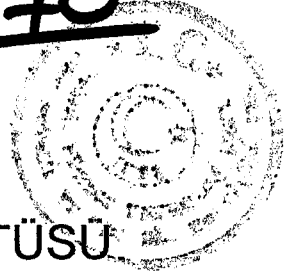


171070



T.C.

GEBZE YÜKSEK TEKNOLOJİ ENSTİTÜSÜ
MÜHENDİSLİK VE FEN BİLİMLERİ ENSTİTÜSÜ

GENETİK HÜCRESEL YAPAY SİNİR
AĞLARI VE JEOFİZİK UYGULAMALARI

Erdem BİLGİLİ
DOKTORA TEZİ

ELEKTRONİK MÜHENDİSLİĞİ
ANABİLİM DALI

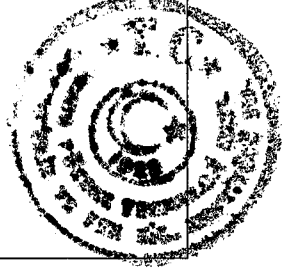
TEZ DANIŞMANI
Prof. Dr. Ali ALKUMRU

GEBZE
2005



**GEBZE YÜKSEK TEKNOLOJİ
ENSTİTÜSÜ**

DOKTORA JÜRİ ONAY FORMU



99102314 numaralı Erdem Bilgili'nin tez çalışması, G.Y.T.E. Mühendislik ve Fen Bilimleri Enstitüsü Yönetim Kurulu'nun 24.06.2005 Tarih ve 2005/23 sayılı kararıyla oluşturulan jüri tarafından ...Elektronik.....Mühendisliği Anabilim Dalında DOKTORA tezi olarak kabul edilmiştir.

JÜRİ

ÜYE

(TEZ DANIŞMANI) :Prof. Dr. Ali ALKUMRU

ÜYE

:Prof. Dr. Osman Nuri UÇAN

ÜYE

:Prof. Dr. İzzet Cem GÖKNAR

ÜYE

:Yrd. Doç. Dr. Arif ERGİN

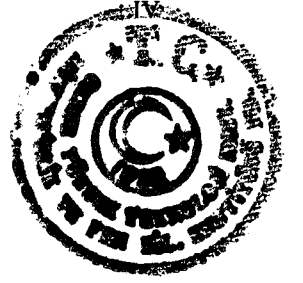
ÜYE

:Doç. Dr. Muhittin ALBORA

ONAY

G.Y.T.E. Mühendislik ve Fen Bilimleri Enstitüsü Yönetim Kurulu'nun 08.09.2005 Tarih ve 2005/32 sayılı kararı.

İMZA/MÜHÜR



ÖZET

Bu çalışmada Hücresel Yapay Sinir Ağlarında (HYSA) standart olarak kullanılan aktivasyon fonksiyonu yerine Trapezoidal Aktivasyon Fonksiyonu (TAF) adı verilen yeni bir lineer olmayan aktivasyon fonksiyonu kullanılmıştır. Yeni HYSA yapısı, XOR gibi Boolean işlemlerini tek katmanla gerçekleştirebilmekte ve aynı zamanda lineer olarak ayrıştırılamayan veri kümelerini sınıflandırabilmektedir. Trapezoidal Aktivasyon Fonksiyonlu Hücresel Yapay Sinir Ağının (TAF-HYSA) kararlılık analizini basitleştirmek için HYSA geribesleme şablonu A yerine W geribesleme matrisi tanımlanmış ve 2 boyutlu durum denklemleri tek boyuta indirginmiştir. HYSA-TAF'ın kararlılık şartları araştırılarak, var olan denge noktaları için kararlılık yeter şartı ve global asimptotik kararlılık kriteri belirlenmiştir. Analitik olarak elde edilen bu kararlılık kriteri çok sayıda sentetik giriş görüntüsü kullanılarak pratik olarak doğrulanmıştır.

Ayrıca, çok seviyeli HYSA (ÇS-HYSA) kullanılarak jeofizik verilerinin analizi için eğitici bir öğrenme algoritması geliştirilmiş ve gerçek verilere uygulanmıştır. ÇS-HYSA, piksellerin birbirleri ile komşuluk ilişkilerini kullanan şablonların optimizasyonuna dayanan stokastik bir görüntü işleme tekniğidir. Önerilen yöntemin jeofizik anomalilerinin ayrıştırılmasında ve jeofizik yapıların sınırlarının belirlenmesinde gösterdiği performans değişik gerçek verilerle gösterilmektedir. HYSA şablonlarının bulunmasında genetik algoritmalar kullanılmıştır. İlk uygulamada jeofizikte temel problemlerden biri olan gravite anomali ayrıştırma klasik yöntemiyle ilgili olarak ülkemizdeki Divriği-Dumluca maden bölgesine ait anomali verileri üzerinde çalışılmıştır. Diğer uygulamada ise Sivas-Altınyayla bölgesindeki Hitit İmparatorluğu kalıntılarına ait arkeolojik yapılara ait anomali haritasına uygulanarak buradaki yer altı yapılarının geometrik sınırlarının tespiti yapılmıştır.

SUMMARY

This study presents a Cellular Neural Network (CNN) scheme employing a new non-linear activation function, called Trapezoidal Activation Function (TAF). The new CNN structure can classify linearly non-separable data points and realize Boolean operations (including XOR) by using only a single-layer CNN. In order to simplify the stability analysis, a feedback matrix W is defined as a function of the feedback template A and 2D equations are converted to 1D equations. The stability conditions of CNN with TAF are investigated and a sufficient condition for the existence of a unique equilibrium and global asymptotic stability is derived. By processing several examples of synthetic images, the analytically derived stability condition is also confirmed.

In addition, a supervised algorithm for the evaluation of geophysical sites using Multi Level-Cellular Neural Network (ML-CNN) is introduced, developed, and then applied to real data. ML-CNN is a stochastic image processing technique that is based on template optimization using neighborhood relationships of the pixels. The separation/enhancement and border detection performance of the proposed method is evaluated by various interesting real applications. Genetic algorithm is used in optimization of CNN templates. The first application is concerned with the separation of potential field data of Dumluca iron region which is one of the rich reserves of Turkey; in this context the classical approach to gravity anomaly separation method, is one of the main problems in geophysics. The other example is the border detection of archeological ruins of the Hittite Empire in Turkey. The Hittite civilization sites located in Sivas-Altinyayla region of Turkey, are among the most important archeological sites in history; one reason among others being that written documentation was first produced by this civilization.

TEŞEKKÜR

Bu çalışma sırasında katkılarını, desteklerini ve ilgilerini esirgemeyen hocalarım Prof. Dr. Ali ALKUMRU'ya, Prof. Dr. Osman Nuri UÇAN'a, Prof. Dr. İzzet Cem GÖKNAR'a ve Doç.Dr. Muhittin ALBORA' ya şükranlarımı sunarım.

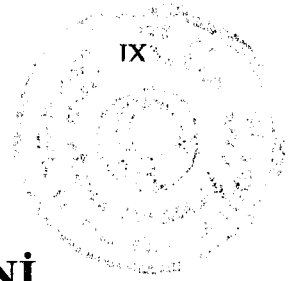
Benden manevi desteklerini esirgemeyen aileme, gösterdikleri anlayıştan dolayı teşekkürlerimi sunarım.



İÇİNDEKİLER DİZİNİ

	<u>Sayfa</u>
ÖZET	iv
SUMMARY	v
TEŞEKKÜR	vi
İÇİNDEKİLER DİZİNİ	vii
SİMGELER VE KISALTMALAR DİZİNİ	ix
ŞEKİLLER DİZİNİ	xi
ÇİZELGELER DİZİNİ	xiii
1.GİRİŞ	1
2.YAPAY SİNİR AĞLARI	5
2.1.Yapay Sinir Ağlarının Üstünlükleri	9
2.2. Basit YSA Modelleri ve Öğrenme Kuralları	10
2.3.Eğitici Öğrenme Yöntemleri	19
2.3.1.Tek Tabakalı Perceptron Ağları	20
2.3.2.Perceptron Öğrenme Kurallı	20
2.3.3.Adaline Öğrenme Kuralları	21
2.4.Geriye Yayılma (Back Propagation)	23
3.HÜCRESEL YAPAY SİNİR AĞLARI	28
3.1. Hücresel Yapay Sinir Ağının Dinamik Davranışı	32
3.2. Hücresel Yapay Sinir Ağının Kararlılığı	35
3.3. Hücresel Yapay Sinir Ağının Eğitimi	38
4. TRAPEZOİDAL AKTİVASYONLU FONKSİYONLU HYSA	40
4.1. Trapezoidal Aktivasyon Fonksiyonlu HYSA'nın Kararlılığı	43
4.1.1.Brouwer Fixed Point Teoremi	47
4.1.2.Global varlık ve teklik teoremi	48
4.1.3. Çözümün Sınırlı Olması (Sonlu Değer İçermesi) Teoremi	49
4.1.4. Lyapunov Anlamında Kararlılık	49
4.2. TAF-HYSA Uygulamaları	54
4.2.1.Kararlılık analizleri	54
4.2.2. Lineer olarak ayrıştırılamayan uygulamalar	67
5.GENETİK ALGORİTMALAR	72

5.1. Genetik Algoritmaların Temel Teoremi	73
5.2. Genetik Algoritmaların Arama Yöntemi	75
5.2.1. Başlangıç Populasyonunun Oluşturulması	76
5.2.2. Uygunluk ve Hata Fonksiyonlarının Belirlenmesi	77
5.2.3. Kopyalama (Çoğalma)	78
5.2.4. Çaprazlama	78
5.2.5. Mutasyon	79
5.3. Genetik Algoritmalarda Önemli Parametreler	80
6. HÜCRESEL YAPAY SİNİR AĞININ JEOFİZİK UYGULAMALARI	81
6.1. Hücresel Yapay Sinir Ağları İle Potansiyel Anomali Ayrıştırma	82
6.1.1. Hücresel Yapay Sinir Ağı'nın Sentetik Verilerle Eğitimi	83
6.1.2. Dumlucu Bölgesindeki Maden Yataklarının Tespiti	89
6.2. Genetik Algoritmalarla Eğitilmiş HYSA İle Magnetik Anomali Ayrıştırma	91
6.2.1. Hücresel Yapay Sinir Ağı'nın Sentetik Verilerle Eğitimi	92
6.2.2. HYSA'nın Hitit Kalıntılarına Uygulanması	98
7.SONUÇ	101
KAYNAKLAR	103
ÖZGEÇMİŞ	108
EKLER (BİLGİSAYAR PROGRAMLARI)	



SİMGELER VE KISALTMALAR DİZİNİ

A : HYSA geribesleme şablonu.

B : HYSA ileribesleme şablonu.

$a(f)$: Aktivasyon fonksiyonu.

$d^{(k)}$: k . sınıfa ait hedef deseni.

E : Hata fonksiyonu.

e : Hata işareti.

$\nabla E(w)$: Hata fonksiyonun gradienti.

F_0 : Yer magnetik alanı

$G(i, j)$: Parlaklık fonksiyonu.

I : HYSA eşik şablonu.

k : Suspektibilite.

k_0 : Genel çekim sabiti.

net_i : Hücrenin toplam giriş işareti büyüklüğü.

$p(m)$: Mutasyon olasılığı.

$p(c)$: Çaprazlama olasılığı.

r : Hata işareti.

U : HYSA giriş matrisi

x_i : Hücrenin i . girişi.

X : HYSA durum matrisi.

$x^{(k)}$: k . sınıfa ait giriş deseni.

W : HYSA geribesleme matrisi, YSA ağırlık matrisi.

w_{ij} : Ağırlık matrisi elemanı.

Y : HYSA çıkış matrisi.

y : Hücre çıkış işareti.

z_q : Saklı tabaka çıkışı (q . hücre).

α : Momentum faktörü.

θ_i : Hücrenin eşik (bias) değeri.



η : Öğrenme sabiti.

δ_{oi} : Çıkış katmanı hata işareti.

μ : Aktivasyon fonksiyonu maksimum eğimi.

τ_i : Aktivasyon fonksiyonu eşik noktası.

AADV: Average Absolute Derivative Value.

BP: Back Propagation.

CNN: Cellular Neural Networks.

HYSA: Hücresel Yapay Sinir Ağları.

ÇS-HYSA: Çok Seviyeli Hücresel Yapay Sinir Ağları

PWL: Piecewise Linear.

TAF: Trapezoidal Aktivasyon Fonksiyonu.

TAF-HYSA: Trapezoidal Aktivasyon Fonksiyonlu Hücresel Yapay Sinir Ağları.





ŞEKİLLER DİZİNİ

Şekil

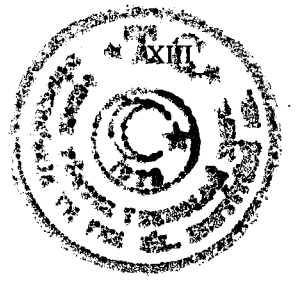
Sayfa

2.1. Perceptron modeli	7
2.2. Farklı ağ bağlantıları	14
2.3. Çok katmanlı ağ yapısı	15
2.4. Eğiticili öğrenme	17
3.1. $r=1$ ve $r=2$ komşuluğundaki hücresel yapay sinir ağları	29
3.2. Hücresel yapay sinir ağı hücresinin matematik modeli	29
3.3. Standart parça parça doğrusal aktivasyon fonksiyonu	31
4.1. Standart HYSA aktivasyon fonksiyonu	41
4.2. Trapezoidal aktivasyon fonksiyonu	42
4.3. Giriş Görüntüleri	55
4.4. Kenar tanıma şablonları için PWL-HYSA çıkışları	56
4.5. Kenar tanıma maskeleri (4.52) kullanılan PWL-HYSA için AADV eğrileri	57
4.6. (4.54) ifadesindeki A şablonu için PWL-HYSA çıkışları	57
4.7. Tanımı (4.54) ifadesinde verilen A şablonu kullanılan PWL-HYSA için AADV eğrileri	58
4.8. (4.55) ifadesindeki non-simetrik A şablonu için PWL-HYSA çıkışlar	59
4.9. Tanımı (4.55) ifadesinde verilen non-simetrik A şablonu kullanılan PWL-HYSA için AADV eğrileri.	60
4.10. (4.56) ifadesindeki non-simetrik A şablonu için PWL-HYSA çıkışları	61
4.11. (4.56) ifadesinde verilen non-simetrik A şablonu kullanılan PWL-HYSA için AADV eğrileri	61
4.12. (4.55) ifadesindeki non-simetrik A şablonu ve (4.57) ifadesindeki aktivasyon fonksiyonu parametreleri kullanılan TAF-HYSA çıkış görüntüleri	62
4.13. (4.55) ifadesindeki non-simetrik A şablonu ve (4.57) ifadesindeki aktivasyon fonksiyonu parametreleri kullanılan TAF-HYSA'ya ait AADV eğrileri	62
4.14. (4.58) deki kararlı non-simetrik A şablonu ve (4.57) ifadesindeki aktivasyon fonksiyonu parametreleri kullanılan TAF-HYSA çıkış görüntüleri	63
4.15. (4.58) ifadesindeki non-simetrik kararlı A şablonu ve (4.57) deki aktivasyon fonksiyonu parametreleri kullanılan TAF-HYSA'nın AADV eğrileri	64



Şekil

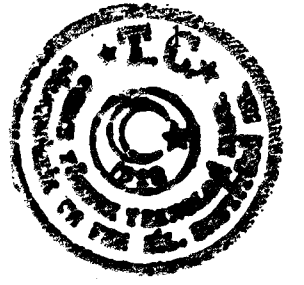
4.16. (4.55) ifadesindeki non-simetrik A şablonu ve (4.59) ifadesindeki aktivasyon fonksiyonu parametreleri kullanılan TAF-HYSA çıkış görüntüleri	65
4.17. ((4.55) ifadesindeki non-simetrik A şablonu ve (4.59) ifadesindeki aktivasyon fonksiyonu parametreleri kullanılan TAF-HYSA'ya ait AADV eğrileri	65
4.18. (4.58) deki kararlı non-simetrik A şablonu ve (4.59) ifadesindeki aktivasyon fonksiyonu parametreleri kullanılan TAF-HYSA çıkış görüntüleri	66
4.19. (4.58) ifadesindeki non-simetrik kararlı A şablonu ve (4.59) deki aktivasyon fonksiyonu parametreleri kullanılan TAF-HYSA'ya ait AADV eğrileri	66
4.20. XOR Problemi	67
4.21. XOR uygulamaları için kullanılan giriş görüntüleri	68
4.22. Yatay parite-2 uygulaması için TAF-HYSA çıkış görüntüleri	69
4.23. Düşey parite-2 uygulaması için TAF-HYSA çıkış görüntüleri	69
4.24. Aynı düşey kenar üzerinde bulunan bitişik 3 piksel üzerinde gerçekleştirilen parite-2 uygulaması için TAF-HYSA çıkış görüntüleri	70
4.25. Lineer olarak ayrıştırılamayan uygulama için (4.62) ifadelerinde tanımlı olan şablonlar ve aktivasyon fonksiyonu için TAF-HYSA çıkış görüntüleri	71
6.1. Sentetik küre modeli	83
6.2. Gravitasyonel veriler için HYSA'nın eğitiminde kullanılan Tablo 6.1. de özellikleri verilen üç sentetik küreye ait anomali haritaları	86
6.3. Gravitasyonel veriler için HYSA'nın testinde kullanılan, özellikleri Tablo 6.3. de verilen 4 sentetik küreye ait anomali haritaları	89
6.4. Dumluca gravite anomalisi ve HYSA çıkış görüntüleri	90
6.5. Magnetik dipol modeli	91
6.6. Magnetik verilerin eğitimi için kullanılan ve özellikleri Tablo 6.4.'de verilen sentetik örnek	94
6.7. Tablo 6.7'de özellikleri verilen 4 magnetik dipole ait anomali haritaları	98
6.8. Hitit İmparatorluğu kalıntılarına ait arkeolojik veriler	99
6.9. Arkeolojik çalışmalar sonucunda çıkartılmış Hitit kalıntılarının görüntüleri	100



ÇİZELGELER DİZİNİ

Çizelge	Sayfa
6.1. Üç küreden oluşmuş gravite eğitim modeline ait parametreler	84
6.2. Genetik algoritma parametreleri	84
6.3. Sentetik gravite test verileri	88
6.4. Aynı (x,y) koordinatlarında farklı derinliklerde bulunan iki dipolün özellikleri	93
6.5. Magnetik verilerin eğitiminde kullanılan genetik algoritma parametreleri	93
6.6. Sentetik dört magnetik dipole ait özellikler	97





1. GİRİŞ

Yapay sinir ağları, günümüzde görüntü işleme ve görüntü tanıma konuları da dahil olmak üzere pek çok konuda hızla büyüyen bir ilgi alanı oluşturmaktadır. 1940'lı yılların başında Mc Culloch, Pitts, Hebb ve Rosenblatt'ın çabalarıyla alışlagelmiş bilgi işleme sistemlerine bir seçenek olarak ilk yapay sinir ağı modelleri geliştirilmeye başlanmıştır [Cichocki A.,1993]. Özellikle 80'li yıllarda bu konudaki çalışmalar hızla artış göstermiş, Hopfield, Backpropagation gibi ağ modellerinin problem çözmedeki başarıları, çeşitli bilim disiplinlerinde çalışan bilim adamlarının yapay sinir ağlarına özel ilgi göstermelerine sebep olmuştur.

Yapay sinir ağlarının özel bir türü olan ve dinamik Hopfield ağlarına benzer bir yapısı olan Hücresel Yapay Sinir Ağları (HYSA) diğer adıyla Cellular Neural Networks (CNN), Chua ve Yang tarafından 1988 yılında tanıtılmıştır [Chua and Yang, 1988]. HYSA ısı yayılımı, dalga yayılımı gibi nonlineer olayların simülasyonunda ve analizinde başarılı bir yöntemdir. Hücresel Yapay Sinir Ağının iki boyutlu bir matris yapısında olması nedeniyle benzer problemlerin iki boyutlu düzlemdeki analizinde ve özellikle görüntü işleme alanında tercih edilmesine sebep olmuştur. Ayrıca ağdaki hücrelerin durum denklemindeki dinamik davranış, hücresel yapay sinir ağını doğal olayların simülasyonunda başarılı kılan bir diğer önemli etkindir.

HYSA dinamik bir yapıya sahip olduğundan dolayı ağın kararlılığı üzerine çok sayıda çalışma yapılmıştır [Gilli, 1994], [Arik, 2002]. Genelde simetrik ileri besleme ve geri besleme şablonları kullanımı tercih edilerek ağın kararlılığı sağlanır. Ancak şablonlardaki simetriklik kısıtlaması HYSA'nın uygulama alanını daraltmaktadır. Çünkü çoğu problemler simetrik olmayan şablonların kullanılmasını zorunlu kılmaktadır. Bu nedenle simetrik olmayan şablonların kullanılması için şablon yapılarına bağlı kararlılık kriterleri araştırılmıştır [Arik ve Tavşanoğlu, 1996,2000].

Yapay sinir ağıları mantık kapılarını gerçekleştirmek için uygulandıklarında AND ve OR işlemleri tek katmanlı bir ağ ile gerçekleşirken XOR işlemi tek katmanlı bir ağ ile gerçekleştirilmektedir [Majorana ve Chua, 1998]. Bu durum HYSA için de geçerlidir. Tez çalışmasında standart PWL aktivasyon fonksiyonu yerine Trapezoidal Aktivasyon Fonksiyonu (TAF) kullanılarak HYSA'nın XOR gibi lineer olarak ayrıştırılamayan problemleri tek katmanda çözebildiği gösterilmektedir. Ayrıca TAF kullanılması durumunda ağıın kararlılık kriteri araştırılmış ve bu kriter standart PWL aktivasyonu kullanılması durumuna uyarlanmıştır. Geliştirilen kararlılık kriteri şablonların simetrik olma kısıtlamasını dikkate almadığı için simetrik olmayan şablonlarla da çalışma imkanı vermektedir.

Tez çalışmasında ayrıca jeofizik verileri kullanılarak HYSA ile magnetik ve gravite anomalileri ayrıştırılmıştır. Jeofizikte, bir bölgedeki magnetik ve gravite haritaları yüzey altındaki farklı yoğunlukta ve farklı magnetik özelliklerdeki kütlelerin yerlerinin belirlenmesinde çok sık kullanılan araçlardır. Keza, iletkenlik haritaları ve uzaktan algılama gibi diğer teknikler de kaya yapılarının yerlerinin belirlenmesinde çok yararlı olurlar. Magnetik ve gravite anomalileri ilk olarak filtrelerden geçirilerek anomalinin yorumunu güçleştiren istenmeyen alan bileşenlerinin elimine edilmesi veya zayıflatılması sağlanır. Filtre işleminden sonra, bileşke alana sebep olan kaya yapıları daha kolay tespit edilebilir. Bu başlangıç filtreleme işlemleri, radyal ağırlıklar yöntemi, en küçük kareler minimizasyonu ve hızlı Fourier dönüşümü filtreleme yöntemi gibi yöntemleri içerir.

Gravite ve magnetik anomali verilerinin frekans analizleri ilk defa 1958 yılında yapıldı ve kontur eğrilerinin türevleri elde edildi [Dean,1958]. Arkasından magnetik alan verilerinin yorumlanması için 2 boyutlu (2B) harmonik analizi gerçekleştirildi [Bhattacharyya,1965]. Daha sonraki çalışmalarda, karesel pikseller halinde kaydedilmiş veriler üzerinde enterpolasyon yapılarak 2B filtreler [Fuller, 1967, Zurflueh,1967], Hankel transformu [Lavin ve Devane, 1970] ve Fourier transformu teknikleri kullanılmıştır [Clement, 1973]. Hankel transformasyonu kullanılan bir sayısal filtre örneği Anderson tarafından verilmiştir [Anderson, 1979]. Missisipi ve Pensilvanya bölgesinden elde edilmiş anomali verileri üzerinde Wiener filtreleme teknikleri kullanılarak anomali ayrıştırma yapılmıştır [Pawlowski ve Hansen,1989]. Gravite anomali ayrıştırmalarında kullanılan bir başka teknik ise

Walsh transformlarıdır. Doğrusal filtreler ve Werner dekonvolüsyon algoritmaları da anomali ayırıştırma işlemlerinde uygulanan yöntemler arasındadır. [Hansen and Simmonds, 1993; Uçan ve Albora, 2003].

Yer altında farklı yapılar arasındaki gömülü cisimlerin sınırlarının anomali haritaları kullanılarak tespit edilmesi bir diğer önemli problemidir [Blakely ve Simpson, 1986]. Bu alanda Wavelet dönüşümünün kullanılma yöntemleri için şu çalışmalar örnek gösterilebilir; i) potansiyel alanların rezidüel ve bölgesel alanlarının ayırıştırılması [Fedi and Quarta, 1998], [Hornby, et al, 1999], ii) aeromagnetik verilerin incelenmesi [Riddsil, et al, 1999], iii) Çok ölçekli kenar analizi kullanılarak jeolojik yapıların geometrisinin modellenmesi [Holden, et al, 2000], iv) Potansiyel alan verilerinde gürültü süzme ve gömülü cisimlerin kenarlarının kestirimi [Boshetti, et al, 2001], [Uçan, et al, 2000, 2001].

Kısaca özetlemek gerekirse, yüzey altında bulunan her kütle yüzey üzerindeki magnetik ve gravitasyonel alan dağılımını, magnetik özellikleri, kütsel yoğunluğu ve kütselinin büyüklüğü nisbetinde etkiler. Bu çalışmada amaç, yer üzerinde ölçülen magnetik ve gravite alanlarının dağılımlarından, bu alanlara sebep olan yoğun kütlelerin yerinin belirlenmesidir.

Problemin çözümünde klasik yöntemlerin yerine bu tür problemlerin çözümü için uygun bir yöntem olan HYSA kullanılmıştır. Hücrel yapay sinir ağının eğitimi genetik algoritmalar ile yapılmıştır. Eğitim işleminde A,B, ve I şablonlarına ait parametrelere ikili kodlanmış genler karşılık düşürülerek, en iyi performansı gerçekleştiren kromozom bulunur. Genetik algoritmalar son yıllarda optimizasyon problemlerinin çözümünde başarıyla kullanılan modern yöntemlerden biridir [Sziranyi and Csapodi, 1994]. Hücrel yapay sinir ağının eğitimi şablon parametrelerinin optimizasyonuna dayandığından genetik algoritmalar eğitim için tercih edilmektedir. Tezin bir diğer özgünlüğü, genetik algoritmalarla eğitilmiş hücrel yapay sinir ağlarının jeofizik alanına ilk defa uygulanmasıdır.

Çalışmada iki farklı veri tipi kullanılmıştır. İlk uygulamada, ülkemizin demir ve krom madeni açısından zengin yataklarından birisi olan Dumluca bölgesine ait gravite verileri üzerinde çalışılmıştır. Jeofizikteki temel problemlerden birisi olan

potansiyel alan ayrıştırma yöntemi, çok katmanlı HYSA ile gerçekleştirilmeye çalışılmıştır. İkinci uygulamada ise Türkiye topraklarında hakimiyet sürmüş olan Hitit İmparatorluğuna ait arkeolojik kalıntıların bulunduğu Sivas-Altınyayla bölgesine ait magnetik veriler kullanılarak, hücresel yapay sinir ağının yeraltındaki görünmeyen bir binanın odalarına ait sınırları tespit etmesi sağlanmıştır. Aslında bu problem, jeofizikteki magnetik anomali ayrıştırmasının bir uygulamasıdır. Her iki uygulama için başlangıçta sentetik magnetik ve gravite verileri üretilerek bu sentetik veriler hücresel yapay sinir ağının eğitiminde kullanılmıştır.



2. YAPAY SINİR AĞLARI

Yapay sinir ağı (YSA) girişine verilen bilgiyi öğrenme, daha önce öğrendiği bilgiyi hatırlama, ve eğitim deseni veya veri kümelerini giriş olarak kabul edip bu bilgilerden yararlanarak genelleştirme yapabilme, yorum yapabilme yeteneği gösteren ümit vaat eden yeni kuşak bir bilgi işleme sistemidir. YSA bilinçli olarak insan beyninin çalışma sistemini taklit etmek amacıyla geliştirilen bir tekniktir. Desen tanıma, sınıflandırma, fonksiyon kestirimi, optimizasyon işlemleri ve verilerin kümelendirilmesi, gruplandırılması işlemlerinin gerçekleştirilmesinde çok iyi bir performans sergilemektedir [Ölmez, 1995]. Biyolojik sistemler, sayısal bilgisayarlar karşısında aritmetik işlemlerde kıyaslanmayacak kadar güçsüz olmalarına karşın, görüntü ve ses işleme gibi bozulmuş, tam tanımlı olmayan yüksek veri yoğunluğuna sahip bilgilerin işlenmesinde sayısal bilgisayarlara göre çok hızlı ve yüksek doğrulukta çalışırlar. Bilgisayarların ses ve görüntü gibi işaretlerin işlenmesinde yapısal olarak güçsüz olduğunun görülmesi 80'li yıllarda çeşitli bilim dallarından büyük bir araştırmacı kitlesinin YSA'na büyük ilgi duymalarına sebep olmuştur.

1943 yılında McCulloch ve Pitts biyolojik sinir hücresinden esinlenerek geliştirdikleri bir yapay sinir hücre modeli ile VE, VEYA gibi mantık işlemlerinin yapılabileceğini gösterdiler [Lee and Lin, 1995]. Böylece canlıların sinir sistemlerinin incelenmesi ve bunlara benzer şekilde çalışan yapay sinir ağı modellerinin geliştirilmesi çok farklı bilim dalından araştırmacıların ortak olarak çalıştığı bir bilim dalı olmuştur.

YSA alanında yapılan araştırmalar 60'lı yılların ortalarından 80'li yılların başına kadar bir durgunluk dönemi yaşadı. Bunun en önemli sebebi bilgi işlemede YSA'nın alternatifi olan bilgisayarların yarı ileken teknolojisindeki gelişmeler ile kapasiteli, ucuz ve güvenilir gerçekleştirme olanağı bulmasıdır. Seri olarak çalışan hızlı birimlerden oluşmuş sayısal bilgisayarlar, aritmetik işlemlerde yüksek hız, kapasite ve güvenilirlik sağlamışlardır.

YSA, birbirleriyle çok sayıda iç bağlantıları olan büyük miktarda, hücre adı verilen işlem elemanlarından oluşur. Bilgi ağı içerisinde paralel işlenir. İnsan

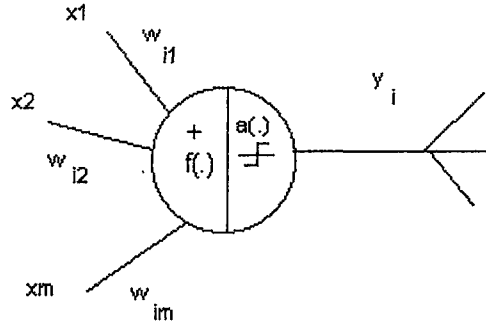
beyninde olduđu gibi YSA'nın da etkileşimli ve kollektif çalışması öğrenme ve öğrenilen bilginin hatırlanmasına, yorum yapmaya imkan tanır. YSA'nın özelliklerinden bahsetmeden önce insan beyninin çalışma sistemine bir gözatmakta fayda vardır. İnsan beyninde yaklaşık 10^{11} civarında farklı yapılarda sinir hücresi bulunduđu bilinmektedir. Biyolojik sinir hücresi üç ana kısımdan oluşur. Bunlar;

1. Soma
2. Dendritler
3. Akson

Soma, hücre çekirdeğinin bulunduđu bölgedir. Her bir sinir hücresine diğ er hücrelerden bilgi ileten sinir tellerine dendrit adı verilir. Hücrenin akson adı verilen tek bir çıkışı vardır ve bu çıkıştan komşu sinir hücrelerine bilgi taşıyan binlerce dendrit bağıdır. Aksona bağı her bir dendritteki iletilen bilgi aynıdır. Yani bilginin iletim sırasında zayıflaması sözkonusu değildir. Hücre girişine bilgi taşıyan dendritler sinaps adı verilen bağlantı noktalarında hücreye bağlanırlar.

Dendritlerde iletilen ve sinapslara ulaşan sinyaller elektrik darbeleridir. Bağlantı noktası (sinaps) tarafından salınan özel iletim maddeleriyle gerçekleştirilen, karmaşık kimyasal işlemleri içeren, bir işaret iletim sistemi söz konusudur. Bu maddeler Na ve K iyonlarıdır. Bu maddelerin konsantrasyonun değ işmesi ile hücre içi elektrik potansiyeli azalır veya artar. Alıcı hücrenin elektriksel potansiyeli, o hücreye ilişkin eşik değ erini geçerse hücre ateşlenir, aktif hale geçer. Hücre çıkışındaki aksonda sabit şiddette ve sabit süreli bir elektrik darbesi oluşur. Bu işaret akson boyunca iletilir ve akson ucuna bağı olan diğ er hücrelere bilgi taşıyan dendritlere aktarılır. Hücreye yeni bir uyarı gelinceye kadar hücre dinlenir. Bu zamana **bekleme veya dinlenme zamanı 'refractory period'** denir. Sinapslar eğer alıcı hücrenin ateşlenmesine sebep olacak sinyalleri geçiriyorsa bu sinapslara **uyarıcı (excitatory)** sinaps adı verilir. Eđer alıcı hücrenin ateşlenmesini engelleyici sinyallerin geçmesine izin veriyorlarsa bu tip sinapslara, engelleyici (inhibitory) sinaps adı verilir.

Şekil 2.1.' de yukarıda anılan özellikleri gerçekleştiren, Mc Culloch ve Pitts tarafından ileri sürülen ve genellikle M-P sinir hücresi adı verilen, biyolojik sinir hücresinin basit bir matematik modeli görölmektedir.



Şekil 2.1. Perceptron Modeli

Bu modelde i işlem elemanı girişlerin ağırlıklı toplamını hesaplar ve bu toplam eşik değerinden büyükse çıkışta $y_i = 1$ olur yani hücre ateşlenir. Aksi halde $y_i = 0$ olur yani hücre ateşlenmez.

Burada aktivasyon fonksiyonu başka bir deyişle hücrenin transfer fonksiyonu ;

$$a(f) = \begin{cases} 1 & f \geq 0 \\ 0 & f < 0 \end{cases} \quad (2.1)$$

şeklinde tanımlı birim basamak fonksiyonudur.

w_{ij} ağırlığı, j . nöronu i . nörona bağlayan sinapsın uyarma şiddetini ya da bağlantının ağırlık katsayısını gösterir. Pozitif ağırlıklar sinapsın uyarıcı olduğunu, negatif ağırlıklar uyarmayı engelleyici bir rol üstlendiğini gösterir. $w_{ij} = 0$ olması iki hücre arasında bir bağlantının olmadığını gösterir.

Mc Culloch ve Pitts modeli, uygun ağırlık katsayıları verilmesi durumunda AND, OR ve NOT işlemlerini gerçekleştirir. Çok değişkenli, kombinasyonel bir lojik fonksiyon, bu basit lojik işlemlerle gerçekleştirildiğinden, bu tip sinir hücrelerinin eşzamanlı çalışmasıyla geleneksel bilgisayarların yaptığı bütün işlemler YSA ile gerçekleştirilebilir.

Bütün bu özellikleri özet olarak aşağıdaki gibi sıralayarak YSA'nın bir tanımını yapabiliriz.

1. Sinir sisteminden esinlenerek oluşturulmuş bir matematik modeldir.
2. Birbirleriyle bağlantıları olan çok sayıda işlem elemanından oluşur.
3. Bağlantılar (ağırlık katsayıları) bilgiyi tutar, hafızayı temsil eder.
4. İşlem elemanı girişindeki uyarıyı hemen yanıtlayabilir. Ve bu çıkış cevabı hücrenin yerel bilgisine bağlıdır. Başka bir deyişle hücrenin çıkış cevabı, gelen bilgiye ve hücredeki yerel bellekteki bilgiye (ağırlık katsayısı ve eşik değeri) bağlıdır.
5. Yeni bağlantı ağırlıkları atayarak ya da bağlantı katsayılarını değiştirerek eğitim verilerini genelleştirebilme, öğrenme ve hatırlama yeteneğine sahiptir.
6. Hücrelerin kollektif, birbirleriyle etkileşimli çalışması hesaplama gücünü artırır. Belirli bir bilgiyi taşıyan özel bir hücre yoktur. Bilgi bütün sisteme dağılmıştır. Buna "Dağılmış yapı -distributed representation özelliği" denir.

Bu özelliklerden dolayı YSA için *paralel, dağıtılmış bilgi işleme modeli*, (*parallel distributed processing models*), *bağlantılı model* (*connectionist model*), *kendi kendine öğrenen model* (*self-organizing systems*), *neurocomputing* ve *neuromorphic systems* adları da yaygın olarak kullanılır.

C. Güzelış bugüne kadar verilen çeşitli YSA modellerinin ortak özelliklerini aşağıdaki gibi sıralamıştır [Güzelış, 1993].

1. Hücreler, genelde çok girişli ve tek çıkışlı yüksek dereceden lineer olmayan alt devrelerdir.
2. Tek olan hücre çıkış işareti, çoğullanarak dışarıya çıkış olarak alınabilir veya diğer hücrelere giriş olacak şekilde (genelde gecikmeler ile) kullanılabilir.
3. Bir hücre; dış girişleri (gecikme ile ulaşmış olabilir), komşu hücre çıkışlarını, ve bir eşik değerini giriş olarak kabul eder.
4. Hücrenin lineer dinamik birimi herhangi bir dereceden olabilir.
5. Hücre çıkışında lineer olamayan cebrik işlev 1-giriş 1-çıkışlıdır ve özel olarak lineer olabilen herhangi bir işlevdir.
6. Hücre parametreleri hücreden hücreye değişebilir.
7. Ağ, genelde çok boyutlu katmanların ileri veya geribeslemeli olacak şekilde bağlanmasından oluşmuştur.
8. Ağdaki her bir katman (layer), hücrelerin çok boyutlu bir dizisidir.

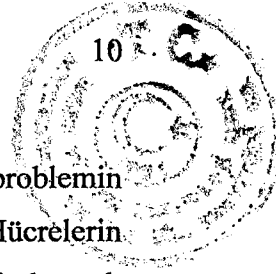


9. Katman içinde, herbir hücre komşularına (belirli bir metrik uyarınca) aynı şekilde bağlıdır. Böylece katman içi bağlantı geometrisi düzgündür.
10. Katmanlar arası bağlantı bir metrik ile tanımlı düzgün bir geometriye sahiptir.
11. Ağda yalnızca bağlantı ağırlıkları değişebilir.
12. Bağlantı ağırlıkları; ya eğitici bir öğrenme kuralı ile ya da eğitici bir öğrenme kuralı ile değiştirilir yada önceden tasarlanan değerlerde sabit tutulur.
13. Eğitici bir öğrenme kuralı (eğer varsa); giriş - (istenilen) çıkış örnek çiftlerinin alındığı, bilinmeyen bir giriş işlevine bu örnekler yardımıyla yaklaşılmayı sağlayan bir bağlantı ağırlıkları değişim kuralıdır.
14. Eğitici bir öğrenme kuralı(eğer varsa); giriş örneklerinin kümelendirilmesini sağlayan bir bağlantı ağırlıklarının değişim kuralıdır.

2.1. Yapay Sinir Ağlarının Üstünlükleri

YSA modelleri biyolojik sinir ağlarının çalışma biçimini açıklayan modeller olmaktan çok, biyolojik sinir ağlarından esinlenerek yaratılmıştır. YSA' lar, biyolojik olmayan yapı taşlarının düzgün bir mimari ile birbirine yoğun olarak bağlanmalarından oluşmaktadırlar. Sinir ağlarında bilgi işlemenin alternatif bilgi işleme yöntemlerine göre üstünlükleri maddeler halinde aşağıda verilmiştir [Zurada, J.M., 1989].

Paralellik: Alışıl gelmiş bilgi işlem yöntemlerinin çoğu seri işlemlerden oluşmaktadır. Bu da hız ve güvenilirlik sorunlarını beraberinde getirmektedir. Seri bir işlem gerçekleşirken herhangi bir birimin yavaş oluşu tüm sistemi doğrudan yavaşlatırken, paralel bir sistemde yavaş bir birimin etkisi çok azdır. Nitekim, seri bir bilgisayarın bir işlem elemanı beyine göre binlerce kez daha hızlı işlem görmesine rağmen, beynin toplam işlem hızı seri bir bilgisayara göre kıyaslanmayacak kadar yüksektir. YSA' nın paralel çalışması hız avantajıyla birlikte yüksek hata toleransı da sağlamaktadır. Seri bilgi işlem yapan bir sistemde herhangi bir birimin hatalı çalışması, hatta bozulmuş olması tüm sistemin hatalı çalışmasına veya bozulmasına sebep olacaktır. Halbuki paralel bilgi işleme yapan bir sistemin ayrı ayrı işlem elemanlarından meydana gelecek olan hatalı çalışma veya hasar, sistemin performansında keskin bir düşüşe yol açmadan, performansın sadece hatalı birimlerin bir oranınca düşmesine sebep olur.



Yerel Bilgi İşleme: YSA'da her bir bilgi işleme birimi, çözülecek problemin tümü ile ilgilenmek yerine, sadece problemin bir parçasını işlemektedir. Hücrelerin çok basit işlemler yapmalarına rağmen, sağlanan görev paylaşımı sayesinde, çok karmaşık ve zor problemler çözülebilmektedir.

Gerçeklenme Kolaylığı: YSA'nın basit işlemler gerçekleyen türdeş hücrelerden oluşması ve bağlantıların düzgün olması, bu ağların fiziksel gerçeklenmesinde bu açıdan kolaylık sağlamaktadır.

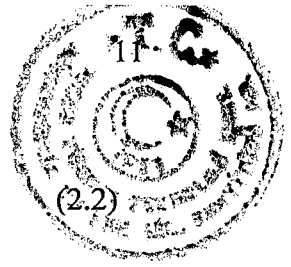
Öğrenebilirlik: Alışlagelmiş veri işleme yöntemlerinin çoğu programlama yolu ile hesaplamaya dayanmaktadır. Bu yöntemler ile tam tanımlı olmayan bir problem çözülemez. Ayrıca herhangi bir problemin çözümü için probleme yönelik bir algoritma geliştirmek gerekir. Oysa YSA' larda örneğin eğitici öğrenmede giriş ve çıkış çifleri verildikten sonra bunlar arasındaki ilişkiyi belirlemek üzere bir algoritma geliştirmeye gerek yoktur. İlgili bağıntıyı YSA bağlantı katsayılarını değiştirerek belirler. Zaten öğrenme dediğimiz işlemde başka bir şey değildir. YSA' da bilgi, bağlantı katsayıları ile tutulur.

2.2. Basit YSA Modelleri ve Öğrenme Kuralları

Bir YSA modelinin karakteri üç temel kriterle belirlenir.

1. İşlem elemanları, sinir modelleri
2. Sinaptik iç bağlantılar ve bağlantı yapıları
3. Bağlantı katsayılarını güncellemek için kullanılan eğitim ve öğrenme algoritmaları

İşlem Elemanları : Daha önce açıklanan M-P sinir modelinin çalışma şekli, gerçekleştirdiği fonksiyon diğer yapay sinir modellerini içine alacak şekilde geliştirilebilir. Bir işlem elemanı girişler ve çıkışlar olmak üzere iki kısımdan oluşur. Hücre girişindeki bilgi diğer hücrelerden gelen bilgilerin ağırlıklı toplamıdır. Bu giriş işareti M-P modelinde olduğu gibi genelde x_j girişlerinin doğrusal bir fonksiyonudur.



$$f_i = net_i = \sum_{j=1}^m w_{ij} x_j(t) - \theta_i \quad (2.2)$$

Burada θ_i hücrenin eşik değeri m ise hücre sayısıdır. Daha karmaşık toplama fonksiyonlarının olması da söz konusudur. Sık kullanılan fonksiyonlar şunlardır:

1. Karesel Fonksiyon (Quadratic function):

$$f_i = \sum_{j=1}^m w_{ij} x_j^2 - \theta_i \quad (2.3)$$

Karesel fonksiyonun kullanılması durumunda, giriş işareti komşu sinir hücrelerinden gelen işaretlerin karelerinin ağırlıklı toplamından elde edilir.

2. Silindirik Fonksiyon (Spherical function)

$$f_i = \rho^{-2} \sum_{j=1}^m (x_j - w_{ij})^2 - \theta_i \quad (2.4)$$

Burada ρ , silindirin yarıçapı, w_{ij} ise merkezini göstermektedir.

3. Polinom fonksiyon (Polynomial Function)

$$f_i = \sum_{j=1}^m \sum_{k=1}^m w_{ijk} x_j x_k + x_j^\alpha + x_k^\beta - \theta_i \quad (2.5)$$

Burada w_{ijk} , j . ve k . hücreyi, i . hücreye bağlayan bağlantının ağırlık katsayısı, α ve β reel sabitlerdir. Bu eşitlikte daha yüksek dereceden terimlerin olması da mümkündür. Bu fonksiyonu gerçekleştiren hücreye *sigma-pi* ($\Sigma\Pi$) birimi adı verilir.

İşlem elemanının (hücrenin) ikinci bir işlevi, giriş işaretini transfer fonksiyonundan diğer adıyla aktivasyon fonksiyonundan $a(f)$ geçirerek, girişin fonksiyonu olan bir çıkış değeri üretmektir. Sıkça kullanılan transfer fonksiyonları ise aşağıdaki fonksiyonlardır.



1. Basamak fonksiyonu (Step function)

$$a(f) = \begin{cases} 1 & f \geq 0 \\ 0 & f < 0 \end{cases} \quad (2.6.)$$

2. Keskin sınırlayıcı fonksiyonu (Hard limiter, threshold function)

$$a(f) = \text{sgn}(f) = \begin{cases} 1 & f \geq 0 \\ -1 & f < 0 \end{cases} \quad (2.7)$$

Burada $\text{sgn}(\bullet)$, signum fonksiyonudur.

3. Rampa fonksiyonu (Ramp function)

$$a(f) = \begin{cases} 1 & f \geq 1 \\ f & 0 \leq f < 1 \\ 0 & f < 0 \end{cases} \quad (2.8.)$$

3. Unipolar sigmoid fonksiyonu

$$a(f) = \frac{1}{1 + e^{-\lambda f}} \quad (2.9)$$

4. Bipolar sigmoid fonksiyonu

$$a(f) = \frac{2}{1 + e^{-\lambda f}} - 1 \quad (2.10)$$

Eşitlik (2.9) ve eşitlik (2.10) da $\lambda > 0$ olduğunda, $f = 0$ civarında, $a(f)$ sürekli fonksiyonunun adım aralığını belirler. $\lambda \rightarrow \infty$ olması durumunda eşitlik (2.9), (2.6) eşitliğine; yani unipolar sigmoid fonksiyonu, basamak fonksiyonuna, eşitlik (2.10) ise (2.7) eşitliğine; yani bipolar sigmoid fonksiyonu, keskin sınırlayıcı fonksiyonuna dönüşür. Keskin sınırlayıcı transfer fonksiyonuna ve lineer toplayıcı birimine sahip olan bir işlem elemanına, doğrusal eşik birimi (linear threshold unit, LTU), sigmoid transfer fonksiyonuna ve doğrusal toplayıcıya sahip işlem elemanına



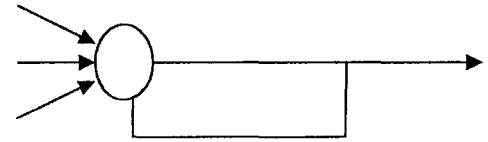
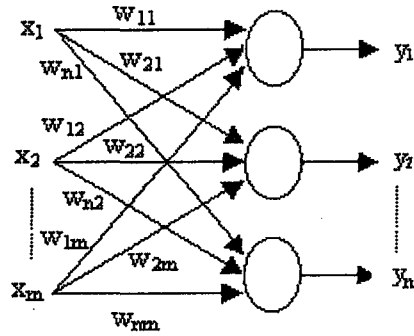
doğrusal eğimli birim (linear graded unit, LGU) adı verilir. LTU ve LGU çok yaygın kullanılan işlem elemanlarıdır.

Girişlerin toplanmasında doğrusal olmayan fonksiyonların kullanılması, karar sınırları adı verilen sınırlarla öznitelik uzayının karmaşık bölgeler şeklinde belirlenmesine imkan tanır. Bununla birlikte, toplama fonksiyonu karmaşıktıkça ve güçlendikçe, ağ yapısının da karmaşık olması kaçınılmazdır. Nasıl bir transfer fonksiyonu kullanılması gerektiği ve toplama fonksiyonun lineer olup olmaması, tamamen problemin yapısına bağlıdır.

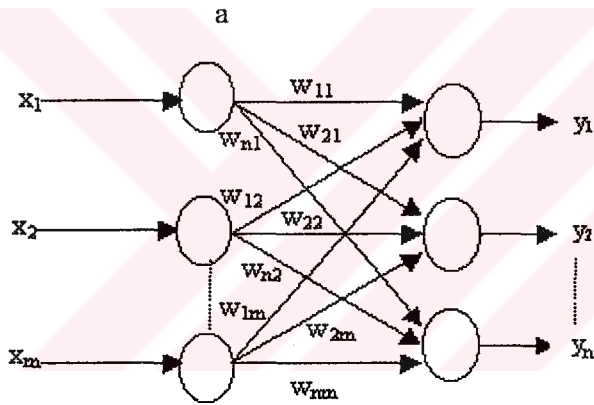
Bağlantılar: Bir YSA çok yoğun bağlantıları olan işlem elemanları (hücreler) kümesinden oluşur. Hücreler birbirlerine veya kendilerine bağlantı ağırlıkları üzerinden bağlıdır. Bağlantı yapısı gecikme ve geri kalmaya imkan tanır. Bu nedenle ağı bağlantı yapısının ve hücreler arası bağlantı geometrisinin, çok büyük bir önemi vardır. İşlem elemanın transfer fonksiyonunun ve toplama birimi türünün yanında bağlantıların nerede başladığı ve nerede sonlandığı, bağlantı biçimi, ağı karakterini belirlemektedir. Mümkün beş farklı ağ geometrisi söz konusudur. Ağ yapısı Şekil (2.2.a)'da görüldüğü gibi bir katmandan oluşuyorsa, bu tip bir ağa tek tabakalı ileri beslemeli ağ adı verilir. Girişler işlem elemanlarına bağlantı katsayıları üzerinden bağlıdır. Görüldüğü gibi her bir hücrenin bir tek çıkışı vardır. Bu şekilde daha fazla katman arasında bağlantılar oluşturarak çok tabakalı ileri beslemeli ağ elde edilir. Girişlerin bağlı olduğu tabakaya giriş katmanı adı verilir. Giriş tabakası veriler üzerinde herhangi bir işlemde bulunmaz. Sadece verilerin ağa girilmesinde arayüz görevi yapar. Ağı çıkışları çıkış tabakası adı verilen katmandan alınır. Giriş ve çıkış katmanları arasında bir veya birden fazla katman bulunabilir. Bu katmanlara saklı, gizli katman adı verilir. Bu katmandaki verilere doğrudan erişme imkanı yoktur. Bir katmandaki her bir hücrenin bir sonraki katmandaki bütün hücrelerle bağlantısı varsa bu tip bir ağa tam bağlantılı (fully connected) ağ adı verilir. Aksi durumda bir katmandaki hücrelerden en az birinin, bir sonraki katmandaki hücrelerden biriyle veya birkaçıyla bağlantısı yoksa bu tip bir bağlantı tam bağlantılı değildir.

Eğer bir katmandaki hücrelerin çıkışları bir önceki katmandaki hücrelere giriş olarak bağlanmıyorsa, böyle bir ağa, ileri beslemeli ağ (feedforward networks) adı

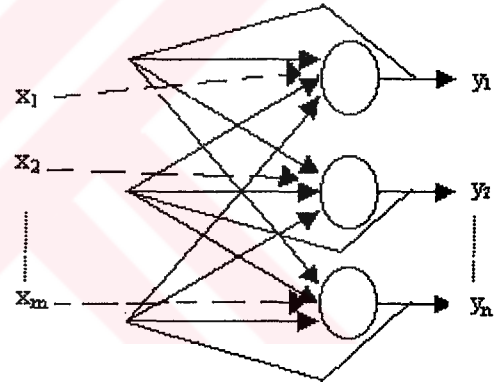
verilir. Geriye doğru bağlantıların olması geri besleme anlamına geldiğinden böyle bir ağı ise, geribeslemeli ağ (recurrent networks) denilmektedir. Eğer bir katmandaki çıkış, aynı katmana giriş olarak geri dönüyorsa, bu tip geribeslemeye yanal geri besleme veya daha doğru bir tabir ile kendi üstüne geri besleme adı verilir.



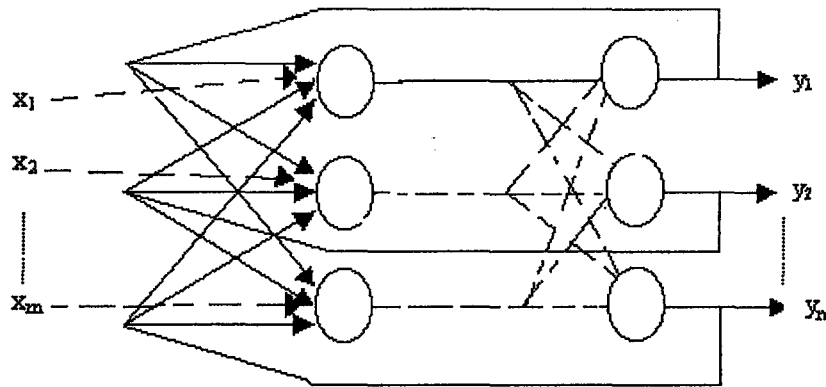
b



c

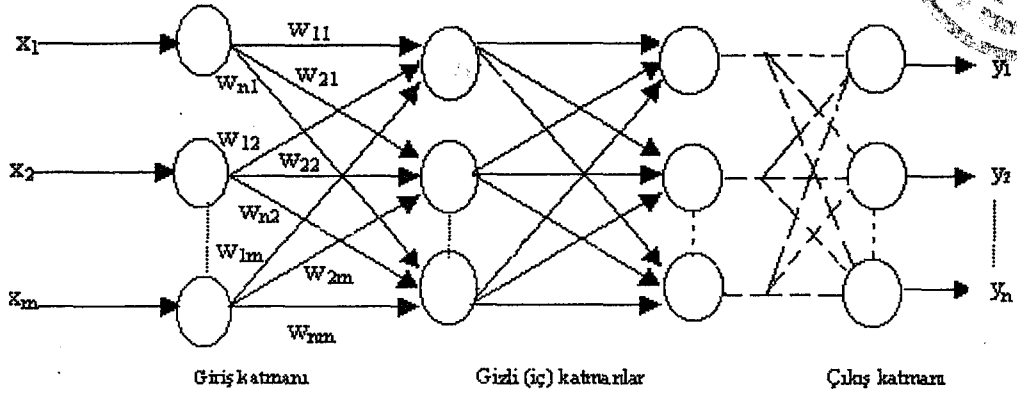
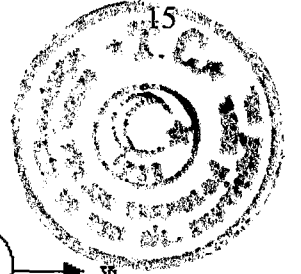


d



e

Şekil 2.2. Farklı Ağ Bağlantıları



Şekil 2.3. Çok katmanlı ağ yapısı

Öğrenme Kuralları: Bir YSA'yı tanımlamada üçüncü önemli unsur öğrenme kuralıdır. Öğrenme kuralları iki ana gruba ayrılır. Bunlardan birincisi parametrik öğrenme dediğimiz ağ parametrelerinin değişmesiyle gerçekleştirilen öğrenme yöntemleri, diğeri ise yapısal öğrenme denilen ağ yapısının değiştirilmesi ile gerçekleştirilen öğrenme yöntemleridir. Yapısal öğrenmede, her bir eğitim aşamasında, ağa yeni hücreler katılır veya o an için mevcut olan, fakat ağ performansını etkilemeyen hücreler ağdan çıkarılır. Bir başka yöntem de, ağda yeni bağlantıların oluşturulması veya varolan bağlantıların kaldırılmasına dayanmaktadır. Bu işlemleri gerçekleştiren algoritmalar ise genetik algoritmalarlardır. Tez içerisinde parametrik öğrenme yöntemleri kullanıldığından ve genelde uygulamalarda parametrik öğrenme yöntemleri yaygın olduğundan, yapısal algoritmalar üzerinde daha fazla durulmayacaktır.

Parametrik öğrenme kurallarında değiştirilen ağ paramatereleri ağıın bağlantı katsayılarıdır. Bağlantı katsayılarını içeren ağırlık matrisi aşağıdaki gibi tanımlanır.

$$W = \begin{bmatrix} w_1^T \\ w_2^T \\ \vdots \\ w_n^T \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1m} \\ w_{21} & w_{22} & \cdots & w_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nm} \end{bmatrix} \quad (2.11)$$

Öyleki $i = 1, 2, \dots, n$ olmak üzere, $w_i = (w_{i1}, w_{i2}, \dots, w_{im})^T$, i . işlem elemanının



ağırlık vektörüdür. w_{ij} 'ler ise j . hücrenin (kaynak), i . hücreye (hedef) olan bağlantısının ağırlık katsayısıdır. Öğrenme; girişlerle, çıkışlar arasında cebrik bir ilişki sağlayan, uygun ağırlık matrisinin bulunması işlemidir. Dolayısıyla öğrenme kuralı, ağırlık matrisinin, her eğitim aşamasında güncellenmesini sağlayan bir algoritmadır. Parametrik öğrenme kurallarını üç gruba ayırmak mümkündür. Bunlar:

1. Eğitici Öğrenme (Supervised learning)
2. Eğitici Öğrenme (Unsupervised learning)
3. İşaretçi vasıtası ile Öğrenme (Reinforcement learning)

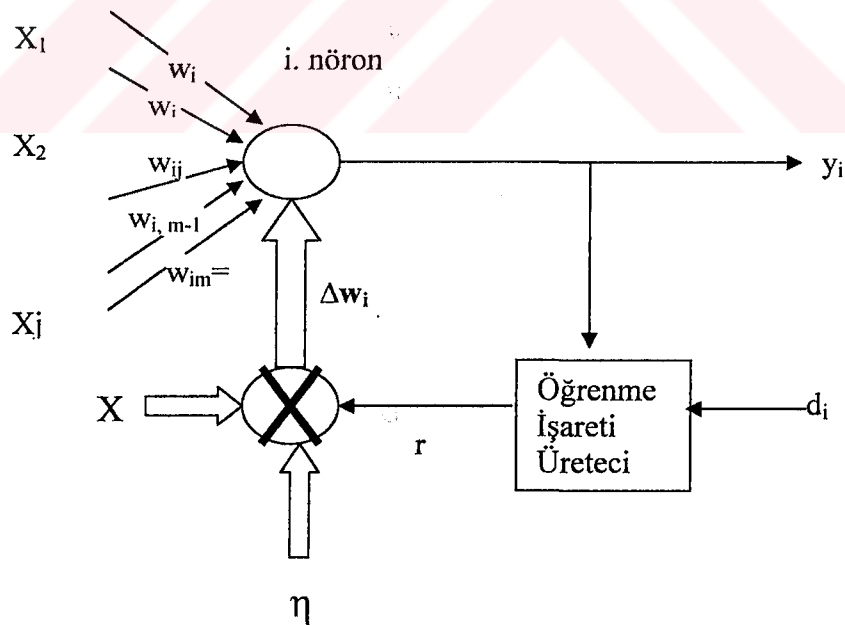
Eğitici Öğrenme: Eğitici öğrenmede her bir adımda giriş verilerini içeren bir X giriş vektörü ile bu giriş vektörüne ilişkin çıkışta görülmesi beklenen çıkış değerlerini içeren d çıkış vektörü ağa yüklenir. Dolayısıyla ağ çıkışı, ilgili girişe ilişkin çıkış değerini tam olarak bilmektedir. Daha açıkçası, eğitici öğrenmede, YSA örnek giriş-çıkış çiftleri $(x^{(1)}, d^{(1)})$, $(x^{(2)}, d^{(2)})$, ..., $(x^{(k)}, d^{(k)})$, ..., ile eğitilir. Her bir $x^{(k)}$ vektörü YSA girişine verilirken aynı anda bu vektöre ilişkin çıkışta görülmesi beklenen $d^{(k)}$ vektörü de YSA'ya yüklenir. Her bir eğitim aşamasında YSA ürettiği mevcut çıkış değeri ile beklenen çıkış değerini karşılaştırarak bir hata sinyali üretir. Bu hata sinyaline göre de ağ parametreleri yani, ağırlık matrisi güncellenir. Eğitim işlemi, çıkışta arzu edilen değer belirinceye kadar ya da sözkonusu değere istenilen tolerans da ulaşıncaya kadar devam edilir.

İşaretçi vasıtasıyla öğrenme: Eğitici öğrenmede her giriş deseni için çıkışta ulaşılacak hedef desen, tam doğru bir şekilde bilinmektedir. Fakat bazı durumlarda, oldukça kısıtlı bilgiye sahip olunması sözkonusudur. Uç bir durum olarak, ulaşılan sonucun sadece doğru yada yanlış olduğunu gösteren, yalnızca bir bitlik geribesleme bilgisi, ağ parametrelerini güncellemek için kullanılır. Böyle bir kritik bilgiyi temel olarak gerçekleştirilen öğrenme algoritmasına işaretçi vasıtalı (reinforced) öğrenme algoritması denir. Bu algoritma, ağın bir geribesleme sinyali ile ağırlık katsayılarını güncellemesinden dolayı eğitici öğrenme yöntemine benzemektedir. Ancak eğitici öğrenmede hedef deseninin kesin bir biçimde bilinmesine rağmen burada sadece o anki ağ çıkışındaki bilginin doğru mu yoksa yanlış mı olduğu bilinmektedir.



Eğiticişiz öğrenme: Eğiticişiz öğrenme de ise, eğiticiş öğrenmenin aksine bir öğretici yoktur. Eğiticiş öğrenmede örnek giriş-çıkış desen çiftleri bir öğrenci gibi ağa verilerek ağın eğitimi yapılır. Bu öğrenme yönteminde ise herhangi bir giriş desenine karşı çıkış deseninin ne olması gerektiği belli değildir. Ağ böyle bir bilgiyle donatılmamıştır. Dolayısıyla ağın o andaki sahip olduğu ağırlık matrisiyle giriş desenine karşılık elde edilen çıkış deseninin doğru olup olmadığı bilinmemektedir. Ağ girişindeki verilere ait çıkış desenini, bu girişlere ait öznelilikleri, her bir desen arasındaki ilişkiyi, verilerin dağılımını ve söz konusu örnekleri gruplandırmayı kendi kendine keşfetmek zorundadır. Bu işlemleri yapmak, gerçekleştirmek için sürekli belli bir kural dahilinde ağ parametrelerini değiştirir. Bu işleve kendi kendini yönetme denir. Giriş desenlerinin başka bir bilgiye ihtiyaç duyulmadan gruplandırılması, kümelendirilmesi bu tür uygulamalara örnek olarak verilebilir. Desenler arasındaki benzerlikler ve farklılıklar ağ tarafından keşfedilerek her bir desenin ait olduğu sınıf belirlenir [Lin and Lee, 1995].

Genelleştirilmiş Öğrenme Kuralı: Üç temel öğrenme kuralı göz önüne alınarak genelleştirilmiş bir öğrenme kuralı çıkarılabilir. Genelleştirilmiş öğrenme



Şekil 2.4. Eğiticiş Öğrenme

kuralında w ağırlık matrisindeki değişme miktarı Δw , $r(t)$ öğrenme işareti ile $x(t)$ giriş işaretinin çarpımıyla orantılıdır. Yani

$$\Delta w_i(t) \propto r \cdot x(t) \quad \text{veya} \quad \Delta w_i(t) = \eta \cdot r \cdot x(t) \quad (2.12)$$

şeklinde ifade edilir. Burada η öğrenme oranını gösteren pozitif bir sayıdır ve öğrenme sabiti denir. r ise, w_i, x ve eğer eğitici öğrenme sözkonusu ise d 'ye bağlı olan öğrenme işaretini gösterir. Yani;

$$r = f_r(w_i, x, d_i) \quad (2.13)$$

Dolayısıyla, genelleştirilmiş öğrenme kuralına göre her hangi bir andaki w_i ağırlık matrisi ;

$$w_i(t+1) = w_i(t) + \eta \cdot f_r(w_i(t), x(t), d_i(t)) \cdot x(t) \quad (2.14)$$

dolayısıyla;

$$w_i^{(t+1)} = w_i^{(t)} + \eta \cdot f_r(w_i^{(t)}, x^{(t)}, d_i^{(t)}) \cdot x^{(t)} \quad (2.15)$$

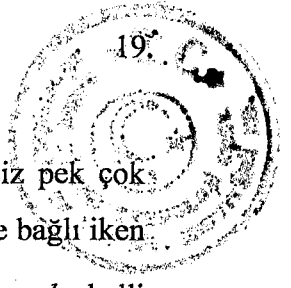
yazılabilir.

Eşitlik (2.14) öğrenmenin ayrık zamandaki ifadesidir. Sürekli zaman için herhangi bir t anındaki ağırlık matrisindeki değişim;

$$\frac{dw_i(t)}{dt} = \eta \cdot r \cdot x(t) \quad (2.16)$$

ifadesiyle gösterilir.

Eşitlik (2.15) ve (2.16) de görüldüğü gibi öğrenme işlemine başlamadan önce bir başlangıç ağırlık matrisi (genelde rastgele ağırlıklar içeren) belirlenmelidir.



Genelleştirilmiş öğrenme kuralı kullanılarak eğitici veya eğitici olmayan pek çok öğrenme kuralı türetilir. Eğitici öğrenmede, öğrenme işareti r , d_i 'ye bağlı iken eğitici olmayan öğrenmede beklenen çıkış değerleri sözkonusu olmadığından d_i belli değildir, dolayısıyla öğrenme işareti, x giriş işareti ve y çıkış işaretine bağlıdır.

Hebb öğrenme yasasının genelleştirilmiş öğrenme kuralına uyarlırsak, Hebb öğrenme kuralında öğrenme işareti r 'nin tanımı aşağıdaki gibidir.

$$r = a(w_i^T x) = y_i \quad (2.17)$$

Burada $a(\bullet)$ hücrenin transfer fonksiyonunu göstermektedir. Görüldüğü gibi Hebb öğrenme yasasında, öğrenme işareti hücrenin çıkış işaretidir. Dolayısıyla ağırlık matrisindeki değişim için;

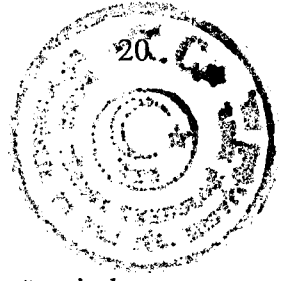
$$\Delta w_i = \eta a(w_i^T x) x_j = \eta y_i x_j \quad i = 1, 2, \dots, n; \quad j = 1, 2, \dots, m \quad (2.18)$$

Bu bağıntıda Hebb öğrenme yasasının eğitici olmayan bir öğrenme kuralı olduğu görülmektedir. Çünkü öğrenme işareti ağırlık giriş değerleri ile çıkış değerlerine bağlıdır. Şayet giriş ile çıkış işaretinin çarpımı pozitif ise, ağırlık katsayıları artarken, negatif olması durumunda ise ağırlık katsayıları azalacaktır.

Bu çalışmada yalnızca eğitici öğrenme yöntemleri kullanıldığından diğer öğrenme yöntemleri üzerinde daha fazla durulmayacaktır.

2.3. Eğitici Öğrenme Yöntemleri

YSA örnek giriş-çıkış çiftleri ile eğitilerek bir eğitim aşamasından geçiyorsa bu tip bir öğrenme yöntemine eğitici öğrenme adı verildiği daha önceki bölümlerde belirtilmişti. Bu bölümde sıklıkla kullanılan eğitici öğrenme yöntemine haiz YSA modelleri tanıtılacaktır. Özellikle tez içerisinde kullanılan geriye yayılma algoritması (backpropagation) ve çağrışımlı bellek yapısı gösteren Hopfield ağ modelleri üzerinde durulacaktır.



2.3.1 Tek Tabakalı Perceptron Ağları

Tek tabakalı perceptron ağlarında girişler ağırlık katsayıları üzerinden hücrelere bağlıdır. Eğiticili öğrenme söz konusu olduğundan bazı örnek giriş değerleri için çıkış değerleri bilinmektedir ve ağ bu giriş çıkış çiftleri ile eğitilir. Şekilde görüldüğü gibi $x^k = [x_1^{(k)}, x_2^{(k)}, \dots, x_m^{(k)}]^T$, giriş değerleri için beklenen çıkış deseni $d^k = [d_1^{(k)}, d_2^{(k)}, \dots, d_n^{(k)}]^T$, olmalıdır. Eğitim işlemi sonucunda ulaşılan $y^k = [y_1^{(k)}, y_2^{(k)}, \dots, y_n^{(k)}]^T$, aktif çıkış değerleri, beklenen $d^{(k)}$, çıkış desenine eşit olmalıdır. Burada m girişlerin sayısı, n çıkışların sayısı, k ise örnek sayısıdır.

$$y_i^{(k)} = a(w_i^T x^{(k)}) = a\left(\sum_{j=1}^m w_{ij} x_j\right) = d_i^{(k)} \quad i = 1, 2, \dots, n; \quad k = 1, 2, \dots, p \quad (2.19)$$

$a(\bullet)$ transfer fonksiyonu bir eşik fonksiyonu veya sürekli sigmoid fonksiyonu olabilir. İşlem elemanının türüne göre uygulanacak öğrenme kuralı da değişir. İşlem elemanı LTU ise perceptron öğrenme kuralı, LGU ise Widrow-Hoff öğrenme kuralı kullanılır.

2.3.2. Perceptron Öğrenme Kuralı

Transfer fonksiyonu olarak signum fonksiyonunu kullanan perceptron ağının çıkışında beklenen değerler $d_i^{(k)}$, ancak ± 1 olabilir. Dolayısıyla çıkış için;

$$y_i^{(k)} = \text{sgn}(w_i^T x^{(k)}) = d_i^{(k)}, \quad i = 1, 2, \dots, n; \quad k = 1, 2, \dots, p \text{ yazılır.} \quad (2.20)$$

Bu ilişkiyi sağlayabilmek için oluşturulacak w ağırlık matrisini güncelleyebilmek için gerekli öğrenme işareti, beklenen çıkış ile aktif çıkış arasındaki farktan elde edilir.

$$r = d_i - y_i \quad (2.21)$$



Beklenen çıkış d_i değerleri ∓ 1 olacağından ağırlık matrisindeki değişim için

$$\Delta w_{ij} = \eta [d_i - \text{sgn}(w_i^T x)] x_j = \begin{cases} 2\eta \cdot d_i x_j & y_i \neq d_i \\ 0 & \text{diğer} \end{cases} \quad j = 1, 2, \dots, m \quad (2.22)$$

bağıntısı geçerlidir.

Buradaki η öğrenme sabitinin büyük değerler alması öğrenmeyi hızlandırır. Çok fazla büyük değerler alması durumunda ise eğitim işleminin, yerel optimumlara takılması durumu ile karşılaşılır. Bu nedenle öğrenme sabitinin alacağı en uygun değer her problemde farklılıklar gösterir.

2.3.3. Adaline Öğrenme Kuralı

Eğitim kümesinin ve çıkışta beklenen değerlerin ayrık değerlerden değilde sürekli işaretlerden oluşması durumunda işlem elemanlarının (hücrelerin) aktivasyon fonksiyonları da buna uygun seçilmelidir. Bilindiği gibi sigmoid fonksiyonu gibi operasyonlar belli bir aralıkta sürekli değerler üretilmesine imkan tanırırlar.

Tek tabakalı lineer aktivasyon fonksiyonuna sahip ağa adaptif lineer element denir ve kısaca ADALINE ismiyle anılır. ADALINE ağında girişlerle çıkışlar arasındaki ilişki lineerdir. Eğitici öğrenmede $(x^{(1)}, x^{(2)}, \dots, x^{(p)})$ giriş örneklerine ait çıkışta $(d^{(1)}, d^{(2)}, \dots, d^{(p)})$ çıkış desenlerinin olacağı ağa eğitim aşamasında yüklenir. Burada ulaşılması istenen hedef, bu ilişkiyi sağlayacak w_i matrisinin bulunmasıdır. Öyleki

$$\sum_{j=1}^m w_j x_j^{(k)} = d^{(k)} \quad k = 1, 2, \dots, p. \quad (2.23)$$

eşitliği sağlanmalıdır.



Eğer giriş desenleri $(x^{(1)}, x^{(2)}, \dots, x^{(p)})$ lineer bağımsız olursa, eşitlik sağlanarak w^* matrisi bulunur. Giriş desenleri $x^{(i)}$ lerin boyu m ise, p adet farklı desen arasındaki lineer bağımsızlık şartı için gerek şart $p \leq m$ olmalıdır.

Eşitlik (2.23) den ağırlık matrisini elde etmek için önce sistemin hatasını ölçen bir hata fonksiyonu $E(w)$ tanımlamak gereklidir.

$$E(w) = \frac{1}{2} \sum_{k=1}^p (d^{(k)} - y^{(k)})^2 = \frac{1}{2} \sum_{k=1}^p (d^{(k)} - w^T x^{(k)})^2 \text{ ve}$$

$$E(w) = \frac{1}{2} \sum_{k=1}^p \left(d^{(k)} - \sum_{j=1}^m w_j x_j^{(k)} \right)^2 \quad (2.24)$$

Açıkça görüldüğü gibi $E(w)$ normalde pozitifdir ve $y^{(k)}, x^{(k)}$ 'ya yaklaştıkça sıfıra yaklaşır. $E(w)$ küçüldükçe ideal w değerine ulaşılabileceği görülmektedir. O halde yapılması gereken ortalama karesel hatayı yani $E(w)$ değerini minimize etmektir. Yapay sinir ağlarındaki öğrenme işlemini diğer geleneksel işlemlerden ayıran belki de en önemli özellik budur. Yani w matrisini bulmak için analitik çözümler yapmak yerine w 'nın rasgele bir başlangıç değerinden yola çıkarak en uygun çözümü buluncaya kadar w matrisini güncelleme işlemini gerçekleştirecek bir öğrenme algoritması kullanılmalıdır. Sık kullanılan algoritmalarından birisi Eğilim düşme (Gradient -descent) Öğrenme Algoritmasıdır. Bu yöntemle ağırlık matrisinin güncellenmesinde kullanılan Δw değişim miktarının bulunmasında, $E(w)$ fonksiyonun gradientinin ters işaretlisi kullanılır.

$$\Delta w = \eta \nabla_w E(w). \quad (2.25)$$

$$\Delta w_j = \eta \frac{\partial E}{\partial w_j} = \eta \sum_{k=1}^p (d^{(k)} - w^T x^{(k)}) x_j^{(k)} \quad j = 1, 2, \dots, m \quad (2.26)$$

Eğer bütün sınıflar topluca değil de sırayla ağın eğitimi için ağa sunulursa, herhangi bir $x^{(k)}$ giriş deseni için ağırlık katsayılarındaki değişim;

$$\Delta w_j = \eta (d^{(k)} - w^T x^{(k)}) x_j^{(k)} \quad (2.27)$$

eşitliğindeki gibi olacaktır.

Yukarıdaki eşitliğe Adaline öğrenme kuralı veya Widrow-Hoff öğrenme kuralı adı verilir. Bu aynı zamanda endüyük kareler ortalaması (LMS) yöntemi adını alır. Genelleştirilmiş öğrenme kuralı ile karşılaştırıldığında burada öğrenme sinyali;

$$r = d - y$$

alınmıştır.

2.4. Geriye Yayılma (Back Propagation)

Sinir ağlarının tarihsel gelişimi içinde en önemli öğrenme algoritmalarından birisi geriye yayılma (back propagation) algoritmasıdır. [Bryson and Ho, 1969; Werbos, 1974; LeCun, 1985; Parker, 1985; Rumelhart et al., 1986a,b]. Bu öğrenme algoritması çok tabakalı ileri beslemeli ağlar ile birleştirildiğinde elde edilen yapay sinir ağına geriye yayılma ağı adı verilir. Eğitim kümesindeki $\{x^{(k)}, d^{(k)}\}$ giriş-çıkış çiftleri verildiğinde geriye yayılma ağının giriş desenlerini doğru bir şekilde sınıflandırabilmesi için öğrenme kuralı vasıtasıyla ağırlık matrisi değiştirilir. Ağırlıkların güncellenmesi, perceptron ağlarında kullanılan eğim düşümü (gradient-descent) algoritmasına dayanır. Geriye yayılma ağında iki yönlü bilgi akışı söz konusudur. İlk olarak giriş deseni $x^{(k)}$, ağ girişinden ağ çıkışına doğru gönderilir. Bu giriş desenine karşılık çıkışta oluşan $y^{(k)}$ çıkış deseni ile çıkışta görülmesi beklenen $d^{(k)}$ deseni arasındaki fark, yani hata sinyali çıkıştan geriye, giriş tabakasına doğru iletilir. Ağ parametrelerinin güncellenmesi bu hata sinyali ile yapılır. Bunları daha açık olarak görmek amacıyla şekildeki geriye yayılma ağını dikkate alınırsa, sürekli çizgiler giriş işaretinin ileriye doğru yayılmasını, kesikli çizgiler ise hata işaretinin geriye doğru yayılmasını göstermektedir.

İfadeleri basitleştirmek için k indisi kaldırılmış bir giriş- çıkış eğitim çifti (x, d) ele alınırsa, verilen x giriş deseni için saklı tabakadaki q işlem elemanın alacağı toplam giriş işareti aşağıdaki gibi olur.

$$net_q = \sum_{j=1}^m v_{qj} x_j \quad (2.28)$$

ve hücre çıkışında üretilen işaret;

$$z_q = a(net_q) = a\left(\sum_{j=1}^m v_{qj} x_j\right) \quad (2.29)$$

olacaktır. Çıkış katmanındaki i . hücrenin girişine ulaşan toplam giriş işareti ise;

$$net_i = \sum_{q=1}^l w_{iq} z_q = \sum_{q=1}^l w_{iq} a\left(\sum_{j=1}^m v_{qj} x_j\right) \quad (2.30)$$

olur ve çıkış işareti;

$$y_i = a(net_i) = a\left(\sum_{q=1}^l w_{iq} z_q\right) = a\left(\sum_{q=1}^l w_{iq} a\left(\sum_{j=1}^m v_{qj} x_j\right)\right) \quad (2.31)$$

Yukarıdaki eşitlikler back propagation ağında ileri doğru yayılma denklemleridir. Geriye doğru yayılma eşitliklerini elde etmek için ilk olarak hatayı belirlemek gerekir. Toplam karesel hatayı belirleyen hata fonksiyonu;

$$E(w) = \frac{1}{2} \sum_{i=1}^n [d_i - a(net_i)]^2 \quad \text{ve}$$

$$E(w) = \frac{1}{2} \sum_{i=1}^n \left[d_i - a\left(\sum_{q=1}^l w_{iq} z_q\right) \right]^2 \quad (2.32)$$

Eğim düşümü algoritmasına göre, saklı tabakadan çıkış tabakasına doğru bütün ağırlık katsayıları aşağıdaki kuralla güncellenir.

$$\Delta w_{iq} = -\frac{\partial E}{\partial w_{iq}} \quad (2.33)$$

(2.30)-(2.32) eşitlikleri kullanılarak zincir kuralıyla $\partial E/\partial w_{iq}$ için;

$$\Delta w_{iq} = -\eta \left[\frac{\partial E}{\partial y_i} \right] \left[\frac{\partial y_i}{\partial net_i} \right] \left[\frac{\partial net_i}{\partial w_{iq}} \right] = \eta [d_i - y_i] [a'(net_i)] [z_q] = \eta \delta_{oi} z_q \quad (2.34)$$

Bu son eşitlikteki δ_{oi} hata işaretidir ve ikinci indis çıkış tabakasındaki i . hücreyi gösterir. Hata işareti aşağıdaki gibi tanımlanır;

$$\delta_{oi} = -\frac{\partial E}{\partial net_i} = -\left[\frac{\partial E}{\partial y_i} \right] \left[\frac{\partial y_i}{\partial net_i} \right] = [d_i - y_i] [a'(net_i)] \quad (2.35)$$

Bu tanım aynı zamanda delta öğrenme kuralında kullanılan hata işaretinin tanımıdır. Çıkış katmanındaki i . hücreye giren toplam işaret net_i ile gösterilirken $a'(net_i)$ için aşağıdaki tanım yapılır.

$$a'(net_i) = \frac{\partial a(net_i)}{\partial net_i} \quad (2.36)$$

Eşitlikteki z_q ise saklı tabakanın çıkış değeridir. Giriş katmanı ile, saklı tabaka arasındaki ağırlıkların güncellenmesi için de eğim düşümü yöntemi kullanılır. Giriş tabakasındaki j . hücre ile, saklı tabakadaki q . hücre arasındaki bağlantı ağırlığındaki değişim miktarı;

$$\Delta v_{qj} = -\eta \left[\frac{\partial E}{\partial v_{qj}} \right] = -\eta \left[\frac{\partial E}{\partial net_q} \right] \left[\frac{\partial net_q}{\partial v_{qj}} \right] \text{ ve}$$

$$\Delta v_{qj} = -\eta \left[\frac{\partial E}{\partial z_q} \right] \left[\frac{\partial z_q}{\partial net_q} \right] \left[\frac{\partial net_q}{\partial v_{qj}} \right] \quad (2.37)$$

zincir kuralıyla elde edilir.

Eşitlik (2.32)'den açıkça görüldüğü gibi her bir hata terimi $[d_i - y_i]$, z_q 'nin fonksiyonudur. Zincir kuralının uygulanmasıyla aşağıdaki ifade elde edilir.

$$\Delta v_{qj} = \eta \sum_{i=1}^n [(d_i - y_i) \cdot a'(net_i) w_{iq}] a'(net_q) x_j \quad (2.38)$$

Eşitlik (2.35) kullanılarak, bu son eşitlik yeniden yazılabilir;

$$\Delta v_{qj} = \eta \sum_{i=1}^n [d_{oi} w_{iq}] a'(net_q) x_j = \eta \delta_{hq} x_j \quad (2.39)$$

Buradaki δ_{hq} , saklı tabakadaki q . elemana ait hata işaretidir ve aşağıdaki gibi tanımlanır;

$$\delta_{hq} = -\frac{\partial E}{\partial net_q} = -\left[\frac{\partial E}{\partial z_q} \right] \left[\frac{\partial z_q}{\partial net_q} \right] = a'(net_q) \sum_{i=1}^n \delta_{oi} w_{iq} \quad (2.40)$$

eşitlikteki net_q saklı tabakadaki q . elemana ait toplam giriş işaretini göstermektedir. Saklı tabakadaki bir hücreye ait hata işaretinin, çıkış tabakasındaki bir hücreye ait hata sinyalinden farklı olduğu görülmektedir. Bundan dolayı (2.40) da tanımlı olan hata işaretini kullanan ağırlık güncelleme eşitliği, eşitlik (2.38) ve (2.39)'a genelleştirilmiş delta öğrenme kuralı ismi verilir. Görüldüğü gibi saklı tabakadaki hata işareti δ_{hq} , çıkış tabakasındaki hata işareti δ_{oi} 'nin fonksiyonudur. Daha açıkçası hata işareti çıkış tabakasından girişe doğru yayılmaktadır. Eşitlik (2.30) ve (2.32) genelleştirilmiş öğrenme kuralı açısından yaklaşılmca birbirinin aynıdır. Aralarındaki fark $r = \delta$ hata işaretinin tanımlarındadır. Dolayısıyla geriye



yayılma ağlarında, herhangi bir katman için genelleştirilmiş ağırlık güncelleme ifadesi;

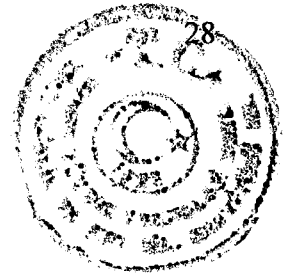
$$\Delta w_{ij} = \eta \delta_i x_j = \eta \delta_{i.cikis} x_{j.giris} \quad (2.41)$$

Bu ifadedeki "i. çıkış" ve "j. giriş" ifadeleri j. hücreden i. hücreye doğru olan bağlantının uç noktalarını gösterir. Saklı tabakadaki veya giriş tabakasındaki giriş ucu x_j ile temsil edilirken, δ_i ise son çıkış tabakası için eşitlik (2.35), diğer katmanlar için ise eşitlik (2.40) ile tanımlanan hata işaretidir. Hücrelerde transfer fonksiyonu olarak bipolar sigmoid fonksiyonunun kullanılması durumunda hata işaretine ilişkin ifadeler aşağıdaki ifadelere dönüşür.

$$\delta_{oi} = \frac{1}{2} (1 - y_i^2) [d_i - y_i] \quad (2.42)$$

ve

$$\delta_{hq} = \frac{1}{2} (1 - z_q^2) \sum_{i=1}^n \delta_{oi} w_{iq} \quad (2.43)$$



3. HÜCRESEL YAPAY SİNİR AĞLARI

Hücresel Yapay Sinir Ağları (HYSA) bir yapay sinir ağı modeli olup, özellikle görüntü işleme uygulamalarında oldukça başarılı bir performans sergiler [Wang, et al, 1998]. CNN iki boyutlu hücre dizilerinden oluştuğu için görüntüdeki her bir piksele karşı bir hücre düşürülerek işlemler yapılır. Hücre çıkışları $[-1, 1]$ aralığında değerler alır. Görüntüdeki siyah noktalar '-1' ve beyaz noktalar '1' değerleri ile sembolize edildiğinde, herhangi bir hücrenin çıkışı, o hücreye karşı düşen pikselin parlaklık değerini gösterir.

Chua ve Yang tarafından 1988 yılında tanıtılan Hücresel Yapay Sinir Ağları ısı yayılımı, dalga yayılımı gibi nonlinear olayların simülasyonunda ve analizinde başarılı bir yöntemdir [Chua and Yang, 1988]. Hücresel Yapay Sinir Ağının iki boyutlu bir matris yapısında olması nedeniyle benzer problemlerin iki boyutlu düzlemdeki analizinde ve özellikle görüntü işleme alanında tercih edilmesine sebep olmuştur. Ayrıca ağdaki hücrelerin durum denklemindeki dinamik davranış, hücresel yapay sinir ağını doğal olayların simülasyonunda başarılı kılan bir diğer önemli etkidir. Hücresel yapay sinir ağına ait eşitlikler aşağıdaki gibidir.

Durum Denklemi:

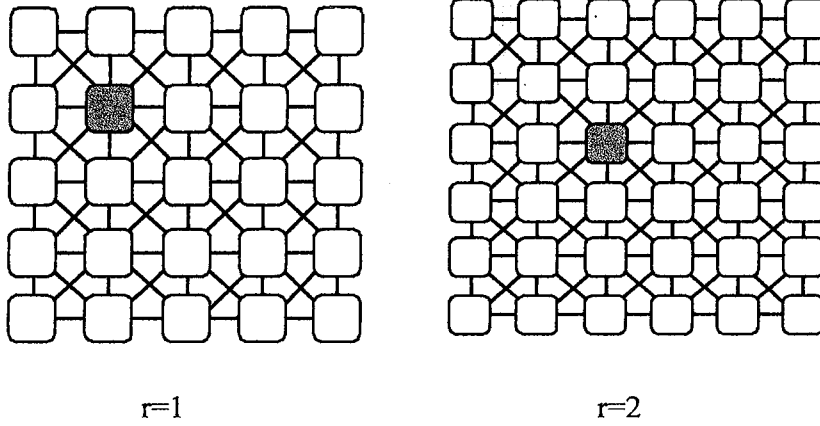
$$C \frac{dx_{ij}(t)}{dt} = -\frac{1}{R_x} x_{ij}(t) + \sum_{C(k,l) \in N_r(i,j)} A(i,j;k,l) y_{kl}(t) + \sum_{C(k,l) \in N_r(i,j)} B(i,j;k,l) u_{kl}(t) + I \quad (3.1)$$

Bu denklem, 'r' komşuluğundaki bir hücresel yapay sinir ağı için durum denklemdir. $M \times N$ boyutundaki bir ağ için r komşuluğu içerisindeki hücreler aşağıdaki tanımlanır (Şekil 3.1).

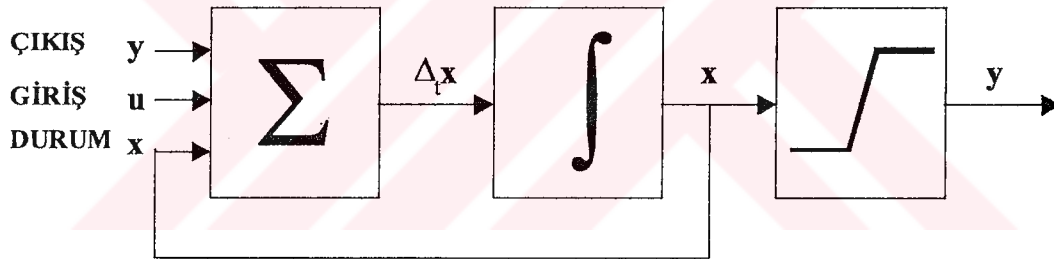
$$N_r(i,j) = \{C(k,l) \mid \max\{|k-i|, |l-j|\} \leq r, 1 \leq k \leq M, 1 \leq l \leq N\} \quad (3.2)$$



Şekil 3.1' de aynı hücrenin $r=1$ ve $r=2$ komşuluğu şartı altında, komşu olduğu hücreler gösterilmektedir. Burada görüldüğü gibi $r=1$ durumunda merkezdeki hücreye komşu olan hücre sayısı kendisi dahil 9, $r=2$ olması durumunda komşu hücre sayısı 25, $r=3$ için 49 ve en genel halde r komşuluğu altında, komşuluk sınırları içerisindeki hücre sayısı $(2r+1)^2$ olacaktır.

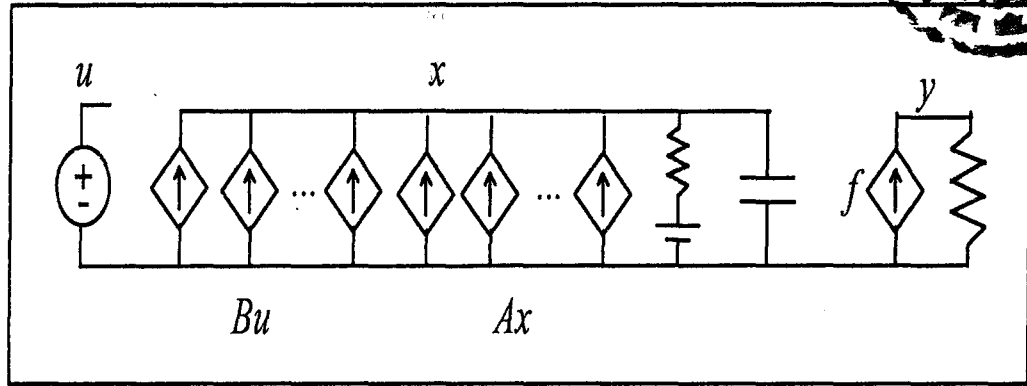
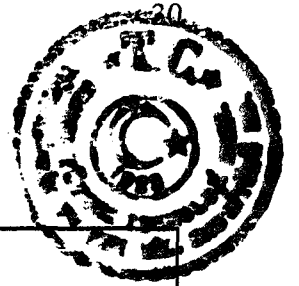


Şekil 3.1. $r=1$ ve $r=2$ komşuluğundaki hücresel yapay sinir ağları



Şekil 3.2. Hücresel yapay sinir ağı hücresinin matematik modeli

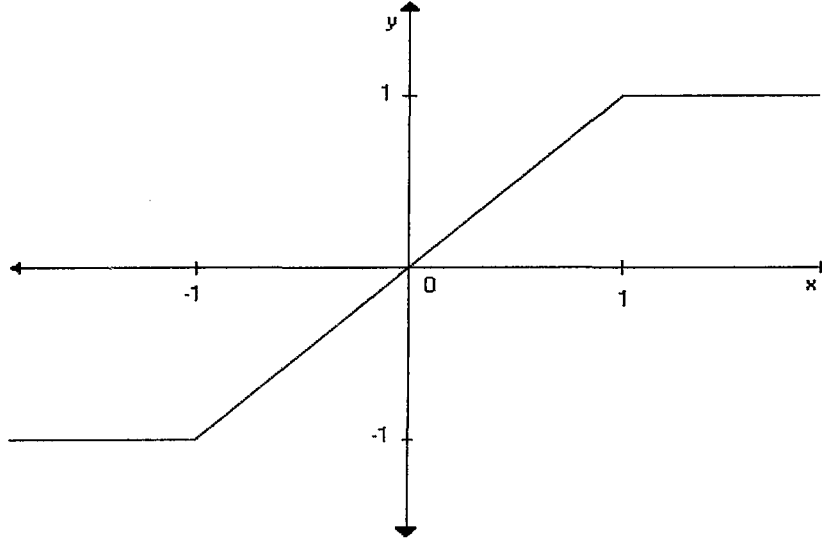
Burada x_{ij}, y_{ij} ve u_{ij} , $C(i, j)$ hücresine ait sırasıyla durum, çıkış ve giriş değişkenidir. A, B ve I ise sırasıyla, geri yöndeki geri besleme şablonu, ileri yöndeki ileri besleme şablonu ve eşik (offset) şablonudur. Bu şablonların boyutu ağı boyutundan bağımsız olup, ağıdaki komşuluğa bağlıdır ve şablonlar ağı her yerinde aynıdır, yani, yerel değişmez (local invariant) özelliktedirler. Denklemdaki C ve R_x ise hücre eşdeğerindeki kapasitenin sığasını ve direncin değerini gösterir ve genel olarak bu değerler 1 seçilir. Şekil 3.3.'de ağ içerisinde bir hücrenin eşdeğer devresi şekilde gösterilmektedir.



Şekil 3.3. HYSA içerisindeki hücrelerin eşdeğer devresi

Şekilde görüldüğü gibi bir hücresel yapay sinir ağı hücresi, bir gerilim kaynağı u_{ij} , bağımsız bir akım kaynağı (I), bir kapasite (C), iki lineer direnç, (R_x, R_y) ve diğer hücrelerin çıkışı ile kontrol edilen, m komşu sayısı olmak üzere en fazla $2m$ tane lineer gerilim kontrollü akım kaynağından oluşmaktadır. Bu kaynakların karakteristikleri; $C(k,l) \in N_r(i,j)$ için $I_{xy}(i,j;k,l) = A(i,j;k,l)V_{ykl}$ ve $I_{xu}(i,j;k,l) = B(i,j;k,l)V_{ukl}$ şeklinde tanımlanmaktadır. Burada $A(i,j;k,l)$ ve $B(i,j;k,l)$ katsayıları $C(i,j)$ ve $C(k,l)$ hücreleri arasındaki bağlantı ağırlık katsayılarıdır. $A(i,j;k,l)$ katsayılarına geri besleme ağırlık katsayıları $B(i,j;k,l)$ katsayılarına giriş ağırlık katsayıları veya ileri yön bağlantı katsayıları adı verilmektedir. Hücresinin çıkışında ise parça parça doğrusal olan, $I_{yz} = (1/R_y)f(V_{xij})$ şeklinde tanımlı olan gerilim kontrollü bir akım kaynağı vardır. Parça parça doğrusal olan $f(\bullet)$ fonksiyonunun davranışı Şekil 3.4' de gösterilmiş olup tanım bağıntısı aşağıdaki eşitlikteki gibi olan bir sigmoid fonksiyonu tipindedir.

$$y_{ij}(t) = \frac{1}{2} \left(|x_{ij}(t) + 1| - |x_{ij}(t) - 1| \right) \quad (3.3)$$



Şekil 3.4. Standart parça parça doğrusal aktivasyon fonksiyonu

Eşitlik (1)'de görüldüğü gibi HYSA dinamik bir davranışa sahiptir. Bundan dolayı da girişine bir işaret uygulandığında çıkış, belli bir zaman içerisinde sürekli değişir. Bu dinamik davranış hücre eşdeğerindeki kondansatörden kaynaklanmaktadır. HYSA kararlı duruma ulaştığında, kararlı hücre çıkışı Eşitlik 3.4' deki gibi olacaktır. Ağın kararlı duruma ulaşma süresi ise $\tau = R_x C$ zaman sabiti ile ilgili olacaktır. HYSA'nın dinamik davranışı sonraki bölümlerde daha geniş olarak açıklanacaktır.

Kararlı hücre çıkışı:

$$y_{ij}^{\infty} = y_{ij}(t) \Big|_{\frac{dx}{dt}=0} \quad (3.4)$$

HYSA' nın davranışını daha iyi analiz edebilmek için R_x direnci ve C kapasitesi 1 seçilirse, Eşitlik (1)' de verilen durum denklemini ifadesi aşağıdaki gibi basitleştirilebilir.

$$-\frac{dx_y(t)}{dt} = -x_y(t) + \sum_{C(k,l) \in N_r(i,j)} A(i,j;k,l)y_{kl}(t) + \sum_{C(k,l) \in N_r(i,j)} B(i,j;k,l)u_{kl}(t) + I \quad (3.5)$$

Bu ifade de konvolüsyon operatörleri yardımıyla en genel halde aşağıdaki gibi yazılabilir.

$$\frac{dX}{dt} = -X + A * Y + B * U + I \quad (3.6)$$

3.1. Hücresel Yapay Sinir Ağı'nın Dinamik Davranışı

Bu bölümde HYSA eşdeğer devresinde bulunan C kondansatöründen dolayı ağda ortaya çıkan dinamik davranış aşağıdaki belirtilen şartlar altında incelenecektir. Başlangıçta $M * N$ boyutlarında bir HYSA göz önüne alınsın ve bu ağ aşağıdaki şartları sağlasın;

Durum Denklemi;

$$C \frac{dx_{ij}(t)}{dt} = -\frac{1}{R_X} x_{ij}(t) + \sum_{C(k,l) \in N_p(i,j)} A(i,j;k,l) y_{kl}(t) + \sum_{C(k,l) \in N_e(i,j)} B(i,j;k,l) u_{kl}(t) + I, \quad (3.7.a.)$$

$$1 \leq i \leq M; 1 \leq j \leq N.$$

Çıkış Denklemi;

$$Y_{ij}(t) = \frac{1}{2} \left(|x_{ij}(t) + 1| - |x_{ij}(t) - 1| \right), \quad 1 \leq i \leq M; 1 \leq j \leq N. \quad (3.7.b.)$$

Giriş Denklemi;

$$u_{ij} = E_{ij}, \quad 1 \leq i \leq M; 1 \leq j \leq N. \quad (3.7.c.)$$

Sınırlamalar:

$$|x_{ij}(0)| \leq 1, \quad 1 \leq i \leq M; 1 \leq j \leq N. \quad (3.7.d.)$$

$$|u_{ij}| \leq 1, \quad 1 \leq i \leq M; 1 \leq j \leq N. \quad (3.7.e.)$$

Parametre değerleri:

$$A(i, j; k, l) = A(k, l; i, j), \quad 1 \leq i \leq M; 1 \leq j \leq N. \quad (3.7.f.)$$

$$C > 0, R_x > 0. \quad (3.7.g.)$$

Teorem I: HYSA' da bütün x_{ij} durumları, her $t > 0$ anında sonludur ve $X_{\max}$ aşağıdaki formül ile hesaplanabilir [Chua and Yang, 1988] ;

$$x_{\max} = 1 + R_x |I| + R_x \max_{1 \leq i \leq M; 1 \leq j \leq N} \left[\sum_{C(k,l) \in N_r(i,j)} (|A(i, j; k, l)| + |B(i, j; k, l)|) \right] \quad (3.8)$$

İspat: Eşitlik (7.a)'yı aşağıdaki gibi yazalım;

$$\frac{dx_{ij}(t)}{dt} = -\frac{1}{R_x C} x_{ij}(t) + f_{ij}(t) + g_{ij}(t) + \hat{I}, \quad 1 \leq i \leq M; 1 \leq j \leq N. \quad (3.9.a)$$

Burada

$$f_{ij}(t) = \frac{1}{C} \sum_{C(k,l) \in N_r(i,j)} A(i, j; k, l) y_{kl}(t), \quad 1 \leq i \leq M; 1 \leq j \leq N. \quad (3.9.b)$$

$$g_{ij}(t) = \frac{1}{C} \sum_{C(k,l) \in N_r(i,j)} B(i, j; k, l) u_{kl}(t), \quad 1 \leq i \leq M; 1 \leq j \leq N. \quad (3.9.c)$$

ve

$$\hat{I} = \frac{I}{C} \quad (3.9.d.)$$

şeklindedir. $u = [E_{ij}]_{1 \times MN}$, MN boyutlu sabit giriş vektörünü göstermektedir. Eşitlik (9.a) birinci dereceden adi diferansiyel bir denklemdir ve çözümü;

$$x_{ij}(t) = x_{ij}(0)e^{-t/R_x C} + \int_0^t e^{-(t-\tau)/R_x C} [f_{ij}(\tau) + g_{ij}(u) + \hat{I}] d\tau \quad (3.10)$$

ifadesi ile verilir. Bu denklem aşağıdaki gibi bir mutlak değer eşitsizliğine dönüştürülerek,

$$\begin{aligned} |x_{ij}(t)| &\leq |x_{ij}(0)e^{-t/R_x C}| + \left| \int_0^t e^{-(t-\tau)/R_x C} [f_{ij}(\tau) + g_{ij}(u) + \hat{I}] d\tau \right| \\ &\leq |x_{ij}(0)e^{-t/R_x C}| + \int_0^t e^{-(t-\tau)/R_x C} [|f_{ij}(\tau)| + |g_{ij}(u)| + |\hat{I}|] d\tau \\ &\leq |x_{ij}(0)e^{-t/R_x C}| + \int_0^t e^{-(t-\tau)/R_x C} d\tau \\ &\leq |x_{ij}(0)| + R_x C [F_{ij} + G_{ij} + |\hat{I}|] \end{aligned} \quad (3.11)$$

ifadesi elde edilir. Burada;

$$F_{ij} \equiv \max |f_{ij}| \leq \frac{1}{C} \sum_{C(k,l) \in N_r(i,j)} |A(i,j;k,l)| \cdot \max_t (|y_{kl}(t)|), \quad 1 \leq i \leq M; 1 \leq j \leq N \quad (3.12.a.)$$

ve

$$G_{ij} \equiv \max |g_{ij}| \leq \frac{1}{C} \sum_{C(k,l) \in N_r(i,j)} |B(i,j;k,l)| \cdot \max_u (|U_{kl}|), \quad 1 \leq i \leq M; 1 \leq j \leq N \quad (3.12.b.)$$

şeklinde tanımlanmıştır. (7.d.) ve (7.e.) sınırlamaları sağlandığında ve her zaman $|y_{ij}(t)| < 1$ olduğundan (11) ve (12) denklemlerinden;

$$\begin{aligned} |x_{ij}(t)| &\leq |x_{ij}(0)| + R_x \left[\sum_{C(k,l) \in N_r(i,j)} |A(i,j;k,l)| \cdot \max_t (|Y_{kl}(t)|) + \sum_{C(k,l) \in N_r(i,j)} |B(i,j;k,l)| \cdot \max_u (|U_{kl}|) + |\hat{I}| \right] \\ &\leq |x_{ij}(0)| + R_x \left[\sum_{C(k,l) \in N_r(i,j)} (|A(i,j;k,l)| + |B(i,j;k,l)| + |\hat{I}|) \right], \quad 1 \leq i \leq M; 1 \leq j \leq N \end{aligned} \quad (3.13)$$

Şimdi

$$x_{\max} = \max_{(i,j)} \left\{ 1 + R_x |I| + R_x \sum_{C(k,l) \in N_r(i,j)} (|A(i,j;k,l)| + |B(i,j;k,l)|) \right\} \quad (3.14)$$

olsun. $x_{\max}$, zamana ve hücreye bağlı olmadığından, yani t ve $C(i,j)$ 'den bağımsız olduğundan ;

$$\max_i |x_{ij}| \leq x_{\max}, \quad 1 \leq i \leq M; \quad 1 \leq j \leq N \quad (3.15)$$

elde edilir. Bu sonuç da, ağ içerisindeki her hangi bir hücrenin sahip olabileceği durum değişkeninin mutlak değerinin sınırlarının Eşitlik (3.8)' deki gibi olacağını göstermektedir.

3.2. Hücresel Yapay Sinir Ağı'nın Kararlılığı

Hücresel yapay sinir ağlarının yaptığı temel işlev, verilen gri seviyeli, yani pikselleri $[-1,1]$ aralığında değerler alan bir görüntüyü, yine bu aralıkta başka bir görüntüye dönüştürmektir. Bu aşamada çıkış görüntüsünün yalnızca -1 ve 1 değerlerini alan ikili (binary) bir görüntü olacağı varsayılacaktır. Bu bölümde, ağdaki her bir hücrenin başlangıçta sahip olduğu gri seviyeli durumdan bu sabit -1 ve 1 değerlerine yakınsamasını garantileyebilmek için ağ şablonlarında yapılması gereken kısıtlamalar belirlenmeye çalışılacaktır. Dinamik doğrusal olmayan lineer devrelerin yakınsama özelliklerini incelemek için kullanılan en etkin yöntemlerden birisi Lyapunov yöntemidir. Bu nedenle ilk önce HYSA için aşağıdaki gibi tanımlanan Lyapunov fonksiyonu belirlenir.

$$E(t) = -\frac{1}{2} \sum_{(i,j)} \sum_{(k,l)} A(i,j;k,l) y_{ij}(t) y_{kl}(t) + \frac{1}{2R_x} \sum_{(i,j)} y_{ij}(t)^2 - \sum_{(i,j)} \sum_{(k,l)} B(i,j;k,l) y_{ij} u_{kl} - \sum_{(i,j)} I y_{ij}(t) \quad (3.16)$$

Burada $E(t)$ Lyapunov fonksiyonunun fiziksel anlamı pek açık olmamakla birlikte, bu fonksiyon ile ağın genelleştirilmiş enerjisi modellenmektedir. Görüldüğü

gibi Lyapunov fonksiyonu x_{ij} durum değişkenlerinden bağımsız olarak yalnızca U_{ij} girişleri ve y_{ij} çıkışları cinsinden yazılmıştır. Aşağıdaki teoremden $E(t)$ 'nin sonlu olduğu ispatlanmaktadır.

Teorem II : Eşitlik (3.16) da verilen $E(t)$ ifadesi $M \times N$ boyutlarında bir ağ için aşağıda gösterildiği gibi sonlu değerler alır [Chua and Yang, 1988].

$$\max_t (|E(t)|) \leq E_{\max} \quad (3.17.a.)$$

öyle ki;

$$E_{\max} = \frac{1}{2} \sum_{(i,j)} \sum_{(k,l)} |A(i,j;k,l)| + \sum_{(i,j)} \sum_{(k,l)} |B(i,j;k,l)| + MN \left(\frac{1}{2R_x} + |I| \right) \quad (3.17.b.)$$

İspat:

$$\begin{aligned} |E(t)| &= \frac{1}{2} \sum_{(i,j)} \sum_{(k,l)} |A(i,j;k,l)| |y_{ij}(t)| |y_{kl}(t)| \\ &\quad + \frac{1}{2R_x} \sum_{(i,j)} y_{ij}(t)^2 + \sum_{(i,j)} \sum_{(k,l)} |B(i,j;k,l)| |y_{ij}| |u_{kl}| + \sum_{(i,j)} |I| |y_{ij}(t)| \end{aligned} \quad (3.18)$$

$y_{ij}(t)$ ve u_{ij} (3.7.d.) ve (3.7.e.) gereğince sınırlı olduğundan

$$|E(t)| = \frac{1}{2} \sum_{(i,j)} \sum_{(k,l)} |A(i,j;k,l)| + \frac{MN}{2R_x} + \sum_{(i,j)} \sum_{(k,l)} |B(i,j;k,l)| + MN |I| \quad (3.19)$$

elde edilir. Bu sonuç da (3.17.a.) ve (3.17.b.) de iddia edildiği gibi $E(t)$ nin sonlu olduğunu gösterir.

Şimdi de $E(t)$ 'nin monoton olarak azalan bir fonksiyon olduğunu gösterelim.



Teorem III: $E(t)$ skaler fonksiyonu aşağıdaki gibi monoton azalan bir fonksiyondur [Chua and Yang, 1988].

$$\frac{dE(t)}{dt} \leq 0 \quad (3.20)$$

İspat: (3.7.f.) de ki simetriklik koşulunu dikkate alarak $E(t)$ 'nin zamana göre türevi alınır;

$$\begin{aligned} \frac{dE(t)}{dt} = & - \sum_{(i,j)} \sum_{(k,l)} A(i,j;k,l) \frac{dy_{ij}(t)}{dx_{ij}(t)} \frac{dx_{ij}(t)}{dt} y_{kl}(t) \\ & + \frac{1}{R_x} \sum_{(i,j)} \frac{dy_{ij}(t)}{dx_{ij}(t)} \frac{dx_{ij}(t)}{dt} y_{ij}(t) - \sum_{(i,j)} \sum_{(k,l)} B(i,j;k,l) \frac{dy_{ij}(t)}{dx_{ij}(t)} \frac{dx_{ij}(t)}{dt} u_{kl} - \sum_{(i,j)} I \frac{dy_{ij}(t)}{dx_{ij}(t)} \frac{dx_{ij}(t)}{dt} \end{aligned} \quad (3.21)$$

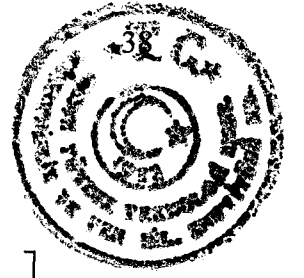
ifadesi elde edilir. Burada parça parça sürekli olan aktivasyon fonksiyonunun türevi ve kendisi için;

$$\frac{dy_{ij}(t)}{dx_{ij}(t)} = \begin{cases} 1, & |x_{ij}(t)| < 1 \\ 0, & |x_{ij}(t)| \geq 1 \end{cases} \quad (3.22.a.)$$

$$x_{ij} = y_{ij}, \quad |x_{ij}| < 1 \quad (3.22.b.)$$

tanımları yapılarak ayrıca komşuluk sınırları dışındaki hücrelerin bağlantı katsayılarını da sıfır seçerek;

$$C(k,l) \notin N_r(i,j) \text{ için } A(i,j;k,l) = 0, \quad B(i,j;k,l) = 0 \quad (3.22.c.)$$



$$\begin{aligned}
 \frac{dE(t)}{dt} &= - \sum_{(i,j)} \frac{dy_{ij}(t)}{dx_{ij}(t)} \frac{dx_{ij}(t)}{dt} \\
 &= \sum_{|x_{ij}| < 1} \frac{dx_{ij}(t)}{dt} \left[\sum_{C(k,l) \in N_r(i,j)} A(i,j;k,l) y_{kl}(t) - \frac{1}{R_x} y_{ij}(t) + \sum_{C(k,l) \in N_r(i,j)} B(i,j;k,l) u_{kl}(t) + I \right] \\
 &= \sum_{|x_{ij}| < 1} \frac{dx_{ij}(t)}{dt} \left[\sum_{C(k,l) \in N_r(i,j)} A(i,j;k,l) y_{kl}(t) - \frac{1}{R_x} y_{ij}(t) + \sum_{C(k,l) \in N_r(i,j)} B(i,j;k,l) u_{kl}(t) + I \right]
 \end{aligned} \tag{3.23}$$

(3.7.a.) durum denkleminde $C > 0$ alınarak (3.23) denkleminde yazılırsa;

$$\frac{dE(t)}{dt} = - \sum_{|x_{ij}| < 1} C \left[\frac{dx_{ij}(t)}{dt} \right]^2 \leq 0 \tag{3.24}$$

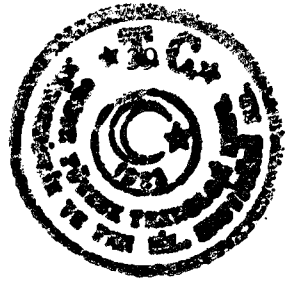
elde edilir. O halde bu denklem göstermektedir ki enerji monoton olarak azalan bir fonksiyondur. Buradan şu sonuca ulaşılır.

Teorem IV: Simetrik şablonlar kullanıldığında ve sonlu girişler uygulandığında HYSA her zaman kararlıdır.

HYSA çıkışlarının dijital olması istendiğinde geri besleme şablonu olan A şablonunun merkez elemanı $A_{22} \geq 1$ olmalıdır [Chua and Yang, 1988].

HYSA'nın dinamik davranışı iki hücreli bir ağ yapısı için durum uzayı 9 bölgeye ayrılarak Civaleri tarafından analiz edilmiştir [Civaleri and Gilli, 1993, 1994].

Güzeliş C., HYSA'yı genelleştirerek, genelleştirilmiş HYSA modeli üzerinde kararlılık analizi yapmıştır [Guzelis and Chua, 1993].



3.3.Hücreesel Yapay Sinir Ağının Eğitimi

Dinamik bir ağ olan HYSA için çeşitli öğrenme algoritmaları geliştirilmiştir. Schuler'in geliştirdiği algoritma sürekli zaman HYSA'larının istenen yörüngelerini öğrenmeye yöneliktir. [Schuler, et al, 1992]. Bir başka çalışmada ise ayrık zamanlı CNN'in kararlı-durum çıkışlarını öğrenmek amacıyla bir öğrenme algoritması geliştirilmiştir [Magnussen, et al,1992]. İki algoritma da tam kararlılık ve iki kutupluluk için gerekli olan kısıtlamaları göz ardı etmektedir. Bu kısıtlamalar da göz önüne alındığında HYSA'nın kararlı durum çıkışlarını öğrenmeye yönelik bir öğrenme algoritması aslında kısıtlamalı bir enazlama algoritmasıdır. Karamahmut görüntü işleme uygulamalarında Dinamik Geriye Yayılım Öğrenme Algoritması ("Recurrent Backpropagation Learning Algorithm:RBLA") ve Dinamik Algılayıcı Öğrenme Algoritmasını ("Recurrent Perceptron Learning Algorithm:RPLA") kullanmıştır [Karamahmut, 1994].

Literatür incelendiğinde HYSA eğitiminde en uygun yöntemin genetik algoritmalar olduğu görülmektedir [T. Kozek et al, 1988], [Sziranyi and Csapodi, 1994]. Bu nedenle tez çalışmasında hücreesel yapay sinir ağının eğitimi genetik algoritmalar ile yapılmıştır. Eğitim işleminde A,B, ve I şablonlarına ait parametrelere ikili kodlanmış genler karşılık düşürülerek, en iyi performansı gerçekleştiren kromozom bulunur. Genetik algoritmalar son yıllarda optimizasyon problemlerinin çözümünde başarıyla kullanılan modern yöntemlerden biridir. Hücreesel yapay sinir ağının eğitimi şablon parametrelerinin optimizasyonuna dayandığından genetik algoritmalar eğitim için tercih edilmektedir.

Çok seviyeli HYSA kaskat bağlı HYSA katmanlarından oluşur. Bir katmanın çıkışı, takip eden katmanın girişine bağlanır. Tez çalışmasında çok seviyeli HYSA şablonları bütün seviyeler için eşit seçilmiştir. Birinci seviye çıkışı kararlı bir konuma oturduktan sonra ikinci seviye girişi uyarılmaktadır ve HYSA seviyeleri arasında geri besleme uygulanmaktadır.

4. TRAPEZOİDAL AKTİVASYON FONKSİYONLU HÜCRESEL YAPAY SİNİR AĞI

Bu bölümde standart hücre sel yapay sinir ağı nda (HYSA) kullanılan parça parça lineer aktivasyon fonksiyonu yerine trapezoidal aktivasyon fonksiyonu kullanılacaktır. Bu fonksiyonun kullanılması durumunda HYSA'nın kararlılık kriterleri araştırılacaktır. HYSA, Chua ve Yang tarafından aşağıdaki gibi modellenmiştir [Chua and Yang, 1988].

$$\frac{dX}{dt} = -X + A * Y + B * U + I$$

$$\begin{aligned} \frac{dx_{ij}(t)}{dt} = & -x_{ij}(t) + \sum_{k=1}^{2r+1} \sum_{l=1}^{2r+1} A_{kl} \cdot y_{i+k-r-1, j+l-r-1}(t) \\ & + \sum_{k=1}^{2r+1} \sum_{l=1}^{2r+1} B_{kl} \cdot u_{i+k-r-1, j+l-r-1}(t) + I \end{aligned} \quad (4.1)$$

Burada u_{ij} , x_{ij} ve y_{ij} sırasıyla $C(i, j)$ hücresinin giriş, durum ve çıkışını göstermektedir. A geri besleme yönündeki geri besleme şablonu, B ileri yön şablonu olup boyutları ağda seçilen komşuluğa bağlıdır. Komşuluk r ise, A ve B şablonları $(2.r+1) \times (2.r+1)$ boyutundadır. I ise bütün hücreler için sabit bir değerde olan eşik değeridir. Bu ifadede kullanılan herhangi bir $C(i, j)$ hücresinin r komşuluğundaki hücrelerin kümesi ise aşağıda ifade edilmiştir.

$$N_r(i, j) = \{C(k, l) \mid \max\{|k-i|, |l-j|\} \leq r\}, \quad 1 \leq k \leq m, \quad 1 \leq l \leq n \quad (4.2)$$

Yukarıdaki komşuluk şartı altında, aşağıdaki şartların birisi olduğunda; $y_{i+k-r-1, j+l-r-1} = 0$ ve $u_{i+k-r-1, j+l-r-1} = 0$ kabul edilecektir.

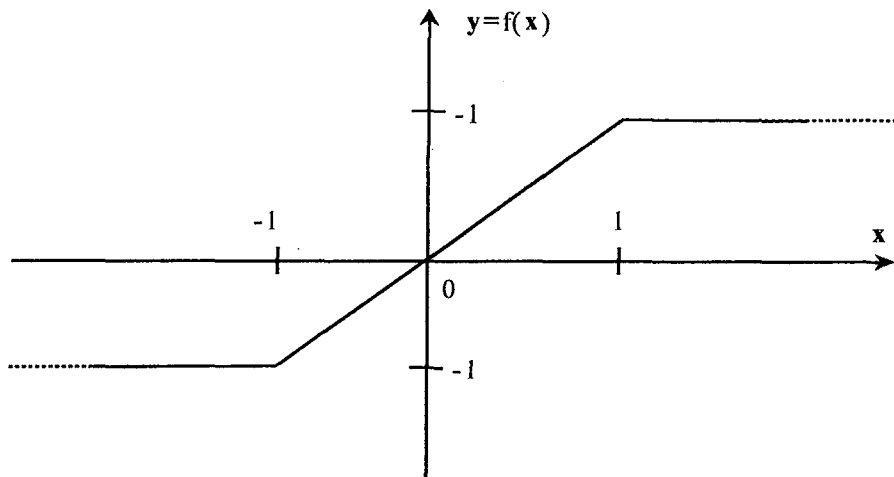
- (i) $i+k-r-1 > M$, (ii) $i+k-r-1 < 1$, (iii) $j+l-r-1 > N$, (iv) $j+l-r-1 < 1$.

Eşitlik (4.1)'deki, geri besleme yolundaki A şablonunu $r = s$ komşuluğundaki bir ağ için en genel halde (4.3) ifadesinde olduğu gibi yazabiliriz.

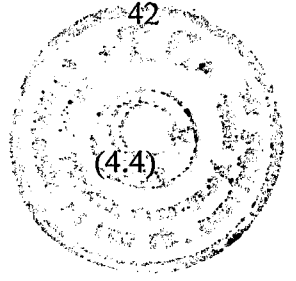
$$A = \begin{bmatrix} A_{1,1} & \cdots & A_{1,s} & A_{1,s+1} & A_{1,s+2} & \cdots & A_{1,2s-1} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{s,1} & \cdots & A_{s,s} & A_{s,s+1} & A_{s,s+2} & \cdots & A_{s,2s-1} \\ A_{s+1,1} & \cdots & A_{s+1,s} & A_{s+1,s+1} & A_{s+1,s+2} & \cdots & A_{s+1,2s-1} \\ A_{s+2,1} & \cdots & A_{s+2,s} & A_{s+2,s+1} & A_{s+2,s+2} & \cdots & A_{s+2,2s-1} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{2s-1,1} & \cdots & A_{2s-1,s} & A_{2s-1,s+1} & A_{2s-1,s+2} & \cdots & A_{2s-1,2s-1} \end{bmatrix} \quad (4.3)$$

$r = 1$ durumunda 3×3 lük bir A şablonu için şablonun merkezindeki eleman $A_{s+1,s+1}$, $A_{2,2}$ elemanına karşılık gelir, aynı şekilde $r = 2$ durumunda 5×5 lik bir A şablonu için $A_{s+1,s+1}$, $A_{3,3}$ 'e karşılık gelir. Yani bütün komşuluk değerleri için şablon merkezindeki eleman $A_{s+1,s+1}$ olur. Bu gösterilimde komşuluk büyüdükçe A şablonunun merkezindeki elemanların indisleri ' s ' cinsinden aynı kalır. Komşuluk değerinin 1 artmasıyla, A şablonuna 2 satır ve 2 sütun eklenmiş olur.

Hücrenin durumu ile çıkışı arasındaki ilişki aşağıdaki fonksiyonla tanımlanmaktadır.



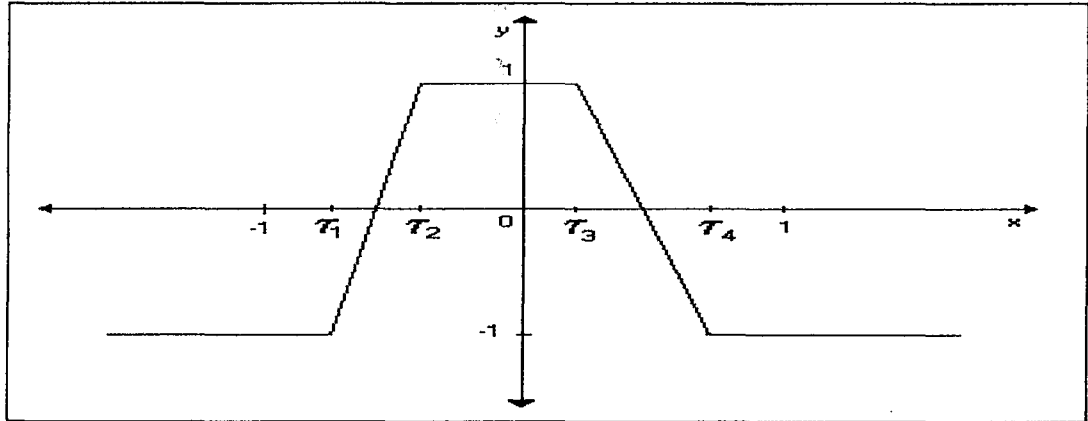
Şekil 4.1. Standart HYSA aktivasyon fonksiyonu



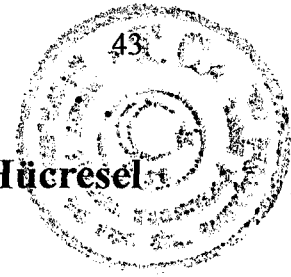
$$y_{ij}(t) = \frac{1}{2} \left(\left| x_{ij}(t) + 1 \right| - \left| x_{ij}(t) - 1 \right| \right) \quad (4.4)$$

Tezin bu bölümünde yukarıda gösterilen aktivasyon fonksiyonu yerine aşağıdaki fonksiyon kullanılacaktır (Şekil 4.2). Fonksiyondaki $\tau_1, \tau_2, \tau_3, \tau_4$ değerleri değiştirilerek uyarlanabilir bir aktivasyon fonksiyonu ile çalışmak mümkündür.

$$y = f(x) = \begin{cases} -1 & x < \tau_1 \\ \frac{2}{\tau_2 - \tau_1} x - \frac{\tau_1 + \tau_2}{\tau_2 - \tau_1} & \tau_1 \leq x < \tau_2 \\ 1 & \tau_2 \leq x < \tau_3 \\ \frac{2}{\tau_3 - \tau_4} x - \frac{\tau_3 + \tau_4}{\tau_3 - \tau_4} & \tau_3 \leq x < \tau_4 \\ -1 & x > \tau_4 \end{cases}, \quad -1 \leq \tau_1 < \tau_2 < \tau_3 < \tau_4 \leq 1 \quad (4.5)$$



Şekil 4.2. Trapezoidal aktivasyon fonksiyonu



4.1. Trapezoidal Aktivasyon Fonksiyonlu Hücresel Yapay Sinir Ağının Kararlılığı

Komşuluk değeri $1 \leq r \leq \min\{m, n\}$ için (4.1)'de verilen HYSA durum denklemini aşağıdaki gibi yazabiliriz.

$$\frac{dx_{ij}(t)}{dt} = -x_{ij}(t) + \sum_{k=1}^{2r+1} \sum_{l=1}^{2r+1} A_{kl} f(x_{i+k-r-1, j+l-r-1}(t)) + s_{ij} \quad (4.6)$$

Burada s_{ij} ,

$$s_{ij} = \sum_{k=1}^{2r+1} \sum_{l=1}^{2r+1} B_{kl} u_{i+k-r-1, j+l-r-1}(t) + I_{ij} \quad (4.7)$$

olup, HYSA'da giriş, u , zamanla değişmediğinden ve B_{kl} ve $I_{ij} = I$ zamanla değişmeyen sabit değerler olduğundan, s_{ij} değerleri ağdaki her bir hücre için farklı sabit değerlere eşit olup zamanla değişmemektedir.

Eşitlik (4.6) daki HYSA denkleminde kullanılan $f(\bullet)$ fonksiyonu Eşitlik (4.5)'de tanımlanan trapezoidal aktivasyon fonksiyonudur. Bir HYSA'da $M \times N$ adet hücre olduğu için, bu hücrelere ait $M \times N$ adet durum değişkeni vardır. HYSA'nın davranışının analizini basitleştirmek için, HYSA içerisindeki hücrelerin matrisyel notasyonla değilde referansda [Chua and Roska, 2002] verildiği gibi vektörel notasyonla gösterilmesi için, iki boyutlu bir düzlem içerisinde yerleştirilmiş olan C_{ij} hücreleri, (4.8) ifadesinde olduğu gibi uygun bir dönüşüm ile bir boyutlu C_p formunda bir dizinin elemanları olarak gösterilmelidir.

$$\begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{21} & \cdots & C_{2n} \\ \vdots & \vdots & C_{ij} & \vdots \\ C_{m1} & C_{m2} & \cdots & C_{mn} \end{bmatrix} \xrightarrow{F} [C_{11} \ C_{12} \ \cdots \ C_{1n} \ \cdots \ C_{ij} \ \cdots \ C_{m1} \ C_{m2} \ \cdots \ C_{mn}]' = C \quad (4.8)$$



Bu dönüşüm ile $M \times N$ boyutlu HYSA içerisindeki hücreler C_{ij} gibi satır ve sütun numaraları ile değil C_p şeklinde gösterilecektir. Burada p indisi ile i, j indisleri arasında (4.9) bağıntısı bulunmaktadır.

$$p = j + (i-1).N \quad (4.9)$$

bağıntısı vardır. Bunun ters dönüşümü de aşağıdaki bağıntılarda verilmiştir.

$$i(p) = \left\lfloor \frac{p}{N} \right\rfloor, \quad j(p) = p \bmod(N) \quad (4.10)$$

(4.6) eşitliği, (4.9) ve (4.10) bağıntıları kullanılarak (4.11) de gösterildiği gibi yeniden yazılabilir.

$$\frac{d}{dt} x_p(t) = -x_p(t) + \sum_{q=1}^{MN} w_{pq} f(x_q(t)) + s_p, \quad p = 1, 2, \dots, MN \quad (4.11)$$

(4.11) ve (4.6) karşılaştırılarak w_{pq} elemanları ile A_{kl} şablon elemanları arasında bir ilişki kurulabilir. Bunun için aşağıdaki (4.12) gibi v ve z indisleri tanımlanarak elemanları genelde sıfır olan (sparse matrix) $MN \times MN$ boyutlarında bir W matrisi oluşturulur (4.13).

$$v = i + k - r - 1, \quad z = j + l - r - 1 \quad (4.12.a)$$

$$k = v - i + r + 1, \quad l = z - j + r + 1 \quad (4.12.b)$$

$$p = j + (i-1)N, \quad q = z + (v-1)N \quad (4.12.c)$$

$$w_{pq} = \begin{cases} A_{i(q)-i(p)+r+1, j(q)-j(p)+r+1}, & 1 < i(q)-i(p)+r+1 < 2r+1 \\ & 1 < j(q)-j(p)+r+1 < 2r+1 \\ 0, & \text{diğer} \end{cases} \quad (4.13)$$



Bu matris incelendiğinde (4.14)'deki gibi bir simetriklik yapısı içerdiği görülmektedir.

$$w_{pq} = w_{MN+1-q, MN+1-p} \quad (4.14)$$

(4.12) ve (4.13) ifadeleri dikkate alınarak ve $p = j + (i-1)N$ için $x_{ij}(t) = x_p(t)$ ve $s_{ij} = s_p$ yazılarak;

$$-x_{ij}(t) + \sum_{k=1}^{2r+1} \sum_{l=1}^{2r+1} A_{kl} f(x_{i+k-r-1, j+l-r-1}(t)) = -x_p(t) + \sum_{q=1}^{MN} w_{pq} f(x_q(t)) \quad (4.15)$$

ifadesi elde edilir.

Eşitlik (4.11) ifadesini tüm ağ için genelleştirilirse,

$$\frac{dX}{dt} = -X + W.F(X) + S \quad (4.16)$$

ifadesi elde edilir. Burada

$$X = [x_1(t), x_2(t), \dots, x_p(t)]^T, S = [s_1, s_2, \dots, s_p]^T, (p = MN) \text{ ve } W, (MN \times MN)$$

boyutunda ağırlık matrisidir.

Bir HYSA'nın çıkışlarının kararlı olabilmesi için HYSA bir denge noktasına ulaşmalıdır. Denge noktasında, her bir hücre için $\frac{d}{dt}x_p(t) = \dot{x}_p(t) = \dot{x}_{pe} = 0$ olacağından (4.11) denkleminde,

$$\dot{x}_{pe}(t) = -x_{pe}(t) + \sum_{q=1}^{MN} w_{pq} f(x_{qe}(t)) + s_p = 0 \quad (4.17)$$

ve



$$x_{pe}(t) = \sum_{q=1}^{MN} w_{pq} f(x_{qe}(t)) + s_p, \quad p = 1, 2, \dots, MN$$

elde edilir. Böylece denge noktasının bulunabilmesi için (4.19) doğrusal denklem takımının çözümünün bulunması gerektiği görülür.

$$X = W.F(X) + S, \text{ öyle ki } X = [x_1, x_2, \dots, x_p]', \quad p = MN \quad (4.19)$$

Aşağıdaki gibi bir $G(\bullet)$ fonksiyonu tanımlanırsa,

$$G(X) = W.F(X) + S \quad (4.20)$$

Aşağıdaki şartlar sağlanırsa $G(X)$ fonksiyonu sürekli olacaktır;

- (i) Aktivasyon fonksiyonu f sürekli ve sonlu değerler alan bir fonksiyondur.
- (ii) Bütün A, B, I şabon elemanları sonlu değerler almaktadır.
- (iii) Bütün girişler ve durum değişkenlerinin başlangıç değerleri sonlu değerler almaktadır.

$$\forall |x_{ij}(0)| \leq x_{\max} < \infty, \quad \forall |u_{ij}(t)| \leq u_{\max} < \infty$$

(4.20) ifadesi, (4.19) denkleminde sağ taraf için kullanılırsa,

$$X = G(X) \quad (4.21)$$

ifadesine ulaşılır. Bu durumda $X_e, G: \mathcal{R}^{m^n} \rightarrow \mathcal{R}^{m^n}$ fonksiyonunu sabit noktasıdır (fixed point). Buradan şu hükmü vermemiz için hiç bir engel yoktur.

Eğer G fonksiyonunun en az bir sabit noktası (fixed point) bulunabiliyorsa (4.19) denkleminin bir denge noktası vardır ve bu çözüm X_e dir. G fonksiyonunun

değişim aralığının sınırlarını bulmak için, (4.20) ifadesi HYSA'da bir p hücresi için yazılır ve mutlak değeri alınır (4.22) ifadesi elde edilir.

$$|(G(x))_p| = \left| \sum_{q=1}^{MN} w_{pq} f(x_q) + s_p \right| \quad (4.22)$$

Üçgen eşitsizliğinden ve aktivasyon fonksiyonun maksimum değerinin 1 olduğunu dikkate alarak;

$$|(G(x))_p| \leq \sum_{q=1}^{MN} |w_{pq} f(x_q)| + |s_p| \leq \sum_{q=1}^{MN} |w_{pq}| |f(x_q)| + |s_p| \leq \sum_{q=1}^{MN} |w_{pq}| + |s_p| \quad (4.23)$$

eşitsizliği elde edilir. Burada $G(x)$ in alabileceği maksimum değer,

$$\Psi = \max_{1 \leq p \leq m \cdot n} \left(\sum_{q=1}^{MN} |w_{pq}| + |s_p| \right) \quad (4.24)$$

olacaktır. Denklem (4.21) ve (4.24) den

$$G(\mathcal{R}^{m \cdot n}) \subset \mathcal{Q} = \{[x_1, x_2, \dots, x_{m \cdot n}] \in \mathcal{R}^{m \cdot n} \mid |x_p| \leq \Psi\}, \quad p = 1, 2, \dots, m \cdot n \quad (4.25)$$

koşulu elde edilmiş olur. Buradan görülmektedir ki $G(\bullet)$ fonksiyonu sınırlı ve sürekli.

4.1.1. Brouwer Fixed Point Teoremi

D üzerinde bir f fonksiyonu aşağıdaki gibi tanımlandığı kapalı bir bölgede sürekli olsun.

$$f = D^n \rightarrow D^n, \quad D^n = \{x \in \mathcal{R}^n, |x| \leq k < \infty\} \quad (4.26)$$

Fonksiyon sınırlı bir bölgede tanımlı ve sürekli ise f fonksiyonu, $f(a) = a$ şartını sağlayan $a \in D^n$ sabit noktaya (fixed point) sahiptir [Vidyasagar, 1978], [Larsen, 1973], [Hille, 1972].

Görüldüğü gibi bu teoreme göre bir fonksiyonun sabit noktaya (fixed point) sahip olabilmesi için, kapalı bir bölgede tanımlı olmalı, sonlu değerler almalı ve bu bölgede fonksiyon sürekli olmalıdır. Eşitlik (4.20) da tanımlanan G fonksiyonu, (4.25) eşitliğinde gösterildiği gibi sınırlı ve sürekli olduğundan, Brouwer fixed point teoremine göre, bir sabit noktaya (fixed point) sahiptir. Bu da (4.19) denkleminin çözümünün kesinlikle bir denge noktasına sahip olacağını göstermektedir. Şimdi, varlığı belirlenen bu denge noktasının asimptotik anlamda kararlılık şartlarını incelemek için aşağıdaki (4.1.2) ve (4.1.3) teoremlerine ihtiyaç duyulmaktadır [Chua and Roska, 2002].

4.1.2. Global varlık ve teklik teoremi

(4.1) denklemi ile tanımlanan standart HYSA aşağıda verilen A1-A3 şartlarını sağlıyorsa, herhangi bir $x_{ij}(0) \in R$ başlangıç değeri ve zamanla değişmeyen sabit giriş değeri için, HYSA'nın tek bir çözümü vardır.

A1: Sinaptik operatörler lineer ve belleksizdir: $A_{kl} y_{i+k-r-1, j+l-r-1}$ ve $B_{kl} u_{i+k-r-1, j+l-r-1}$ terimleri skaler çarpımlardır ve $A(k, l)$ ve $B(k, l)$ şablon elemanları reel sayıdır.

A2: $u_{ij}(t)$ ve $I_{ij}(t)$ zamana göre süreklidir. (Her iki değerde bütün hücreler için sabit olduğundan bu şart sağlanmaktadır.)

A3: Nonlineerlik (aktivasyon fonksiyonu) $f(x)$ Lipschitz sürekliliğine sahiptir. Bütün x', x'' değerleri için;

$$\|f(x') - f(x'')\| \leq L \|x' - x''\| \quad (4.27)$$

şartını sağlayan sabit ve reel bir L değeri vardır.

4.1.3. Çözümün Sınırlı Olması (Sonlu Değer İçermesi) Teoremi

Eğer, $|x_{ij}(0)| \leq 1$, $|u_{ij}(t)| \leq 1$, $|I_{ij}(t)| \leq I_{\max}$ ise $A(k, l)$ ve $B(k, l)$ gibi lineer ve belleksiz sinaptik operatörleri olan HYSA düzgün sınırlıdır. Her bir hücrenin $x_{ij}(t)$ durum değişkeni için $|x_{ij}(t)| \leq x_{\max} \quad \forall t \geq 0, \quad 1 \leq i \leq M, \quad 1 \leq j \leq N$ şartını sağlayan sabit bir $x_{\max}$ değeri vardır ve bu değer (4.28) eşitliğinde gösterildiği gibi şablon elemanlarının mutlak değerlerinin toplamının 1 fazlasına eşittir.

$$x_{\max} = 1 + I_{\max} + \sum_{k=1}^{2r+1} \sum_{l=1}^{2r+1} (|A(k, l)| + |B(k, l)|) \quad (4.28)$$

Bu teoremler herhangi bir değişikliğe ihtiyaç duyulmadan trapezoid aktivasyon fonksiyonlu (TAF) hücresel yapay sinir ağına uygulanabilir. Çünkü, TAF standart PWL aktivasyon fonksiyonu gibi *Lipschitz* sürekliliğine haizdir ve aynı aralıkta değerler almaktadır; $|f_{TAF}(x)| \leq 1$, $|f_{SPWL}(x)| \leq 1$. Ayrıca, TAF fonksiyonlu HYSA (TAF-HYSA) teorem (4.1.2) de verilen A1-A3 şartlarını da sağlamaktadır.

4.1.4. Lyapunov Anlamında Kararlılık

(4.28) ifadesi, (4.6) denkleminin ve dolayısıyla (4.16) denkleminin bütün çözümlerinin sonlu olduğunu göstermektedir. Brouwer teoremine göre de (4.16) HYSA denkleminin en az bir denge noktasının olduğu görülmektedir. Bu bölümde bu noktanın kararlılığını incelemek için Lyapunov kararlılık kriteri uygulanacaktır.

Eğer $X_e = [x_{1e}, x_{2e}, \dots, x_{ie}]^T$, $i = MN$, bir denge noktası ise, bu denge noktası ile herhangi bir anda HYSA'nın sahip olduğu durum arasındaki sapma;

$$e_i = x_i(t) - x_{ie} \quad i = 1, 2, \dots, MN \quad (4.29)$$

olacaktır. Her bir hücre için bu sapmanın türevi;



$$\dot{e}_i(t) = \dot{x}_i(t)$$

olacaktır. Burada $\dot{x}_i(t)$ için (4.11) denklemindeki ifadenin sağ tarafı yazılırsa,

$$\dot{e}_i(t) = \dot{x}_i(t) = -x_i(t) + \sum_{q=1}^{MN} w_{iq} f(x_q(t)) + s_i \quad (4.31)$$

elde edilir. Bu ifade de $x_i(t)$ için (4.29) ifadesinden faydalanarak

$$x_i(t) = e_i + x_{ie} \quad (4.32)$$

yazılır ve (4.18) denklemindeki s_p ifadesi için

$$s_p = x_{pe}(t) - \sum_{q=1}^{MN} w_{pq} f(x_{qe}(t)) \quad p = 1, 2, \dots, MN \quad (4.33)$$

yazılırsa

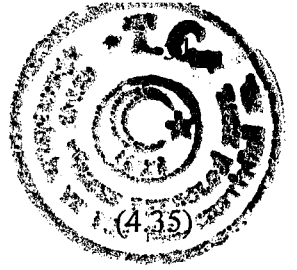
$$\dot{e}_i(t) = -e_i(t) + \sum_{q=1}^{MN} w_{iq} [f(e_q(t) + x_{qe}) - f(x_{qe})] \quad p = 1, 2, \dots, MN \quad (4.34)$$

elde edilir.

$e_e = [0, 0, \dots, 0]'$ noktası, (4.34) denkleminin bir denge noktasıdır. (4.16)

denkleminle tanımlı HYSA'nın global asimptotik kararlılığını ispatlamak için (4.34) denkleminin global asimptotik kararlılığını ispatlamak yeterli olacaktır [Zhou and Zhang, 2001].

Şimdi aşağıdaki ifadeyi Lyapunov fonksiyonu olarak seçelim.



$$P(x) = \frac{1}{2} \sum_{i=1}^{M.N} c_i (x_i - x_{ie})^2 \quad c_i > 0, \quad i = 1, 2, \dots, MN \quad (4.35)$$

Bu fonksiyon kesin pozitif bir fonksiyon olduğu için Lyapunov fonksiyonu olarak kullanılabilir. Şimdi bu fonksiyonu türetelim;

$$\frac{dP}{dt} = \sum_{i=1}^{MN} c_i e_i(t) \left\{ -e_i(t) + \sum_{q=1}^{MN} w_{iq} \left[f(e_q(t) + x_{qe}) - f(x_{qe}) \right] \right\} \quad (4.36)$$

Şimdi f fonksiyonu için (4.5) ifadesinde verilen Lipschitz şartını sağlayan fonksiyon kuralına göre

$$|f(x) - f(y)| \leq \mu |x - y| \quad \mu > 0 \quad (4.37)$$

şartını yazabiliriz. Buradaki μ için daha kesin olarak,

$$\mu \geq \max\{m_1, -m_2\} = \max\left\{ \frac{2}{\tau_2 - \tau_1}, \frac{2}{\tau_4 - \tau_3} \right\} \quad (4.38)$$

yazabiliriz. Aktivasyon fonksiyonu $f(x)$ in (4.37) eşitsizliğinde verilen özelliğe sahip olması, $G(x)$ fonksiyonunun Lipschitz şartına uyan sürekli bir fonksiyon olmasını sağlamaktadır.

Şimdi (4.37) ifadesinden faydalananarak, (4.36) ifadesi tekrar düzenlenirse,

$$\frac{dP}{dt} \leq \sum_{i=1}^{MN} \left[-c_i e_i^2(t) + c_i \sum_{q=1}^{MN} |w_{iq}| |e_i(t)| \mu |e_q(t)| \right] \quad (4.39)$$

elde edilir. Bu ifade düzenlenirse ;

$$\frac{dP}{dt} \leq \sum_{i=1}^{MN} \left[-c_i e_i^2(t) + \frac{1}{2} \sum_{q=1}^{MN} |w_{iq}| \left(e_i^2(t) + (c_i \mu e_q(t))^2 \right) \right] \quad (4.40)$$



yazılabilir. Eşitsizliğin sağ tarafındaki ikinci toplam ifadesi tekrar düzenlenirse;

$$\frac{dP}{dt} \leq \sum_{i=1}^{MN} \left[-c_i e_i^2(t) + \frac{1}{2} \sum_{q=1}^{MN} |w_{iq}| e_i^2(t) + \frac{1}{2} \sum_{q=1}^{MN} |w_{iq}| (c_i \mu e_q(t))^2 \right] \quad (4.41)$$

ifadesi elde edilir. Bu ifade açılırsa;

$$\frac{dP}{dt} \leq \sum_{i=1}^{MN} -c_i e_i^2(t) + \frac{1}{2} \sum_{i=1}^{MN} \sum_{q=1}^{MN} |w_{iq}| e_i^2(t) + \frac{1}{2} \sum_{i=1}^{MN} \sum_{q=1}^{MN} |w_{iq}| (c_i \mu e_q(t))^2 \quad (4.42)$$

elde edilir. En sağdaki toplam ifadesinde indisler değiştirilirse;

$$\frac{dP}{dt} \leq \sum_{i=1}^{MN} -c_i e_i^2(t) + \frac{1}{2} \sum_{i=1}^{MN} \sum_{q=1}^{MN} |w_{iq}| e_i^2(t) + \frac{1}{2} \sum_{i=1}^{MN} \sum_{q=1}^{MN} |w_{qi}| (c_q \mu e_i(t))^2 \quad (4.43)$$

ifadesi bulunur. Bu ifade $e_i^2(t)$ parantezi altında yazılırsa

$$\frac{dP}{dt} \leq - \sum_{i=1}^{MN} \left[c_i - \frac{1}{2} \sum_{q=1}^{MN} |w_{iq}| - \frac{1}{2} \sum_{q=1}^{MN} |w_{qi}| c_q \mu^2 \right] e_i^2(t) \quad (4.44)$$

elde edilir. Sistemin kararlı olabilmesi için (4.44) ifadesinin sağ tarafının sıfırdan küçük olması gerekir. $e_i^2(t) \geq 0$ olduğundan kararlılık için aşağıdaki (4.45) şartının sağlanması yeterlidir.

$$\left[c_i - \frac{1}{2} \sum_{q=1}^{MN} |w_{iq}| - \frac{1}{2} \sum_{q=1}^{MN} |w_{qi}| c_q \mu^2 \right] > 0 \quad (4.45)$$

Eşitlik (4.13) ve (4.14) denklemlerinden W matrisine ilişkin aşağıdaki özellikler çıkarılabilir.

$$\sum_{q=1}^{MN} w_{iq} = \sum_{q=1}^{MN} w_{qi} \leq \sum_{k=1}^{2r+1} \sum_{l=1}^{2r+1} A_{kl} \quad (4.46)$$

(4.45)'de bütün c_i sabitleri $c_i = c$, $c > 0$ $i = 1, 2, \dots, MN$ olacak şekilde sabit bir c değeri seçilerek, W matrisinin (4.46) ifadesinde verilen özelliği kullanılırsa

$$c - \left(\frac{1 + c^2 \mu^2}{2} \right) \sum_{q=1}^{MN} |w_{iq}| > 0 \quad (4.47)$$

ve

$$\sum_{q=1}^{MN} |w_{iq}| < \frac{2c}{c^2 \mu^2 + 1} \quad (4.48)$$

koşulu elde edilir. Sonuç olarak

$$\sum_{k=1}^{2r+1} \sum_{q=1}^{2r+1} |A_{kl}| < \frac{2}{c^2 \mu^2 + 1} c \quad (4.49)$$

kararlılık kriterine ulaşılır. Bu ifade, HYSA'nın kararlılığının, A şablonunun elemanlarına ve aktivasyon fonskiyonunun maksimum mutlak eğimi olan μ parametresine bağlı olduğunu göstermektedir. Kararlılık bölgesini genişletmek için (4.49) ifadesinin sağ tarafı mümkün olduğunca büyük tutulmalıdır. Bu bize A şablonu elemanlarını daha geniş bir aralıkta seçebilme imkanı verecektir. İfadenin sağ tarafı analiz edilirse c sabiti $1/\mu$ değerini aldığı anda, ifade maksimum değerini almaktadır. Eşitsizlik $c = 1/\mu$ şartı altında tekrar düzenlenirse

$$\sum_{k=1}^{2r+1} \sum_{q=1}^{2r+1} |A_{kl}| < \frac{1}{\mu} \quad (4.50)$$

kararlılık kriteri bulunur. Bu kriter TAF-HYSA sisteminin kararlılığı için yeter şarttır. Sonuç olarak TAF-HYSA için kararlılık kriteri aşağıdaki teoremle özetlenebilir.

Teorem: Eğer, sabit girişler için (4.1) denklemi ile tanımlanan TAF-HYSA sisteminin şablon elemanları (4.50) şartını sağlıyorsa, sistemin durum değişkenlerinin yörüngesi asimptotik olarak tek bir denge noktasına yakınsar.

TAF-HYSA için (4.50) ifadesinde elde edilen kararlılık kriteri standart PWL aktivasyon fonksiyonu kullanılan PWL-HYSA için de uygulanabilir. Standart PWL aktivasyon fonksiyonunda değeri $\mu = 1$ olan tek bir eğim sözkonusu olduğundan, PWL-HYSA için kararlılık kriteri

$$\sum_{k=1}^{2r+1} \sum_{q=1}^{2r+1} |A_{kl}| < 1 \quad (4.51)$$

şeklinde yazılabilir. Bu sonuç [Liao, et al ,2002] , [Zhang J, 2002] ve [Zhou and Zhang, 2001] referanslarında elde edilen kararlılık kriterleri ile karşılaştırıldığında daha pratiktir.

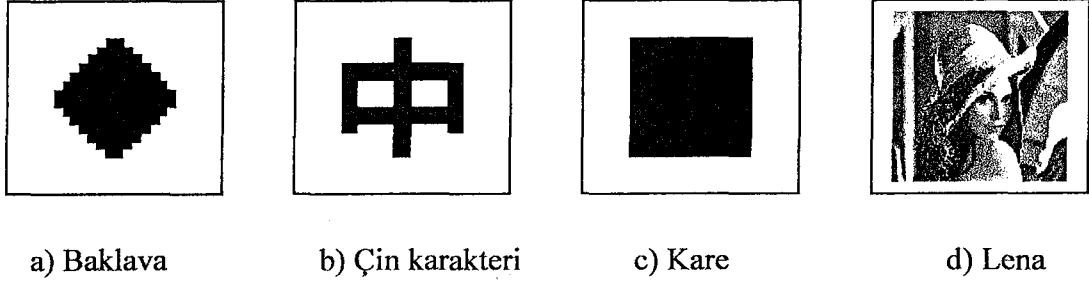
4.2. TAF-HYSA Uygulamaları

Bu bölümde, ilk önce hücrel yapay sinir ağlarının kararlılık şartını sağlayan ve sağlamayan farklı şablonlar kullanılarak, bir önceki bölümde verilen kararlılık kriterlerinin doğruluğu pratik örneklerle gösterilerek, kararlılık analizleri yapılmaktadır. İkinci bölümde ise tek katmanlı standart hücrel yapay sinir ağları ile çözülemeyen XOR gibi lineer olarak ayrıştırılamayan problemlerin trapezoidal aktivasyonu kullanıldığında çözülebildiği gösterilmektedir.

4.2.1. Kararlılık analizleri

Bu çalışmada sıklıkla kullanılan standart PWL aktivasyon fonksiyonu yerine TAF adı verilen trapezoid aktivasyon fonksiyonu kullanıldı. Bunun nedeni; TAF'lu HYSA'nın EXOR ve parite problemleri gibi lineer olarak ayrıştırılamayan işlerde kullanılmasının amaçlanmasıdır . Bir önceki bölümde, şablon parametrelerinin (4.50) eşitsizliğini sağlaması durumunda HYSA-TAF'ın eksponansiyel olarak kararlı olduğu gösterildi.

Çalışmada, HYS A'nın simülasyonu için Pentium III-450 Mhz lik bilgisayarda MATLAB 6.5 kullanıldı. Sonuçların karşılaştırılması açısından başlangıçta standart PWL aktivasyon fonksiyonlu HYS A, dört farklı giriş görüntüsü için çalıştırıldı. (Şekil 4.3). Kullanılan giriş görüntüleri literature uygunluk açısından, üç tanesi Chua ve Yang'ın makalesinde kullandıkları görüntülerden seçilirken diğer görüntü, görüntü işleme alanında sıklıkla kullanılan standart Lena resmi olarak seçildi.



Şekil 4.3. Giriş Görüntüleri

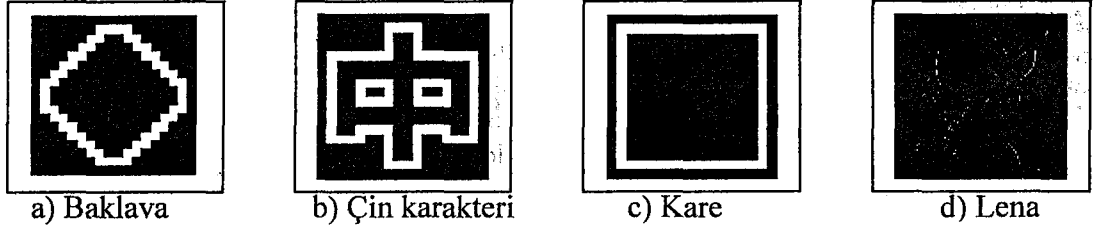
A , B ve I şablonları Chua ve Yang'ın kenar tanıma için kullandıkları maskelerden seçildi [Chua and Yang, 1988]. Söz konusu şablonlar aşağıdaki gibidir;

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} -0.25 & -0.25 & -0.25 \\ -0.25 & 2 & -0.25 \\ -0.25 & -0.25 & -0.25 \end{bmatrix} \quad I = -1.5 \quad (4.52)$$

Burada kullanılan şablonlar simetriktir ve dolayısıyla bu şablonlarla standart PWL-HYS A kararlı olacaktır. HYS A'nın kararlı hale geçiş davranışını izlemek için her bir hücrenin durum değişkeninin türevinin mutlak değeri gözlemlendi ve bütün ağ için ortalama mutlak türev (AADV) değeri hesaplandı.

$$AADV(t) = \frac{1}{MN} \sum_{p=1}^{MN} \left| \frac{dx_p}{dt} \right| \quad (4.53)$$

Açıktır ki kararlı HYS A için (AADV) değeri transient sonunda sıfır olacaktır. Çünkü, kararlı halde bütün hücrelerin durum değişkenlerinin türevleri sıfırdır. Keza eğer HYS A ekponansiyel olarak kararlı ise AADV değeri de zamanla ekponensiyel olarak azalacaktır. Simülasyon sonuçları Şekil 4.4'de görülmektedir.



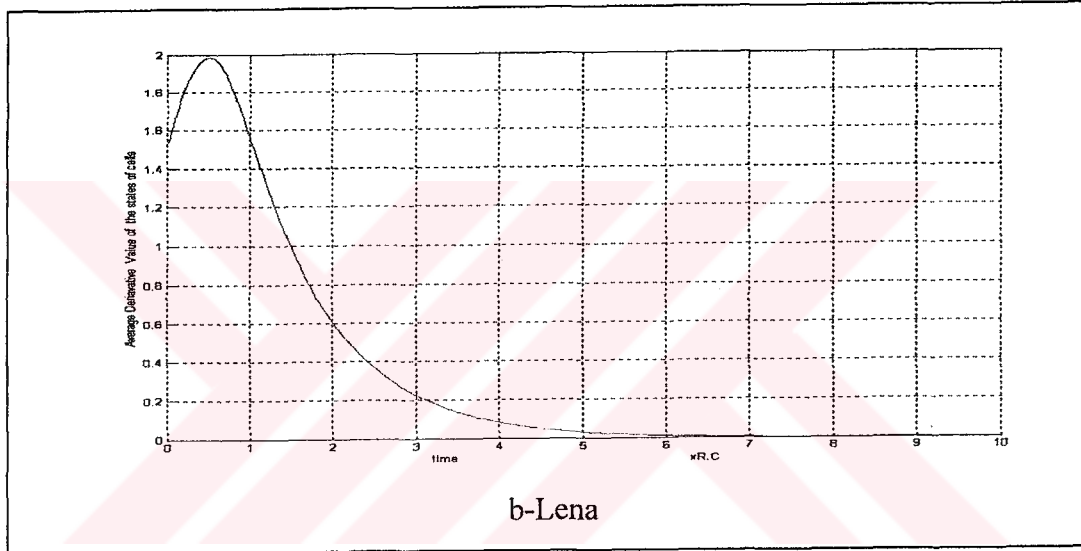
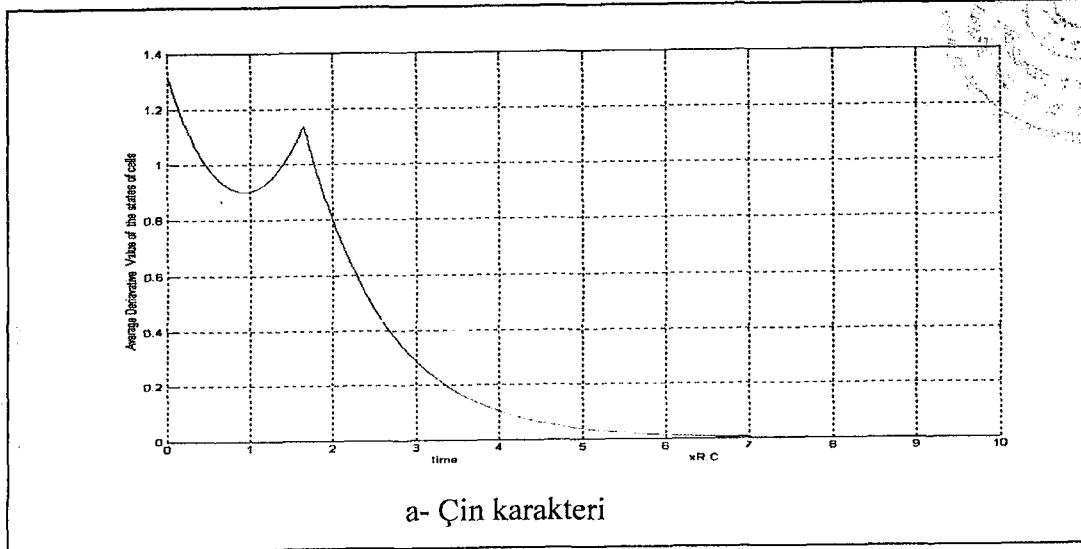
Şekil 4.4. Kenar tanıma şablonları için PWL-HYSA çıkışları

Burada görülmektedir ki kenar tanıma maskeleri ikili giriş görüntüleri için çok başarılıdır fakat giriş görüntüsü Lena görüntüsü gibi ikili değil, gri seviyeli ise PWL-HYSA bu şablonlarla bütün kenarları tanıyamamaktadır. Baklava dilimi, Çin karakteri ve kare için AADV performansının benzer olduğu gözlemlendiğinden, yalnızca Çin karakterine ait grafiklerle Lena görüntüsüne ait grafikler verilecektir. Şekil 4.5.'te Çin karakteri ile Lena görüntüsüne ait AADV grafikleri verilmektedir.

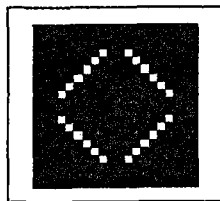
Bu ve bundan sonraki benzer grafiklerde zaman eksenini $R \times C$ zaman sabitine göre ölçeklendirilmiştir. Burada R ve C , HYSA eşdeğer devresindeki direnç ve kondansatörü temsil etmektedir. Beklendiği gibi zaman $5 \times R \times C$ geçtikten sonra HYSA kalıcı duruma ulaşır.

Eşitlik (4.52) de ki şablonlar simetrik olduğundan HYSA kararlı olur. Şimdi (4.51) şartını sağlamak için, A şablonu 0.475 ile çarpılarak, A şablonunun elemanlarının mutlak değerlerinin toplamı 1 den küçük olması sağlandı. Yeni A şablonu;

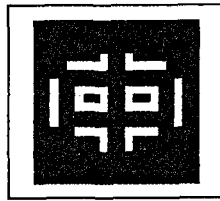
$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.95 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (4.54)$$



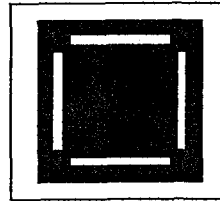
Şekil 4.5. Kenar tanıma maskeleri (4.52) kullanılan PWL-HYSA için AADV eğrileri. Giriş görüntüsü ; a-) Çin karakteri, b-) Lena.



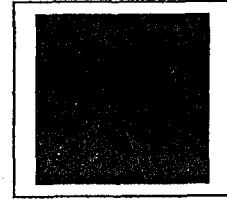
a) Baklava



b) Çin karakteri



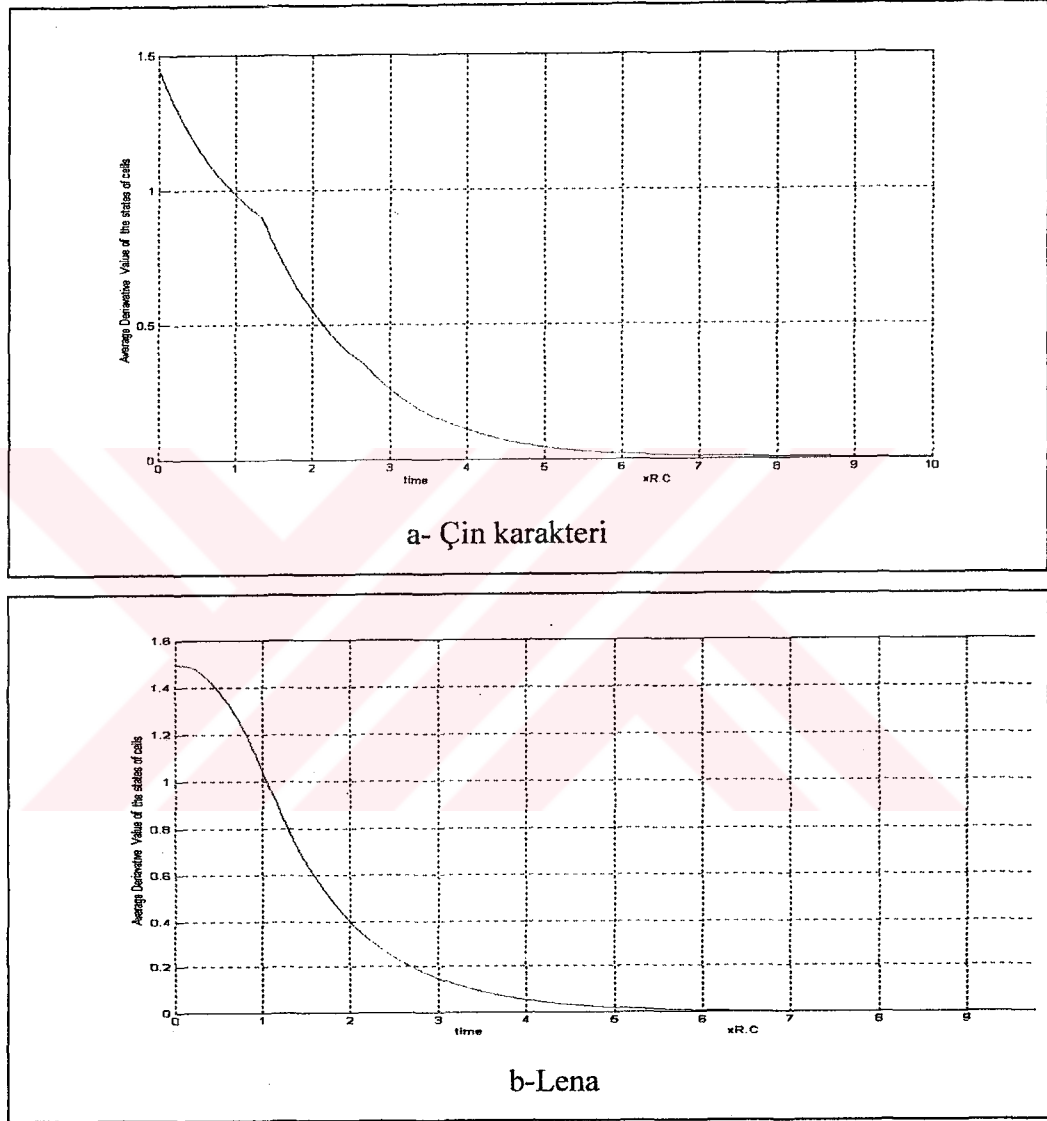
c) Kare



d) Lena

Şekil 4.6. (4.54) ifadesindeki A şablonu için PWL-HYSA çıkışları

Burada, HYSA'nın yalnızca A şablonu değiştirilirken B ve I şablonları önceki örnekteki gibi seçildi. A şablonunun merkez elemanı 1 den küçük olduğundan Şekil.4.6.'da görüldüğü gibi çıkış ikili değildir. Ama HYSA Şekil 4.7.'de görüldüğü gibi kararlıdır. Çin karakterine ve Lena görüntüsüne ait AADV değerleri birbirine çok yakın gözlemlendi.



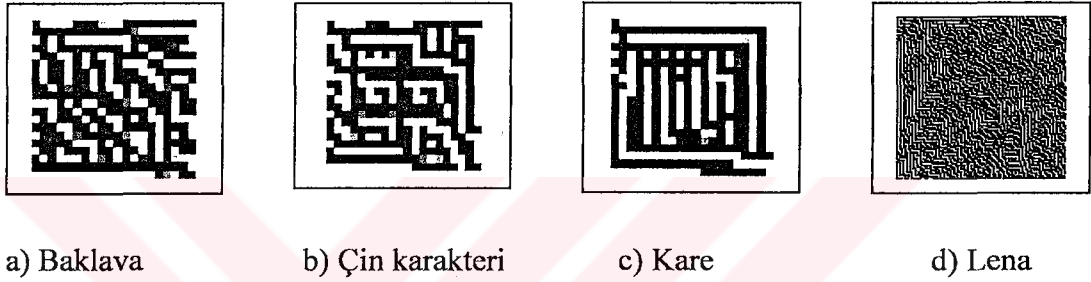
Şekil. 4.7. Tanımı (4.54) ifadesinde verilen A şablonu kullanılan PWL-HYSA için AADV eğrileri. Giriş görüntüsü ; a-) Çin karakteri, b-) Lena.

A değiştiğinden artık yeni şablonlar kenar tanıma şablonları değildir. Bu uygulamanın asıl amacı HYSA'nın kararlılığının araştırılmasıdır. Bu nedenle, Şekil 4.5 ve Şekil 4.7 karşılaştırılırsa Şekil 4.7'de AADV monoton azalırken, Şekil 5'te bazı bölgelerde yükseldiği görülür. Her iki şekilde de HYSA'lar kararlıdır.

Bir sonraki uygulamada kararsız HYSA' nın AADV değerleri gözlenmektedir. Bu örnekte de B ve I aynı kalırken, A şablonu non-simetrik bir yapıya dönüştürülmüştür. Yeni A şablonu aşağıdaki gibi düzenlendi

$$A = \begin{bmatrix} -0.25 & 0.50 & -1.00 \\ 1.50 & 2.00 & -1.75 \\ -1.25 & -0.75 & 0.25 \end{bmatrix} \quad (4.55)$$

Bu yeni şablona ait çıkış görüntüleri ve AADV eğrileri Şekil 4.8 ve Şekil 4.9'da görülmektedir.

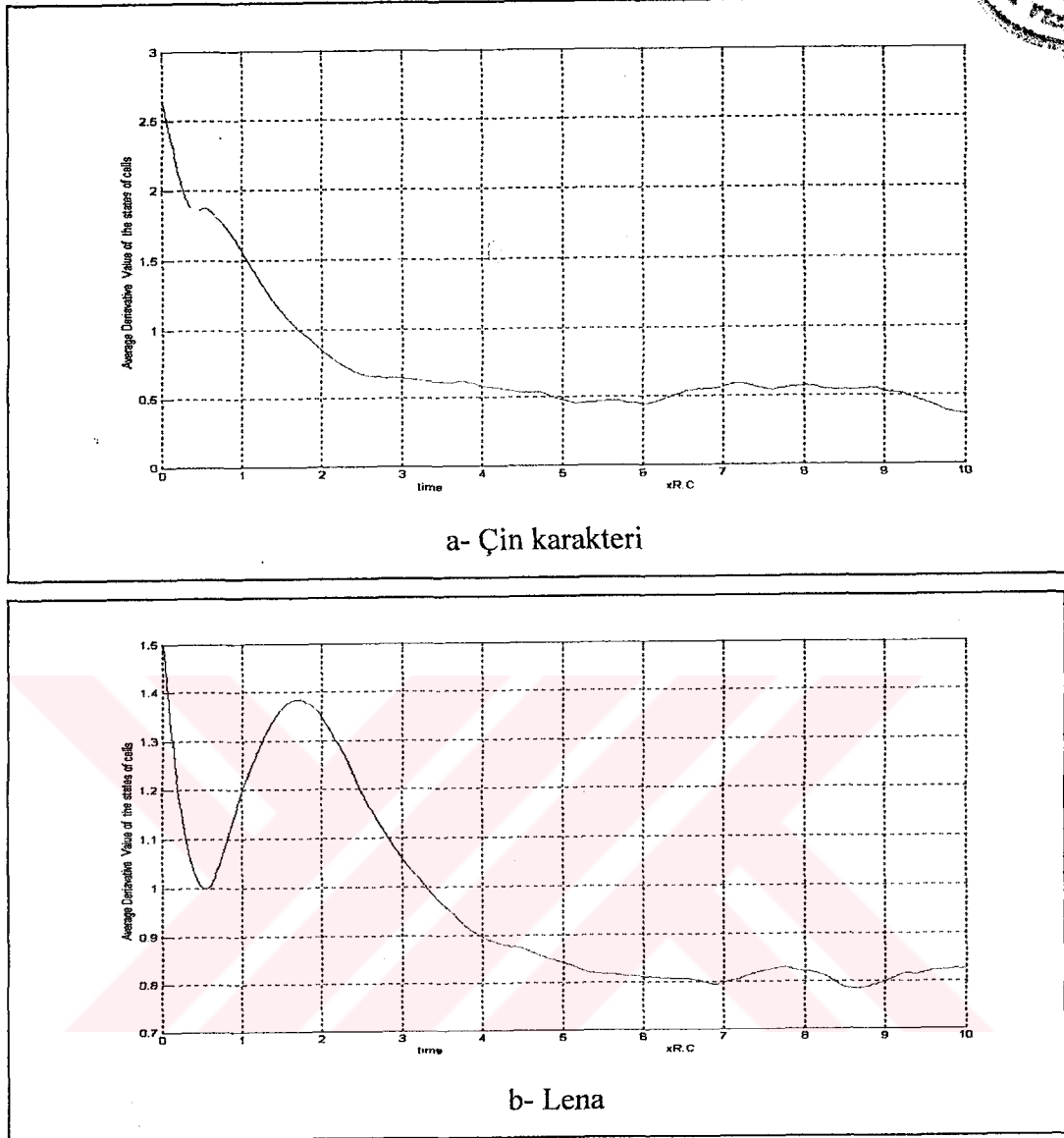


Şekil. 4.8. (4.55) ifadesindeki non-simetrik A şablonu için PWL-HYSA çıkışları.

Şekil 4.9' da görüldüğü gibi AADV monoton azalan değildir. Başlangıçtan $10 \times R \times C$ sonra, sıfıra gitme eğilimi görülmemektedir. Ağın çıkışları gözlemlendiğinde, kararsız olduğu gözlenmektedir. Bir sonraki uygulamada HYSA'yı, non-simetrik şablonla kararlı hale getirebilmek için, (4.51) ifadesindeki koşul gereğince, şablon elemanlarının mutlak değerleri toplamı 9.25 olduğundan bütün elemanlar 10'a bölünerek şablonun non-simetrikliği bozulmadan, elemanlarının mutlak değerleri toplamı 1'den küçük yapılmış olur. Yeni şablon;

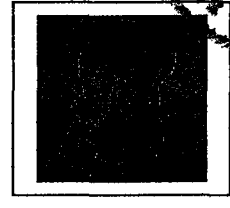
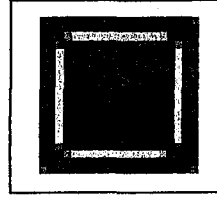
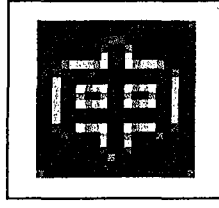
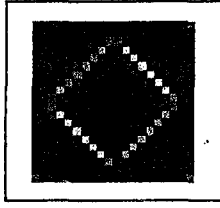
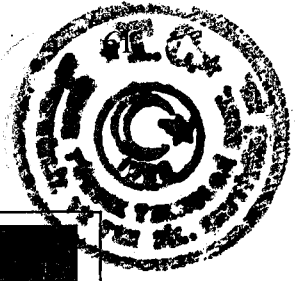
$$A = \begin{bmatrix} -0.025 & 0.050 & -0.100 \\ 0.150 & 0.200 & -0.175 \\ -0.125 & -0.075 & 0.025 \end{bmatrix} \quad (4.56)$$

ifadesinde tanımlandığı gibi oluşturulur.



Şekil. 4.9. Tanımı (4.55) ifadesinde verilen non-simetrik A şablonu kullanılan PWL-HYSA için AADV eğrileri. Giriş görüntüsü ; a-) Çin karakteri, b-) Lena.

Şekil 4.10 ve Şekil 4.11’de (4.56) ifadesinde tanımlanan A şablonu için PWL-HYSA çıkış görüntüleri ve AADV eğrileri gösterilmektedir. Şekil 4.11’de verilen AADV grafikleri HYSA’nın (4.56) ifadesindeki A şablonu için kararlı olduğunu göstermektedir.



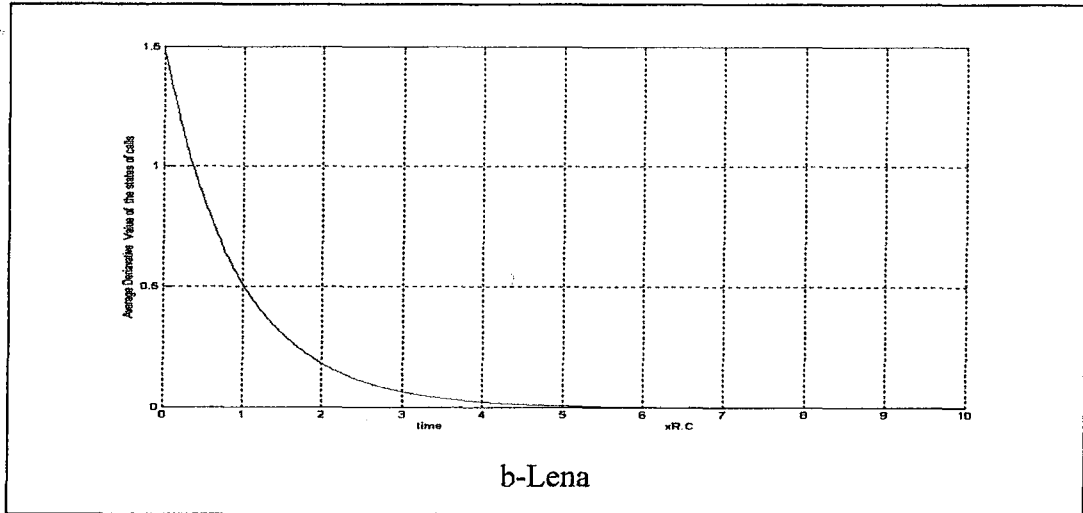
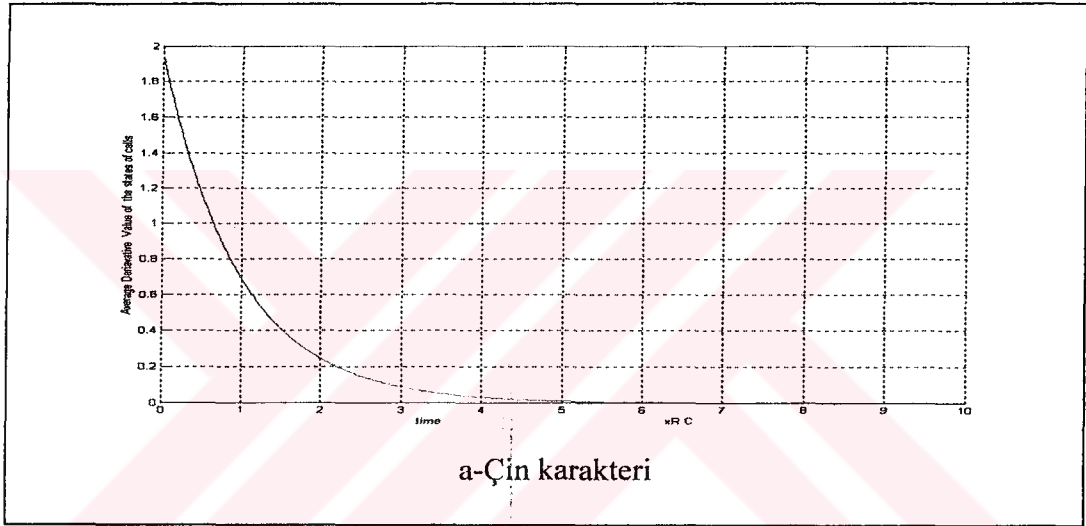
a) Baklava

b) Çin karakteri

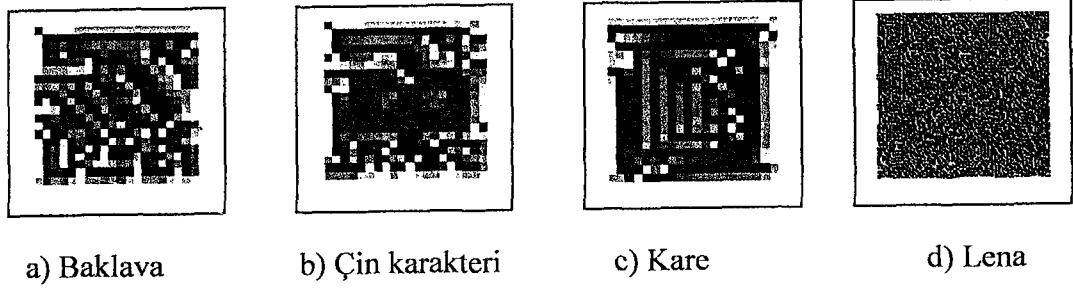
c) Kare

d) Lena

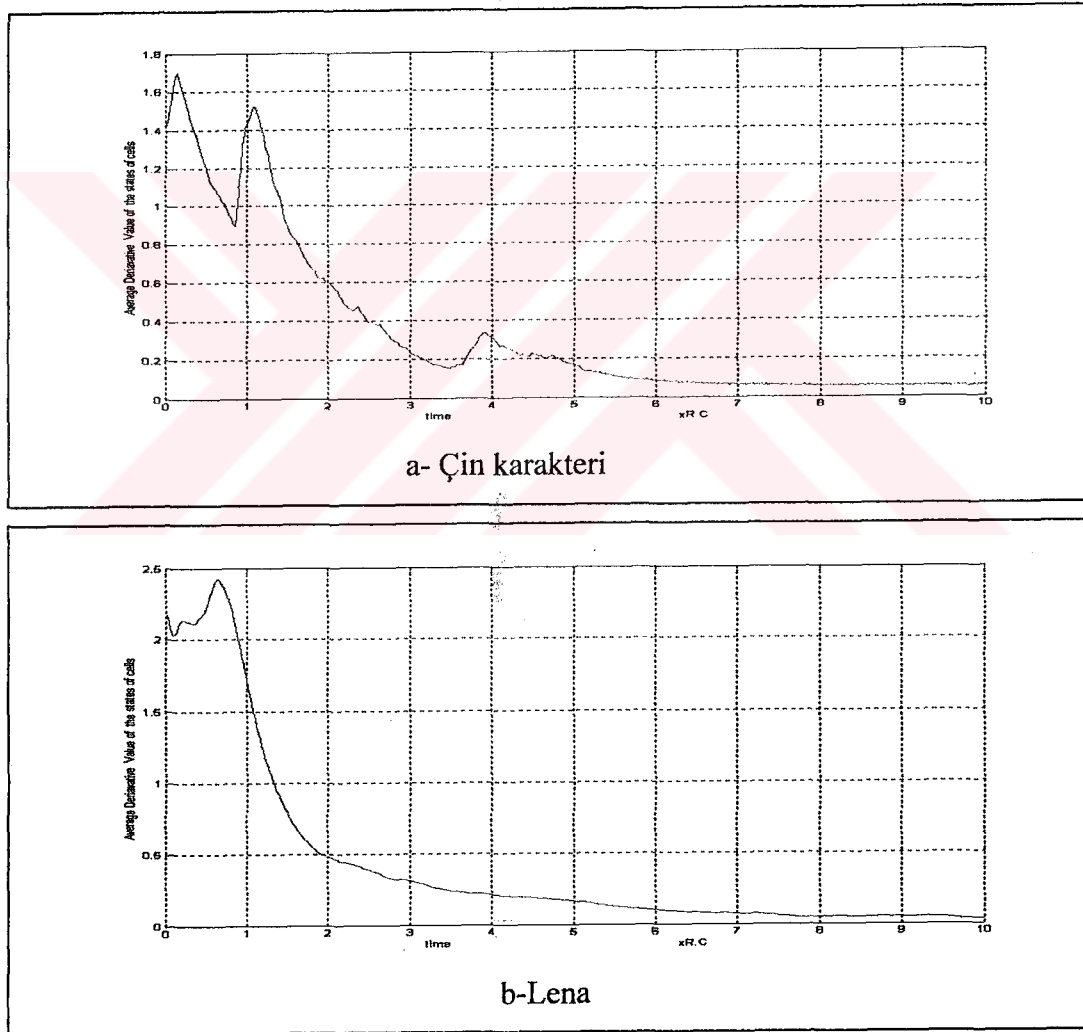
Şekil. 4.10. (4.56) ifadesindeki non-simetrik A şablonu için PWL-HYSA çıkışları.



Şekil. 4.11. Tanımı (4.56) ifadesinde verilen non-simetrik A şablonu kullanılan PWL-HYSA için AADV eğrileri. Giriş görüntüsü; a-) Çin karakteri, b-) Lena.



Şekil.4.12. Eşitlik (4.55) deki non-simetrik A şablonu ve (4.57) ifadesindeki aktivasyon fonksiyonu parametreleri kullanılan TAF-HYSA çıkış görüntüleri.



Şekil.4.13. Eşitlik (4.55) deki non-simetrik A şablonu ve (4.57) ifadesindeki aktivasyon fonksiyonu parametreleri kullanılan TAF-HYSA'ya ait AADV eğrileri. Giriş görüntüsü ; a-) Çin karakteri, b-) Lena

Trapezoid aktivasyon fonksiyonlu HYSA'nın kararlılığı aynı giriş görüntüleri kullanılarak aşağıdaki örneklerde incelenmektedir. İlk olarak A şablonu (4.55) deki gibi seçilir. B ve I şablonları önceki örneklerdeki gibidir. Aktivasyon fonksiyonu parametreleri aşağıdaki gibi belirlenir;

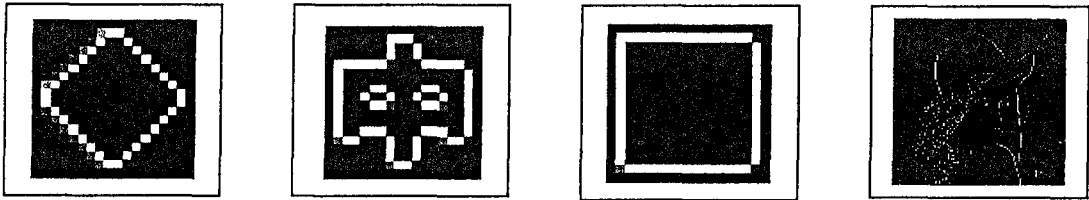
$$[\tau_1, \tau_2, \tau_3, \tau_4] = [-1.0, -0.75, 0.75, 1.0] \quad (4.57)$$

Bu durumda elde edilen çıkış görüntüleri Şekil 4.12'de ve AADV değişimi Şekil 4.13' de gösterilmektedir. Bu şablon ve aktivasyon fonksiyonu değerleri ile HYSA'nın kararsız olduğu gözlenmektedir.

Aktivasyon fonksiyonu parametrelerinin eşitlik (4.57) de verilen yukarıdaki değerlerle aynı olması durumunda (4.50) koşulunun sağlanması için, eğer A şablonu 40' a bölünürse, aşağıdaki gibi elde edilir.

$$A = \begin{bmatrix} -0.0063 & 0.0125 & -0.0250 \\ 0.0375 & 0.0500 & -0.0437 \\ -0.0313 & -0.0187 & 0.0063 \end{bmatrix} \quad (4.58)$$

Bu kararlı non-simetrik A şablonu için TAF-HYSA çıkış görüntüleri Şekil 4.14.' de ve buna ilişkin AADV değişimi eğrileri Şekil 4.15.'de verilmiştir. Beklendiği gibi , bu şablonla HYSA kararlıdır.



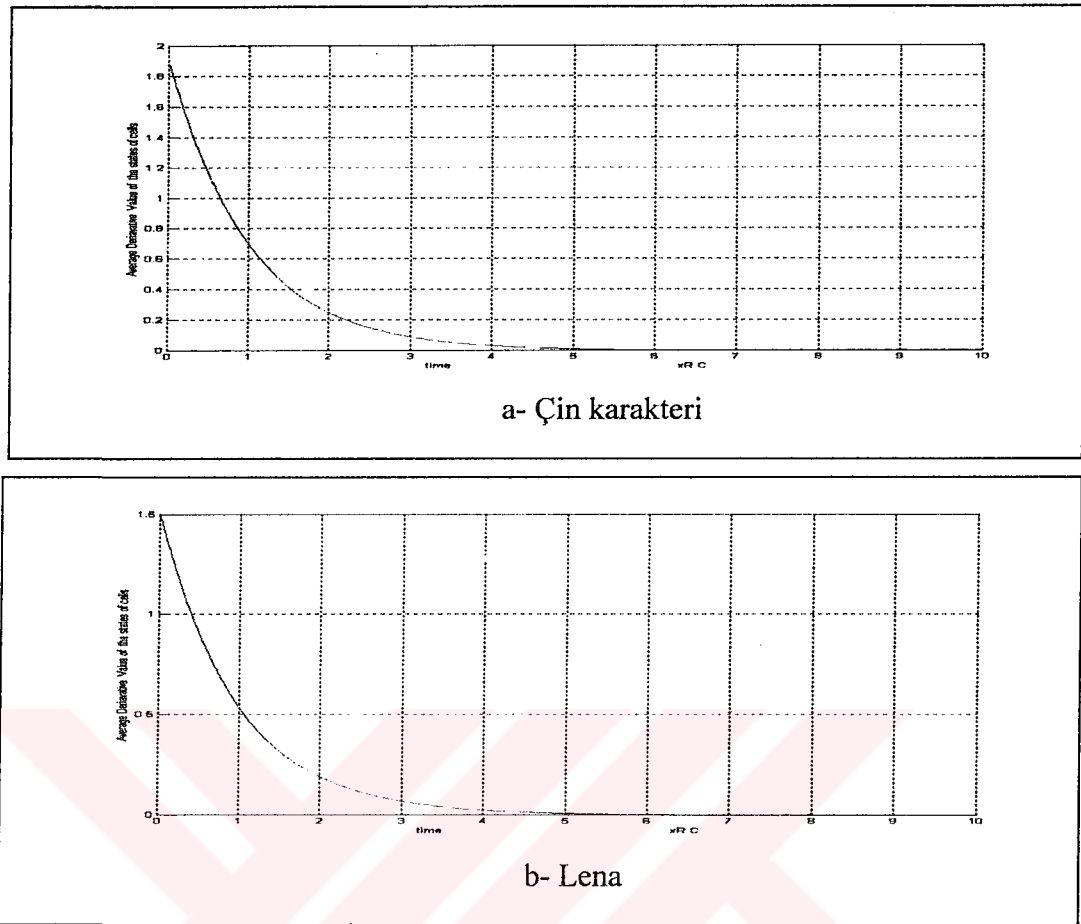
a) Baklava

b) Çin karakteri

c) Kare

d) Lena

Şekil. 4.14. Eşitlik (4.58) deki kararlı non-simetrik A şablonu ve (4.57) ifadesindeki aktivasyon fonksiyonu parametreleri kullanılan TAF-HYSA çıkış görüntüleri.

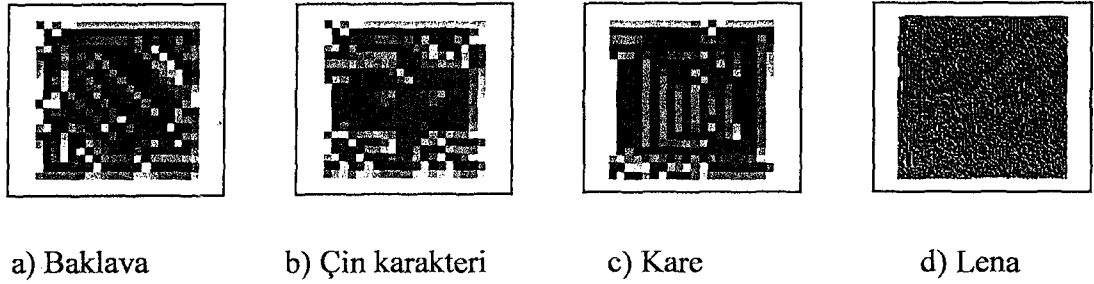


Şekil.4.15. Eşitlik (4.58) deki non-simetrik kararlı A şablonu ve (4.57) ifadesindeki aktivasyon fonksiyonu parametreleri kullanılan TAF-HYSA'ya ait AADV eğrileri. Giriş görüntüsü ; a-) Çin karakteri, b-) Lena

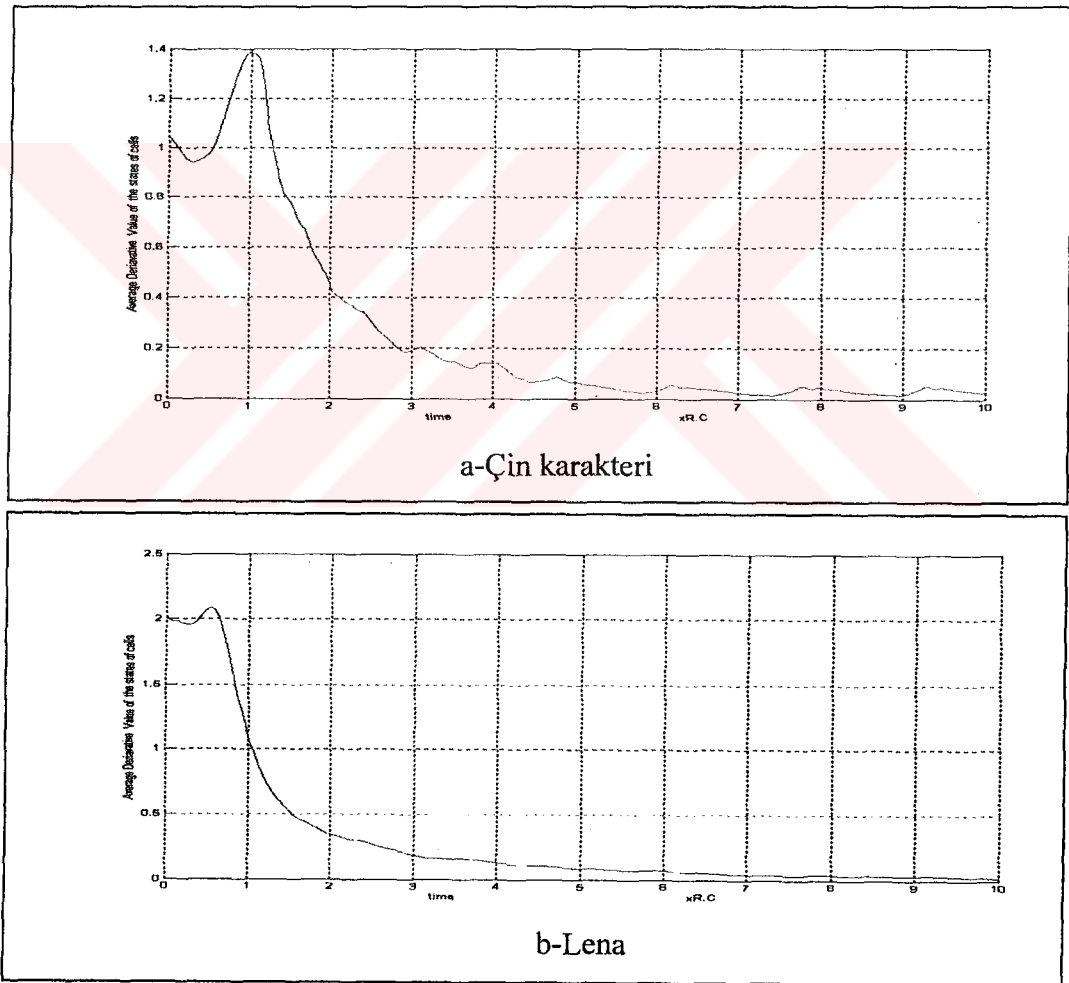
Diğer iki örnek aktivasyon fonksiyonunun parametreleri değiştirilerek kararlılık gözlenmiştir. Aktivasyon fonksiyonu parametreleri aşağıdaki gibi seçilmiştir.

$$[\tau_1, \tau_2, \tau_3, \tau_4] = [-1.0, -0.25, 0.25, 1.0] \quad (4.59)$$

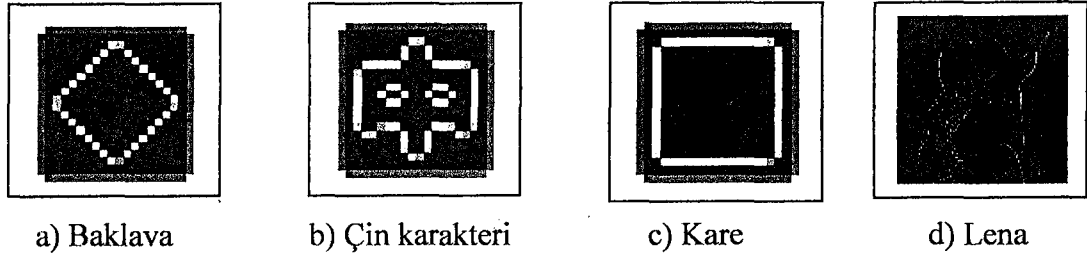
(4.55) ifadesinde verilen non simetrik A şablonu kullanıldığında, beklendiği gibi TAF-HYSA'nın kararsız olduğu gözlenmiştir. Çıkış görüntüleri ve AADV değişimleri Şekil.4.16. ve Şekil.4.17'de verilmektedir



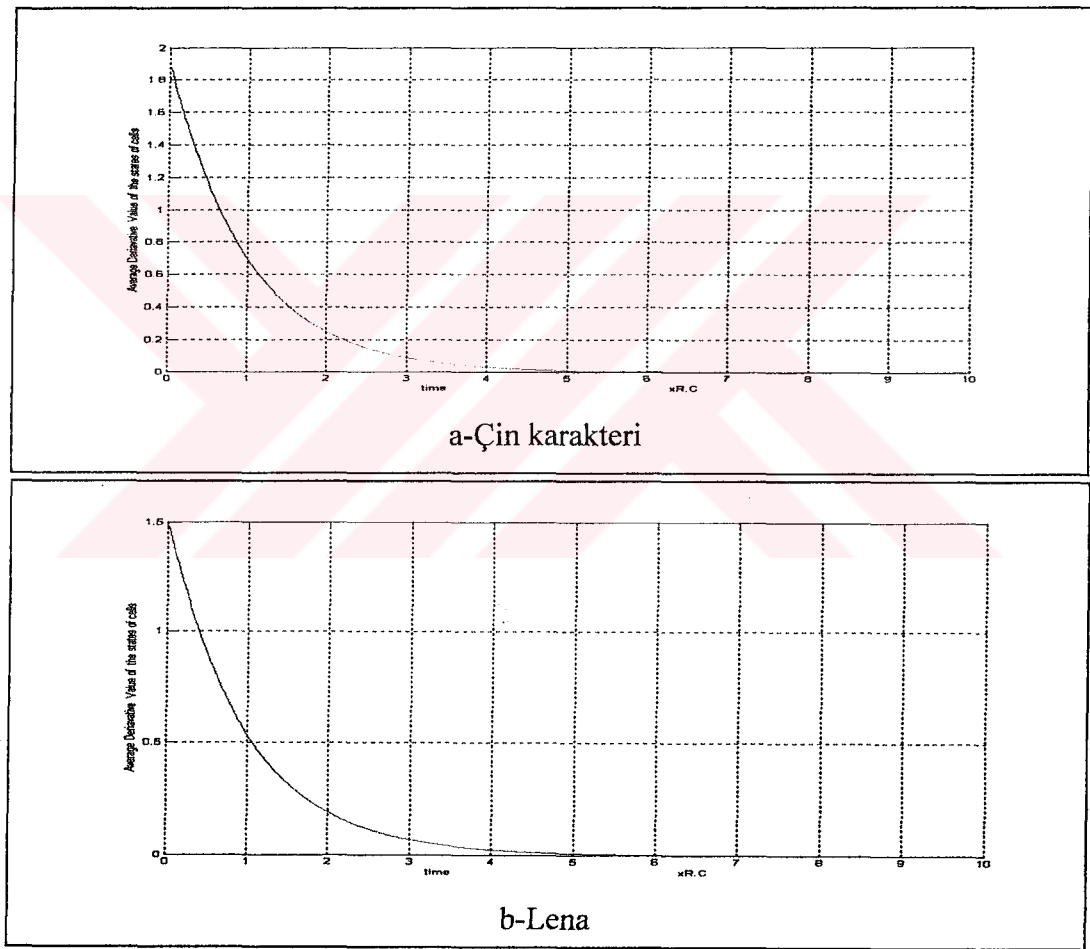
Şekil. 4.16. Eşitlik (4.55) deki non-simetrik A şablonu ve (4.59) ifadesindeki aktivasyon fonksiyonu parametreleri kullanılan TAF-HYSA çıkış görüntüleri.



Şekil.4.17. Eşitlik (4.55) deki non-simetrik A şablonu ve (4.59) ifadesindeki aktivasyon fonksiyonu parametreleri kullanılan TAF-HYSA'ya ait AADV eğrileri. Giriş görüntüsü ; a-) Çin karakteri, b-) Lena



Şekil. 4.18. Eşitlik (4.58) deki non-simetrik kararlı A şablonu ve (4.59) ifadesindeki aktivasyon fonksiyonu parametreleri kullanılan TAF-HYSA çıkış görüntüleri

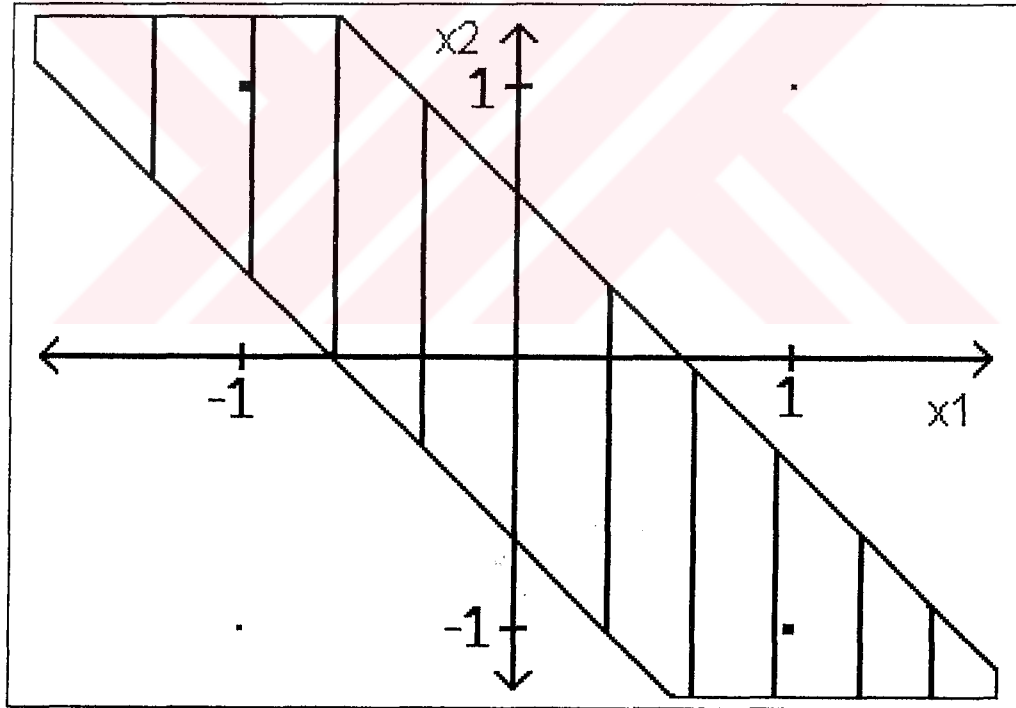


Şekil.4.19. Eşitlik (4.58) deki non-simetrik kararlı A şablonu ve (4.59) ifadesindeki aktivasyon fonksiyonu parametreleri kullanılan TAF-HYSA'ya ait AADV eğrileri. Giriş görüntüsü ; a-) Çin karakteri, b-) Lena

Önceki örneklerdeki benzer yaklaşımla (4.58) ifadesinde tanımlanan A şablonu kullanılırsa kararlılık için yeter şart sağlanmış olur. Bu durumda çıkış görüntüleri ve AADV değişimleri Şekil 4.18 ve Şekil 4.19’da gösterilmektedir.

4.2.2. Lineer olarak ayrıştırılamayan uygulamalar

Daha önce belirtildiği gibi tek katmanlı öz geribeslemesiz standart PWL-HYSA Boolean operasyonları için çok kullanışlıdır, ancak XOR problemini çözme kapasitesi yoktur.(Şekil.4.20). XOR problemini ancak çok katmanlı olarak tasarlandığında çözebilir [Majorana and Chua, 1998], [Chua and Roska, 2002]. Bu çalışmada tek katmanlı TAF-HYSA yapısı ile XOR fonksiyonunu gerçekleştirebilmek için trapezoid aktivasyon fonksiyonu kullanılmıştır.



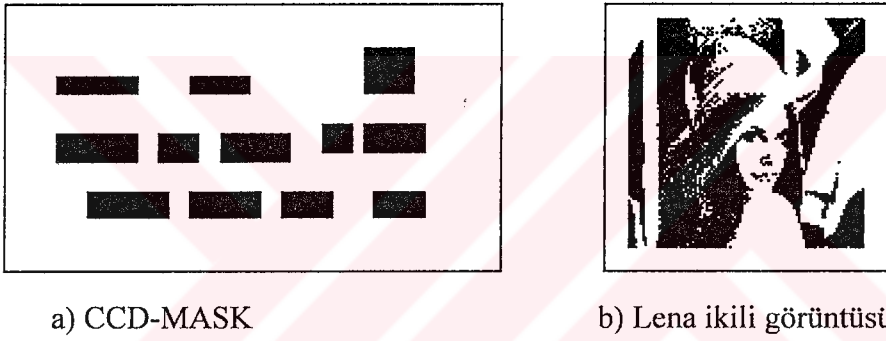
Şekil.4.20. XOR Problemi

4.2.2.1. Parite 2: Bu HYSA türü, bir satır üzerinde yanyana bulunan iki piksel için XOR fonksiyonunu gerçekleştirir. İki pikselin mümkün olan 4 farklı kombinasyonuna göre, piksellerin ikisinde aynı türden değilse, yani biri siyah diğeri

beyaz ise, çıkış beyaz olacak aksi halde çıkış siyah olacaktır. Bu işlem lineer olarak ayrıştırılamayan bir işlemdir ve tek katmanlı standart HYSA ile gerçekleştirilemez. Şablon ve aktivasyon fonksiyonu parametreleri aşağıdaki gibidir.

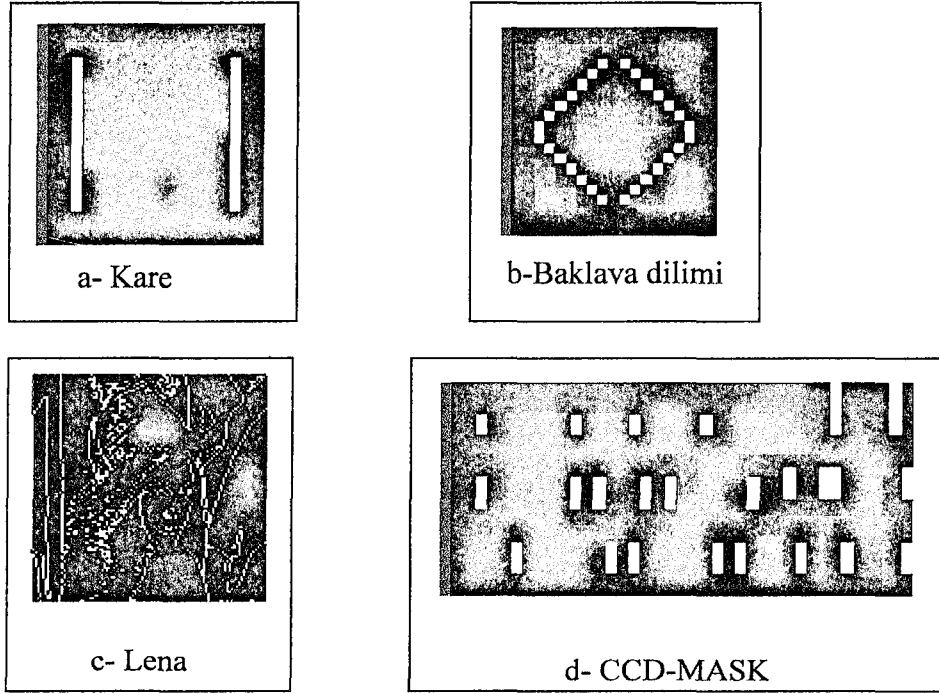
$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.125 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.25 & 0.1 \\ 0 & 0 & 0 \end{bmatrix} \quad I = -0.075 \quad (4.60.a)$$

$$[\tau_1, \tau_2, \tau_3, \tau_4] = [-0.2, -0.1, 0.0, 0.1] \quad (4.60.b)$$

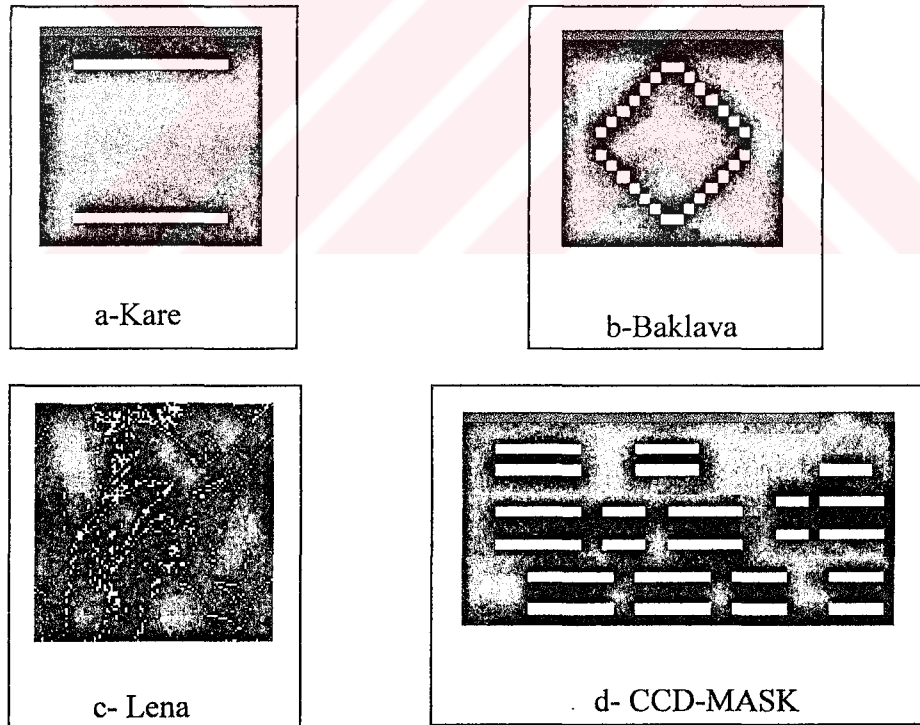


Şekil.4.21. XOR uygulamaları için kullanılan giriş görüntüleri

Daha önce kullanılmayan iki yeni giriş görüntüsü Şekil 4.21’de görülmektedir. Bu verilen giriş görüntüleri için HYSA çıkışı Şekil 4.22’ de gösterilmektedir. Eğer B şablonunun transpozesi alınırsa aynı aktivasyon fonksiyonuyla düşey olarak yanyana bulunan pikseller üzerinde XOR işlemi gerçekleştirilir.(Şekil 4.23).



Şekil.4.22. Yatay parite-2 uygulaması için TAF-HYSA çıkış görüntüleri



Şekil.4.23. Düşey parite-2 uygulaması için TAF-HYSA çıkış görüntüleri

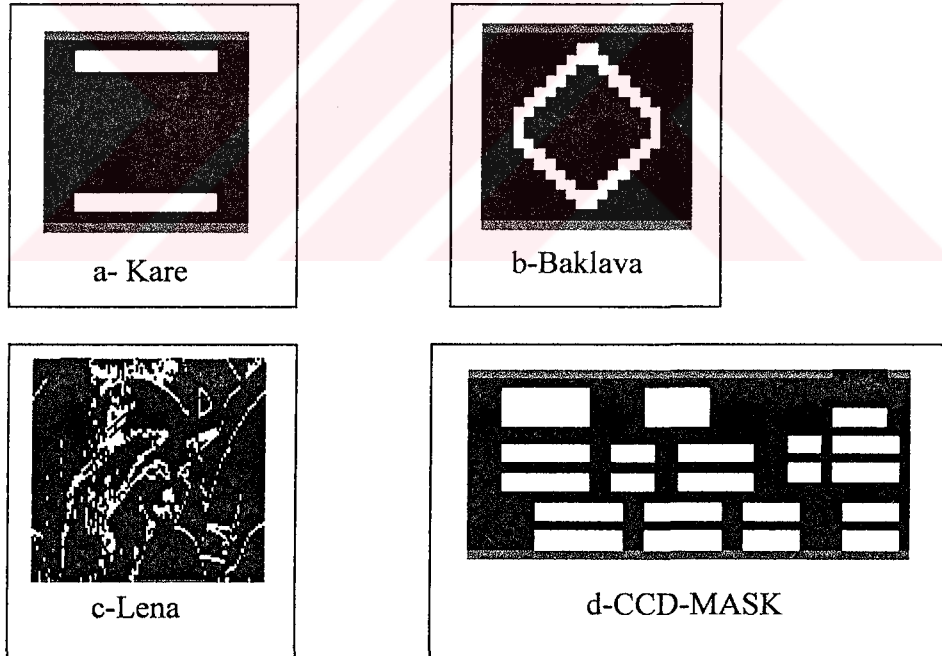
4.2.2.2. Üç Piksel arasında Parite2: Bu tür HYSA, aşağıdaki kurala göre yanyana bulunan üç piksel arasında XOR işlemi gerçekleştirir. Bir satır üzerinde

yanyana üç piksel verildiğinde 8 farklı kombinasyon söz konusudur. Ortadaki pikselin sağ ve sol tarafındaki piksellerden enaz bir tanesi kendisinden farklı ise çıkış beyaz olur. Diğer durumlarda yani her üç pikselde aynı ise çıkış siyah olur. Bu işlem de lineer olarak ayrıştırılamayan bir işlemdir. Bu işlem için şablon ve aktivasyon fonksiyonu değerleri aşağıdaki gibi olacaktır.

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.125 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 & 0 \\ 0.1 & 0.25 & 0.1 \\ 0 & 0 & 0 \end{bmatrix} \quad I = -0.075 \quad (4.61.a)$$

$$[\tau_1, \tau_2, \tau_3, \tau_4] = [-0.3, -0.2, 0.1, 0.2] \quad (4.61.b)$$

Eğer B şablonu transpoze edilirse aynı işlem düşey olarak yanyana bulunan piksellere uygulanabilir. Bu durumda çıkışlar Şekil 4.24 deki gibi olacaktır.



Şekil. 4.24. Aynı düşey kenar üzerinde bulunan bitişik 3 piksel üzerinde gerçekleştirilen parite-2 uygulaması için TAF-HYSA çıkış görüntüleri.

4.2.2.3. Homojen yapının bozulduğu yerin tespiti: Son bir uygulama olarak, lineer olarak ayrıştırılamayan problemlere örnek olarak şu uygulama yapılabilir.

Dokuz komşu piksel verilsin ve birisi merkezde diğer sekiz piksel merkez pikselin etrafına kartezyen olarak yerleşmiş olsun. Eğer bütün pikseller aynı ise çıkış siyah diğer bütün hallerde çıkış beyaz olacaktır. Bu uygulama için kullanılan şablonlar ve aktivasyon fonksiyonu değerleri aşağıdaki gibidir.

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.125 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0.1 & 0.1 & 0.1 \\ 0.1 & 0.25 & 0.1 \\ 0.1 & 0.1 & 0.1 \end{bmatrix} \quad I = -0.075 \quad (4.62.a)$$

$$[\tau_1, \tau_2, \tau_3, \tau_4] = [-0.9, -0.8, 0.7, 0.8] \quad (4.62.b)$$

Bu uygulamaya ilişkin çıkış Şekil 4.25' de gösterilmektedir.



Şekil.4.25. Lineer olarak ayrıştırılamayan uygulama için (4.62) ifadelerinde tanımlı olan şablonlar ve aktivasyon fonksiyonu için TAF-HYSA çıkış görüntüleri.

5. GENETİK ALGORİTMALAR

Genetik algoritmalar, son yıllarda çeşitli mühendislik problemlerin çözümünde sıklıkla kullanılan yapay zeka yöntemlerinden birisi haline gelmiştir. Özellikle çok sayıda parametrenin sonuca etkisinin olduğu durumlarda evrimsel yaklaşımlardan olan genetik algoritma, türev ve integral gibi karmaşık matematiksel hesaplamalara ihtiyaç duymadan, parametre optimizasyonu yapan çok kullanışlı bir algoritmadır.

Genetik algoritmalar, doğal seçim ilkelerine dayanan bir arama ve optimizasyon yöntemidir. Temel ilkeleri John Holland tarafından ortaya atılmıştır [Holland, 1975]. Bilindiği gibi doğada canlıların temel özellikleri, hücre çekirdeğinde bulunan DNA molekülü tarafından belirlenmektedir. Çift sarmallı çok büyük bir molekül olan DNA içerisinde, canlının çeşitli özelliklerini belirleyen genler bulunmaktadır. Tabiattaki bütün canlıların genleri, 4 çeşit nükleotit adı verilen kimyasal yapılardan oluşmaktadır. Bunlar; Adenin, Timin, Guanin ve Citosin diye adlandırılır ve {A,T,G,C} baş harfleri ile sembolize edilirler. Bir bakıma bu 4 nükleotit genetik biliminin alfabesidir. Canlıların kalıtsal bilgilerinin sürekliliği, genlerinin alt kuşaklara geçirilmesi ile sağlanmaktadır. Bir popülasyonda çevre şartlarına uyum sağlayan bireyler daha uzun yaşayabileceklerinden, bu bireylerin diğerlerine oranla daha fazla yavruları olacaktır. Bu durum zaman içerisinde, çevre şartlarına uyumlu genlerin popülasyon içerisindeki görülme frekansının artmasını sağlar. Sonuç olarak, uzun dönemde popülasyon, çevre şartlarına azami ölçüde uyum sağlamış, çevresel zorluklara karşı güçlü bireylerden oluşacaktır. Genetik algoritmalar, işte bu noktada tabiatta yaşanan sürecin benzetiminin yapılarak, bir problemin çözümü için en optimum parametre setinin bulunmasını sağlamaktadır. Doğada genetiğin alfabesi 4 harften oluşurken, genetik algoritmalarda dijital bilgi işlemeye uygun olarak {1,0} ikili seti kullanılmaktadır. Her bir parametre sabit uzunluklu ikili bir kodla kodlanmaktadır. Bir parametreye karşılık düşen kod parçasına gen adı verilmektedir. Parametre sayısı kadar oluşturulan bu ikili kodlar birleştirilerek canlılarda olduğu gibi kromozom adı verilen daha büyük bir yapı içerisinde saklanmaktadır.

Genetik algoritmelerde, doğadaki süreçte olduğu gibi, temel olarak kopyalama, çaprazlama ve mutasyon operatörleri kromozomlara uygulanmaktadır. Bunların dışında evirme (kromozomun evriğinin alınması), tersleme (ikili dizinin tersten yazılması) ve kaydırma (bilginin sağa veya sola kaydırılması) gibi doğada olmayan ilave operatörlerde kullanılmaktadır.

Genetik algoritmaların temel ilkelerinin ortaya atılmasından sonra, genetik algoritmalar hakkında bir çok bilimsel çalışma yayınlanmıştır. Ayrıca, genetik algoritmaların teorik kısmı ve uygulamaları hakkında bir çok uluslararası konferans da düzenlenmektedir. Genetik algoritmaların, fonksiyon optimizasyonu, çizelgeleme, mekanik öğrenme, tasarım, hücresel üretim gibi alanlarda başarılı uygulamaları bulunmaktadır. Geleneksel optimizasyon yöntemlerine göre farklılıkları olan genetik algoritmalar, parametre kümesini değil kodlanmış biçimlerini kullanırlar. Olasılık kurallarına göre çalışan genetik algoritmalar, yalnızca amaç fonksiyonuna gereksinim duyar. Çözüm uzayının tamamını değil belirli bir kısmını tararlar. Böylece, etkin arama yaparak çok daha kısa bir sürede çözüme ulaşırlar [Goldberg, 1989].

Genetik algoritmaların arama yapısı, alt diziler teoremi ve yapı blokları hipoteziyle açıklanmaktadır.

5.1. Genetik Algoritmaların Temel Teoremi

Genetik algoritmaların nasıl arama yaptığı alt dizi kavramıyla açıklanmaktadır. Alt diziler, genetik algoritmaların davranışlarını açıklamak için kullanılan teorik yapılardır. Bir alt dizi, belirli dizi kümeleri arasındaki benzerliği tanımlayan bir dizidir. Alt diziler, $\{0, 1, *\}$ alfabesi kullanılarak tanımlanır. Örneğin H alt dizisi, ilk konumunda 0, ikinci ve dördüncü konumunda 1 değeri olan kromozomlar kümesi içindir.

$$H = 0 \ 1 \ * \ 1 \ *$$

* sembolü dizinin o konumunun hangi değeri alıp almadığının önemli olmadığı anlamındadır. Dizi o konumda 0 veya 1 değeri alabilir. Eğer bir x dizisi, alt dizinin

kalıbına uyarsa x dizisine “H’nin bir örneğidir” denir. Alt dizilerin iki özelliği mevcuttur. Bu özellikler aşağıda verilmiştir [Goldberg, 1989].

1. **Alt dizi derecesi:** Bir H alt dizisinin derecesi $o(H)$ ile gösterilir ve mevcut alt dizi kalıbında bulunan sabit konumların sayısıdır. Bu sayı ikili alfabede 0 ve 1 değerlerinin sayısının toplamına eşittir.
2. **Alt dizi uzunluğu:** Bir H alt dizisinin uzunluğu $s(H)$ ile gösterilir ve mevcut alt dizi kalıbında bulunan belirli ilk ve son konumlar arasındaki uzaklıktır.

Bu iki kavramın genetik algoritmaların temel teoreminde son derece önemli bir yeri vardır. Alt dizi derecesi düşük, alt dizi uzunluğu kısa olan diziler “yapı blokları” olarak adlandırılır. John Holland, genetik algoritmaların işleyişinde uygun yapı bloklarının tanımlanmasını ve bu yapı bloklarının daha uygun yapı blokları elde etmek amacıyla birleştirilmesini önermektedir. Bu fikir yapı blokları hipotezi olarak bilinmektedir. Genetik algoritmanın temel teoremi ise şöyle açıklanmaktadır.

Populasyon ortalamasının üstünde uyum gücü gösteren, kısa uzunluğa ve düşük dereceye sahip alt diziler zamanın ilerlemesiyle üstsel olarak çoğalırlar. Bu çoğalma, genetik işlemler aracılığı ile gerçekleşmektedir ve sonucunda ana-babadan daha üstün özellikler taşıyan bireyler ortaya çıkmaktadır. Bu çözüm kalitesinin kuşaktan kuşağa artması iki nedene bağlanmaktadır. Bu nedenler şöyle açıklanabilir [Emel ve Taşkın, 2002].

Başarısız olan bireylerin üreme şansları azaltıldığı için kötüye gidiş zorlaşmaktadır. Genetik algoritmaların yapısı kötüye gidişi engellemekle kalmamakta, genetik algoritmaların temel teoremi uyarınca, zaman içinde hızlı bir iyiye gidiş de sağlayabilmektedir.

Genetik algoritmaların işleme adımları incelendiğinde bu nedenler daha iyi anlaşılmaktadır. Genetik algoritmalar yapısı gereği, kötü bireyleri yani uygun olmayan çözümleri, operatörleri sayesinde elemektedir. Bu işlemler bir döngü içerisinde durdurma kriteri sağlanana kadar devam etmektedir.



5.2. Genetik Algoritmaların Arama Yöntemi

Genetik algoritmalarla bir problemin çözümü yapılırken, başlangıçta optimizasyonu yapılacak parametre sayısı ve her bir parametrenin kaç bitlik olacağına karar verilmesi gerekir. Parametre sayısı ve parametre çözünürlüğü belirlendikten sonra, bu iki parametrenin çarpımıyla, popülasyonda bulunabilecek her bir kromozomun boyu belirlenir. Belirlenen kromozom uzunluğunda rasgele ikili diziler üretilerek başlangıç popülasyonu oluşturulur. Popülasyonda bulunan ikili dizi (kromozom) sayısı popülasyon nüfusunu vermektedir. Genelde algoritmanın sonuna kadar popülasyon nüfusunun sabit tutulması istenir [Engin, 2001].

Başlangıç popülasyonu oluşturulduktan sonra her bir kromozomun içerdiği ikili koddan parametre değerleri hesaplanarak, bu parametre değerlerinin ulaşılması istenen sonucu sağlayıp sağlamadığı ölçülür. Ulaşılmak istenilen amaca olan uzaklık, veya daha genel tanımla mevcut parametre setiyle elde edilen hata, bir hata (cost) fonksiyonu tanımlanarak ölçülmektedir. Doğal olarak hatanın sürekli azalması arzu edilmektedir. Problemin yapısına bağlı olarak bir uygunluk (fitness) fonksiyonu tanımlanarak, her bir kromozom için hesaplanan hata değerinden, kromozomun uygunluk değeri hesaplanır. Uygunluk fonksiyonu belirlenirken, bir kromozomun hata(cost) değeri yüksekse, uygunluğunun düşük; hata değeri düşükse, uygunluğunun yüksek olmasını sağlayacak bir fonksiyon seçilmelidir. Kromozomun uygunluk değeri, kromozomun çevre faktörlerine (hedefe) uyumluluğun bir ölçüsüdür.

Kromozomların uygunluk değerleri dikkate alınarak popülasyonda sırasıyla, kopyalama, çaprazlama ve mutasyon işlemleri uygulanır. Bu işlemlerin detayların ilerleyen bölümlerde anlatılacaktır. Mutasyon işlemi sonunda, popülasyon yapısı artık değişmiştir. Popülasyonda daha önce bulunmayan genetik bilgiler taşıyan yeni bireyler oluşmuştur. Uygunluk değeri yüksek olan bireylerin sayısı artmışken, uygunluk değeri düşük olanlar popülasyondan atılmıştır.

Başlangıç popülasyonu oluşturulduktan sonra yapılan işlemler, önce her bir kromozomun uygunluk değerinin bulunması ve bu uygunluk değerleri dikkate alınarak genetik operatörler (kopyalama, çaprazlama, mutasyon) bu yeni



populasyona da uygulanır. Her defasında yeni bir populasyon oluşturularak döngü, istenen hedefe ulaşılan kadar devam eder. Algoritmanın durma kriteri sağlandığında, populasyon içerisindeki uygunluk değeri en yüksek olan kromozomun içerdiği parametre değerleri problem için aranan çözüm olacaktır.

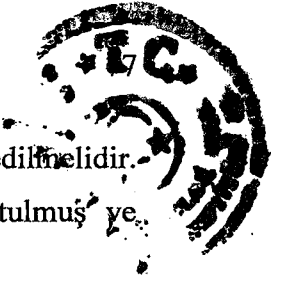
5.2.1. Başlangıç Populasyonunun Oluşturulması

Bir problemin çözümü için genetik algoritma geliştirmenin ilk adımı, tüm çözümlerin aynı boyutlara sahip bitler dizisi biçiminde gösterilmesidir. Dizilerden her biri bir kromozomdur ve problemin muhtemel çözümler uzayındaki rasgele bir noktayı simgeler. Parametrelerin kodlanması, probleme özgü bilgilerin genetik algoritmanın kullanacağı şekle çevrilmesine olanak tanır [Emel ve Taşkın, 2002]. Başlangıç populasyonu oluşturulurken kromozom uzunlukları, populasyon nüfusu (populasyondaki kromozom sayısı) ve kodlama türü önceden belirlenmiş olmalıdır. Kromozom uzunlukları belirlenirken, problemdeki parametre sayısı zaten belli olduğundan, parametre başına bit sayısı; yani her bir parametrenin kaç bitle kodlanacağını kararının verilmesi gerekir. Bütün kromozom uzunluklarının sabit tutulması zorunlu iken bütün parametrelerin eşit uzunluklu ikili kodla kodlanması zorunluluğu yoktur. Hassas olarak ayarlanmak istenen parametreler daha uzun bit dizileri ile kodlanırken, yüksek hassasiyet derecesi gerektirmeyen parametreler daha az bitle temsil edilebilirler.

Kodlamada bir diğer husus ise kodlama algoritmasının seçimidir. Genetik algoritmalarda ikili kodlama yöntemi dışında, çözümü aranan probleme bağlı olarak farklı kodlama yöntemleri de kullanılmaktadır [Goldberg, 1989]. Genelde ikili sayı sistemine göre kodlama yapılırken, Gray kodu kullanıldığında her bir seviye arasında bir bit değiştiğinden dolayı daha iyi sonuçlar alınabilmektedir. Bu nedenle tez çalışmasında Gray kodu kullanılmıştır.

Populasyon nüfusu, populasyonda bulunan kromozomların sayısıdır. Her bir kromozom bir parametre setini temsil ettiğinden, populasyon nüfusu aynı zamanda parametre uzayındaki muhtemel çözüm noktalarının sayısını göstermektedir. Bu nedenle en iyi çözümün bulunabilmesi için kromozom sayısının çok fazla seçilmesi yerine, başlangıç populasyonunda seçilen kromozomların içerdiği parametre

değerleri, çözüm uzayının farklı bölgelerinden örneklenmiş olması tercih edilmelidir. Tez içerisinde populasyon nüfusu 20 birey ile 30 birey arasında tutulmuş ve algoritma boyunca değiştirilmemiştir.



5.2.3. Uygunluk ve Hata Fonksiyonlarının Belirlenmesi

Genetik algoritmanın başarısını etkileyen en önemli unsur seçilen uygunluk fonksiyonudur. Eğer uygunluk fonksiyonun seçiminde hata yapılırsa, bu algoritmayı yanlış çözüm noktalarına yönlendirebilir. Uygunluk fonksiyonun seçiminde önerilebilecek çok genel bir fonksiyon yoktur. Uygunluk fonksiyonu seçiminden önce, çözümü aranan problemde, ulaşılmak istenen çözümün özellikleri çok iyi tanımlanmalı ve parametre uzayında herhangi bir noktadan bu noktaya doğru gelirken karşılaşılabilecek muhtemel durumlar dikkate alınmalı ve bu hususlar ışığında hatayı çok iyi ölçebilen bir hata fonksiyonu tanımlanmalıdır. Hata fonksiyonu tanımlandıktan sonra, değişimi hata fonksiyonunun tersi yönde olan, yani hata fonksiyonu değeri düşerken kendisi artan bir uygunluk fonksiyonu seçilmelidir. Mesela, bir $f(x)$ fonksiyonun maksimumunun bulunmasının istendiği bir problemde genetik algoritmanın, bu fonksiyonun maksimumun bulunduğu noktanın koordinatlarını vermesi beklenir. Bu durumda kromozomlar rasgele seçilmiş x değerlerini gösterecektir. Başlangıçta fonksiyonun alabileceği değer bilinemeyeceği için maksimum değeri 10 civarında bir değer bekleniyorsa, bundan daha büyük bir değer örneğin 20 seçilerek, her bir kromozomun temsil ettiği x değeri için fonksiyonun $f(x)$ değeri hesaplanarak bulunan sonuç 20 değerinden çıkarıldığında bulunan sonuç o kromozomun hata değeri seçilebilir. Böylece herhangi bir kromozom için fonksiyon düşük bir değer ürettiğinde, bu kromozoma ilişkin hata (cost) büyüyecektir. Fonksiyon 20 değerinden daha büyük bir sayı üretemeyeceği için de hata her zaman pozitif kalacaktır. Buna karşılık uygunluk fonksiyonu hatanın tersi olarak yani;

$$Fitness(x) = 1 / E(x) \quad (6.1)$$

şeklinde tanımlanabilir. Bu durumda uygunluk fonksiyonu, $f(x)$ fonksiyonunun yüksek değerler alması durumunda artacaktır. Bir çözümün uygunluk değeri ne kadar

yüksek ise, yaşama ve çoğalma şansı o kadar fazladır ve bir sonraki kuşakta temsil edilme oranı da o kadar yüksektir.

5.2.4. Kopyalama (Çoğalma)

Kopyalama işleminde, populasyon içindeki mevcut kromozomların uygunluk değerlerine göre, bazı kromozomların birden fazla kopyaları popülasyona dahil edilirken, bazı kromozomlar da tamamen popülasyondan silinir. Doğal olarak uygunluk değeri yüksek olan kromozomların sayıları artırılır. Fakat sonuçta toplam kromozom sayısının, yani popülasyon nüfusunun sabit kalmasına dikkat edilir.. Kopyalama işlemi ile, bir sonraki kuşak için yavru üretmek amacıyla hangi ailelerin yer alması gerektiğine karar verilmektedir. Bu doğal seçimdeki en uygunun yaşaması durumuna benzerdir. Bir kromozomun kopyalanma şansı, uygunluk fonksiyonuyla hesaplanan uygunluk değerine bağlıdır [Emel ve Taşkın, 1997]. Seçim yöntemlerine rulet tekerleği seçimi, turnuva seçimi ve sıralama seçimi gibi seçim yöntemleri örnek verilebilir.

5.2.5. Çaprazlama

Mevcut gen havuzunun potansiyelini araştırmak üzere, bir önceki kuşaktan daha iyi nitelikler içeren yeni kromozomlar yaratmak amacıyla çaprazlama operatörü kullanılmaktadır.

Genetik algoritmanın performansını etkileyen önemli parametrelerden biri olan çaprazlama operatörü doğal popülasyonlardaki çaprazlamaya karşılık gelmektedir. Çoğalma işlemi sonucunda elde edilen yeni popülasyondan rasgele iki kromozom seçilmekte ve karşılıklı çaprazlama işlemine tabi tutulmaktadır. Çaprazlama işleminde dizi uzunluğu N olmak üzere, $1 \leq n \leq N-1$ aralığında n tamsayısı seçilmektedir. Bu tamsayı değerine göre dizi çaprazlamaya uğratılır. En basit çaprazlama yöntemi tek noktalı çaprazlama yöntemidir. Tek noktalı çaprazlama yapılabilmesi için her iki kromozomun da aynı gen uzunluğunda olması gerekir. İki noktalı çaprazlamada ise kromozom iki noktadan kesilir ve karşılıklı olarak

pozisyonlar yer değiştirilir [Figlali ve Engin, 2002]. Tez çalışmasında iki noktalı çaprazlama kullanılmıştır.

5.2.6. Mutasyon

Çaprazlama mevcut gen potansiyellerini araştırmak üzere kullanılır. Fakat populasyon gerekli tüm kodlanmış bilgiyi içermez ise, çaprazlama tatmin edici bir çözüm üretemez. Algoritma bir yerel minimuma takıldığında gen havuzundaki kromozomların çoğu bu yerel minimum civarında toplanabilir. Bundan dolayı, mevcut kromozomlardan yeni kromozomlar üretme yeteneğine sahip bir operatör gerekmektedir. Mutasyon, düşük bir olasılık değeri altında populasyon içerisindeki bazı bitleri değiştirir. Mutasyona uğrayan bit değeri 0 ise mutasyon sonunda 1 olur. Veya tersi durumda 1 olan bit değeri mutasyon sonunda 0 olur. Mutasyonun kullanımının temel amacı, genetik çeşitliliği sağlamak veya korumaktır.

5.3. Genetik Algoritmalarda Önemli Parametreler

Parametreler, genetik algoritma performansı üzerinde önemli etkiye sahiptir. Optimal kontrol parametreleri bulmak için bir çok çalışma yapılmıştır fakat tüm problemler için genel olarak kullanılabilecek parametreler bulunamamıştır [Altıparmak, Dengiz ve Smith, 2000]. Bu parametreler, kontrol parametreleri olarak adlandırılmaktadır. Kontrol parametreleri populasyon büyüklüğü, çaprazlama olasılığı, mutasyon olasılığı, kuşak aralığı, seçim stratejisi ve fonksiyon ölçeklemesi olarak sayılabilir. Bu parametreler aşağıda açıklanmıştır .

- **Populasyon Büyüklüğü:** Genetik algoritma kullanıcısı tarafından verilen en önemli kararlardan birisidir. Bu değer çok küçük olduğunda, genetik algoritma yerel bir optimuma takılabilmektedir. Populasyonun çok büyük olması ise çözüme ulaşma zamanını arttırmaktadır. Schaffer ve arkadaşları 1989'da yaptıkları araştırmalar sonucunda, 20-30 arası bir populasyon büyüklüğünün iyi sonuçlar verdiğini belirtmişlerdir [Emel ve Taşkın, 2002].

• **Çaprazlama Olasılığı:** Çaprazlamanın amacı, mevcut iyi kromozomların özelliklerini birleştirerek daha uygun kromozomlar yaratmaktır. Kromozom çiftleri $P(c)$ olasılığı ile çaprazlamaya uğramak üzere seçilirler. Çaprazlamanın artması, yapı bloklarının artmasına neden olmakta fakat aynı zamanda bazı iyi kromozomların da bozulma olasılığını arttırmaktadır.

• **Mutasyon Olasılığı:** Mutasyonun amacı popülasyondaki genetik çeşitliliği korumaktır. Eğer $P(c)$ mutasyon olasılığı artarsa, genetik arama rasgele bir aramaya dönüşür. Fakat bu aynı zamanda kayıp genetik malzemeyi tekrar bulmada yardımcı olmaktadır.

• **Kuşak Aralığı:** Her kuşaktaki yeni kromozom oranına kuşak aralığı denmektedir.

• **Seçim Stratejisi:** Eski kuşağı yenilemenin çeşitli yöntemleri mevcuttur. Kuşaksal stratejide, mevcut popülasyondaki kromozomlar tamamen yavrular ile yer değiştirir. Popülasyonun en iyi kromozomu da yenilendiğinden dolayı bir sonraki kuşağa aktarılamaz ve bu yüzden bu strateji en uygun (elitist) stratejisiyle beraber kullanılmaktadır. En uygun stratejisinde, popülasyondaki en iyi kromozomlar hiçbir zaman yenilenmemektedir, bundan dolayı çoğalma için en iyi çözüm her zaman elverişlidir. Denge durumu stratejisinde ise, her kuşakta yalnızca birkaç kromozom yenilenmektedir. Genellikle, yeni kromozomlar popülasyona katıldığında en kötü kromozomlar yenilenir.

• **Fonksiyon Ölçeklemesi:** Doğrusal ölçekleme, üstsel ölçekleme gibi yöntemler mevcuttur. Probleme göre en uygun ölçekleme yönteminin seçilmesi genetik algoritmanın etkin işlemesi açısından önem taşımaktadır.

6. HÜCRESEL SİNİR AĞLARININ JEOFİZİK UYGULAMALARI

Jeofizikte, bir bölgedeki magnetik ve gravite haritaları yüzey altındaki farklı yoğunlukta ve farklı magnetik özelliklerdeki kütlelerin yerlerinin belirlenmesinde çok sık kullanılan araçlardır. Keza, iletkenlik haritaları ve uzaktan algılama gibi diğer teknikler de kaya yapılarının yerlerinin belirlenmesinde çok yararlı olurlar. Magnetik ve gravite anomalileri ilk olarak filtrelerden geçirilerek anomalinin yorumunu güçleştiren istenmeyen alan bileşenlerinin elimine edilmesi veya zayıflatılması sağlanır. Filtre işleminden sonra, bileşke alana sebep olan kaya yapıları daha kolay tesbit edilebilir. Bu başlangıç filtreleme işlemleri, radyal ağırlıklar yöntemi, en küçük kareler minimizasyonu ve hızlı fourier dönüşümü filtreleme yöntemi gibi yöntemleri içerir.

Kısaca özetlemek gerekirse, yüzey altında bulunan her kütle yüzey üzerindeki magnetik ve gravitasyonel alan dağılımını, magnetik özellikleri, kütleli yoğunluğu ve kütleli büyüklüğü nisbetinde etkiler. Bu çalışmada amaç, yer üzerinde ölçülen magnetik ve gravite alanlarının dağılımlarından, bu alanlara sebep olan yoğun kütleli yerinin belirlenmesidir.

Problemin çözümünde klasik yöntemlerin yerine bu tür problemlerin çözümü için uygun bir yöntem olan HYSA kullanılmıştır.

Tez çalışmasında HYSA eğitimi genetik algoritmalar ile yapılmıştır. Eğitim işleminde A,B, ve I şablonlarına ait parametrelere ikili kodlanmış genler karşılık düşürülerek, en iyi performansı gerçekleştiren kromozom bulunur. Genetik algoritmalar son yıllarda optimizasyon problemlerinin çözümünde başarıyla kullanılan modern yöntemlerden biridir. Hücreli yapay sinir ağı eğitimi şablon parametrelerinin optimizasyonuna dayandığından genetik algoritmalar eğitim için tercih edilmektedir. Tezin özgünlüğü, genetik algoritmalarla eğitilmiş hücreli yapay sinir ağlarının jeofizik alanına ilk defa uygulanmasıdır.



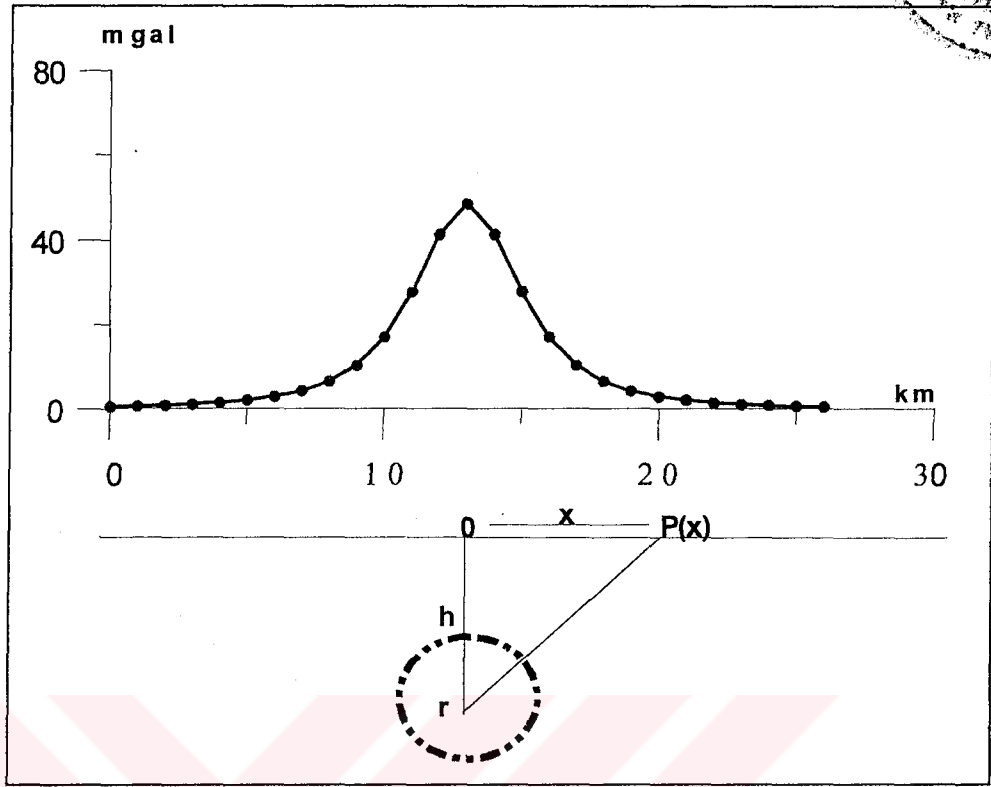
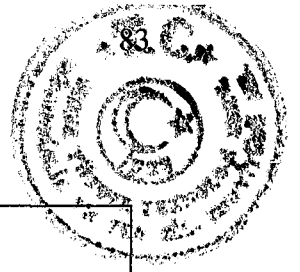
Çalışmada iki farklı veri tipi kullanılmıştır. İlk uygulamada , ülkemizin demir madeni açısından zengin yataklarından birisi olan Dumluca bölgesine ait gravite verileri üzerinde çalışılmıştır. Jeofizikteki temel problemlerden birisi olan potansiyel alan ayrıştırma yöntemi, çok katmanlı hücresel yapay sinir ağları ile gerçekleştirilmeye çalışılmıştır. İkinci uygulamada ise Türkiye topraklarında hakimiyet sürmüş olan Hitit İmparatorluğuna ait arkeolojik kalıntıların bulunduğu Sivas-Altınyayla bölgesine ait magnetik veriler kullanılarak, hücresel yapay sinir ağının yeraltındaki görünmeyen bir binanın odalarına ait sınırları tespit etmesi sağlanmıştır. Aslında bu problem, jeofizikteki magnetik anomali ayrıştırmasının bir uygulamasıdır. Her iki uygulama için başlangıçta sentetik magnetik ve gravite verileri üretilerek bu sentetik veriler hücresel yapay sinir ağının eğitiminde kullanılmıştır.

6.1.Hücresel Yapay Sinir Ağları İle Potansiyel Anomali Ayrıştırma

Potansiyel kaynaklı alanlardan ilk örnek, gravite yöntemindendir. Bu örnekte çeşitli derinliklerde, ebatlarda ve yarıçaplarda 3 adet küre kullanılmıştır. Herhangi bir kürenin yeryüzündeki herhangi bir P noktasındaki gravite etkilerini hesaplayabilmek için,

$$G = \frac{k_0 m h}{(x^2 + y^2 + h^2)^{3/2}} \quad m = \frac{4}{3} \pi r^3 \rho \quad (6.1)$$

formülü kullanılır. Burada, k_0 Gravitasyon sabiti olup, $6.67 \cdot 10^{-8} \text{ cm}^3/\text{grsn}^2$ dir, r , kurenin yarıçapı, ρ , kürenin yoğunluğu, x ve y koordinatlar, h kürenin orta noktasından yeryüzüne kadar olan uzaklık (kürenin derinliği), G , gravite anomali değeridir. Burada derinde bulunan küre rejyonal etki yaratarak yüzeye daha yakın kürelerin gravite anomalilerini bastırmaktadır. Şekil 6.1. gravite anomali değerinin x eksenini boyunca uzaklığa olan değişimini göstermektedir.



Şekil. 6.1. Sentetik küre modeli

6.1.1. Hücresel Yapay Sinir Ağının Sentetik Verilerle Eğitimi

Eğitim verisi olarak farklı boyutlarda ve yoğunluklarda bulunan üç küre kullanıldı. Küreler farklı koordinatlara yerleştirildi. Sentetik kürelere ait parametreler Tablo 6.1.'de gösterilmektedir. Konumu açısından en derinde bulunan küre yüzeye yakın olan kürelerin gravite anomali etkisini bozmakta bir başka deyişle gölgelemektedir.

Hücresel yapay sinir ağının eğitiminde amaç, hücresel yapay sinir ağına bir giriş verisi uygulandığında hücresel sinir ağının çıkışında arzu edilen istenilen çıkış görüntüsünün elde edilebilmesi için şablon parametrelerinin optimum değerlerinin bulunması demektir. Problemin doğası gereği gravite alan dağılımı yönden bağımsız olduğundan A ve B şablonları merkez elemanlarına göre simetrik seçilmiştir. HYSA komşuluğu 1 seçildiğinden A ve B şablon elemanlarının belirlenmesi için her ikisinin ayrı ayrı iki parametresinin belirlenmesi yeterlidir. I şablonu da 1 elemandan oluştuğu için ağı eğitimi toplam 5 parametrenin



optimize edilmesi gereklidir. Bu nedenle genetik algoritmada her bir kromozom 5 genden oluşmaktadır. Hücresel yapay sinir ağının eğitimi sırasında kullanılan genetik algoritma parametreleri Tablo. 6.2.'de gösterilmektedir. Ağın eğitiminde genetik algoritma aşağıdaki adımlar takip edilerek uygulanmıştır.

Tablo.6.1. Üç küreden oluşmuş gravite eğitim modeline ait parametreler.

Parametreler	(x,y) Koordinatlar	h (Derinlik) (m)	r (Yarıçap) (m)	ρ Yoğunluk gr / cm ³
Küre 1	31,31	15	10	1
Küre 2	19,16	8	3	1.2
Küre 3	11,16	6	2	1.3

Tablo.6.2. Genetik algoritma parametreleri

PARAMETRELER	DEĞER
Kromozom sayısı	30
Parametre başına bit sayısı	16
Şablon parametresi sayısı	5
Kromozom uzunluğu	80
Populasyondaki toplam bit sayısı	1600
Mutasyon olasılığı	1 %
Evrilen bit sayısı	24
Şablon parametrelerinin değişim aralığı	[-8, 8]

Adım 1. Başlangıç popülasyonunun oluşturulması: Popülasyon matrisi adı verilen bir matris oluşturulur. Matrisin her satırı bir kromozoma ait bilgileri içerir. 30 kromozom seçildiğinden popülasyon matrisinde 30 satır vardır. Her bir kromozom



80 bitlik (ikili) binary dizilerden oluştuğu için popülasyon matrisinde 80 sızma vardır. Başlangıçta popülasyon matrisi rasgele(random) oluşturulur.

Adım 2. HYSA şablonlarının oluşturulması: Kromozomlar HYSA şablonlarının- A, B, I - ikili kodlarını temsil ederler. Bu adımda kromozomlarda kodlanmış olan şablon parametrelerinin ikili kodları çözülür ve şablon parametreleri $[-8,8]$ aralığında gerçek değerleri ile oluşturulur. Parametreler Gray koduna göre kodlanmıştır. Her bir parametre 16 bit seçildiğinden, 5 şablon parametresi için 80 bitlik kromozom kullanılır. İlk 5 bit şablon elemanlarının 1. bitini, ikinci 5 bit şablon elemanlarının 2. bitlerini gösterecek şekilde kromozom aşağıdaki gibi kodlanmıştır.

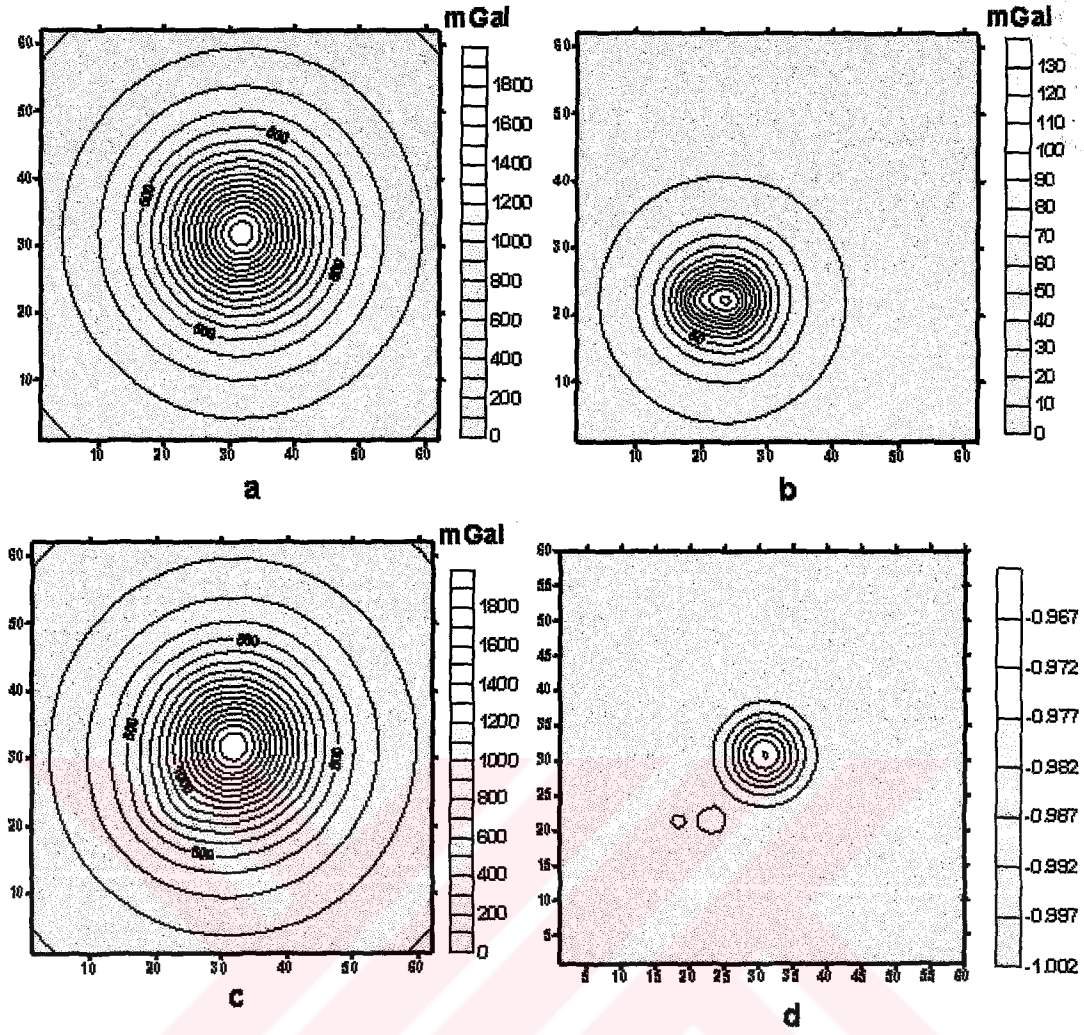
$$S = [A_{1,1}, A_{2,2}, B_{1,1}, B_{2,2}, I] \quad (6.2)$$

Adım 3. Her bir kromozom için hata fonksiyonunun değerinin hesaplanması; Bu adımda, eğitimde kullanılan eğitim görüntüsü HYSA'na giriş olarak verilir. HYSA'nın durum değişkenleri de bu giriş görüntüsüne eşitlenir. Adım 2.' de bulunan şablon değerleri kullanılarak HYSA çıkışı izlenir. HYSA çıkışında oluşan görüntü ile, olması istenilen hedef görüntü karşılaştırılır. Aradaki farka göre hata fonksiyonu (cost function) değeri hesaplanır. Bu işlem bütün kromozomlar için tekrarlanarak bulunan hata fonksiyonu değerleri saklanır. Hata fonksiyonu aşağıdaki gibi seçilmiştir.

$$\text{cost}(A, B, I) = \sum_i^m \sum_j^n P_{i,j} \oplus T_{i,j} \quad (6.3)$$

Burada A, B, I şablon elemanları, m ve n görüntünün piksel numaralarını, P ve T sırasıyla giriş görüntüsünü ve hedef görüntüyü göstermektedir. Şekil 6.2.a. giriş verisini Şekil 6.2.d. ise çıkış verisini göstermektedir. “ $\oplus$ ” sembolü P ve T . arasında yapılan XOR operasyonunu göstermektedir. Paha fonksiyonu değerleri bulunduktan sonra bu değerlere göre uygunluk (fitness) fonksiyonu değerleri aşağıdaki kural uyarınca yapılır.

$$\text{fitness}(A, B, I) = m * n - \text{cost}(A, B, I) \quad (6.4)$$



Şekil 6.2. Gravitasyonel veriler için HYSA'nın eğitiminde kullanılan Tablo 6.1. de özellikleri verilen üç sentetik küreye ait ; (a) Bölgesel anomali alanı, (b) Rezidüel anomali haritası, (c) Toplam gravite anomali haritası, (d) ÇS-HYSA çıkışı.

Başka bir kural da durma kriteri için belirlenir. Söz konusu durma kriteri aşağıda gösterilmiştir.

$$stcriterion = 0.99 * m * n \quad (6.5)$$

Burada m görüntünün satır, n ise sütun numarasını göstermektedir. Eğer maksimum uygunluk değeri durma eşiğinden büyükse algoritma durdurulur ve uygunluk değeri en yüksek olan kromozom seçilir. Bu seçilen en iyi kromozomdan

oluşturulan şablonlar bu uygulamada istenilen hedefi gerçekleştirmek için kullanılacak en uygun şablonlardır.

Adım 4. Yeni Populasyon Oluşturma : Bir sonra ki jenerasyonu oluşturmadan önce popülasyondaki uygunluk değerleri azalan şekilde sıralanır ve popülasyondaki bütün kromozomlara ait uygunluk değerleri toplamı alınarak normalize edilir. 0 ile 1 arasında rasgele bir r değeri seçilir. Sonra uygunluk değerleri normalize edilmiş popülasyon bireylerinden uygunluk değeri r ' ye eşit ve veya daha büyük olan en yakın kromozom bir sonraki popülasyona eklenir. Bu işlem 30 kere tekrarlanır ve uygunluk değeri iyi olmayan kromozomlar popülasyondan silinir. Bu bahsedilen işleme genetik algoritmalarda kopyalama (reproduction) işlemi adı verilir. Kopyalama işleminin ardından 15 kromozom çifti bir sonraki jenerasyondaki kromozomlara ebeveyn olmak üzere rasgele seçilir. "s1" ve "s2" gibi 1 ile kromozom uzunluğu 80 arasında iki sayı seçilir. "s1" ve "s2" arasındaki bit dizisi çaprazlama bölgesi olarak adlandırılır. Çaprazlama işlemi sırasında, kromozom çiftlerinin "s1" ve "s2" arasındaki bit dizileri aralarında değiş tokuş edilir. Böylece önceki kromozomların parçalarından oluşan iki yeni kromozom oluşur. Sonuç olarak, bir sonraki popülasyonu oluşturmak üzere 30 yeni kromozom oluşturulur. Bu kromozomlar üzerinde mutasyon işlemi uygulanır. Bu uygulamada mutasyon olasılığı %1, seçildiğinden bütün popülasyon toplam 24 bit rasgele seçilerek, bit değerinin evriği alınır. Kopyalama işleminden önceki en iyi kromozom olduğu gibi yeni popülasyona katılır. Popülasyon nüfusunu sabit tutmak için bu eklenen eski kromozoma karşılık son jenerasyondaki kromozomlardan birisi rasgele seçilerek popülasyondan silinir. Önceki jenerasyondaki en iyi kromozomun aynen yeni popülasyona katılmasından maksat, bir önceki jenerasyondaki en iyi bilgiyi saklama ihtiyacındandır. Artık yeni popülasyon, yani yeni kuşak oluşturulmuş olur. Yeni kuşak oluşturulduktan sonra optimum sonucu aramak için 2. adıma dönülür ve algoritma durma kriterini gerçekleninceye kadar bu şekilde devam eder.

Eğitim işleminin sonucunda 403 jenerasyon sonunda istenilen sonuca ulaşıldı. Aşağıdaki şablonlar optimum şablon olarak bulundu.

$$A = \begin{bmatrix} -0.2246 & -0.2246 & -0.2246 \\ -0.2246 & 0.5000 & -0.2246 \\ -0.2246 & -0.2246 & -0.2246 \end{bmatrix} \quad B = \begin{bmatrix} -0.9513 & -0.9513 & -0.9513 \\ -0.9513 & 7.7656 & -0.9513 \\ -0.9513 & -0.9513 & -0.9513 \end{bmatrix} \quad I = -2.6409$$

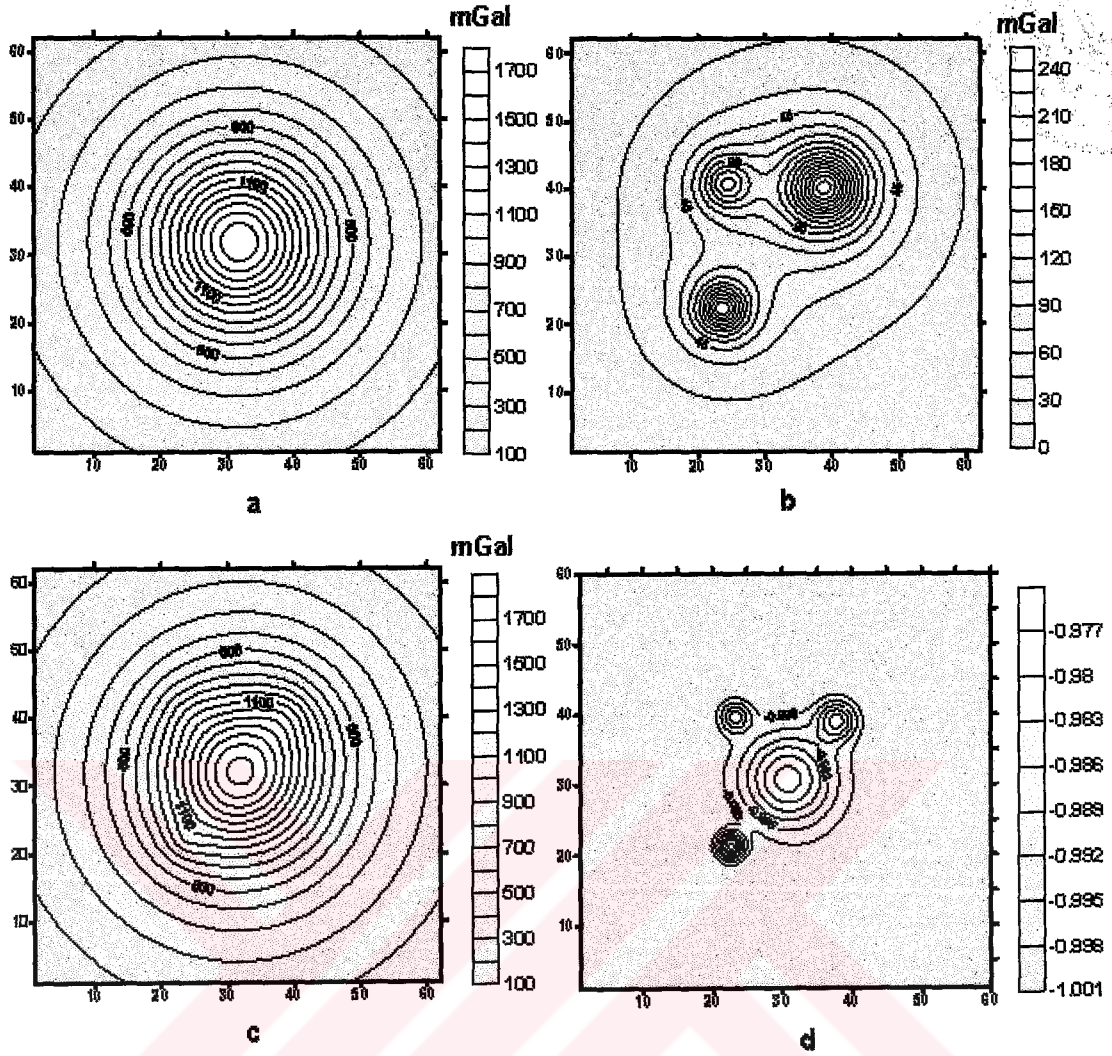
Bu şablonlara ait en iyi kromozom ise aşağıdaki yapıda elde edildi.

BC = [000101011110110101111001010011000000010100001101101000100101101010
00010010100101]. (6.6)

HYSA eğitildikten sonra bu şablon parametreleri kullanılarak Tablo 6.3. de özellikleri verilen eğitim verisinden farklı 4 sentetik küreden hasil olan gravite alanı verisi normalize edilerek sınır ağı girişine uygulandı . Şekil.6.3.'de bu test verisine ilişkin klasik yöntemlerle elde edilen sonuçlar ve HYSA çıkış görüntüsü görülmektedir. Görüldüğü gibi HYSA çok net bir şekilde kürelerin konumlarını belirleyebilmektedir.

Tablo. 6.3 Sentetik gravite test verileri.

Parametreler	x,y (koordinatlar)	h (Derinlik) (m)	r (Yarıçap) (m)	ρ Yoğunluk gr / cm^3
Küre 1	31,31	20	10	1.2
Küre 2	41,43	8	4	1
Küre 3	18,44	6	3	0.9
Küre 4	19,16	5	3	1.1

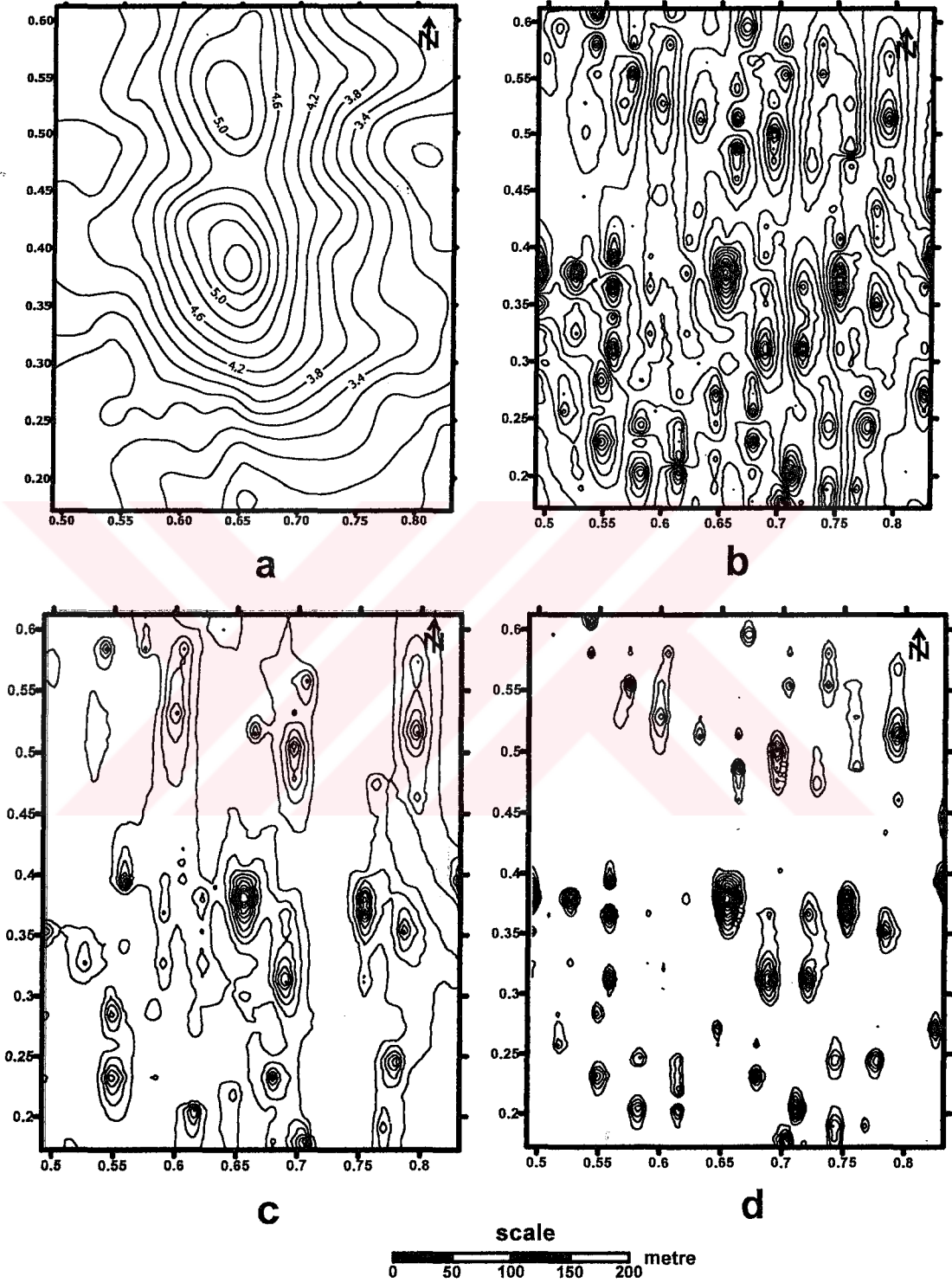


Şekil 6.3. Gravitasyonel veriler için HYSA'nın testinde kullanılan, özellikleri Tablo 6.3. de verilen 4 sentetik küreye ait; (a) Bölgesel anomali alanı (b) Rezidüel anomali haritası (c) Bouguer anomali alanı (d) ÇS-HYSA çıkışı.

6.1.2. Dumluca Bölgesindeki Maden Yataklarının Tespiti

Türkiye'de Divriği civarında ve yakın bölgede üçüncü jeolojik devire ait kuartzdan oluşmuş metmorfitler ve Akdağ bölgesinde limonit yapıları bulunmaktadır. Bölgede demir yatakları ultrabazik tektonikler içinde podiform şekilde görülmektedir. Ultrabazik tektonik bölgelerdeki kayalar altın, kobalt ve bakır yönünden fakirdir. Dumluca demir yatakları yalnızca tektonik kayalar içerir ve tektonik aktiveyle orijinal yeri belirlenir. Dumluca demir yatağına ait gravite anomali haritası şekilde gösterilmektedir. Bu uygulamada hücresel sinir ağı sentetik gravite verileri ile eğitildiğinde elde edilen şablon parametreleri kullanılarak

Dumluca bölgesinden alınan gerçek verilere uygulandı. Şekil 6.4. uygulanan



Şekil 6.4. (a) Dumluca gravite anomali haritası (kontor aralığı; 0.02) b) Halck'ın ikinci türev yöntemine göre elde edilmiş çıkış haritası (kontor

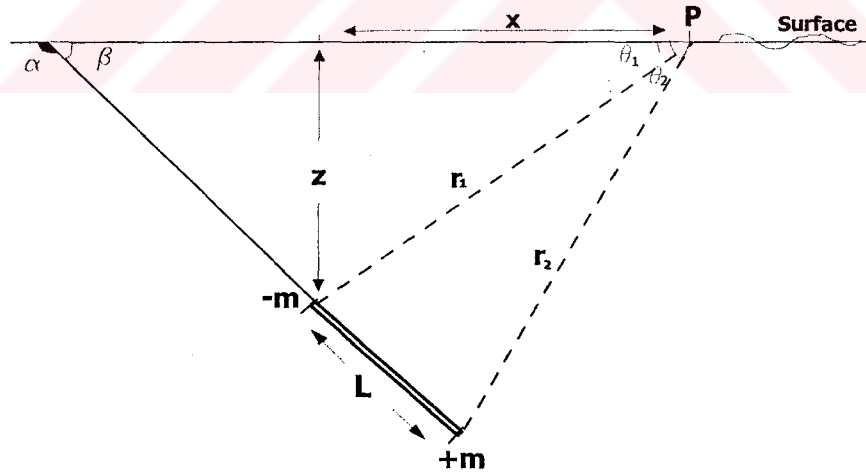
aralığı.01). (c) Üç katmanlı ÇS-HYSA çıkışı. (d) Beş katmanlı ÇS-HYSA çıkışı

giriş verisini ve HYSA çıkışında elde edilen çıkış verisini göstermektedir. HYSA çıkışı klasik yöntemle elde edilen çıkış haritası ile karşılaştırıldığında HYSA çıkışında gravite potansiyel alan kaynaklarının yerlerinin daha belirgin olarak tespit edildiği açıkça görülmektedir.

6.2. Genetik Algoritmalarla Eğitilmiş Hücresel Sinir Ağları İle Magnetik Anomali Ayırıştırma

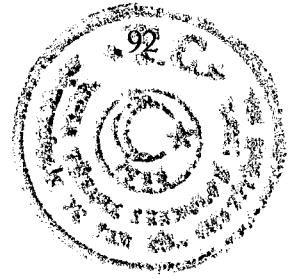
Potansiyel kaynaklı alanlardan ikinci örnek ise magnetik anomali ayırıştırma. Yer altındaki bir magnetik dipolden hasıl olan alanın düşey bileşeni Z aşağıdaki gibi hesaplanır (Telford et al. 1981).

$$Z = kF_0S \left\{ \left(\frac{1}{r_1^2} \right) \left(\frac{z}{r_1} \right) - \left(\frac{1}{r_2^2} \right) \left(\frac{z + L \sin \alpha}{r_2} \right) \right\} = kF_0S \left\{ \left(\frac{z}{r_1^3} \right) - \left(\frac{z + L \sin \alpha}{r_2^3} \right) \right\} \quad (6.7)$$



Şekil 6.5. Magnetik dipol modeli

Burada k suspektibilite, F_0 yerin toplam alanı ve S yüzey alanıdır. Çubuğun kendi eksenini boyunca magnetize olduğunu kabul ettiğimizden, bu ifade yalnızca aşağıdaki şartlardan birisi veya her ikisi birden gerçekleştiğinde geçerlidir.

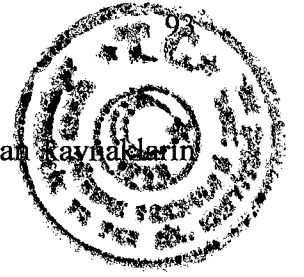


- (i) Çubuktan hasıl olan dış alandan çok büyüktür.
- (ii) Çubuk alan doğrultusunda yönlenmiştir.

Potansiyel alan haritaları, alan çizgilerinin birbirleri üzerine binerek haritayı bulanıklaştırmasından dolayı çok sayıda anomali içermektedir. Mesela, bir magnetik harita bölgesel ve mikro anomalilerden oluşabilir. Bu durumda, alanı oluşturan kaynakların sınırlarının tesbiti birbirine yakın kaynakların birbirlerine olan etkilerinden dolayı güçleşir. Bu da kaynakların yerlerinin yanlış belirlenmesine yol açar. Bir anomali türü bir diğerini maskeleydiğinden farklı anomalilerin birbirlerinden ayırt edilmesi gereklidir. Jeofizik haritalarında incelenen potansiyel alan verileri yer altındaki bütün kaynakların ürettiği alanlarının bileşkesinden oluşur. Bazı haritalarda aranılan hedef yapılar sıklıkla küçük ölçekli yapılardır ve yüzeye yakında gömülüdür. Bu hedeflerin etkileri, hedeften daha büyük olan veya daha derinde bulunan rezidüel kaynaklardan ortaya çıkan bölgesel alanlar içerisinde saklanır. Doğru kestirim ve başlangıç alan gözlemlerinden bölgesel alanların çıkarılması hedef kaynaklar tarafından üretilen artık alan değerlerini verir. Sayısal modelleme ve yorum rezidüel alan verileri üzerinde yapılır ve yorumun güvenilirliği çok büyük ölçüde bölgesel rezidü ayrıştırılmasının başarısına bağlıdır. İlk adım olarak magnetik ve gravite anomalilerin yorumlanması, istenilen anomaliyi izole etmek için istenmeyen alanları elimine etmek veya zayıflatmak için haritanın daha kullanışlı hale gelmesini sağlar. Bu başlangıç filtreleme işlemi için, FFT filtreleme yöntemleri, rekürsif filtreler ve rasyonel yaklaşım teknikleri kullanılır. Aranılan hedefin (sinyalin) jeolojik niteliğinin gravite cevabı incelenen gravite alanının bir bölgesinde veya baskınsa veya alanın güç spektrumunun bir spektral bandında yoğunlaşmışsa gravite anomalinin ayrıştırılması mümkündür. Bu uygulamada arkeolojik yapılardan elde edilen verilerin işlenmesi için çok katmanlı hücresel yapay sinir ağı (ÇS-HYSA) kullanılmaktadır.

6.2.1. Hücresel Yapay Sinir Ağı'nın Sentetik Verilerle Eğitimi

Bu uygulamada aşağıda Tablo 6.4.'te verilen konum ve büyüklüğe sahip magnetik dipollerin, bilgisayar simülasyonu ile magnetik anomali haritaları



oluşturuldu (Şekil 6.6). Daha sonra bu magnetik anomali haritasından kaynakların yerini belirlemek için hücresel sinir ağı genetik algoritmalarla eğitildi.

Tablo 6.4. Aynı (x,y) koordinatlarında farklı derinliklerde bulunan iki dipolün özellikleri.

PARAMETRELER	DİPOL 1	DİPOL 2
(x,y) koordinatları	(32,32)	(32,32)
Z (Derinlik)	5	25
L (Boy)	8	30
α (açı)	90	90

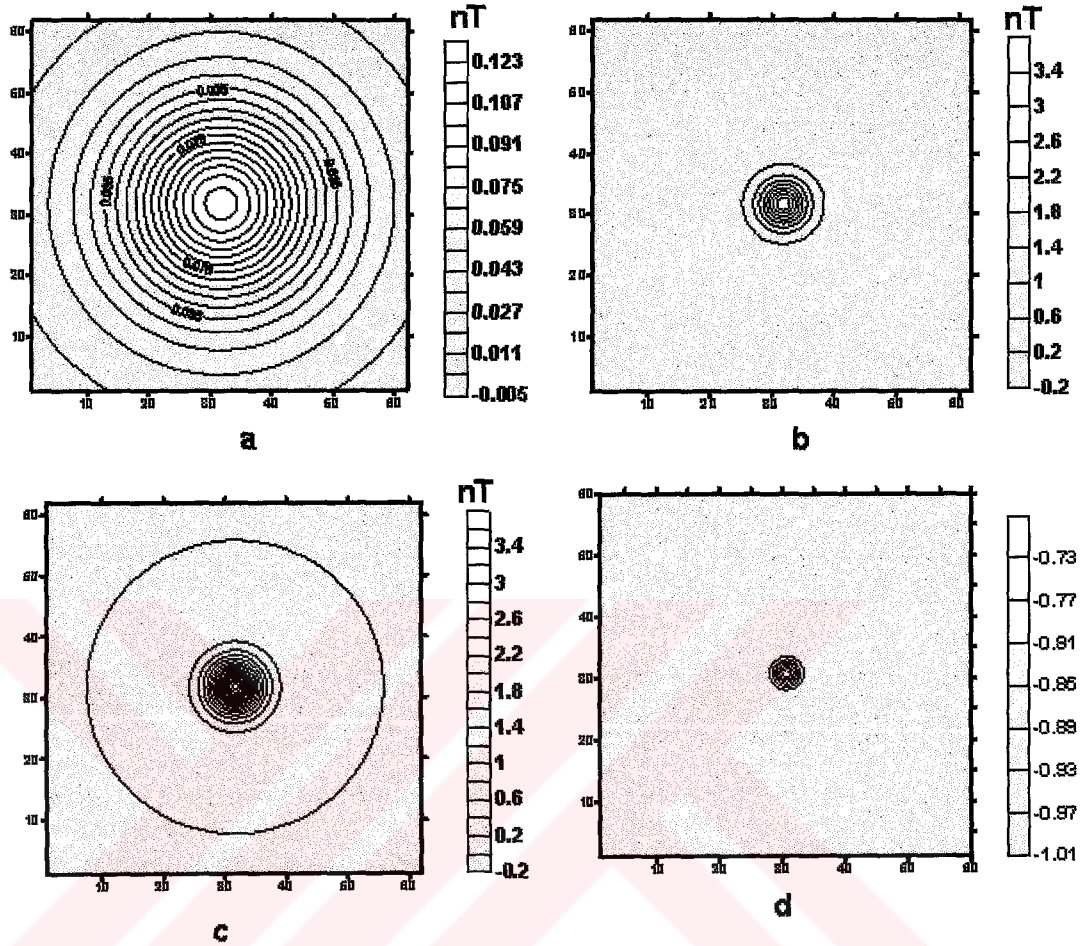
Genetik algoritmada kullanılan parametreler Tablo 6.5.'de verilmektedir.

Tablo 6.5. Magnetik verilerin eğitiminde kullanılan genetik algoritma parametreleri.

PARAMETRELER	DEĞER
Kromozom sayısı	20
Parametre başına bit sayısı	16
Şablon parametresi sayısı	5
Kromozom uzunluğu	80
Populasyondaki toplam bit sayısı	1600
Mutasyon olasılığı	1 %
Evrilen bit sayısı	16
Şablon parametrelerinin değişim aralığı	[-5, 5]



Bu uygulama için genetik algoritma aşağıdaki adımlar takip edilerek uygulanmıştır.



Şekil 6.6. Magnetik verilerin eğitimi için kullanılan ve özellikleri Tablo 6.4.'de verilen sentetik örnek : (a) Bölgesel magnetik anomali haritası (kontor aralığı 0.008 nT) , (b) Rezidüel anomali haritası(kontor aralığı 0.4 nT) (c) Toplam düşey magnetik anomali haritası (kontor aralığı 0.2 nT) (d) ÇS-HYSA çıkışı

Adım 1. Başlangıç popülasyonunun oluşturulması: Popülasyon matrisi adı verilen bir matris oluşturulur. Matrisin her bir satırı kromozomlara ait bilgileri içerir. 20 kromozom seçildiğinden popülasyon matrisinde 20 satır vardır. Her bir kromozom 80 bitlik (ikili) binary dizilerden oluştuğu için popülasyon matrisinde 80 sütun vardır. Başlangıçta popülasyon matrisi rastgele oluşturulur.

Adım 2. HYSA şablonlarının oluşturulması: Kromozomlar HYSA şablonlarının- A, B, I - ikili kodlarını temsil ederler. Bu adımda kromozomlarda kodlanmış olan şablon parametrelerinin ikili kodları çözülür ve şablon parametreleri

$[-5,5]$ aralığında gerçek değerleri ile oluşturulur. Her bir parametre 16 bit seçildiğinden, 5 şablon parametresi için 80 bitlik kromozom kullanılır. İlk 5 bit şablon elemanlarının 1. bitini, ikinci 5 bit şablon elemanlarının 2. bitlerini gösterecek şekilde kromozom aşağıdaki gibi kodlanmıştır.

$$S = [A_{1,1}, A_{2,2}, B_{1,1}, B_{2,2}, I] \quad (6.8)$$

Adım 3. Her bir kromozom için hata fonksiyonunun değerinin hesaplanması; Bu adımda, eğitimde kullanılan eğitim görüntüsü HYSA'a giriş olarak verilir. HYSA'ın durum değişkenleri de bu giriş görüntüsüne eşitlenir (Şekil 6.6.c). Adım 2.'de bulunan şablon değerleri kullanılarak HYSA çıkışı izlenir. HYSA çıkışında oluşan görüntü ile, olması istenilen hedef görüntü (Şekil 6.6.d) karşılaştırılır. Aradaki farka göre hata fonksiyonu (cost function) değeri hesaplanır. Bu işlem bütün kromozomlar için tekrarlanarak bulunan hata fonksiyonu değerleri saklanır. Hata fonksiyonu aşağıdaki gibi seçilmiştir.

$$\text{cost}(A, B, I) = \sum_i^m \sum_j^n P_{i,j} \oplus T_{i,j} \quad (6.9)$$

Burada A, B, I şablon elemanları, m ve n görüntünün piksel numaralarını, P ve T sırasıyla giriş görüntüsünü ve hedef görüntüyü göstermektedir. " $\oplus$ " sembolü P ve T . arasında yapılan XOR operasyonunu göstermektedir. Paha fonksiyonu değerleri bulunduktan sonra bu değerlere göre uygunluk (fitness) fonksiyonu değerleri aşağıdaki kural uyarınca yapılır.

$$\text{fitness}(A, B, I) = m * n - \text{cost}(A, B, I) \quad (6.10)$$

Başka bir kural da durma kriteri için belirlenir. Söz konusu durma kriteri aşağıda gösterilmiştir.

$$\text{stcriterion} = 0.99 * m * n \quad (6.11)$$

Burada m görüntünün satır, n ise sütun numarasını göstermektedir. Eğer maksimum uygunluk değeri durma eşiğinden büyükse algoritma durdurulur

uygunluk değeri en yüksek olan kromozom seçilir. Bu seçilen en iyi kromozomdan oluşturulan şablonlar, bu uygulamada istenilen hedefi gerçekleştirmek için kullanılacak en uygun şablonlardır.

Adım 4. Yeni Popülasyon Oluşturma : Bir sonra ki jenerasyonu oluşturmadan önce popülasyondaki uygunluk değerleri azalan şekilde sıralanır ve popülasyondaki bütün kromozomlara ait uygunluk değerleri toplamı alınarak normalize edilir. 0 ile 1 arasında rasgele bir r değeri seçilir. Sonra uygunluk değerleri normalize edilmiş popülasyon bireylerinden uygunluk değeri r' ye eşit ve veya daha büyük olan en yakın kromozom bir sonraki gen havuzuna eklenir. Bu işlem 20 kere tekrarlanır ve uygunluk değeri iyi olmayan kromozomlar popülasyondan silinir.

Kopyalama işleminin ardından 10 kromozom çifti bir sonraki jenerasyondaki kromozomlara ebeveyn olmak üzere rasgele seçilir. “s1” ve “s2” gibi 1 ile kromozom uzunluğu 80 arasında iki sayı seçilir. “s1” ve “s2” arasındaki bit dizisi çaprazlama bölgesi olarak adlandırılır. Çaprazlama işlemi sırasında, kromozom çiftlerinin “s1” ve “s2” arasındaki bit dizileri aralarında değiş tokuş edilir. Böylece önceki kromozomların parçalarından oluşan iki yeni kromozom oluşur. Bu uygulamada mutasyon olasılığı %1, seçildiğinden bütün popülasyonda toplam 16 bit rasgele seçilerek, bit değerinin evriği alınır. Kopyalama işleminden önceki en iyi kromozom olduğu gibi yeni popülasyona katılır. Popülasyon nüfusunu sabit tutmak için bu eklenen eski kromozoma karşılık son jenerasyondaki kromozomlardan birisi rasgele seçilerek popülasyondan silinir. Artık yeni popülasyon, yani yeni kuşak oluşturulmuş olur. Yeni kuşak oluşturulduktan sonra optimum sonucu aramak için 2. adıma dönülür ve algoritma durma kriterini gerçekleninceye kadar bu şekilde devam eder.

Eğitim işleminin sonucunda 108 jenerasyon sonunda istenilen sonuca ulaşıldı. Aşağıdaki şablonlar optimum şablon olarak bulundu.

$$A = \begin{bmatrix} 0.0020 & 0.0020 & 0.0020 \\ 0.0020 & 0.5015 & 0.0020 \\ 0.0020 & 0.0020 & 0.0020 \end{bmatrix} B = \begin{bmatrix} 1.0032 & 1.0032 & 1.0032 \\ 1.0032 & -3.7542 & 1.0032 \\ 1.0032 & 1.0032 & 1.0032 \end{bmatrix} I = -0.3718 \quad (6.12)$$

Bu şablona karşı gelen en iyi kromozom ise aşağıdaki gibidir.

BC=

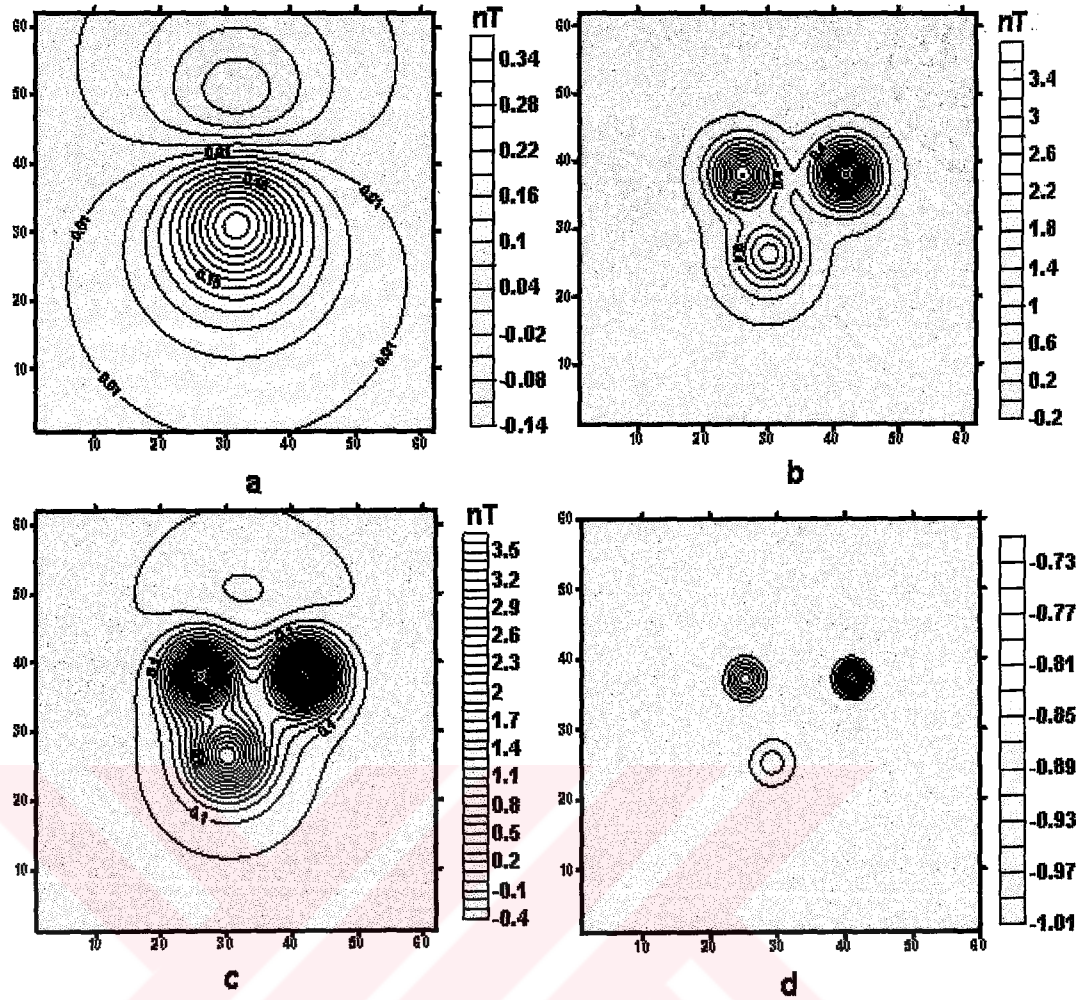
[101000001100001001010000100000000110001000010000100010010001111100
1110001001111] (6.13)

Bulunan bu şablon parametreleri kullanılarak Tablo 6.6.'de özellikleri verilen 4 sentetik dipolün oluşturduğu magnetik anomali haritasına HYSA uygulandı. Şekil 6.7.'de bu veriye ait giriş görüntüsü ve HYSA çıkışı görülmektedir.

Tablo 6.6. Sentetik dört magnetik dipole ait özellikler.

PARAMETRELER	(x,y) Koordinat	Z (derinlik)	L (boy)	α (açı)
Dipol 1	32,32	15	20	20
Dipol 2	45,40	5	10	90
Dipol 3	25,38	6	8	90
Dipol 4	30,25	8	8	90

Bir sonraki bölümde aynı şablon parametreleri kullanılarak HYSA Hitit kalıntılarına uygulanmaktadır.

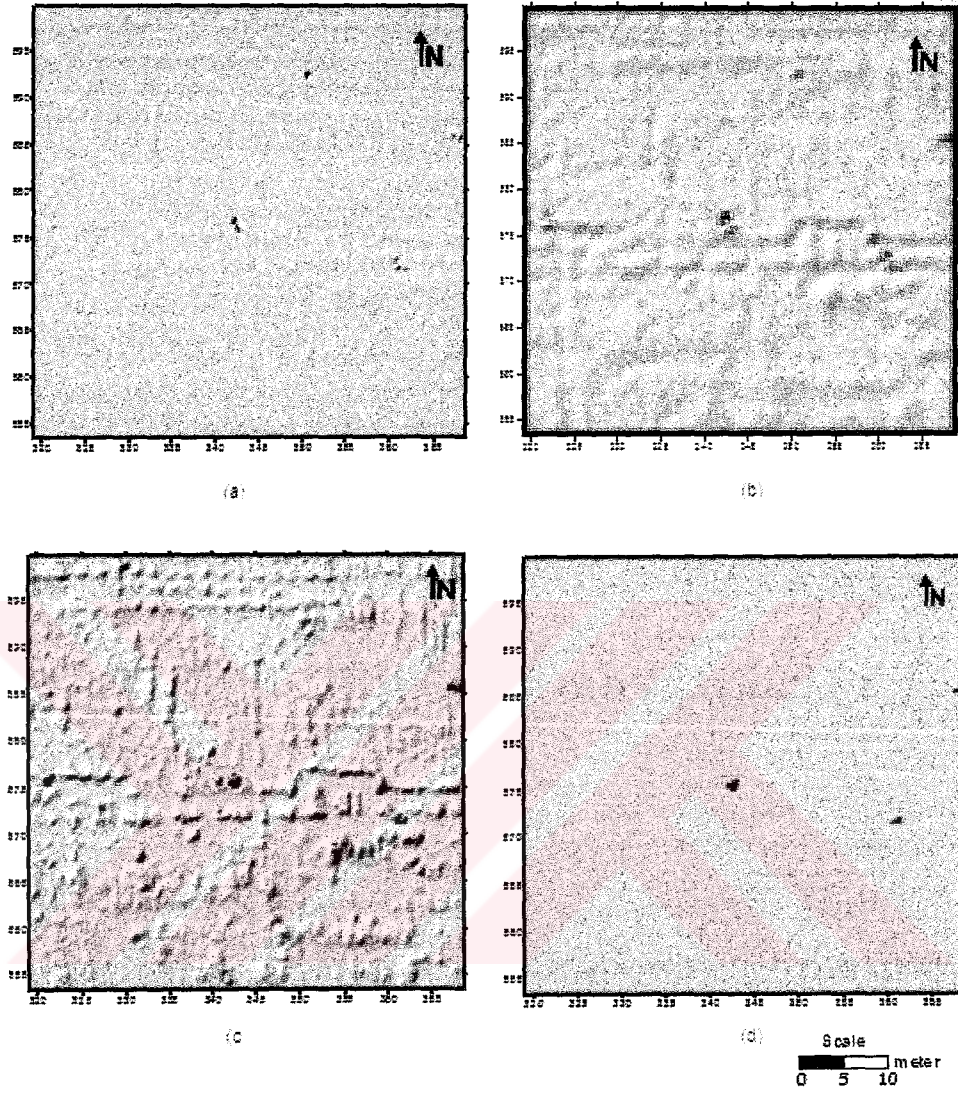


Şekil 6.7. Tablo.6.6’de özellikleri verilen 4 magnetik dipole ait anomali haritaları: (a) Bölgesel anomali (kontor aralığı .003 nT) (b) Rezidüel anomali (kontor aralığı .02 nT) (c) Toplam magnetik anomali (kontor aralığı 0.1nT) (d) ÇS-HYSA çıkışı.

6.2.2. HYSA’nın Hitit Kalıntılarına Uygulanması

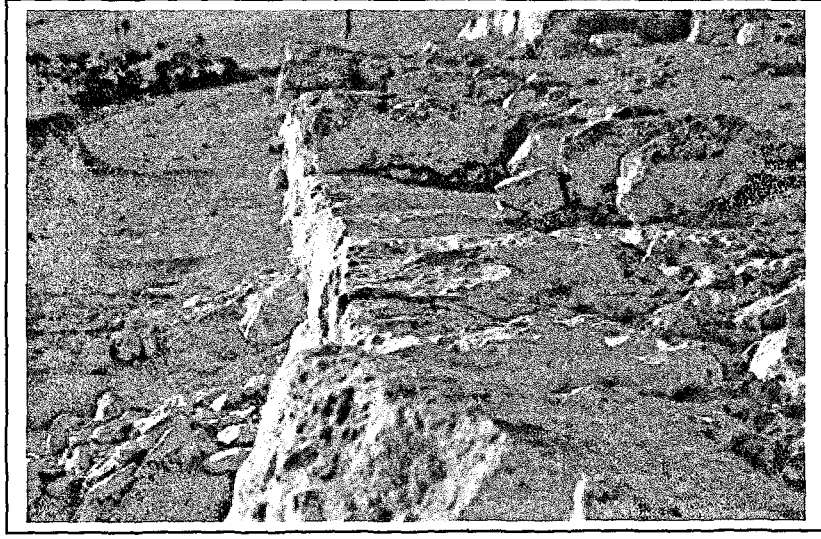
İnsanlık tarihinde önemli bir yer tutmasından dolayı bu uygulamada Hitit yerleşkeleri seçildi. Hitit yerleşkeleri Prof.Dr. Andreas Muller gibi arkeologlar ve Kiel Üniversitesinden Prof.Dr. Herald Stümpel gibi jeofizikçiler tarafından 1966’dan beri araştırılmaktadır [Stümpel, 1997,1998]. Bu araştırmalardan elde edilen gerçek saha verileri bu çalışmada değerlendirilmektedir. Hitit imparatorluğu kalıntılarına ait magnetik anomali haritası ve HYSA uygulandığında elde edilen çıkış görüntüsü Şekil 6.8.’ de gösterilmektedir. Klasik bir yöntem olan Halck yöntemiyle elde edilen Şekil 6.8.b. görüntüsü ve HYSA ile elde edilen Şekil 6.8.c karşılaştırıldığında HYSA görüntüsünde arkeolojik yapının odaları ve koridorları çok daha belirgin

olarak ortaya çıkmaktadır. Şekil 6.9'da bölgede yapılan kazılar sonucu çıkarılan duvar kalıntılarının fotoğrafı görülmektedir.



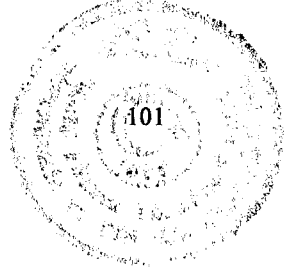
Şekil.6.8. Hitit İmparatorluğu kalıntılarına ait arkeolojik veriler (a). Sivas-Altınayla Bölgesindeki Hitit İmparatorluğu kalıntılarına ait magnetik anomali haritası. (b) Halck'ın ikinci türev yöntemiyle elde edilmiş çıkış görüntüsü (c) ÇS-HYSA çıkışı (d) Üç-katmanlı ÇS-HYSA çıkışı.

Bu sonuçlar HYSA'nın arkeolojik yapıların haritalarının çıkarılmasında çok faydalı olacağını göstermekte ve cesaret vermektedir.



Şekil 6.9. Arkeolojik çalışmalar sonucunda ortaya çıkartılmış Hitit kalıntılarının görüntüleri.



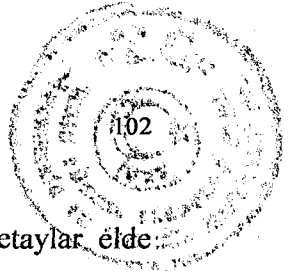


7. SONUÇ

Bu çalışmada, HYSA, Trapezoidal Aktivasyon Fonksiyonu (TAF) adı verilen yeni bir nonlinear aktivasyon fonksiyonu ile çalıştırılarak Trapezoidal Aktivasyon Fonksiyonlu Hücresel Yapay Sinir Ağı (TAF-HYSA) tanıtılmıştır. TAF-HYSA, tek katmanlı perceptron ağı için Aksin ve arkadaşları tarafından donanım olarak gerçekleştirilen çift eşikli aktivasyon fonksiyonunun genelleştirilmiş bir uygulamasıdır. Geliştirilen TAF-HYSA için iki boyutlu durum denklemleri tek boyutlu hale dönüştürülerek ve Lyapunov kararlılık yöntemi uygulanarak, A şablonu elemanları ve TAF parametrelerine bağlı olan bir kararlılık kriteri bulunmuştur. Bu kararlılık kriteri aynı zamanda standart PWL aktivasyon fonksiyonu kullanan standart CNN'e uyarlanmıştır. Teorik sonuçlar çok bilinen standart giriş görüntüleri kullanılarak pratik olarak desteklenmiştir. Kararlılık kriteri HYSA şablonlarının simetrik olma kısıtlamasını gerektirmediğinden, özellikle örüntü tanıma ve sınıflandırma işlemleri için simetrik olmayan şablon tasarımına imkan tanıyarak çok büyük avantaj sağlamaktadır. HYSA-TAF'ın bir diğer üstünlüğü, aktivasyon fonksiyonu parametrelerinin şablon elemanları ile birlikte veya bağımsız olarak ayarlanara büyük bir tasarım kolaylığı sağlamasıdır.

Ayrıca HYSA-TAF'ın, tek katmanlı HYSA ile gerçekleştirilemeyen XOR ve parite problemleri gibi lineer olarak ayrıştırılamayan verilerin ayrıştırmasını tek katmanda yaparak önemli bir üstünlük sağladığı çok sayıda örnek uygulamayla gösterilmiştir.

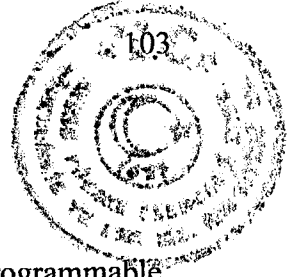
Tez çalışmasında ayrıca HYSA genetik algoritmalarla eğitilerek gravite ve magnetik anomalilerin ayrıştırma işlemleri yapılmıştır. HYSA eğitimi için sentetik veriler üretilip, bu verilerle HYSA şablonları elde edilmiş ve bu şablonlar başka sentetik verilere ve gerçek verilere uygulanmıştır. Ülkemizdeki Divriği-Dumluca bölgesine ait anomali verileri kullanılarak bölgedeki maden yataklarının yerleri tespit edilmiş ve Sivas-Altınyayla bölgesindeki Hitit İmparatorluğu dönemine ait arkeolojik kalıntılara ait



anomali verileri kullanılarak, yer altındaki yapının geometrisine ait detaylar elde edilmiştir. Bulunan HYSa sonuçları klasik yöntemlerle elde edilen sonuçlarla karşılaştırıldığında, HYSa'nın anomali ayırtırmadaki bariz üstünlüğü gözlenmiştir.

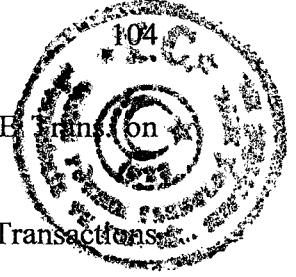
Tez içerisinde jeofizik verilerin iki boyutlu analizleri yapılmıştır. Bundan sonraki çalışmalarda uygun trapezoidal aktivasyonu parametreleri bulunarak magnetik ve gravite anomali haritalarında üçüncü boyut bilgisi elde etmek amacıyla farklı derinliklerde bulunan katmanlara ait anomali ayırtırma uygulamaları yapılabilir.





KAYNAKLAR

- [1] Aksin ,Y.D., Aras, S., Goknar IC. "CMOS Realization of User Programmable, Single-Level, Double-Threshold Generalized Perceptron.", Proceedings of Turkish Artificial Intelligence and Neural Networks Conference, TAINN-2000, Izmir, Turkey, June 2000
- [2] Aksin,Y.D., Aras S, Goknar IC. "Realization and Extensions of User Programmable, Single- Level, Double-Threshold Generalized Perceptron" , Istanbul University Journal of Electrical&Electronics Engineering (IU-JEEE); v.1(1): s.123-129. 2001
- [3] Altıparmak F., Dengiz B. ve Smith A.E. (), "An Evolutionary Approach For Reliability Optimization in Fixed Topology Computer Networks",Transactions On Operational Research, Volume: 12, Number: 1-2, pp. 57-75, 2000.
- [4] Anderson, W. L. "Computer Program Numerical Integration of Related Hankel Transforms of Orders 0 and 1 by Adaptive Digital Filtering," Geophysics, vol. 44, no. 7, pp. 1287-1305, 1979.
- [5] Arik ,S. and Tavsanoğlu, V. "Absolute stability of nonsymmetric neural networks," Proc. IEEE Int. Symp. Circuits and Systems, ISCAS'96, pp. 441-444, May 1996.
- [6] Arik ,S. and Tavsanoğlu, V. "On the global asymptotic stability of delayed cellular neural networks," IEEE Trans. Circuits Syst. I, vol. 47,pp. 571-574, Apr. 2000.
- [7] Arik ,S. "An Improved Global Stability Result for Delayed Cellular Neural Networks" , IEEE Trans. Circuits Syst. I, vol. 49:8, pp. 1221-1214, Apr. 2002.
- [8] Bhattacharyya, B.K. "Two-Dimensional Harmonic Analysis as a tool for Magnetic Interpretation" Geophysics, vol. 30, no. 5, pp.829-857, Oct. 1965.
- [9] Blakely, R.J. and Simpson, R.W. "Approximating Edges of Source Bodies from Magnetic or Gravity Anomalies," Geophysics, vol. 51, no. 7, pp. 1494-1498, 1986.
- [10] Boschetti, F. , Hornby, P. and Horowitz, F.G. "Wavelet Based Inversion of Gravity Data", Exploration Geophysics, vol. 32, pp. 48-55, 2001.
- [11] Bryson A.E., Ho Y.C. "Applied optimal control", Blaisdell, New York, 1969.
- [12] Chen CW, Chen L. "Cellular Neural Network Architecture for Gibbs Random Field Based Image Segmentation". Journal of Electronic Imaging; 7(1):45-51, 1998.



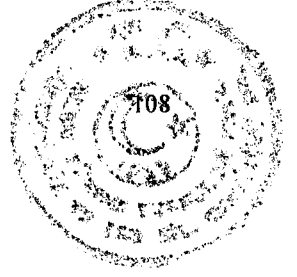
- [13] Chua, L.O., Yang, L. "Cellular Neural Networks: Theory," IEEE Trans. on Circuit and Systems, vol.35, pp. 1257-1272, 1988.
- [14] Chua LO, Yang L. Cellular Neural Networks: Applications. IEEE Transactions on Circuit and Systems 1988; **35**(10): 1273-1288.
- [15]. Chua L, Roska T, Cellular Neural Networks and Visual Computing, Cambridge University Press, Cambridge, 2002.
- [16] Cichocki A. "Neural network for optimization and signal processing", New York: JohnWiley, 1993.
- [17] Civalleri PP, Gilli M. Some Dynamic Phenomena in Delayed Cellular Neural Networks. International Journal of Circuit Theory and Applications 1994; **22**:77-105.
- [18] Civalleri PP, Gilli M. On the Dynamic Behavior of Two-Cell Cellular Neural Networks. International Journal of Circuit Theory and Applications, 1993; **21**: 451-471.
- [19] Clement, W. G. "Basic Principles of Two-Dimensional Digital Filtering", Geophysical Prospecting, vol. 21, 1973, pp. 125-145.
- [20] Davis, L. "Handbook of Genetic Algorithms", Van Nostrand Reinhold, New York, 1991.
- [21] Dean, W.C. "Frequency Analysis for Gravity and Magnetic Interpretation", Geophysics, vol. 23, 1958, pp. 97-127.
- [22] Emel, G.K. ve Taşkın Ç., "Genetik Algoritmalar ve Uygulama Alanları", Uludağ Üni., İkt. ve İdari Bil. Fak. Dergisi, Cilt:21, Sayı:1, s.129-152, 2002.
- [23] Engin O., "Akış Tipi Çizelgeleme Problemlerinin Genetik Algoritma ile Çözüm Performansının Arttırılmasında Parametre Optimizasyonu", İTÜ, Fen Bilimleri Enstitüsü, Yayınlanmamış Doktora Tezi, 2001.
- [24] Fedi , M. and Quarta, T. "Wavelet Analysis for the Regional Residual and Local Separation of Potential Field Anomalies," Geophysical Prospecting, vol. 46, 1988, pp.507-525.
- [25] Figlali A. ve Engin O., "Genetik Algoritmalarla Akış Tipi Çizelgelemede Üreme Yöntemi Optimizasyonu", İTÜ Dergisi, s. 1-6. ,2002.
- [26] Fuller, D. "Two-Dimensional Frequency Analysis and Design of Grid Operators," Mining Geophysics, vol. 2, 1967, pp. 658-708.
- [27] Gilli, M. "Stability of cellular neural networks and delayed cellular neural networks with nonpositive templates and nonmonotonic output functions," IEEE Trans. Circuits Syst. I, vol. 41, pp. 518-528, Aug.1994.



- [28] Goldberg D.E., "Genetic Algorithms in Search, Optimization and Machine Learning", Addison-Wesley, USA, 1989.
- [29] Güzeliş C. "Genelleştirilmiş hücresel yapay sinir ağları", EMO 5. Ulusal Kongresi, Trabzon, pp:7-12,1993.
- [30] Güzelis C, Chua LO. "Stability Analysis of Generalized Cellular Neural Networks". International Journal of Circuit Theory and App. 1993;21(1):1-33.
- [31] Hansen, R.O. and Pawlowski, R.S. "Reduction to the Pole at Low Latitudes by Wiener Filtering," Geophysics, vol. 54, 1989, pp. 1607-1613.
- [32] Hansen, R.O. and Simmonds, M. "Multiple-Source Werner Deconvolution," Geophysics, vol. 58, 1993, pp. 1792-1800.
- [33] Hille E. "Methods in Classical and Functional Analysis", Adisson Wesley; US; 1972
- [34] Holden, D., Archibald, N., Boschetti, F. and Jessell, M. "Inferring Geological Structures Using Wavelet-Based Multiscale Edge Analysis and Forward Models," Exploration Geophysics, vol. 31, 2000, pp. 617-621.
- [35] Holland, J.H. "Adaptation in Neural and Artificial Systems", University of the Michigan Press, Ann Arbor, MI, 1975.
- [36] Hornby, P. , Boschetti, F. , and Horowitz, F.G. "Analysis of Potential Field Data in the Wavelet Domain," Geophysical Journal International, vol. 137, 1999, pp. 175-196.
- [37] Karamahmut S. "Hücresel yapay sinir ağları için iki öğrenme algoritması ve görüntü işleme uygulamaları", İTÜ Merkez Kitaplığı, no: 39003, 1994.
- [38] Kozek, T., Roska, T. and Chua, L.O. "Genetic Algorithms for CNN Template Learning", IEEE Trans. on Circuit and Systems, vol. 40, no. 6, 1988, pp. 392-402.
- [39] Larsen R. "Functional Analysis"; Marcel Dekker Inc: New York, 1973.
- [40] Lavin, P.M. and. Devane, J. F. "Direct Design of Two-Dimensional Digital Wavenumber Filters," Geophysics, vol. 35, 1970, pp.1073-1078.
- [41] Lee G.C.S., Lin C.T. "Neural fuzzy systems", Prentice Hall, New Jersey, 1995.
- [42] LeeCun Y. "A learning procedure for assymetric network", Cognitiva 85: 599-604, Paris, 1985.
- [43] Liao X, Yu J, Chen G. "Novel Stability Criteria for Bidirectional Associative Memory Neural Networks with Time Delays", International Journal of Circuit Theory and Applications 2002; 30: 519-546.

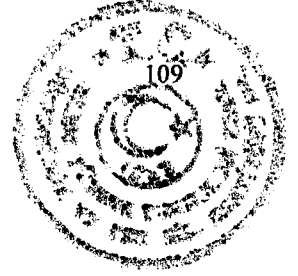
- [44] Magnussen, H. and Nossek, J.A. "Towards a Learning Algorithm for Discrete Time Cellular Neural Networks", In Proc. IEEE 2nd International Workshop on Cellular Neural Networks, pp.80-85, 1992.
- [45] Majorana S, Chua LO. "A Unified Framework for Multilayer High Order CNN", International Journal of Circuit Theory and Applications 1998; 26(6): 567-592.
- [46] Ölmez T. "Yapay sinir ağları yardımıyla biyomedikal dokuların sınıflandırılması", İTÜ Merkez Kitaplığı no:44082, 1995.
- [47] Parker D.B. "Learning logic", Technial Report, no:TR-47, Center for Comp. Research in Economics and Management Science., MIT, 1985.
- [48] Ridsdill, T.A., "Smith and Dentith, M.C. "Wavelet Transform in Aeromagnetic Processing," Geophysics, vol. 64, no. 4, 1999, pp. 1003-1013.
- [49] Rumelhart D.E., Hinton G.E., Willias R.J. "Learning represantions by back propagation errors", Nature 323:533-536, 1986(a).
- [50] Rumelhart D.E., Hinton G.E., Willias R.J. "Learning internal represations by error propagation", Paralell distrubuted processing Volume 1: Chapter 8, Cambridge, MA, 1986(b).
- [51] Schuler, A.J., Nachbar, P., Nossek, J.A. and Chua, L.O. "Learning State Space Trajectories in Cellular Neural Networks", In Proc. IEEE 2nd International Workshop on Cellular Neural Networks, pp.68-73, 1992.
- [52] Stümpel, H. "Untersuchungen in Kusakli: Geophysikalische Prospektion," Mitteilungen der Deutschen Orient-Gesellschaft, Berlin, vol. 129, 1997, pp. 134-140.
- [53] Stümpel, H. "Untersuchungen in Kusakli: Geophysikalische Prospektion", Mitteilungen der Deutschen Orient-Gesellschaft, Berlin, vol. 130, 1998, pp. 144-153.
- [54] Sziranyi T, Csapodi M. "Texture Classification and Segmentation by Cellular Neural Network using Genetic Learning". Research Report, Analogical and Neural Computing Systems Research Lab., Computer and Automation Inst. Hungarian Academy of Sciences (MTA-SZTAKI), Budapest, Hungary, Nov.1, 1994.
- [55] Ucan, O.N., Sen, B., Albora, A.M. and Özmen, A. "A New Gravity Anomaly Separation Approach: Differential Markov Random Field (DMRF)," Electronic Geosciences, vol. 5, no. 1, 2000
- [56] Ucan, O.N., Şeker, S., Albora, A. M. and Özmen, A. "Separation of Magnetic Field Data using 2-D Wavelet approach," Journal of The Balkan Geophysical Society, vol. 3, 2001, pp. 53-58.

- [57] Ucan, O.N., Albora, A.M. "İşaret ve Görüntü İşlemede Yeni Yaklaşımlar", İ.U. Mühendislik Fak. Yayınları, İstanbul, 2003.
- [58] Vidyasagar M. "Nonlinear System Analysis"; Electrical Engineering Series, Prentice Hall: New Jersey, 1978.
- [59] Wang L, Pineda de Gyvez J, Sinencio ES. "Time Multiplexed Color Image Processing Based on a CNN with Cell-State Outputs". IEEE Trans. of VLSI Sytems 1998; 6(2): 314-322.
- [60] Yıldız, N. "Drilling Report of Divriği-Dumluca Iron Ore," MTA Report, no. 315, Ankara, 1977.
- [61] Zhang J. "Absolutely Exponential Stability in Delayed Cellular Neural Networks." International Journal of Circuit Theory and Applications 2002; 30: 395-409.
- [62] Zhou D.M, Zhang L.M. "Global Exponential Stability of Cellular Neural Networks". ICONIP 2001, 8. Proceedings of International Conference on Neural Information Processing, Shangai, China, November 14 -18, 2001.
- [63] Zurada J.M. "Introduction to artifical neural systems", West Publishing Co., St. Paul,1991.
- [64] Zurflueh E.G. "Applications of Two-Dimensional Linear Wave-Length Filtering," Geophysics, vol.32, no. 6, 1967, pp. 1015-1035.



ÖZGEÇMİŞ

Erdem BİLGİLİ, 1972 yılında Şenkaya'da doğdu. İlk,orta ve lise öğrenimini Erzurum'da tamamladı. 1996 yılında Yıldız Teknik Üniversitesi Elektronik ve Haberleşme Mühendisliği Bölümünden mezun oldu. Aynı yıl Atatürk Üniversitesi Tarım Makinaları Bölümünde araştırma görevlisi olarak çalışmaya başladı. 1996 yılı güz yarısında Gebze Yüksek Teknoloji Enstitüsünde yüksek lisans öğrenimine başladı. 1998 yılında Gebze Yüksek Teknoloji Enstitüsünde araştırma görevliliğine atandı.1999 yılında “Yapay Sinir Ağları ile Bitkilerin Şekillerinin Sınıflandırılması” adlı yüksek lisans tezini tamamladı. 2000-2004 yılları arasında TÜBİTAK Marmara Araştırma Merkezinde yarı zamanlı araştırmacı olarak çalıştı. 2004 yılından itibaren GYTE’ deki araştırma görevliliği görevinden ayrılarak TÜBİTAK Marmara Araştırma Merkezinde tam zamanlı araştırmacı olarak çalışmaktadır.



EKLER

BİLGİSAYAR PROGRAMLARI

%%TRAPEZOİD AKTİVASYON FONKSİYONLU CNN-GA ile EĞİTİMİ %%%

```
%%[A B I]=gncnneg(ornekler,sboyut,gennuf,bit)
%%Gk:desen türleri(parlaklık ortalamaları)sütun matris;
%%test resmi
%%sboyut:A ve B şablonları boyut(3,4,5 gibi)
%%gennuf:toplam kromozom sayısı
%%parametre başına bit sayısı
%%n parametre sayısı
function [A, B, I,ghavuz,gdeg,uygunluk,Gk2]=myenigncnneggr1k2(Gk,test,gennuf,n,bit,ekpop)

ghavuz=mhavuz(n,gennuf,bit);
ghavuz=[ghavuz ;
        ekpop];

[erd lem]=size(ghavuz);
gennuf=erd;
%display('gendeg');
genlik=2;                                %maksimum genlik değeri
gdeg=mgendeggr(ghavuz,n,genlik); %gecti
[h r]=size(Gk);
Gk2=Gk;
usinir=0.01;                             % gecti
for i=1:gennuf
    [A B I TTes]=msablonlar(gdeg(i,:),genlik); %gecti
    es1=TTes(1);
    esh=TTes(2);
    es2=TTes(3);
    esl=TTes(4);
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%kararlılık kriteri%%%%%%%%%%5
    f1=esh-es1;
    f2=esl-es2;
    dff=min(f1,f2)/2.1;
    mutlaktopa=sum(abs(A(:)));
    A=A*dff/mutlaktopa
    B=B
    I=I
    TTes=TTes

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
cik=mcnnseg(test,A,B,I,es1,esh,es2,esl); %gecti
yer=cik;
imshow(yer)
[uyg hata1]=mkpahal1(Gk2,yer);           %gecti
uygunluk(1,i)=uyg;
hata(1,i)=hata1;
i
```



```
end
figure(7)
plot(1:gennuf,uygunluk)
title('uygunluk');

nes=1
uygmax=max(uygunluk);
topuyg(nes)=sum(uygunluk);
figure(8)
plot(topuyg)
title('toplamuygunluk');
hold on
newplot;
figure(9)
plot(nes,uygmax)
title('maxuygunluk');
hold on
newplot;
while min(hata)>usinir
    nes=nes+1
    ghavuz=mgkopya306(uygunluk,ghavuz); %gecti
    [ghavuz iyiler]=mcaprazlama(ghavuz,nes); %gecti
    ghavuz=mmutasyon(ghavuz,nes); %gecti
    ghavuz(gennuf-1:gennuf,:)=iyiler; %gecti
    gdeg=mgendeggr(ghavuz,n,genlik); %gecti
    %size(gdeg)
    %pause
    figure(7)
    plot(1:gennuf,uygunluk)
    uygmax=max(uygunluk);
    for i=1:gennuf

        [A B I TTes]=msablonlar(gdeg(i,:),genlik); %gecti
        es1=TTes(1);
        esh=TTes(2);
        es2=TTes(3);
        esl=TTes(4);
        %%%%%%%%%%%%%%%kararlılık kriteri%%%%%%%%%%5
        f1=esh-es1;
        f2=esl-es2;
        dff=min(f1,f2)/2.1;
        mutlaktopa=sum(abs(A(:)));
        A=A*dff/mutlaktopa
        B=B
        I=I
        TTes=TTes
        %%%%%%%%%%%%%%%

        cik=mcnnseg(test,A,B,I,es1,esh,es2,esl); %gecti
        yer=cik;
        [uyg hata1]=mkpahal1(Gk2,yer);
        uygunluk(1,i)=uyg;
        hata(1,i)=hata1;
        i
    end
end
```

```

[h bul]=find(uygunluk==max(uygunluk));
[A B I TTes]=msablonlar(gdeg(bul(1),:),genlik);           %gecti
    es1=TTes(1);
    esh=TTes(2);
    es2=TTes(3);
    esl=TTes(4);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%kararlilik kriteri%%%%%%%%%5
    f1=esh-es1;
    f2=esl-es2;
    dff=min(f1,f2)/2.1;
    mutlaktopa=sum(abs(A(:)));
    A=A*dff/mutlaktopa
    B=B
    I=I
    TTes=TTes
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

save 'd:\matlabR12\work\erdem\mcnnccikiskkenar6',A,B,I,TTes,gdeg,ghavuz,uygunluk,bul;
cik=mcnnseg(test,A,B,I,es1,esh,es2,esl);                 %gecti
figure(6)
imshow(cik);
topuyg=sum(uygunluk);
figure(8)
plot(nes,topuyg)
hold on
newplot;
figure(9)
plot(nes,uygmax)
hold on
newplot;

end

[h bul]=find(uygunluk==max(uygunluk));
[A B I TTes]=msablonlar(gdeg(bul(1),:),genlik);           %gecti
    es1=TTes(1);
    esh=TTes(2);
    es2=TTes(3);
    esl=TTes(4);

    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%kararlilik kriteri%%%%%%%%%5
    f1=esh-es1;
    f2=esl-es2;
    dff=min(f1,f2)/2.1;
    mutlaktopa=sum(abs(A(:)));
    A=A*dff/mutlaktopa
    I=I
    TTes=TTes
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

cik=mcnnseg(test,A,B,I,es1,esh,es2,esl);                 %gecti

```

%%%%%%%%%%GEN HAVUZU OLUŞTURMA%%%%%%%%%

```

%%%GENETİK havuz
%%havuz=havuz(p,b)
%%p:kromozom sayısı (nüfus)
%%n: parametre sayısı şablon katsayıları+aktivasyon fonksiyonun eşik değerleri 4adet
%%b: bit sayısı/kromozom parametre başına bit sayısı
%%havuz=[paramatre1(b bit),param2(b bit)... 1.kromozom
%%      paramatre1(b bit),param2(b bit)... 2.kromozom
function havuz=mhavuz(n,p,b)
b=n*b;
havuz=round(rand(p,b));

```

%%%%%%%%%%

%%%%%%%%%%KROMOZOMN KOPYALAMA%%%%%%%%%

```

function gkop=mgkopya(uygunluk,ghavuz)
[h n]=size(uygunluk)
[p r]=sort(uygunluk);

%p küçükten büyüğe doğru uygunluk değerleri
%r ise indeksleri veriyor
toplamuygunk=sum(uygunluk);
ortuyg=toplamuygunk/n;
[k l]=size(ghavuz);
%ghavuz(r(1),:)=round(rand(1,l));
%ghavuz(r(2),:)=round(rand(1,l));
tr=(n*p/toplamuygunk);           % 2*p/ortuyg ile aynı
tr=sort(tr)

%tr
[s1 s2]=find(tr>=1);
[s3 s4]=size(s1);
k1=0;
inx=0;
for i=1:s3
    for k1=1:round(tr(n-i+1))
        inx=inx+1;
        ghavuz1(inx,1:l)=ghavuz(r(n-inx+1),1:l);
    end

    %k1=k1+tr(n-i+1);
end

for kontenjan=1:k-inx;
    ghavuz1(kontenjan+inx,:)=ghavuz(r(k-kontenjan+1),:);

```

```

end

%for kontenjan=inx+1:k;
%;
%ghavuz1(inx+1:k,:)=ghavuz(inx:k-1,:);

%ghavuz1(k-1,:)=ghavuz(r(n),:);
%ghavuz1(k,:)=ghavuz(r(n-1),:);

%ghavuz1(k+1:k,:)=round(rand(k-k1,l));
gkop=ghavuz1;
%inx=inx
%%%%%%%%%%%%%%yukarı ilave aşağı iptal

%top(1,1)=uygunluk(1,1);

%for i=2:n
% top(1,i)=top(1,i-1)+uygunluk(1,i);
%end
%top=top/top(1,n);
%for j=1:n
%at=rand;
%i=0;
%fark=-1;
%while fark<=0
% i=i+1;
% fark=top(1,i)-at;
%end
%gkop(j,:)=ghavuz(i,:);
%end
%%%%%%%%%%%%%%

%%%%%%%%%%%%%%ÇAPRAZLAMA (CROSS-OVER) %%%%%%%%%%%%%%%

function [ynesil, iyiler]=mcaprazlama(ghavuz)
[m n]=size(ghavuz);
sira=randperm(m);
sira;
for i=1:m/2
p1=sira(i); %ilk bireylerin numaraları
p2=sira(m/2+i); %ikincil bireylerin numaraları
baba=ghavuz(p1,:);
ana=ghavuz(p2,:);
%buraya bak!!!!!!!!!!!!!!
ky=fix((n-1)*rand(1,2))+1; %simetrik için 2 ye bölündü
ky=sort(ky);
if ky(1)==ky(2);
ky(2)=fix((n-1)*rand)+1;
ky=sort(ky);
end

ogul1=[baba(1:ky(1)) ana(ky(1)+1:ky(2)) baba(ky(2)+1:n)]; % ana(1,ky(3):n)];
ogul2=[ana(1:ky(1)) baba(ky(1)+1:ky(2)) ana(ky(2)+1:n)]; % baba(1,ky(3):n)];

```

```

% ogul1=[baba(1,1:ky(1)) ana(1,ky(1)+1:ky(2)) baba(1,ky(2)+1:n)];% ana(1,ky(3):n)];
% ogul2=[ana(1,1:ky(1)) baba(1,ky(1)+1:ky(2)) ana(1,ky(2)+1:n)];% baba(1,ky(3):n)];
%if sum(ogul1)==0;
    % ogul1=round(rand(1,n));
%end
% if sum(ogul2)==0;
    % ogul2=round(rand(1,n));
%end
yenesil(i,:)=ogul1;
yenesil(i+m/2,:)=ogul2;
end
%yenesil(m,:)=ghavuz(1,:);
%yenesil(m-1,:)=ghavuz(2,:);
%yenesil(m-2,:)=ghavuz(3,:);
iyiler(1:2,:)=ghavuz(1:2,:);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

```

%%MUTASYON %%%

```

function ghavuz=mmutasyon(ghavuz,nes)
[m,n]=size(ghavuz);
%pm=0.10*exp(-50./nes);
pm=0.001*exp(-nes/50);
msay=round(m*n*pm);
if msay<1
    msay=5
end
indeks=(m-1)*n*rand(1,msay);
sat=[fix(indeks/n)+1;
    fix(mod(indeks,n))+1];
for i=1:msay
    ghavuz(sat(1,i),sat(2,i))=~ghavuz(sat(1,i),sat(2,i));
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

```

%%FITNESS VE COST HESAPLAMA%%

```

%%%%%%%%Gk:ortalama parlaklık deęerleri
%m:satır
%n:sütun
%k:desen sayısı
%g:yerel ortalama parlaklık resmi
%Gk:desenlerin parlaklık ortalamaları satır vektörü
%[uygunluk paha]=kpaha1(Gk,g)
function [uygunluk,fark]=mkpaha11(hedef,gir)
%format long;
[m,n]=size(gir);
%gir=1-gir; %sonradan eklendi 1, 0 uyumsuzluğu
%%%%%%%%%%%%%son=gir-hedef;
%%%%%%%%%%%%%dist=20*(hedef+.05).*son;
%snk=min(uygunluk);
%uygunluk=min(uygunluk)+uygunluk;

```

```

son=gir-hedef;
dist=son;
fark=sum(sum(dist.^2))
fark1=(1.01).^(fark+1);
uygunluk=1000./(fark1+1)
toplamlam=sum(sum(gir));
if uygunluk<0
    uygunluk=0;
end
if toplamlam/(m*n)<=.01
    uygunluk=0;
end
%uygunluk=10.^(10*uygunluk);

```

```

%if sum(sum(gir))>.35*m*n
%    uygunluk=-1;
%end

```

```

%eski kısım iptal edildi
%alan=m*n;
%r=hedef-gir;
%fark=sum(sum(abs(r)));
%uygunluk=alan-1.5*fark
%if sum(sum(gir))>alan*.99
%    uygunluk=-100;
%end

```

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

```

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%GRAY KODU%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

```

```

%graykodunda verilen bir ikili dizinin onluk karşılığını verir

```

```

%ikili dizinin uzunluğu çift sayı olmalıdır

```

```

%KONTROL EDİLEEK

```

```

function onluk=graycode(kod)

```

```

onluk=0;

```

```

deg=0;

```

```

deg1=0;

```

```

uzunluk=length(kod)/2;

```

```

for i=1:uzunluk

```

```

    kod1=kod(2*i-1:2*i);

```

```

    deg1=mod(deg+deg1,2);

```

```

    if kod1==[0 0]

```

```

        deg=0;

```

```

    elseif kod1==[0 1]

```

```

        deg=1;

```

```

    elseif kod1==[1 1]

```

```

        deg=2;

```

```

    elseif kod1==[1 0]

```

```

        deg=3;

```

```

    end

```

```

    switch deg1

```

```

        case 0,degs=deg;

```

```

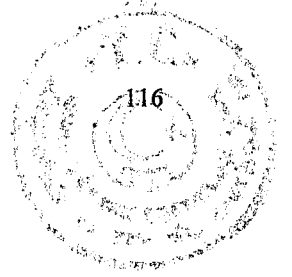
        case 1,degs=3-deg;

```

```

        %case 2,degs=deg;

```



```
%case 3,degs=3-deg;
end
onluk=onluk+degs*4^(uzunluk-i);
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

%%ŞABLON OLUŞTURMA%%

```
%gdeg: parametre değerleri
%%k: komşuluk boyutu(şablonların boyutu)n=1 için 3;
function [A, B, I, TTes]=msablonlar(gdeg,genlik)
[m,n1]=size(gdeg);
n=n1-4; %4 tane eşik parametresi var
s1=(n-1)/2; % s1 tane A ve B şablonun elemanı var
s=sqrt(s1); % şablonların boyutu s*s
for i=1:s
    for c=1:s
        A(i,c)=gdeg(c+(i-1)*s);
        B(i,c)=gdeg(s1+c+(i-1)*s);
    end
end
I=gdeg(1,n-4);

%*****Eşik parametreleri hesaplanıyor*****
Tes=gdeg(n-3:n);
es1=Tes(1)/genlik;
gt1=(1-es1)/2;
esh=es1+gt1+Tes(2)*gt1/genlik;
gt2=(1-es1)/2;
es2=esh+gt2+Tes(3)*gt2/genlik;
gt3=(1-es2)/2;
esL=es2+gt3+Tes(4)*gt3/genlik;
TTes=[es1 esh es2 esL];
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

%%GENLERİN OKUNMASI%%

```
%gendeg=gendeg(genhav,pmtsayisi)
%%genhav:genhavuzu
%%pmtsayisi:bir gendeki parametre sayısı
%%gendeg:parametrelerin değerleri n.satır n. gene ait
%% parametreler

function gendegs=mgendeggr(genhav,pmtsayisi,genlik)
pmtsayisi=fix(pmtsayisi); %parametre sayısı
[m n]=size(genhav);
b=fix(n/pmtsayisi);
```



```
for i=1:m %m adet kromozom
    for j=1:pmtsayisi
        gen=genhav(i,j); %parametre sayısı
        for k=1:b-1 %b bitlik parametreye ait bitlerin toplanması
            inds=j+k*pmtsayisi; % j. parametrenin 1. biti asıl genin J.biti, 2. biti ise j+pmtsayisi
            indisindeki bit
            gen=[gen genhav(i,inds)];
        end % j. parametreye ait gen parçası oluşturuldu
        %baslangic=(j-1)*b+1;
        %bitis=j*b;
        %gendeg(i,j)=ikion(genhav(i,baslangic:bitis));
        gendegs(i,j)=graycode(gen); %genetik kodun onluk sistemde değeri
        clear gen;
    end
end
gendegs=2*(genlik/2^b)*(gendegs-2^(b-1)); %biraz düşün +genlik, -genlik aralığında değişiyor.
```

%%%

%%TAF-CNN%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%CNN SİMÜLASYONU%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function V1=mcnns(parlaklik1,A,B,I,es1,esh,es2,esl)

```
hataşiniri=10^(-16);
zamanaralik=0.1;
guncelleme=6/zamanaralik;
```

```
[m,n]=size(parlaklik1);
parlaklik=parlaklik1(1:m,:);
[m1 n1]=size(A);
C=1;
Rx=1;
Vu=2*parlaklik-1;
[m,n]=size(Vu);
alan=m*n;
I=I*ones(m,n);
Vx=Vu;
%Vx=parlaklik1(m+1:2*m,:);
```

```
for r=1:m
    for s=1:n
        if Vx(r,s)>es1 & Vx(r,s)<esh
            Vy(r,s)=(2/(esh-es1))*(Vx(r,s)-es1)-1;
        elseif Vx(r,s)>=esh & Vx(r,s)<=es2
            Vy(r,s)=1;
        elseif Vx(r,s)>=es2 & Vx(r,s)<=es1
            Vy(r,s)=1-(2/(es1-es2))*(Vx(r,s)-es2);
        else
            Vy(r,s)=-1;
        end
    end
end
end
```



```
Vyi=Vy+.21;
tek=0;
ilr=conv4(B,Vu);
[im in]=size(ilr)
sabit=ilr(m1-1:im-1,n1-1:in-1)+1;
error=2*alan;
while (error>hata siniri*alan & tek<guncelleme)
    %pause
    tek=tek+1
    Vyi=Vy;
    ger=conv4(A,Vy);
    %pause
    dVx=-Vx/Rx+ger(m1-1:im-1,n1-1:in-1)+sabit;

    %pause
    topd=dVx/alan;
    topd1=sum(abs(topd(:)))
    tt(tek)=topd1;

    Vxi=Vx;
    Vx=Vx+zaman aralik*dVx;
    for r=1:m
        for s=1:n
            if Vx(r,s)>esh & Vx(r,s)<esh
                Vy(r,s)=(2/(esh-es1))*(Vx(r,s)-es1)-1;
            elseif Vx(r,s)>=esh & Vx(r,s)<=es2
                Vy(r,s)=1;
            elseif Vx(r,s)>=es2 & Vx(r,s)<=es1
                Vy(r,s)=1-(2/(es1-es2))*(Vx(r,s)-es2);
            else
                Vy(r,s)=-1;
            end
        end
    end

    error=sum(sum((Vx-Vxi).^2));
    V1=(Vy+1)/2;
    imshow(V1)
end
error=error
imshow(V1)

figure
plot(tt)
%V1=(Vy-1)/2;
%figure(mod(tek,4)+1);
%imshow(V1);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
İKİLİK-ONLUK DON.%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%ikiliyi onluğa dönüştür
function on=ikion(ikili)
[m n]=size(ikili);
```



```
on=0;
for i=1:n
    on=on+ikili(1,n-i+1)*(2^(i-1));
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%STANDART CNN.%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%CNN SİMÜLASYONU%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function V1=cnns(parlaklik,A,B,I)

hataasiniri=10^(-15);
zamanaralik=0.1;
guncelleme=5/zamanaralik;

%format long;
C=1;
Rx=1;
Vu=2*parlaklik-I;
[m,n]=size(Vu);
alan=m*n;
I=I*ones(m,n);
Vx=Vu;
[sbboy,sben]=size(A);
Vy=.5*(abs(Vx+1)-abs(Vx-1));
%k=25;
%Vy=.5*(log(cosh(k*(Vx+1)))-log(cosh(k*(Vx-1))))/k;
Vyi=Vy+.1;
tek=0;
ilr=conv4(B,Vu);
[im in]=size(ilr);
sabit=ilr(sbboy-1:im-1,sbboy-1:in-1)+I;
error=2*alan;
while (error>hataasiniri*alan & tek<guncelleme)
    %pause
    tek=tek+1;
    Vyi=Vy;
    ger=conv4(A,Vy);
    %pause
    dVx=-Vx/Rx+ger(sbboy-1:im-1,sbboy-1:in-1)+sabit;

    topd=dVx/alan;
    topd1=sum(abs(topd(:)));
    tt(tek)=topd1;

    %topd1=sum(abs(topd(:)))
    %tt(tek)=topd1;
    %pause
    Vxi=Vx;
    Vx=Vx+zamanaralik*dVx;
    Vy=.5*(abs(Vx+1)-abs(Vx-1));

    %Vy=.5*(log(cosh(k*(Vx+1)))-log(cosh(k*(Vx-1))))/k;
    error=sum(sum((Vx-Vxi).^2));
    error1(tek)=error/alan;
```

