

Harflerden Oluşan Genişletilmiş MNİST Veri Kümesinin Derin Öğrenme Tabanlı Tasarlanmış Sinir Ağı Modeli ile Sınıflandırılması

Mesut TOĞAÇAR*¹

¹Fırat Üniversitesi, Teknik Bilimler Meslek Yüksekokulu, Bilgisayar Teknolojileri Bölümü, Elazığ

Geliş tarihi: 01.06.2021

Kabul tarihi: 13.09.2021

Öz

MNİST veri kümesi, model öğrenmesi, görüntü işleme, sınıflandırma süreçlerinin gerçekleştirilmesinde standart bir ölçüt olarak kullanılmaktadır. MNİST veri kümesi içerisinde; el yazısı formatında hazırlanmış büyük harf, küçük harf ve rakam görüntülerinden oluşmaktadır. Genişletilmiş MNİST veri kümesi, MNİST veri kümesi ile aynı değerler dizisi formatında hazırlanmış daha kapsamlı ve sınıflandırma süreci daha zor bir veri seti türüdür. Günümüzde birçok alanda yapay zekâ tabanlı çalışmalar ilgi görmeye başlamıştır. Bu çalışmada, Genişletilmiş MNİST veri kümesinin eğitilmesi ve sınıflandırması amacıyla Python dilinde tasarlanmış yeni bir sinir ağı modeli önerilmektedir. Önerilen modelde, ön işlem adımı olarak veri büyütme yöntemi eğitim verileri için uygulanmıştır ve 26 harf kategorik olarak sınıflandırıldı. Sınıflandırma sürecinde genel doğruluk başarısı %94,73 olarak elde edildi. Önerdiğimiz model, el yazısı görüntülerinin sınıflandırılmasında başarılı bir analiz gerçekleştirdiği gözlemlendi.

Anahtar Kelimeler: Derin öğrenme, Genişletilmiş MNIST veri kümesi, Harf sınıflandırması, Sinir ağları

Classification of Extended MNIST Dataset Consisting of Letters with Deep Learning-Based Designed Neural Network Model

Abstract

MNIST dataset is used as a standard criterion for model learning, image processing and classification processes. In the MNIST dataset; it consists of uppercase, lowercase letters and numbers in handwriting format. The expanded MNIST data set is a more comprehensive type of data set prepared in the same paradigm format as the MNIST dataset, and the classification process is more difficult. Today, artificial intelligence-based studies in many areas have started to attract attention. In this study, a new neural network model designed in Python language is proposed in order to train and classify the extended MNIST dataset. In the proposed model, data enlargement method as a preprocess step was applied for training data and 26 letters were categorically classified. The overall accuracy success achieved in the classification process was %94.73. The proposed model we are observed to perform a successful analysis in classifying handwritten images.

Keywords: Deep learning, Extended MNIST dataset, Letters classification, Neural networks

*Sorumlu yazar (Corresponding author): Mesut TOĞAÇAR, mtogacar@firat.edu.tr

1. GİRİŞ

Görüntü işleme, makine öğrenmesi, derin öğrenme gibi yaklaşımların başarı ölçütlerinden biri de standartlaşmış veri kümelerini kullanmasıdır. Standartlaşmış veri kümeleri farklı öğrenme yaklaşımları veya yöntemlerinden elde edilen analizleri karşılaştırmak için adil bir yol sağlar. Araştırmacılar tarafından geliştirilmiş yaklaşım veya yöntemler bu veri kümelerini kullanarak yayınladıkları makalelerde daha hızlı bir şekilde diğer araştırmacılar tarafından iç göri kazanmasına imkân sağlar [1,2]. Standartlaşmış veri kümeleri arasında bazıları şunlardır; MNİST, Coco, İmagenet, Cifar, Flicker, Cityscapes [3–5] vb. İlk standartlaşmış veri kümeleri arasında MNİST, 1998 yılında Lecun ve arkadaşları tarafından oluşturulmuştur. MNİST veri kümesi, el yazısı karakterlerden oluşan 28x8 piksel boyutuna sığdırılan gri seviyeli görüntülerden oluşmaktadır [6]. Genişletilmiş MNİST, 2017 yılında Gregory Cohen ve arkadaşları tarafından oluşturulmuştur [7]. Genişletilmiş MNİST veri kümesi, MNİST veri kümesinden daha geniş kapsamlı el yazısı görüntülerini içeren görüntülerden oluşmaktadır [8].

El yazılı harf karakterlerinden oluşan görüntüleri tanımada bilgisayar görmesi için klasikleşmiş bir problemdir. Yapay zekâ tabanlı birçok evrimsel sinir ağı (ESA) modeli ile harf karakteri tanıma ve sınıflandırma işleminde başarılı sonuçlar elde edilmiştir. Bu çalışmada önerilen yaklaşım ile genişletilmiş MNİST veri kümesinde tasarlanan ESA modeli sınıflandırılması hedeflenmektedir. Önerilen yaklaşımın başarımını artırmak için ise eğitim verisindeki her bir görüntü büyütme tekniği ile işlenecektir. Harf karakterlerinin sınıflandırılmasında birçok çalışma gerçekleştirilmiştir. Genişletilmiş MNİST veri kümesini kullanarak analizler gerçekleştirmiş bazı çalışmalar incelenirse; A. Amelia [9], çalışmada 26 harf görüntüsünden oluşan veri kümesini kendi tasarladığı ESA modeli ile sınıflandırmıştır. Sınıflandırma sürecinde Softmax fonksiyonunu kullanmış ve iki formatta bu süreci gerçekleştirmiştir. Birincisi, veri kümesini

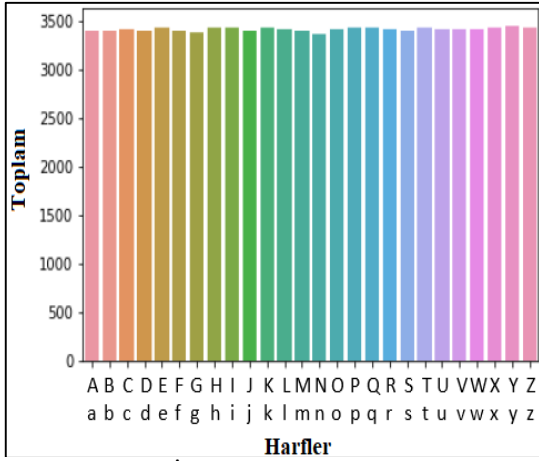
doğrudan ESA modeli ile gerçekleştirerek %92 oranında genel doğruluk başarısı elde etmiştir. İkincisi ESA modelinden çıkartılan paylaşılmış ağırlıkları kullanarak %74,29 oranında bir genel doğruluk başarısı elde etmiştir. Dawid Grzelak ve arkadaşları [10], 26 karakterli harflerden oluşan veri kümesinin sınıflandırma sürecinde, tasarladıkları ESA modelini kullanmışlar. ESA modelinde ayrıca görüntü kaydırma ve ölçekleme işlemlerini gerçekleştiren Rampa fonksiyonunu (Ramp) kullanmışlar. Onların ESA modelinin son katmanında kullandıkları Softmax fonksiyonu ile elde ettikleri genel doğruluk oranı %92,40 olmuştur. V. Jayasundara ve arkadaşları [11], 26 karakterli veri kümesini ön işlem adımlarından geçirerek iyileştirmişlerdir. Onlar, kapsül ağlarını ve kod çözücü ağları birlikte kullanarak melez bir model sunmuşlardır. Onların çalışmasında elde ettikleri genel doğruluk başarısı %92,79 olmuştur. Ruthvik Vaila ve arkadaşları [12], veri kümesini 26 harf türüne göre başarılı bir şekilde sınıflandırmışlar. Onlar, çalışmada melez bir model önermişler ve sivri sinir ağlar ile dereceli alçalma yöntemini birlikte kullanmışlar. Onların çalışmasında elde ettikleri genel doğruluk başarısı %85,35 olmuştur.

Bu makale kısaca şöyle düzenlenmiştir; veri kümesi hakkında bilgiler Bölüm 2’de verilmiştir. Ön işlem adımı (veri büyütme) ve ESA modeli hakkında bilgiler Bölüm 3’te verilmiştir. Çalışmanın deneysel analizleri ile ilgili bilgiler Bölüm 4’te verilmiştir. Tartışma ve Sonuç hakkında bilgiler sırasıyla son iki bölümde yer almıştır.

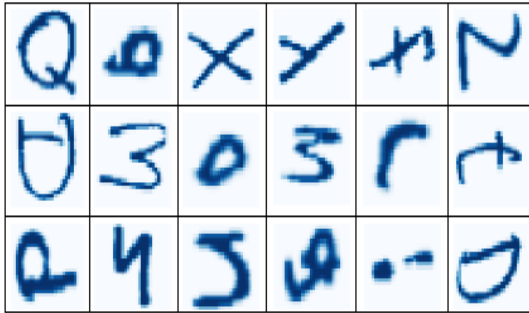
2. VERİ KÜMESİ

Genişletilmiş MNİST (GMNİST) veri kümesi, MNİST veri kümesine göre daha fazla görüntülerden oluşmuş ve zorluk derecesi artırılmış standartlaşmış bir veri kümesidir. GMNİST’teki her bir görüntü çözünürlüğünün düşük olması, rahat depolanabilir olması, erişim kolaylığı ve kullanım kolaylığı bu veri kümesini standartlaştırılmasını sağlamıştır. Bu çalışmada kullanılan GMNİST veri kümesi, küçük ve büyük

el yazısından oluşan 26 harf görüntülerinden oluşmaktadır. Her bir görüntü 28x28 piksel çözünürlüğündedir. Ek olarak, her bir görüntü 8 bit gri seviyeli çözünürlüğe sahiptir ve toplam 103,600 karakterden oluşmaktadır. Veri kümesinin; 88,800 görüntüsü eğitim verileri olarak ayrılmıştır ve 14,800 görüntüsü ise test verileri için ayrılmıştır [7]. Eğitim verilerinin harf tabanlı istatistik bilgisi Şekil 1’de gösterildi. GMNİST veri kümesinin örnek görüntülerinden oluşan alt kümesi ise Şekil 2’de gösterildi.



Şekil 1. GMNİST veri kümesinde, eğitim verileri için ayrılmış her bir harf miktarı



Şekil 2. GMNİST veri kümesinin alt kümesi örnek görünümü

Veri kümesi, Şekil 2 incelendiğinde dengeli bir dağılım gösterdi. İstatistiksel olarak veri kümesinin %30’u test verisi ve %70’i eğitim verisidir [13]. Bu çalışmanın analizinde de mevcut durum korundu.

3. ÖNERİLEN YAKLAŞIM

3.1. Veri Büyütme ve Optimizasyon

Birçok derin öğrenme modellerinde ve makine öğrenme yöntemlerinde veri büyütme tekniği kullanılmaktadır. Veri büyütme tekniği ile ESA modellerinin performanslarına katkı sağlanmakta ve modellerin aşırı uyumunu önlemektedir. Bu teknik sayesinde, her bir görüntünün orijinal halinden biraz değiştirilmiş formatta kopyalarının oluşturulmaktadır veya orijinal görüntüler yerine kullanılmaları amaçlanmaktadır. Böylece her bir görüntünün kalitesi artırılmış olur [14]. Veri büyütme yönteminde kullanılan parametreler vardır. Bu çalışmada veri büyütme tekniği python dilinde tasarlanmıştır ve keras kütüphanesinin “Image Data Generator” fonksiyonu kullanılarak yalnızca eğitim verilerine veri büyütme tekniği uygulandı [15]. Çalışmanın deneyinde tercih edilen önemli veri büyütme parametreleri ve önerilen yaklaşımın diğer parametre değerleri Çizelge 1’de verildi. Önerilen yaklaşımda ESA modelinin mini seri boyutu 64 tercih edildi. Mini seri boyutu sayesinde veri kümeleri alt veri kümelerine ayrılarak ESA modelleri tarafından daha rahat eğitilmesine katkıda bulunmaktadır [16]. Deneyin analizinde tercih edilen devir sayısı 50 tercih edildi ve tercih edilen optimizasyon algoritması Adam yöntemi idi.

Adam yöntemi, her partideki ağırlık parametrelerini ve öğrenme katsayılarını güncelleyen etkili bir optimizasyon yöntemidir. Modelin eğitimi sırasında Adam yöntemi, yön türevi (gradyan) değerlerini tahmin ederek ağırlık parametrelerinin öğrenme oranlarının ayarlanmasına yardımcı olur. Adam, yön türevin inişlerinde (g_t) hesaplanan üstel hareketli ortalamaları kullanır. Geçmiş yön türevi (m_t) ve geçmiş kare gradyan (V_t) ortalamaları sırasıyla Eşitlik 1 ve Eşitlik 2’ye göre hesaplanır. Bu denklemlerde; β değişkeni hiper parametre değerlerini hesaplamak için kullanılır ve bu değişkenin değeri $0,9 < \beta < 0,999$ arasındadır. Zaman sürecini ise, t değişkeni temsil eder.

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (1)$$

$$V_t = \beta_2 V_{t-1} + (1 - \beta_2) g_t^2 \quad (2)$$

Çizelge 1. Bu çalışmada kullanılan veri büyüme tekniği ile optimizasyon yönteminin parametreleri

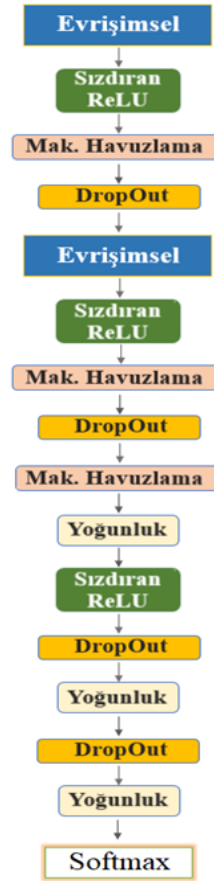
Önerilen veri büyüme tekniği		Optimizasyon ve parametreler	Veri büyüme parametreleri ve değerleri
Yazılım	Python	Optimizasyon= Adam Öğrenme oranı= 0,001 Gecikme=0,0 $\beta_1=0,9$ $\beta=0,999$	Yakınlaştırma aralığı = 0,1 Yükseklik kaydırma aralığı = 0,1 Genişlik kaydırma aralığı = 0,1 Döndürme sınırı = 10°
Kütüphane	Keras		
Mini seri boyutu	64		
Devir sayısı	50		
Kayıp türü	Kategorik çapraz entropi		

3.2. Önerilen Model

Önerilen yaklaşımda kullanılan ESA modeli python yazılımında tasarlanmış ve GMNİST veri kümesine uyarlanmış özgün bir modeldir. Model genel yapısıyla; evrişimsel katmanlar, havuzlama katmanları ve tam bağlantılı katmanlardan oluşmaktadır [17]. Evrişimsel katman, girdi görüntüsü üzerinde filtre çerçevesini dolaştırarak hem görüntü boyutunu azaltır hem de bir sonraki katmana görüntü özelliklerinin yer aldığı aktivasyon haritalarını aktarır. Havuzlama katmanı, girdi boyutunun azaltılmasını sağlar ve bu durum derinlik boyutunu etkilemez. Tam bağlantılı katman / yoğunluk katman, modelin sınıflandırma sürecinde kullanılan fonksiyonuna tüm girdi değerlerini katarak değerler üretilmesini sağlar. Genelde tam bağlantılı katmanlar ESA modellerinin son katmanlarında tercih edilmektedir [18]. Bu çalışmada tam bağlantılı katmanın işlevini yerine getirecek yoğunluk katmanı (Dense layer) kullanıldı.

Önerilen modelde kullanılan diğer fonksiyonlar ise; Softmax, Sızdıran ReLU, DropOut. Softmax yöntemi, sınıflandırma işlemi gerçekleştirmek için genelde ESA modelleri tarafından en son katmanda tercih edilen bir yöntemdir [19]. Softmax, bir önceki yoğunluk/ tam bağlantılı katmandan gelen girdi değerlerini işleyerek olasılık değerleri üretilmesini sağlar. Böylece, modelin girdi görüntüsü olasılık değeri en yüksek olan sınıfa aktarılır [20]. Sızdıran ReLU fonksiyonu, pozitif girdi değerlerini işleyerek pozitif yeni bir çıktı değeri üretilmesini sağlar; aynı zamanda negatif girdi değerlerini işleyerek, negatif yeni bir çıktı değeri üretilmesine engel olur. Negatif bir çıktı değeri üretmeden sıfır değerine

yakın pozitif bir değer üretilmesini sağlar [21]. DropOut fonksiyonu, aşırı uyumun önüne geçebilmek için girdi değerlerini seyrelterek çıkışa aktarır [22]. Bu çalışmada kullanılan ESA modelinin genel tasarımı Şekil 3'te gösterildi. Önerilen modelin tasarım detayları ise Çizelge 2'de verildi.



Şekil 3. Önerilen modelin genel tasarımı

Çizelge 2. Önerilen modelin genel tasarımı ve parametre değerleri

Katman /fonksiyon	Filtre	Parametre/Değer	Adım boyutu	Çıkış boyutu
Evrişimsel	32	-	5 × 5	28 × 28
Sızdıran ReLU	-	Alfa= 0,02	-	-
Maksimum havuzlama	-	-	2 × 2	-
DropOut	-	0,2	-	-
Evrişimsel	196	-	5 × 5	-
Sızdıran ReLU	-	Alfa= 0,02	-	-
Maksimum havuzlama	-	-	2 × 2	-
DropOut	-	0,2	-	-
Maksimum havuzlama	-	-	2 × 2	-
Yoğunluk	1024	-	-	-
Sızdıran ReLU	-	Alfa= 0,02	-	-
DropOut	-	0,25	-	-
Yoğunluk	128	ReLU	-	-
DropOut	-	0,25	-	-
Yoğunluk	-	-	-	26
Softmax				Sınıf sayısı: 26

4. BULGULAR

Bu çalışmada kullanılan veri büyütme yöntemi ve önerilen model Python 3.6 yazılım dilinde tasarlanmıştır. Yazılımın derlenmesinde Jupyter Notebook ara yüzü kullanıldı. Deneyin analizi için kullanılan donanım özellikleri şu şekildedir; işlemci Intel © Xeon © Gold 6132 CPU @ 2,6 GHz, grafik kartı NVIDIA Quadro P6000 24 GB, bellek kartı 64 GB ve 64 bitlik Windows 10 işletim sistemi kullanıldı. Bu çalışmanın geçerliliğinin ölçülmesi sürecinde karmaşıklık matrisi tercih edildi. Karmaşıklık matrisinin metrikleri; duyarlılık, özgüllük, hassasiyet, f-skor ve doğruluktur. Karmaşıklık matrisindeki metrik değerlerinin hesaplanmasında aşağıda belirtilen denklemler (Eşitlik 3-7) kullanılmaktadır. Bu denklemlerde kullanılan değişkenler; doğru pozitif (DP), doğru negatif (DN), yanlış pozitif (YP), yanlış negatif (YN) anlamına gelmektedir [23,24].

$$\text{Duyarlılık (Duy)} = \frac{(DP)}{(DP+YN)} \quad (3)$$

$$\text{Özgüllük (Özg)} = \frac{(DN)}{(DN+YP)} \quad (4)$$

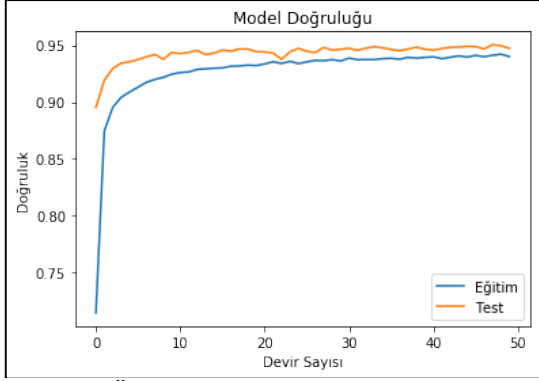
$$\text{Hassasiyet (Has)} = \frac{(DN)}{(DN+YP)} \quad (5)$$

$$\text{F-skor (f-skr)} = \frac{(2 \times DP)}{(2 \times DP + YP + YN)} \quad (6)$$

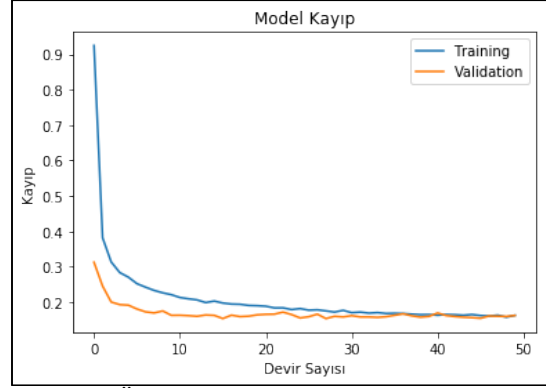
$$\text{Doğruluk (Dğr)} = \frac{(DP+DN)}{(DP+YN)+(YP+DN)} \quad (7)$$

Çalışmanın deneysel analizinde veri kümesi görüntüleri “csv” uzantılı dosya formatına dönüştürüldü. Ardından, eğitim verileri ve test verileri girdi olarak önerilen yaklaşıma verildi. Eğitimde tercih edilen devir sayısı 50’ydi. Analizler sonucunda, elde edilen genel doğruluk başarıları %94,73 ve genel kayıp oranı ise %16,23’tü. Modelin her bir devir sayısı ortalama 57 saniye sürdü. Tüm devir sayısı dikkate alındığında ortalama geçen süre 2850 saniyeydi. Önerilen yaklaşımda elde edilen genel doğruluk grafiği Şekil 4’te gösterildi ve genel kayıp grafiği Şekil 5’te gösterildi. Ek olarak, analiz sonucunda elde edilen karmaşıklık matris tablosu Şekil 6’ta gösterildi ve bu karmaşıklık matrisi tablosundan elde edilen analiz sonuçları Çizelge 3’te verildi. Analiz sonuçları incelendiğinde en iyi sınıf başarıları %99,90 oranında “S, s - Z, z” harflerinde elde edildi ve en kötü sınıf başarıları ise %97,98 oranında “I, i” harfinden elde edildi.

Harflerden Oluşan Genişletilmiş MNİST Veri Kümesinin Derin Öğrenme Tabanlı Tasarlanmış Sinir Ağı Modeli ile Sınıflandırılması



Şekil 4. Önerilen yaklaşımın eğitim ve test doğruluk başarıları grafiği



Şekil 5. Önerilen yaklaşımın eğitim ve test doğruluk kayıp grafiği

Çizelge 3. Deneysel analizden elde edilen sonuçlar (%)

Veri kümesi	Sınıf	Duy.	Özg.	Has.	f-skr	Dğr.	Genel dğr.
GMNİST	A, a	95,62	99,88	97,29	96,45	99,70	94,73
	B, b	97,74	99,89	97,54	97,64	99,81	
	C, c	96,81	99,92	98,08	97,44	99,79	
	D, d	98,19	99,74	94,42	96,27	99,68	
	E, e	98,15	99,87	97,11	97,63	99,80	
	F, f	97,71	99,91	97,91	97,81	99,82	
	G, g	82,93	99,73	92,87	87,62	99,04	
	H, h	97,01	99,88	97,38	97,20	99,77	
	I, i	64,46	99,38	81,27	71,90	97,98	
	J, j	95,17	99,84	96,21	95,69	99,65	
	K, k	97,78	99,95	98,77	98,27	99,86	
	L, l	86,67	98,47	70,12	77,52	98,00	
	M, m	98,78	99,90	97,59	98,18	99,85	
	N, n	96,04	99,87	97,00	96,52	99,72	
	O, o	95,84	99,95	98,85	97,32	99,79	
	P, p	98,30	99,94	98,68	98,49	99,87	
	Q, q	93,84	99,28	84,85	89,12	99,06	
	R, r	95,20	99,90	97,59	96,38	99,71	
	S, s	98,30	99,96	99,19	98,74	99,90	
	T, t	97,41	99,88	97,32	97,36	99,78	
	U, u	92,48	99,82	95,75	94,08	99,53	
V, v	96,41	99,59	91,02	93,64	99,46		
W, w	99,30	99,87	97,09	98,18	99,85		
X, x	98,20	99,89	97,65	97,92	99,82		
Y, y	96,62	99,86	96,71	96,67	99,73		
Z, z	98,20	99,97	99,42	98,80	99,90		

Karmaşıklık Matrisi

0	1006	1	1	2	0	0	3	1	0	0	1	0	1	1	2	0	21	2	0	1	3	0	4	1	1	0
1	1	995	0	2	1	0	1	4	0	1	0	5	0	0	1	1	0	2	0	1	0	0	3	0	0	0
2	0	0	974	1	14	0	2	0	0	0	0	7	0	1	1	0	2	1	0	2	1	0	0	0	0	0
3	1	3	0	1033	1	0	0	0	1	2	0	0	0	0	3	7	0	0	0	0	0	0	0	0	0	1
4	1	1	9	0	1011	1	1	0	0	0	0	1	0	0	1	0	1	1	0	0	0	0	0	0	0	2
5	0	0	0	2	2	985	1	1	0	2	0	1	0	0	0	2	1	0	2	8	0	0	0	1	0	0
6	8	11	2	3	3	0	860	3	0	2	0	0	0	1	0	0	136	1	2	1	0	0	0	0	4	0
7	2	0	0	0	0	0	1007	0	0	2	5	1	13	0	0	0	1	0	1	2	1	1	2	0	0	0
8	0	0	0	1	1	0	0	1	664	14	0	346	0	0	0	0	0	0	0	1	0	0	0	1	1	0
9	0	2	0	3	0	1	3	1	19	966	0	3	0	0	0	0	1	0	3	3	3	5	1	0	1	0
10	0	2	0	0	0	1	0	4	0	0	969	2	0	0	0	0	1	0	1	2	0	2	6	1	0	0
11	0	0	3	1	0	0	1	1	126	3	0	885	0	0	0	0	1	0	0	0	0	0	0	0	0	0
12	1	0	0	0	0	0	0	2	0	0	0	0	973	2	0	0	0	0	0	0	0	3	1	3	0	0
13	1	0	0	0	0	0	0	5	0	0	0	0	16	972	0	1	0	6	0	0	0	0	8	1	1	1
14	1	1	0	32	0	0	0	0	0	0	1	0	0	0	1	946	0	2	0	0	0	3	0	0	0	0
15	0	1	0	8	0	2	0	0	0	1	1	1	0	1	0	1047	2	0	0	0	1	0	0	0	0	0
16	7	0	0	1	0	0	47	0	0	0	0	0	0	0	2	2	975	1	1	0	1	1	0	0	1	0
17	0	0	1	2	5	1	0	1	1	1	6	2	0	6	0	0	1	972	0	2	0	14	0	1	5	0
18	2	1	1	1	1	0	6	0	1	2	0	0	0	0	0	0	1	0	987	0	0	0	1	0	0	0
19	0	0	0	0	2	10	0	2	2	3	0	0	0	0	0	0	2	0	1017	0	0	0	2	3	1	0
20	1	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	1	0	0	0	947	64	6	0	2	0
21	0	1	0	1	0	1	0	0	0	3	0	2	0	0	0	0	4	0	0	18	994	0	0	7	0	0
22	0	0	0	0	0	0	0	0	0	0	2	2	1	0	0	0	0	0	0	2	0	1002	0	0	0	0
23	0	0	0	0	0	3	0	0	0	0	2	0	1	2	0	0	0	0	2	0	2	1	1039	5	1	0
24	0	0	0	1	0	0	1	1	0	0	0	0	2	0	0	1	1	2	0	4	3	9	0	9	973	0
25	2	1	2	0	0	1	0	0	3	2	0	0	1	0	0	0	3	0	0	1	0	1	0	1	1	1037
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	

Şekil 6. Bu çalışmanın analizinde elde edilmiş karmaşıklık matrisi

5. TARTIŞMA

GMNİST veri kümesini oluşturan harfleri sınıflandırmak oldukça zor bir süreçtir. Çünkü MNİST veri kümesine göre veriler artırılmış ve anlaşılması zor el yazısı imgeleri veri kümesine eklenmiştir. Bu tür standart veri kümeleri

tasarlanmış ESA modellerinin başarılarını ölçmek için ilk etapta kullanılmaktadır. Böylece aynı veri kümesini kullanan diğer ESA modelleri ile performans karşılaştırılması yapılabilecektir. Önerilen yaklaşımdaki başarı ön işlem adımı olarak gerçekleştirdiğimiz ve eğitim verilerine uyguladığımız veri büyütme yöntemine bağlıdır.

Diğer yönüyle önerilen modelimde kullandığım katman ve bu katmanlara bağlı parametre değerleri de başarılı bir sonuç almamı sağladı. Önerilen yaklaşımın eksik olarak görebileceğim yönü ise ön işlem adımlarının çeşitliliğini artırabilirdim. Örneğin; veri büyütme yöntemi ile

beraber, görüntü imgelerinde bölgesel odaklanma ve piksel artırma yöntemleri de denenebilirdi. GMNİST (26 sınıflı) veri kümesini kullanarak analizler gerçekleştirildi ve literatürde yer alan çalışmalar Çizelge 4’te verildi.

Çizelge 4. Aynı veri kümesini kullanan çalışmaların sonuçları

Makale	Yıl	Model/Yöntem	Genel dğr. (%)
A. Amelia [9]	2018	ESA	92
Dawid Grzelak ve arkadaşları [10]	2019	ESA + Rampa	92,40
V. Jayasundara ve arkadaşları [11]	2019	Kapsül ağ + Kod çözücü	92,79
Ruthvik Vaila ve arkadaşları [12]	2020	Sivri ağ + Dereceli alçalma	85,35
Önerilen yaklaşım	2021	ESA + Veri büyütme	94,73

A. Amelia [9] çalışmasında, yalnızca ESA modeli kullandı. O, ESA modeli ile birlikte bir ön işlem adımı veri kümesine uygulaması önerdiği yaklaşımın başarımını artırabilirdi. Dawid Grzelak ve arkadaşları [10], ESA modeli birlikte bir ön işlem adımı her bir evrimsel katmandan sonra (Rampa) uygulamışlardır. Bu yöntemin model içerisinde uygulanmaktansa, model eğitiminden önce görüntü verisine uygulanması daha verimli olacaktır. V. Jayasundara ve arkadaşları [11], farklı bir ESA mimarisi olan Kapsül ağlar ile birlikte otomatik kod çözücü yöntemini birlikte kullanmışlardır. Onlar çalışmasında, ön işlem adımı olarak veri büyütme yöntemini de kullanmışlardır. Ruthvik Vaila ve arkadaşları [12], ESA modellerinden farklı bir mimari yapıda olan Sivri ağları veri kümesinin eğitiminde kullanmışlar. Fakat onların kullandığı sivri ağlar, sınıflandırma performansında çok fazla katkı sunmamıştır. Onlar ESA modeli üzerinde analizlerini gerçekleştirdiler, sınıflandırma performansları %85,35’in üzerine çıkarabilirlerdi. Sonuç olarak önerilen yaklaşım, Çizelge 4’te belirtilen diğer çalışmalara göre daha iyi sonuç vermiştir.

6. SONUÇ

Standartlaşmış veri kümelerinde ESA modellerinin başarı ölçütlerinin değerlendirilmesi, modellerin geçerliliğini ve güvenilirliğini artırmaktadır. Bu çalışmada standartlaşmış veri kümesi olan GMNİST kullanılmıştır ve 26 sınıftan

oluşan harf görüntüleri sınıflandırılmıştır. Sınıf kümelerinde küçük ve büyük el yazısı harflerin bulunması sınıflandırma sürecini daha da zor hale getirmiştir. Ancak, önerilen yaklaşım ile elde edilen genel doğruluk oranı %94,73’tü. Bu çalışma ile el yazısı karakterlerinin sınıflandırılmasında önerilen yaklaşımın başarılı bir şekilde sonuç aldığı gözlemlenmiştir.

Gelecek çalışmalarda rakamlar, özel karakterler, harflerden oluşan standartlaşmış veri kümeleri üzerinde analizler yapılması düşünülmektedir. Ayrıca, önerilen yaklaşımın performans artışını sağlamak için çeşitli ön işlem adımları ile yeniden tasarlanması planlanmaktadır.

7. KAYNAKLAR

1. Lundervold, A.S., Lundervold, A., 2019. An Overview of Deep Learning in Medical Imaging Focusing on MRI. *Z Med Phys.* 29(2), 102–127. <https://doi.org/10.1016/j.zemedi.2018.11.002>.
2. AlQuraishi, M., 2019. ProteinNet: a Standardized Data Set for Machine Learning of Protein Structure. *BMC Bioinformatics*, 20, 311. <https://doi.org/10.1186/s12859-019-2932-0>.
3. Zhang, J.M., Harman, M., Ma, L., Liu, Y., 2020. Machine Learning Testing: Survey, Landscapes and Horizons. *IEEE Trans Softw Eng.* 99, 1–1. <https://doi.org/10.1109/tse.2019.2962027>.
4. Studer, L., Alberti, M., Pondenkandath, V.,

- Goktepe, P., Kolonko, T., Fischer, A., Liwicki, M., Ingold, R., 2019. A Comprehensive Study of Imagenet Pre-training for Historical Document Image Analysis. Proc Int Conf Doc Anal Recognition, ICDAR. 720–725. <https://doi.org/10.1109/ICDAR.2019.00120>.
5. Wei, Z., Wang, F., 2019. Adaptive Cascade Single-shot Detector on Wireless Sensor Networks. EURASIP J Wirel Commun Netw. 150. <https://doi.org/10.1186/s13638-019-1440-2>.
 6. Lecun, Y., MNIST handwritten digit database, Corinna Cortes and Chris Burges. <http://yann.lecun.com/exdb/mnist/>. Erişim tarihi: 7 Haziran 2021.
 7. Cohen, G., Afshar, S., Tapson, J., Van Schaik, A., 2017. EMNIST: Extending MNIST to Handwritten Letters. Proc Int Jt Conf Neural Networks 2017-May:2921–2926. <https://doi.org/10.1109/IJCNN.2017.7966217>
 8. Jiang W., 2020. MNIST-MIX: A Multi-Language Handwritten Digit Recognition Dataset. IOP SciNotes 1:025002. <https://doi.org/10.1088/2633-1357/abad0e>
 9. Amelia, A., 2018. Convolution Neural Network to Solve Letter Recognition Problem.
 10. Grzelak, D., Podlaski, K., Wiatrowski, G., 2019. Analyze the Effectiveness of an Algorithm for Identifying Polish Characters in Handwriting Based on Neural Machine Learning Technologies. J King Saud Univ-Comput Inf. Sci. <https://doi.org/10.1016/j.jksuci.2019.08.001>.
 11. Jayasundara V., Jayasekara S., Jayasekara, H., Jayasekara, H., Rajasegaran, J., Seneviratne, S., Rodrigo, R., 2019. TextCaps: Handwritten Character Recognition with Very Small Datasets. In: 2019 IEEE Winter Conference on Applications of Computer Vision (WACV). 254–262.
 12. Vaila, R., Chiasson, J., Saxena, V., 2020. A Deep Unsupervised Feature Learning Spiking Neural Network with Binarized Classification Layers for EMNIST Classification using Spyke Flow. 2002.11843.
 13. Crawford C., 2017. Extended MNIST Letter Dataset. In: Kaggle. <https://www.kaggle.com/crawford/emnist>. Erişim tarihi: 7 Haziran 2021.
 14. Shorten, C., Khoshgoftaar, T.M., 2019. A Survey on Image Data Augmentation for Deep Learning. J Big Data. 6, 60. <https://doi.org/10.1186/s40537-019-0197-0>.
 15. Image Data Augmentation. In: Keras Blog. <https://keras.io/api/preprocessing/image/>. Erişim tarihi: 6 Haziran 2021
 16. Dokuz, Y., Tufekci, Z., 2021. Mini-batch Sample Selection Strategies for Deep Learning Based Speech Recognition. Appl Acoust 171:107573. <https://doi.org/10.1016/j.apacoust.2020.107573>.
 17. Gadekallu, T.R., Khare, N., Bhattacharya, S., Singh, S., Maddikunta, P.K.R., Srivastava, G., 2020. Deep Neural Networks to Predict Diabetic Retinopathy. J Ambient Intell Humaniz Comput. <https://doi.org/10.1007/s12652-020-01963-7>.
 18. Suárez-Paniagua, V., Segura-Bedmar, I., 2018. Evaluation of Pooling Operations in Convolutional Architectures for Drug-drug Interaction Extraction. BMC Bioinformatics 19(S8), 209. <https://doi.org/10.1186/s12859-018-2195-1>.
 19. Lu, J., Ye, Y., Xu, X., Li, Q., 2019. Application Research of Convolution Neural Network in Image Classification of Icing Monitoring in Power Grid. EURASIP J Image Video Process. 49, 1-11. <https://doi.org/10.1186/s13640-019-0439-2>.
 20. Luo, Y., Wong, Y., Kankanhalli, M., Zhao, Q., 2020. Softmax: Improving Intra-class Compactness and Inter-class Separability of Features. IEEE Trans Neural Networks Learn Syst. 31, 685–699. <https://doi.org/10.1109/tnnls.2019.2909737>
 21. Zhang, X., Zou, Y., Shi, W., 2017. Dilated Convolution Neural Network with LeakyReLU for Environmental Sound Classification. In: 2017 22nd International Conference on Digital Signal Processing (DSP). 1–5.
 22. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. J Mach Learn Res 15(1), 1929-1958
 23. Toğaçar, M., Ergen, B., Sertkaya, M.E., 2019. Zatiürre Hastalığının Derin Öğrenme Modeli ile

- Tespiti. Fırat Üniversitesi Mühendislik Bilim Dergisi, 31(1), 223–230.
24. Sertkaya, M.E., Ergen, B., Togacar, M., 2019. Diagnosis of Eye Retinal Diseases Based on Convolutional Neural Networks Using Optical Coherence Images. In: 2019 23rd International Conference Electronics. 1–5.