Cilt 9, Sayı 2, Sayfa 105-111, 2020 Volume 9, Issue 2, Page 105-111, 2020

Araştırma Makalesi https://doi.org/10.46810/tdfd.755233 Research Article

Türk Doğa ve Fen Dergisi
Turkish Journal of Nature and Science
www.dergipark.gov.tr/tdfd

Örtüşme ve Deforme Durumlarında Kapsül Ağ ile Evrişimsel Sinir Ağ Sınıflama Performanslarının Karşılaştırılması

Kazım FIRILDAK[1*], Muhammed Fatih TALU[2]

[1] Fırat Üniversitesi, Kovancılar Meslek Yüksek Okulu, Bilgisayar Teknolojileri Bölümü, Elazığ, Türkiye
[2] İnönü Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Malatya, Türkiye
Kazım FIRILDAK ORCID No: 0000-0002-1958-3627
Muhammed Fatih TALU ORCID No: 0000-0003-1166-8404

**Sorumlu yazar: kfirildak@firat.edu.tr*

(Alınış: 19.06.2020, Kabul: 02.12.2020, Online Yayınlanma: 30.12.2020)

Anahtar Kelimeler
ESA,
KA,
ESA ile KA kıyaslanması

Öz: Evrişimsel sinir ağı (ESA) ve Kapsül Ağ (KA) önemli derin öğrenme mimarileridir. Bu makalede, ESA ve KA mimarilerinin MNIST ve Fashion MNIST veri kümelerindeki örtüşme ve deformasyon durumlarında sınıflama doğrulukları incelenmiştir. Bu veri kümelerinin her bir test verisi, rastgele alınan 7x7 lik blokların iki, üç ve dört tanesinin kendi aralarından yer değişmesiyle deforme edilmiştir. Bunun yanında veri kümelerine ait farklı sınıftaki veriler birbirilerini kısmi olarak örtecek şekilde test veri kümeleri oluşturulmuştur. Yapılan deneysel sonuçlarda deformasyon durumunda ESA'nın sınıflama doğruluğunun MNIST veri kümesi için %5 ila %13,5, Fashion MNIST için %7,1 ila %25,4 oranında azaldığı, KA da ise MNIST veri kümesi için %31 ila %43, Fashion MNIST için %34,7 ila %53,2 oranında azaldığı gözlemlenmiştir. Bu durum ESA mimarisinin, deforme edilmiş test kümelerinde KA' ya göre yüksek doğrulukta sınıflama yaptığı gösterilmiştir. Bunun yanında KA' nın deforme edilen sahte verilere karşı duyarlılığı ve sınıflama başarımına tepkisi ESA'ya göre daha doğru olduğu görülmüştür. Deforme durumlarının önemsenmediği uygulamalarda ESA'nın, deforme durumuna hassasiyet gösterecek güvenlik uygulamalarında KA kullanılması sonucuna varılmıştır. Örtüşme durumlarında ise ESA'nın sınıflama doğruluğu MNIST için %63, Fashion MNIST için % 58, KA ise doğruluk oranı MNIST için %88, Fashion MNIST için %81 olarak elde edilmiştir. Bu durum örtüşme durumlarında KA' nın ESA'ya göre daha yüksek sınıflama doğruluğuna ulaştığı gözlemlenmiştir.

Comparison of Capsule Network and Convolutional Neural Network Classification Performances in Overlap and Deformation Conditions

Keywords
CNN,
CapNet,
Comparison of CapNet and CNN

Abstract: Convolution Neural Network (CNN) and Capsule Network (CapsNET) are important deep learning architectures. In this paper, the classification accuracy of CNN and CapsNET architectures in the overlap and deformation states of MNIST and Fashion MNIST data sets are examined. Each test data of these data sets are deformed by the interchange of two, three, and four of the 7x7 blocks taken randomly. In addition, test data sets are created to overlap each other with different classes of data clusters partially. In experimental results, in case of deformation, the classification accuracy of CNN for the MNIST data set decreased to between 5% and 13.5% and for the Fashion MNIST to between 7.1 % and 25.4%, on the other hand, it was observed that the classification accuracy of CapsNET for the MNIST data set decreased to between 31% and 43% and for the Fashion MNIST to between 34.7% and 53.2%. This situation shows that CNN architecture makes high accuracy classification in deformed test sets compared to CapsNET. In addition, it was seen that the sensitivity of CapsNET to deformed fake data and its response to classification performance is more accurate than CNN. It is concluded that CNN should be used in applications where deformed situations are not considered, and CapsNET should be used in security applications that may show sensitivity in deformed situations. In the case of overlap, the classification accuracy of CNN is obtained 63% for MNIST, 58% for Fashion MNIST, on the other hand, CapsNET is obtained 88% for MNIST and 81% for Fashion MNIST. In this case of overlap, it was observed that CapsNET obtained higher classification accuracy than CNN.

1. GİRİŞ

Bilgisayar bilimleriyle ilgilenen birçok araştırmacının ilgi duyduğu güncel konulardan biri yapay öğrenmedir [1]. İnsan gibi düşünebilen, hareket edebilen veya muhakeme edebilen bilgisayarlar geliştirmek yapay öğrenmenin temel amaçları arasındadır. Derin öğrenme, son yıllarda yapay öğrenmenin popüler bir araştırma alanıdır. Derin öğrenme, klasik öğrenme mimarilerinden farklı olarak daha fazla öğrenme katmanı barındırabilir(GoogleNet 22 katman [2]), giriş parametresi olarak imge, ses sinyalleri gibi veriler alabilir ve daha fazla sınıf barındıran büyük veriler giriş olarak uygulanabilir (ImageNet [3] 1000 sınıf ve 100000 imge). Bu durum yüksek hesaplama maliyeti gerektirir. Fakat öğrenme üzerindeki başarımının yüksek olduğu bilimsel çalışmalarla ortaya konulmaktadır [4]. Evrişimsel Sinir Ağları (ESA), Boltzman Makinaları, Otomatik Kodlayıcılar (Autoencoder), Çekişmeli Üretici Ağlar (ÇÜA) ve Kapsül Ağlar (KA) en sık kullanılan derin öğrenme mimarileridir.

ESA, medikal imgelerin sınıflanması, nesnelerin tanınması ve tespiti gibi uygulama alanlarından sıklıkla kullanılan bir mimaridir [4–7]. ESA temelde evrişim (konvülasyon), aktivasyon fonksiyonu (Relu) ve havuzlama (pooling) katmanlarının belirli sayıda uygulanmasıyla ve sonunda tam bağlı bir sinir ağını barındıran bir yapay öğrenme modelidir. Bu yapının havuzlama katmanı 3x3, 5x5 gibi boyutlarda imgeyi bloklara ayırmaktadır. Bu blokların her biri seçilen havuzlama modeline (maksimum havuzlama, ortalama havuzlama ve minumum havuzlama gibi) göre tek bir skaler değere indirgenmektedir. Yapılan havuzlama işlemi evrişim uygulanmış imgenin büyük bir veri ve özellik kaybına neden olur. Birçok problemde yüksek doğruluk başarımı gösteren ESA'nın bu yönüyle sağladığı güvenilirlik tartışılmaktadır.

KA mimarisi, imgenin temel özelliklerini (konum, yönelim, kalınlık) öğrenebilen derin öğrenme mimarisidir [8,9]. ESA'nın literatürde vurgulanan eksik yönlerine alternatif olarak önerilen KA, temelde evrişim katmanı ile başlamaktadır. Daha sonra sınıf özelliklerine göre boyutlanan vektör tabanlı bir kapsül katmanıyla imgelerin özelliklerini modellemektedir. Kapsül katmanına anlaşmalı yönlendirme algoritması uygulanarak çıkış kestirimi yapılmaktadır. Ayrıca öğrenme ağında çözücü katman kullanılmaktadır. Bu katman, kapsül ağın vektörel çıktılarını giriş olarak alır ve giriş verisinin tekrar yapılanmasını öğrenir.

Literatürde KA'ın hesaplama yükünü azaltmak için temel mimarisindeki gereksiz ağırlıkları budama metoduyla tespit eden mimari sunulmaktadır [10]. Bu çalışmada yapılan analiz sonucunda gereksiz yapılar KA çıkarılarak işlem karmaşası ve hesaplama yükü azaltılmıştır. Diğer bir çalışmada KA mimarisinde anlaşmalı yönlendirme algoritmasına alternatif bir yöntem önerilmiştir. Sunulan yöntem farklı veri kümeleri için başarımı gösterilmiştir [11].

KA literatürde göğüs kanseri imgelerinin sınıflanmasında [12], beyin tümör imgelerinin sınıflanması [13], diyabetik retinopati ve mitoz tespiti [14] gibi medikal alanda sıklıkla kullanılmıştır. Bunun yanında kapsül tabanlı çekişmeli üretici ağ, veri üretimindeki performansı literatürde detaylandırılmıştır [15]. Ayrıca işaret dili tanıma uygulamasında yüksek başarım gösterdiği gözlemlenmiştir [16]. KA, CIFAR [17] gibi birden fazla nesnenin bulunduğu karmaşık veri kümelerinde başarımı düşük olduğu literatürde sunulmuştur [18].

Bu makalede, son dönemlerde popülerliği giderek artan ESA ve KA'ın bozulmuş veya üst üste örtüşme durumlarını içeren imgelerdeki sınıflandırma performansları kıyaslanmaktadır. Bu bağlamda çalışmada öncelikle ESA mimarisi anlatılmaktadır. Daha sonra KA mimarisi sunulmaktadır. Çalışmanın devamında deneysel sonuçlarda kullanılacak veri kümeleri ve özellikleri verilmektedir. Bu veri kümeleri önce bloklarına ayrılmaktadır. Bloklar 1,2 ve 3 parça olarak yer değiştirilerek veri kümesi yeniden düzenlenmektedir. Bunun yanında veri kümesindeki farklı sınıflara ait imgeler, rastgele seçilerek iki imge üst üste kaydırılarak örtüşme oluşturulan yeni test veri kümesi elde edilmektedir. Son olarak ESA ve KA performansları bu farklılaştırılmış iki veri kümesi için detaylandırılarak tartışılmaktadır.

2. EVRİŞİMSEL SİNİR AĞI

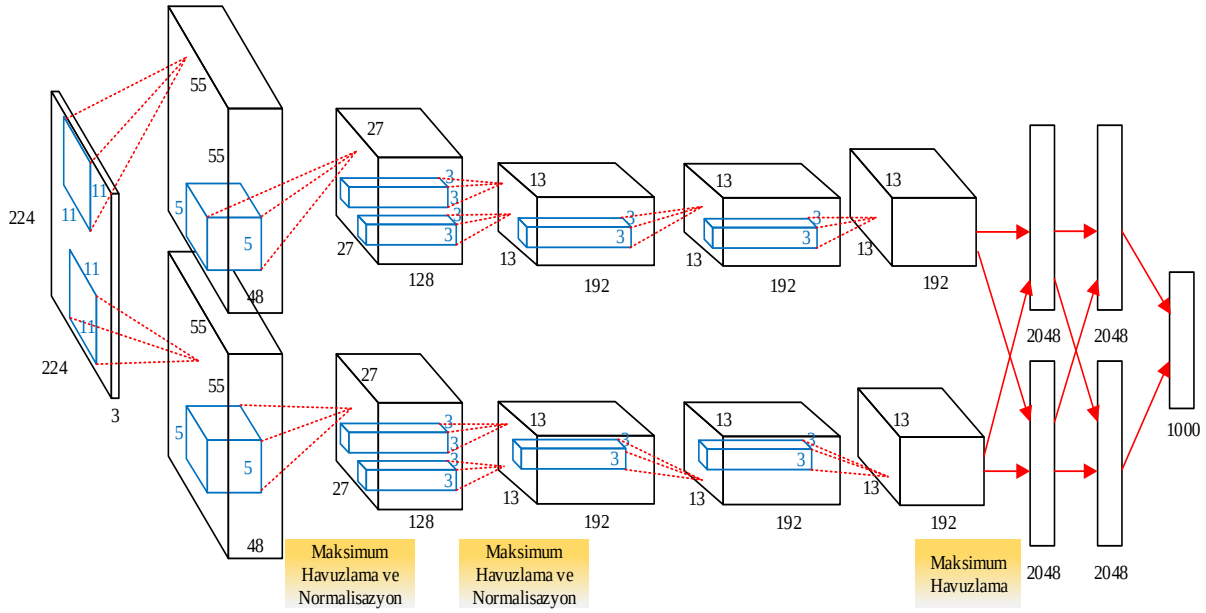
ESA, klasik sinir ağından farklı olarak evrişim katmanı içermektedir. Evrişim katmanın ilk adımı belirlenen sayıda ve boyuttaki filtreyi, kaydırma ile padding değerine göre imgeye uygulanan konvolüsyon işlemidir. Denklem 1' de konvolüsyon işleminin matematiksel formülü verilmektedir.

$$x_{i,j}^{l} = \sum_{a}^{n}\sum_{b}^{n} w_{ab}\, y_{(i+a)(i+b)}^{l-1} \qquad (1)$$

Denklem 1' de de y giriş imgesi w ise uygulanacak konvolüsyon filtresini göstermektedir. Evrişim sonucu yatay kenarlar, dikey, ve açısal kenarlar gibi uygulanan imge ile aynı veya yakın boyutta imge türevleri elde edilmektedir [5]. ESA 'larda her bir konvolüsyon sonucu, relu veya hiperbolik tanjant gibi bir aktivasyon fonksiyonuna verilmektedir. Böylelikle konvolüsyon çıkış değerlerinin belirli bir sayı aralığında olduğu garantilenmektedir. Denklem 2' de Relu fonksiyonun matematiksel modeli gösterilmektedir.

$$\mathrm{z} = \max(0, x) \qquad (2)$$

Denklem 2' de x, konvolüsyon çıkış değerlerini göstermektedir. z ise bu çıkış değerlerinin Relu



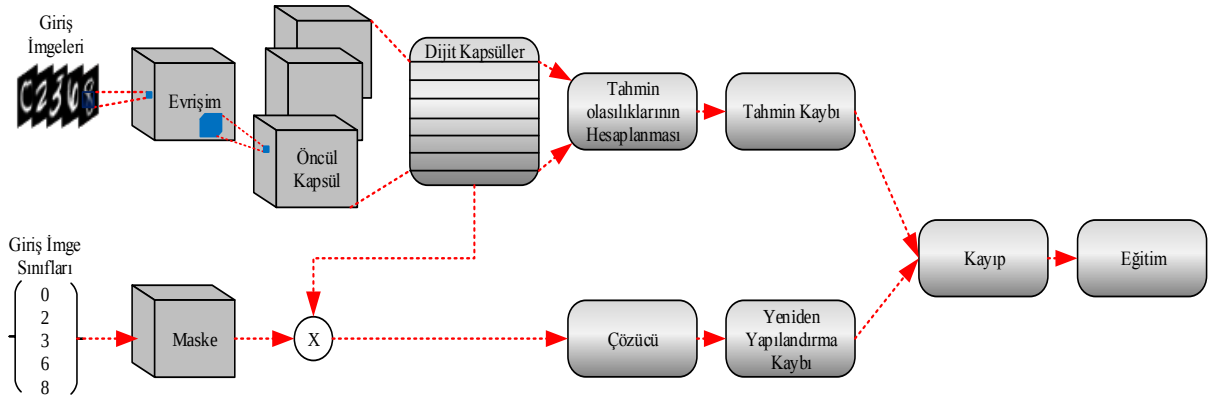
Şekil 1. AlexNET mimarisi [4].

aktivasyon fonksiyonu sonucudur. ESA evrişim işleminin son aşaması havuzlama katmanıdır. Havuzlama 3x3, 5x5 gibi maskeleri imgeye uygulayarak 3x3 filtre için 9 tane, 5x5 filtre için 25 tane değeri maksimum, minimum veya ortalama değerini alarak tek bir skaler değere indirerek boyut küçülten bir katmandır.

ESA'nın son katmanında genel olarak tam bağlı yapay sinir ağı kullanılmaktadır. Bu yapısıyla ESA güçlü bir derin öğrenme mimarisidir. Literatürde birçok çalışmada kullanılan bir ESA ağı olan Alexnet ağının örnek şema modeli Şekil 1' de gösterilmektedir.

3. KAPSÜL AĞ

Sabour ve diğ tarafından ilk olarak literatürde sunulan KA, temelde ESA'nın havuzlama katmanındaki özellik kaybına dikkat çekerek, ESA' ya alternatif olarak tasarlanmıştır. Sekil 2' de KA aşamaları gösterilmektedir [9].



Şekil 2. Kapsül ağ işlem aşamaları [9].

KA ilk olarak bir evrişim katmanıyla imgeleri filtrelemektedir. Bu filtre ağırlıkları başlangıçta rastgele verilmektedir. Şekil 2' de görüldüğü üzere evrişim işleminin sonucu ezme fonksiyonuna giriş olarak verilmektedir. Ezme fonksiyonu denklem 3' te gösterilmektedir.

$$v_j = \frac{\|s_j\|^2}{1+\|s_j\|^2}\frac{s_j}{\|s_j\|} \qquad (3)$$

Burada v_j, j. kapsül çıkış vektörüdür. v_j ise toplam giriş verilerini temsil etmektedir. Bu işlem aşamasından sonra dijit kapsül aşaması gelmektedir. Dijit kapsül aşamasına giriş olarak gelen her vektör ilgili w ağırlığı ile çarpılarak çıkış tahmini yapılmaktadır. Denklem 4' te ağırlıklandırılmış çıkış tahmini gösterilmektedir.

$$\hat{u}_{j|i} = w_{j|i}u_i \qquad (4)$$

Bu aşamadan sonra anlaşmalı yönlendirme algoritması uygulanarak dijit kapsül katmanı çıkış kestirimi yapılmaktadır. Anlaşmalı yönlendirme algoritmasının sözde kodu Tablo 1' de verilmektedir.

Anlaşmalı yönlendirme algoritması uygulandıktan sonra dijit kapsül çıkışlarının uzunlukları hesaplanmaktadır.

Her bir dijit kapsülün olasılıksal çıkışları, kapsülü oluşturan ve veri kümesinin sınıf sayısına eşit olan her bir elemanın uzunluğuna eşittir. Bunun anlamı giriş değerinin hangi sınıfa ait olduğu uzunluk değerine göre tespit edilmektedir. Tespit hatası herhangi bir k dijit kapsülü için L_k, Denklem 5'te verilmektedir.

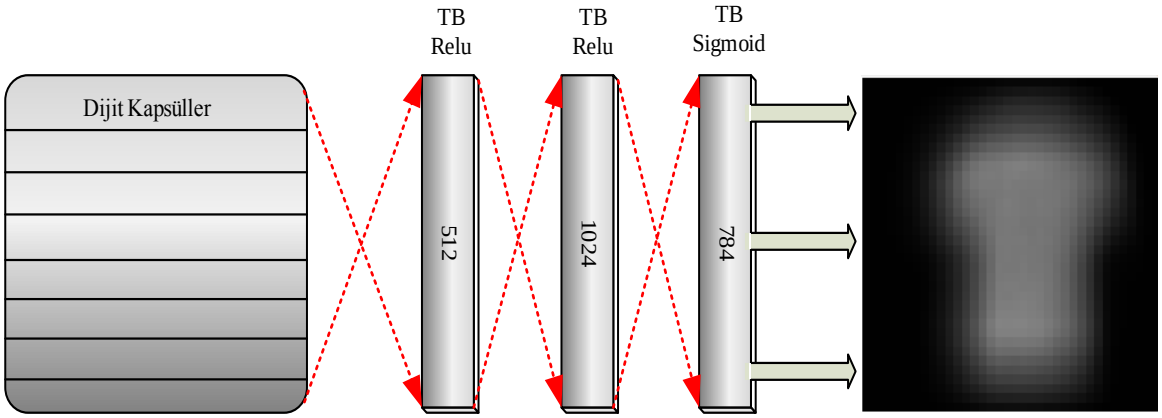
Tablo 1. Anlaşmalı yönlendirme algoritması sözde kodu [9].

Anlaşmalı Yönlendirme
1: fonksiyon Yonlendirme $(\hat{\mathbf{u}}_{j\|i}, \mathbf{r}, \mathbf{l})$
2: for $\mathbf{l}$ katmanındaki bütün $\mathbf{i}$ kapsülleri ve $\mathbf{l}+\mathbf{1}$ katmandaki kapsül j ler için $\mathbf{b}_{ij} \leftarrow \mathbf{0}$
3: for $\mathbf{r}$ iterasyon sayısınca
4: for $\mathbf{l}$ katmanındaki bütün $\mathbf{i}$ kapsülleri için: $\mathbf{c}_i \leftarrow \mathbf{softmax}(\mathbf{b}_i)$
5: for $\mathbf{l}+\mathbf{1}$ katmandaki kapsül j lerin için: $\mathbf{s}_j \leftarrow \sum_i \mathbf{c}_{ij}\hat{\mathbf{u}}_{j\|i}$
6: for $\boldsymbol{l}+\mathbf{1}$ katmandaki kapsül j lerin için: $\boldsymbol{v}_j = \boldsymbol{ezme}(\boldsymbol{s}_j)$
7: for $\mathbf{l}$ katmanındaki bütün $\mathbf{i}$ kapsülleri ve $\mathbf{l}+\mathbf{1}$ katmandaki kapsül j lerin için
8: $\mathbf{b}_{ij} \leftarrow \mathbf{b}_{ij} + \hat{\mathbf{u}}_{j\|i}\mathbf{v}_j$
9: return $\mathbf{v}_j$

$$L_k = T_k \max(0, m^+ - \|v_k\|)^2 + \Delta(1 - T_k) \max(0, \|v_k\| - m^-)^2 \quad (5)$$

Burada T_k değeri imgedeki farklı sınıflara ait tek bir değer olması durumunda 1 olarak belirlenmektedir. m^-, m^+ ve Δ değeri veri kümesinin göre belirlenen sabitlerdir. Kapsül ağın kayıp fonksiyonu, L_k ile dijit kapsülün değerlerini giriş alan çözücü yapısının maliyetinin sabit bir katsayı ile çarpımlarının toplamı olarak belirlenmektedir. Şekil 3' te örnek bir çözücü ağ yapısı modellenmektedir. Çözücü ağ, giriş olarak dijit kapsül verilerini alarak tekrardan imgeyi inşa etmeye çalışmaktadır. Çözücünün giriş verisi dijit kapsül değerleridir. Çıkışı ise imgenin kapsül değerlerine göre yeniden üretilmiş halidir.

Şekil3'te TB, tam bağlı sinir ağıdır. Ayırca Şekil 3'te Fashion MNIST veri kümesi için örnek bir KA ve TB yapılarının nöron sayıları göstermektedir. KA son aşaması eğitim safhasıdır. Kayıp maliyeti hesaplandıktan sonra ağın bütün ağırlıkları güncellenmektedir.



Şekil 3. Çözücü ağ yapısı.

4. VERİ KÜMELERİ

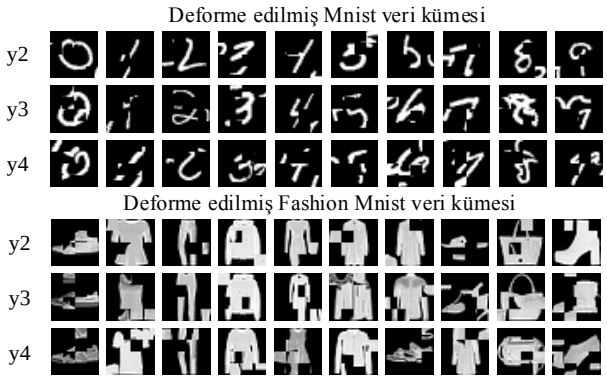
Bu makale çalışmasında MNIST [19] ve Fashion MNIST [20] veri kümeleri kullanılmaktadır. Bu veri kümelerinin her ikisi de 10 adet sınıfa sahip 28x28 boyutundaki 70000 gri formatta imgeden oluşmaktadır. MNIST veri kümesi, el yazması rakamlardan meydana gelmektedir. Fashion mnist veri kümesi ise ayakkabı, elbise, çanta, tişört, pantolon, mont ve bunların farklı türevlerinden oluşmaktadır. MNIST ve Fashion MNIST veri kümesindeki farklı sınıflara ait örnek imgeler Şekil 4' te gösterilmektedir.



Şekil 4. (a) Fashion MNIST (b) MNIST için seçilmiş örnek imgeler.

4.1. NxM Boyutundaki Blokların Yer Değişmesiyle İmgelerin Deformasyonu

ESA ve KA doğruluk ve test güvenliği performanslarını karşılaştırmak için MNIST ve Fashion MNIST veri setleri 7x7 boyutunda bloklara bölünmektedir. Daha sonra bu bloklar arasından rastgele seçilen 2, 3 ve 4 blok kendi arasında yer değiştirilerek yeni örnek imgeleri elde edilmiştir. Deforme edilmiş bu görüntüler Şekil 5' te gösterilmektedir. Bu yeni imge kümesi için y2, iki bloğun, y3 üç bloğun, y4 dört bloğun yer değişmesi anlamına gelmektedir.

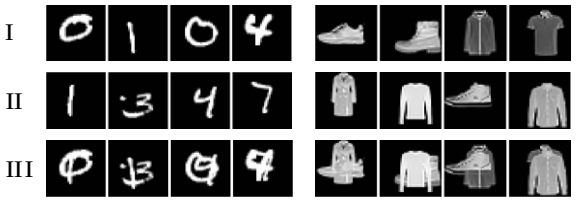


Şekil 5. Deforme edilmiş veri kümeleri

Deformasyon işlemi, sınıflandırılacak nesnelerin bütünlüğünün bozulması durumunda derin mimarilerin performanslarının kıyaslanması için uygulanmaktadır.

4.2. İmgelerin Örtüşmesi

Örtüşme, veri küme sınıfına ait birden fazla üyenin imge üzerinde görünümlerini kısmi perdeleyecek şekilde kapatmasıdır. MNIST ve Fashion MNIST veri kümesi için farklı sınıflara ait imgeler birbirleriyle aynı sahnede kaydırarak örtüşme oluşturulmaktadır. Örtüşme işlemi ile ilgili literatürde çeşitli çözüm önerileri sunulmaktadır [21]. Örtüşme durumunda ESA ve KA performanslarının karşılaştırılması için tam örtüşme durumunda veriler Şekil 6' da sunulmaktadır.



Şekil 6. I. İlk imge, II. İkinci imge, III. İlk ve ikinci imgenin örtüşmüş halleri a. MNIST veri kümesi için örtüşmeli imgelerin oluşumu b. Fashion MNIST için örtüşmeli imgelerin oluşumu

4. DENEYSEL SONUÇLAR

Deneysel sonuçlar, Matlab 2019b ortamında ve Python programlama dili ile gerçekleştirilmektedir. Geliştirilen yazılımlar, 20 çekirdekli 2 işlemcili, 64 gb ram sahip bir sunucu bilgisayarda bulunan Nvidia Quard 4000 gpu hızlandırıcıda yürütülmektedir. Bunun yanında deneysel çalışmalarda kullanılan KA'ın katmanları ve özellikleri Tablo 2' de, ESA'nınki de Tablo 3'te sunulmaktadır. Deneysel çalışmalarda yöntemlerin toplam doğruluk kıyaslanmaktadır. Toplam doğruluk hesabı denklem 6'da sunulmaktadır.

$$Doğruluk\ (D) = \frac{\sum TP}{Toplam\ Veri\ Sayısı} \qquad (6)$$

TP değişkeni, sınıflardaki doğru tahmin edilen veri sayısını ifade etmektedir. Doğruluk değeri, her sınıfta tahmin edilen doğru veri sayısının toplam veri sayısına oranıdır.

Tablo 2. KA katman yapısı.

Katman	Çıkış boyutu	Parametre sayısı	Bağlantısı
İmge giriş (input_1)	(?,28,28,1)	0	-
Konvülasyon1 (conv1)	(?,20,20,256)	20992	input_1
Öncül kapsül (convp)	(?,6,6,256)	5308672	conv1
Öncül kapsül tekrar boyutlama (reshape)	(?,1152,8)	0	convp
Öncül ezme fonksiyonu (lambda)	(?,1152,8)	0	reshape
Dijit kapsül (digitcaps)	(?,10,16)	1474560	lambda
Giriş (input_2)	(?,10)	0	-
Maske (mask)	(?,160)	0	digitcaps input_2
Kapsül ağ (length)	(?,10)	0	digitcaps
Çözücü (decoder)	(?,28,28,1)	1411344	mask

Tablo 2' de verilen KA mimarisinde giriş örnek boyutunun olarak 28x28 olduğu görülmektedir. Eğitim veri kümesindeki eleman sayısı ve parametrelerine göre çıkış boyutundaki sütunun ilk değerleri belirlenmektedir. Bu nedenle o sütundaki ilk boyut '?' ile ifade edilmektedir. Giriş verilerine 9x9 luk 256 adet filtre ile evrişim uygulanmaktadır. Evrişim katmanın ardından kaydırma boyutu 2 olarak belirlenen öncül kapsül katmanı ile resim boyutu 6x6x256 indirgenmektedir. Daha sonra tekrar boyutlama yapılarak 8 elemanlı 1152 adet öncül kapsül çıkışı belirlenmektedir. Dijit kapsül katmanın her bir elemanı sınıf sayısı kadar çıkış içeren 16 adet vektör ile temsil edilmektedir. Dijit kapsülün çıkışı, çözücü katmana giriş olarak verilmektedir. Böylelikle ağın kaybı hesaplanarak eğitim işlemi gerçekleştirilmektedir.

Tablo 3. ESA katman yapısı.

Katman	Filtre boyutu	Filtre sayısı	Padding	Stride
Resim girişi	-	-	-	-
Konvolüsyon1	3x3	8	Same	-
Batch normalisazyonu	-	-	-	-
Relu aktivasyon Fonksiyonu	-	-	-	-
maxPooling	2x2	-	-	2x2
Konvolüsyon2	3x3	16	Same	-
Batch normalisazyonu	-	-	-	-
Relu aktivasyon Fonksiyonu	-	-	-	-
maxPooling	2x2	-	-	2x2
Konvolüsyon3	3x3	32	Same	
Batch normalisazyonu	-	-	-	-
Relu aktivasyon Fonksiyonu	-	-	-	-
Tam bağlı sinir ağı	-	-	-	-
Softmax	-	-	-	-
Sınıflama katmanı	-	-	-	-

Tablo 3' te ESA yapısında öncelikle 3x3 lük 3 adet evrişim sırasıyla giriş veri kümesine uygulanmaktadır. Daha sonra evrişim katmanları arasında toplu normalisazyon ve Relu aktivasyon işlemleri yapılarak veriler normalize edilmektedir. Bunun yanında her evrişim işleminden sonra maksimum havuzlama işlemi yapılarak boyut indirgenmektedir. ESA ağının son aşamasında sınıflama işlemi için tam bağlı sinir ağı, softmax ve sınıflama katmanı uygulanmaktadır.

Deneysel çalışmaların ilkinde, NxM boyutunda blokların yerleri değişmesiyle oluşan test veri kümelerini önceden

eğitim verileriyle eğitilmiş KA ve ESA'ya verilerek sahte imgeler için performansları incelenmektedir. Tablo 4' te KA ve ESA'nın sahte resimlerde toplam doğruluk ve deforme durumuna göre yüzde değişimi gösterilmektedir.

Tablo 4' de YD, başarımın yüzde azalma oranını göstermektedir. Deforme edilmiş Mnsit ve Fashion MNIST imgelerinde ESA'nın KA'ya göre sınıflama başarımının yüksek olduğu görülmektedir. Ancak KA sahte resimlere karşı duyarlılığı ve sınıflama başarımına tepkisi ESA'ya göre daha doğru olduğu gözlemlenmektedir. KA iki veri kümesi için rastgele seçilen test verilerinin tekrar yapılanma sonuçları ve tahmin çıktıları Şekil 7'de sunulmaktadır.

Tablo 4. Y2,Y3 ve Y4 deforme durumunda performans kıyaslanması.

	Veri Kümesi	D	Y2	Yd2	Y3	Yd3	Y4	Yd4
Ka	Mnist	99.2	68	31	61	38	56	43
Esa	Mnist	99	91	5	81	15	83	13.5
Ka	F- Mnist	92	60	34.7	49	46.7	43	53.2
Esa	F- Mnist	90	83,6	7.1	76,8	14.6	67,1	25.4

Tablo 4' te YD, başarımın yüzde azalma oranını göstermektedir. Deforme edilmiş Mnsit ve Fashion MNIST imgelerinde ESA'nın KA' ya göre sınıflama başarımının yüksek olduğu görülmektedir. Ancak KA sahte resimlere karşı duyarlılığı ve sınıflama başarımına tepkisi ESA'ya göre daha doğru olduğu gözlemlenmektedir. KA iki veri kümesi için rastgele seçilen test verilerinin tekrar yapılanma sonuçları ve tahmin çıktıları Şekil 7' de sunulmaktadır.



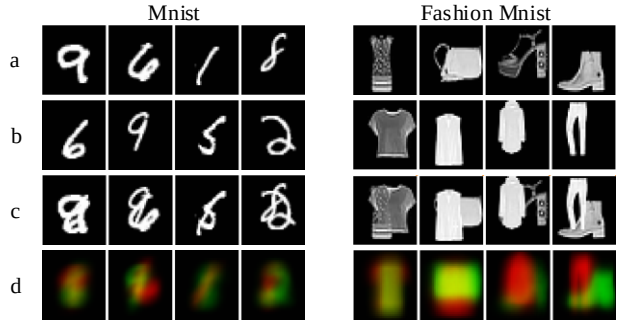
Şekil 7. Deforme edilmiş görüntülerin KA çözücü sonuçları a. MNIST veri kümesi, b. MNIST çözücü sonuçları, c. Fashion MNIST veri kümesi, d. Fashion MNIST çözücü sonuçları.

KA sanal verileri gerçeğinden ayırt etme yeteneği ile ESA'nın bir adım önüne geçmektedir. ESA ağ yapısı verilerin konum, yönelim, kalınlık gibi önemli özniteliklerini KA' ya göre daha sınırlı miktarda öğrendiği gözlemlenmektedir. ESA sinir ağının örnek sanal imgeleri ayırt etme yeteneği KA göre daha düşüktür. ESA, yapısı büyük ölçüde deforme olmuş imgeleri orijinal imgenin parçalarının yer değişmesi olarak algılayarak deformasyon durumunda hassasiyet göstermemektedir. KA ise tekrar yapılanma kaybı ve dijit kapsül çıkış kestirim kaybını göz önüne alarak imgenin şekil bütünlüğüne karşı hassasiyet göstermektedir.

Deneysel çalışmaların ikincisinde örtüşme durumunda KA ve ESA performansları karşılaştırılmaktadır. Öncelikle iki veri kümesine ait veriler kendi aralarında farklı sınıflardan birer tane olmak üzere rastgele seçilerek örtüşme işlemi uygulanmaktadır. Örtüşme durumunda ağların doğruluk performansları Tablo 5' te gösterilmektedir. KA örtüşme durumu ile başa çıkmak için ağın kayıp fonksiyonunda T_k değerini değiştirerek ve dijit kapsül sayısını artırarak imgeden birden fazla eleman çıkarımı yapmaktadır. Bunun yanında çözücü katman yardımıyla imgede var olan farklı sınıflara ait bilgileri çözmeye çalışmaktadır. Şekil 8' de örtüşme durumunda KA' ın çözücü sonucu verilmektedir.

Tablo 5. Örtüşme durumlarında doğruluk kıyaslanması.

Mimari	Veri Kümesi	Tek İmge Doğruluk(%)	Örtüşme Doğruluğu(%)
Ka	Mnist	99.2	88
Esa	Mnist	97	63
Ka	F-Mnist	92	81
Esa	F-Mnist	90	58



Şekil 8. MNIST ve Fashion MNIST veri kümeleri için seçilen örnek örtüşme imgeleri ve KA sonuçları. a. ilk imge, b. ikinci imge, c. örtüşme durumu, d. KA çözücü sonucu.

Şekil 8' de örtüşme durumunda kapsül ağın çözücü katmanının 2 adet maske ürettiği görülmektedir. Bu maskelerin kırmız ve yeşi renk ile üst üste konumu gösterilir. KA, ESA' ya göre örtüşme durumlarında eğitim veri setinde yüksek başarım göstermektedir. Bu durum KA, kapsül yapısı ve çözücü katmanın ağa kazandırdığı önemli bir yetenektir. Sınıflara ait yönelim bilgisini öğrenebilen KA, örtüşme durumunda bile imgenin hangi sınıfa ait olduğunu yüksek başarımla tahmin edebilmektedir. ESA, örtüşme durumunda 2 farklı sınıfa ait bilgileri aynı imge üzerinde öğrenmede zorlanmaktadır. Bunun temel nedeni ESA ağında bulunan havuzlama katmanından dolayı imgeye ait

konum ve yönelim bilgilerinin kaybolması olarak yorumlanmaktadır.

5. TARTIŞMA VE SONUÇLAR

Bu çalışmada KA ile ESA gibi iki derin öğrenme mimarisini farklı veri kümelerindeki performansları değerlendirilmiştir. Veri kümesindeki örneklerde örtüşme ve deformasyon durumları gözetilmiş ve hassasiyet ve doğruluk kıstasları irdelenmiştir. KA yaklaşımının deformasyon durumunda ESA'ya göre daha fazla hassasiyet gösterdiği gözlemlenmiştir. ESA'nın yapısal bileşenleri deformasyon durumunu anlamayı engellemektedir. ESA'nın bu davranışı nesne tanıma gibi bazı uygulamalarda olumlu olarak değerlendirileceği gibi güvenlik açısından bir takım zafiyetlerin ortaya çıkmasına neden olmaktadır. Örtüşme durumunda KA ve ESA performanslarına bakıldığı zaman, KA, MNIST veri kümesi için %88, Fashion MNIST veri kümesi içinde %81 doğrulukla ESA'ya karşı büyük bir üstünlük göstermektedir. Bu açıdan bakıldığında örtüşme durumlarının sıklıkla gerçekleştiği veri kümeleri için KA önemli bir yapay öğrenme mimarisi olarak öne çıkmaktadır.

İleriki çalışmalarda, makaledeki tecrübelerden faydalanılarak biyomedikal imge sınıflama ve bölütleme gibi problemlere farklı çözümler sunulması hedeflenmektedir.

KAYNAKLAR

[1] Bishop CM. Pattern Recognition and Machine Learning. New York: Springer; 2006.

[2] Szegedy C, Wei Liu, Yangqing Jia, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. 2015 IEEE Conf. Comput. Vis. Pattern Recognit., IEEE; 2015, p. 1–9. https://doi.org/10.1109/CVPR.2015.7298594.

[3] İnternet. Image.net 2019. http://www.image-net.org (accessed December 20, 2019).

[4] Krizhevsky A, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. Adv Neural Inf Process Syst 2012;1907–1105:1–9.

[5] Firildak K, Talu MF. Evrişimsel Sinir Ağlarında Kullanılan Transfer Öğrenme Yaklaşımlarının İncelenmesi. Bilgi Bilim 2019;4:88–95.

[6] Şengör, Abdülkadir; Yaman, Akbulut; Yanhui, Guo; Varun B. Classification of amyotrophic lateral sclerosis disease based on convolutional neural network and reinforcement sample learning algorithm. Heal Inf Sci Syst 2017. https://doi.org/https://doi.org/10.1007/s13755-017-0029-6.

[7] ÇALIŞAN M, TALU MF. Comparison of Methods for Determining Activity from Physical Movements. J Polytech 2020. https://doi.org/10.2339/politeknik.632070.

[8] Hinton G, Sabour S, Frosst N. Matrix capsules with EM routing. 6th Int. Conf. Learn. Represent. ICLR 2018 - Conf. Track Proc., 2018.

[9] Sabour S, Frosst N, Hinton GE. Dynamic routing between capsules. Adv Neural Inf Process Syst 2017;2017-Decem:3857–67.

[10] Jeong T, Lee Y, Kim H. Ladder Capsule Network. Proc 36th Int Conf Mach Learn 2019:3071--3079.

[11] Heinsen FA. An Algorithm for Routing Capsules in All Domains 2019.

[12] Anupama MA, Sowmya V, Soman KP. Breast cancer classification using capsule network with preprocessed histology images. Proc 2019 IEEE Int Conf Commun Signal Process ICCSP 2019 2019:143–7. https://doi.org/10.1109/ICCSP.2019.8698043.

[13] Afshar P, Mohammadi A, Plataniotis KN. Brain Tumor Type Classification via Capsule Networks. Proc - Int Conf Image Process ICIP 2018:3129–33. https://doi.org/10.1109/ICIP.2018.8451379.

[14] Jiménez-Sánchez A, Albarqouni S, Mateus D. Capsule Networks Against Medical Imaging Data Challenges, 2018, p. 150–60. https://doi.org/10.1007/978-3-030-01364-6_17.

[15] Saqur R, Vivona S. CapsGAN: Using Dynamic Routing for Generative Adversarial Networks 2018.

[16] Beser F, Kizrak MA, Bolat B, Yildirim T. Recognition of sign language using capsule networks. 26th IEEE Signal Process Commun Appl Conf SIU 2018 2018:1–4. https://doi.org/10.1109/SIU.2018.8404385.

[17] Krizhevsky A, Hinton GE. Learning multiple layers of features from tiny images. University of Toronto, 2009.

[18] Xi E, Bing S, Jin Y. Capsule Network Performance on Complex Data 2017;10707:1–7.

[19] Bengio Y, Haffner P. Gradient-Based Learning Applied to Document Recognition 1998;86.

[20] Xiao H, Rasul K, Vollgraf R. Fashion-MNIST: a Novel Image Dataset for Benchmarking Machine Learning Algorithms 2017.

[21] Greff K, Rasmus A, Berglund M, Hao TH, Schmidhuber J, Valpola H. Tagger: Deep Unsupervised Perceptual Grouping 2016.