



T.C.
KONYA TEKNİK ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ



KOMPLEKS DEĞERLİ YAPAY SİNİR
AĞLARI İÇİN YENİ AKTİVASYON
FONKSİYONLARININ TANIMLANMASI

Mehmet ÇELEBİ

YÜKSEK LİSANS TEZİ

Elektrik-Elektronik Mühendisliği Anabilim Dalı

Ağustos-2020
KONYA
Her Hakkı Saklıdır

TEZ KABUL VE ONAYI

Mehmet ÇELEBİ tarafından hazırlanan “Kompleks Değerli Yapay Sinir Ağları İçin Yeni Aktivasyon Fonksiyonlarının Tanımlanması” adlı tez çalışması 05/08/2020 tarihinde aşağıdaki jüri tarafından oy birliği / oy çokluğu ile Konya Teknik Üniversitesi Lisansüstü Eğitim Enstitüsü Elektrik-Elektronik Mühendisliği Anabilim Dalı’nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Jüri Üyeleri

Başkan

Doç Dr. Bayram AKDEMİR

Danışman

Doç. Dr. Murat CEYLAN

Üye

Dr. Öğr. Üyesi Muhammet Üsame ÖZİÇ

İmza

.....

.....

.....

Yukarıdaki sonucu onaylarım.

Prof. Dr. Saadettin Erhan KESEN
Enstitü Müdürü

TEZ BİLDİRİMİ

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

DECLARATION PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.



Mehmet ÇELEBİ

Tarih:05.08.2020

ÖZET

YÜKSEK LİSANS TEZİ

KOMPLEKS DEĞERLİ YAPAY SİNİR AĞLARI İÇİN YENİ AKTİVASYON FONKSİYONLARININ TANIMLANMASI

Mehmet ÇELEBİ

**Konya Teknik Üniversitesi
Lisansüstü Eğitim Enstitüsü
Elektrik-Elektronik Mühendisliği Anabilim Dalı**

Danışman: Doç. Dr. Murat CEYLAN

2020, 70 Sayfa

Jüri

**Doç Dr. Bayram AKDEMİR
Doç. Dr. Murat CEYLAN
Dr. Öğr. Üyesi Muhammet Üsame ÖZİÇ**

Kompleks değerli yapay sinir ağı (KDYSA), ağırlıkları, eşik değerleri, giriş, çıkış sinyallerinin tamamı kompleks sayılardan oluşan sinir ağlarıdır. KDYSA, kompleks sayılara sahip verileri doğrudan işlemek için geliştirilmiştir. Kompleks sayılara sahip verileri içeren problemlerin çözümünde, bilinen yöntem kullanıldığında, kompleks olan verinin reel ve imajiner kısımları için ayrı ayrı YSA uygulanması gerekmektedir. Hâlbuki aynı problem için KDYSA uygulandığında veriler üzerinde, reel ve imajiner kısımları ayırmaya gerek kalmadan, doğrudan işlem yapılabilmektedir. Böylelikle problemlerin işlem zamanının azaldığı gibi ağına da doğruluk oranının arttırıldığı görülmüştür.

KDYSA radar görüntüleme, telekomünikasyon, Fourier dönüşümü ile görüntü işleme ve karmaşık sayılarla uğraşan anten tasarımı gibi çeşitli alanlarda yaygın olarak kullanılmaktadır. Bu işlemleri yapan yapay sinir ağlarının performansı bazı faktörlere göre değişme göstermektedir. Bunlar; verilerin normalizasyonu, öğrenme oranı, başlangıç ağırlıkları ve uygun aktivasyon fonksiyonunu seçimi olarak sıralanabilir. Bu faktörlerden en önemlisi uygun aktivasyon fonksiyonunun seçimidir. Bir aktivasyon fonksiyonunun seçimi, kompleks geriye yayılım algoritmasının yakınsamasını ve genel oluşum özelliklerini belirler. Yapay sinir ağlarında aktivasyon fonksiyonunun seçimi, eğitim dinamikleri ve görev performansı üzerinde önemli bir etkiye sahiptir.

Bu tez çalışmasında, KDYSA performansını arttırmak ve eğitim süresini kısaltmak için yeni kompleks aktivasyon fonksiyonları tanımlanmıştır. Bu fonksiyonlar; kompleks swish, kompleks modifiye swish, kompleks e-swish ve kompleks düzleştirilmiş t-swish olarak isimlendirilmiştir. Yeni tanımlanan bu aktivasyon fonksiyonlarını kullanan ağların performansı, literatürde sıklıkla çözülen Exclusive-OR (XOR), Simetri ve Kanal dengeleme problemleri üzerinde değerlendirilmiştir. Elde edilen sonuçlar karşılaştırmalı olarak sunulmuştur.

Anahtar Kelimeler: Kompleks Değerli Yapay Sinir Ağları, Aktivasyon Fonksiyonu, Swish, Modifiye Swish, E-Swish, Düzleştirilmiş t-swish

ABSTRACT

MS THESIS

THE NEW ACTIVATION FUNCTIONS FOR COMPLEX VALUED NEURAL NETWORKS

Mehmet ÇELEBİ

**Konya Technical University
Institute of Graduate Studies
Department of Electrical-Electronic Engineering**

Advisor: Assoc. Prof. Dr. Murat CEYLAN

2020, 70 Pages

Jury

**Assoc. Prof. Dr. Bayram AKDEMİR
Assoc. Prof. Dr. Murat CEYLAN
Asst. Prof. Dr. Muhammet Üsame ÖZİÇ**

Complex-valued artificial neural network (CVANN), whose parameters (weights, threshold values, input and output signals) are all complex numbers, was developed to process complex valued data directly. In the solution of problems involving data with complex numbers, ANN should be applied separately for real and imaginary parts of complex data when known method is used. However, when CVANN is applied for the same problem, data can be processed directly without having to separate real and imaginary parts. Thus, it has been observed that the processing time is reduced and the accuracy rate of network is increased.

CVANN have become widely used in various fields such as radar imaging, communication signal processing, image processing with the Fourier transformation and antenna designing which dealing with complex numbers. The performance of the CVANN performing these processes varies depending on some factors. These factors are; minimization criterion, learning rate, initial bias and weights and activation function. The most important of these factors is activation function. The selection of the appropriate activation function determines the convergence and general formation characteristics of the complex back propagation algorithm.

In this thesis, new complex activation functions are defined to increase the performance of our complex-valued neural network and shorten the training period. These functions are; complex swish, complex modified swish, complex e-swish and complex flatten t-swish. The convergence performance of networks using these newly defined activation functions has been evaluated on Exclusive-OR (XOR), Symmetry and fading equalization problems which are frequently solved in the literature. The results obtained are presented comparatively.

Keywords: Complex valued artificial neural network, Activation function, Swish, Modified swish, E-swish, Flatten T-swish

ÖNSÖZ

Yüksek lisans tez çalışmalarım boyunca değerli katkılarıyla ve tecrübeleriyle beni yönlendiren danışmanım Doç. Dr. Murat CEYLAN' a, bölümümüzün değerli öğretim elemanlarına ve maddi manevi katkılarını benden esirgemeyen aileme teşekkür ederim.

Mehmet ÇELEBİ
KONYA-2020



İÇİNDEKİLER

ÖZET	iv
ABSTRACT	v
ÖNSÖZ	vi
İÇİNDEKİLER	vii
SİMGELER VE KISALTMALAR	ix
1. GİRİŞ	1
2. KAYNAK ARAŞTIRMASI	3
3. MATERYAL VE YÖNTEM	6
3.1. Yapay Sinir Ağları	6
3.1.1. Yapay sinir ağları yapısı	6
3.1.2. Yapay sinir ağı modelleri.....	7
3.2. Kompleks Değerli Yapay Sinir Ağları.....	9
3.2.1. Kompleks değerli nöron modeli	10
3.2.2. KDYSA ileri yayılım işlemleri	13
3.2.3. KDYSA geri yayılım işlemleri ve hata fonksiyonu.....	15
3.2.4. Kompleks değerli aktivasyon fonksiyonu.....	17
3.3. Kompleks Değerli Yeni Aktivasyon Fonksiyonlarının Tanımlanması.....	18
3.3.1. Kompleks değerli swish aktivasyon fonksiyonu	18
3.3.2. Kompleks değerli modifiye swish aktivasyon fonksiyonu	20
3.3.3. Kompleks değerli e-swish aktivasyon fonksiyonu	21
3.3.4. Kompleks değerli düzleştirilmiş t-swish aktivasyon fonksiyonu	22
3.4. Kompleks Değerli Veriler.....	23
3.4.1. Kompleks değerli XOR problemi	23
3.4.2. Kompleks değerli simetri problemi	24
3.4.3. Kanal dengeleme problemi	26
4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA	27
4.1. Kompleks Swish Aktivasyon Fonksiyonun Diğer Aktivasyon Fonksiyonlar ile Karşılaştırılması	27
4.1.1. Kompleks değerli XOR problemi	27
4.1.2. Kompleks değerli simetri problemi	29
4.1.3. Kanal dengeleme problemi	31
4.2. Kompleks Modifiye swish, E-Swish ve DTS Aktivasyon Fonksiyonlarının Sonuca En Hızlı Yakınsadığı Sabit Değerlerin Bulunması	33
4.2.1. Kompleks değerli XOR problemi	33
4.2.2. Kompleks değerli simetri problemi	37
4.2.3. Kanal dengeleme problemi	40
4.3. Yeni Tanımlanan Aktivasyon Fonksiyonlarının Performanslarının Birbirleri ile Karşılaştırılması	43

4.3.1. Ortalama karesel hata baz alınarak kıyaslama	43
4.3.2. Sabit iterasyon sayıları için aktivasyon fonksiyonlarının karşılaştırması.....	48
5. SONUÇLAR VE ÖNERİLER	52
5.1 Sonuçlar	52
5.2 Öneriler	53
KAYNAKLAR	54
EKLER	58
ÖZGEÇMİŞ	70



SİMGELER VE KISALTMALAR

Simgeler

σ	: Sigmoid aktivasyon fonksiyonu
ε	: Öğrenme sabiti
e	: Euler sabiti

Kısaltmalar

BBA	:Beyin Bilgisayar Arayüzü
EEG	:Elektroensefalografi
DTS	:Düzleştirilmiş T Swish
GBYSA	:Geri Beslemeli Yapay Sinir Ağı
İBYSA	:İleri Beslemeli Yapay Sinir Ağı
KGAE	: Kanada Gelişmiş Araştırma Enstitüsü
KDYSA	: Kompleks Değerli Yapay Sinir Ağı
MUSTE	: Modifiye Ulusal Standartlar ve Teknoloji Enstitüsü
OKH	:Ortalama Karesel Hata
OKHK	:Ortalama Karesel Hata Kökü
PReLB	: Parametrik Rektifiyeli Lineer Birim
ReLB	: Rektifiyeli Lineer Birim
ÜLB	: Üstel Lineer Birim
YSA	: Yapay Sinir Ağı
XOR	: Özel Veya

1. GİRİŞ

Yapay Sinir Ağları (YSA), insan beyni gibi çalışarak öğrenme yolu ile yeni bilgiler üretebilme ve keşfedebilme yeteneklerini, dışardan herhangi bir yardım olmaksızın kendi başına gerçekleştirebilen bir yapay zekâ metodudur (Ceylan, 2004). YSA, hızlı çalışabilme yeteneği, öğrenme yeteneği, genelleme yapabilmesi, farklı problemlere karşı kolayca uyarlanabilmesi, giriş ve çıkış arasındaki ilişkiyi tarif etme mecburiyetinin olmaması gibi avantajlarıyla geniş bir uygulama alanına sahiptir. YSA, belirlenen bir sisteme ait girişler ve çıkışlar arasında ilişki kurabilme yeteneğine sahiptir (Yaşar, 2015).

Kompleks Değerli Yapay Sinir Ağı (KDSYA), kompleks değerli sayılar içeren verilerin herhangi bir ön işlem gerekmeksizin doğrudan işlenmesi için geliştirilmiştir. KDYSYA'da girişler, eşik değerleri, ağırlıklar ve çıkışların hepsi kompleks sayılardan oluşur. Kompleks değerli verilerin işlenmesi için bilinen reel değerli yapay sinir ağları kullanıldığında, verilerin hem reel hem de imajiner kısımları için ayrı ayrı YSA uygulanması gerekmektedir. Halbuki kompleks değerli verilerin işlenmesi için KDYSYA kullanıldığında, veriler doğrudan işlenebildiği için hem işlem zamanının azaldığı hem de doğruluk oranının arttığı görülmüştür (Ceylan, 2009).

KDYSYA'da en önemli sorunlardan biri düğümler için uygun aktivasyon fonksiyonunun seçimidir. Uygun aktivasyon fonksiyonunun seçimi, kompleks geriye yayılım algoritmasının yakınsaması ve ağıın performansı üzerinde önemli etkisi vardır. Tez çalışmasında, Google araştırmacıları Ramachandra ve ark. (2017) tarafından keşfedilen ve reel düzlemde iyi sonuçlar veren swish aktivasyon fonksiyonu kompleks düzlemde ele alınmıştır. Ayrıca swish fonksiyonu üzerinden türetilen modifiye swish, e-swish ve düzleştirilmiş t-swish (DTS) aktivasyon fonksiyonları da kompleks düzlemde incelenmiştir.

Önerilen aktivasyon fonksiyonlarının kullanılabilirliğini ve geçerliliğini kontrol etmek için, bu fonksiyonları kullanan ağlar kompleks değerli XOR (Özel Veya), simetri belirleme ve kanal dengeleme problemleri üzerinde test edilmiştir. İlk olarak kompleks swish aktivasyon fonksiyonu literatürde sıklıkla kullanılan logaritmik sigmoid ve tanjant sigmoid aktivasyon fonksiyonlarını kullanan ağ ile XOR, simetri ve kanal dengeleme problemleri üzerinde test edilmiştir. Daha sonra kompleks swish fonksiyonu üzerinden türetilen modifiye swish, kompleks e-swish ve kompleks DTS aktivasyon fonksiyonlarının en iyi performans gösterdiği sabit değerler bulunmuştur. Son olarak

kompleks modifiye swish, kompleks e-swish, kompleks düzleştirilmiş t-swish ve kompleks swish aktivasyon fonksiyonlarını kullanan ağlar XOR, simetri ve kanal dengeleme problemleri üzerinde test edilmiş ve elde edilen sonuçlar karşılaştırmalı olarak sunulmuştur.



2. KAYNAK ARAŞTIRMASI

Ceylan (2004) “Kompleks Değerli Yapay Sinir Ağı İle Algoritma Geliştirilmesi ve Uygulamaları” adlı tezinde kompleks değerli yapay sinir ağlarının örüntü sınıflamadaki performanslarını karşılaştırabilmek için logaritmik sigmoid ve tanjant sigmoid aktivasyon fonksiyonu üzerinde çalışmalar gerçekleştirmiştir. Sekiz bitlik kompleks XOR problemi üzerinde yapılan çalışmalar sonucunda, tanjant aktivasyon fonksiyonunun kompleks değerli dataların sınıflandırılmasında logaritmik sigmoid fonksiyonundan daha uygun olduğunu göstermiştir.

Ceylan ve ark. (2007) yaptıkları çalışmada, Mexican hat ve Haar dalgacık kompleks aktivasyon fonksiyonunu kullanan yeni bir ağ önermişlerdir. Önerilen bu ağ 4 bitlik kompleks XOR verisi üzerinde test edilmiştir. Yapılan testler sonucunda yeni ağın performansının gerçek ve kompleks değerli yapay ağlarından daha iyi olduğunu ve sonuca daha erken yakınsadığını ayrıca önerilen yeni ağın belirlenen hata kriterlerine diğer ağlardan daha az iterasyonda ulaştığını göstermişlerdir.

Amin ve Murase (2009) yayınladıkları makalede kompleks değerli veriler üzerinde iki yeni aktivasyon fonksiyonunu kullanan KDYSA'larını test etmişlerdir. Önerilen aktivasyon fonksiyonlarının ilki kompleks değerlerin büyüklüğünü gösterirken, ikincisi kompleks değerlerin karekökünü vermektedir. Önerilen iki aktivasyon fonksiyonunu, literatürde Micheal ve Awwal (1999) tarafından çözülen üç girişli ve 256 farklı değere sahip Boolean problemi üzerinde test etmişlerdir. Michel ve Awwal (1999) yaptıkları çalışmada 3 girişli ve 256 farklı değere sahip Boolean probleminin 245'ini çözebilirken, Amin ve Murase (2009) önerdikleri ilk aktivasyon fonksiyonunu (kompleks değerlerin büyüklüğü) kullanan ağ 256 farklı değerli problemin 250'sini, ikinci aktivasyon fonksiyonunu (kompleks değerlerin karekökü) kullanan ağ 253'ünü çözmüştür.

Burse ve ark. (2011), kompleks değerli 16 bit XOR problemi üzerinde çeşitli aktivasyon fonksiyonlarının performanslarını karşılaştırmışlardır. Bu aktivasyon fonksiyonları; hiperbolik tanjant, sigmoid, exponential, arksinüs, arkatanjant ve arkhiperbolik sinüs'tür. Yaptıkları testler sonucunda en iyi sonuçları veren aktivasyon fonksiyonunun sigmoid ve hiperbolik tanjant olduğunu göstermişlerdir, ayrıca yakınsamanın öğrenme oranı ve ağ mimarisine bağlı olduğunu belirtmişlerdir.

Jalab ve Ibrahim (2011) çok katmanlı KDYSA için yeni bir kompleks değerli sigmoid aktivasyon fonksiyonu önermişlerdir. Sınıflandırıcı olarak gerçek dünyadaki bir problem örneği kullanılmıştır. KDYSA, yengeçlerin cinsiyetini fiziksel boyutlarından

tanımlamak için bir sınıflandırıcı olarak kullanılmıştır. Simülasyonlar sonucunda yeni aktivasyon fonksiyonunu kullanan tamamen kompleks değerli ağı, geleneksel sigmoid aktivasyon fonksiyonu kullanan ağa kıyasla test süresini % 54 oranında azalttığını ortaya koymuşlardır.

Li ve Jiang (2014) kompleks değerli yapay sinir ağı için Morlet (Gabor) aktivasyon fonksiyonu ile yeni bir hata fonksiyonu önermişlerdir. Önerdikleri yöntemin uygulanabilirliğini doğrulamak için, XOR problemini kullanmışlardır. Ağ için öğrenme oranı ve maksimum iterasyon sayısı sırasıyla 0.1 ve 5000 olarak seçilmiştir. Önerilen Morlet aktivasyon fonksiyonu, Mexican hat ve Haar fonksiyonu ile karşılaştırılmıştır. Yapılan testler sonucunda hata kriteri 0.1'e, Morlet aktivasyon fonksiyonunu kullanan ağ 153. iterasyonda ulaşırken, Mexican hat ve Haar fonksiyonunu kullanan KDYSA ise bu hata değerine ulaşmak için 475 iterasyona ihtiyaç duymuştur.

Peker ve ark. 2016 yılında elektroensefalografi (EEG) sinyallerinden epilepsi tanısı için yeni bir KDYSA yöntemi önermişlerdir. Önerdikleri yöntemde EEG verilerinin sınıflandırılmasında ve istatistiksel özelliklerin çıkarılmasında çift ağaç kompleks dalgacık dönüşümü kullanarak KDYSA'ların verimliliğini artırmayı amaçlamışlardır. Verilerin sınıflandırılması için ağı gizli katmanında kompleks sigmoid ve kompleks mexican hat aktivasyon fonksiyonları kullanmışlardır. Yaptıkları testler sonucunda kompleks sigmoid fonksiyonunu kullanan ağ 1.9×10^{-13} ortalama karesel hata kökü (OKHK) elde ederken, mexican hat fonksiyonunu kullanan ağ 6.19×10^{-4} OKHK elde etmiştir.

Mönnig ve Manandhar (2018), Modifiye Ulusal Standartlar ve Teknoloji Enstitüsü (MUSTE) ve Kanada Gelişmiş Araştırma Enstitüsü (KGAE) veri tabanı üzerinde çeşitli kompleks aktivasyon fonksiyonları kullanarak çalışmalar yapmışlar. Tasarladıkları ağda, kompleks hiperbolik tanjant, Rektifiyeli Lineer Birim (ReLB), mutlak değer kareleri toplamı ve mutlak değer aktivasyon fonksiyonlarını kullanmışlardır. Bu fonksiyonları kullanan ağları farklı gizli düğüm sayıları KGAE-10, KGAE-100 ve MUSTE basamak sınıflandırma verileri üzerinde test etmişlerdir. Yapılan testler sonucundan ReLB aktivasyon fonksiyonunu kullanan ağın diğer aktivasyon fonksiyonunu kullanan ağlardan daha kararlı çalıştığını belirtilmiştir.

Ikeda ve Washizawa (2019), yaptıkları çalışmada KDYSA için mutlak değer ve bölünmüş ReLB mutlak değer aktivasyon fonksiyonlarını önermişlerdir. ReLB aktivasyon fonksiyonunu kullanan ağ, BBA2000 (Beyin Bilgisayar Arayüzü) tarafından sağlanan, açık-kapalı gözler için EEG sinyalleri üzerinde test etmişlerdir. Üç kanallı EEG

verileri üzerinde yapılan deneylerde, giriş boyutu 192, 384 ve 768 olarak alınmıştır. Gizli katman sayıları değiştirilerek yapılan testler sonucunda mutlak değer aktivasyon fonksiyonunu kullanan ağın doğruluk oranının, bölünmüş ReLB mutlak değeri ve sigmoid fonksiyonu kullanan ağdan daha yüksek olduğu göstermişlerdir. Ayrıca önerilen aktivasyon fonksiyonunu kullanan ağın düşük gizli katman sayısına rağmen yüksek test doğruluğuna sahip olduğunu belirtilmişlerdir.

Enomoto ve ark. (2019) yaptıkları çalışmada parametrik softign'ı negatif bölgeye uygulayarak iki tür aktivasyon fonksiyonu önermişlerdir. Önerilen bu fonksiyonlar ReLB, Üstel Lineer Birim (ÜLB) ve Parametrik Rektifiyeli Lineer Birim (PReLB) aktivasyon fonksiyonları kullanan KDSYA ile kıyaslanmıştır. Kıyaslama için KGAE-10 veri tabanı kullanılmıştır. Sonuçlara bakıldığında, önerilen KDYSA'nın, nesne görüntülerini diğerlerine göre doğru bir şekilde tanıyabildiğini gösterilmiştir.

Özgür ve ark. (2020) KDYSA için yeni bir tip aktivasyon fonksiyonu sunmuşlardır. Önerdikleri aktivasyon fonksiyonu, belirli bir elipsi sabitleyecek şekilde yapılandırılmıştır. Aktivasyon fonksiyonu olarak tanıtılan kompleks fonksiyonların özel bir formunu kullanarak kompleks değerli bir Hopfield sinir ağına uygulamışlardır. Düzlem eğrisi elipsinin odaklama özelliği gibi ilginç geometrik özellikleri göz önüne alındığında, bu özelliklerin çeşitli sinir ağlarında olası uygulamalara sahip olabileceğini vurgulamışlardır.

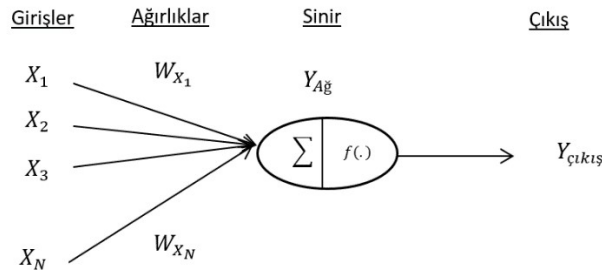
3. MATERYAL VE YÖNTEM

3.1. Yapay Sinir Ağları

YSA, insan beyninin özelliklerinden olan öğrenme, öğrendiği bilgilerden yeni bilgiler türetebilme ve karar verme gibi yetenekleri dışarıdan herhangi bir yardım olmaksızın kendi başına gerçekleştirmeye çalışan yaklaşımlardır. YSA insan beyninin oldukça basitleştirilmiş bir modelidir. YSA yapıları biyolojik sinir ağlarının çalışma prensibine çok benzemektedir. Fakat YSA'lar gerçek sinirlerin sadece bazı fonksiyonlarına benzerlik gösterir. Bu yüzden, YSA gerçek sinirlere göre daha basit yapıda olup analiz edilmeleri daha kolay olmaktadır. İlk YSA modeli Warren McCulloch ve Walter Pitts tarafından ortaya atılmıştır. McCulloch ve Pitts (1946), insan beyninin hesaplama yeteneğinden yola çıkarak, elektrik devreleri kullanarak basit bir sinir ağı tasarlamışlardır. Daha sonra Hebb (1949), Roseblant (1958), Minsky ve Papert (1969), Grossberg (1976), Hopfield (1984) ve Kohonen (1986) gibi araştırmacıların yaptıkları çalışmalar sonucunda YSA yapısı günümüzdeki halini almıştır.

3.1.1. Yapay sinir ağları yapısı

YSA'ların temelini sinir yapıları oluşturur. Her bir sinir yapısı birden fazla girişe ve bir adet çıkışa sahiptir. Sinir yapıları bir araya gelerek yapay sinir ağını oluşturmaktadır. Şekil 3.1'de N adet girişe sahip basit bir sinir yapısı gösterilmiştir. Şekilde gösterilen X 1'den N'e kadar olan sinir girişlerini, W 1'den N'e kadar olan girişlere ait ağırlıkları, f(.) ise kullanılan aktivasyon fonksiyonunu ifade etmektedir.



Şekil 3.1. Temel sinir yapısı

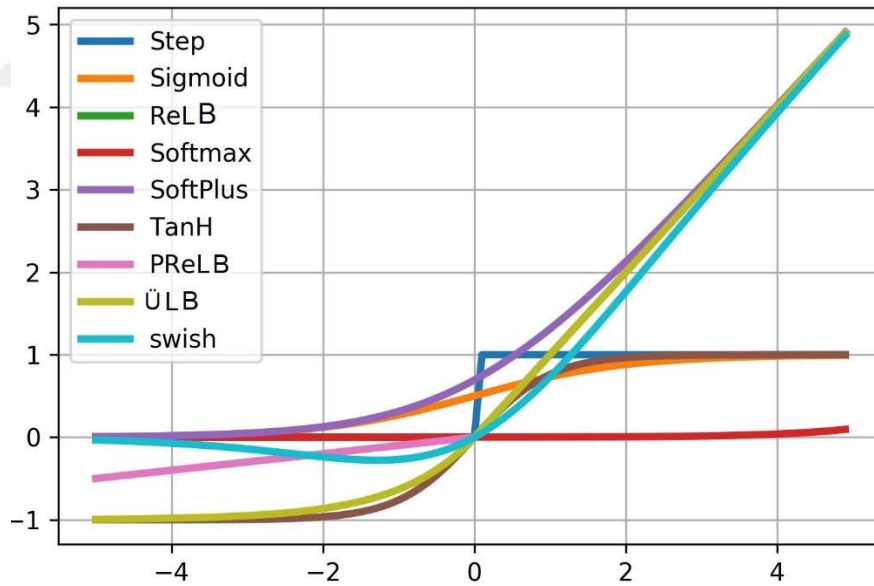
Şekil 3.1’de gösterilen $Y_{Ağ}$ sinir girişleri ile ağırlıkları çarpımlarının toplamını ifade etmektedir. $Y_{Ağ}$ matematiksel olarak Eşitlik 3.1’deki gibi ifade edilmektedir.

$$Y_{Ağ} = \sum_{i=1}^M W_{XN} * X_N \quad (3.1)$$

Sinir yapısına ait çıkışı gösteren $Y_{Çıkış}$ fonksiyonu ise Eşitlik 3.2’deki gibi ifade edilmektedir (Ceylan, 2004).

$$Y_{Çıkış} = f(Y_{Ağ}) \quad (3.2)$$

Aktivasyon fonksiyonu olarak tanımlanan $f(\cdot)$; step, sigmoid, softmax, softplus ve hiperbolik tanjant olmak üzere çeşitli fonksiyonlardan oluşabilir. YSA’da sıklıkla kullanılan aktivasyon fonksiyonları Şekil 3.2’de verilmiştir.



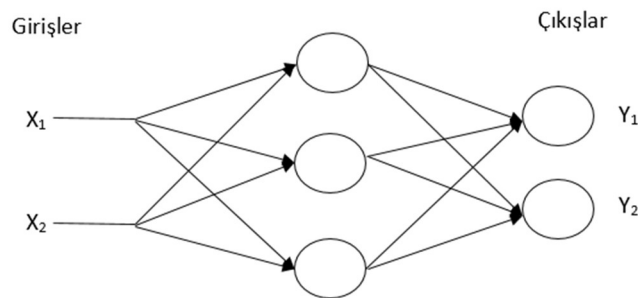
Şekil 3.2. Farklı aktivasyon fonksiyonları

3.1.2. Yapay sinir ağı modelleri

Yapay sinir ağları içerisinde bulduklarını sinirler arası bağlantılar açısından incelendiğinde, temel olarak iki tip model karşımıza çıkar. Bunlar, İleri Beslemeli Yapay Sinir Ağları (İBYSA) ve Geri Beslemeli Yapay Sinir Ağları (GBYSA) 'dır.

İleri beslemeli yapay sinir ağları:

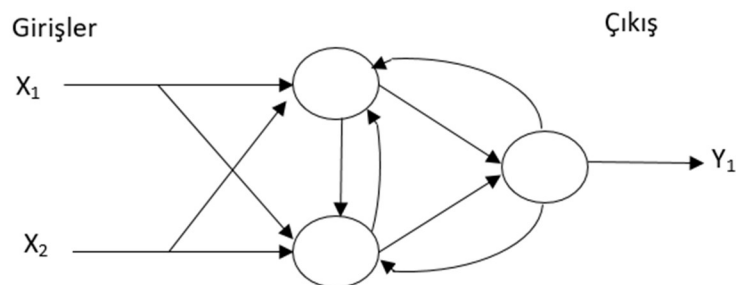
İBYSA'da her bir katmandaki sinirler sadece önceki katmandaki hücreler tarafında beslenir. Aynı katman içerisinde bulunan nöronlar birbiri ile bağlantılı değildir ve diğer katmanlardaki nöronlar ile ağırlık değeri içeren bir bağlantı biçimi ile bağlanmıştırlar. Aynı katmanda bulunan nöronların birbirine bağlantılı olmaması sebebi ile ağ ileriye doğru iletim sağlamaktadır. X_1 ve X_2 girişler, Y_1 ve Y_2 çıkışlar olmak üzere İBYSA yapısı Şekil 3.3'te verilmiştir (Gödekoğlu, 2019).



Şekil 3.3. İleri beslemeli YSA yapısı

Geri beslemeli yapay sinir ağları:

GBYSA'da çıkış ve ara katmanlardaki çıkışlar, giriş veya daha önceki ara katmanlara geri beslenmektedir. Yani nöronlar kendi katmanlarındaki başka herhangi bir nörona veya kendinden önceki katmandaki başka bir nörona bağlanabilirler. Böylece girişler, ileri ve geri yönde aktarılırlar. X_1 ve X_2 girişler, Y_1 çıkış olmak üzere GBYSA yapısı Şekil 3.4'te verilmiştir (Gödekoğlu,2019).



Şekil 3.4. Geri beslemeli YSA yapısı

Yapay sinir ağı eğitiminde kullanan birçok yöntem vardır. Bu yöntemlerden en önemlileri; eğitici (supervised) öğrenme, eğitici (unsupervised) öğrenme ve takviyeli (destekleyici) öğrenmedir (Özkurt ve ark., 2000).

Eğitici öğrenme algoritmasında hata, ağa ait çıkışlar ve hedef arasındaki farka göre hesaplanır. YSA tarafından üretilen çıkış değeri ile beklenen çıkış değeri karşılaştırılarak bir hata sinyali üretilir. Üretilen bu hata değerine göre YSA parametreleri güncellenir. Eğitim işlemi, ağdan istenilen değer üretilinceye kadar devam eder (Şen, 2018).

Eğitici öğrenme, danışmansız öğrenme veya kendi kendini yönetme olarak da bilinir. Bu algorithmada girişlere karşılık çıkışın ne olması gerektiği belli değildir. Ağ giriş ve çıkış arasındaki bağlantıyı kendi kendisine keşfetmek zorundadır (Haykin, 1999).

Takviyeli (destekleyici) öğrenme algoritmasında ise YSA gerçekleştirdiği her iterasyon için sonucun iyi veya kötü olduğu bilgisini vererek sonuçlara göre kendisini yeniler (Özkurt ve ark., 2000).

3.2. Kompleks Değerli Yapay Sinir Ağları

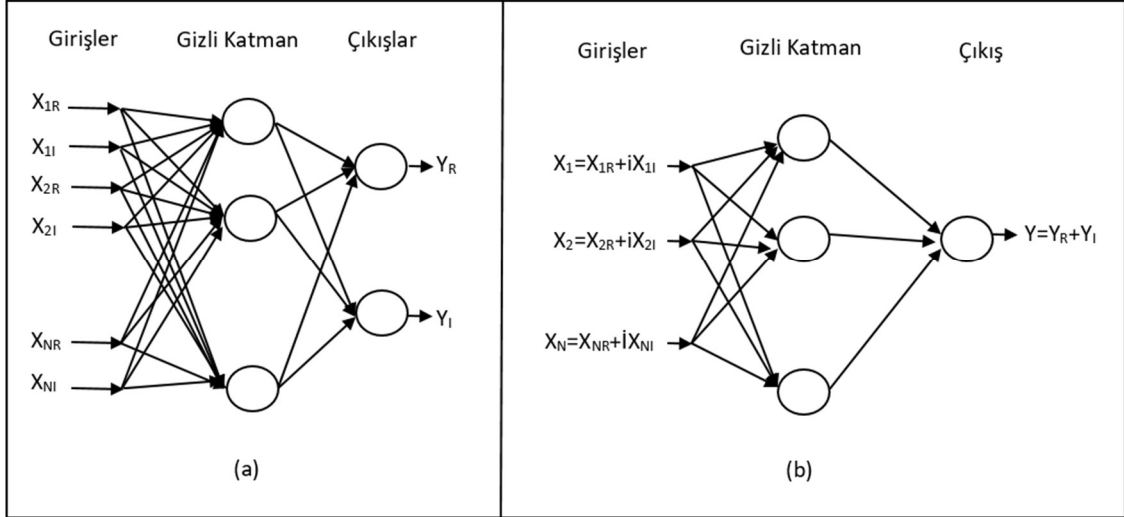
KDYSA resmi olarak ilk defa Clarke (1990) tarafından tanımlanmıştır. Daha sonraki yıllarda gradyanlı iniş fikrine dayanan kompleks değerli geri yayılım algoritması üzerine çeşitli araştırmacılar tarafından çalışmalar yapılmıştır (Benvenuto ve Piazza, 1992; Nitta, 1993; Georgiou & Koutsougeras, 1992). 2000'li yıllarda ise, gerçek değerli sinir ağları gibi kompleks sinir ağları, çeşitli görevlere başarıyla uygulanmıştır. Bu görevler, kompleks değerli verilerin işlenmesi ve analizini içermektedir. Özellikle, dalga formlarındaki sinyaller KDYSA'ya girdi verileri olarak kullanılmıştır (Hirose, 2009).

KDYSA ağırlıkları, eşik değerleri, giriş, çıkış sinyallerinin tümü kompleks sayılardan oluşan sinir ağlarıdır. KDYSA zaman serisi analizi, sinyal genliği ve fazı gibi gerçek dünya olaylarını ifade etmek ve çeşitli matematiksel geometrik ilişkileri analiz etmek için kullanılırlar. Ayrıca, KDYSA gerçek değerli sinyallerin işlenmesinde, gerçek değerli sinir ağlarına göre daha yüksek performans göstermiştir (Jalab ve Ibrahim, 2014).

Kompleks sayılara sahip verileri içeren problemler gerçek değerli YSA ile çözülmek istendiğinde, verinin reel ve imajiner kısımları için ayrı ayrı YSA uygulanması gerekmektedir. Hâlbuki aynı problemler KDYSA kullanılarak çözülmek istendiğinde böyle bir işleme gerek kalmaz. KDYSA kompleks sayılara sahip verileri, hiçbir ön işlem

gerektirmeksizin doğrudan işleyebilmektedir. Böylece veriler işlenirken hem işlem zamanı düşürülmüş hem de doğruluk oranı arttırılmıştır (Ceylan,2009).

Kompleks değerli ve reel değerli YSA'ların çalışma prensipleri ile ilgili farklılıklar Şekil 3.5'te verilmiştir.



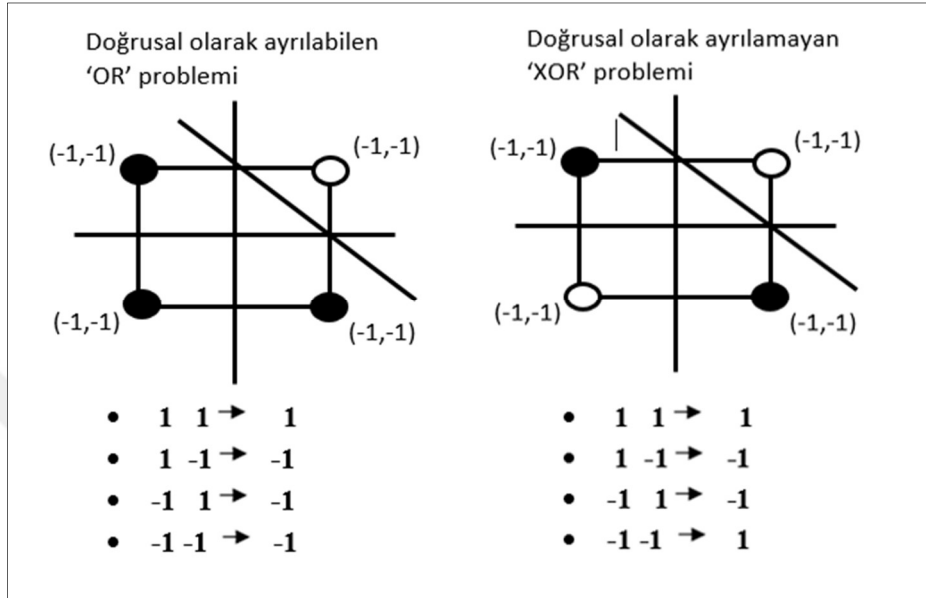
Şekil 3.5. Kompleks değerli verilerin (a) gerçek değerli YSA ve (b) kompleks değerli YSA ile işlenmesi

Parametrelerinin (ağırlıklar ve eşik değerlerin) hepsi kompleks olan KDYSA'nın; telekomünikasyon, konuşma tanıma ve Fourier dönüşümü ile görüntü işleme gibi kompleks sayılarla ilgilenen bir çok alanda uygulamaları vardır (Leung ve Haykin, 1991). Ayrıca görüntü işleme ve EEG sinyallerinin sınıflandırılmasında yapılan çalışmalar da literatürde yer almaktadır (Baba ve ark., 2001; Damiani ve ark., 2002, Peker ve Şen, 2013). Örneğin, kanal dengeleme problemi, kompleks değerli sinir ağı için uygun bir uygulama alanıdır. Sayısal bir iletişim sistemindeki kanal eşitleme, model sınıflandırma problemi olarak görülebilir. Sayısal iletişim sistemi, ilave gürültü ile iletilen bir sinyal dizisi alır ve gerçek iletilen diziyi tahmin etmeye çalışır. İletilen bir sinyal, dört olası kompleks değerden birini alabilir: $-1 -i$; $-1 + i$; $1 -i$ ve $1 + i$ ($i = \sqrt{-1}$). Bu nedenle, kompleks değerli sinir ağı bu alan için uygundur (Nitta, 2003a).

3.2.1. Kompleks değerli nöron modeli

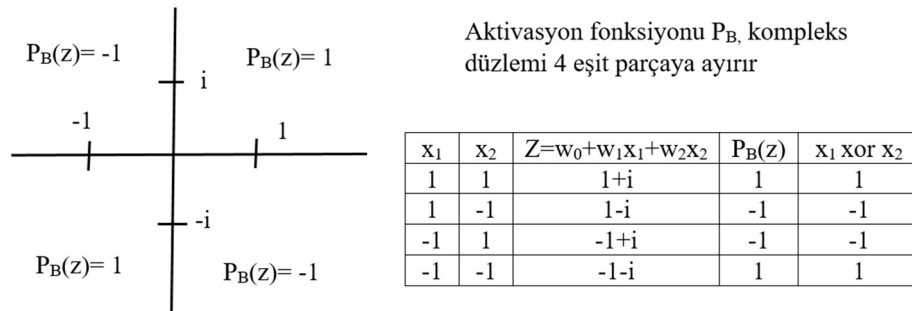
Kompleks değerli nöron modelinde temel amaç, nöronların işlevselliğini arttırmak ve öğrenmeyi daha basit hale getirmektir. Bilinen reel değerli nöron yalnızca doğrusal olarak ayrılabilen giriş-çıkış eşlemelerini öğrenebilir ve doğrusal olmayan ayrılabilir

giriş-çıkış eşlemelerini öğrenemez. Doğrusal olarak ayrılabilen OR ve doğrusal olarak ayrılabilen XOR problemi Şekil 3.6'da verilmiştir. Doğrusal olarak ayrılabilen OR kapısında düzlemde çizilen tek bir çizgi verileri ayrabilirken, doğrusal olmayan XOR kapısında veriler düzlemde tek bir çizgi ile ayrılamaz (Oppenheim, 1981).



Şekil 3.6. Doğrusal olarak ayrılabilen OR ve doğrusal olarak ayrılabilen XOR kapıları

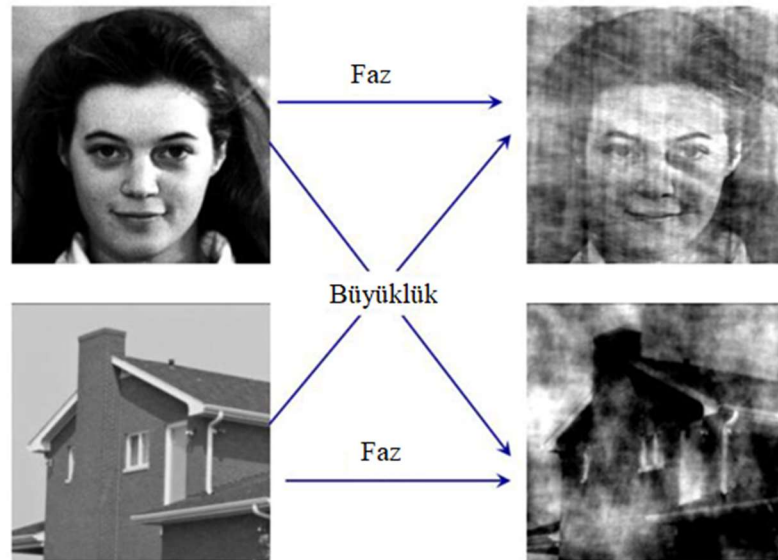
Gerçek değerli YSA kullanılarak doğrusal olarak ayrılabilen XOR fonksiyonunu öğrenmek için en azından üç nörondan oluşan bir ağ gereklidir. Fakat bu kompleks değerli nöronlar için geçerli değildir. Kompleks alana geçerek XOR problemi tek bir nöron ile öğrenilebilir. Şekil 3.7'de kompleks değerli nöron ile XOR probleminin çözümü gösterilmiştir. Kompleks değerli aktivasyon fonksiyonu (P_B) kullanan nöron düzlemi 4 eşit parça ayırarak XOR problemi çözmektedir.



Şekil 3.7. Kompleks değerli nöron ile XOR probleminin çözümü

Kompleks değerli nöronların en önemli avantajı, ağa giriş olarak sunulan sinyallerin ve görüntülerin faz bilgisi ile çalışabilme yeteneğidir. Gerçek değerli sinyallerin analizinde dahi, en verimli yaklaşımlardan biri, kompleks sayıları içeren frekans bölgesi analizidir. Frekans alanındaki sinyal analiz edilerek, her sinyalin, sinyal hakkında farklı bilgiler taşıyan büyüklük ve faz ile karakterize edildiği görülür. Oppenheim (1981) yayınladığı makalede sinyallerin Fourier spektrumundaki fazın, büyüklüğe göre çok daha bilgilendirici olduğunu ve görüntülerin Fourier spektrumunda faz bilgisinin şekil, kenar ve yönelim hakkında bilgiler içerdiğini göstermiştir. Yani Fourier Faz Spektrumunun, nesnelere tanımlayan objektif özelliklerin kaynağı olabileceğini belirtmiştir. Güç Spektrumu (büyüklük) ise görüntüler için bazı global özellikleri (bulanıklaştırma, gürültü, temizlik, kontrast, parlaklık, vb.) tanımlamaktadır (Oppenheim, 1981).

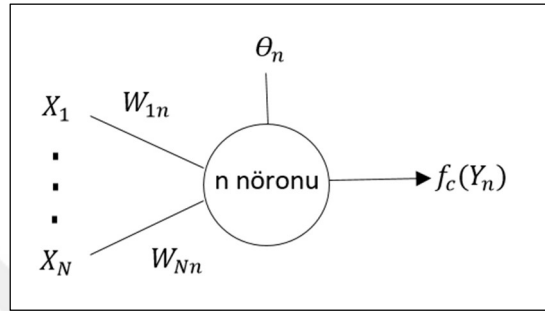
Aşağıda Şekil 3.8’de verilen görüntüde ise iki resmin faz ve büyüklük bilgileri üst üste getirilmiştir. Üstteki nesneye ait görüntünün faz bilgisi ile alttaki resme ait büyüklük bilgisi üst üste getirilince ortaya çıkan resimde üstteki resmin belirgin olduğu görülmüştür. Yine aynı şekilde üstteki nesneye ait büyüklük ve alttaki resme ait faz bilgisi üst üste çakıştırılınca ortaya çıkan resimde alttaki resmin belirgin olduğu görülür (Oppenheim, 1981).



Şekil 3.8. Faz bilgisi ve güç spektrumunun önemi

3.2.2. KDYSA ileri yayılım işlemleri

Kompleks değerli YSA yapısının, gerçek değerli YSA'dan en önemli farkı giriş ve çıkış değerlerinin, ağırlıkların ve eşik değerlerinin tamamının kompleks değerli sayılardan oluşmadır. Bu bölümde, kompleks geriye yayılım algoritmasının çok katmanlı KDYSA'da uygulaması açıklanmıştır. Şekil 3.9'da kompleks değerli nöron modeli verilmiştir.



Şekil 3.9. Kompleks değerli nöron modeli

Giriş değeri, ağırlığı ve eşik değeri kompleks olmak üzere Y_n nöronunun iç potansiyel değeri Eşitlik 3.3'te verildiği gibi tanımlanmıştır.

$$Y_n = \sum_N W_{nN} X_N + \Theta_n \quad (3.3)$$

W_{nN} 'nin, nöron n ve N'yi bağlayan ağırlığı, X_N nöron N'den gelen giriş sinyalini ve Θ_n ise nöronun eşik değerini göstermektedir. Her parametre reel ve imajiner kısma ayrılabilir. Y_n , Eşitlik 3.4'te verildiği gibi tarif edilebilir:

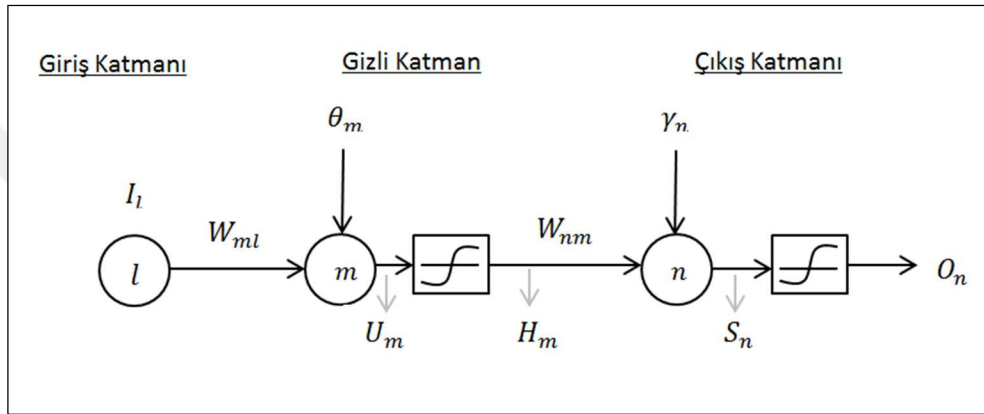
$$Y_n = x + iy \quad (3.4)$$

Nöron çıkış değerini hesaplamak için, iç potansiyel değerinin bir aktivasyon fonksiyonundan geçirilmesi gerekir. Aktivasyon fonksiyonu f_c olarak kabul edilirse, x nöronu için iç potansiyel değerinin reel kısmı, y nöronu için iç potansiyel değerinin imajiner kısmı olmak üzere kompleks değerli nöron çıkışı O_n , Eşitlik 3.5 kullanılarak tanımlanabilir.

$$O_n = f_c(Y_n) = f_R(x) + i * f_R(y) \quad (3.5)$$

Burada f_c kompleks değerli aktivasyon fonksiyonu, f_R ise gerçek değerli aktivasyon fonksiyonudur.

Tez çalışmasında giriş, gizli katman ve çıkış olmak üzere üç katmanlı ağ yapısı kullanılmıştır. Yapıyı anlatmak üzere kullanılan ifadeler aşağıda verilmiş olup YSA modeli Şekil 3.10'da gösterilmiştir (Chen, 2005).



Şekil 3.10. Giriş, gizli ve çıkış katmanına sahip kompleks değerli YSA modeli

- W_{ml} : giriş katman nöronu l ve gizli katman nöronu m arasındaki ağırlık
- W_{nm} : m gizli katman nöronu n ile çıkış katman nöronu n arasındaki ağırlık
- θ_m : m gizli katman nöronu için eşik değeri
- γ_n : n çıkış katman nöronu eşik değeri
- I_l : giriş katman nöronu l için çıkış değeri
- H_m : gizli katman nöronu m için çıkış değeri
- O_n : çıkış katman nöronu n için çıkış değeri
- U_m : gizli katman nöronu m 'nin aktif değeri
- S_n : çıkış katman nöronu n 'nin aktif değeri
- $\delta[n]$: Gerçek çıkış O_n ve hedef çıkış T_n arasında ki hata
- f_c : kompleks aktivasyon fonksiyonu

Gizli ve çıkış katmanlarının çıkış değerleri Eşitlik 3.6 ve 3.7 kullanılarak hesaplanabilmektedir.

$$U_m = \sum_l W_{ml} I_l + \Theta_m \quad (3.6)$$

$$S_n = \sum_m V_{nm} H_m + \gamma_n \quad (3.7)$$

Elde edilen çıkış değerleri kompleks aktivasyon fonksiyonundan geçirilerek gizli katman m nöronu için çıkış değeri H_m ve çıkış katman nöronu n için çıkış değeri O_n değeri Eşitlik 3.8 ve 3.9'da gösterildiği gibi hesaplanır (Chen, 2005).

$$H_m = f_c(U_m) \quad (3.8)$$

$$O_n = f_c(S_n) \quad (3.9)$$

Çıkış hatası $\delta[n]$ Eşitlik 3.10 kullanılarak tanımlanabilir:

$$\delta[n] = O_n - T_n \quad (3.10)$$

Formülde gösterilen T_n , hedeflenen çıkış değerini; O_n ise çıkış katman nöronu n için çıkış değerini ifade eder (Chen, 2005).

3.2.3. KDYSA geri yayılım işlemleri ve hata fonksiyonu

KDYSA'da hata gerçek değerli YSA'da olduğu gibi son katmandan ilk katmana kadar yayılır. Hata Eşitlik 3.11'de verildiği gibi tanımlanabilir.

$$E_p = (1/2) \sum_{n=1}^N |T_n - O_n|^2 \quad (3.11)$$

Formülde verilen N, çıkış katmanına ait nöron sayısını ifade etmektedir. Ağırlık ve eşik değer güncellemeleri Eşitlik 3.12, 3.13, 3.14 ve 3.15 kullanılarak hatanın geriye

doğru yayılması yoluyla gerçekleştirilebilmektedir. Formülde gösterilen ε değeri sıfırdan büyük olmak şartı ile seçilen küçük bir öğrenme sabitidir (Acar ve ark., 2018).

$$\Delta V_{nm} = -\varepsilon \cdot \frac{\partial E_p}{\partial \text{Re}[V_{nm}]} - i \cdot \varepsilon \cdot \frac{\partial E_p}{\partial \text{Im}[V_{nm}]} \quad (3.12)$$

$$\Delta \gamma_n = -\varepsilon \cdot \frac{\partial E_p}{\partial \text{Re}[\gamma_n]} - i \cdot \varepsilon \cdot \frac{\partial E_p}{\partial \text{Im}[\gamma_n]} \quad (3.13)$$

$$\Delta V_{ml} = -\varepsilon \cdot \frac{\partial E_p}{\partial \text{Re}[W_n]} - i \cdot \varepsilon \cdot \frac{E_p}{\partial \text{Im}[V_{ml}]} \quad (3.14)$$

$$\Delta \theta_m = \varepsilon \cdot \frac{\partial E_p}{\partial \text{Re}[\theta_m]} - i \cdot \varepsilon \cdot \frac{\partial E_p}{\partial \text{Im}[\theta V_m]} \quad (3.15)$$

Ağırlıkların ve eşik değerlerinin E_p hata değerine göre kısmi türevleri alındığında formüller aşağıdaki gibi yazılabilir.

$$\Delta V_{nm} = \overline{H_m} \Delta \gamma_n \quad (3.16)$$

$$\Delta \gamma_n = \varepsilon (\text{Re}[\delta^n] (1 - \text{Re}[O_n]) \text{Re}[O_n] + i \cdot \text{Im}[\delta^n] (1 - \text{Im}[O_n] \text{Im}[O_n])) \quad (3.17)$$

$$\Delta W_{m1} = \overline{I_1} \cdot \Delta \theta_m \quad (3.18)$$

$$\begin{aligned} \Delta \theta_m = \varepsilon \left[(1 - \text{Re}[H_m]) \text{Re}[H_m] x \sum_n (\text{Re}[\delta^n] (1 - \text{Re}[O_n]) \text{Re}[O_n] \text{Re}[V_{nm}] \right. \\ \left. + \text{Im}[\delta^n] (1 - \text{Im}[O_n] \text{Im}[V_{nm}])) \right] \\ - i \cdot \varepsilon \left[(1 - \text{Im}[H_m]) \text{Im}[H_m] x \sum_n (\text{Re}[\delta^n] (1 - \text{Re}[O_n]) \text{Re}[O_n] \text{Im}[V_{nm}] \right. \\ \left. - \text{Im}[\delta^n] (1 - \text{Im}[O_n] \text{Im}[O_n] \text{Re}[V_{nm}])) \right] \end{aligned} \quad (3.19)$$

KDYSA'da ağdan istenen değerler ile elde edilen çıkışlar arasındaki oran hesaplanarak ağın performansı ölçülür. Bu ölçüm için çeşitli yöntemler kullanılmakta olup, en yaygını Ortalama Karesel Hata (OKH)'dır. Tez çalışmasında performans kriteri olarak OKH kullanılmıştır. Ağın istenen çıkışı T_n ve ağın gerçek çıkışı O_n olmak üzere, OKH, Eşitlik 3.20'de ifade edilmiştir.

$$OKH = (1/N) \sum_{n=1}^N |T_n - O_n|^2 \quad (3.20)$$

3.2.4. Kompleks değerli aktivasyon fonksiyonu

Kompleks geri yayılım algoritmasında karşılaşılan başlıca güçlüklerden biri uygun aktivasyon fonksiyonunun seçimidir. Literatürde hangi aktivasyon fonksiyonunun seçilmesi gerektiği konusunda fikir birliği yoktur. Aktivasyon fonksiyonu seçimi, giriş eğitimi verisinin türüne ve durumuna bağlıdır (Peker ve ark., 2016).

Girişler, çıkışlar ve ağırlıklar gibi bütün parametrelerin kompleks olduğu KDYSA'da, seçilen aktivasyon fonksiyonunda kompleks düzlemde olması gerekir. Seçilecek olan kompleks aktivasyon fonksiyonu $\varphi(z)=u(x,y)+iv(x,y)$ 'nin aşağıdaki özellikleri karşılaması gerekmektedir (Georgiou ve Koutsougeras, 1992);

-Aktivasyon fonksiyonu $\varphi(z)$ 'in, girişin reel ve imajiner kısımları olan x ve y 'nin her ikisi de doğrusal olmamalıdır. Eğer doğrusal olursa, ağın yetenekleri ciddi şekilde sınırlandırılmış olur. Yani doğrusal olmayan ayrılabilir bir sınıflandırma sorunu gerçekleştirmek için ileri beslemeli bir sinir ağını eğitmek imkânsız olur.

- $\varphi(z)$ fonksiyonu sınırlı olmalıdır. Bu şart hem u hem de v sınırlı ise geçerlidir. Hem u hem de v eğitim için kullanıldığından, bunların sınırlandırılması gerekir. İkisinden biri sınırlandırılmıyorsa, yazılımda hata meydana gelebilir. Algoritma donanımda uygulanırsa benzer sonuçlar görülebilir.

- Kısmi türevler u_x, u_y, v_x ve v_y mevcut ve sınırlı olmalıdır. Öğrenim aşaması, çok katmanlı perseptronların kompleks ağırlıklarının kısmi türevlere orantılı olarak günceller. Bu sebeple, kısmi türevlerin sınırlı olmasına ihtiyaç duyulur.

- $\varphi(z)$ fonksiyonu, kompleks düzlemin her yerinde analitik olan kompleks bir fonksiyon gibi tanımlanmış olmalıdır.

- Kısmi türevler $u_R.v_I = u_I.v_R$ şartını sağlamıyorsa, $\phi(z)$ fonksiyonu aşağıdaki Eşitlik 3.21'de belirtilen durumlar haricinde uygun bir aktivasyon fonksiyonu değildir.

$$\begin{aligned} u_x = v_x = 0 \text{ ve } u_y \neq 0, v_y \neq 0 \\ u_y = v_y = 0 \text{ ve } u_x \neq 0, v_x \neq 0 \end{aligned} \quad (3.21)$$

3.3. Kompleks Değerli Yeni Aktivasyon Fonksiyonlarının Tanımlanması

3.3.1. Kompleks değerli swish aktivasyon fonksiyonu

Swish fonksiyonu, Google araştırmacıları Ramachandra ve ark. (2017) tarafından tanımlanan bir aktivasyon fonksiyonudur. Ramachandra ve ark. (2017) yaptıkları çalışmalarda Swish'in bir dizi zorlu veri setindeki daha derin modellerde, en başarılı ve en çok kullanılan aktivasyon fonksiyonu ReLB'dan daha iyi çalışma eğiliminde olduğunu göstermişlerdir.

Ramachandra ve ark. (2017) ImageNet veri seti üzerinde bulunan Mobile NASNet-A ve Inception-ResNet-v2 verileri üzerinde yaptıkları çalışmalar sonucunda, aktivasyon fonksiyonu olarak ReLB yerine Swish kullanıldığında sınıflandırma doğruluğunun; Mobile NASNet-A için %0,9 ve Inception-ResNet-v2 için %0,6 arttığını göstermişlerdir.

Swish fonksiyonu Eşitlik 3.22'de gösterildiği gibi girdiler ile sigmoid fonksiyonunun çarpımı olarak tanımlanır;

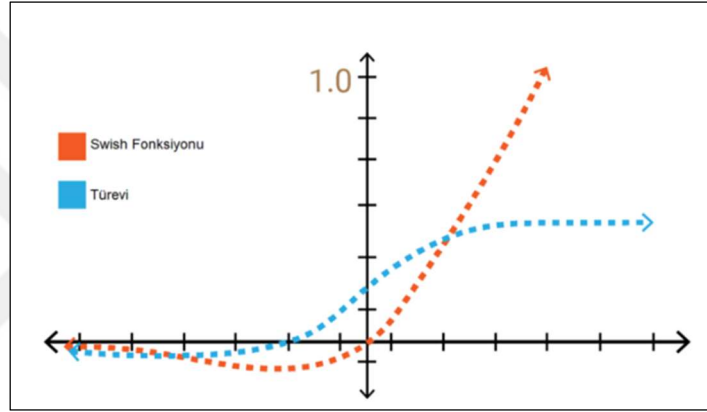
$$y = x \cdot \text{sigmoid}(x) \quad (3.22)$$

Literatürde sıklıkla kullanılan ve Şekil 3.2'de verilen sigmoid fonksiyonunun uç noktalarına dikkatlice bakarsak, y değerleri x'teki değişikliklere çok az tepki vermektedir. Yani bu bölgelerde türev değerleri çok küçük olur ve 0'a yakınsar. Buna gradyanların ölmesi/kaybolması (vanishing gradient) denir ve öğrenme olayı minimum düzeyde gerçekleşir. Sigmoid fonksiyonu kaybolan gradyan probleminden dolayı eski bir fonksiyon haline gelmiştir. Küçük bir değişiklik ise bu eski etkinleştirme işlevini yeniden önemli kılmıştır. Sigmoid fonksiyonu giriş değeri ile çarpılarak daha etkin hale

getirilmiştir. Sigomid fonksiyonunun matematiksel ifadesi yerine yazılacak olursa Eşitlik 3.23 elde edilir.

$$y = x \left(\frac{1}{1 + e^{-x}} \right) = \frac{x}{1 + e^{-x}} \quad (3.23)$$

Swish fonksiyonunun, literatürde sıklıkla kullanılan ReLB aktivasyon fonksiyonundan farkı, negatif bölgede değer almasıdır. Negatif bölgelerde de değer alan Leaky ReLB'den farkı ise negatif bölgede doğrusal olmamasıdır. Swish aktivasyon fonksiyonu ve türevine ait grafik Şekil 3.11'de verilmiştir.



Şekil 3.11. Swish fonksiyonu ve türevi

Çoğu aktivasyon fonksiyonu (sigmoid, tanh ve ReLB gibi) monotoniktir, yani giriş arttıkça değerleri asla düşmez ($x < 0$ olduğunda ReLB'lar için olduğu gibi değer aynı kalabilir). Fakat bu durum swish için geçerli değildir. Swish fonksiyonun çıktısı girdi arttığında bile düşebilir. Örneğin negatif bölgede giriş arttığında çıkış azalmaktadır.

Geriye yayılım algoritması kullanan YSA'larda seçilecek olan aktivasyon fonksiyonunun türevi alınabilmelidir. Çünkü modelin eğitimi ve geriye yayılımı esnasında eğrinin hangi yönde ne kadar değiştireceğinin hesaplanması gerekmektedir. Bu nedenle önerilen swish fonksiyonunun türevlenebilir bir fonksiyon olması gerekir. σ , sigmoid fonksiyonu olmak üzere Eşitlik 3.22'de verilen swish fonksiyonu aşağıdaki gibi formüle edilir.

$$y = x \cdot \sigma(x) \quad (3.24)$$

y fonksiyonunun türevi Eşitlik 3.25'teki gibi formüle edilir.

$$y' = x'.\sigma(x) + x.\sigma(x)' \quad (3.25)$$

Sigmoid fonksiyonunun türevi aşağıdaki gibi olduğuna göre;

$$\sigma(x)' = \sigma(x).(1-\sigma(x)) \quad (3.26)$$

Buradan Eşitlik 3.26 kullanılarak Eşitlik 3.25 yeniden yazılabilir.

$$\begin{aligned} y' &= \sigma(x) + x.(\sigma(x).(1-\sigma(x))) \\ y' &= \sigma(x) + x.\sigma(x) - x.\sigma^2(x) \\ y' &= x.\sigma(x) + \sigma(x) - x.\sigma^2(x) \\ y' &= x.\sigma(x) + \sigma(x)(1-x.\sigma(x)) \end{aligned} \quad (3.27)$$

Swish fonksiyonunun $y = x.\sigma(x)$ olarak tanımlandığı göz önünde bulundurulduğunda, Eşitlik 3.28 elde edilir.

$$y' = y + \sigma(x)(1-y) = \frac{x.e^{-x}(x+1) + 1}{(1+e^{-x})^2} \quad (3.28)$$

Tez çalışmasında, swish aktivasyon fonksiyonu kompleks düzlemde ele alınmıştır. Kompleks sayı olan z giriş olmak üzere, swish KDYSA'ya Eşitlik 3.29'da gösterildiği gibi uygulanmıştır (Çelebi ve Ceylan, 2019).

$$y = swish(Re(z)) + i.swish(Im(z)) \quad (3.29)$$

3.3.2. Kompleks değerli modifiye swish aktivasyon fonksiyonu

Modifiye swish aktivasyon fonksiyonu Ramachandra ve ark. (2017) tarafından swish fonksiyonunun performansını arttırmak için önerilmiştir. Modifiye swish aktivasyon fonksiyonu Eşitlik 3.30'daki gibi tanımlanmaktadır.

$$y = x. \sigma(\alpha x) \quad (3.30)$$

Modifiye swish aktivasyon fonksiyonunun swish fonksiyonundan farkı, giriş değerine α sabitinin eklenmesidir. Modifiye swish aktivasyon fonksiyonunun türevi Eşitlik 3.31’de verilmiştir.

$$y' = \alpha y + \sigma(\alpha x)(1 - \alpha y) \quad (3.31)$$

Modifiye swish aktivasyon fonksiyonu tez çalışmasında KDYSA’da kullanılmıştır. KDYSA’da tüm parametreler kompleks sayılardan oluştuğu için, girişin ve aktivasyon fonksiyonunun kompleks sayı olması gerekmektedir. Kompleks sayı olan z giriş olmak üzere, modifiye swish fonksiyonu z ’nin reel ve imajiner kısmına Eşitlik 3.32’de gösterildiği gibi uygulanmıştır (Çelebi ve Ceylan, 2020).

$$y = \text{Modifiyeswish}(Re(z)) + i. \text{Modifiyeswish}(Im(z)) \quad (3.32)$$

3.3.3. Kompleks değerli e-swish aktivasyon fonksiyonu

E-swish aktivasyon fonksiyonu Alcaide (2018) tarafından önerilmiştir. Alcaide, swish aktivasyon fonksiyonu içerisine β sabitini ilave ederek yaptığı çalışmalarda e-swish’in, Swish aktivasyon fonksiyonundan daha iyi performans elde ettiğini göstermiştir. E-swish aktivasyon fonksiyonu Eşitlik 3.33’te verilmiştir.

$$y = \beta x. \sigma(x) \quad (3.33)$$

E-swish fonksiyonunun türevi Eşitlik 3.34’te verilmiştir.

$$y' = y + \sigma(x)(\beta - y) \quad (3.34)$$

E-swish aktivasyon fonksiyonu swish fonksiyonu ile çok benzerdir çünkü sabit β değeri 1 olarak alındığında e-swish aktivasyon fonksiyonu, swish aktivasyon fonksiyonu ile aynı olur. Sabit değer olan β ’ya istenilen değer verilerek, daha iyi performans elde edebilmek adına çalışmalar yapılabilir (Alcaide, 2018).

E-swish fonksiyonu kompleks düzleme Eşitlik 3.35'teki gibi taşınmıştır. Burada z kompleks sayıya sahip girişi ifade etmektedir (Çelebi ve Ceylan, 2020).

$$y = Eswish(Re(z)) + i.Eswish(Im(z)) \quad (3.35)$$

3.3.4. Kompleks değerli düzleştirilmiş t-swish aktivasyon fonksiyonu

DTS olarak bilinen bu yeni aktivasyon fonksiyonu Chieng ve ark. (2018) tarafından önerilmiştir. DTS, ReLB ve swish aktivasyonlarının birleşimi olarak da düşünülebilir. Chieng ve ark. (2018) yaptıkları çalışmalarda swish ve ReLB'i birleştirerek elde ettikleri yeni fonksiyon DTS'in, swish aktivasyon fonksiyonundan daha iyi performans elde ettiğini göstermişlerdir. DTS Eşitlik 3.36'daki gibi ifade edilir.

$$DTS(x) = \begin{cases} \frac{x}{1 + e^{-x}} + T & , \quad x \geq 0 \\ T & , \quad x < 0 \end{cases} \quad (3.36)$$

Yukarıda formülde belirtilen T sabit değeri, eşik değeri olarak isimlendirilir. Girişin negatif olduğu kısımlarda ağ, T sabit değerini verirken, pozitif kısımlarda $swish(x)+T$ değerini verir. YSA, ağırlık ve biasların güncellenmesi için gradyan iniş metodu kullanacağından, formülün türevlenebilir olması gerekmektedir. DTS aktivasyon fonksiyonunun, pozitif değerlerdeki türevi Eşitlik 3.37'de verilmiştir (Chieng ve ark., 2018).

$$y' = \sigma(x)(1 - y) + y \quad (3.37)$$

Formülde belirtilen T sabit bir değer olduğu için türevi 0 olmaktadır yani fonksiyon negatif bölgede sıfır değerini üretir. DTS fonksiyonunun türevi Eşitlik 3.38'de belirtildiği şekilde kullanılır.

$$DTS'(x) = \begin{cases} \sigma(x)(1 - y) + y & , \quad x \geq 0 \\ 0 & , \quad x < 0 \end{cases} \quad (3.38)$$

DTS fonksiyonunu kompleks değerli girişi (z) ile KDYSA'da kullanabilmek için kompleks düzleme Eşitlik 3.39'da gösterildiği gibi taşınmıştır.

$$y = DTS(Re(z)) + i.DTS(Im(z)) \quad (3.39)$$

3.4. Kompleks Değerli Veriler

Bu bölümde tez çalışmasında önerilen aktivasyon fonksiyonlarının (kompleks swish, e-swish, modifiye swish ve düzleştirilmiş t-swish) başarımını belirlemek için kullanılan; 4 bit XOR, 3 giriş 1 çıkışlı simetri ve kanal dengeleme problemleri ele alınmıştır.

KDYSA'da tüm parametreler (çıkışlar, ağırlıklar, aktivasyon fonksiyonu) kompleks düzlemde olduğu için girişlerinde kompleks düzlemde olması gerekir. Bu yüzden bu bölümde gerçek düzlemde bulunan 4 bit XOR ve 3 giriş 1 çıkışlı simetri problemlerinin kompleks düzleme taşınması ele alınmıştır.

3.4.1. Kompleks değerli XOR problemi

XOR problemi sahip olduğu doğrusal olmayan yapı sayesinde sınıflama algoritmalarının test edilmesi ve başarılarının ölçülmesinde önemli bir ölçüt haline gelmiştir. XOR problemine ait giriş çıkış eşleşmesi Çizelge 3.1'de verilmiştir. Çizelgeden de anlaşılacağı gibi veriler düzlemde tek bir çizgi ile ayrılamaz.

Çizelge 3.1. XOR problemi giriş çıkış eşleşmesi

Girişler		Çıkış
X_1	X_2	Y
0	0	0
0	1	1
1	0	1
1	1	0

Yukarıda çizelgede belirtilen X_1 ve X_2 girişler olup Y ise çıkışı göstermektedir. Girişlerin aynı olması durumunda çıkış 0, girişlerin farklı olması durumunda ise çıkış 1 olmaktadır.

KDYSA'da tüm parametreler kompleks sayılardan oluştuğu için girişlerinde kompleks sayı olması gerekmektedir. Kompleks düzleme kodlanmış XOR problemi giriş çıkış eşleşmesi Çizelge 3.2'de verilmiştir.

Çizelge 3.2. Kompleks değerli XOR problemi giriş çıkış eşleşmesi

Giriş	Çıkış
X_1	Y
0	0
i	1
1	$1+i$
$1+i$	i

Kompleks düzlemdeki XOR problemi aşağıda belirtilen iki madde temel alınarak kodlanmıştır (Nitta, 2008).

- Çıkışın reel kısmı, girişin reel ve imajiner kısmının XOR'una eşittir.
- Çıkışın imajiner kısmı, girişin reel kısmına eşittir.

3.4.2. Kompleks değerli simetri problemi

Simetri belirleme problemi, ikili verilerin simetrik olup olmadığını tespit etmeyi amaçlar. Bit sayısı arttıkça simetrik olma olasılığı yarı yarıya azalır. Bu yüzden simetri tespit problemi dengesiz verileri araştırmak için çok uygun bir problemdir. Çizelge 3.3'te gerçek düzlemde üç giriş ve bir çıkışlı simetri algılama problemine ait giriş çıkış eşleşmesi verilmiştir (Chen ve Tang, 2005).

Çizelge 3.3. Simetri problemi giriş çıkış eşleşmesi

Girişler			Çıkış
X_1	X_2	X_3	Y
0	0	0	1
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	0
1	0	1	1
1	1	0	0
1	1	1	1

Gerçek değerli problemlerinin KDYSA ile çözülmesi istendiğinde, gerçek değerli girdi verilerinin kompleks değerli verilere dönüştürülmesi gerekir. Bu dönüştürme işlemi basit bir açı tabanlı kodlama ile yapılır. a ve b gerçek sayılar olmak üzere giriş verisi $[a,b]$ aralığında değişiyorsa veri kompleks düzleme Eşitlik 3.40 kullanılarak taşınır (Acar ve ark., 2018).

$$\varphi = \frac{\pi(x - a)}{b - a} \quad (3.40)$$

Hesaplanan φ faz açısı ile reel düzlemdeki sayılar, Eşitlik 3.41 kullanılarak kompleks düzleme taşınmış olur.

$$Z = e^{i\varphi} = \cos\varphi + i\sin\varphi \quad (3.41)$$

Yukarıda verilen eşitlikler kullanılarak gerçek değerli simetri problemi kompleks düzleme taşınarak, KDYSA’da kullanılmaya hazır hale getirilmiştir. Tez çalışmasında a ve b , giriş verileri 0 ile 1 arasında değiştiği için sırasıyla 0 ve 1 olarak alınmıştır. Veriler $\varphi = \pi/4$ faz açısına sahip olacak şekilde kompleks düzleme taşınmıştır. Elde edilen üç giriş ve bir çıkışlı kompleks değerli simetri problemine ait giriş çıkış eşleşmesi Çizelge 3.4’te verilmiştir.

Çizelge 3.4. Kompleks değerli simetri problemi giriş çıkış eşleşmesi

Girişler			Çıkış
X_1	X_2	X_3	Y
1	1	1	$0.7+0.7i$
1	1	$0.7+0.7i$	1
1	$0.7+0.7i$	1	$0.7+0.7i$
1	$0.7+0.7i$	$0.7+0.7i$	1
$0.7+0.7i$	1	1	1
$0.7+0.7i$	1	$0.7+0.7i$	$0.7+0.7i$
$0.7+0.7i$	$0.7+0.7i$	1	1
$0.7+0.7i$	$0.7+0.7i$	$0.7+0.7i$	$0.7+0.7i$

3.4.3. Kanal dengeleme problemi

Sayısal iletişim sistemindeki kanal dengeleme problemi, örüntü sınıflandırma problemi olarak görülebilir. Sayısal iletişim sistemi, ilave gürültü ile iletilen bir sinyal dizisini alır ve bu sinyallerden gerçek iletilen diziyi tahmin etmeye çalışır. İletilen bir sinyal aşağıdaki dört kompleks değerden birini alabilir: $-1-i$, $-1+i$, $1-i$ ve $1+i$ ($i = \sqrt{-1}$). Böylece, alınan sinyal $-1-i$, $-1+i$, $1-i$ ve $1+i$ civarında değer alacaktır, çünkü iletim esnasında bazı gürültü sinyalleri eklenmiştir. Gürültü ile bu kompleks değerlerden gerçek kompleks değerlerin tahmin edilmesi gerekmektedir. Dolayısıyla, tahmin için genelleme kabiliyetine sahip bir yönteme ihtiyaç duyulmaktadır. Problemdeki giriş-çıkış eşleşmesi Çizelge 3.5'te gösterilmiştir (Nitta, 2004).

Çizelge 3.5. Kanal dengeleme problemine ait giriş çıkış eşleşmesi

Giriş	Çıkış
X_1	Y
$-1 - i$	$-1 - i$
$-1 + i$	$-1 + i$
$1 - i$	$1 - i$
$1 + i$	$1 + i$

KDYSA için kodlanmış kanal dengeleme problemi giriş çıkış eşleşmesi Çizelge 3.6'da verilmiştir (Nitta, 2003b).

Çizelge 3.6. KDYSA için kodlanmış kanal dengeleme problemi giriş çıkış eşleşmesi

Giriş	Çıkış
X_1	Y
$-1 - i$	0
$-1 + i$	i
$1 - i$	1
$1 + i$	$1 + i$

4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA

Bu bölümde ilk olarak önerilen swish aktivasyon fonksiyonu literatürde sıklıkla kullanılan sigmoid ve hiperbolik tanjant fonksiyonları ile karşılaştırılmıştır. Sonraki bölümde swish fonksiyonu üzerinden türetilen modifiye swish, e-swish ve DTS aktivasyon fonksiyonlarının en iyi performans gösterdiği değerler bulunmuştur. Son olarak önerilen swish, modifiye swish, e-swish ve DTS aktivasyon fonksiyonlarının performansları birbirleri ile karşılaştırılmıştır.

4.1. Kompleks Swish Aktivasyon Fonksiyonun Diğer Aktivasyon Fonksiyonlar ile Karşılaştırılması

Tez çalışmasında kompleks swish aktivasyon fonksiyonu ilk olarak kompleks (hiperbolik tanjant ve sigmoid) aktivasyon fonksiyonları ile karşılaştırılmıştır. Karşılaştırma işlemi 4 bitlik kompleks XOR, 3 giriş 1 çıkışlı kompleks simetri ve kanal dengeleme problemleri üzerinde yapılmıştır. Yapılan testler ve sonuçları karşılaştırmalı olarak sunulmuştur.

4.1.1. Kompleks değerli XOR problemi

Önerilen kompleks swish aktivasyon fonksiyonunu kullanan ağ, ilk olarak kompleks değerli 4 bit XOR problemi çözümünde kullanılmıştır. Ayrıca karşılaştırma yapabilmek için kompleks hiperbolik tanjant ve sigmoid aktivasyon fonksiyonlarını kullanan ağlar içinde aynı problem çözülmüştür. Her bir aktivasyon fonksiyonunu kullanan ağın sonuçları karşılaştırmalı olarak sunulmuştur.

Kompleks değerli 4 bit XOR probleminin çözümünde literatürdeki diğer çalışmalar (Nitta, 2013a; Hadim ve Rabha, 2011; Enomoto, 2019; Ceylan, 2011; Gürüler ve Peker, 2015) ile adil bir karşılaştırma yapabilmek için, bir giriş bir çıkış ve gizli katman içinde iki düğüme sahip KDSYA mimarisi kullanılmıştır. Bu mimaride öğrenme oranı 0.5 olarak alınmıştır.

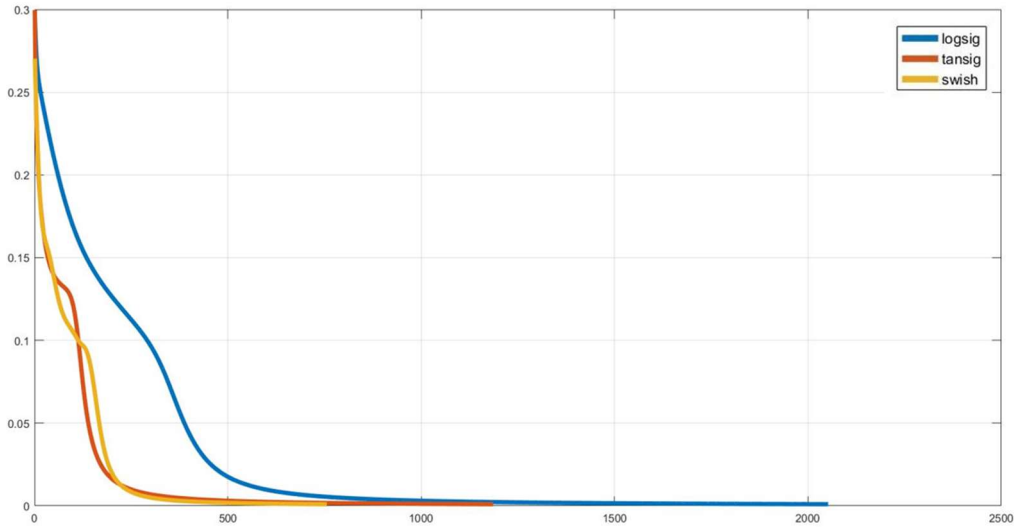
Oluşturulan KDYSA, 0.001 hata (OKH) değerine ulaştığı anda durdurulmuş ve bu hata oranına ulaştığı bütün iterasyonlara ait hata değerleri kaydedilmiştir. Hata oranı OKH olarak hesaplanmış olup formülü Eşitlik 3.11'de verilmiştir.

Yapılan testler sonucunda kompleks swish, hiperbolik tanjant ve sigmoid aktivasyon fonksiyonlarını kullanan ağların 0.1, 0.01 ve 0.001 hata oranına (OKH) ulaştığı iterasyonlar ve bu iterasyonlarda ki ağın gerçek çıkışları ve hedefleri Çizelge 4.1’de verilmiştir.

Çizelge 4.1. XOR problemi için test sonuçları (logsig, tansig, swish)

AKTİVASYON FONKSİYONU: LOGSIG	0.1 (OKH) Hata oranına ulaşılan iterasyon:		0.01 (OKH) Hata oranına ulaşılan iterasyon:		0.001 (OKH) Hata oranına ulaşılan iterasyon:	
	292		597		2052	
	Hedef	Çıkış	Hedef	Çıkış	Hedef	Çıkış
	0	$0.4506 + 0.0652i$	0	$0.1520 + 0.0119i$	0	$0.0489 + 0.0036i$
	1	$0.3706 + 0.1643i$	1	$0.8372 + 0.0607i$	1	$0.9497 + 0.0190i$
	$1+i$	$0.8130 + 0.8601i$	$1+i$	$0.9284 + 0.9166i$	$1+i$	$0.9756 + 0.9715i$
i	$0.3180 + 0.8873i$	i	$0.1049 + 0.9471i$	i	$0.0314 + 0.9824i$	
AKTİVASYON FONKSİYONU: TANSIG	0.1 (OKH) Hata oranına ulaşılan iterasyon:		0.01 (OKH) Hata oranına ulaşılan iterasyon:		0.001 (OKH) Hata oranına ulaşılan iterasyon:	
	114		250		1186	
	Hedef	Çıkış	Hedef	Çıkış	Hedef	Çıkış
	0	$0.4633 + 0.0999i$	0	$0.1237 + 0.0118i$	0	$0.0378 + 0.0241i$
	1	$0.4559 + 0.2414i$	1	$0.8462 + 0.0881i$	1	$0.9510 + 0.0258i$
	$1+i$	$0.7424 + 0.7032i$	$1+i$	$0.9350 + 0.8999i$	$1+i$	$0.9812 + 0.9692i$
i	$0.2026 + 0.8719i$	i	$0.0630 + 0.9047i$	i	$0.0241 + 0.9678i$	
AKTİVASYON FONKSİYONU: SWISH	0.1 (OKH) Hata oranına ulaşılan iterasyon:		0.01 (OKH) Hata oranına ulaşılan iterasyon:		0.001 (OKH) Hata oranına ulaşılan iterasyon:	
	114		238		756	
	Hedef	Çıkış	Hedef	Çıkış	Hedef	Çıkış
	0	$0.2935 + 0.2724i$	0	$0.1585 + 0.1026i$	0	$0.0526 + 0.0352i$
	1	$0.3761 + 0.0911i$	1	$0.8567 + 0.0003i$	1	$0.9582 + 0.0000i$
	$1+i$	$0.8546 + 1.0000i$	$1+i$	$0.9210 + 1.0000i$	$1+i$	$0.9742 + 1.0000i$
i	$0.3967 + 0.7515i$	i	$0.1094 + 0.9288i$	i	$0.0332 + 0.9782i$	

Ayrıca her iterasyon için hata oranı kaydedilmiş olup, Şekil 4.1’de gösterilmiştir. Önerilen KDYSA, 0.001 hata oranına ulaştığı anda ağ durdurulmuştur.



Şekil 4.1. XOR problemi için hata grafiği (logsig, tansig, swish)

Yapılan testler sonucunda swish aktivasyon fonksiyonunu kullanan ağın diğer aktivasyon fonksiyonlarını kullanan ağlara göre hedef değere daha erken yakınsadığı görülmüştür. Gizli katmanında logaritmik sigmoid aktivasyon fonksiyonu kullanan ağ 0.001 OKH hata oranına 2052. iterasyonda ulaşmış olup, tanjant sigmoid kullanan ağ 1186., swish fonksiyonu kullanan ağ ise 756. iterasyonda ulaşmıştır.

4.1.2. Kompleks değerli simetri problemi

Kompleks swish aktivasyon fonksiyonunu kullanan ağ, kompleks simetri probleminin çözümünde kullanılmıştır. Seçilen simetri problemi 3 giriş ve 1 çıkışa sahiptir.

Kompleks simetri probleminin çözümünde literatürdeki diğer çalışmalar (Nitta, 2013a; Chen ve ark., 2005) ile adil bir karşılaştırma yapabilmek için, bir giriş üç çıkış ve iki gizli katmana sahip KDSYA kullanılmış olup öğrenme oranı bütün mimarilerde 0.5 olarak alınmıştır.

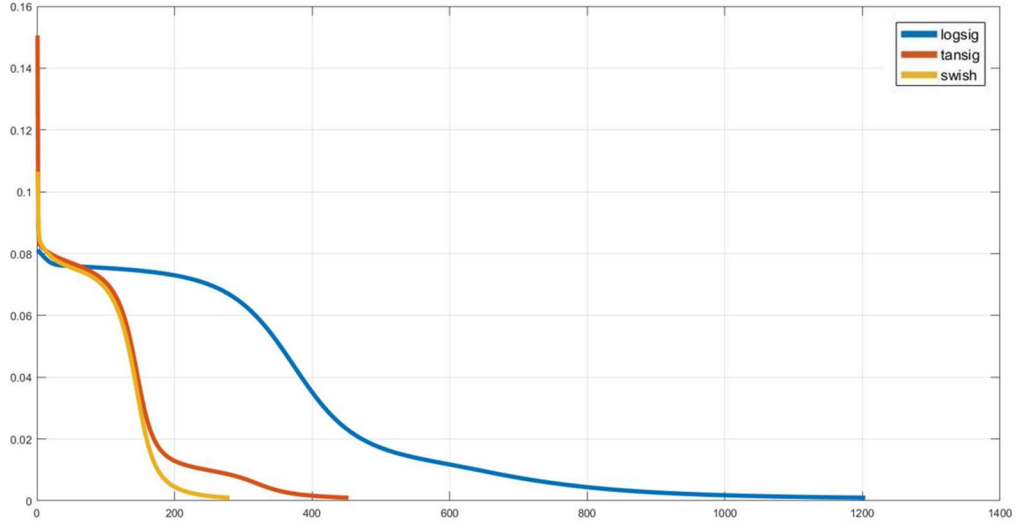
Tasarlanan ağlar 0.001 hata (OKH) oranına ulaştığı anda durdurulmuş ve bu hata oranına ulaştığı bütün iterasyonlara ait hata değerleri kaydedilmiştir.

Simetri probleminin çözümünde kullanılan KDYSYA'larının (kompleks swish, hiperbolik tanjant ve sigmoid aktivasyon fonksiyonlarını kullanan), sırasıyla 0.1, 0.01 ve 0.001 hata (OKH) oranında ulaştığı iterasyonlar ve bu iterasyonlarda ki ağın çıkışları ve hedefleri Çizelge 4.2'de verilmiştir.

Çizelge 4.2. Simetri problemi için test sonuçları (logsig, tansig, swish)

AKTİVASYON FONKSİYONU: LOGSIG	0.1 (OKH) Hata oranına ulaşılan iterasyon:		0.01 (OKH) Hata oranına ulaşılan iterasyon:		0.001 (OKH) Hata oranına ulaşılan iterasyon:	
	2		641		1204	
	Hedef	Çıkış	Hedef	Çıkış	Hedef	Çıkış
	$0.7+0.7i$	$0.96+0.34i$	$0.7+0.7i$	$0.93+0.67i$	$0.7+0.7i$	$0.72+0.69i$
	1	$0.96+0.34i$	1	$0.92+0.09i$	1	$0.95+0.03i$
	$0.7+0.7i$	$0.94+0.35i$	$0.7+0.7i$	$0.90+0.66i$	$0.7+0.7i$	$0.71+0.70i$
	1	$0.94+0.35i$	1	$0.89+0.09i$	1	$0.94+0.04i$
	1	$0.94+0.33i$	1	$0.93+0.09i$	1	$0.95+0.03i$
	$0.7+0.7i$	$0.94+0.29i$	$0.7+0.7i$	$0.72+0.66i$	$0.7+0.7i$	$0.71+0.70i$
	1	$0.92+0.34i$	1	$0.91+0.08i$	1	$0.95+0.03i$
$0.7+0.7i$	$0.92+0.31i$	$0.7+0.7i$	$0.67+0.65i$	$0.7+0.7i$	$0.71+0.69i$	
AKTİVASYON FONKSİYONU: TANSIG	0.1 (OKH) Hata oranına ulaşılan iterasyon:		0.01 (OKH) Hata oranına ulaşılan iterasyon:		0.001 (OKH) Hata oranına ulaşılan iterasyon:	
	2		249		453	
	Hedef	Çıkış	Hedef	Çıkış	Hedef	Çıkış
	$0.7+0.7i$	$0.92+0.53i$	$0.7+0.7i$	$0.93+0.70i$	$0.7+0.7i$	$0.74+0.72i$
	1	$0.94+0.50i$	1	$0.93+0.08i$	1	$0.95+0.03i$
	$0.7+0.7i$	$0.85+0.54i$	$0.7+0.7i$	$0.93+0.19i$	$0.7+0.7i$	$0.70+0.68i$
	1	$0.87+0.50i$	1	$0.32+0.66i$	1	$0.95+0.03i$
	1	$0.85+0.43i$	1	$0.74+0.07i$	1	$0.95+0.03i$
	$0.7+0.7i$	$0.92+0.35i$	$0.7+0.7i$	$0.71+0.64i$	$0.7+0.7i$	$0.72+0.69i$
	1	$0.75+0.45i$	1	$0.93+0.07i$	1	$0.95+0.03i$
$0.7+0.7i$	$0.87+0.38i$	$0.7+0.7i$	$0.71+0.66i$	$0.7+0.7i$	$0.72+0.69i$	
AKTİVASYON FONKSİYONU: SWISH	0.1 (OKH) Hata oranına ulaşılan iterasyon:		0.01 (OKH) Hata oranına ulaşılan iterasyon:		0.001 (OKH) Hata oranına ulaşılan iterasyon:	
	2		176		280	
	Hedef	Çıkış	Hedef	Çıkış	Hedef	Çıkış
	$0.7+0.7i$	$0.92+0.56i$	$0.7+0.7i$	$0.82+0.60i$	$0.7+0.7i$	$0.73+0.70i$
	1	$0.91+0.46i$	1	$0.90+0.10i$	1	$0.95+0.03i$
	$0.7+0.7i$	$0.86+0.58i$	$0.7+0.7i$	$0.80+0.63i$	$0.7+0.7i$	$0.72+0.69i$
	1	$0.84+0.48i$	1	$0.89+0.13i$	1	$0.95+0.03i$
	1	$0.86+0.47i$	1	$0.87+0.13i$	1	$0.94+0.03i$
	$0.7+0.7i$	$0.86+0.27i$	$0.7+0.7i$	$0.75+0.62i$	$0.7+0.7i$	$0.71+0.69i$
	1	$0.81+0.51i$	1	$0.88+0.09i$	1	$0.94+0.02i$
$0.7+0.7i$	$0.81+0.33i$	$0.7+0.7i$	$0.74+0.64i$	$0.7+0.7i$	$0.70+0.69i$	

Ağların 0.001 (OKH) hata oranına ulaştığı ana kadar ki hata değişim grafiği Şekil 4.2’de verilmiştir.



Şekil 4.2. Simetri problemi için hata grafiği (logsig, tansig, swish)

Simetri problemi çözümü için gerçekleştirilen test sonuçları incelendiğinde, swish aktivasyon fonksiyonu kullanan KDYSA'nın, kompleks hiperbolik tanjant ve sigmoid aktivasyon fonksiyonlarını kullanan KDYSA'lara göre sonuca daha erken yakınsadığı görülmüştür. Swish aktivasyon fonksiyonunu kullanan ağ 280. iterasyonda 0.001 hata oranına ulaşırken, hiperbolik tanjant aktivasyon fonksiyonunu kullanan ağ 453. iterasyonda, kompleks sigmoid aktivasyon fonksiyonunu kullanan ağ ise 1204. iterasyonda ulaşmaktadır.

4.1.3. Kanal dengeleme problemi

Son olarak kompleks swish aktivasyon fonksiyonunu kullanan ağ, kompleks hiperbolik tanjant ve sigmoid aktivasyon fonksiyonlarını kullanan KDYSA ile kanal dengeleme problemi çözümünde karşılaştırılmıştır.

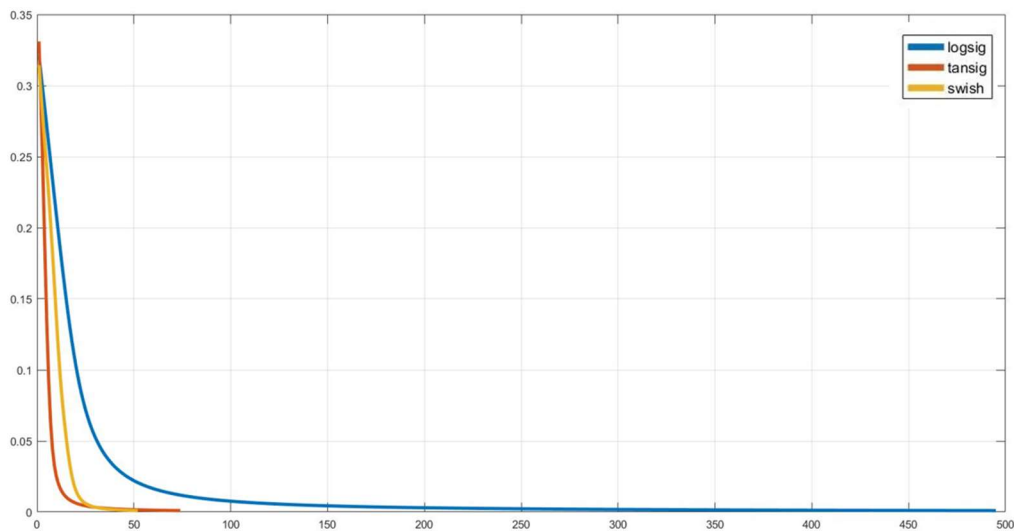
Kanal dengeleme probleminin çözümünde, bir giriş, bir çıkış ve gizli katman içinde iki düğüme sahip KDSYA mimarisi kullanılmış olup bütün ağlarda öğrenme oranı 0.5 olarak alınmıştır. Tasarlanan KDYSA 0.001 hata oranına (OKH) ulaştığı anda durdurulmuş ve bu hata oranına ulaştığı bütün iterasyonlara ait hata değerleri kaydedilmiştir.

Yapılan testler sonucunda ağların sırasıyla 0.1, 0.01 ve 0.001 hata oranına (OKH) ulaştığı iterasyonlar ve bu iterasyonlara ait ağın hedef ve gerçek çıkışları Çizelge 4.3'te verilmiştir.

Çizelge 4.3. Kanal dengeleme problemi için test sonuçları (logsig, tansig, swish)

AKTİVASYON FONKSİYONU: LOGSIG	0.1 (OKH) Hata oranına ulaşılan iterasyon:		0.01 (OKH) Hata oranına ulaşılan iterasyon:		0.001 (OKH) Hata oranına ulaşılan iterasyon:	
	21		54		495	
	Hedef	Çıkış	Hedef	Çıkış	Hedef	Çıkış
	0	$0.3127 + 0.2803i$	0	$0.1070 + 0.1122i$	0	$0.0230 + 0.0271i$
	i	$0.3544 + 0.7117i$	i	$0.1636 + 0.9016i$	i	$0.0359 + 0.9784i$
	1	$0.7170 + 0.3485i$	1	$0.8616 + 0.1540i$	1	$0.9651 + 0.0300i$
	$1+i$	$0.6903 + 0.6602i$	$1+i$	$0.8424 + 0.8335i$	$1+i$	$0.9613 + 0.9613i$
AKTİVASYON FONKSİYONU: TANSIG	0.1 (OKH) Hata oranına ulaşılan iterasyon:		0.01 (OKH) Hata oranına ulaşılan iterasyon:		0.001 (OKH) Hata oranına ulaşılan iterasyon:	
	1		6		74	
	Hedef	Çıkış	Hedef	Çıkış	Hedef	Çıkış
	0	$0.9012 + 0.3222i$	0	$0.3817 + 0.1732i$	0	$0.0384 + 0.0211i$
	i	$0.8490 + 0.3792i$	i	$0.4652 + 0.7957i$	i	$0.0343 + 0.9740i$
	1	$0.9029 + 0.5504i$	1	$0.9239 + 0.4188i$	1	$0.9851 + 0.0385i$
	$1+i$	$0.8756 + 0.4534i$	$1+i$	$0.8504 + 0.6671i$	$1+i$	$0.9745 + 0.9568i$
AKTİVASYON FONKSİYONU: SWISH	0.1 (OKH) Hata oranına ulaşılan iterasyon:		0.01 (OKH) Hata oranına ulaşılan iterasyon:		0.001 (OKH) Hata oranına ulaşılan iterasyon:	
	1		12		52	
	Hedef	Çıkış	Hedef	Çıkış	Hedef	Çıkış
	0	$0.8475 + 0.4323i$	0	$0.5067 + 0.2481i$	0	$0.0495 + 0.0130i$
	i	$0.8407 + 0.4854i$	i	$0.3105 + 0.7763i$	i	$0.0366 + 0.9716i$
	1	$0.8870 + 0.6360i$	1	$0.6860 + 0.4401i$	1	$0.9525 + 0.0262i$
	$1+i$	$0.9131 + 0.6182i$	$1+i$	$0.9743 + 0.9342i$	$1+i$	$0.9975 + 0.9881i$

KDYSA'lar 0.001 hata (OKH) oranına ulaştığı anda ağ durdurulmuş olup her bir iterasyon için hata oranı kaydedilerek, Şekil 4.3'te gösterilmiştir.



Şekil 4.3. Kanal dengeleme problemi için hata grafiği (logsig, tansig, swish)

Yapılan testler sonucunda swish aktivasyon fonksiyonunu kullanan ağın kanal dengeleme problemi çözümünde diğer aktivasyon fonksiyonlarını kullanan ağlara göre hedef değere daha erken yakınsadığı görülmüştür. Gizli katmanında kompleks swish aktivasyon fonksiyonu kullanan ağın 0.001 OKH hata oranına 52. iterasyonda ulaşırken, kompleks hiperbolik tanjant kullanan ağ 74., kompleks sigmoid fonksiyonu kullanan ağ ise 495. iterasyonda ulaşmıştır. Sonuç olarak kanal dengeleme problemi çözümünde gizli katmanında kompleks swish aktivasyon fonksiyonunu kullanan ağın, diğer bilinen aktivasyon fonksiyonlarını (logsig, tansig) kullanan ağlara göre sonuca daha erken yakınsadığı ve daha hızlı çalıştığı görülmüştür.

4.2. Kompleks Modifiye swish, E-Swish ve DTS Aktivasyon Fonksiyonlarının Sonuca En Hızlı Yakınsadığı Sabit Değerlerin Bulunması

Swish fonksiyonu üzerinden türetilen modifiye swish, e-swish ve DTS aktivasyon fonksiyonları içerisinde sırasıyla α , β ve T sabitlerini içermektedirler. Bu bölümde α , β ve T için farklı sayılar verilerek, kompleks 4 bit XOR, 3 giriş 1 çıkış simetri problemi ve kanal dengeleme problemi üzerinde testler yapılmıştır. Yapılan bu testler sonucunda aktivasyon fonksiyonlarının kompleks veriler üzerinde en iyi performans gösterdiği α , β ve T değerleri bulunmuştur.

4.2.1. Kompleks değerli XOR problemi

Kompleks modifiye swish, e-swish ve DTS aktivasyon fonksiyonlarının en iyi performans gösterdikleri sabit değerleri bulabilmek için ilk olarak XOR problemi çözülmüştür.

Kompleks e-Swish aktivasyon fonksiyonu içerisinde bulunan β sabitine 0.1 ve 4 aralığında 18 farklı değer verilerek, XOR problemi üzerinde 4 adet test yapılmıştır. KDYSA için 1 giriş, 1 çıkış ve 2 gizli katmana sahip ağ mimarisi kullanılmış olup öğrenme oranı 0.5 olarak alınmıştır. Hata oranı (OKH) 0.001'e ulaştığı anda ağ durdurulmuş olup, iterasyon numaraları kaydedilerek Çizelge 4.4'te verilmiştir. Maksimum iterasyon sayısı 3000 alınmış olup, ağın 0.001 hata oranına ulaşamaması durumunda iterasyon sayısı 3000+ olarak girilmiştir. Yapılan 4 test sonucundan hedefe en erken iterasyonda ulaşan 4 adet β değeri koyu renkle işaretlenmiş olup, skor kısmında koyu renk ile işaret edilen test sayılarının toplamı verilmiştir.

Çizelge 4.4. E-swish aktivasyon fonksiyonunun XOR problemi üzerinde en iyi performans gösterdiği β değerleri

Aktivasyon Fonksiyonu	0.001(OKH) hata oranına ulaşılan iterasyon				
	Test-I	Test-II	Test-III	Test-IV	Skor
E-Swish ($\beta = 0.1$)	3000+	3000+	3000+	3000+	-
E-Swish ($\beta = 0.2$)	2031	3000+	2821	2903	-
E-Swish ($\beta = 0.3$)	1218	3000+	1830	1765	-
E-Swish ($\beta = 0.4$)	930	2489	1371	1273	-
E-Swish ($\beta = 0.5$)	774	1756	1108	1004	-
E-Swish ($\beta = 0.6$)	669	1416	938	837	-
E-Swish ($\beta = 0.7$)	585	1201	821	722	-
E-Swish ($\beta = 0.8$)	512	1036	737	639	-
E-Swish ($\beta = 0.9$)	447	912	677	575	-
E-Swish ($\beta = 1.0$)	391	846	634	525	-
E-Swish ($\beta = 1.1$)	343	940	602	484	-
E-Swish ($\beta = 1.2$)	304	1066	579	449	1
E-Swish ($\beta = 1.3$)	272	735	570	420	2
E-Swish ($\beta = 1.4$)	245	665	353	395	4
E-Swish ($\beta = 1.5$)	223	920	507	372	3
E-Swish ($\beta = 2$)	203	738	3000+	293	3
E-Swish ($\beta = 3$)	100	792	3000+	223	3
E-Swish ($\beta = 4$)	368	1000	3000+	3000+	-

Yapılan testler sonucunda kompleks e-swish aktivasyon fonksiyonunu kullanan ağın XOR problemi çözümünde en iyi performans gösterdiği β değeri 1.4 olarak belirlenmiştir. β değeri 1.4 olarak alındığında yapılan tüm testlerde, ağın hedefe daha erken yakınsadığı görülmüştür. β değeri 1.4 alındığında e-swish aktivasyon fonksiyonunun formülü Eşitlik 4.1’de verilmiştir.

$$y = (1.4).x.\sigma(x) \quad (4.1)$$

Kompleks DTS aktivasyon fonksiyonunun en iyi performans gösterdiği T sabit değerini bulabilmek için, T sabitine 1 ile -2 aralığında 18 farklı değer verilerek testler yapılmıştır. Ağ mimarisi olarak 1 giriş, 1 çıkış ve 2 gizli katmanlı yapı kullanılmış olup öğrenme oranı 0.5 olarak alınmıştır. Ağ 0.001 hata oranı (OKH) ulaştığı anda durdurulmuş ve iterasyon sayıları kaydedilerek Çizelge 4.5’te gösterilmiştir. Yapılan testler sonucunda hedefe en erken yakınsayan 4 adet T değeri koyu renkle işaretlenmiş olup, skor kısmında koyu renkli testlerin toplamı verilmiştir.

Çizelge 4.5. DTS aktivasyon fonksiyonunun XOR problemi üzerinde en iyi performans gösterdiği T değerleri

Aktivasyon Fonksiyonu	0.001(OKH) hata oranına ulaşılan iterasyon				
	Test-I	Test-II	Test-III	Test-IV	Skor
DTS ($T = 1$)	3000+	3000+	3000+	3000+	-
DTS ($T = 0.5$)	3000+	3000+	3000+	3000+	-
DTS ($T = 0.4$)	3000+	3000+	3000+	3000+	-
DTS ($T = 0.3$)	3000+	3000+	2576	1611	-
DTS ($T = 0.2$)	2587	3000+	1826	1421	-
DTS ($T = 0.1$)	1628	3000+	1337	953	-
DTS ($T = 0$)	993	3000+	984	569	1
DTS ($T = -0.1$)	518	563	687	563	3
DTS ($T = -0.2$)	538	593	645	594	3
DTS ($T = -0.3$)	560	669	461	625	4
DTS ($T = -0.4$)	623	2385	433	646	3
DTS ($T = -0.5$)	3000+	3000+	434	3000+	1
DTS ($T = -0.6$)	3000+	3000+	432	3000+	1
DTS ($T = -0.7$)	3000+	3000+	486	3000+	-
DTS ($T = -0.8$)	3000+	3000+	661	3000+	-
DTS ($T = -0.9$)	3000+	3000+	882	3000+	-
DTS ($T = -1$)	3000+	3000+	1249	3000+	-
DTS ($T = -2$)	3000+	3000+	3000+	3000+	-

Yapılan testler sonucunda DTS aktivasyon fonksiyonunun $T = -0.3$ sabit değerinde en iyi performansı gösterdiği görülmüştür. Yapılan testlerin tamamında DTS, $T = -0.3$ sabit değeri ile hedefe en erken yakınsayan fonksiyon olduğu görülmüştür. $T = -0.3$ sabit değeri ile formüle edilen düzleştirilmiş t-swish Eşitlik 4.2’de verilmiştir.

$$DTS(x) = \begin{cases} \frac{x}{1 + e^{-x}} + -0.3 & , x \geq 0 \\ -0.3 & , x < 0 \end{cases} \quad (4.2)$$

Son olarak kompleks modifiye swish aktivasyon fonksiyonun hedefe en erken yakınsadığı α sabit değerini bulabilmek için XOR problemi kullanılmıştır. Yapılan testlerde 0.1 ile 4 aralığında 18 farklı α sabit değeri kullanılmıştır. KDYSA için 1 giriş, 1 çıkış ve 2 gizli katmana sahip ağ mimarisi kullanılmış olup öğrenme oranı 0.5 olarak alınmıştır. Hata oranı (OKH) 0.001’e ulaştığı anda ağ durdurulmuş olup, iterasyon sayıları Çizelge 4.5’te verilmiştir. Ağın hedefe en erken yakınsadığı α değerleri koyu renkle işaretlenmiş olup, skor kısmında koyu renkli olan test sayılarının toplamı verilmiştir.

Çizelge 4.6. Modifiye swish aktivasyon fonksiyonunun XOR problemi üzerinde en iyi performans gösterdiği α değerleri

Aktivasyon Fonksiyonu	0.001(OKH) hata oranına ulaşılan iterasyon				Skor
	Test-I	Test-II	Test-III	Test-IV	
Modifiye Swish ($\alpha = 0.1$)	3000+	3000+	3000+	3000+	-
Modifiye Swish ($\alpha = 0.2$)	3000+	3000+	3000+	3000+	-
Modifiye Swish ($\alpha = 0.3$)	3000+	3000+	3000+	3000+	-
Modifiye Swish ($\alpha = 0.4$)	3000+	2762	3000+	3000+	-
Modifiye Swish ($\alpha = 0.5$)	3000+	3000+	3000+	3000+	-
Modifiye Swish ($\alpha = 0.6$)	968	3000+	1794	896	-
Modifiye Swish ($\alpha = 0.7$)	769	2542	1680	722	-
Modifiye Swish ($\alpha = 0.8$)	688	3000+	1602	639	-
Modifiye Swish ($\alpha = 0.9$)	633	1221	1106	584	-
Modifiye Swish ($\alpha = 1.0$)	586	1427	872	544	-
Modifiye Swish ($\alpha = 1.1$)	545	1000	741	513	1
Modifiye Swish ($\alpha = 1.2$)	515	797	702	487	2
Modifiye Swish ($\alpha = 1.3$)	504	712	688	665	3
Modifiye Swish ($\alpha = 1.4$)	523	674	686	445	4
Modifiye Swish ($\alpha = 1.5$)	618	652	692	427	3
Modifiye Swish ($\alpha = 2$)	3000+	879	1042	365	1
Modifiye Swish ($\alpha = 3$)	3000+	3000+	931	367	1
Modifiye Swish ($\alpha = 4$)	3000+	3000+	578	525	1

Yapılan testler sonucunda kompleks modifiye swish aktivasyon fonksiyonu $\alpha=1.4$ değeri ile sonuca en erken yakınsadığı görülmüş olup modifiye swish aktivasyon fonksiyonu Eşitlik 4.3'te verilmiştir.

$$y = x \cdot \sigma((1.4) \cdot x) \quad (4.3)$$

Kompleks XOR problemi üzerinde 1 giriş, 1 çıkış ve 2 gizli katmana sahip ağ mimarisi ile yapılan testler sonucunda, kompleks e-swish aktivasyon fonksiyonunu $\beta=1.4$ değeri ile, DTS aktivasyon fonksiyonunun $T= -0.3$ sabit değeri ile ve kompleks modifiye swish aktivasyon fonksiyonunun $\alpha=1.4$ değeri ile en iyi performans gösterdiği belirlenmiştir. Önerdiğimiz aktivasyon fonksiyonlarının XOR problemi çözümünde en iyi performans gösterdiği sabit değerler belirlenmiş olup bir sonraki testlerde bu sabit değerleri kullanan fonksiyonlar arasında XOR problemi üzerinde karşılaştırma yapılmıştır.

4.2.2. Kompleks değerli simetri problemi

Kompleks modifiye swish, e-swish ve DTS aktivasyon fonksiyonlarının en iyi performans gösterdikleri sabit değerleri bulabilmek için ikinci olarak simetri problemi çözülmüştür.

İlk olarak kompleks e-swish aktivasyon fonksiyonunun en iyi performans gösterdiği β değerini bulabilmek için, 18 farklı değer ile simetri problemi üzerinde 4 adet test yapılmıştır. KDYSA için 3 giriş, 1 çıkış ve 1 gizli katmana sahip ağ mimarisi kullanılmıştır. Bu ağ için öğrenme oranı 0.5 olarak alınmıştır. Hata değeri (OKH) 0.001 oranına ulaşınca iterasyonlar durdurulup kaydedilerek Çizelge 4.7’de verilmiştir. Maksimum iterasyon sayısı 3000 olarak alınmıştır. Sonuca en erken yaklaşan dört β değeri koyu renkle işaretlenmiş olup, skor kısmında koyu renkli testlerin toplamı verilmiştir.

Çizelge 4.7. E-swish aktivasyon fonksiyonunun simetri problemi üzerinde en iyi performans gösterdiği β değerleri

Aktivasyon Fonksiyonu	0.001(OKH) hata oranına ulaşılan iterasyon				
	Test-I	Test-II	Test-III	Test-IV	Skor
E-Swish ($\beta = 0.1$)	3000+	3000+	3000+	3000+	-
E-Swish ($\beta = 0.2$)	3000+	3000+	3000+	3000+	-
E-Swish ($\beta = 0.3$)	3000+	1813	1899	3000+	-
E-Swish ($\beta = 0.4$)	3000+	1141	1349	3000+	-
E-Swish ($\beta = 0.5$)	981	792	1041	2661	-
E-Swish ($\beta = 0.6$)	679	595	850	3000+	-
E-Swish ($\beta = 0.7$)	518	474	721	1136	-
E-Swish ($\beta = 0.8$)	417	394	629	906	-
E-Swish ($\beta = 0.9$)	349	338	560	756	-
E-Swish ($\beta = 1.0$)	300	296	506	663	-
E-Swish ($\beta = 1.1$)	264	264	463	574	1
E-Swish ($\beta = 1.2$)	236	239	428	535	1
E-Swish ($\beta = 1.3$)	215	218	398	540	3
E-Swish ($\beta = 1.4$)	198	202	373	464	4
E-Swish ($\beta = 1.5$)	185	189	352	602	3
E-Swish ($\beta = 2$)	146	150	321	3000+	3
E-Swish ($\beta = 3$)	131	3000+	3000+	3000+	1
E-Swish ($\beta = 4$)	3000+	3000+	3000+	3000+	-

Yapılan testler sonucunda kompleks e-swish aktivasyon fonksiyonunu kullanan ağın simetri problemi çözümünde en iyi performans gösterdiği β değeri 1.4 olarak belirlenmiştir. β değeri 1.4 alınarak yapılan tüm testlerde, ağın sonuca diğer ağlardan daha

erken yakınsadığı görülmüştür. β değeri 1.4 alınarak elde edilen e-swish fonksiyonu Eşitlik 4.4'te verilmiştir.

$$y = (1.4).x.\sigma(x) \quad (4.4)$$

Simetri problemi üzerinde ikinci olarak kompleks DTS aktivasyon fonksiyonu denenmiştir. DTS fonksiyonuna 1 ve -2 aralığında 18 farklı değer verilerek, fonksiyonun en iyi performans gösterdiği sabit değer belirlenmeye çalışılmıştır. Ağ mimarisi olarak diğer ağlar ile aynı şekilde 3 giriş, 1 çıkış, 2 gizli katman ve 0.5 öğrenme oranlı yapı kullanılmıştır. Maksimum iterasyon 3000 seçilmiş olup, ağların 0.001 hata oranına ulaştığı iterasyonlar kaydedilerek, Çizelge 4.8'de verilmiştir. Her test için en iyi performans gösteren 4 sabit değer koyu renk ile işaretlenmiş olup, skor kısmında bu koyu renkli değerlerin toplamı verilmiştir.

Çizelge 4.8. DTS aktivasyon fonksiyonunun simetri problemi üzerinde en iyi performans gösterdiği T değerleri

Aktivasyon Fonksiyonu	0.001(OXH) hata oranına ulaşılan iterasyon				Skor
	Test-I	Test-II	Test-III	Test-IV	
DTS (T = 1)	3000+	3000+	3000+	3000+	-
DTS (T = 0.5)	3000+	3000+	3000+	3000+	-
DTS (T = 0.4)	3000+	3000+	3000+	3000+	-
DTS (T = 0.3)	824	3000+	3000+	3000+	-
DTS (T = 0.2)	338	3000+	973	3000+	-
DTS (T = 0.1)	278	3000+	554	2182	1
DTS (T = 0)	292	3000+	274	629	3
DTS (T = -0.1)	192	3000+	151	177	3
DTS (T = -0.2)	290	612	157	256	4
DTS (T = -0.3)	402	764	270	745	3
DTS (T = -0.4)	922	3000+	3000+	3000+	-
DTS (T = -0.5)	3000+	540	3000+	3000+	1
DTS (T = -0.6)	3000+	584	3000+	3000+	1
DTS (T = -0.7)	3000+	1499	3000+	3000+	-
DTS (T = -0.8)	3000+	3000+	3000+	3000+	-
DTS (T = -0.9)	3000+	3000+	3000+	3000+	-
DTS (T = -1)	3000+	3000+	3000+	3000+	-
DTS (T = -2)	3000+	3000+	3000+	3000+	-

Simetri problemi üzerinde DTS aktivasyon fonksiyonu ile yapılan testler sonucunda, $T = -0.2$ sabit değerinde fonksiyonun en iyi performansı gösterdiği görülmüştür. Yapılan testlerin tamamında DTS'nin $T = -0.2$ sabit değeri ile simetri problemi üzerinde hedefe en erken yakınsayan fonksiyon olduğu görülmüştür. $T = -0.2$

sabit değeri ile formüle edilmiş düzleştirilmiş t-swish fonksiyonu Eşitlik 4.5'te verilmiştir.

$$DTS(x) = \begin{cases} \frac{x}{1 + e^{-x}} + -0.2 & , x \geq 0 \\ -0.2 & , x < 0 \end{cases} \quad (4.5)$$

Son olarak simetri problemi üzerinde kompleks modifiye swish aktivasyon fonksiyonu ele alınmıştır. Modifiye swish fonksiyonu içerisinde bulunan α 'a 0.1 ve 4 aralığında 18 farklı değer verilerek, simetri problemi üzerinde en iyi performans gösterdiği değer bulunmaya çalışılmıştır. Ağ mimarisi olarak bir önceki ağlarının aynıları kullanılmış olup sonuçlar Çizelge 4.9'da verilmiştir.

Çizelge 4.9. Modifiye Swish aktivasyon fonksiyonunun simetri problemi üzerinde en iyi performans gösterdiği α değerleri

Aktivasyon Fonksiyonu	0.001(OXH) hata oranına ulaşılan iterasyon				
	Test-I	Test-II	Test-III	Test-IV	Skor
Modifiye Swish ($\alpha = 0.1$)	3000+	3000+	3000+	3000+	-
Modifiye Swish ($\alpha = 0.2$)	3000+	3000+	3000+	3000+	-
Modifiye Swish ($\alpha = 0.3$)	3000+	3000+	3000+	3000+	-
Modifiye Swish ($\alpha = 0.4$)	3000+	3000+	3000+	3000+	-
Modifiye Swish ($\alpha = 0.5$)	3000+	3000+	3000+	896	-
Modifiye Swish ($\alpha = 0.6$)	3000+	3000+	3000+	574	-
Modifiye Swish ($\alpha = 0.7$)	924	886	3000+	437	1
Modifiye Swish ($\alpha = 0.8$)	642	734	3000+	361	1
Modifiye Swish ($\alpha = 0.9$)	520	1275	476	314	1
Modifiye Swish ($\alpha = 1.0$)	446	1354	426	280	1
Modifiye Swish ($\alpha = 1.1$)	397	1656	613	255	-
Modifiye Swish ($\alpha = 1.2$)	361	1518	743	235	-
Modifiye Swish ($\alpha = 1.3$)	335	1099	593	219	1
Modifiye Swish ($\alpha = 1.4$)	315	975	500	207	3
Modifiye Swish ($\alpha = 1.5$)	300	901	448	196	4
Modifiye Swish ($\alpha = 2$)	262	530	576	170	3
Modifiye Swish ($\alpha = 3$)	266	3000+	3000+	3000+	1
Modifiye Swish ($\alpha = 4$)	3000+	3000+	3000+	3000+	-

Yapılan testler sonucunda kompleks modifiye swish aktivasyon fonksiyonu $\alpha=1.5$ değeri ile sonuca diğer sabit değerleri kullanan ağlara göre daha erken yakınsadığı görülmüş olup fonksiyon Eşitlik 4.6'da verilmiştir.

$$y = x \cdot \sigma((1.5) \cdot x) \quad (4.6)$$

Simetri problemi üzerinde yapılan testler sonucunda, $\beta=1.4$ sabit değerli e-swish, $T=-0.2$ sabit değerine sahip DTS ve $\alpha=1.5$ sabit değerine sahip modifiye swish'in sonuca daha erken yakınsadığı görülmüştür. Bundan sonraki testlerde bu sabit değerlere sahip aktivasyon fonksiyonlarının performansları birbirleri ile karşılaştırılacaktır.

4.2.3. Kanal dengeleme problemi

Önerdiğimiz kompleks modifiye swish, e-swish ve DTS aktivasyon fonksiyonlarının hedefe en erken yakınsadığı sabit değerleri bulabilmek için son olarak kanal dengeleme problemi çözülmüştür.

Kanal dengeleme problemi çözümünde ilk olarak e-swish fonksiyonu denenmiştir. E-swish fonksiyonu içerisine bulunan β sabitine 18 farklı değer verilerek, 0.001 hata oranına ulaştığı iterasyonlar kaydedilerek Çizelge 4.10'da verilmiştir. Ağ mimarisi olarak 1 giriş, 1 çıkış ve 2 gizli katman seçilmiş olup öğrenme oranı 0.5 alınmıştır. Yapılan 4 test sonucunda sonuca en erken yakınsayan 4 değer koyu renk ile boyanmış olup, skor kısmında koyu renkli değerlerin toplamı verilmiştir.

Çizelge 4.10. E-swish aktivasyon fonksiyonunun kanal dengeleme problemi üzerinde en iyi performans gösterdiği β değerleri

Aktivasyon Fonksiyonu	0.001(OXH) hata oranına ulaşılan iterasyon				Skor
	Test-I	Test-II	Test-III	Test-IV	
E-Swish ($\beta = 0.1$)	852	922	711	723	-
E-Swish ($\beta = 0.2$)	533	483	470	448	-
E-Swish ($\beta = 0.3$)	347	337	399	344	-
E-Swish ($\beta = 0.4$)	232	260	370	242	-
E-Swish ($\beta = 0.5$)	166	209	360	164	-
E-Swish ($\beta = 0.6$)	127	170	359	117	-
E-Swish ($\beta = 0.7$)	102	141	362	88	-
E-Swish ($\beta = 0.8$)	85	118	366	71	-
E-Swish ($\beta = 0.9$)	72	101	366	59	-
E-Swish ($\beta = 1.0$)	62	88	364	50	-
E-Swish ($\beta = 1.1$)	54	78	359	35	1
E-Swish ($\beta = 1.2$)	46	71	354	33	1
E-Swish ($\beta = 1.3$)	40	65	350	32	2
E-Swish ($\beta = 1.4$)	34	60	347	26	4
E-Swish ($\beta = 1.5$)	29	56	345	36	3
E-Swish ($\beta = 2$)	20	46	343	37	3
E-Swish ($\beta = 3$)	27	44	366	38	2
E-Swish ($\beta = 4$)	571	63	525	417	-

Kanal dengeleme problemi üzerinde e-swish aktivasyon fonksiyonunu ile yapılan testler sonucunda kompleks e-swish aktivasyon fonksiyonunun en iyi performans gösterdiği β değeri 1.4 olarak belirlenmiştir. β değeri 1.4 alınarak yapılan tüm testlerde, ağın sonuca diğer ağlardan daha erken yakınsadığı görülmüş olup e-swish fonksiyonun formüle edilmiş hali Eşitlik 4.7’de verilmiştir.

$$y = (1.4).x.\sigma(x) \quad (4.7)$$

Kompleks DTS aktivasyon fonksiyonunun kanal dengeleme problemi üzerinde hedefe en erken yakınsadığı T sabit değerini bulabilmek için, T sabitine 18 farklı değer verilerek testler yapılmıştır. Ağ mimarisi olarak 1 giriş, 1 çıkış ve 2 gizli katmanlı yapı kullanılmış olup öğrenme oranı 0.5 olarak alınmıştır. Ağ 0.001 hata oranı (OKH) ulaştığı anda durdurulmuş ve iterasyon sayıları kaydedilerek Çizelge 4.11’de verilmiştir. Yapılan testler sonucunda hedefe en erken yakınsayan 4 adet T değeri koyu renkle işaretlenmiş olup, skor kısmında koyu renkli testlerin toplamı verilmiştir.

Çizelge 4.11. DTS aktivasyon fonksiyonunun kanal dengeleme problemi üzerinde en iyi performans gösterdiği T değerleri

Aktivasyon Fonksiyonu	0.001(OKH) hata oranına ulaşılan iterasyon				
	Test-I	Test-II	Test-III	Test-IV	Skor
DTS (T = 1)	42	49	65	57	4
DTS (T = 0.5)	48	51	84	70	3
DTS (T = 0.4)	50	53	88	71	3
DTS (T = 0.3)	54	54	93	73	3
DTS (T = 0.2)	41	58	101	75	-
DTS (T = 0.1)	77	66	113	78	-
DTS (T = 0)	169	86	134	82	-
DTS (T = -0.1)	54	307	101	99	-
DTS (T = -0.2)	57	269	150	183	-
DTS (T = -0.3)	61	226	181	196	-
DTS (T = -0.4)	65	187	178	163	-
DTS (T = -0.5)	65	154	171	134	-
DTS (T = -0.6)	63	127	153	110	-
DTS (T = -0.7)	61	108	132	91	-
DTS (T = -0.8)	60	94	115	76	-
DTS (T = -0.9)	64	85	166	64	1
DTS (T = -1)	80	81	231	55	1
DTS (T = -2)	97	1036	1884	38	1

Yapılan testler sonucunda DTS aktivasyon fonksiyonunun $T=1$ sabit değerinde en iyi performansı gösterdiği görülmüştür. Yapılan testlerin tamamında DTS'nin $T=1$ sabit değeri ile hedefe en erken yakınsayan fonksiyon olduğu görülmüştür. $T=1$ sabit değerine sahip düzleştirilmiş t-swish fonksiyonu Eşitlik 4.8'de verilmiştir.

$$DTS(x) = \begin{cases} \frac{x}{1 + e^{-x}} + 1, & x \geq 0 \\ 1, & x < 0 \end{cases} \quad (4.8)$$

Son olarak kanal dengeleme problemi modifiye swish aktivasyon fonksiyonu kullanarak çözülmüştür. Ağ mimarisi olarak diğer kanal dengeleme problemleri ile aynı şekilde; 1 giriş, 1 çıkış ve 2 gizli katmana sahip yapı kullanılmış olup öğrenme oranı 0.5 alınmıştır. KDYSA 0.001 hata oranına (OKH) ulaştığı anda iterasyonlar durdurulmuş ve sonuçlar kaydedilerek Çizelge 4.12'de verilmiştir.

Çizelge 4.12. Modifiye Swish aktivasyon fonksiyonunun kanal dengeleme problemi üzerinde en iyi performans gösterdiği α değerleri

Aktivasyon Fonksiyonu	0.001(OKH) hata oranına ulaşılan iterasyon				
	Test-I	Test-II	Test-III	Test-IV	Skor
Modifiye Swish ($\alpha = 0.1$)	22	53	55	45	3
Modifiye Swish ($\alpha = 0.2$)	64	69	64	56	2
Modifiye Swish ($\alpha = 0.3$)	76	90	76	79	1
Modifiye Swish ($\alpha = 0.4$)	87	122	87	1081	-
Modifiye Swish ($\alpha = 0.5$)	96	187	96	542	-
Modifiye Swish ($\alpha = 0.6$)	101	328	101	456	-
Modifiye Swish ($\alpha = 0.7$)	102	365	102	412	-
Modifiye Swish ($\alpha = 0.8$)	101	343	101	361	-
Modifiye Swish ($\alpha = 0.9$)	93	326	93	338	-
Modifiye Swish ($\alpha = 1.0$)	86	315	86	328	-
Modifiye Swish ($\alpha = 1.1$)	79	308	79	324	-
Modifiye Swish ($\alpha = 1.2$)	72	303	72	323	-
Modifiye Swish ($\alpha = 1.3$)	65	302	65	322	-
Modifiye Swish ($\alpha = 1.4$)	59	312	59	320	-
Modifiye Swish ($\alpha = 1.5$)	54	337	54	319	1
Modifiye Swish ($\alpha = 2$)	39	273	39	94	2
Modifiye Swish ($\alpha = 3$)	32	166	32	52	3
Modifiye Swish ($\alpha = 4$)	47	83	47	41	4

Kanal dengeleme problemi üzerinde yapılan testler sonucunda kompleks modifiye swish aktivasyon fonksiyonu $\alpha=4$ değeri ile sonuca diğer sabit değerleri kullanan ağlara göre daha erken yakınsadığı görülmüştür. $\alpha=4$ değerine sahip modifiye swish fonksiyonu Eşitlik 4.9'da verilmiştir.

$$y = x. \sigma((4).x) \quad (4.9)$$

Önerilen aktivasyon fonksiyonları 18 farklı sabit değer ile kanal dengeleme problemi üzerinde test edilmiş olup yapılan testler sonucunda, $\beta=1.4$ sabit değerli e-swish, $T=1$ sabit değerine sahip DTS ve $\alpha=4$ sabit değerine sahip modifiye swish'in sonuca daha erken yakınsadığı görülmüştür.

4.3. Yeni Tanımlanan Aktivasyon Fonksiyonlarının Performanslarının Birbirleri ile Karşılaştırılması

Bu bölümde kompleks swish fonksiyonun ve yine swish fonksiyonu üzerinden türetilen kompleks e-swish, DTS ve modifiye swish aktivasyon fonksiyonlarının (bir önceki bölümde belirlenen sabit değerleri ile) performansları, kompleks XOR, simetri ve kanal dengeleme problemleri üzerinde karşılaştırılmıştır. İlk bölümde kompleks swish aktivasyon fonksiyonunun bilinen kompleks aktivasyon fonksiyonlarından (hiperbolik tanjant ve sigmoid) sonuca daha erken yakınsadığı gösterilmişti. Sonraki bölümde ise swish aktivasyon fonksiyonu üzerinden türetilen yeni aktivasyon fonksiyonlarının en iyi performans gösterdiği sabit değerler elde edilmiştir.

Önerilen kompleks swish, e-swish, DTS ve modifiye swish aktivasyon fonksiyonlarını kullanan ağlar karşılaştırılmış, en iyi performans gösteren aktivasyon fonksiyonu belirlenmiştir. Bu kapsamda ilk olarak ağların 0.001 hata oranına (OKH) ulaştığı iterasyon sayıları baz alınarak karşılaştırma yapılmıştır. İkinci olarak ise belirlenen iterasyon sayılarına gelen ağların hata oranları ve çıkışları verilerek karşılaştırma yapılmıştır.

4.3.1. Ortalama karesel hata baz alınarak kıyaslama

Kompleks değerli XOR problemi:

Kompleks swish, e-swish ($\beta=1.4$), DTS ($T=-0.3$) ve modifiye swish ($\alpha=1.4$) aktivasyon fonksiyonlarını kullanan KDYSA'lar ilk olarak XOR problemi üzerinde test edilmiştir. Ağ mimarisi olarak 1 giriş, 1 çıkış ve 2 gizli düğüme sahip yapı kullanılmıştır.

Ağ için öğrenme oranı 0.5 olarak alınmıştır. Toplamda 5 adet test yapılmış olup ağın 0.001 hata oranına (OKH) ulaştığı iterasyon sayıları Çizelge 4.13'te gösterilmiştir.

Çizelge 4.13. XOR problemi için test sonuçları (swish, e-swish, DTS, modifiye swish)

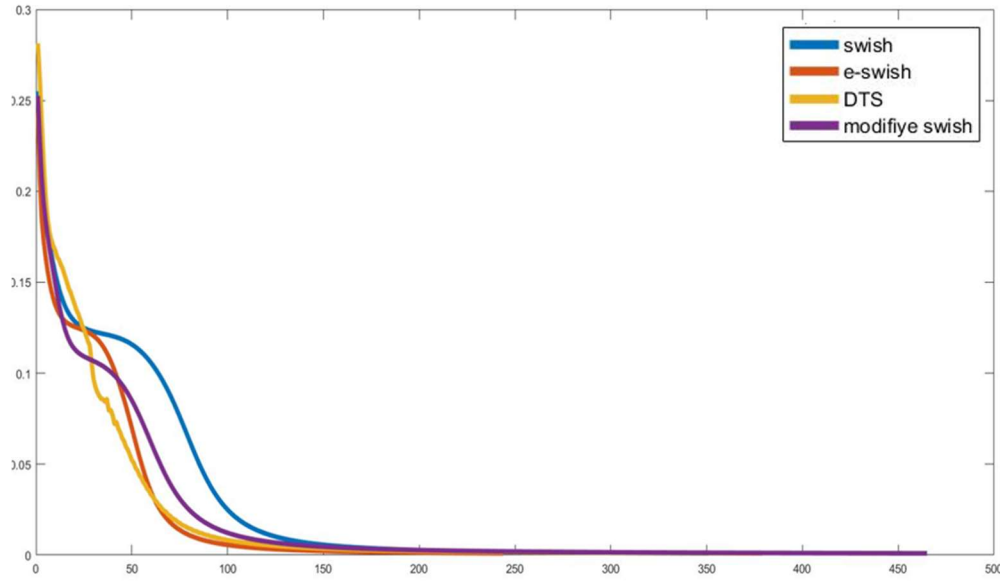
Aktivasyon Fonksiyonu	0.001(OKH) hata oranına ulaşılan iterasyon				
	Test-I	Test-II	Test-III	Test-IV	Test-V
Swish	351	387	594	1062	1619
*E- Swish ($\beta = 1.4$)	244	261	552	471	188
DTS (T = -0.3)	379	796	594	575	541
Modifiye Swish ($\alpha = 1.4$)	465	399	651	1014	668

Yukarıda tabloda da görüldüğü gibi $\beta=1.4$ sabit değerine sahip kompleks e-swish ($\beta=1.4$) aktivasyon fonksiyonunu kullanan ağ, sonuca diğer ağlardan daha erken yakınsamıştır. Yapılan 5 test sonucunda, tüm testlerde en iyi performans gösteren ağ, kompleks e-swish ($\beta=1.4$) aktivasyon fonksiyonunu kullanan ağ olmuştur.

Toplamda 5 adet test yapılmış olup; I no'lu test için hedef ve çıkış tablosu Çizelge 4.14'te, hata oranları grafiği Şekil 4.4'te verilmiştir.

Çizelge 4.14. XOR problemi I no'lu test için hedef çıkış tablosu (swish, e-swish, DTS, modifiye swish)

Aktivasyon fonksiyonu SWISH		Aktivasyon fonksiyonu E-SWISH ($\beta = 1.4$)		Aktivasyon Fonksiyonu DTS (T = -0.3)		Aktivasyon Fonksiyonu MODİFİYE SWISH ($\alpha = 1.4$)	
0.001(OKH) hata oranına ulaşılan iterasyon: 351		0.001(OKH) hata oranına ulaşılan iterasyon: 244		0.001(OKH) hata oranına ulaşılan iterasyon: 379		0.001(OKH) hata oranına ulaşılan iterasyon: 465	
<i>Hedef</i>	<i>Çıkış</i>	<i>Hedef</i>	<i>Çıkış</i>	<i>Hedef</i>	<i>Çıkış</i>	<i>Hedef</i>	<i>Çıkış</i>
0	0.057+0.035i	0	0.058+0.036i	0	0.062+0.050i	0	0.054+0.041i
1	0.969+0.021i	1	0.970+0.021i	1	0.975+0.002i	1	0.974+0.017i
1+i	0.964+0.990i	1+i	0.966+0.989i	1+i	0.980+0.985i	1+i	0.973+0.978i
i	0.024+0.997i	i	0.022+0.997i	i	0.016+0.991i	i	0.014+0.991i



Şekil 4.4. XOR problemi I no'lu test için hata grafiği (swish, e-swish, DTS, modifiye swish)

Kompleks değerli simetri problemi:

Önerdiğimiz aktivasyon fonksiyonlarını kullanan KDYSA ikinci olarak simetri problemi üzerinde test edilmiştir. Kompleks swish aktivasyon fonksiyonu ile bir önceki bölümde bulunan sabit değerlere sahip kompleks e-swish, modifiye swish ve DTS fonksiyonları kıyaslanmıştır. Simetri problemi üzerinde e-swish $\beta=1.4$ sabit değeri ile en iyi sonuçları vermiş olup bu bölümde de aynı değer ile test edilmiştir. Yine aynı şekilde DTS $T= -0.2$, modifiye swish ise $\alpha=1.5$ sabit değeri ile test edilmiştir. KDYSA ağ mimarisi 3 giriş, 1 çıkış ve gizli katman içinde 2 düğüm olacak şekilde tasarlanmıştır. Ağ için öğrenme oranı 0.5 olarak alınmıştır. Toplamda 5 adet test yapılmış olup; ağ, 0.001 hata oranına (OKH) ulaştığı anda iterasyonlar durdurularak sonuçlar kaydedilmiştir. Simetri problemi için 0.001 hata oranında ulaşılan iterasyonlar Çizelge 4.15'te verilmiştir.

Çizelge 4.15. Simetri problemi için test sonuçları (swish, e-swish, DTS, modifiye swish)

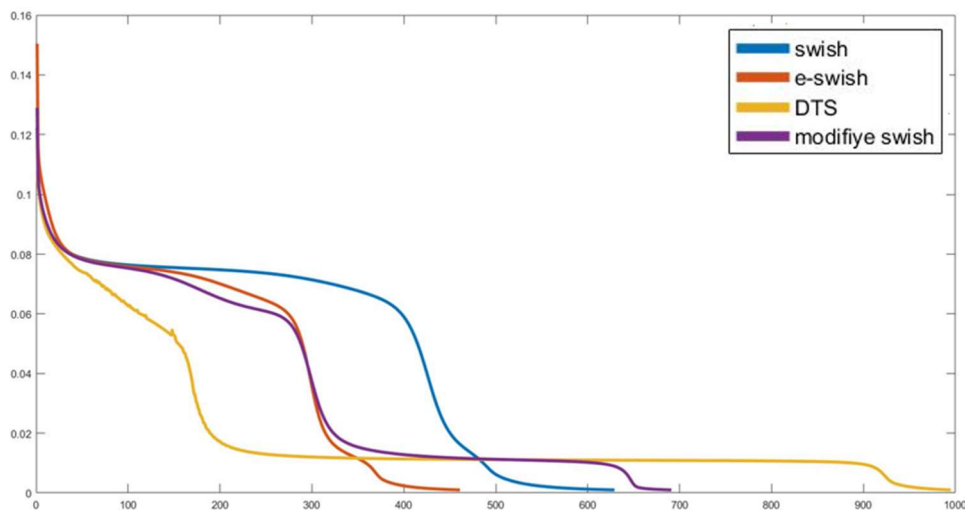
Aktivasyon Fonksiyonu	0.001(OKH) hata oranına ulaşılan iterasyon				
	Test-I	Test-II	Test-III	Test-IV	Test-V
Swish	629	175	215	245	261
*E- Swish ($\beta = 1.4$)	461	112	174	146	165
DTS (T = -0.2)	691	133	239	259	214
Modifiye Swish($\alpha = 1.5$)	995	158	193	199	250

Simetri problemi üzerinde yapılan testler sonucunda, $\beta=1.4$ sabit değerine sahip kompleks e-swish aktivasyon fonksiyonunu kullanan ağı, diğer ağlara göre sonuca daha erken yakınsadığı görülmüştür. Yapılan 5 adet testin tamamında, simetri problemi üzerinde en iyi performans gösteren ağ, e-swish ($\beta=1.4$) aktivasyon fonksiyonunu kullanan ağ olmuştur.

Yapılan 5 adet testten ilki olan I no'lu test için hedef-çıkış tablosu Çizelge 4.16'da, hata oranları grafiği ise Şekil 4.5'te verilmiştir.

Çizelge 4.16. Simetri problemi I no'lu test için hedef çıkış tablosu (swish, e-swish, DTS, modifiye swish)

Aktivasyon fonksiyonu SWISH		Aktivasyon fonksiyonu E-SWISH ($\beta = 1.4$)		Aktivasyon Fonksiyonu DTS ($T = -0.2$)		Aktivasyon Fonksiyonu MODİFİYE SWISH ($\alpha = 1.5$)	
0.001(OKH) hata oranına ulaşılan iterasyon: 629		0.001(OKH) hata oranına ulaşılan iterasyon: 461		0.001(OKH) hata oranına ulaşılan iterasyon: 691		0.001(OKH) hata oranına ulaşılan iterasyon: 995	
<i>Hedef</i>	<i>Çıkış</i>	<i>Hedef</i>	<i>Çıkış</i>	<i>Hedef</i>	<i>Çıkış</i>	<i>Hedef</i>	<i>Çıkış</i>
$0.7+0.7i$	$0.71+0.69i$	$0.7+0.7i$	$0.71+0.68i$	$0.7+0.7i$	$0.72+0.70i$	$0.7+0.7i$	$0.72+0.70i$
1	$0.95+0.04i$	1	$0.97+0.04i$	1	$0.94+0.02i$	1	$0.94+0.02i$
$0.7+0.7i$	$0.71+0.70i$	$0.7+0.7i$	$0.71+0.70i$	$0.7+0.7i$	$0.71+0.71i$	$0.7+0.7i$	$0.71+0.71i$
1	$0.95+0.03i$	1	$0.94+0.03i$	1	$0.94+0.02i$	1	$0.94+0.02i$
1	$0.95+0.03i$	1	$0.98+0.02i$	1	$0.94+0.02i$	1	$0.94+0.02i$
$0.7+0.7i$	$0.71+0.70i$	$0.7+0.7i$	$0.71+0.70i$	$0.7+0.7i$	$0.71+0.70i$	$0.7+0.7i$	$0.71+0.70i$
1	$0.95+0.04i$	1	$0.96+0.04i$	1	$0.94+0.03i$	1	$0.94+0.03i$
$0.7+0.7i$	$0.71+0.70i$	$0.7+0.7i$	$0.71+0.71i$	$0.7+0.7i$	$0.71+0.70i$	$0.7+0.7i$	$0.71+0.70i$



Şekil 4.5. Simetri problemi I no'lu test için hata grafiği (swish, e-swish, DTS, modifiye swish)

Kanal Dengeleme Problemi:

Son olarak kompleks swish, e-swish, DTS ve modifiye swish fonksiyonları kanal dengeleme problemi üzerinde test edilmiştir. Daha önce kanal dengeleme problemi üzerinde bulunan e-swish için $\beta=1.4$, DTS için $T=1$ ve modifiye swish için $\alpha=4$ sabit değerleri kullanılmıştır. KDYSA için 1 giriş, 1 çıkış ve 2 gizli katmana sahip ağ mimarisi kullanılmış olup öğrenme oranı 0.5 olarak alınmıştır. Toplamda 5 adet test yapılmış olup, ağın 0.001 hata oranına (OKH) ulaştığı anda iterasyonlar Çizelge 4.17’de verilmiştir.

Çizelge 4.17. Kanal dengeleme problemi için test sonuçları (swish, e-swish, DTS, modifiye swish)

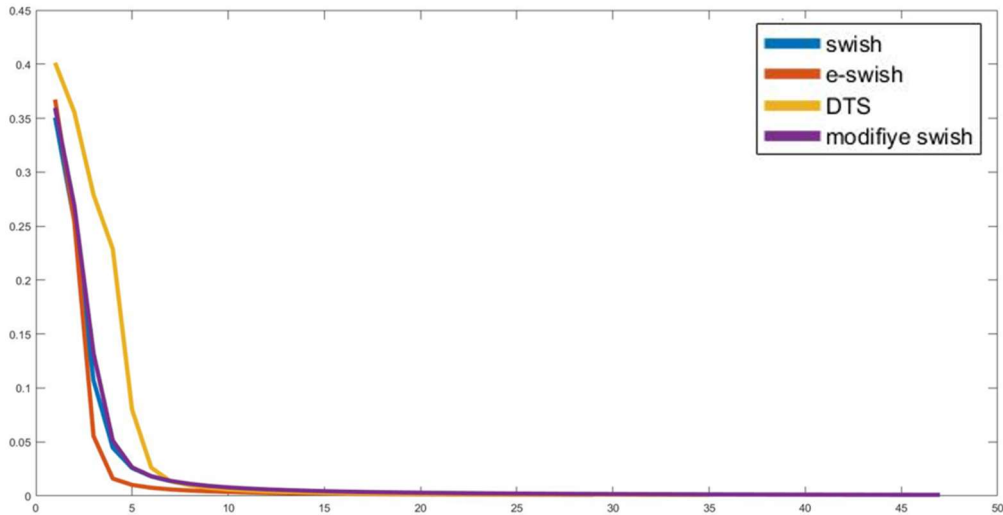
Aktivasyon Fonksiyonu	0.001(OKH) hata oranına ulaşılan iterasyon				
	Test-I	Test-II	Test-III	Test-IV	Test-V
Swish	46	40	31	38	48
*E- Swish ($\beta = 1.4$)	29	27	11	23	34
DTS ($T = 1$)	35	35	19	29	43
Modifiye Swish($\alpha = 4$)	47	37	24	37	44

Yapılan 5 adet test sonucunda, sonuca en erken yakınsayan ağ, kompleks e-swish ($\beta=1.4$) aktivasyon fonksiyonunu kullanan ağ olmuştur. Yapılan testlerin tamamında en iyi performansı e-swish fonksiyonunu kullanan ağ göstermiştir.

Yapılan I no’lu teste ait hedef çıkış tablosu Çizelge 4.18’de, hata oranları grafiği ise Şekil 4.6’da verilmiştir.

Çizelge 4.18. Kanal dengeleme problemi I no’lu için hedef çıkış tablosu (swish, e-swish, DTS, modifiye swish)

Aktivasyon fonksiyonu SWISH		Aktivasyon fonksiyonu E-SWISH ($\beta = 1.4$)		Aktivasyon Fonksiyonu DTS ($T = 1$)		Aktivasyon Fonksiyonu MODİFİYE SWISH ($\alpha = 4$)	
0.001(OKH) hata oranına ulaşılan iterasyon: 46		0.001(OKH) hata oranına ulaşılan iterasyon: 29		0.001(OKH) hata oranına ulaşılan iterasyon: 47		0.001(OKH) hata oranına ulaşılan iterasyon: 35	
<i>Hedef</i>	<i>Çıkış</i>	<i>Hedef</i>	<i>Çıkış</i>	<i>Hedef</i>	<i>Çıkış</i>	<i>Hedef</i>	<i>Çıkış</i>
<i>0</i>	<i>0.047+0.004i</i>	<i>0</i>	<i>0.044+0.002i</i>	<i>0</i>	<i>0.051+0.007i</i>	<i>0</i>	<i>0.053+0.004i</i>
<i>i</i>	<i>0.044+0.990i</i>	<i>i</i>	<i>0.041+0.990i</i>	<i>i</i>	<i>0.032+0.993i</i>	<i>i</i>	<i>0.018+0.997i</i>
<i>1</i>	<i>0.971+0.003i</i>	<i>1</i>	<i>0.972+0.002i</i>	<i>1</i>	<i>0.966+0.003i</i>	<i>1</i>	<i>0.961+0.001i</i>
<i>1+i</i>	<i>0.961+0.961i</i>	<i>1+i</i>	<i>0.959+0.959i</i>	<i>1+i</i>	<i>0.965+0.958i</i>	<i>1+i</i>	<i>0.963+0.959i</i>



Şekil 4.6. Kanal dengeleme problemi I no'lu test için hata grafiği (swish, e-swish, DTS, modifiye swish)

4.3.2. Sabit iterasyon sayıları için aktivasyon fonksiyonlarının karşılaştırması

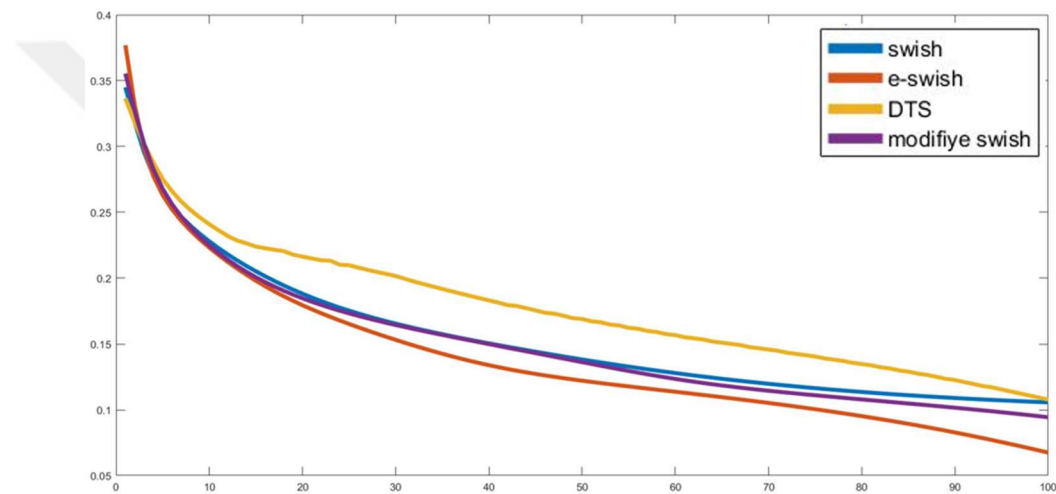
Ağın öğrenmesinin gösterilmesinin en güzel yolu hata grafiğini çizmektir. Her iterasyonda oluşan hatanın grafiği çizilirse hatanın zaman içinde düştüğü gözlemlenebilir. Hata oranının hızlı düşmesi ise bize öğrenmenin hızlı olduğunu gösterir ve buda ağın hızlı çalışması anlamına gelir. Bu bölümde ağlar için sabit bir iterasyon sayısı belirlenerek bu sayıya gelen ağlar durdurulmuştur. Maksimum iterasyon sayısına ulaşan ağlara ait hata oranları ve hata grafikleri karşılaştırılarak, en iyi performans gösteren aktivasyon fonksiyonu belirlenmiştir.

Kompleks değerli XOR problemi:

Kompleks swish, e-swish, DTS ve modifiye swish aktivasyon fonksiyonunu kullanan ağların, bir önceki bölümde verilen aynı ağ mimarisini kullanarak XOR problemi üzerinden test edilmiştir. Bütün ağlar için maksimum iterasyon sayısı 100 alınarak, ağların 100. iterasyonda elde ettiği çıkışlar ve hata oranları Çizelge 4.19'da verilmiştir. Ayrıca ağların 100. iterasyona kadar elde ettiği hatanın grafiği ise Şekil 4.7'de verilmiştir.

Çizelge 4.19. XOR problemi 100. iterasyon için hedef çıkış tablosu (swish, e-swish, DTS, modifiye swish)

Aktivasyon fonksiyonu SWISH		Aktivasyon fonksiyonu E-SWISH ($\beta = 1.4$)		Aktivasyon Fonksiyonu DTS ($T = -0.3$)		Aktivasyon Fonksiyonu MODİFİYE SWISH ($\alpha = 1.4$)	
Hata oranı %0.105		Hata oranı %0.067		Hata oranı %0.107		Hata oranı %0.094	
<i>Hedef</i>	<i>Çıkış</i>	<i>Hedef</i>	<i>Çıkış</i>	<i>Hedef</i>	<i>Çıkış</i>	<i>Hedef</i>	<i>Çıkış</i>
0	0.662+0.094i	0	0.547+0.078i	0	0.662+0.242i	0	0.673+0.102i
1	0.541+0.171i	1	0.684+0.167i	1	0.535+0.117i	1	0.614+0.142i
1+i	0.736+0.830i	1+i	0.771+0.844i	1+i	0.771+0.985i	1+i	0.769+0.841i
i	0.243+0.979i	i	0.173+0.998i	i	0.242+0.848i	i	0.204+0.973i



Şekil 4.7. XOR problemi 100. iterasyon için hata grafiği (swish, e-swish, DTS, modifiye swish)

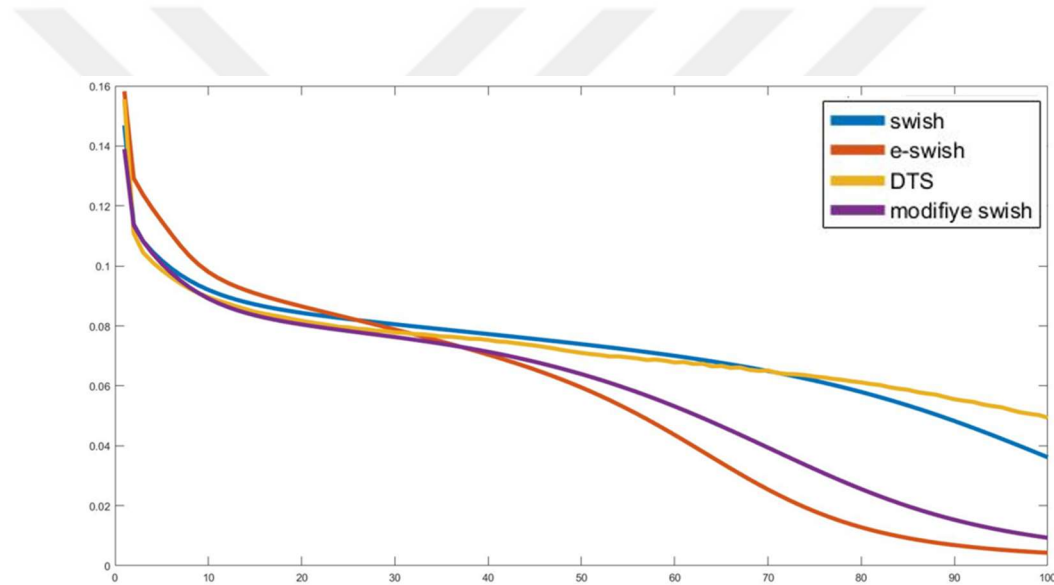
Yapılan 100 iterasyon sonucunda kompleks e-swish ($\beta=1.4$) fonksiyonu 0.067 hata oranı ile en düşük değere ulaşırken, sırasıyla modifiye swish ($\alpha=1.5$) 0.094, swish 0.0105 ve DTS ($T= -0.3$) 0.107 hata oranlarına ulaşmıştır.

Kompleks değerli simetri problemi:

Ağ mimarisi bir önceki bölümde çözülen simetri problemi ile aynı kalacak şekilde sadece maksimum iterasyon sayısı 100 alınarak Kompleks swish, e-swish, DTS ve modifiye swish aktivasyon fonksiyonunu kullanan ağlar ikinci olarak simetri problemi üzerinde test edilmiştir. Yapılan testlerde elde edilen hedef-çıkış tablosu ve hata grafiği Çizelge 4.20’de verilmiştir. Ağların 100. iterasyona kadar elde ettiği hatalar ise Şekil 4.8’de karşılaştırmalı olarak verilmiştir.

Çizelge 4.20. Simetri problemi 100. iterasyon için hedef çıkış tablosu (swish, e-swish, DTS, modifiye swish)

Aktivasyon fonksiyonu SWISH		Aktivasyon fonksiyonu E-SWISH ($\beta = 1.4$)		Aktivasyon Fonksiyonu DTS ($T = -0.3$)		Aktivasyon Fonksiyonu MODİFİYE SWISH ($\alpha = 1.5$)	
Hata oranı %0.036		Hata oranı %0.004		Hata oranı %0.049		Hata oranı %0.009	
<i>Hedef</i>	<i>Çıkış</i>	<i>Hedef</i>	<i>Çıkış</i>	<i>Hedef</i>	<i>Çıkış</i>	<i>Hedef</i>	<i>Çıkış</i>
$0.7+0.7i$	$0.79+0.35i$	$0.7+0.7i$	$0.67+0.58i$	$0.7+0.7i$	$0.91+0.41i$	$0.7+0.7i$	$0.70+0.53i$
1	$0.92+0.24i$	1	$0.95+0.07i$	1	$0.98+0.30i$	1	$0.90+0.11i$
$0.7+0.7i$	$0.83+0.51i$	$0.7+0.7i$	$0.77+0.73i$	$0.7+0.7i$	$0.92+0.51i$	$0.7+0.7i$	$0.80+0.67i$
1	$0.88+0.31i$	1	$0.94+0.09i$	1	$0.98+0.31i$	1	$0.93+0.15i$
1	$0.80+0.18i$	1	$0.90+0.07i$	1	$0.71+0.18i$	1	$0.86+0.11i$
$0.7+0.7i$	$0.85+0.50i$	$0.7+0.7i$	$0.74+0.71i$	$0.7+0.7i$	$0.91+0.50i$	$0.7+0.7i$	$0.77+0.66i$
1	$0.91+0.24i$	1	$0.96+0.08i$	1	$0.79+0.25i$	1	$0.95+0.12i$
$0.7+0.7i$	$0.81+0.51i$	$0.7+0.7i$	$0.73+0.66i$	$0.7+0.7i$	$0.91+0.50i$	$0.7+0.7i$	$0.78+0.64i$



Şekil 4.8. Simetri problemi 100. iterasyon için hata grafiği (swish, e-swish, DTS, modifiye swish)

Kompleks e-swish ($\beta=1.4$) fonksiyonunu kullanan ağ 100 iterasyon sonucunda 0.004 hata oranı ile diğer ağlar arasında en düşük hata oranını elde eden ağ olmuştur. Ağ çıkışlarından da görüldüğü üzere sonuca en yakın değerleri kompleks e-swish ($\beta=1.4$) fonksiyonunu kullanan ağ vermiştir.

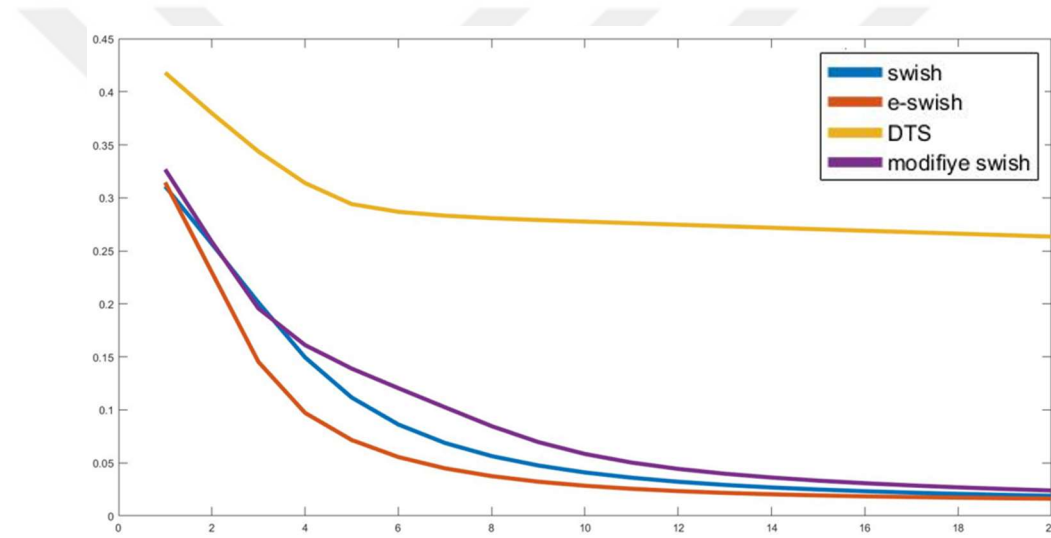
Kanal dengeleme problemi:

Son olarak kompleks swish, e-swish ($\beta=1.4$), DTS ($T=1$) ve modifiye swish ($\alpha=4$) aktivasyon fonksiyonlarını kullanan ağlar için maksimum iterasyon sayısı 20 alınarak, ağların bu iterasyonda ürettikleri çıkışlar ve hata oranları karşılaştırılmıştır. Ağların 20.

iterasyonda elde ettiği çıkışlar ve hata oranları Çizelge 4.21’de, hata grafiği ise Şekil 4.9’da karşılaştırmalı olarak verilmiştir.

Çizelge 4.21. Kanal dengeleme problemi 20. iterasyon için hedef çıkış tablosu (swish, e-swish, DTS, modifiye swish)

Aktivasyon fonksiyonu SWISH		Aktivasyon fonksiyonu E-SWISH ($\beta = 1.4$)		Aktivasyon Fonksiyonu DTS ($T = 1$)		Aktivasyon Fonksiyonu MODİFİYE SWISH ($\alpha = 4$)	
Hata oranı %0.019		Hata oranı %0.016		Hata oranı %0.263		Hata oranı %0.024	
<i>Hedef</i>	<i>Çıkış</i>	<i>Hedef</i>	<i>Çıkış</i>	<i>Hedef</i>	<i>Çıkış</i>	<i>Hedef</i>	<i>Çıkış</i>
<i>0</i>	<i>0.016+0.103i</i>	<i>0</i>	<i>0.005+0.082i</i>	<i>0</i>	<i>0.244+0.986i</i>	<i>0</i>	<i>0.014+0.068i</i>
<i>i</i>	<i>0.025+0.977i</i>	<i>i</i>	<i>0.011+0.991i</i>	<i>i</i>	<i>0.006+1.000i</i>	<i>i</i>	<i>0.025+0.971i</i>
<i>1</i>	<i>0.810+0.310i</i>	<i>1</i>	<i>0.831+0.305i</i>	<i>1</i>	<i>0.662+0.978i</i>	<i>1</i>	<i>0.778+0.363i</i>
<i>1+i</i>	<i>0.914+0.980i</i>	<i>1+i</i>	<i>0.953+0.995i</i>	<i>1+i</i>	<i>0.999+1.000i</i>	<i>1+i</i>	<i>0.948+0.957i</i>



Şekil 4.9. Kanal dengeleme problemi 20. iterasyon için hata grafiği (swish, e-swish, DTS, modifiye swish)

Yapılan 20 iterasyon sonucunda kompleks e-swish ($\beta=1.4$) aktivasyon fonksiyonunu kullanan ağın, bütün iterasyonlarda sahip olduğu hata oranının diğer ağlardan daha düşük olduğu grafik üzerine gösterilmiştir. Ayrıca ağların 20. iterasyonda elde ettiği çıkışlar incelendiğinde sonuca en yakın değerleri kompleks e-swish ($\beta=1.4$) aktivasyon fonksiyonunu kullanan ağın verdiği görülmektedir.

5. SONUÇLAR VE ÖNERİLER

5.1 Sonuçlar

Yapılan tez çalışmasında KDYSA için yeni aktivasyon fonksiyonları tanımlanmıştır.

İlk olarak, önerilen kompleks swish aktivasyon fonksiyonun geçerliliğini ve uygulanabilirliğini test etmek için XOR, simetri ve kanal dengeleme problemleri üzerinde testler yapılmıştır. Yapılan testler literatürde yer alan aktivasyon fonksiyonları hiperbolik tanjant ve sigmoid aktivasyon fonksiyonları ile karşılaştırılmıştır. XOR problemi üzerinde yapılan testler sonucunda, 0.001 hata oranına (OKH) sigmoid fonksiyonunu kullanan ağ 2052., hiperbolik tanjant fonksiyonunu kullanan ağ 1186. iterasyonda ulaşırken, swish fonksiyonunu kullanan ağ 756. iterasyonda ulaşmıştır. Simetri problemi üzerinde yapılan testler sonucunda ise, 0.001 hata oranına (OKH) sigmoid fonksiyonunu kullanan ağ 1204. iterasyonda, hiperbolik tanjant fonksiyonunu kullanan ağ 453. iterasyonda ulaşırken, swish fonksiyonunu kullanan ağ ise 280. iterasyonda ulaşmıştır. Son olarak ise kanal dengeleme problemi üzerinde yapılan testlerde 0.001 hata oranına (OKH) sigmoid fonksiyonunu kullanan ağ 495. iterasyonda, hiperbolik tanjant fonksiyonunu kullanan ağ 74. iterasyonda ulaşırken, swish fonksiyonunu kullanan ağ 52. iterasyonda ulaşmıştır. Sonuç olarak swish aktivasyon fonksiyonunu kullanan ağın, tüm problemler üzerinde sonuca en erken yakınsayan ağ olduğu görülmüştür.

Tez çalışmasında önerilen kompleks e-swish, DTS ve modifiye swish aktivasyon fonksiyonlarının en iyi performans gösterdiği değerler XOR, simetri ve kanal dengeleme problemleri üzerinde belirlenmiştir. Yapılan testler sonucunda kompleks e-swish fonksiyonun; XOR, simetri ve kanal dengeleme problemleri için $\beta=1.4$ sabit değeri ile, kompleks DTS aktivasyon fonksiyonunun; XOR probleminde $T=-0.3$, simetri problemi üzerinde $T=-0.2$ ve kanal dengeleme probleminde $T=1$ sabit değeri ile en iyi performansı gösterdiği belirlenmiştir. Kompleks Modifiye swish aktivasyon fonksiyonunun ise XOR probleminde $\alpha=1.4$, simetri probleminde $\alpha=1.5$ ve kanal dengeleme probleminde $\alpha=4$ sabit değeri ile hedefe daha hızlı yakınsadığı belirlenmiştir.

Son olarak, önerilen kompleks swish, E-swish, DTS ve Modifiye swish aktivasyon fonksiyonları XOR, simetri ve kanal dengeleme problemleri üzerinde birbirleri ile karşılaştırılmıştır. Karşılaştırma yapılırken kompleks e-swish, DTS ve modifiye swish fonksiyonları için bir önceki bölümde belirlenen sabit değerler kullanılmıştır. XOR, simetri ve kanal dengeleme üzerinde yapılan testlerin tamamında

kompleks $\beta=1.4$ sabit değerli E-swish aktivasyon fonksiyonunu kullanan ağı, diğer aktivasyon fonksiyonlarını kullanan ağlara göre sonuca daha erken yakınsadığı görülmüştür. Ayrıca önerilen aktivasyon fonksiyonlarını kullanan ağların sabit iterasyonda elde ettiği çıkışlar ve hata grafikleri karşılaştırmalı olarak incelenmiştir. Sonuçlardan da görüldüğü üzere, kompleks E-swish ($\beta=1.4$) aktivasyon fonksiyonunu kullanan ağı aynı iterasyonda diğer ağlara göre sonuca daha yakın değerler verdiği ve daha az hata oranına sahip olduğu görülmüştür.

Buna göre önerilen aktivasyon fonksiyonlarının kompleks sayılarla işlem yapmayı gerektiren problemlerin çözümünde kullanılarak, YSA'ların performansını arttıracığı sonucuna varılmıştır.

5.2 Öneriler

KDYSA'larının performansı ve kompleks geriye yayılım algoritmasının yakınsaması; başlangıç ağırlıkları, öğrenme oranı, verilerin normalizasyonu ve aktivasyon fonksiyonunun seçimi başta olmak üzere çeşitli etkenlere bağlıdır. Tez çalışması kapsamında, KDYSA'ların performansını ve kararlılığını arttırmak amacıyla yeni aktivasyon fonksiyonları tanımlanarak, ağı performansı üzerindeki etkisi incelenmiştir. Bu amaçla reel düzlemde iyi sonuçlar veren swish aktivasyon fonksiyonu kompleks düzlemde ele alınmıştır. Ayrıca swish fonksiyonu üzerinden türetilen ve yine reel düzlemde iyi sonuçlar veren modifiye swish, e-swish ve DTS aktivasyon fonksiyonları yine kompleks düzlemde incelenmiş ve ağı performansı üzerindeki etkileri gösterilmiştir. İleride yapılacak çalışmalarda KDYSA'ların başlangıç ağırlıkları, hata fonksiyonu, öğrenme oranı ve verilerin normalizasyonu alanında çalışmalar yapılarak, ağı performansı üzerindeki etkisi incelenebilir.

Yapılan çalışmalarda veri olarak kompleks XOR, simetri ve kanal dengeleme problemleri kullanılmış olup yeni yapılacak çalışmalarda daha kapsamlı ve gerçek dünya verileri kullanılarak, önerilen fonksiyonların sınıflandırma problemleri üzerindeki etkileri incelenebilir.

KAYNAKLAR

- Acar, Y. E., Ceylan, M. and Yaldız, E., 2018, An examination on the effect of CVNN parameters while classifying the real-valued balanced and unbalanced data, International Conference on Artificial Intelligence and Data Processing (IDAP), Malatya, 1023-2016
- Alcaide, E., 2018, E-swish: Adjusting activations to different network depths, ArXiv Prepr. ArXiv 1801.07145, 2018, available at: <https://arxiv.org/abs/1801.07145v1>.
- Amin, F. and Murase, K., 2009, Single-layered complex-valued neural network for real valued classification problems, *Neurocomputing* 72(4-6), 945-955.
- Baba, N., Jain, L. C. and Howlett, R. J., 2001, Knowledge-based Intelligent Information Engineering Systems & Allied Technologies, Part I, 550-580, IOS Press, Tokyo.
- Benvenuto, N. and Piazza. F., 1992, On the complex backpropagation algorithm. *IEEE Transactions on Signal Processing*, 40(4):967–969, ISSN 1053-587X.
- Burse, K., Pandey, A. and Somkuwar, A., 2011, Convergence analysis of complex valued multiplicative neural networks for various activation functions, International Conference on Computational Intelligence and Communication Systems, 279-282
- Ceylan, M., 2004, Kompleks değerli yapay sinir ağı ile algoritma geliştirilmesi ve uygulamaları, Yüksek lisans tezi, *Selçuk Üniversitesi Fen Bilimleri Enstitüsü*, Konya, 32-42.
- Ceylan, M., 2009, Bilgisayar tomografili akciğer görüntülerinin değerlendirilmesinde kompleks değerli yeni bir akıllı sistem tasarımı, Doktora tezi, *Selçuk Üniversitesi Fen Bilimleri Enstitüsü*, Konya, 53-58.
- Ceylan, M., 2011, Combined complex-valued artificial neural network (CCVANN), Proceedings of the World Congress on Engineering 2011 Vol II WCE 2011, London, U.K.
- Chen, X. and Tang, Z., 2005, An Modified error function for the complex-value backpropagation neural networks, *Neural Information Processing - Letters and Reviews* Vol.8, No.1,
- Chen, X., Tang, Z. and Variappan, C., 2005, A Modified error backpropagation algorithm for complex-value neural networks, *International Journal of Neural Systems*, Vol. 15, No. 6, 435-443
- Chieng, H. H., Wahid, N., Pauline, O. and Perla, S. R. K., 2018, Flatten T-Swish: a thresholded ReLU-Swish-like activation function for deep learning, *International Journal of Advances in Intelligent Informatics*, Vol. 4, No. 2, pp. 76-86
- Clarke, T. L., 1990, Generalization of neural networks to the complex plane, *IJCNN International Joint Conference on Neural Networks*, pages 435–440 vol.2.

- Çelebi, M. and Ceylan, M., 2019, The New Activation Function for Complex Valued Neural Networks: Complex Swish Function, 4th International Symposium on Innovative Approaches in Engineering and Natural Sciences, 4 (6), 169-173.
- Çelebi, M. and Ceylan, M., 2020, New complex valued Activation functions: complex modified swish, complex e-swish and complex flatten t-swish, International Journal of Advanced Research in Computer Science, Volume 11, No. 2, 16-23.
- Damiani, E., Howlett, R. J., Jain, L. C. and Ichalkaranje, N., 2002, Knowledge based Intelligent Information Engineering Systems & Allied Technologies, Part I, 623-647, IOS Press, Amsterdam.
- Enomoto, T., Kakuda, K. and Miura, S., 2019, New Activation Functions in CNN and Its Applications, ICCES, vol.1, no.2, pp.36-39.
- Georgiou, G. M. and Koutsougeras, C., 1992, Complex domain backpropagation, IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, 39(5):330–334, ISSN 1057-7130.
- Gödekoğlu, D. A., 2009, Yapay sinir ağları kullanılarak solunum fonksiyon testleri ile cinsiyet, obezite ve sigara kullanımının ilişkilendirilmesi, Yüksek lisans tezi, *Sakarya Üniversitesi Fen Bilimleri Enstitüsü*, Sakarya, 17-40.
- Gürüler, H. and Peker, M., 2015, A software tool for complex-valued neural network: CV-ANN, 23rd Signal Processing and Communications Applications Conference (SIU), Malatya.
- Haykin, S., 1999, Neural Networks, Second Edition, Prentice Hall
- Hirose, A., 2009, Complex-valued neural networks: The merits and their origins, International Joint Conference on Neural Networks, pp. 1237–1244.
- Ikeda, A. and Washizawa, Y., 2019, Spontaneous EEG Classification Using Complex Valued Neural Network , 26th International Conference, ICONIP 2019 Sydney, NSW, Australia, Part IV, pp. 495-503.
- Jalab, H. A. and Ibrahim, R. W., 2011, New activation functions for complex-valued neural network, International Journal of the Physical Sciences Vol. 6(7), pp. 1766-1772
- Kara, S., Latifoğlu, F., Ceylan, R. and Ceylan, M., 2007, Complex-valued wavelet artificial neural network for Doppler signals classifying, Artificial Intelligence in Medicine , 143-156
- Leung, H. and Haykin, S., 1991, The complex backpropagation algorithm, IEEE Transactions On Signal Processing, Vol. 39, No. 9.
- Li, S. and Jiang, M., 2013, The new complex-valued wavelet neural network, TELKOMNIKA, Vol.12, No.3, 613-622.

- Michel, H. E. and Awwal, A. A. S., 1999, Enhanced artificial neural networks using complex numbers, IJCNN'99 International Joint Conference on Neural Networks, Vol. 1 (1999) 456–461.
- Mönning, N. and Manandhar, S., 2018, Evaluation of complex-valued neural networks on real-valued classification tasks, ArXiv Prepr. arXiv:1811.12351v1.
- Nitta, T., 2008, Complex-valued neural network and complex-valued backpropagation learning algorithm, Advances in Imaging and Electron Physics, Volume 152, ISSN 1076-5670.
- Nitta, T., 2004, Orthogonality of decision boundaries in complex-valued neural networks, Neural Computation 16, 73–97.
- Nitta, T., 2003a, Solving the XOR problem and the detection of symmetry using a single complex valued neuron, Neural Networks, 16:1101-1105.
- Nitta, T., 2003b, The computational power of complex-valued neuron, ICANN/ICONIP, LNCS 2714, pp. 993–1000.
- Nitta, T., 1993, A back-propagation algorithm for complex numbered neural networks, International Conference on Neural Networks IJCNN-93, Nagoya, Japan, volume 2, pp. 1649–1652.
- Oppenheim, A.V.; Lim, J.S., The importance of phase in signals, IEEE Proceedings, v. 69, No 5, 1981, pp.: 529- 541
- Özgür, N., Taş, N. and Peters, J. F., 2020, New complex-valued activation functions, An International Journal of Optimization and Control: Theories & Applications Vol.10, No.1, pp.66-72.
- Özkurt, S., Bostancı, M., Altın R., Özşahin, A. ve Akdağ, B., 2000, Tıp fakültesi çalışanlarında sigara içme prevalansı, Nikotin Bağımlılığı ve Solunum Fonksiyon Testleri, Tüberküloz ve Toraks Dergisi, 48(2): 140-147.
- Peker, M. and Sen, B., 2013, A new complex-valued intelligent system for automated epilepsy diagnosis using EEG signals, 3rd World Conference on Information Technology (WCIT), Vol 03, 1121-1128.
- Peker, M., Sen, B., and Delen, D., 2016, A novel method for automated diagnosis of epilepsy using complex-valued classifiers, IEEE Journal of Biomedical and Health Informatics, 20(1), 108-118.
- Ramachandran, P., Zoph, B. and Le, Q. V., 2017, Searching for activation functions, In International Conference on Learning Representations Q. V. L.
- Şen, C., 2018, İçme Suyu Tesislerinde SCADA Sistemlerinde Sezgisel Yaklaşımların Uygulanması, *Sakarya Üniversitesi Fen Bilimleri Enstitüsü*, Yüksek Lisans Tezi, 83-86.

Yaşar, H., 2015, Medikal görüntülerin çoklu çözünürlük metotları ile analizi, Yüksek lisans tezi, *Selçuk Üniversitesi Fen Bilimleri Enstitüsü*, Konya, 37-41.



EKLER

EK-1 Swish aktivasyon fonksiyonu ve türevinin tanımlanmasına ait MATLAB kodları

```
function y=swish(x)  
y=x.*logsig(x);
```

```
function y=dswish(x)  
y=swish(x)+logsig(x).*(1-swish(x));
```



EK-2 E-Swish aktivasyon fonksiyonu ve türevinin tanımlanmasına ait MATLAB kodları

```
function y=eswish(x)  
B=1.4;  
y=B.*x.*logsig(x);
```

```
function y=deswish(x)  
B=1.4;  
y=eswish(x)+logsig(x).*(B-eswish(x));
```



EK-3 Düzleştirilmiş t-swish aktivasyon fonksiyonu ve türevinin tanımlanmasına ait MATLAB kodları

```
function y=DTS(x)
T=-0.3;
if x<0
    y=T;
else
    y=swish(x)+T;
end
```

```
function y=dDTS(x)
T=-0.3;
if x<0
    y=0;
else
    y=DTS(x)+logsig(x).*(1-DTS(x));
end
```

EK-4 Modifiye Swish aktivasyon fonksiyonu ve türevinin tanımlanmasına ait MATLAB kodları

```
function y=modswish(x)
M=1.4;
y=x.*logsig(M.*x);
```

```
function y=dmodswish(x)
M=1.4;
y=M.*modswish(x)+logsig(M.*x).*(1-M.*modswish(x));
```



EK-5 Swish aktivasyon fonksiyonunu kullanan KDYSA MATLAB kodları

```

function [ysa_out1,Errors1] =
cvnn_swish(data,w1,w2,bias1,bias2,l_in,max_it,lr,n_n,row,n_
training_data,l_out,do)

for N=1:max_it
    Ep=0;
    k=1;
    cRatet=0;
for T=1:n_training_data

    x=data(T,1);
    y_desired=data(T,2);
    %% operations in section 1

    for j=1:n_n
        y(j)=activate_swish(x*w1(:,j)+bias1(j));
    end

    %% operations between hidden layer and output
    for j=1:l_out
        out(j)=activate(y*w2(:,j)+bias2(j));
    end

    %% evaluating the output error and gradient
    e_out=y_desired-out;
    Ep=Ep+(sum(abs(e_out)).^2)./2;

    db2r=(real(e_out)).*(1-real(out)).*real(out);
    db2i=(imag(e_out)).*(1-imag(out)).*imag(out);
    db2=lr*db2r+lr*1i*db2i;%%delta_bias
    %% evaluating the hidden layers gradients
    for i=1:n_n
        rr_sum=0;
        ii_sum=0;
        ri_sum=0;
        ir_sum=0;
        for k=1:l_out
            rr_sum=rr_sum+db2r(k)*real(w2(i,k));
            ii_sum=ii_sum+db2i(k)*imag(w2(i,k));
            ri_sum=ri_sum+db2r(k)*imag(w2(i,k));
            ir_sum=ir_sum+db2i(k)*real(w2(i,k));
        end
        db1r(i)=dswish(real(y(i)))*(rr_sum+ii_sum);
        db1i(i)=-dswish(imag(y(i)))*(ri_sum-ir_sum);
        db1=lr*db1r+1i*lr*db1i;
    end
end
end

```

```

        end
        %% evaluating the deltas in the section 2 (hidden
layer-output)
        for i=1:n_n
            for j=1:l_out
                delta_w2(i,j)=conj(y(i))*db2(j);
            end
        end
        end
        %% evaluating the deltas in the section 1 (input-
hidden layer)
        for i=1:l_in
            for j=1:n_n
                delta_w1(i,j)=conj(x(i))*db1(j);
            end
        end
        end

        %% updating all weights and biases
        w1=w1+delta_w1;
        w2=w2+delta_w2;
        bias1=bias1+db1;
        bias2=bias2+db2;

        ysa_out3(T, :, N)=out;

    end

    Errors3(N, :)=Ep./row;
    if N>1
        if Errors3(N)<do, break, end; %durdurma kriteri

    end

end
end

```

EK-6 E-Swish aktivasyon fonksiyonunu kullanan KDYSA MATLAB kodları

```

function [ysa_out2,Errors2] =
cvnn_eswish(data,w1,w2,bias1,bias2,l_in,max_it,lr,n_n,row,n
_training_data,l_out,do)

for N=1:max_it
    Ep=0;
    k=1;
    cRatet=0;
for T=1:n_training_data

    x=data(T,1);
    y_desired=data(T,2);
    %% operations in section 1

for j=1:n_n
    y(j)=activate_eswish(x*w1(:,j)+bias1(j));
end

    %% operations between hidden layer and output
for j=1:l_out
    out(j)=activate(y*w2(:,j)+bias2(j));
end

    %% evaluating the output error and gradient
    e_out=y_desired-out;
    Ep=Ep+(sum(abs(e_out)).^2)./2;

    db2r=(real(e_out)).*(1-real(out)).*real(out);
    db2i=(imag(e_out)).*(1-imag(out)).*imag(out);
    db2=lr*db2r+lr*1i*db2i;%%delta_bias
    %% evaluating the hidden layers gradients
for i=1:n_n
    rr_sum=0;
    ii_sum=0;
    ri_sum=0;
    ir_sum=0;
for k=1:l_out
    rr_sum=rr_sum+db2r(k)*real(w2(i,k));
    ii_sum=ii_sum+db2i(k)*imag(w2(i,k));
    ri_sum=ri_sum+db2r(k)*imag(w2(i,k));
    ir_sum=ir_sum+db2i(k)*real(w2(i,k));
end
    db1r(i)=deswish(real(y(i)))*(rr_sum+ii_sum);
    db1i(i)=-deswish(imag(y(i)))*(ri_sum-ir_sum);
    db1=lr*db1r+1i*lr*db1i;

```

```

        end
        %% evaluating the deltas in the section 2 (hidden
layer-output)
        for i=1:n_n
            for j=1:l_out
                delta_w2(i,j)=conj(y(i))*db2(j);
            end
        end
        end
        %% evaluating the deltas in the section 1 (input-
hidden layer)
        for i=1:l_in
            for j=1:n_n
                delta_w1(i,j)=conj(x(i))*db1(j);
            end
        end
        end

        %% updating all weights and biases
        w1=w1+delta_w1;
        w2=w2+delta_w2;
        bias1=bias1+db1;
        bias2=bias2+db2;

        ysa_out3(T, :, N)=out;

    end

    Errors3(N, :)=Ep./row;
    if N>1
        if Errors3(N)<do, break, end; %durdurma kriteri

    end

end

```

EK-7 Düzleştirilmiş t-swish aktivasyon fonksiyonunu kullanan KDYSA MATLAB kodları

```

function [ysa_out3,Errors3] =
cvnn_DTS(data,w1,w2,bias1,bias2,l_in,max_it,lr,n_n,row,n_training_data,l_out,do)

for N=1:max_it
    Ep=0;
    k=1;
    cRatet=0;
for T=1:n_training_data

    x=data(T,1);
    y_desired=data(T,2);
    %% operations in section 1

    for j=1:n_n
        y(j)=activate_DTS(x*w1(:,j)+bias1(j));
    end

    %% operations between hidden layer and output
    for j=1:l_out
        out(j)=activate(y*w2(:,j)+bias2(j));
    end

    %% evaluating the output error and gradient
    e_out=y_desired-out;
    Ep=Ep+(sum(abs(e_out)).^2)./2;

    db2r=(real(e_out)).*(1-real(out)).*real(out);
    db2i=(imag(e_out)).*(1-imag(out)).*imag(out);
    db2=lr*db2r+lr*1i*db2i;%%delta_bias
    %% evaluating the hidden layers gradients
    for i=1:n_n
        rr_sum=0;
        ii_sum=0;
        ri_sum=0;
        ir_sum=0;
        for k=1:l_out
            rr_sum=rr_sum+db2r(k)*real(w2(i,k));
            ii_sum=ii_sum+db2i(k)*imag(w2(i,k));
            ri_sum=ri_sum+db2r(k)*imag(w2(i,k));
            ir_sum=ir_sum+db2i(k)*real(w2(i,k));
        end
        db1r(i)=dDTS(real(y(i)))*(rr_sum+ii_sum);
        db1i(i)=-dDTS(imag(y(i)))*(ri_sum-ir_sum);
        db1=lr*db1r+1i*lr*db1i;
    end
end
end

```

```

        end
        %% evaluating the deltas in the section 2 (hidden
layer-output)
        for i=1:n_n
            for j=1:l_out
                delta_w2(i,j)=conj(y(i))*db2(j);
            end
        end
        %% evaluating the deltas in the section 1 (input-
hidden layer)
        for i=1:l_in
            for j=1:n_n
                delta_w1(i,j)=conj(x(i))*db1(j);
            end
        end

        %% updating all weights and biases
        w1=w1+delta_w1;
        w2=w2+delta_w2;
        bias1=bias1+db1;
        bias2=bias2+db2;

        ysa_out3(T, :,N)=out;

    end

    Errors3(N, :)=Ep./row;
    if N>1
        if Errors3(N)<do, break, end; %durdurma kriteri

    end

end

```

EK-8 Modifiye-Swish aktivasyon fonksiyonunu kullanan KDYSA MATLAB kodları

```

function [ysa_out4,Errors4] =
cvnn_modswish(data,w1,w2,bias1,bias2,l_in,max_it,lr,n_n,row
,n_training_data,l_out,do)

for N=1:max_it
    Ep=0;
    k=1;
    cRatet=0;
for T=1:n_training_data

    x=data(T,1);
    y_desired=data(T,2);
    %% operations in section 1

for j=1:n_n
    y(j)=activate_modswish(x*w1(:,j)+bias1(j));
end

    %% operations between hidden layer and output
for j=1:l_out
    out(j)=activate(y*w2(:,j)+bias2(j));
end

    %% evaluating the output error and gradient
    e_out=y_desired-out;
    Ep=Ep+(sum(abs(e_out)).^2)./2;

    db2r=(real(e_out)).*(1-real(out)).*real(out);
    db2i=(imag(e_out)).*(1-imag(out)).*imag(out);
    db2=lr*db2r+lr*1i*db2i;%%delta_bias
    %% evaluating the hidden layers gradients
for i=1:n_n
    rr_sum=0;
    ii_sum=0;
    ri_sum=0;
    ir_sum=0;
for k=1:l_out
    rr_sum=rr_sum+db2r(k)*real(w2(i,k));
    ii_sum=ii_sum+db2i(k)*imag(w2(i,k));
    ri_sum=ri_sum+db2r(k)*imag(w2(i,k));
    ir_sum=ir_sum+db2i(k)*real(w2(i,k));
end
    db1r(i)=dmodswish(real(y(i)))*(rr_sum+ii_sum);
    db1i(i)=-dmodswish(imag(y(i)))*(ri_sum-ir_sum);
    db1=lr*db1r+1i*lr*db1i;

```

```

        end
        %% evaluating the deltas in the section 2 (hidden
layer-output)
        for i=1:n_n
            for j=1:l_out
                delta_w2(i,j)=conj(y(i))*db2(j);
            end
        end
        end
        %% evaluating the deltas in the section 1 (input-
hidden layer)
        for i=1:l_in
            for j=1:n_n
                delta_w1(i,j)=conj(x(i))*db1(j);
            end
        end
        end

        %% updating all weights and biases
        w1=w1+delta_w1;
        w2=w2+delta_w2;
        bias1=bias1+db1;
        bias2=bias2+db2;

        ysa_out4(T, :, N)=out;

    end

    Errors4(N, :)=Ep./row;
    if N>1
        if Errors4(N)<do, break, end; %durdurma kriteri

    end

end
end

```

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : Mehmet ÇELEBİ
Uyruğu : Türkiye Cumhuriyeti
Doğum Yeri ve Tarihi : Ankara – 09.06.1987
Telefon : (+90) 543 978 11 45
E-Posta : mehmetc@ilbank.gov.tr

EĞİTİM

Derece	Adı	İlçe	İl	Bitirme Yılı
Lise	: Atatürk Lisesi, Çankaya, Ankara			2005
Üniversite	: Selçuk Üniversitesi, Selçuklu, Konya			2010
Yüksek Lisans	: Konya Teknik Üniversitesi, Selçuklu, Konya			Devam Ediyor
Doktora	:			

İŞ DENEYİMLERİ

Yıl	Kurum	Görevi
2011-2014	Yılmaz Elektrik San.Tic.Ltd.Şti.	Saha Mühendisi
2014-.....	İller Bankası A.Ş.	Teknik Uzman

UZMANLIK ALANI

Yapay Sinir Ağları, İçmesuyu ve Atıksu Tesisleri Elektrik Otomasyon Projeleri

YABANCI DİLLER

İngilizce

YAYINLAR

Çelebi, M. and Ceylan, M., 2019, The New Activation Function for Complex Valued Neural Networks: Complex Swish Function, 4th International Symposium on Innovative Approaches in Engineering and Natural Sciences, 4 (6), Samsun.

Çelebi, M. and Ceylan, M., 2020, New complex valued activation functions: complex modified swish, complex e-swish and complex flatten t-swish, International Journal of Advanced Research in Computer Science, Volume 11, No. 2.