

Yapay Sinir Ağlarında Yeni Bir Ön Değer Atama Yöntemi: Laplasyen

A New Initialization Method for Artificial Neural Networks: Laplacian

Burak Çatalbaş, Bahadır Çatalbaş, Ömer Morgül
Elektrik ve Elektronik Mühendisliği Bölümü
Bilkent Üniversitesi
Ankara, Türkiye
{catalbas,cbahadir,morgul}@ee.bilkent.edu.tr

Özetçe —Makine öğrenimi alanında yapay sinir ağlarının popülerliği 2006 yılında derin öğrenme kavramının yerleşmesi sonrasında günden güne artmaktadır. Derin sinir ağlarının eğitiminde başarı yüzdelerini önemli oranda etkileyen etkenlerden biri de ön değerlerin atanmasıdır. Bu makalede Laplasyen dağılımını temel alan yeni ön değer atama yöntemleri önerilmiştir. Bu yeni ön değer atama yöntemlerinin kullanımı ile ağ parametrelerine, sinir ağının daha iyi eğitilmesini sağlayacak şekilde, uygun ön değerlerin atanması amaçlanmaktadır. Yöntemlerimizin Kaliforniya Üniversitesi Irvine Kampüsü (UCI) İnsan Hareketi Tanımlama ve CIFAR-10 veri setlerinde aldıkları sonuçlar ağ formasyonu ve katman yapısı değiştirilmeden Tekdüze ve Gauss gibi bilinen ön değer atama yöntemlerinin aldığı sonuçlar ile karşılaştırılmıştır. Bu karşılaştırma ile Laplasyen temelli ön değer atama yöntemlerinin var olan diğer yöntemlere kıyasla avantajları test başarıları göz önüne alınarak tartışıldı.

Anahtar Kelimeler—yapay sinir ağları; derin öğrenme; sınıflandırıcı; ön değer atama.

Abstract—Artificial neural networks' popularity in the field of machine learning increases day by day since 2006, foundation date of deep learning. One of the factors which greatly affects the success percentages of deep neural networks is their initialization. In this article, new initialization methods based on Laplacian distribution is proposed. With the use of these new initialization methods, it is aimed to assign appropriate initial values to the network parameters so as to better train the network. Results of our methods on University of California, Irvine (UCI) Human Activity Recognition and CIFAR-10 datasets are compared with the networks which are initialized with well-known methods, such as Gaussian and Uniform initialization, while network formation and layer structure are left unchanged. With this comparison, the advantages of Laplacian-based initialization methods compared to existing methods were discussed considering the test success.

Keywords—artificial neural networks; deep learning; classifier; initialization.

I. GİRİŞ VE MOTİVASYON

Yapay Sinir Ağları (Artificial Neural Networks) günümüzde endüstriden eğitime, savunmadan ekonomiye kadar birçok sektörde, gerek sınıflandırma görevlerinde,

gerekse tahmin algoritmalarında geniş bir kullanım alanına ulaşmıştır. Literatürde farklı yapay sinir ağı yapıları ve farklı amaçlı kullanımları bulunmaktadır. Bunlara ön-beslemeli sinir ağları (feedforward neural networks) ile karakter tanımlamadan [1] düşme tespitine [2], evrimsel sinir ağları (convolutional neural networks) ile resim sınıflandırmadan [3] nesne algılamaya [4], tekrarlı sinir ağları ile robot hareketinin sağlanmasından [5] konuşma tanımlanmasına [6] kadar farklı uygulamalar örnek verilebilir. Tüm bu uygulamaların ortak noktası ağıdaki serbest parametrelerin eğitim yoluyla verilen görevi gerçekleştirebilecek şekilde uygun hale getirilmesidir. Bu noktada parametre ilk değerleri kritik bir roledir.

Geçtiğimiz yüzyılda basitleştirilmiş sinir modelinin öne sürülmesi [7], çok katmanlı sinir ağları ile sınıflandırılmanın yapılabileceğinin gösterilmesi ve sinir ağlarının eğitimini kolaylaştıran geri yayılma (Backpropagation) metodunun geliştirilmesi gibi önemli çalışmalar ile temelleri atılan yapay sinir ağları, 21. yüzyılda artan veri miktarının ve işleme kapasitesi ile önemli ilerleme kaydetmiş ve insan performansını geçen sınıflandırma başarıları göstererek [8] makine öğrenimi alanının vazgeçilmez bir parçası haline gelmiştir. Bu ilerlemeler arasında "Dropout" [9] ve rezervuar katmanı [10] gibi yapısal eklemeler bulunurken, halihazırda kullanılan yapıların optimizasyonu da hedeflenmiştir. Çapraz entropi ve momentum gibi tekniklerin başarı oranlarını arttırmada ve eğitim süresini azaltmada sinir ağının yapısında büyük değişiklikler yapmadan faydalı olduğu görülmüştür.

Bu çalışmada sınıflandırma görevinde yaygın şekilde kullanılan Ön-beslemeli Sinir Ağı ve Evrimsel Sinir Ağı yapıları, Kaliforniya Üniversitesi Irvine Kampüsü (UCI) İnsan Hareketi Tanımlama [11] ve CIFAR-10 [12] veri setleri üzerinde sırasıyla eğitilmiştir. Eğitimler sırasında literatürde var olan ön değer atama (initialization) yöntemlerine ek olarak önerdiğimiz Laplasyen (Laplacian) temelli iki farklı ön değer atama yöntemi de ayrıca uygulanmıştır. Sinir ağı formasyonu ve katman yapısı değiştirilmeden yapılan bu denemelerde elde edilen sonuçlar her iki eğitim seti üzerinde karşılaştırılmış, önerilen ön değer atama tekniğinin başarı oranları literatürdeki yöntemlerin sonuçlarıyla kıyaslanmıştır.

II. VERİ SETİ VE MEVCUT LİTERATÜR

A. Veri Setleri

Bu çalışmada araştırma alanında tanınan iki farklı veri seti kullanılmıştır. İnsan Hareketi Tanımlama veri seti durma, oturma, yatma, yürüme, aşağı ve yukarı yürüme hareket sınıflarının hız, ivme gibi 561 elemanlı öznitelik vektörlerini içermekte, CIFAR-10 veri seti ise uçak, araba, kuş, kedi, geyik, köpek, kurbağa, kamyon, at veya gemi içeren renkli obje resimlerinden oluşmaktadır. İnsan Hareketi Tanımlama veri seti 7352 eğitim örneğine ve 2947 test örneğine sahipken, CIFAR-10 veri seti ise 50000 eğitim örneği ve 10000 test örneği içermektedir. Bununla beraber verinin kompleksliği dolayısıyla CIFAR-10 seti zorluk olarak öne çıkmaktadır.

B. Literatürdeki Sonuçlar

İnsan Hareketi Tanımlama veri setinde en başarılı sonuç %96,4 olarak elde edilmiştir [13]. CIFAR-10 veri setinde ise en başarılı sonuç %97,14 olarak elde edilmiştir [14]. Her ne kadar CIFAR-10 veri setinde rapor edilen başarı oranları daha yüksek olsa da, bunun nedeni görüntü verilerinin daha zengin öznitelikler içermesinin yanı sıra kullanılabilen yapay sinir ağlarının daha gelişkin olmasıdır. Literatürde rapor edilen diğer sonuçlar da göz önüne alındığında CIFAR-10 veri seti için %90 ve İnsan Hareketi Tanımlama veri seti için %95 civarında sonuçlar veren yapay sinir ağı modelleri metod iyileştirmelerinin sonuçlarını göstermesi bakımından yeterli görülmüştür.

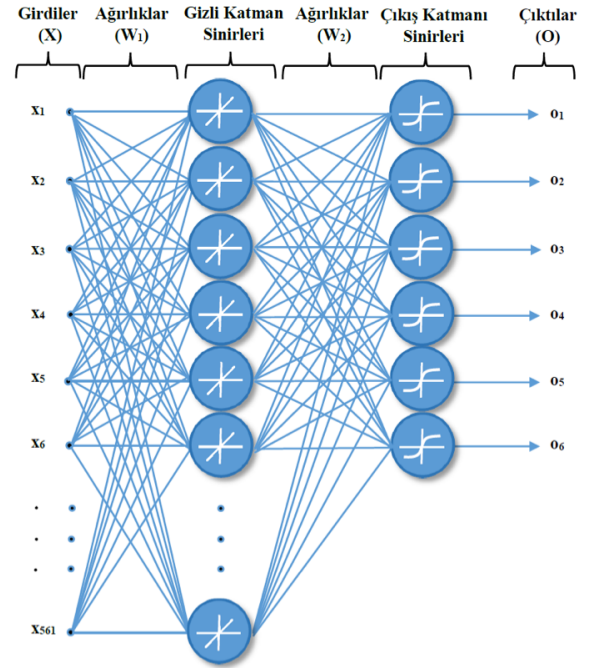
III. KULLANILAN METOTLAR

A. Yapay Sinir Ağı Yapıları ve Kullanılan Yöntemler

İlk veri setinde ön-beslemeli bir yapay sinir ağı yapısı tercih edilirken, CIFAR-10 veri seti için daha gelişkin yapıda, "tek kule" (one-tower) türü evrişimsel yapay sinir ağı kullanılmıştır. Veri setlerinin zorluğuna göre kullanılan metotlarda değişiklikler yapılmıştır. Ayrıca, sadece CIFAR-10 veri setinde veri artırma (data augmentation) uygulandıktan sonra sinir ağına girdiler verilmiştir, böylece bu eğitimde aşırı uyum (overfitting) engellenmeye çalışılmıştır. Veri artırma metotlarıysa farklı oranda yaklaştırma, döndürme, kesme, yatay veya dikey kaydırma ve yatay ters çevirme olarak seçilmiştir.

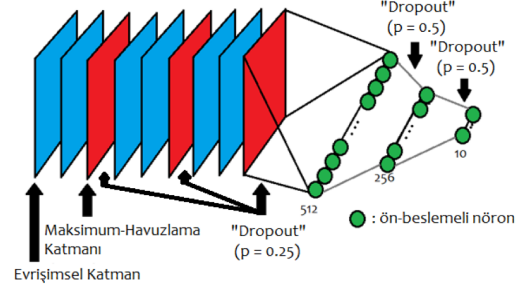
İnsan Hareketi Tanımlama veri seti için tasarlanan yapay sinir ağı iki katmanlı ve ön-beslemeli olup Şekil 1'deki verilen yapıya sahiptir. 6 çıktı sınıfa sahip olan bu sinir ağına girdi nöronu sayısı ise veri setindeki özellik sayısı gibi 561 tanedir. Eğitim tekniği olarak Gradyan Alçalma (Gradient Descent) prensibi ile çalışan Geri Yayılma kullanılmış, kısaca çevrimiçi öğrenim (online learning) stokastik gradyan alçalma algoritmasıyla gerçekleştirilmiştir. Kayıp fonksiyonu olarak ortalama karesel hata formülü kullanılırken aktivasyon fonksiyonu olarak ise hiperbolik tanjant belirlenmiştir. Farklı öğrenim katsayıları da denenilen bu eğitimde rezervuar katmanı olan yapılar denenmiştir. Asıl değişken olarak ise farklı ön değer atama teknikleri seçilmiş ve en başarılı yöntem aranmıştır.

CIFAR-10 veri seti için Kaur tarafından tasarlanan yapay sinir ağı yapısını temel alan, arasında evrişimsel, maksimum-havuzlama ve ön-beslemeli katmanların bulunduğu, 12 katmanlı, evrişimsel bir sinir ağı modeli geliştirilmiştir [15]. Eğitim tekniği olarak AdaMax temelli geri yayılma kullanılırken [16] öğrenimde çevrimiçi öğrenimden farklı olarak



Şekil 1: Ön-beslemeli rezervuar katmanlı yapay sinir ağı yapısı

128 örnek içeren mini-yığın öğrenimi yapılmıştır. Kayıp fonksiyonu olarak kategorik çapraz entropi (categorical cross entropy) seçilmiştir, aktivasyon fonksiyonları ise ReLU ve Softmax'tir. Eğitimi güçlendirmek için bazı katmanlarda farklı olasılık değerleriyle "Dropout" uygulanmıştır. Veri artırma tekniklerinin de kullanıldığı bu ağı yapısı Şekil 2'dedir.



Şekil 2: Evrişimsel yapay sinir ağı yapısı.

Bu yapay sinir ağına ise tüm katmanlardaki nöron sayıları aynı tutulmakla beraber, rezervuar katmanı bulunmamaktadır. Dolayısıyla bu veri seti için oluşturulan yapay sinir ağlarındaki tek farklılık ön değer atama yöntemleri olmakta ve bu veri setinde en başarılı olan yöntemin bulunması hedeflenmektedir.

Bu yapay sinir ağlarında denenecek olan dağılımlar her iki veri seti için de farklı ön değer atama yöntemleri olmakla beraber, İnsan Hareketi Tanımlama veri seti için Tekdüze (Uniform) dağılım (1), Gauss (Gaussian) dağılımı (2) ve önerdiğimiz Laplasyen (Laplacian) dağılımı (3) kullanılmıştır. CIFAR-10 veri setinde ise bu dağılımlara ek olarak, ilgili katmandaki nöronların girdi sayısına (#) göre budanmış He tipi Tekdüze dağılım (He Uniform Distribution) (4) ve budanmış He tipi Gauss dağılımı (He Normal Distribution) (5) ve önerdiğimiz budanmış Laplasyen dağılımı (Truncated Laplacian Distribution) (6) ile (7) eklenmiş ve bunların başarıları karşılaştırılmıştır. Budanmayan dağılım fonksiyonları Şekil 3'te, ilgili katmandaki nöronların girdi sayısına göre budanmış He tipi benzeri fonksiyonlar ise Şekil 4'te gösterilmiştir [17].

Budanmış dağılımların sınır değerleri ilgili katmandaki nöronların girdi sayısına bağlı bir şekilde bulunur. Oluşturulan dağılımlardan rastgele alınan ön değerler sinir ağının ağırlık katsayıları olarak atanır. Dağılım fonksiyonlarımız budanmış Tekdüze dağılım (4), budanmış Gauss dağılımı (5) ile $TC_1 = 12$ ve $TC_2 = 13,5$ değerlerine göre $T = \sqrt{TC/\#}$ formülü ile bulunan T_1 ve T_2 budama limitleri belirlenmiş budanmış olan Laplasyen dağılımları (6), (7) olarak seçilmiştir. Nöron girdi sayıları dışında (4) ve (5)'de olan 6 ve 2 budama sabitleri uygun bulunmuştur [17]. Önerilen budanmış Laplasyen yönteminde bu değerlerin yerine $TC_1 = 12$ ve $TC_2 = 13,5$ budanma sabitleri denenip hem ortalama hem de maksimum başarı oranlarındaki sonuçlar araştırılmıştır.

$$P_1(x) = \begin{cases} \frac{1}{2L} & , \text{eğer } -L \leq x < L \\ 0 & , \text{diğer} \end{cases} \quad (1)$$

$$P_2(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \mu = 0, \sigma = 1 \quad (2)$$

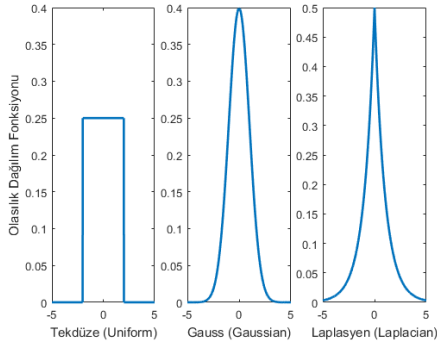
$$P_3(x) = \frac{1}{2\sigma} e^{-\frac{|x-\mu|}{\sigma}}, \mu = 0, \sigma = 1 \quad (3)$$

$$P_4(x) = \begin{cases} \frac{1}{2\sqrt{\frac{6}{\#}}} & , \text{eğer } -\sqrt{\frac{6}{\#}} \leq x < \sqrt{\frac{6}{\#}} \\ 0 & , \text{diğer} \end{cases} \quad (4)$$

$$P_5(x) = \begin{cases} \frac{P_2(x)}{\int_{-\sqrt{\frac{2}{\#}}}^{\sqrt{\frac{2}{\#}}} P_2(x) dx} & , \text{eğer } -\sqrt{\frac{2}{\#}} \leq x < \sqrt{\frac{2}{\#}} \\ 0 & , \text{diğer} \end{cases} \quad (5)$$

$$P_6(x) = \begin{cases} \frac{P_3(x)}{\int_{-T_1}^{T_1} P_3(x) dx} & , \text{eğer } -T_1 \leq x < T_1 \\ 0 & , \text{diğer} \end{cases} \quad (6)$$

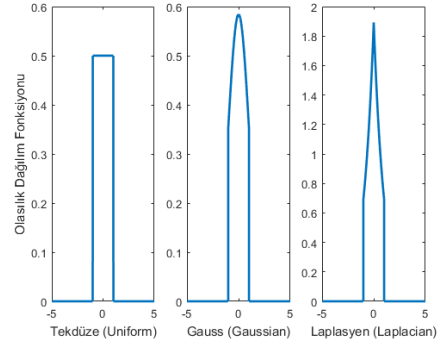
$$P_7(x) = \begin{cases} \frac{P_3(x)}{\int_{-T_2}^{T_2} P_3(x) dx} & , \text{eğer } -T_2 \leq x < T_2 \\ 0 & , \text{diğer} \end{cases} \quad (7)$$



Şekil 3: Budanmamış dağılım fonksiyonları.

B. Eğitim ve Test Uygulamaları

İnsan Hareketi Tanımlama veri seti için kullanılan ön-beslemeli yapay sinir ağı yapısı Şekil 1'de gösterildiği gibi olup, orta katman (rezervuar katmanı) nöron sayısı 1000 olacak şekilde seçilmiştir. Her eğitim için iki farklı öğrenme katsayısı denenmiştir. Eğitimde çevrimiçi öğrenmeyi ön yayılma ve geri yayılma ile gerçekleştiren bir 'döngü' 330000 kez tekrar eder, her döngüde sadece bir örnek kullanılır. Kullandığımız daraltılmış eğitim setinde 6 sınıftan 6969 tane örnek bulunur ve bu veri setine herhangi bir ön işlem uygulanmaz. Her 100



Şekil 4: Budanmış dağılım fonksiyonları.

döngü sonrasında daraltılmış eğitim setinden ayrı tutulan, fakat eğitim veri setinin içinde yer alan 383 tane örneği doğrulama veri seti olarak kullanır. Eğitim sırasında doğrulama setinde en yüksek başarıya sahip ağırlık matrisi (yapay sinir ağı katsayıları) saklanır. Bu eğitim prosedürü birden fazla ön değer atama yöntemi ile ayrı ayrı tekrarlanmış ve bu yöntemlerden hangisinin daha iyi olduğuna her yöntem için yapılan 10 eğitilmiş sinir ağının test sonuçlarına bakılarak karar verilir. Bu aşamada 6 sınıftan 2947 örnek içeren test setinde maksimum başarı alan ağ bulunur. Böylece aynı ağ yapısıyla çalışan ön değer atama yöntemlerinden en başarılı olanı belirlenir.

CIFAR-10 veri seti için kullanılan evrimsel yapay sinir ağının yapısı Şekil 2'de belirtilmiş olup, önceki ön-beslemeli yapay sinir ağına göre daha gelişkin bir formasyona ve fazla sayıda parametreye sahiptir. Gücünü farklı katman çeşitleri ve katsayı fazlalığından alan bu sinir ağı tipinin eğitiminde gerçek eğitim setinin ilk 1000 örneği doğrulama seti olarak, geriye kalan 49000 örneği ise eğitim seti olarak ayrılmıştır. Eğitim prosedürü bu kez eğitim örneklerinin mini yığınlar halinde, önceden bahsedilen veri arttırma işlemlerine tabi tutulmasından sonra başlamaktadır. Sonra toplamda 49000 eğitim seti örneği, 128 örneğe sahip bölümler halinde mini-yığın eğitimine sokulur. Bütün örneklerin bir 'epok' sonunda bu şekilde kullanılmasından sonra diğer epokun başladığı bu eğitimde, toplamda 300 epok bulunmaktadır. Her 50 epok sonrasında -veri arttırma ön işlemine tabi tutulmamış olan-doğrulama setindeki başarıya bakılan bu eğitimde, bu oranlara göre seçilen en iyi sinir ağının 10000 örnek içeren test setinde elde ettiği genel başarı oranı hesaplanır. Bu işlem her ön değer atama yöntemi için on defa tekrarlanarak on adet eğitilmiş sinir ağı elde edilir. Bu 10 test sonucunun ortalama ve maksimum başarı oranı, o ön değer atama yönteminin sonucu olarak atanır. Tüm yöntemler arasından ortalama veya maksimum oranda en başarılı olan ise en iyi ön değer atama yöntemi olarak belirlenir, maksimum başarı oranlarının yakın olması dolayısıyla ortalama başarı oranları da kıyas için hesaplanmıştır.

Her iki veri seti için de farklı ön değer atama yöntemleri denenmiş olmakla beraber, İnsan Hareketi Tanımlama veri seti için Gauss, Tekdüze ve önerdiğimiz Laplasyen dağılımları kullanılmıştır. CIFAR-10'da ise bu dağılımlara ek olarak nöron girdi sayısına göre budanmış He tipi Gauss ve Tekdüze dağılım ile önerimiz budanmış Laplasyen dağılımları eklenmiştir. Bunlar arasından en yüksek başarıya sahip olanları bulunmuştur.

C. Yöntemlerin Karşılaştırılması

Tablo I'de görüldüğü gibi düşük öğrenme katsayısıyla Tekdüze ve Laplasyen yüksek başarı elde ederken, yüksek

katsayı ile Gauss dağılımı en yüksek başarıyı elde etmiştir.

TABLO I: İNSAN HAREKETİ TANIMLAMA VERİ SETİ İÇİN REZERVUAR KATMANI İÇEREN YAPAY SINIR AĞLARI

Kullanılan Yapay Sinir Ağlarının Ön Değer Dağılım Türü	Farklı Öğrenme Katsayılarına (η) Göre 10 Denemede Alınan Maksimum Başarı	
	$\eta = 1, 25 \cdot 10^{-4}$	$\eta = 1, 25 \cdot 10^{-3}$
Tekdüze (Uniform)	%95,55	%95,86
Laplasien (Laplacian)	%95,55	%95,83
Gauss (Gaussian)	%95,11	%96,2

Tablo II’de küçülen öğrenme katsayısına sahip AdaMax optimizasyonu ile öğrenim gören evrişimsel ağ, budanmış Laplasien dağılımıyla Gauss ve Tekdüze’den başarılı olmuştur.

TABLO II: CIFAR-10 İÇİN EVRİŞİMSSEL SINIR AĞLARI

Yapay Sinir Ağına Ön Değer Dağılım Türü	10 Denemede Aldığı Başarı Oranları	
	Ortalama Başarı	Maksimum Başarı
Tekdüze (Uniform)	%88,44	%89,27
Laplasien (Laplacian)	%88,59	%89,29
Gauss (Gaussian)	%88,88	%89,50
Budanmış Tekdüze	%88,80	%89,42
Budanmış Laplasien Tip 1 ^a	%88,75	%89,51
Budanmış Laplasien Tip 2 ^b	%88,89	%89,48
Budanmış Gauss	%88,68	%89,37

^aTip 1 Laplasien, Budanma Sabiti olarak 12 kullanılmaktadır.

^bTip 2 Laplasien, Budanma Sabiti olarak 13,5 kullanılmaktadır.

Çalışma sonunda elde edilen en başarılı sınıflandırıcılar-dan İnsan Hareketi Tanımlama veri setine ait olan rezervuar katmanlı yapay sinir ağı, önceki çalışmamızda belirttiğimiz gibi %96,2 gibi yüksek bir başarı oranına ulaşmıştır [10]. Aynı zamanda kullanılan Tekdüze ve Laplasien ön değer atama yöntemleri Gauss yönteminin %96,2 aldığı öğrenme katsayısında daha başarısız kalırken, öğrenme katsayısı düşürüldüğünde Gauss yönteminden başarılı, %95,55 gibi sonuçlar da verilmektedir. CIFAR-10 veri setine ait sinir ağına ise, diğerlerini az farkla geride bırakan budanmış Laplasien ön değer atama yöntemi için budanma sabiti seçimi, sonuçları değiştiren önemli bir etken olarak görülmektedir. Çalışma sonunda elde edilen oranlar yakın olsa da, önerimiz Laplasien temelli ön değer atamanın kullanılabilirliği açık bir şekilde görülmektedir.

IV. SONUÇLAR

Bu çalışmada yapay sinir ağları için çeşitli ön değer atama yöntemleri karşılaştırılmıştır. Yakın sonuçlar çıkmakla beraber ön-beslemeli sinir ağlarında Gauss temelli, evrişimsel sinir ağlarında ise geliştirdiğimiz Laplasien temelli ön değer atama yöntemleri daha başarılı olmuştur. Bunun sebebi olarak eğitimde AdaMax optimizasyonunun kullanımıyla öğrenme katsayısının düşmesi görülmekte, düşük katsayıyla Gauss ve Tekdüze dağılıma kıyasla Laplasien gibi, birbirine daha yakın ön değerler atayan yöntemler daha başarılı görünmektedir.

Gelecekte çalışmada kullanılan ön-beslemeli ve evrişimsel sinir ağı yapılarını geliştirerek, sıkça kullanılan Gauss ve Tekdüze temelli ön değer atama yöntemleri yerine Laplasien temelli ön değer atama yönteminin kullanımıyla, ayırt edilir şekilde daha başarılı sonuç alınabileceği düşünülmektedir. Budanma sabitinin sinir ağı yapıları için adaptif şekilde değiştirilmesi ve bunun sonuçlara yansımalarının görülmesi de gerekmektedir. Ayrıca tekrarlı yapay sinir ağı gibi farklı türlerdeki

sinir ağlarında da Laplasien temelli ön değer atama yönteminin etkisinin araştırılmasının faydalı olacağı düşünülmektedir.

BİLGİLENDİRME

Bu çalışmada kullanılan Titan X Pascal GPU’nun başışı için NVIDIA Corporation’a teşekkür ederiz. Çalışmamız sürecinde desteğini esirgemeyen İsmail Uyanık ve Hasan Hamzaçebi’ye teşekkür ederiz. Yazarlardan Bahadır ve Burak Çatalbaş Türkiye Bilimsel ve Teknolojik Araştırma Kurumuna (TÜBİTAK) verilen değerli finansal destek için teşekkür eder.

KAYNAKÇA

- [1] M. B. Ali, “Use of dropouts and sparsity for regularization of autoencoders in deep neural networks,” Master’s thesis, Bilkent University, 2015.
- [2] B. Çatalbaş, B. Yücesoy, G. Seçer, and M. Aslan, “A comparative study of classification methods for fall detection,” in *2014 22nd Signal Processing and Communications Applications Conference (SIU)*, pp. 1315–1318, April 2014.
- [3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in neural information processing systems*, pp. 1097–1105, 2012.
- [4] J. Redmon and A. Farhadi, “Yolo9000: better, faster, stronger,” *arXiv preprint*, vol. 1612, 2016.
- [5] B. Çatalbaş, “Recurrent neural network learning with an application to the control of legged locomotion,” Master’s thesis, Bilkent University, 2015.
- [6] A. Graves, A. R. Mohamed, and G. Hinton, “Speech recognition with deep recurrent neural networks,” in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 6645–6649, May 2013.
- [7] W. S. McCulloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” *The bulletin of mathematical biophysics*, vol. 5, no. 4, pp. 115–133, 1943.
- [8] A. Karpathy, “Lessons learned from manually classifying cifar-10,” *Published online at karpathy.github.io/2011/04/27/manually-classifying-cifar10*, 2011.
- [9] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: A simple way to prevent neural networks from overfitting,” *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [10] B. Çatalbaş, B. Çatalbaş, and Ö. Morgül, “Human activity recognition with different artificial neural network based classifiers,” in *2017 25th Signal Processing and Communications Applications Conference (SIU)*, pp. 1–4, May 2017.
- [11] D. Anguita, A. Ghio, L. Oneto, X. Parra Perez, and J. L. Reyes Ortiz, “A public domain dataset for human activity recognition using smartphones,” in *Proceedings of the 21th International European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, pp. 437–442, 2013.
- [12] A. Krizhevsky, “Learning multiple layers of features from tiny images,” 2009.
- [13] B. Romera-Paredes, M. S. Aung, and N. Bianchi-Berthouze, “A one-vs-one classifier ensemble with majority voting for activity recognition,” in *ESANN 2013 proceedings, 21st European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, pp. 443–448, 2013.
- [14] X. Gastaldi, “Shake-shake regularization,” *arXiv preprint arXiv:1705.07485*, 2017.
- [15] P. Kaur, “Convolutional neural networks (cnn) for cifar-10 dataset,” *parneetk.github.io/blog/cnn-cifar10/*, 2017. Accessed: 09 February 2018.
- [16] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.
- [17] K. He, X. Zhang, S. Ren, and J. Sun, “Delving deep into rectifiers: Surpassing human-level performance on imagenet classification,” in *Proceedings of the IEEE international conference on computer vision*, pp. 1026–1034, 2015.