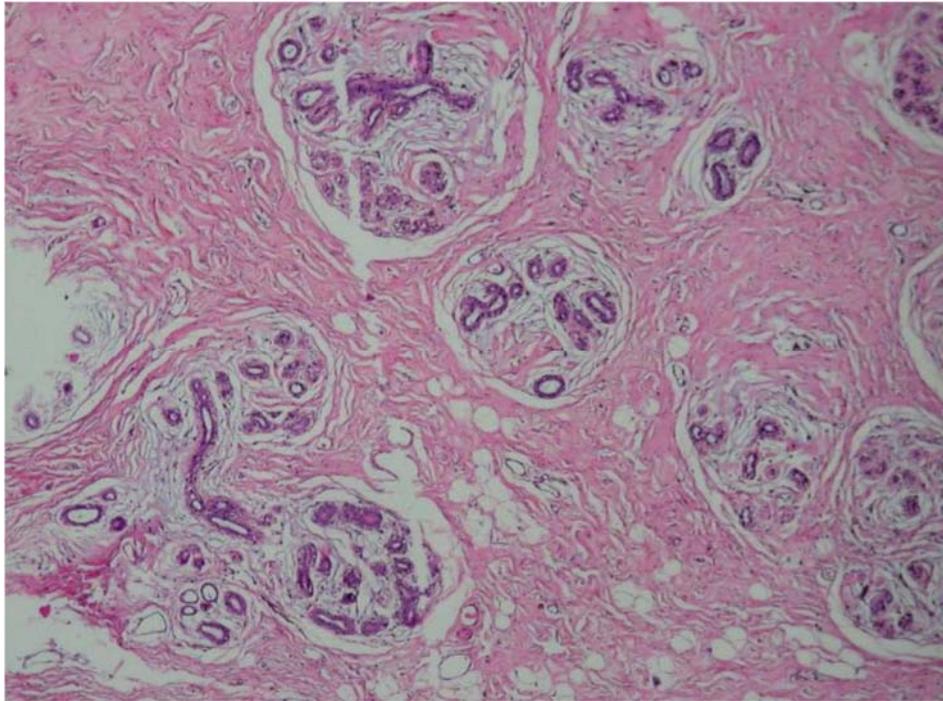
Şekil 2.2: Meme kanserinde tümör.

2.2 Doku Yapısında Meme Kanserinin Yaygın Türleri

Meme kanserinin en yaygın türleri; ya süt taşıyan kanallarda oluşan süt kanalı kanseri ya da süt bezlerinde oluşan süt bezi kanseri olarak bilinir. Kanserin meme dokusu üzerinde nerede olduğu mikroskop incelemesi ile belirlenir. Şekil 3'te, 40x büyütme için mikroskopu altında Hemotoksilin Eosin (HE) boyanmasında normal bir meme dokusuna örnek görüntü verilmiştir [7]. Meme kanserinin adlandırılması yapılırken kanserin nerede olduğunun yanı sıra kanserli hücrenin davranış biçimine de bakılır.

2.2.1 Yayılmacı Olmayan

Latince'de yerinde manasına gelen (in situ) kelimesiyle de tanımlanan bu davranış biçiminde, kanserli hücreler oluştukları yerde kalıcıdır, süt bezleri yahut kanalları dışındaki meme dokusuna yayılmazlar. Yayılmacı olmayan meme kanseri türleri içerisinde en çok görüleni yayılmacı olmayan süt kanalı kanseridir. Bu kanser türünde, kanser hücreleri süt kanallarının iç duvar yüzeyinde hapis kalırlar ve dışarı yayılmazlar (Şekil 2.3). Uygun bir tedavi yöntemiyle bu kanser türünde başarılı sonuçlar alınabilmektedir.



Şekil 2.3: HE boyanmasında normal bir meme dokusunun görünümü [7]

2.2.2 Yayılımacı

Bu tür kanserde, kanserli hücreler süt kanalını veya bezini diğer meme dokusundan ayıran ince zarın dışına geçerek meme dokusu içerisinde yayılırlar. Kanser hücreleri vücudun diğer kesimlerine, örneğin lenf düğümlerine, geçebilirler. Yaygın iki türü bulunmaktadır [2]. Yayılımacı Süt Kanalı Kanseri (YSKK): Meme kanseri vakalarının %70'i bu kanser türündendir. Kanser hücreleri süt kanalının iç yüzey duvarlarında oluşur, süt kanalı duvarı içerisinde yakındaki meme dokusuna geçer. Bu aşamadan sonra kanser hücreleri belirli bölgede kalabilir veya kan dolaşım sistemi ve lenf sistemi vasıtasıyla vücudun diğer bölgelerine yayılabilirler.

Yayılımacı Süt Bezi Kanseri (YSBK): Her ne kadar YSKK den daha az rastlansa da, bu tür kanser süt bezlerinde baslar ve benzer biçimde çevreleyen meme dokusu içerisine yayılır. YSBK vücudun daha uzak bölgelerine de yayılabilir. Bu kanser türünde ayrık sert bir parça hissi yerine meme üzerinde bir bölgenin veya tamamının katılaşması hissi vardır. Bu yaygın türlerin yanı sıra nadir rastlanan farklı bazı meme kanseri türleri de vardır [8].

Meme kanseri kadınlarda oldukça yaygın olup, dünyada çapında ikinci en yüksek ölüm sayısının nedeni olarak gösterilmektedir. Meme kanserinin saptanması için kullanılan ve günümüzde mevcut olan külfetli, pahalı ve zaman alıcı teknikler yeni, özel ve aşırı duyarlı cihazların geliştirilmesi ihtiyacını güçlendirmektedir. Biyosensörler, hasta başında kullanılan araçlar olarak muazzam potansiyele sahip umut verici ve seçici algılama cihazlarıdır. Bu inceleme, bugüne kadar geliştirilen çeşitli meme kanseri biyosensörlerini ve teknik değerlendirmesini seçilen biyoreseptörlerin ve biyotransdüserlerin etkililiği ve potansiyeli bakımından kapsamlı bir şekilde bütün boyutlara ile ele almaktadır. Glikoproteinlerin, DNA biyobelirteçlerinin, mikro-RNAların, sirkülatuar (dolaşımdaki) tümör hücrelerinin ve bazı potansiyel biyolojik belirteçlerin kısa bir tanımı yapılmıştır. İnceleme ayrıca nano-malzemeler, redoks araçları, p19 proteini, çift özel nükleazlar ve redoks döngüsü gibi sinyal yükseltilmesinde kullanılan çeşitli stratejileri tartışmaktadır [9].

Düşük seviyeli elektrik alanlarının, dakikalar veya saat boyunca uygulandığında hücre zarı reseptörlerinin uzamsal olarak yeniden dağılımını başlattığı gösterilmiştir. Bununla birlikte, kanser tedavisi için geri dönüşümsüz elektroporasyonda tipik olarak kullanılan kısa süreli yüksek amplitüdü darbeler ile hücre iletişimi üzerindeki etkisi hakkında sınırlı sayıda literatür bulunmaktadır.

Bununa ilaveten, geri dönüşümsüz elektroporasyondan sonra bağışıklık hücre dolaşımına ilişkin sinyalizasyon ile ilgili literatür çelişkilidir. Darbe parametrelerinin (alan gücü ve maruz kalma süresi) hücre iletişimini etkilediğini ve daha sonra bağışıklık hücresi dolaşımını etkilediğine ilişkin bir varsayım ortaya koyduk. Bu hipotez, geri dönüşümsüz elektroporasyon ile tedavi edilen üçlü negatif meme kanseri hücreleri üzerinde deney ortamında test edildi ve buradaki darbe parametrelerinin anahtar hücre iletişim faktörleri üzerine etkisi araştırıldı. Önemli bir şekilde, gerçek zamanlı PCR mRNA ölçümleri ve ELISA protein analizleri timik stromal lenfopietin (TSLP) sinyalizasyonunun, tipik olarak klinik amaçlı olarak kullanılanların maruz kalma sürelerinden bağımsız olarak, kritik eşğin üzerindeki elektrik alan kuvvetleri tarafından aşağı regüle edildiğini ortaya koymuştur. Diğer tedavilerle (termal şok, kimyasal ayırma, kinaz inhibitörleri) karşılaştırıldığında, geri dönüşümsüz elektroporasyonun (IRE) timik stromal lenfopietin (TSLP) üzerinde benzersiz bir etkisi olduğu ortaya çıkmıştır. TSLP sinyalleme sinin meme ve pankreas kanserlerinde pro-kanserli bağışıklık hücresi fenotiplerini tahrik ettiği gösterildiğinden dolayı, bulgularımız yaşayan canlılar üzerinde bir anti-tümör bağışıklık cevabı indüksiyonu için geri dönüşümsüz elektroporasyonun (IRE) potansiyel kullanımına yönelik daha ileri araştırmaları teşvik etmektedir [10].

Elektroporasyonun insan meme kanseri (MDA-MB-231), insan kolon kanseri (SW-480 ve HCT-116), insan fibroblast hücre dizisi (MRC-5), primer insan aortik düz kas hücreleri (hAoSMC) ve insan umbilikal damar endotel hücreleri (HUVEC) üzerindeki etkileri incelenmiştir. Hücre canlılığı izlemek için gerçek zamanlı teknoloji kullanılmıştır. Hücre ölüm tipi saptaması için akridin turuncu boya/etidyum bromür tahlili uygulanmıştır. Sayısal bir elektroporasyonun modeli önerilmiştir. Elektroporasyon, doz (voltaj) bağımlı şekilde hücre yaşayabilirliğini tetiklemiştir. 375-437.5 V cm⁻¹ elektroporasyon tedavisi, kanser hücrelerinin geri dönüşümsüz elektroporasyonuna ve sağlıklı hücrelerin geri dönüşümlü elektroporasyonuna sebebiyet vermiştir. Daha düşük voltaj oranının (250 V cm⁻¹) uygulanması baskın hücre ölümü tipi olarak apoptosise yol açarken, daha yüksek voltaj (500 V cm⁻¹) kullanımı büyük oranda nekroza neden olmuştur. Elektroporasyon kanser tedavisinde umut verici bir yöntemdir. Farklı kanser hücresi dizileri aynı elektroporasyon tedavisine farklı cevaplar vermiştir. Elektroporasyon 375-437.5 V cm⁻¹ seçici bir şekilde kanser hücrelerinin (SW-480) kalıcı hasarına neden olurken, sağlıklı hücreler

(MRC-5, hAoSM ve HUVEC) 72 saat sonra iyileşmiştir. Hücre ölümü tipi elektroporasyon koşullarına bağlıdır. Önerilen sayısal model elektroporasyon tedavisine ilişkin olguların analizi için faydalıdır [11].

Dijitalleştirilmiş histopatolojide meme kanserinin düşük, orta ve yüksek derecelendirilmesini otomatik olarak ayırt etmek için yeni bir görüntü analizine dayalı yöntem ortaya koymaktayız. Kıvrımlı sinirsel ağlardan (KSA) türetilen piksel-, nesne- ve semantik-düzey özellikleri de dahil olmak üzere çoklu seviye bir özellik seti 106 kadın hastadan alınan 106 hematoksilen ve eozin boyalı meme biyopsisi dokusu çalışmalarından yola çıkılarak elde edilmiştir. Bu çoklu seviye özellikler sadece kanser morfolojisinin karakterize edilmesini değil aynı zamanda histopatolojik görüntüler içinde yapısal ve çevrilebilir bilgilerin elde edilmesine olanak vermektedir. Bu çalışmada çekirdekleri görüntülerden ayırmak için gelişmiş bir hibrit aktif kontur modeli esaslı segmentasyon yöntemi kullanılmıştır. Semantik seviye özellikleri, özelliklerin tek tek alt türlerinden potansiyel olarak daha iyi sonuç verecek bir görüntü öznitelikleri seti oluşturmak için piksel düzeyi (doku) ve nesne düzeyinde (mimari) özelliklerle birlikte farklı derecelere ait çekirdek oranlarını tanımlayan kıvrımlı sinirsel ağ (KSA) yaklaşımı ile çıkarılmıştır. Çoklu seviyelerden çıkarılan farklı özellik setlerini güçlendirerek performansı en üst düzeye çıkarma olanağını oluşturabilmek için, özellik alt türleri kombinasyonlarını kullanarak çoklu destek vektör makinesi (DVM) sınıflandırıcıları eğitmek için kademeli bir yaklaşımdan faydalanılmıştır. Son dönem (kansere derecesi) tek tek destek vektör makinesi sınıflandırıcıları tarafından üretilen puanlar birleştirilerek belirlenmiştir. Hafif (üç katmanlı) bir KSA modeli ve paralel hesaplama kullanarak, sunulan yaklaşımın hesaplama açısından verimli olduğu ve büyük ölçekli veri kümelerine uygulanabilir olduğu ortaya konulmuştur. Yöntem histopatolojik meme kanseri görüntülerinin düşük, orta ve yüksek derecelerini ayırt ederken; düşükle yüksek arasında 0.92, düşük ve orta arasında 0.77 ve orta ile yüksek arasında 0.76 doğruluk oranı, genel olarak ta 0.69 doğruluk oranı sağlamıştır. Bu da, derecelendirme yöntemimizin tanı için objektif, tekrarlanabilir ve alternatif sağlayacak şekilde meme kanseri derecelendirmelerini ayırt etmek için bilgisayarlı bir tanı aracı geliştirmede yararlı olabileceğini düşündürmektedir [12].

Epitel (EP) ve stromal (ST) histolojik görüntülerde kullanılan dokuların iki türüdür. Tümör mikro ortamını analiz etmek için bilgisayarlı bir sistem geliştirirken,

EP ve ST dokularının otomatik segmentasyonu veya sınıflandırılması önemlidir. Bu makalede, dijitalleştirilmiş tümör dokusu mikro dizilimlerinden (TMA'lar) EP ve ST bölgelerini otomatik olarak bölmek veya sınıflandırmak için Derin Kıvrımsal Sinir Ağları (DKSA) tabanlı bir özellik öğrenimi sunulmaktadır. Mevcut yaklaşımlar, iki bölgenin sınıflandırılmasında renk, doku ve Yerel İkili Modeller (YİM) gibi el becerisi özellik sunumuna dayanmaktadır. Göreve bağlı temsili içeren el becerisi özellik tabanlı yaklaşımlarla karşılaştırıldığında, DCNN veri odaklı biçimde doğrudan EP ve ST dokularının ham piksel yoğunluğu değerinden öğrenilebilecek uçtan uca özellik çıkarıcısıdır. Bu üst düzey özellikler iki doku tipini ayırt etmek için denetlenmiş bir sınıflandırıcının oluşturulmasına katkıda bulunmaktadır. Bu çalışmada, DCNN tabanlı modeller ile el becerisi özellik çıkarımı tabanlı yaklaşımları sırası ile meme kanserinin 157 Hematoksilen ve Eozin (H&E) boyalı görüntüsü ve kolorektal kanserin 1376 immünohistolojik (IHC) boyalı görüntülerinden oluşan iki farklı veri kümesi üzerinde karşılaştırdık. DCNN tabanlı özellik öğrenim yaklaşımının sırası ile H&E boyalı (NKI ve VGH) ve IHC boyalı verilerde %85,%89 ve%100 F1 sınıflandırma puanı, %84,%88 ve%100'lük bir doğruluk oranı (ACC) ve %86, %77 ve %100'lük bir Matthews Korelasyon Katsayısı (MCC) gösterdiği bulunmuştur. DNN tabanlı yaklaşımımız EP ve ST bölgelerinin sınıflandırılması açısından üç el becerisi özellik çıkarma temelli yaklaşımdan daha iyi sonuç vermiştir [13].

Mevcut bu çalışmada, mikro besinlerin meme kanserine duyarlılığı nasıl hafiflettiğine ilişkin araştırma yapma amacıyla folat ve ksenobiyotik yol genetik polimorfizmlerinin verilerinden besleyici ve demografik değişkenlerle birlikte yapay sinir ağı tabanlı (YSA) bir meme kanseri tahmini modeli geliştirilmiştir. Gelişmiş ANN modeli meme kanseri tahmininde %94.2 değişkenlik ortaya koydu. Folat (400 µg/gün) ve B12 'nin (6 µg/gün) sabit etki modelleri sırasıyla %33.3 ve %11.3 risk azaltımı göstermiştir. Çok faktörlü boyut boyutluluk indirme analizi folata cevap verenlerde aşağıdaki etkileşimleri gösterdi: RFC1 G80A × MTHFR C677T (birincil), COMT H108L × CYP1A1 m2 (ikincil), MTR A2756G (üçüncü derece). Yanıtlayıcıların B12 ile olan etkileşimleri RFC1G80A × cSHMT C1420T ve CYP1A1 m2 × CYP1A1 m4 'tür. ANN simülasyonları artan folatın ER ve PR ekspresyonunu düzeltebileceğini ve ekstra hücrel süperoksit dismutaz ve BRCA1'in düzenleyici CpG ada metilasyonunu azaltabileceğini ortaya koydu.

Folatın diyetle yardımı ile alımı, EC-SOD ve BRCA1'in metilasyonu ve ER ve PR ekspresyonu etkilerini hafifletirerek meme kanserine karşı koruma sağladığı ortaya çıkarılmıştır [14].

Meme görüntüleme lezyonlarının otomatik sınıflandırılması hali hazırda çözülmemiş bir konudur. Bu makale mamografide meme kanseri tanısı için derin öğrenme teknikleri ve el becerisi ile görüntüleme tabanlı özellik saptayıcılarından sakınan ayırt edici özellikleri otomatik öğrenme ile birleştiren yenilikçi bir gösterimsel öğrenme çerçevesi tanımlamaktadır. Kitlelerle ilişkilendirilmiş el ile ayrılmış lezyonları (426 iyi huylu lezyon ve 310 kötü huylu lezyon) temsil eden toplamda 736 film mamografi (medyolateral yan karın kası ve kraniokaudal) görüntüsünü içeren 344 meme kanseri olgusundan biyopsi ile kanıtlanmış yeni bir kıyaslama veri seti oluşturulmuştur. Geliştirilen yöntem iki ana aşamadan oluşmaktadır: (i) görüntü ayrıntılarını geliştirmek için ön işlemler ve (ii) hem özellikleri hem de meme görüntüleme lezyonları sınıflandırıcısını öğrenmek için denetlenen eğitim. Önceki çalışmaların aksine, mamografi görüntülerinin içeriğini açıklamak için belirli tanımlayıcılar dizayn etmek yerine kıvrımlı sinir ağlarının denetlenen bir şekilde temsilini öğrenmek için kullanıldığı hibrit bir yaklaşımı benimsemekteyiz. Geliştirilmiş kıyaslama meme kanseri veri seti kullanan deneysel sonuçlar yöntemimizin yönlendirilmiş gradyanların histogramı (YGH) ve gradyan farklılıklarının histogramı (GFH) gibi en son teknoloji görüntü tanımlayıcılarına kıyasla önemli ölçüde geliştirilmiş performans sergilediğini göstermiş, ROC eğrisi (AUC) altındaki alan açısından performansı 0.787'den 0.822'ye kadar arttırmıştır. İlginç bir şekilde bu model radyoloji uzmanı tarafından uygulanan segmentasyondan elde edilen ek bilgilerin avantajlarından yararlanan bir dizi el becerisi özelliklerinden de daha iyi performans göstermektedir. Son olarak hem öğrenilen hem de el becerisi ile üretilen her iki gösterimin kombinasyonu kitle lezyonu sınıflaması için en iyi tanımlayıcı ile sonuçlanarak AUC skorunda 0.826 elde etmiştir. Mamografide meme kitlesi lezyonlarının sınıflandırılmasına otomatik olarak atıfta bulunan yeni bir derin öğrenme tabanlı çerçeve geliştirildi [15].

Bu çalışmada meme kanseri tanısı için derin kıvrımsal sinir ağı (CNN) kullanarak grafik tabanlı bir yarı denetimli öğrenme (SSL) şeması geliştirdik. CNN genellikle parametreleri eğitmek ve ince ayar yapmak için büyük miktarda etiketli verilere ihtiyaç duymaktadır, ancak bizim önerdiğimiz şemamız eğitim setinde

yalnızca küçük miktarda etiketli veri gerektirmektedir. Tanı sistemine dört modül dahil edilmiştir: veri ağırlığı hesaplama, özellik seçimi, birlikte eğitim veri etiketlemesine bölünmesi ve CNN. Bu çalışmada her biri, 1874 çift mamogram görüntüsünden çıkarılmış bir kitle içeren 3158 ilgi alanı (ROI'ler) kullanılmıştır. Bunların arasında 100 ROI etiketli veri olarak değerlendirilmiş, geri kalanları ise etiketsiz olarak değerlendirilmiştir. Çalışmamızda gözlenen eğri altındaki alan (AUC) 0.8818 'dir ve karışık etiketli verileri ve etiketsiz verileri kullanan CNN'nin doğruluğu 0.8243'tür [16].

Bu çalışmada, mikro besinlerin meme kanserine duyarlılığı nasıl hafiflettiğine ilişkin araştırma yapma amacıyla folat ve ksenobiyotik yol genetik polimorfizmlerinin verilerinden besleyici ve demografik değişkenlerle birlikte yapay sinir ağı tabanlı (ANN) bir meme kanseri tahmini modeli geliştirilmiştir. Gelişmiş ANN modeli meme kanseri tahmininde %94.2 değişkenlik ortaya koydu. Folat (400 µg/gün) ve B12 'nin (6 µg/gün) sabit etki modelleri sırasıyla %33.3 ve %11.3 risk azaltımı göstermiştir. Çok faktörlü boyut boyutluluk indirme analizi folata cevap verenlerde aşağıdaki etkileşimleri gösterdi: RFC1 G80A × MTHFR C677T (birincil), COMT H108L × CYP1A1 m2 (ikincil), MTR A2756G (üçüncü derece). Yanıtlayıcıların B12 ile olan etkileşimleri RFC1G80A × cSHMT C1420T ve CYP1A1 m2 × CYP1A1 m4 'tür. ANN simülasyonları artan folatın ER ve PR ekspresyonunu düzeltebileceğini ve ekstra hücrel süperoksit dismutaz ve BRCA1'in düzenleyici CpG ada metilasyonunu azaltabileceğini ortaya koydu. Folatın diyetle yardımı ile alımı, EC-SOD ve BRCA1'in metilasyonu ve ER ve PR ekspresyonu etkilerini hafifletirerek meme kanserine karşı koruma sağladığı ortaya çıkarılmıştır [14].

Her sekiz kadından biri hayatının herhangi bir döneminde meme kanserine duyarlı ve yatkın hale gelmektedir. Erken teşhis ve etkili tedavi meme kanserinden kaynaklanan ölüm oranlarını azaltmak için tek kurtarma araçlardır. Tıbbi tanıda meme kanseri tümörünün doğru bir şekilde sınıflandırılması önem arz etmektedir. Makine öğrenme teknikleri sınıflandırma yetenekleri nedeniyle tıbbi tanıda önemli bir yer tutmaya başlamıştır. Bu makalede sınıflandırma problemlerini çözmek için, yeni Genetik Olarak Optimize Edilmiş Sinir Ağı (GOOESA) algoritması önermekteyiz. Sınıflandırma için mimarisini (yapı ve ağırlık) optimize etmek amacı ile genetik bir sinir ağı geliştirdik. Standart çaprazlama ve mutasyon operatörlerinden

farklı olarak yeni çaprazlama ve mutasyon operatörlerini bu operatörlerin yıkıcı niteliklerini azaltmak için sunduk. Meme kanseri tümörlerini iyi huylu veya kötü huylu olarak sınıflandırmak için GONN algoritmasını kullanmaktayız. Sonuçlarımızı göstermek için, UCI Makine Öğrenme veri deposundan WBCD veri tabanını aldık ve GONN sınıflandırma doğruluğu, hassasiyeti, özgüllüğü, karışıklık matrisi, ROC eğrileri ve ROC altındaki AUC klasik model ve klasik geri yayılım modeli ile karşılaştırmıştır. Algoritmamız sırasıyla 50-50, 60-40, 70-30 eğitim-test bölümleri için %98.24, %99.63 ve %100 sınıflandırma doğruluğu ve 10 kat çapraz doğrulama için %100'lük bir sınıflandırma doğruluğu vermiştir. Sonuçlar yaklaşımımızın meme kanseri veri tabanı ile birlikte iyi çalıştığını ve bilinen makine öğrenme yöntemlerine iyi bir alternatif olabileceğini göstermiştir [17].

Rajib Kumar ve arkadaşları, düşük kontrastlı ve karanlık görüntüleri geliştirmek için doğrusal ve dinamik olmayan stokastik rezonansa dayalı bir teknik önerdi. Düşük bir kontrastı alt-öncesi sinyal olarak gördüler, daha sonra görüntü kontrastını iyileştirmek için gürültülü bir sinyal işleme kullandılar [18].

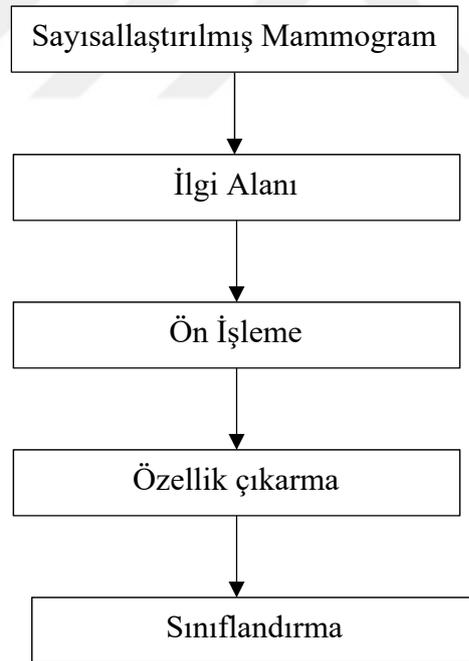
R. Mekle ve diğerleri. Al. Kolombiya Presbiteryen Tıp Merkezi ve Kolombiya Üniversitesi Biyomedikal Görüntüleme Laboratuvarındaki Meme Görüntüleme Merkezi'nde radyologlarla işbirliği içinde gerçekleştirilen bir çalışmada, ilk ROC çalışmasının araçlarla kontrast artışının çıkarlarını değerlendirebileceği bildirildi Multiscale uzantısı için bir taban olarak "dinamik spline dalgacık" kullanarak mamografinin tam ölçekli geniş ölçekli genişletilmesinden ve "ilişkili nonlinear sigmoid" iyileştirme fonksiyonunu [19] tasarlamış olmasındır.

Papadopoulos A. ve diğerleri. Al. Çalışmalarında görüntü iyileştirme işlem aşamasının etkisini ve mamografide mikro-kalsifikasyonlar tespiti için bir bilgisayar destekli algılama (BDA) sisteminin parametre ayarlamasını değerlendirdi. Kontrast sınırlandırılmalı uyarlamalı histogram eşitlemesi (CLAHE), yerel aralık modifikasyonu (LRM) ve fazlalıklı ayırık dalgacık (RDW) doğrusal germe ve çekme algoritmaları [20] getirilerek beş görüntü iyileştirme algoritması kullanılmıştır.

Bilgisayar Destekli Algılama (BDA) optimum, en iyi kontrastlı piksellerin yüzdesine ve mikroskatalizasyonu tatminkar bir biçimde temsil edebilen minimum algılanabilir nesne boyutuna hedeflenmiştir. LRM ve dalgacık tabanlı doğrusal gerdirme, LRM elde edildiğinde ($A(Z) = 0.932$) en yüksek performans oranına ulaşmış ve dalgacık tabanlı doğrusal gerilme ($A(Z) = 0.926$) [21] sağlamıştır.

Mamografinin sınıflandırılması, erken bir evrede meme kanseri keşfetmek için popüler yöntem olduğundan devam eden çok önemli bir araştırmadır. [22] amacı, özellik çıkarım algoritmalarını kullanarak mamografi sınıflandırma sürecini otomatikleştirmenin yollarını önermek ve daha sonra makine öğrenme algoritmaları kullanarak sınıflandırmaktır. Denetlenen makine öğrenmesinde araştırmanın en aktif alanlarından biri, iyi öğrenen toplulukları oluşturmak için yöntemler araştırmaktır. Bu yazıda [22] meme kanseri verileri Destek Vektör Makinesi (SVM) ve Yapay Sinir Ağı (YSA) sınıflandırıcıları ile sınıflandırılarak sınıflandırma süreci tamamlanmıştır. ANN, sınıflandırma problemleri için denetlenen öğrenme algoritması olarak öne çıkmaktadır. Amaç mamogramların sınıflandırılmasında yüksek kalitede sonuçlar veren sınıflandırma algoritmalarının iyi kombinasyonunu önermektir [22].

Mamogramlar, lezyonları ve ilgili bölgeleri bulmak için sayısallaştırılmıştır. Tümör bulabilmek ve görüntüyü geliştirmek için ön işlem tamamlanacaktır. Spesifik özellikler tümörleri benign veya habis olarak daha da sınıflandırmak için çıkarılır. Metodoloji grafiksel olarak şekil 2.4'te gösterilmektedir.



Şekil 2.4: [22]'deki önerilen yöntem.

Bu akış diyagramı meme kanserinin nasıl görüntü üzerinden sınıflandırma yapılacağını gösteriyor. İlk başta Sayısallaştırılmış Mammogram görüntüsü sisteme tanımlanıyor ve ilgi alanı seçiliyor. Sonra ön işleme metotları üzerinde uygulanır ve

ondan sonra özellik çıkarma metotları uygulanır. En sonda özelliklere göre sınıflandırma yapılır. K .Usha ran İleri besleme, Geri yayılım metodunu uyguladılar ve % 92 doğruluk oranını elde ettiler [23]. Destek vektor makinesi Liao, R. Tarafından uygulandı ve % 86.92 sonucu elde edildi [5]. 2005 yılında Bevilacqua, V. Xyct sistemi, Tek Yön Bırakma yöntemini kullanarak % 91 elde etti [24]. ANFIS metodunu meme kanseri taşıması için Land, W 2003 yılında uyguladı ve % 59.90 sonucunu elde etti [25]. Jordan-Elman neural network (JENN) Chunekar, V.N. tarafından uygulandı ve % 70.725 elde edildi [26]. Guruprasad Bhat ve arkadaşları 2012 yılında yapay sinir ağlarını kullanarak doğruluk oranını % 92.8571 elde ettiler [27]. Çevrimiçi eğitilmiş sinir ağı ilk kez FANG Yi ve arkadaşları tarafından 2006 yılında uygulandı ve % 94.9 oranda iyi bir miktara ulaştılar [28]. T. Kıyan ve arkadaşları 2004 yılında istatistiksel sinir ağları üzerinde çalıştılar ve % 95.74 sonucunu elde ettiler [29]. Y. Hu ve Ark. 1994 yılında Tek gizli katman, hata geri yayılımı ile ileri sinir ağları kullanarak % 96.9 doğruluk sonucunu elde ettiler [30].

Tablo 2.1’de araştırmaların özeti ve sonuçları gösterilmektedir

Tablo 2.1: Meme kanser tesbiti hakkında çalışmalar

Yazarlar	Kullanılan Teknik	Elde edilen sonuçlar
K .Usha ran 2010 [23]	İleri besleme, Geri yayılım	% 92
Liao, R. 2011 [5]	SVM	% 86.92
Bevilacqua, V. 2005 [24]	Xyct sistemi, Tek Yön Bırakma yöntemini kullanarak	% 91
Land, W 2003 [25]	ANFIS	% 59.90
Chunekar, V.N. 2009 [26]	JENN	% 70.725
Guruprasad Bhat ve arkadaşları 2012 [27]	ANN	% 92.8571
FANG Yi ve arkadaşları 2006 [28]	çevrimiçi eğitilmiş sinir ağı	% 94.9
T. Kıyan ve arkadaşları 2004 [29]	istatistiksel sinir ağları	% 95.74
A. Papadopoulos 2008 [20]	Kontrast sınırlandırılmalı uyarlamalı histogram eşitlemesi (RDWveLRM)	% 86.6
Raif Mekle ve Ark. 2000 [19]	spline dalgacık(very tabanı)	% 91.36
Y. Hu ve Ark. 1994 [30]	Tek gizli katman, hata geri yayılımı ile ileri NN	% 96.9

Bu tezde iki farklı veri tabanı kullandık. İlk veri tabanında 194 kişi vardır ve bu kişilerden 148 kişi sağlam ve 46 kişi hastadır. İkinci veri tabanında 569 kişi vardır ve bu kişilerden 212 sağlam ve 357 hastadır.

Bu tezde MATLAB 2017 versiyonunu kullandık. Tablo 2.1de gösteren sonuçlara göre bizim sonuçlar daha iyi oldu ve sonuçları 4. Bölümde göstereceğiz.

Bu tezde GRNN, YSA, DVM ve Wisconsin Prognostic Breast Cancer (WPBC) kullanarak bir sınıflandırma yöntemi önerilecek. [22] referansında göleme metodunu kullanarak meme kanseri teşhis edilmiştir. Bu tezde görüntü işleme üzerinde çalışmayacağız ve veri tabanı üzerinde çalışacağız.

Bu tezde sınıflandırma yöntemlerin üzerinde çalışacağız ve en iyi senaryoyu bulacağız. GRNN ve YSA ve DVM yöntemlerini kullanarak en iyi sınıflandırma yöntemini ortaya çıkaracağız. Bu tezde Wisconsin Meme Kanseri Prognozu (WPBC) veri tabanı kullanacağız. Bu veri tabanında 569 kişi var ve toplam 212 kişi hasta ve 357 kişi sağlam. Her kişiye ait 30 parametre var ve bu parametreler sınıflandırmada özellik olarak uygulanacak. YSA ve GRNN ve DVM’ında girdi vektörü 30 olacak ve çıktı vektörü sadece 1 olacak. Bu tezin amacı sınıflandırmaktır ve özellik çıkarma üzerinde çalışmayacağız. [22]’deki referansta sistemin doğruluğu YSA için %93 ve DVM için %95 elde edildi. Bu tezde bu yöntemleri ve GRNN yöntemini kullanarak en iyi performansı olan yöntemi ortaya çıkaracağız.

ÜÇÜNCÜ BÖLÜM

YÖNTEM

Bu bölümde önerilen yöntem hakkında bahsedilecek ve algoritmanın akış diyagramı anlatılacak. Veri tabanı hakkında bahsedilecektir ve sınıflandırma yöntemleri anlatılacaktır.

3.1 Wisconsin Prognostic Breast Cancer (WPBC)

Bu tezde Wisconsin Meme Kanseri Prognozu (WPBC) veri tabanı kullanılmıştır. Doktor William H. Wolberg, Wisconsin Üniversitesi Genel Cerrahi Bölümü tarafından hazırlanan bu veritabanı, Madison Bilim Merkezleri Klinik Merkezi'ni yarattı. Doğru tahmin için iki kategori vardır ve sık tekrarlar = R = N. veri kümesi belirli bir son noktayı yansıtacak şekilde önce süzülmalıdır, örneğin, 24 aydan önce nüks = pozitif, 24 aydan sonra tekrarlamama = negatif. Verilerin önceki versiyonunu kullanarak doğruluk %86.3, tekrarlamamanın doğruluğunu 2 yıl tahmin etmektedir.

Bu sefer nüksü öngörmek için 3. Alan tekrar suçludur. Relaps kullanımı ile ortalama tahmin hatası 13.9 ay düzeyindeydi.

Her kayıt, bir olgu için meme kanseri takip verilerini gösterir. 1984 ardışık hastalar bir doktor tarafından görülebilir ve anında sadece invaziv meme kanseri metastazları ve tanısı olmayan Wolberg vakalarını teşhis eder. Memedeki kitlelerin ince iğne aspirasyonunda 30'un ilk hesaplanan dijital görüntüsü bulunur. Onlar, hücre çekirdeğinde imgenin bir açıklamasıdır.

Her bir hücre çekirdeği için on özellik doğru olarak hesaplanır:

- A) Yarıçap (çevre merkezinden ortalama mesafe)
- B) Doku (gri tonlamanın standart sapması)
- C) Çevre
- D) Bölge
- E) Filtre (yarıçapın yerel çeşitliliği)
- C) Yoğunluk ($\text{çevre}^2 / \text{Bölge} - 1.0$)
- G) İçbükey (kesinlikle içbükey kesit konturu)
- H) İçbükey (konturun içbükey parça numarası)
- I) Simetri
- J) Fraktal boyut ("yaklaşık kıyı şeridi" - 1)

Yukarıda bahsedilen makalenin bazıları, bu özelliklerin nasıl hesaplanacağı hakkında ayrıntılı açıklamalar içermektedir. Ortalama, standart hata ve özellik "en kötü" veya (üç büyük değerlerin ortalaması) her bir resim için hesaplandı ve sonuçta 30 özellik elde edildi.

Orijinal veri seti Wisconsin Üniversitesi hastanelerinden Dr. William H. Wolberg tarafından elde edilmiş olup, veri setinde 16'sı kayıp değerler içeren 683'ü tam toplam 699 adet örnek içermektedir. Her bir kayıt 11 farklı nitelik değeri ile sunulmaktadır. Bu niteliklerden ikisi hasta numarası ve sınıf etiketidir. Meme kanseri verisinde sınıf etiketi için iki değer bulunur: iyi huylu (kansersiz olmayan) için 2 değeri ve habis (kansersiz) için 4 değeri vardır. Geriye kalan nitelikler meme kütesinin ince iğne aspirasyonu yardımıyla elde edilmiş resminden hesap edilen değerlerdir. Bu değerler; yığın kalınlığı, hücre boyutu homojenliği, hücre şekli homojenliği, marjinal yapışma, epitel hücre boyutu, çıplak çekirdekler, yumuşak kromatin, normal çekirdekçik ve mitoz ismi verilen niteliklerle 1-10 arası değerlerle hücre çekirdeği karakteristiklerini açıklar [31]. Diagnostic veri seti ve prognostic veri seti de orijinal veri seti gibi meme kitlesinin ince iğne aspirasyonundan elde edilen resmin sayısallaştırılmasına dayalı bilgiler içerir. Bununla birlikte Diagnostic veri seti toplam 32 adet özellikten meydana gelir. Bu özellikler; ID, Diagnosis (M=malignant, B=benign) ve 30 adet gerçek değerli özelliktir. Veri setinde toplam 357 benign, 212 adet malignant kayıt verisi bulunur. Kayıp değer içeren kayıt bulunmaz. Prognostic veri seti invaziv meme kanseri verileri içerir. Veride yer alan 30 özellik meme kitlesinin ince iğne aspirasyonundan elde edilen resmin sayısallaştırılmasından elde

edilmiştir. Bu özellikler görüntüde bulunan hücre çekirdeğinin karakteristiklerini açıklar. Toplam 198 adet kayıt bulunur. Her kayıt 34 özellik ile (ID, sonuç (outcome), 32 gerçek değerli giriş özelliği) sunulur. 151 nonrecur ve 47 recur kayıt vardır. 4 adet kayıp veri vardır. Gerek centroid tabanlı sınıflayıcılarla gerekse diğer sınıflayıcılarla çalışırken verinin %90'ı eğitim %10'u test olarak kullanılmış ve bütün deneylerde 10-katlı çapraz doğrulama çalıştırılmıştır. Sınıflayıcıları karşılaştırırken şekil 3.1'de detayları verilen doğruluk ölçümü kullanılmıştır.

Tablo 3.1: Katıştırma matrisi ve doğruluk değeri [31].

Gerçek	Tahmin edilen	
	İyi huylu	Kötü huylu
İyi huylu	A	B
Kötü huylu	C	D

A: TP (Gerçek Pozitif)

B: FN (Hatalı Negatif)

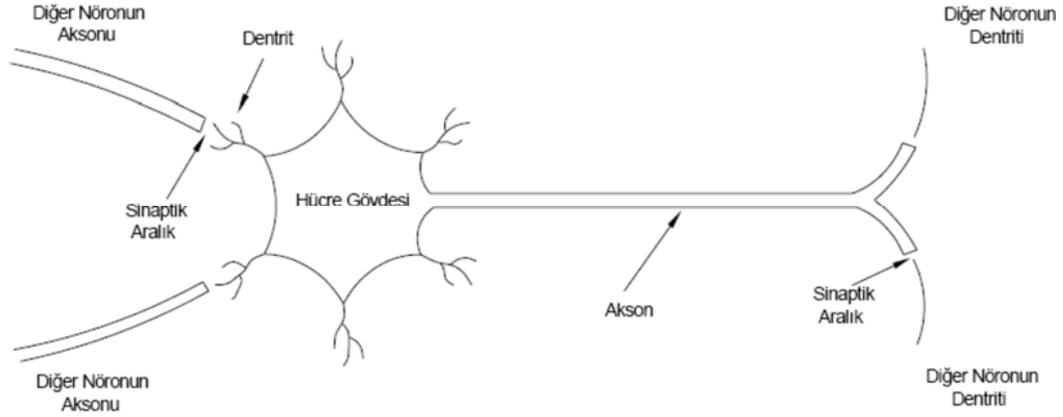
C: FP (Hatalı Pozitif)

D: TN (Gerçek Negatif)

$$dogruluk = \frac{a + d}{a + d + c + d} = \frac{TP + TN}{TP + FN + FP + TN} \quad (\text{Denklem 3.1})$$

3.2 Biyolojik Sinir Hücresi

İnsan beyni, nöron olarak adlandırılan yaklaşık 10^{11} hesap elemanından oluşmaktadır. Biyolojik sinir ağını oluşturan nöronlar; hücre gövdesi, akson, dentrit denilen üç temel elemandan oluşurlar (Şekil 3.1).

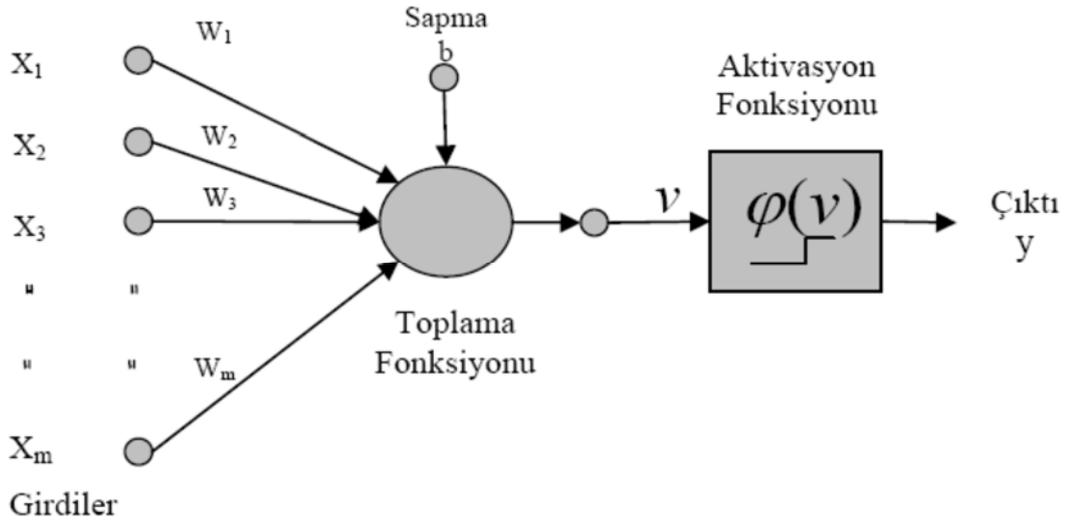


Şekil 3.1: Biyolojik sinir hücresi [32].

Soma (hücre gövdesi), hücreyi denetler ve hücre etkinliklerinin tümünü yönetir. Hücre gövdesinden dendritler ve aksonlar olarak adlandırılan iki uzantı çıkmaktadır. Dendritler bilgiyi iletim hatları olarak kullanılan uzun fiberlerden oluşmuş aksonlar boyunca diğer nöronlardan alır ve hücre gövdesine taşırlar. Aksonlar gövdedeki bilgiyi diğer nöronların dendritlerine taşımakla sorumludurlar. Akson sonları ile dendritler arasında bilgilerin uzun süre saklandığı bilgi saklama yerleri oldukları düşünülen sinaps denilen küçük boşluklar bulunur. Sinapslar, nöronun kendi sinyalini komşu nörona tanıttığı bağlantı noktasıdır. Sinaptik bağlantılar mesajların nöronlar arası taşınmasını sağlarlar. Birden çok sinaps yoluyla ulaşan uyarım, belirli bir eşiği aştığında ateşlenir ya da boşalır. İnsan beyinde öğrenme, yeni aksonlar üreterek, aksonların uyarılmasıyla, mevcut aksonların güçlerini değiştirerek gerçekleşmektedir [32].

3.2.1 Yapay Sinir Hücresi

Biyolojik sinir ağlarının yapı bileşenleri sinir hücreleridir benzer şekilde yapay sinir ağlarının da yapay sinir hücreleri bulunmaktadır (Şekil 3.2). YSA, insan sinir ağındaki gibi nöronlardan ve onlar arasındaki bağlantılardan oluşur. Bilgi, ağ tarafından bir öğrenme süreciyle çevreden elde edilir. Elde edilen bilgileri biriktirmek için sinaptik ağırlıklar olarak da bilinen hücreler arası bağlantı güçleri kullanılır [33].



Şekil 3.2: Basit algılayıcı modeli [34].

YSA'ya bilgi sayısal olarak dış dünyadan, diğer hücrelerden ya da kendi kendisinden gelebilir. Sinir hücresine bilgiler ağırlıklar yoluyla taşınırlar. Ağırlık değerleri bilginin önemini ifade eder, değişken ya da sabit değerler olabilirler, pozitif ya da negatif değerler alabilirler. Bir sinir hücresine gelen net bilgi yaygın olarak toplama fonksiyonu aracılığıyla hesaplanır. Her girdi değeri kendi ağırlığı ile çarpılır. Toplama fonksiyonu tüm girdiler için gelen bu değerleri toplayarak net hücre çıktısını hesaplar. Her hücre diğer hücrelerden bağımsız olarak bu net değerini hesaplar. Sapma (bias- b_k) değerinin aktivasyon fonksiyonuna giren değeri yükseltme ya da düşürme etkisi vardır. Denklem 3.1'de kullanılan x_j gelen bilgileri, w_{kj} her girdi değerine ait ağırlıkları, b_k sapma değerini, v_k nöronun çıktı değini ifade etmektedir.

$$v_k = \sum_{j=1}^m w_{kj} x_j + b_k \quad (\text{Denklem 3.1})$$

“Eş. 3.2”de görüldüğü gibi her sinir hücresinin net bilgisi eşik değerine sahip bir aktivasyon fonksiyondan geçirilerek gerçek bir çıktı oluşturulur. Genellikle kullanılan aktivasyon fonksiyonları eşik, sigmoid, hiperbolik tanjant vb. fonksiyonlardır. Aktivasyon fonksiyonu ($\phi(.)$) genellikle doğrusal olmayan bir fonksiyondur.

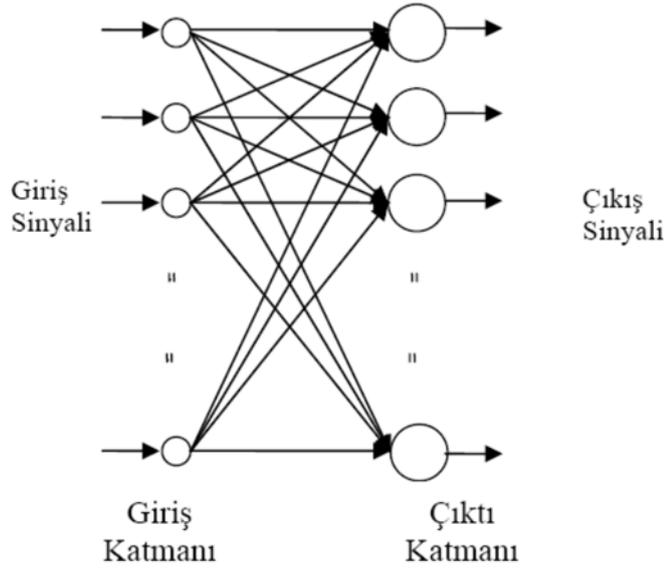
$$y_k = \phi(v_k) \quad (\text{Denklem 3.2})$$

YSA herhangi bir konu ile ilgili veri setleriyle eğitilirken eğitim algoritmaları kullanırlar. Öğrenilmesi istenen olay için oluşturulan eğitim seti ağa sunulurken

hedef çıktı değerleri de ağı sunulabilir. Sadece girdi seti ağı sunulabilir, sistemin kendi kendine öğrenmesi istenilebilir ya da her girdi seti için sistemin kendisinin bir çıktı üretmesi sağlanabilir. Üretilen çıktının doğru ya da yanlış olduğunu gösteren sinyal üretilerek, bu sinyale göre sistem eğitime devam edilebilir. YSA yapılarında bulunan katman sayılarına göre tek ya da çok katmanlı olarak gruplandırılırlar [8].

3.2.2 Tek Katmanlı Algılayıcılar (TKA) [35]

Tek katmanlı yapay sinir ağları giriş ve çıkış katmanlarından oluşur. Girdi ve çıktı katmanlarında birden fazla giriş ve çıkış değeri bulunmaktadır. Giriş katmanındaki her giriş değerini çıkış katmanına bağlayan sinaptik bağlantılar mevcuttur. Her bağlantı bir ağırlık değerine sahiptir. Aynı zamanda ağın çıktısının sıfır olmasını engelleyen bias sapma değeri mevcuttur (Şekil 3.3).



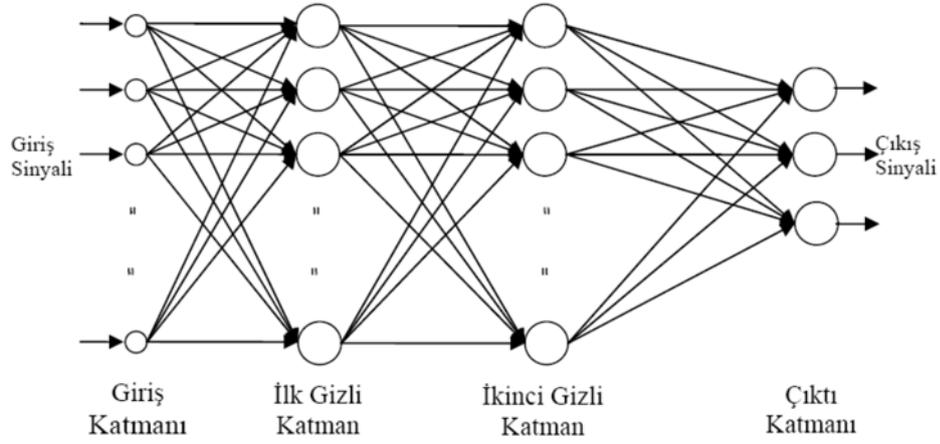
Şekil 3.3: Tek katmanlı algılayıcı [34].

Şekilde görüldü gibi bir giriş sinyali var ve bir de çıkış sinyali var. Buradaki nöron yapısı tek katmanlı oluyor.

3.2.3 Çok Katmanlı Algılayıcılar (ÇKA) [35]

Çok katmanlı algılayıcılar (ÇKA), bilgilerin girildiği girdi katmanı, bir veya daha fazla sinir hücresinden oluşan gizli katmanları ve bir çıktı katmanını içerir.

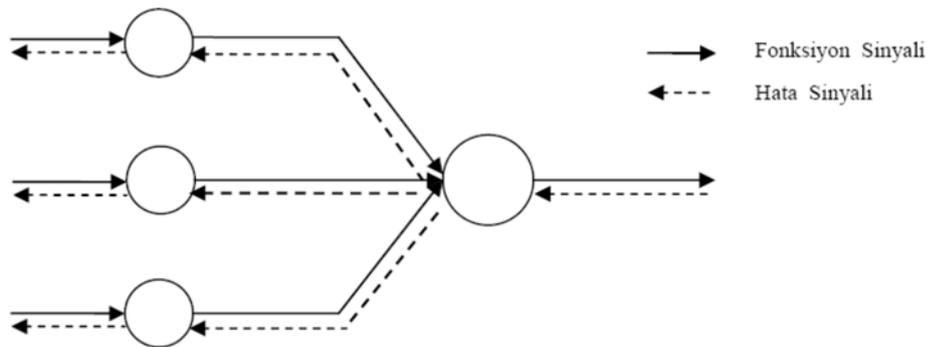
Girdi sinyalleri ağ boyunca bir katmandan diğer katmana ileri yönde yayılırlar (Şekil 3.4).



Şekil 3.4: Çok katmanlı algılayıcı [34].

Şekilde görüldü gibi bu iki tane gizli katman var bundan dolayı bu tip ağlara çok katmanlı yapay sinir ağı denir. Bu tezde bu modeli uyguladık ve 4. Bölümdü net bir şekilde anlatılacak.

Ağın farklı katmanları boyunca ileri ve geri yayılım olarak adlandırılan iki geçiş bulunur. İleri yayılımda, bir girdi vektörü ağın giriş katmanına uygulanır ve bu girdinin etkisi ağda katmandan katmana yayılır. İleri yayılım sırasında ağın sinaptik ağırlıkları belirlidir. Geri yayılımda sinaptik ağırlıkların tümü bir hata-düzeltilme kuralı ile uyumlu olarak düzenlenir. Bir hata sinyali üretmek için ağın gerçek çıktısı istenilen bir çıktıdan çıkartılır. Bu hata sinyali sinaptik bağlantıların tersi yönünde ağda geriye doğru yayılır. Sinaptik ağırlıklar ağın gerçek çıktısını istatistiksel anlamda arzu edilen çıktıya yakın hale getirmek için düzenlenir (Şekil 3.5).



Şekil 3.5: Sinyal akış gösterimi [34].

ÇKA'nın üç ayırt edici özelliği bulunur.

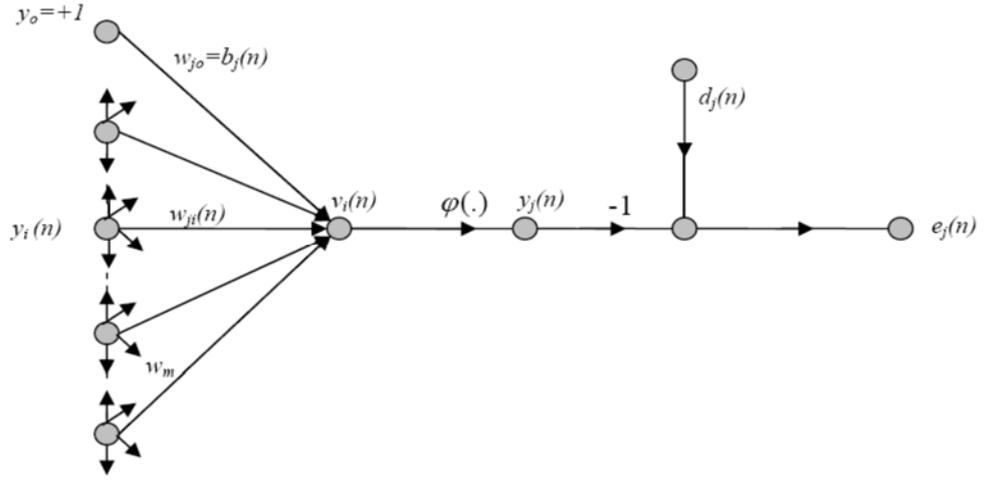
- 1- Ağdaki her nöron modeli doğrusal olmayan bir aktivasyon fonksiyonuna sahiptir. Bu fonksiyon çoğunlukla lojistik fonksiyon $y = 1/(1 + \exp(-v_i))$ kullanılır. Problemin doğrusal olmaması önemlidir, eğer doğrusal ise zaten tek katmanlı algılayıcı ile çözülebilir.
- 2- Bir ağ bir ya da daha fazla gizli katman içerebilir. Bu gizli katmanlar giriş ya da çıkış arasında bulunurlar. Katman sayısını problemin zorluk derecesi belirler.
- 3- Ağdaki nöronlar arası bağlantılar yüksek derecededir, ağ bağlantısındaki bir değişiklik, nöron bağlantı sayılarını ve onların ağırlıklarındaki değişikliği gerektirir. ÇKA için çok farklı öğrenme kuramları geliştirilmiş ve pek çok probleme uygulanmıştır ancak ÇKA bütün problemleri çözebilir diyemeyiz.

Fonksiyon sinyali giriş sinyalidir, ağdaki her şey onunla hesaplanır. Hata Sinyali, çıkış nöronundan başlayarak geriye doğru akmaktadır. Gizli katmanlar ağın giriş ve çıktı katmanları arasında bulunurlar. Gizli katmanlar ve çıktı katmanı giriş sinyallerini çıkış nöronlarına taşımak için kullanılırlar.

ÇKA'da doğrusal olmamanın varlığı ve ağın üst düzeydeki ilişkiselliği ağların teorik analizini üstesinden gelinmesi zor bir durum yapmaktadır. Ayrıca gizli nöronların kullanımı öğrenme sürecinin canlandırılmasını zorlaştırır, çünkü olası fonksiyonların aranması daha geniş bir uzayda gerçekleştirilmelidir.

3.2.4 Geriye Yayılım Algoritması

Yapay sinir ağlarında eğitim işlemine başlarken ağırlık ve bias değerleri rasgele verilir. Ağa örnekler sunulur ve bilgi ileri doğru yayılır, çıktı katmanındaki nöronlarda hata değeri bulunur (Şekil 3.6).



Şekil 3.6: Geri yayılım algoritması [34].

Şekil 3.6’da gösterildiği gibi Geri yayılım algoritmasının nöron yapısını ifade ediyor. Çıktı nöronunda n. eğitim basamağında oluşan hata (e), istenilen çıktıdan (d) hesaplanan çıktının (y) çıkarılması ile hesaplanır (Denklem 3.3).

$$e_j(n) = d_j(n) - y_j(n) \quad (\text{Denklem 3.3})$$

j nöronu bir çıktı düğümüdür.

Elde edilen bu hataların anlık hata enerjileri $e_j^2(n)/2$ ifadesi ile hesaplanır. Toplam hata enerjisi bütün nöronlardaki anlık hata enerjilerinin toplanması ile elde edilir. Hata kareler ortalaması, toplam hata enerjilerinin ε_{ort} bulunup eğitim küme sayısı N’ye göre normalize edilmesi ile hesaplanır (Denklem 3.4).

$$E(n) = \varepsilon_{ort} = \frac{1}{N} \sum_{n=1}^N \frac{1}{2} \sum_{j \in C} e_j^2(n) \quad (\text{Denklem 3.4})$$

Öğrenme sürecinin amacı hatayı minimize edecek ağırlık ve bias değerlerini yani ağırlık ve bias değerlerini ayarlamaktır. Bu işlem yapılırken eğitim setlerinin hepsi ağırlık ve bias değerleri ayarlanıncaya kadar bir kez sunulur, serbest parametreler ondan sonra ayarlanır ya da sunulan her örnekten sonra ağırlık ve bias değerleri güncellenir. Ağırlıklara yapılan düzenlemeler ağırlık ve bias değerleri güncellenen her örnek için hesaplanan hatayla uyumlu olarak yapılır.

Geri besleme algoritmasında ağırlık ve bias değerleri güncellemek için düzeltme değeri ($\Delta w_{ji}(n)$), hatanın ağırlıklara göre kısmi türevinin alınması ($\partial \varepsilon(n) / \partial w_{ji}(n)$) ile hesaplanır (Denklem 3.5).

$$\frac{\partial \varepsilon(n)}{\partial w_{ji}(n)} = -e_j(n)\varphi'(v_j(n))y_i(n) \quad (\text{Denklem 3.5})$$

$$\Delta w_{ji}(n) = -\eta \frac{\partial \varepsilon(n)}{\partial w_{ji}(n)} \quad (\text{Denklem 3.6})$$

$$\delta_j(n) = -e_j(n)\varphi'(v_j(n)) \quad (\text{Denklem 3.7})$$

“Eş. 3.5” de kullanılan η , geri besleme algoritmasının öğrenme oranı parametresidir, eksi işareti ağırlık uzayındaki hatayı düşürecek ağırlık değişimi için yön arayan gradyanı ifade eder. Lokal gradyan $\delta_j(n)$ ilgili nöron için karşılık gelen hata sinyali $e_j(n)$ ve ilgili karşılık gele aktivasyon fonksiyon türevinin $\varphi'(v_j(n))$ çarpımına eşittir (Denklem 3.6).

Düzeltilme değeri, lokal gradyanların kullanımı ile “Eş. 3.8” deki ifade edilir.

$$\Delta w_{ji}(n) = -\eta \delta_j(n) y_i(n) \quad (\text{Denklem 3.8})$$

Ağ eğitim esnasında hata yüzeyi sabit bir noktada takılı kalabilir, minimum hataya yakınsayamayabilir. Bunu engellemek için düzeltme değerine, momentum katsayısı olarak ifade edilen bir dğer eklenir (Denklem 3.9).

$$\Delta w_{ji}(n) = \alpha \Delta w_{ji}(n-1) - \eta \delta_j(n) y_i(n) \quad (\text{Denklem 3.9})$$

α momentum katsayısı

Elde edilen bu düzeltme değeri eski ağırlıklara eklenerek yeni değerleri elde edilir. Bu işlem doğru ağırlık değerleri elde edilene kadar devam eder [36, 37].

DÖRDÜNCÜ BÖLÜM

BULGULAR VE TARTIŞMA

Bu bölümde deneysel sonuçlar açıklanacaktır. Bölüm 4.1'de Normalleştirme Yöntemi, bölüm 4.2'de 4.2 performans ölçümü hakkında olacak ve bölüm 4.3 ve 4.4'te Sonuçların değerlendirilmesinde kullanılan değerlendirme ve doğruluk kriterleri.

4.1 Normalleştirme Yöntemi

Bu tezde normalleştirme için [38] yöntemi kullanılmıştır. İlk adımda tüm girdi verileri -1 ve 1'e normalize edilmiş ve daha sonra bu veriler eğitim algoritmasında kullanılacaktır. Cevabı aldıktan sonra çıktı verileri tekrar uygulanır. Fakat bu zaman zarfında normalleştirmenin tersi kullanılır.

I satırındaki E değişkeni için e_i 'nin normalize edilmiş değeri aşağıdaki şekilde hesaplanır:

$$Normalized(e_i) = \frac{e_i - E_{\min}}{E_{\max} - E_{\min}} \quad (\text{Denklem 4.1})$$

Burada

E_{\min} = E'nin minimum değeri

E_{\max} = E'nin maksimum değeri

Normalde, her satır için min ve max değerlerini hesaplar ve her satırın değerlerini buna göre eşleştirir. Bununla birlikte, tek bir column vektörüyle çağrıldığında sütun için tutarsız olarak min ve max değerlerini hesaplar ve buna göre değerlerini eşleştirir. Bu tezde kullanılan fonksiyon sürekli olarak satırlar için min ve max değerlerini hesaplar.

4.2 Performans Ölçümü

Kullanılan sınıflandırma algoritmasının etkililiğini belirlemek için bir ölçüm gereklidir. Sık kullanılan ölçümler arasında sınıflandırma doğruluğu, F-Measure, hassasiyet, geri çağırma, Alıcı Çalışma Karakteristiği eğrileri ve Eğri Altındaki Alan bulunur. Bu ölçümler, Confusion Matrix olarak adlandırılan bir matris biçiminde yaygın şekilde tablolaştırılan sınıflandırma sonuçları ile hesaplanabilir. Confusion Matrix Klasik bir ikili sınıflandırma probleminde sınıflandırıcı öğeleri pozitif veya negatif olarak etiketlendirir. Bir karışıklık matrisi, matris biçiminde algoritmanın sonucunu özetler. İkili örneğimizde, konfüzyon matrisinin dört sonucu vardır: Gerçek pozitifler (TP) pozitif olarak doğru pozitif öğeler olarak sınıflandırılır. Gerçek negatifler (TN) negatif maddeler olarak doğru olarak tanımlanırlar. Yanlış pozitif (FP) olumlu olarak sınıflandırılan olumsuz maddelerdir. Yanlış negatifler (FN) negatif olarak sınıflandırılan pozitif maddelerdir. Tablo 4.1, bir karışıklık matrisi örneğini göstermektedir.

Tablo 4.1: Confüzyon matrisi.

Confüzyon matrisi		Tahmin edilen	
		Negatif	Pozitif
Gerçek	Negatif	TN	FP
	Pozitif	FN	TP

Aşağıdaki performans ölçümleri, hesaplamalarında karışıklık matrisinin değerlerini kullanmaktadır. Sınıflama Doğruluğu En basit performans ölçüsü doğruluktur. Algoritmanın genel etkinliği, doğru etiketlemenin tüm sınıflandırmalara bölünmesiyle hesaplanır. Negatif vakaların sayısı pozitif vakaların sayısından çok fazla olduğunda, belirlenen doğruluk yeterli bir performans ölçütü olmayabilir.

4.3 Sınıflandırıcı Yöntemi 194 Kişi İçin

4.3.1 Genelleştirilmiş Regresyon Sinir Ağı

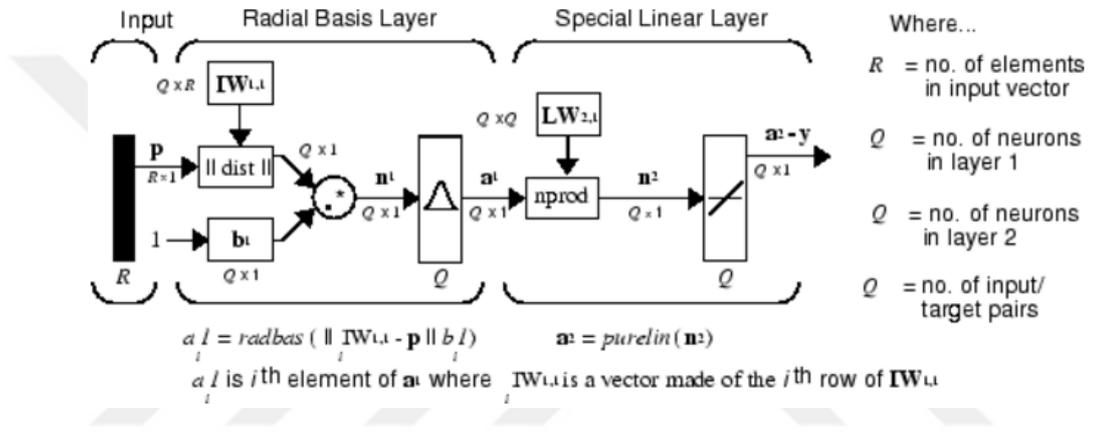
Bu tezde, simülasyon sonucu için MATLAB programlama dilleri kullanılmıştır. MATLAB akademik yazılımdır ve bu nedenle bu iş için kullanılır. Bu

aşamada 148 kişi sağlam, 46 kişi hasta olmak üzere toplam 194 kişi kullandık.

Genelleştirilmiş regresyon sinir ağı (GRSA), fonksiyonu yaklaşık olarak bulmak için sıklıkla kullanılır. Bir yatak katmanı ve özel bir doğrusal katman.

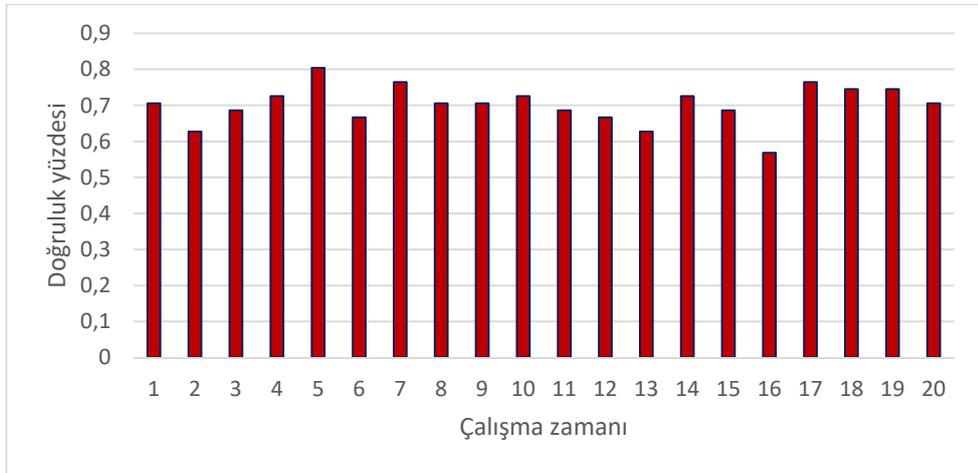
GRNN mimarisi aşağıda gösterilmektedir. Bu radyal taban ağına benzer, ancak ikinci katman biraz farklıdır.

Sadece GRNN değişen ihtiyaçlar belirlendi. Yaygın, kaygan, işlev ne kadar büyükse. Yerine çok yakın, gelişme girdi vektörleri arasındaki tipik mesafeden daha az. Yerinde daha sorunsuz, daha geniş bir yayılma kullanın. Büyükanne eğitim için kullanılır.



Şekil 4.1: Genelleştirilmiş regresyon sinir ağı mimarisi [35]

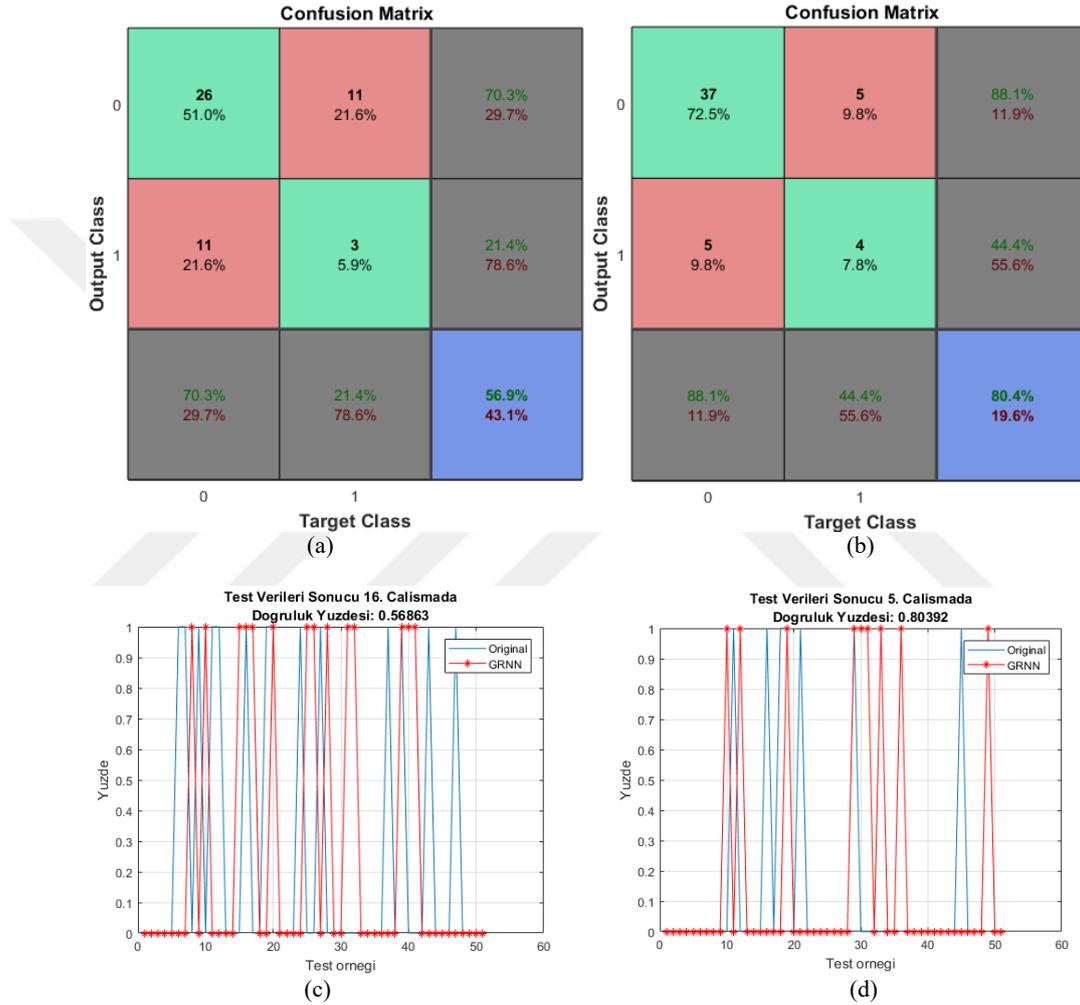
Bu aşamada yaptığımız programı 20 kez çalıştırdık ve her çalıştırdığımızda sonuçları kaydettik. Her çalıştırdığımızda farklı farklı sonuçlar elde ettik. Bu sonuçlar şekil 4.2’de gösterilmektedir.



Şekil 4.2: Doğruluk yüzdesi çalışma zamanına göre.

Resimde gösterildiği gibi ortalama oranı %70.98 elde edildi.

GRNN'nin karışıklık matrisi şekil 4.3'te gösterilmiştir. Bu şekilde görüldüğü gibi doğruluk %80.4'dür. Ayrıca, bu şekilde yaklaşık 37 örnek yaklaşık tekrarlamayan olarak sınıflandırılmıştır ve 4 örnek doğru olarak nüks olarak sınıflandırılmıştır. Yaklaşık 5 numune yanlışlıkla tekrarlamayan olarak sınıflandırılır ve 5 numune yanlışlıkla nüks olarak sınıflandırılır.

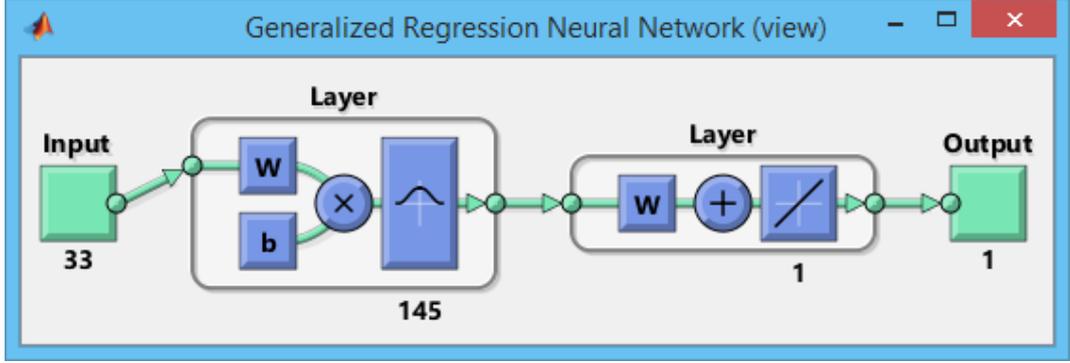


Şekil 4.3: GRNN için karışıklık matrisi, a) en az doğruluk 16'ncı çalışmada, b) en yüksek değer 5. çalışmada, c) test verileri sonucu düşük değer, d) test verileri sonucu yüksek değer.

Bu şekilde gösterildiği gibi toplam 51 kişi test için seçildi ve 42 hastadan 37 kişi doğru ve 5 kişi yanlış bulundu ve 9 kişi sağlam datadan 4 doğru ve 5 kişi yanlış bulundu. Sonuçta %80.4 doğruluk oranı tespit edildi.

Test verileri sonucu şekil 4.3-c gösterilmektedir. Bu şekil test verileri teker teker sonuçlarını gösteriyor.

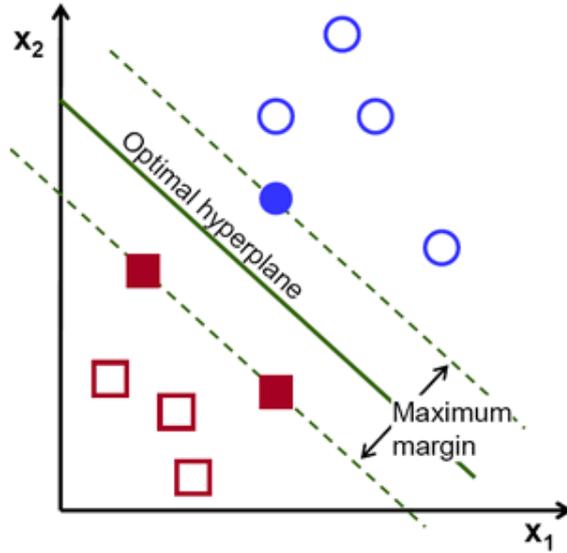
Önerilen Generalized Regression Neural network modeli şekil 4.4'te gösterilmektedir.



Şekil 4.4: Önerilen generalized regression neural network.

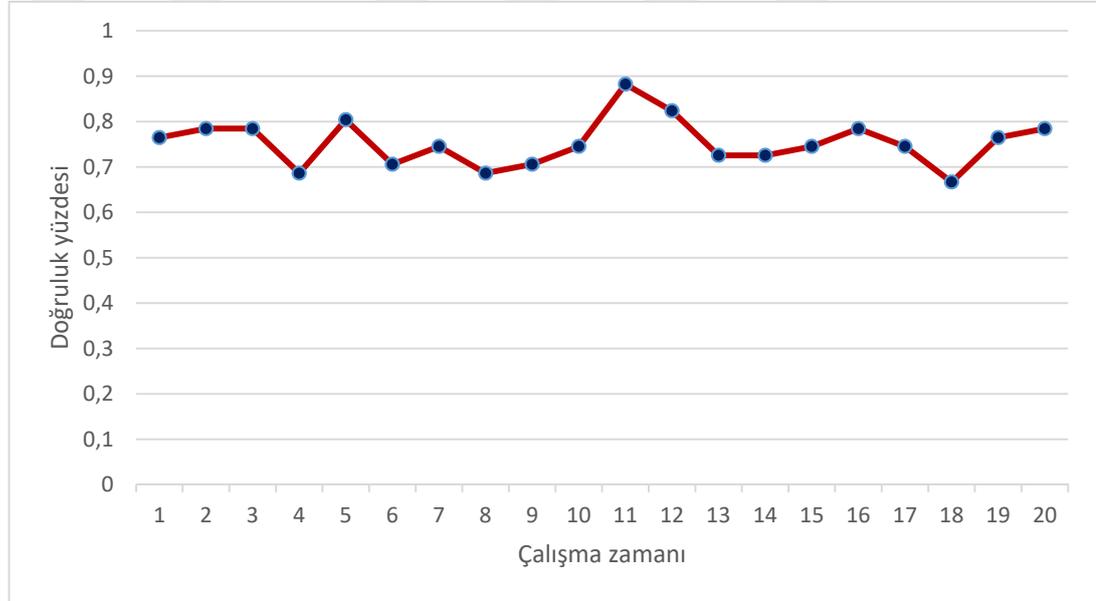
4.3.2 Destek Vektör Makinesi

Bir destek vektör makinesi (SVM) sınıflandırıcısı tanımlanmış bir hiperplane tarafından resmi olarak ayrılmış bir diferansiyel. Başka bir deyişle, etiketin verisine (öğretmen denetimi) göre, algoritma yeni örnekleri sınıflandırarak en uygun bir hiper düzlemi çıkarır. Hiperplane anlamda istenen mi? Şu basit sorunu düşünelim: iki kategoriden birine ait bir koleksiyon için bir 2D-algı algılaması, düz bir çizgi bulmak. SVM'nin bazı örnekleri aşağıdaki şekilde gösterilmiştir.



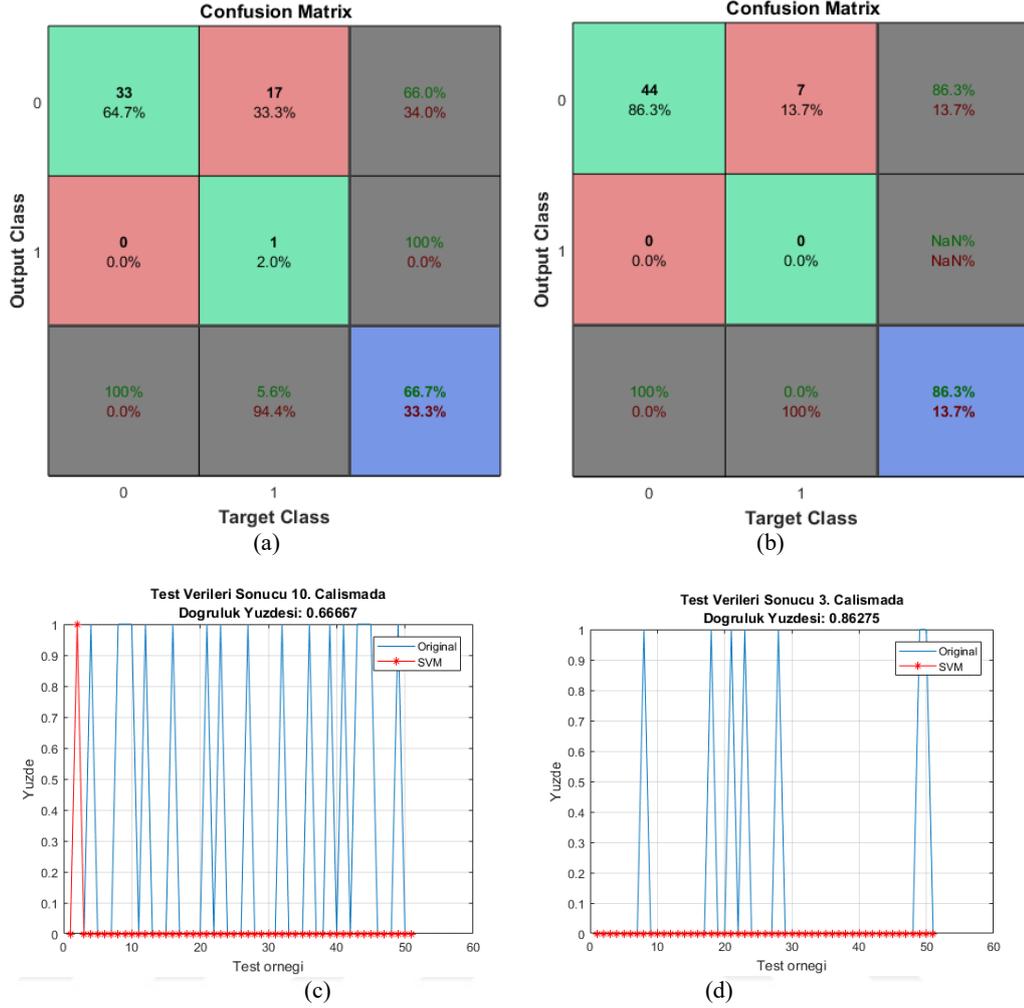
Şekil 4.5: Verilerin sınıflandırılması için SVM [39]

Bu çalışmada Wisconsin Madison Üniversitesi, Klinik Bilimler Merkezi'nin meme kanseri veri tabanı [40] kullanılmıştır. Veri tabanında, ince iğne aspirasyon yöntemi ile elde edilmiş toplam 699 örnek bulunmaktadır. Bunlardan 241'i kötü huylu (kanser), 458'i de iyi huylu tümör örnekleridir. Her örnek uzman tarafından değerlendirilerek dokuz farklı fiziksel özelliği 1-10 arasında puanlanmıştır. Bunlar örneklerin öznitelik vektörünü oluşturmaktadır. Örnekler, kanser olup olmama durumuna göre uzman tarafından ayrıca etiketlenmiştir. Öznitelikler şu şekilde açıklanmaktadır: Hücre topluluğu kalınlığı, topluluktaki hücre boyutlarının eşitliği, hücre şekillerinin aynı olması, hücrelerin çevreye yapışma miktarı, tek epitel hücre boyutu, hücre çekirdeğinin çevre sitoplazmaya oranı, kromatin dağılımı, normal çekirdekçik, mitoz sıklığı.



Şekil 4.6: Ortalama 75.29 elde edildi.

Bu tezde, tekrarlayan ve tekrarlamayan tahmini için doğrusal destek vektör makinesi kullanılmıştır. Doğruluk oranı 10'uncu çalışmada %66.66 ve 3'üncü çalışmada %86.27 elde edildi. SVM'nin karışıklık matrisi Şekil 4.7'da gösterilmiştir. Bu şekilde görüldüğü gibi doğruluk %86.3'dir. Ayrıca, bu şekilde yaklaşık 44 örnek doğru sınıflandırılmış ve 7 örnek doğru sınıflandırılmamıştır.



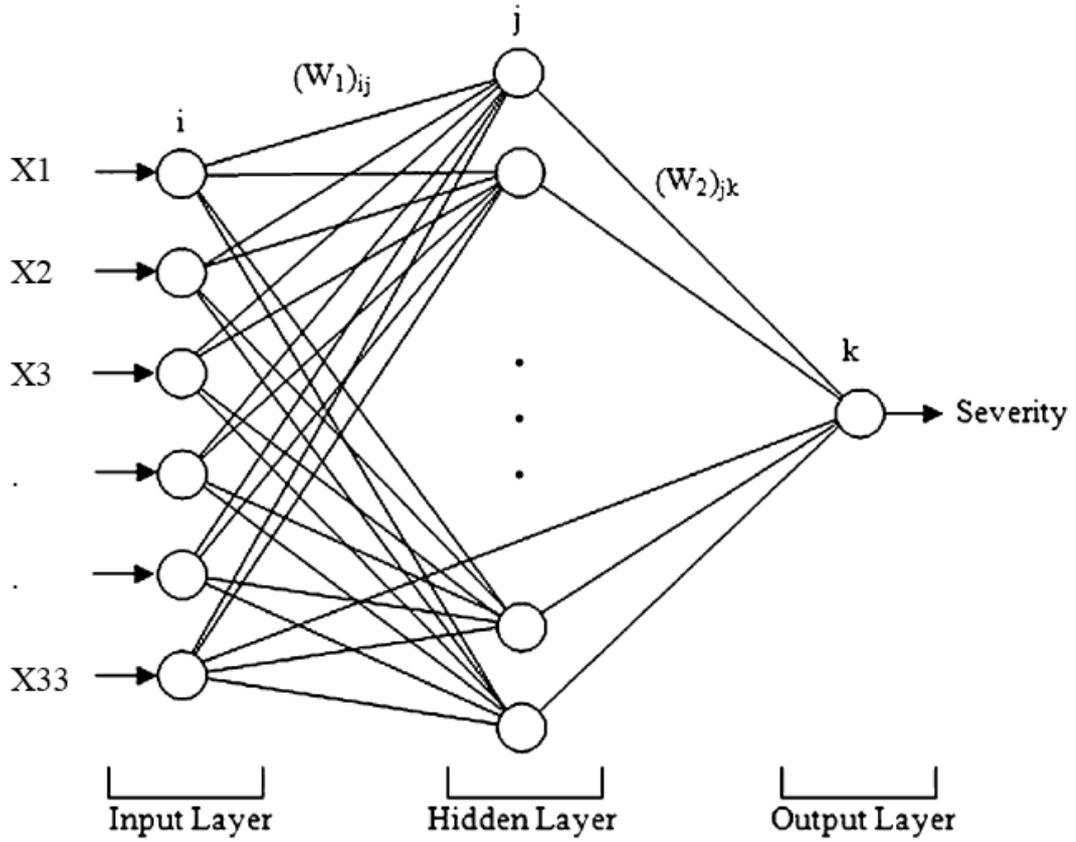
Şekil 4.7: SVM için karışıklık matrisi, a) en az doğruluk 10'uncü çalışmada, b) en yüksek değer 3. Çalışmada, c) test verileri sonucu düşük değer, d) test verileri sonucu yüksek değer.

Bu şekilde gösterildiği gibi toplam 51 kişi test için seçildi ve 51 hastadan 44 kişi doğru ve 7 kişi yanlış bulundu ve 0 kişi sağlam datadan 0 doğru ve 0 kişi yanlış bulundu. Sonuçta %86.3 doğruluk oranı tespit edildi.

4.3.3 Çok Katlı Perceptron

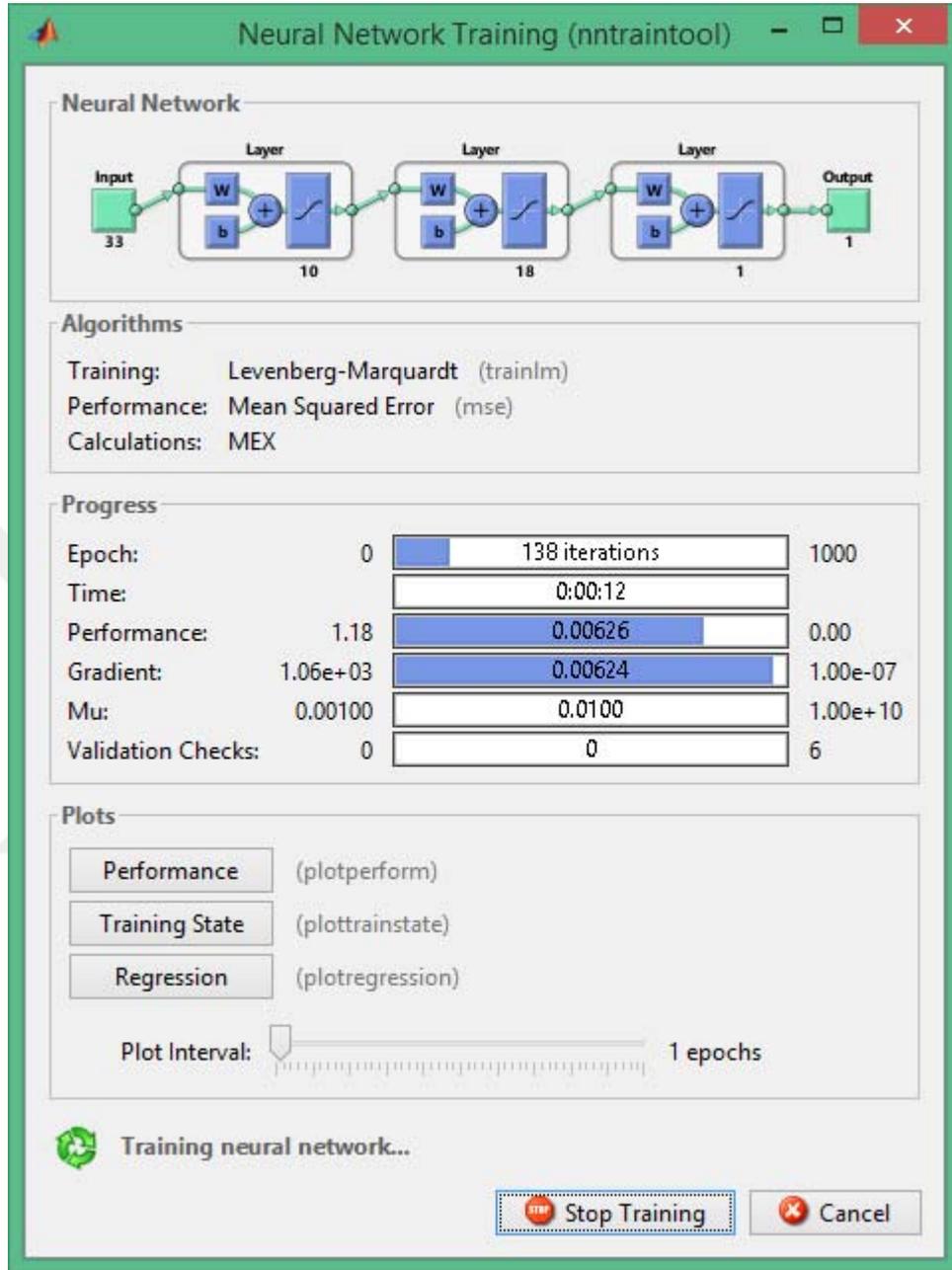
Çok Katlı Perceptron (MLP), girdi katmanına girilen verilerin gizli katman (lar) yoluyla çıktı katmanına bir yönde yayılım yaptığı bir ileri-geri YSA'dır. Gizli ve çıktı katmanları, perceptronlar olarak adlandırılan tek birimlerden oluşur. MLP'deki her algılayıcı, önce ağırlıklandırılan ve daha sonra birlikte eklenen bir girdi seti alır. Elde edilen değer kombine girdileri uygun çıktı tepkisine eşleyecek bir aktivasyon işlevini tetiklemek için kullanılır. Eğitim verileri ve simülasyon verileri için tren ve sim işlevi kullanılır.

Bu tezde, tekrarlayan ve tekrarlamayan tahmini sınıflandırmada Levenberg Marquat eğitim algoritması ile geri yayılım algoritması ile Çok katmanlı perceptron kullanılmıştır. 2 gizli katman için farklı nöronlar kullanılır ve sonuç kaydedilir. Sonraki bölümde sonuç gösterilir. Bu çalışmada, bir gizli katmanın senaryosu kullanılmış ve test edilmiştir. İki gizli katman için toplam 360 senaryo kullanılmıştır.



Şekil 4.8: YSA'nın yapısı.

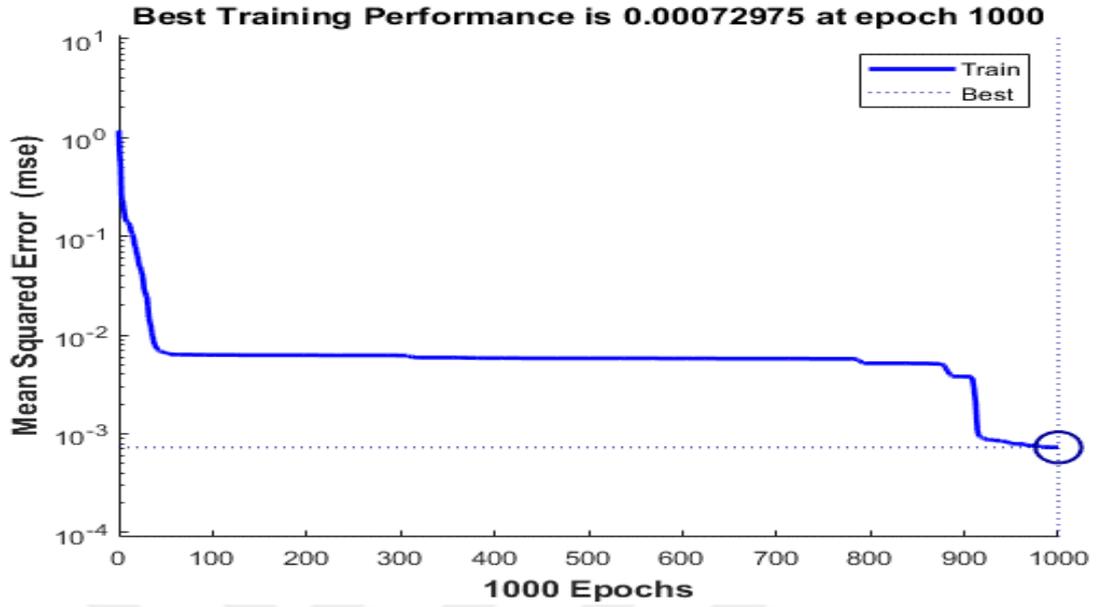
Kullandığımız YSA'nın modeli şekil 4.9'da gösterilmektedir.



Şekil 4.9: Kullanıldığı YSA'nın modeli.

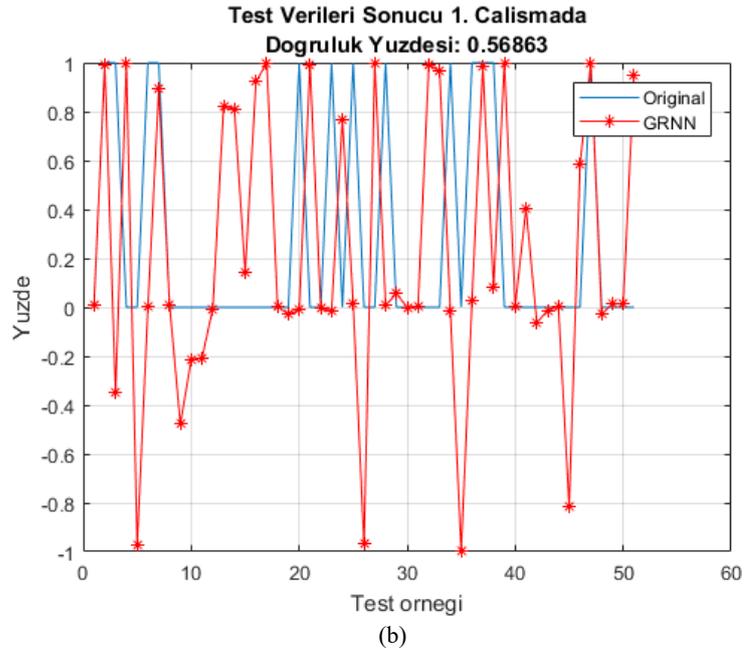
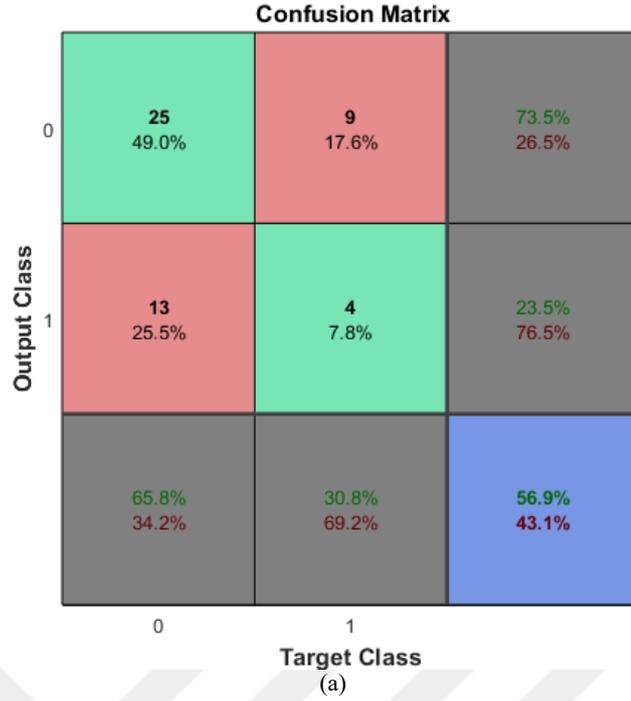
Şekil 4.9 YSA'larının çalışma toolboxını gösteriyor. Bu şekilde gösterildiği gibi 33tane girdi kullandık. İlk gizli katman için 10 tane nöron ve ikinci katman için 18 tane nöron ve çıktıda 1 tane nöron kullandık. Eğitim için Levenberg-Marquardt algoritmasını kullandık. İterasyon sayısı burada 1000 seçildi ve 138'inci iterasyonda eğitim tamamlandı. Eğitim süresi 12 saniye olarak sistemde kaydedildi. Gradyan hatası 0.00626 bulundu.

Eđitim iin hatanın performansı Őekil 4.10'da gsterilmektedir.



Őekil 4.10: Eđitim hatanın performansı.

MLP'nin, birinci gizli katmandaki 10 ve ikinci gizli katmanda 18 nron ve en sondaki katmanda bir nronun bulunduđu karıŐıklık matrisi Őekil 4.11'de gsterilmiŐtir. Bu Őekilde grldđ gibi, dođruluk %56.9'dir. Bu Őekilde, yaklaşık 25 rnek yaklaşık tekrarlamayan olarak sınıflandırılmıŐ ve 4 rnek dođru olarak tekrarlamaya sınıflandırılmıŐtir. Yaklaşık 9 numune yanlıŐlıkla tekrarlamayan olarak sınıflandırılır ve 13 numune yanlıŐlıkla nks olarak sınıflandırılır.



Şekil 4.11: YSA için karışıklık matrisi, a) doğruluk oranı 56.9, b) test verileri sonucu.

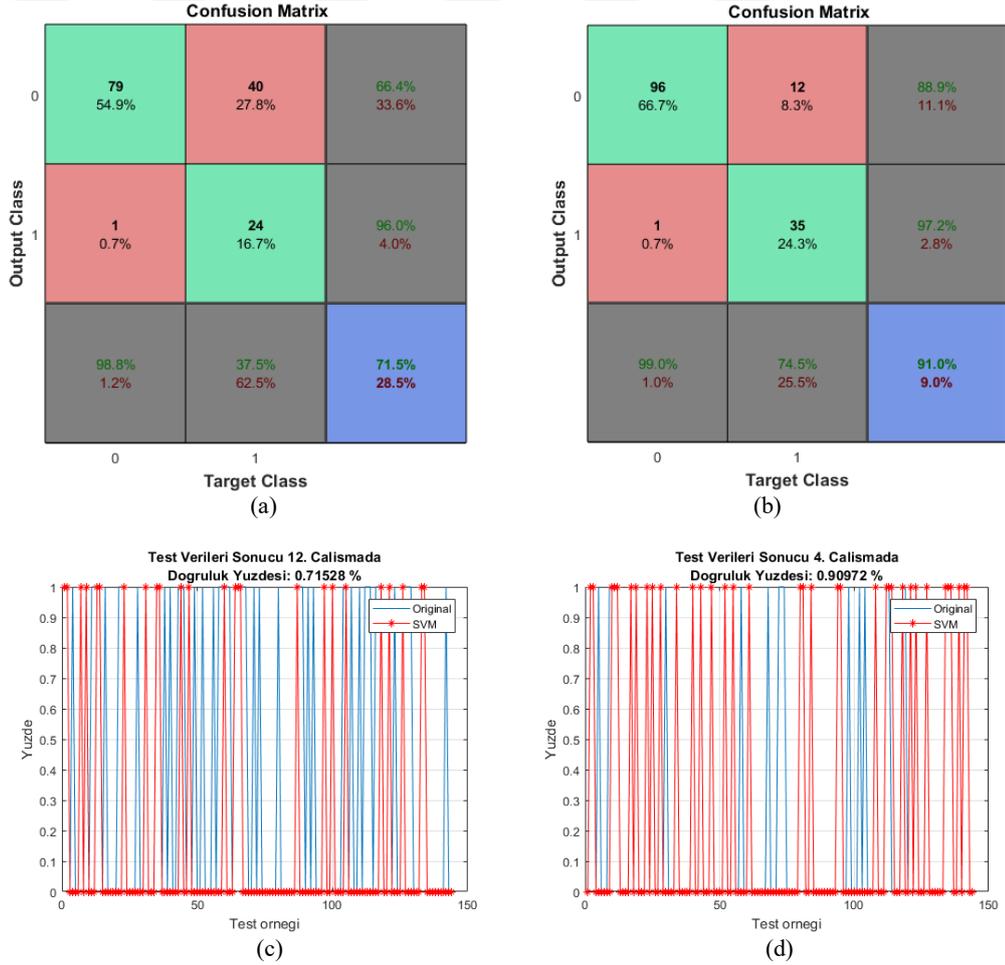
Bu şekilde gösterildiği gibi toplam 51 kişi test için seçildi ve 34 hastadan 25 kişi doğru ve 9 kişi yanlış bulundu ve 17 kişi sağlam datadan 4 doğru ve 13 kişi yanlış bulundu. Sonuçta %56.90 doğruluk oranı tespit edildi.

4.4 Sınıflandırıcı Yöntemi 569 Kişi İçin

Bu aşamada her 3 metodu 569 data için uyguladık. Bu datalardan %25 test için ve %75'i eğitim dataları için seçtik. Yaklaşık 144 sayıda test için seçildi ve 427 sayıda eğitim için seçildi.

4.4.1 SVM Sonuçları

Doğruluk oranı 12'uncü çalışmada %71.5 ve 3'üncü çalışmada %91.0 elde edildi. SVM'nin karışıklık matrisi Şekil 4.12'da gösterilmiştir. Bu şekilde görüldüğü gibi doğruluk %91.0'dir. Ayrıca, bu şekilde yaklaşık 96 örnek doğru sınıflandırılmış ve 35 örnek doğru sınıflandırılmamıştır.

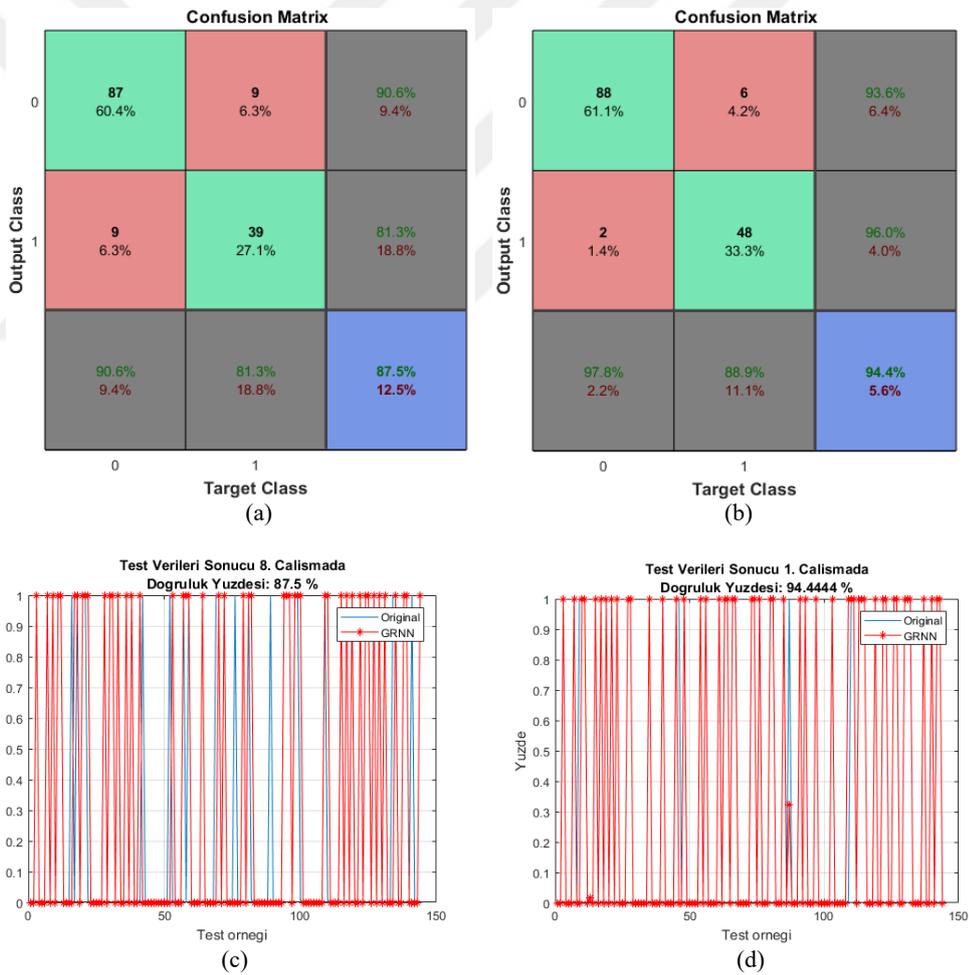


Şekil 4.12: SVM için karışıklık matrisi, a) en az doğruluk 12'ncü çalışmada, b) en yüksek değer 4. çalışmada, c) test verileri sonucu düşük değer, d) test verileri sonucu yüksek değer.

Bu şekilde gösterildiği gibi toplam 144 kişi test için seçildi ve 108 hastadan 96 kişi doğru ve 12 kişi yanlış bulundu ve 36 kişi sağlam datadan 35 doğru ve 1 kişi yanlış bulundu. Sonuçta %91.0 doğruluk oranı tespit edildi.

4.4.2 GRNN Sonuçları

GRNN'nin karışıklık matrisi şekil 4.13'te gösterilmiştir. Bu şekilde görüldüğü gibi doğruluk %94.4'dur. Ayrıca, bu şekilde yaklaşık 88 örnek yaklaşık tekrarlamayan olarak sınıflandırılmıştır ve 48 örnek doğru olarak nüks olarak sınıflandırılmıştır. Yaklaşık 6 numune yanlışlıkla tekrarlamayan olarak sınıflandırılır ve 2 numune yanlışlıkla nüks olarak sınıflandırılır.

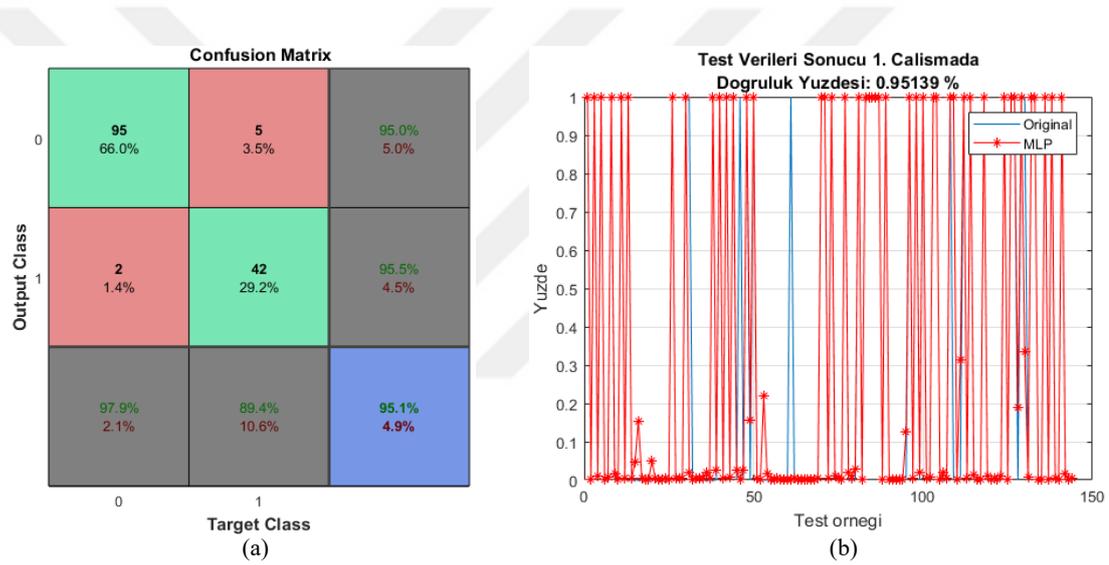


Şekil 4.13: GRNN için karışıklık matrisi, a) en az doğruluk 8'inci çalışmada, b) en yüksek değer 1. çalışmada, c) test verileri sonucu düşük değer, d) test verileri sonucu yüksek değer.

Bu şekilde gösterildiği gibi toplam 144 kişi test için seçildi ve 94 hastadan 88 kişi doğru ve 6 kişi yanlış bulundu ve 50 kişi sağlam datadan 48 doğru ve 2 kişi yanlış bulundu. Sonuçta %94.4 doğruluk oranı tespit edildi.

4.4.3 MLP Algoritması

MLP sonucu şekil 4.14'te gösterilmektedir. Bu şekilde görüldüğü gibi, doğruluk %95.13'dir. Bu şekilde, yaklaşık 95 örnek yaklaşık tekrarlamayan olarak sınıflandırılmış ve 42 örnek doğru olarak tekrarlamaya sınıflandırılmıştır. Yaklaşık 5 numune yanlışlıkla tekrarlamayan olarak sınıflandırılır ve 2 numune yanlışlıkla nüks olarak sınıflandırılır.



Şekil 4.14: YSA için karışıklık matrisi, a) doğruluk oranı 95.139, b) test verileri sonucu.

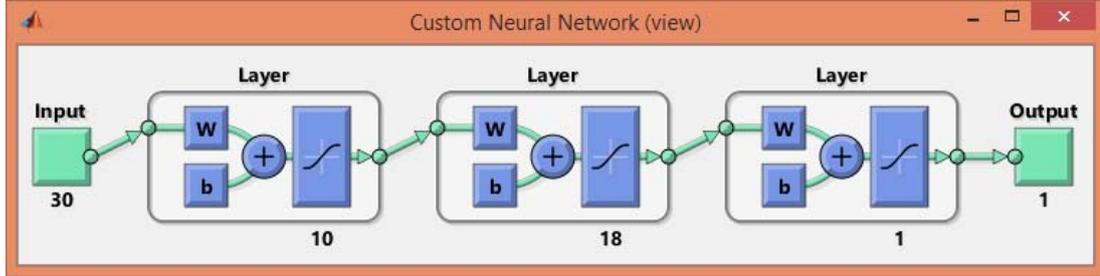
Bu şekilde gösterildiği gibi toplam 144 kişi test için seçildi ve 100 hastadan 95 kişi doğru bulundu ve 44 kişi sağlam datadan 42 doğru ve 2 kişi yanlış bulundu. Sonuçta %95.1 doğruluk oranı tespit edildi.

Bu iki veri tabanı için sonuçlar tablo 4.2'de gösterilmektedir.

Tablo 4.2: SVM, GRNN ve MLP karşılaştırılması.

Veri tabanı	SVM	GRNN	MLP
194	%80.4	%86.3	%56.9
569	%91.0	%94.4	%95.14

Tablo 4.2’de gösterildiği gibi en iyi sonuçlar 569 data için elde edildi. Veri tabanı ne kadar yüksek olursa o kadar sonuçlar iyi olacaktır. MLP için 2 katmanlı nöron kullandık ve bunun mimarisi şekil 4.15’de gösterilmektedir.



Şekil 4.15: Önerilen MLP'nin mimarisi.

Bu şekilde görüldüğü gibi 2 gizli katman kullandık ve birinci gizli katmanda 10 nöron ve ikinci gizli katmanda 18 nöron. Girdi vektöründe 30 parametre kullandık. Bu parametreler Wisconsin veri tabanından alınan değerlerdir. Eğitim için 427 data ve test için 144 data seçildi. Bu nöron sayılarını deneyerek elde ettik.

BEŞİNCİ BÖLÜM

SONUÇLAR VE ÖNERİLER

Bu bölüm tezi sonuçlandırmaktadır. Bölüm 5.1, tezin arka planını, hedeflerini ve katkılarını özetlemektedir. Bölüm 5.2 ileri çalışmalar hakkında bahsedilecektir.

5.1 Sonuç

Bu tez çalışmasında, meme kanseri verilerini sınıflandırmak ve tasarlanan sınıflandırıcının sınıflandırma doğruluğunu tahmin etmek amaçlanmıştır. Özellik seçimi, çıktı sınıflandırmasında her bir özellik potansiyelini araştırmak için tüm özellik kombinasyonlarını düşündüğü için, çalışmanın doğrudan bir amacı olmamıştır. Bu nedenle araştırma, desen sınıflandırma sistemi tasarımının iki aşamasına odaklanmıştır; mevcut veri kümesinin sınırlamaları dikkate alınarak sınıflandırıcı tasarımı ve sınıflandırma hata tahmini - küçük örneklem büyüklüğü (WPDC) ve yüksek karmaşıklık.

Desen sınıflandırıcısını tasarlama aşamaları şunlardır:

1. Özellik seçimi (tüm özellikleri kullanın)
2. Sınıflandırıcı Tasarımı
3. (Küçük örneklem büyüklüğü ve yüksek veri karmaşıklığı göz önüne alındığında)
4. Sınıflandırma Hata Tahmini
5. (Küçük örneklem büyüklüğünü göz önünde bulundurarak)

Bu amaçlara ulaşmak için kullanılan yöntemler ve simülasyon sonuçları önceki bölümlerde açıklanmıştır. ANN, GRNN, MLP ve SVM bu çalışmanın süresi boyunca mevcut tüm belirteçlere ve yüksek doğruluğa ulaşılmasına yönelik kriterlere dayanarak tasarlanmış ve geliştirilmiştir. Bu tezde geri dönen meme kanseri

incelendi. Yine de önceki arařtırmalar, dođrulama için öngörücü belirteçlerin bir kombinasyonu üzerinde anlaşma eksikliđini göstermektedir.

Küçük örneklem büyüklüğü ve karmařık doğrusal olmayan girdi-çıkıı ilişkisi, meme kanserinde metastazı öngörmeye yaygın problemlerdir. İnvaziv olmayan bir şekilde elde edilen birçok biyolojik belirteç meme kanserinde nodal tutulumun durumu ile ilişkili olduđu tespit edildiđinden, invaziv olarak edinilmeyen verileri kullanarak hastaları sınıflandıran bir sınıflandırıcı tasarlamak avantajlıdır. Bu amaçla, tasarlanan sınıflayıcı karmařık doğrusal olmayan giriş-çıkıı ilişkisini yakalayabilmeli ve mevcut verilere aşırı uymayı önlemelidir. Bu, mevcut tüm vakalarla sınıflandırıcıyı tasarlamayı ve daha sonra yeniden örnekleme yöntemi kullanarak tasarlanan sınıflandırıcının genelleme kabiliyetini deđerlendirmeyi gerektirir.

Bu çalışmada, ANN, GRNN ve SVM meme kanseri hastalarında bazı bimoleküler belirteçlere ve bunların farklı kombinasyonlarına dayalı nodal tutulum durumunu öngörmek için bir platform olarak kullanılmıřtır. Bu uygulama için ANN sonuçlarına kıyasla bir karşılařtırma yöntemi olarak MLP sinir ađı da tasarlanmıřtır.

Bu tezde iki farklı veri tabanı kullandık. İlk veri tabanında 194 kiři vardır ve bu kiřilerden 148 kiři sađlam ve 46 kiři hastadır. İkinci veri tabanında 569 kiři vardır ve bu kiřilerden 212 sađlam ve 357 hastadır.

Tablo 5.1: Bizim yöntemimiz ve diđer çalışmaların karşılařtırılması.

Referans	Dođruluk
Guruprasad Bhat ve arkadaşları [27]	%92.8571
FANG Yi ve arkadaşları [28]	%94.9
T. Kıyan ve arkadaşları [29]	%95.74
Bizim yaptığımız yöntem	%95.14

Tablo 5.1'de görüldü gibi bizim yaptığımız yöntem diđerlerine göre daha iyi sonuç vermiřtir. Bu dođruluk oranını başka veri tabanlarını kullanarak daha iyi sonuçlar elde edebiliriz. Başka veri tabanlarında daha fazla kiři olursa daha iyi sonuçlar elde edilebilir.

5.2 İleri Çalışmalar

İleride bu veri tabanlarını farklı farklı metotlarla deneyebiliriz. Bu metotlar logistic regresion ve K-NN ve farklı sınıflandırma yöntemleri olabilir. Bu çalıştığımız yöntemleri cihazlarda kullanarak hastanelerde uygulayabiliriz. Bunun için bu yaptığımız yöntemleri MATLAB kodundan C++ diline çevirip cihazlarda uygulayabiliriz.



KAYNAKÇA

- [1] Türkan Yiğitbaşı, "Pers Kraliçesi Atossa ile geçmişten geleceğe kanserin yolculuğu," *Sağlık Düşüncesi ve Tıp Kültürü Platformu*, 28 AĞUSTOS 2014.
- [2] R. A. Smith, D. Saslow, K. A. Sawyer, W. Burke, M. E. Costanza, W. Evans, *et al.*, "American Cancer Society guidelines for breast cancer screening: update 2003," *CA: a cancer journal for clinicians*, vol. 53, pp. 141-169, 2003.
- [3] L. Çelik and R. Çubuk, "Meme Manyetik Rezonans Görüntüleme: Nasıl, Niçin, Ne Zaman, Kime," *Klinik Gelişim* 23, vol. 2, pp. 6-10, 2010.
- [4] U. Bayrak, E. Gram, E. Mengeş, Z. G. Okumuş, H. C. Sayar, E. Skrijelj, *et al.*, "Üniversite öğrencilerinin sağlıkla ilgili alışkanlıklar ve kanser konusundaki bilgi ve tutumları," *Dokuz Eylül Üniversitesi Tıp Fakültesi Dergisi*, vol. 24, pp. 95-104, 2010.
- [5] R. Liao, T. Wan, and Z. Qin, "Classification of benign and malignant breast tumors in ultrasound images based on multiple sonographic and textural features," in *Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2011 International Conference on*, 2011, pp. 71-74.
- [6] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener, "Histopathological image analysis: A review," *IEEE reviews in biomedical engineering*, vol. 2, pp. 147-171, 2009.
- [7] A. Brook, R. El-Yaniv, E. Isler, R. Kimmel, R. Meir, and D. Peleg, "Breast cancer diagnosis from biopsy images using generic features and SVMs," *IEEE Transactions on Information Technology in Biomedicine*, 2006.
- [8] S. S. Haykin, *Neural networks: a comprehensive foundation*: Tsinghua University Press, 2001.
- [9] S. Mittal, H. Kaur, N. Gautam, and A. K. Mantha, "Biosensors for breast cancer diagnosis: A review of bioreceptors, biotransducers and signal amplification strategies," *Biosensors and Bioelectronics*, vol. 88, pp. 217-231, 2017.

- [10] I. Goswami, S. Coutermarsh-Ott, R. G. Morrison, I. C. Allen, R. V. Davalos, S. S. Verbridge, *et al.*, "Irreversible electroporation inhibits pro-cancer inflammatory signaling in triple negative breast cancer cells," *Bioelectrochemistry*, vol. 113, pp. 42-50, 2017.
- [11] D. M. Cvetković, M. N. Živanović, M. G. Milutinović, T. R. Djukić, M. D. Radović, A. M. Cvetković, *et al.*, "Real-time monitoring of cytotoxic effects of electroporation on breast and colon cancer cell lines," *Bioelectrochemistry*, vol. 113, pp. 85-94, 2017.
- [12] T. Wan, J. Cao, J. Chen, and Z. Qin, "Automated grading of breast cancer histopathology using cascaded ensemble with combination of multi-level image features," *Neurocomputing*, vol. 229, pp. 34-44, 2017.
- [13] J. Xu, X. Luo, G. Wang, H. Gilmore, and A. Madabhushi, "A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images," *Neurocomputing*, vol. 191, pp. 214-223, 2016.
- [14] S. M. Naushad, M. J. Ramaiah, M. Pavithrakumari, J. Jayapriya, T. Hussain, S. A. Alrokayan, *et al.*, "Artificial neural network-based exploration of gene-nutrient interactions in folate and xenobiotic metabolic pathways that modulate susceptibility to breast cancer," *Gene*, vol. 580, pp. 159-168, 2016.
- [15] J. Arevalo, F. A. González, R. Ramos-Pollán, J. L. Oliveira, and M. A. G. Lopez, "Representation learning for mammography mass lesion classification with convolutional neural networks," *Computer methods and programs in biomedicine*, vol. 127, pp. 248-257, 2016.
- [16] W. Sun, T.-L. B. Tseng, J. Zhang, and W. Qian, "Enhancing deep convolutional neural network scheme for breast cancer diagnosis with unlabeled data," *Computerized Medical Imaging and Graphics*, vol. 57, pp. 4-9, 2017.
- [17] A. Bhardwaj and A. Tiwari, "Breast cancer diagnosis using genetically optimized neural network model," *Expert Systems with Applications*, vol. 42, pp. 4611-4620, 2015.
- [18] K. Thangavel, M. Karnan, R. Sivakumar, and A. K. Mohideen, "Automatic detection of microcalcification in mammograms-a review," *International Journal on Graphics Vision and Image Processing*, vol. 5, pp. 31-61, 2005.
- [19] R. Mekle, A. Laine, S. Smith, C. Singer, T. Koenigsberg, and M. Brown, "Evaluation of a multiscale enhancement protocol for digital mammography [4119-124]," in *Proceedings-Spie The International Society For Optical Engineering*, 2000, pp. 1038-1049.

- [20] A. Papadopoulos, D. I. Fotiadis, and L. Costaridou, "Improvement of microcalcification cluster detection in mammography utilizing image enhancement techniques," *Computers in biology and medicine*, vol. 38, pp. 1045-1055, 2008.
- [21] R. Garg, B. Mittal, and S. Garg, "Histogram equalization techniques for image enhancement," *Int. J. Electron. Commun. Technol*, vol. 2, pp. 107-111, 2011.
- [22] I. V. Lakshmi and G. Krishnaveni, "Performance Assessment by using SVM and ANN for Breast cancer Mammography image classification," 2017.
- [23] K. U. Rani, "Parallel approach for diagnosis of breast cancer using neural network technique," *International Journal of Computer Applications*, vol. 10, pp. 1-5, 2010.
- [24] V. Bevilacqua, G. Mastronardi, and F. Menolascina, "Hybrid Data Analysis Methods and Artificial Neural Network Design in Breast Cancer Diagnosis: IDEST Experience," in *Computational Intelligence for Modelling, Control and Automation, 2005 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference on*, 2005, pp. 373-378.
- [25] W. Land and E. Verheggen, "Experiments using an evolutionary programmed neural network with adaptive boosting for computer aided diagnosis of breast cancer," in *Soft Computing in Industrial Applications, 2003. SMCia/03. Proceedings of the 2003 IEEE International Workshop on*, 2003, pp. 167-172.
- [26] V. N. Chuneekar and H. P. Ambulgekar, "Approach of neural network to diagnose breast cancer on three different data set," in *Advances in Recent Technologies in Communication and Computing, 2009. ARTCom'09. International Conference on*, 2009, pp. 893-895.
- [27] G. Bhat, V. G. Biradar, H. Sarojadevi, and N. Nalini, "Artificial neural network based cancer cell classification," *Computer Engineering and Intelligent Systems*, vol. 3, pp. 7-16, 2012.
- [28] F. Yi, Z. Chongxun, P. Chen, and L. Li, "White blood cell image segmentation using on-line trained neural network," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, 2006, pp. 6476-6479.
- [29] T. Kiyan and T. Yildirim, "Breast cancer diagnosis using statistical neural networks," *IU-Journal of Electrical & Electronics Engineering*, vol. 4, pp. 1149-1153, 2004.

- [30] Y. Hu, K. Ashenayi, R. Veltri, G. O'Dowd, G. Miller, R. Hurst, *et al.*, "A comparison of neural network and fuzzy c-means methods in bladder cancer cell classification," in *Neural Networks, 1994. IEEE World Congress on Computational Intelligence., 1994 IEEE International Conference on*, 1994, pp. 3461-3466.
- [31] W. H. Wolberg, W. N. Street, D. M. Heisey, and O. L. Mangasarian, "Computerized breast cancer diagnosis and prognosis from fine-needle aspirates," *Archives of Surgery*, vol. 130, pp. 511-516, 1995.
- [32] N. V. Vasif, "Yapay Zeka; Problemler-Yöntemler-Algorithm," *Seçkin Yayıncılık*, vol. Kitap, 2005.
- [33] E. Taş, "Yapay sinir ağlarında momentumlu dik iniş ve eşlenik gradyan eğitim algoritmalarının karşılaştırılması," *Yüksek Lisans Tezi, Anadolu Üniversitesi Fen Bilimleri Enstitüsü, Eskişehir*, pp. 5-61, 2005.
- [34] S. Haykin and N. Network, "A comprehensive foundation," *Neural Networks*, vol. 2, p. 41, 2004.
- [35] H. Demuth, M. Beale, and M. Hagan, "Neural network toolbox™ 6," *User's guide*, pp. 37-55, 2008.
- [36] E. Binici, "Java ile yapay zeka mekanizmasına sahip bir ağ yönetim sistemi geliştirilmesi," Ege Üniversitesi, 2006.
- [37] L. Fausett and L. Fausett, *Fundamentals of neural networks: architectures, algorithms, and applications*: Prentice-Hall, 1994.
- [38] I. Saritas, I. A. Ozkan, and I. U. Sert, "Prognosis of prostate cancer by artificial neural networks," *Expert Systems with Applications*, vol. 37, pp. 6646-6650, 2010.
- [39] K. R. Foster, R. Koprowski, and J. D. Skufca, "Machine learning, medical diagnosis, and biomedical engineering research-commentary," *Biomedical engineering online*, vol. 13, p. 94, 2014.
- [40] M. Ferris and O. Mangasarian, "Breast-Cancer Diagnosis Via Linear-Programming," vol. 2, ed: IEEE Computer Soc 10662 Los Vaqueros Circle, Po Box 3014, Los Alamitos, CA 90720-1264, 1995, pp. 70-71.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : OMER TAHA AHMED AL-KHALIDI
Uyruđu : Irak
Dođum Yeri ve Tarihi : Diali 14.09.1989
Medeni Hali : Bekler
Adres : Birlik mh. 450 cad. 68/1 Çankaya – Ankara
E-Posta Adresi : omertaha89@yahoo.com
İletişim (Telefon) : 05378731988

EĐİTİM

Lise : khanaqin lisesi / Irak - Diali
Lisans : Bilgisayar Mühendisliđi / Irak - Diali
Yüksek Lisans : Türk Hava Kurumu Üniversitesi

MESLEKİ DENEYİM

2011- : Bilgisayar Mühendisi

DİLLER

Arapça, Türkçe, İngilizce, Kürtçe