

# Evrişimsel Sinir Ağlarında Kullanılan Transfer Öğrenme Yaklaşımlarının İncelenmesi

Investigation of Transfer Learning Approaches Used Convolution Neural Networks

Kazım FIRILDAK<sup>1</sup>, Muhammed Fatih TALU<sup>2</sup>

<sup>1</sup>Fırat Üniversitesi, Elazığ, Türkiye ([kfirildak@firat.edu.tr](mailto:kfirildak@firat.edu.tr))

<sup>2</sup>İnönü Üniversitesi, Malatya, Türkiye ([fatihaltu@gmail.com](mailto:fatihaltu@gmail.com))

Received: Feb.15, 2019

Accepted: May 4, 2019

Published: Dec.1, 2019

**Özetçe** — Evrişimsel sinir ağları (ESA), yapay sinir ağları tabanlı derin öğrenme mimarileridir. Katman ve sinir hücresi sayısının, yapay sinir ağına göre fazlalığından dolayı ESA eğitiminin hesaplama maliyeti yüksek bir işlemdir. Bunun yanında probleme özgü eğitim kümesi her zaman sahip değiliz. Sınıflama başarısı kanıtlanmış, büyük ve kapsamlı eğitim veri setiyle eğitilmiş bir ESA'nın katmanlarından ağırlık transferi yaygın kullanılan yöntemdir. Ön eğitim, özellik çıkarıcı ve kısmi özellik çıkarıcı yaklaşım ESA'larda kullanılan transfer öğrenme yaklaşımlarıdır. Bu çalışmada Cifar, Caltech, Mnist veri kümeleri için AlexNet'den transfer edilen ağırlıklarla ESA'lar için sınıflama başarıları incelenmiştir. AlexNet, yeni veri kümelerinin sınıflamasında kullanmak için ön işlemlere tabi tutularak farklı veri kümeleri için yüksek sınıflama başarısı göstermiştir. Yeni veri kümesi sınıflarının, AlexNet'in eğitim kümesi ile benzerliği arttıkça sınıflama başarısının transfer öğrenmede arttığı gözlemlenmiştir. Bunun yanında kaynak ESA'nın eğitim veri kümesi, hedef veri kümesine olan benzerliği sınıflama doğruluğunu artırmaktadır. Transfer öğrenme ile ESA'nın farklı veri kümeleri için %80-%90 ortalama sınıflama başarısına ulaştığı da makalede sunulmaktadır.

**Anahtar Kelimeler:** Derin öğrenme, Transfer öğrenme, Evrişimsel sinir ağlarında transfer öğrenme.

**Abstract** — Convolutional neural networks (CNN) are deep learning architectures based on artificial neural networks. The cost of calculation of CNN training is high due to the high number of layers and neurons compared to the artificial neural network. Furthermore, We do not always have a training data set for each problem. The transfer of weight from the layers of CNN with a well-defined, large and comprehensive training data set is a widely used method. Pre-training, feature extraction and fine-tuning approach are transfer learning approaches used in CNNs. In this study, classifying achievements for CNNs were analysed for Cifar, Caltech and Mnist datasets with weights transferred from AlexNet. AlexNet has been subjected to pre-processing for use in the classification of new datasets, and has shown a high classification success for different data sets. It is observed that the success of the classification of new dataset classes increases with transfer learning as the similarity with AlexNet's education set increases. Besides, the similarity of source CNN's training data set to the target dataset increases the classification accuracy. It is also reported in the article that CNN achieves an average classification success of 80% to 90% with transfer learning for different data sets.

**Keywords**— Deep learning, Transfer learning, Transfer learning for convolution neural networks.

## 1. Giriş

Eğitim bilimlerinde öğrenme kavramı, “bireyin, yaşantılar sonucu davranışında meydana gelen kalıcı değişiklik” olarak tanımlanmaktadır. Yapay öğrenme, yapay bir sistemin (yazılımın) gerçek bir sistemi taklit edebilmesi, yani benzer girişlere benzer çıktılar üretebilmesi için parametrelerini güncelleme işlemidir. Teknolojik gelişmeler ışığında

bilgisayar bilimciler, popüler bir araştırma alanı olan yapay öğrenme yöntemlerini geliştirerek insan gibi karar verebilen bilgisayar sistemleri oluşturmaya çalışmaktadır [1][2].

Güncel bir yapay öğrenme yaklaşımı olan derin öğrenme, yüksek doğrulukta sınıflandırma kabiliyetine sahip olması nedeniyle birçok araştırmacının ilgisini çekmeyi başarmıştır. Klasik yapay sinir ağı yaklaşımından üç temel farklılığı görülmektedir: 1) Sistem girişi vektör yerine imge olabilir; 2) Ara katman sayısı oldukça fazla olabilir (örnek olarak GoogleNet [3] 22 katman); 3) Eğitim kümesindeki imge ve kategori sayısı oldukça yüksek olabilir (örnek olarak Yahoo Flickr [4] 100M imge ve 2K kategori). Bunun yanında Boltzman makineleri, autoencoders, derin inanç ağları, ve evrimsel sinir ağları (ESA), derin öğrenme alanında kullanılan farklı mimariler olarak öne çıkmaktadır.

Kullanım kolaylığı, yüksek eşleştirme kabiliyeti ve GPU destekli paralel çalışma özellikleri nedeniyle ESA, örüntü sınıflandırma, nesne tanıma ve tespit etmede güncel yöntemler içerisinde yer almaktadır [5][6][7]. Ayrıca ESA mimarisi ilk olarak ImageNet [8] ( 1000 sınıf ve 1000000 resim) gibi büyük bir eğitim kümesinde kullanılmıştır. Giriş imgelerin ilgili sınıflara yüksek bir doğrulukta (eğitim hatası: 0.15) eşleştiğini gösteren bu çalışma birçok araştırmacıya ilham kaynağı olmuştur [9].

Problem çözümlerinde ESA'yı tercih eden araştırmacıların sıklıkla rastladığı sorunlardan biri eğitim probleme özgü eğitim veri kümesinin teminidir. ESA'nın yüksek katman sayısı ile probleme göre genel ve özel özellikleri doğru bir şekilde çıkarabilmesi için nicel ve nitel olarak yeterli örüntüleri içeren bir eğitim kümesinin araştırmacılar tarafından oluşturulması gerekmektedir. Örnek olarak ImageNet veri kümesi, bir milyon imgenin sırayla bin kategoriden biriyle etiketlenmesi sonucu oluşmuştur. Büyük bir emek sonucu hazırlanan bu eğitim kümesi kullanılarak eğitilen bir ESA oldukça yüksek doğrulukta sınıflandırma yapabilmektedir. Ancak probleme özgü (örneğin medikal imge sınıflamada) etiketlenmiş yeterli sayıda veri bulmak oldukça güçtür veya bulunamamaktadır. Yeterli verinin bulunamaması durumunda yüksek sınıflama doğruluğu sağlayan ESA'yı nasıl kullanacağız sorunu karşımıza çıkmaktadır. Bu sorunun çözümü için "bilgi transferi" yaklaşımı kullanılmaktadır. Bilgi transferi ImageNet ve ya eğitim başarısı ispatlanmış bir veri kümesi ile eğitilmiş bir ESA ağından bilgi transferinin gerçekleştirilmesidir. Literatürde "transfer öğrenme" olarak adlandırılan bu yöntem, derin öğrenme yöntemlerinin geliştirilmesiyle önemli bir araştırma konusu haline gelmiştir.

Transfer öğrenme yaklaşımları, insan öğrenme modelinden etkilenmektedir. İnsanlar öğrenme işleminde, daha önce karşılaşmadıkları bir problemi çözmek için hayatında önceden tecrübe ettiği problemlerin çözümünden faydalanır [10][11]. Özetle, büyük bir veri kümesiyle eğitilmiş bir ağı ağırlıkları farklı bir veri kümesi için kullanılması anlamına gelir. Transfer öğrenme yaklaşımları Prat'ın (1993) çalışması ile ivme kazanmıştır [11]. Bu çalışmada etiketlenmiş bir veri kümesiyle eğitim yapıldıktan sonra öğrenilen katsayılar kullanılarak farklı bir veri kümesi sınıflandırılmak istenmiştir. Ayrıca literatürde transfer öğrenmeyi takviyeli öğrenme üzerinde kullanan ve özellik transferi yapan çalışmalar bulunmaktadır [12].

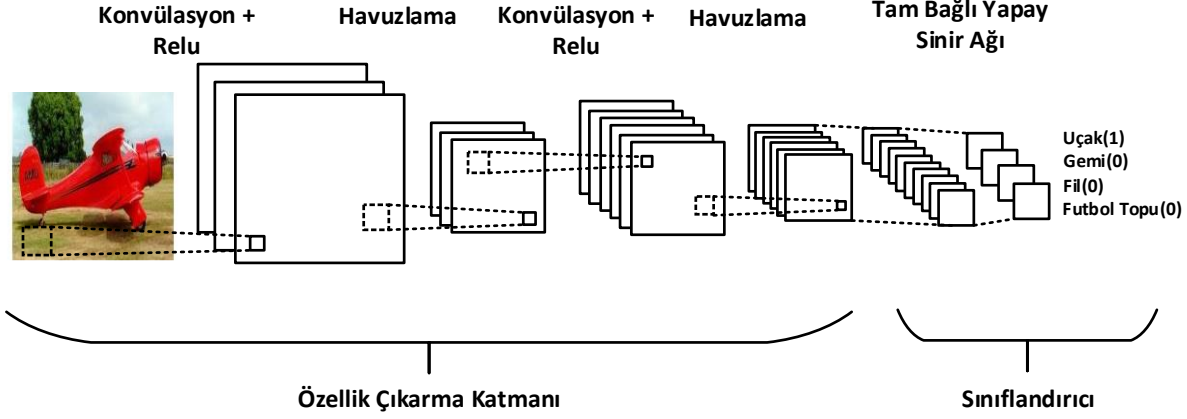
Bilindiği üzere, yapay bir ağı ilk katmanları genel imge özelliklerini çıkaran filtreler (Gabor filtresi, renk etiketi gibi), son katmanlarıysa veri kümesine özgü özellikleri çıkaran filtre katsayılarını içermektedir [13]. Transfer öğrenme alanında yapılan güncel araştırmalar, hangi katmanlarda ve hangi ölçüde bilginin aktarılması gerektiğine odaklanmıştır. Katmanlar arasındaki genelden özele doğru geçişle ilgili aşağıda listelenen problemler hala güncelliğini korumaktadır [13].

1. Belirli bir katmanın genel veya spesifik olma derecesini ölçebilir miyiz?
2. Katmanlar arası geçiş birdenbire tek bir katmanda mı, yoksa birkaç katmana mı yayılıyor?
3. Bu geçiş nereden geçiyor: ağı ilk, orta veya son katmanın yakınında mı?

Bu makalede farklı transfer öğrenme yaklaşımları incelenmiştir. Bu yaklaşımların performans ve doğruluk kriterleri kıyaslamalı bir şekilde karşılaştırılmış ve elde edilen sonuçlar tablolarla sunulmuştur. Bu amaçla, Makalenin ikinci bölümünde ESA yaklaşımı özetle verilirken üçüncü bölümde, transfer öğrenme yaklaşımı ifade edilmektedir. Dördüncü bölümde, ESA kullanılan transfer öğrenme stratejileri açıklanmaktadır. Beşinci bölümde eğitilmiş bir farklı bir veri kümesi üzerindeki başarıyı gözlemlenmiştir. Son bölümde transfer öğrenme hakkında yapılan uygulamalar neticesinde elde edilen genel sonuçlar paylaşılmaktadır.

## 2. Evrişimsel Sinir Ağları

Sınıflandırma başta olmak üzere birçok bilgisayarlı görü alanında kullanılan evrişimsel sinir ağları klasik sinir ağlarından farklı olarak, evrişim özellik çıkarma ve sınıflandırma katmanlarını içermektedir. Klasik bir ESA yapısı Şekil 1 de gösterilmiştir.



Şekil 1. Klasik bir ESA mimarisi [15].

Evrişim katmanı konvolüsyon (convolution), aktivasyon fonksiyonu (relu) ve havuzlama (pooling) aşamalarını içermekte ve yüksek boyutlu veriden düşük boyutlu özniteliklerin çıkarılmasını sağlamaktadır. Probleme özgü olarak evrişim katmanının ardışıl bağlanma derinliği değişmektedir. Sınıflandırma katmanı ise düşük boyutlu öznitelikler ile kategorilerin eşleşmesini sağlar ve genellikle tam bağlı bir yapay sinir ağı mimarisine sahiptir.

Evrişim katmanının ilk aşamasında bulunan konvolüsyon işlemi (denklem 1), imgenin bir maskeyle filtrenmesini sağlar [14]. Konvolüsyon sonucunda yatay kenarlar, dikey kenarlar, açısal kenarlar, yumuşatılmış imge, keskinleştirilmiş imge gibi orjinal imgeyle aynı boyutta imge türevleri elde edilmektedir.

$$x_{i,j}^l = \sum_a \sum_b w_{ab} y_{(i+a)(j+b)}^{l-1} \quad (1)$$

Evrişim katmanında konvolüsyondan sonra aktivasyon aşaması bulunmaktadır. Klasik sinir ağlarında sigmoid, tanjant hiperbolik gibi aktivasyon fonksiyonları kullanılırken ESA'da relu fonksiyonu (denklem 2) tercih edilmektedir. Doğrusal olmayan bu fonksiyon, imgedeki negatif değerlerin elenmesini sağlamaktadır.

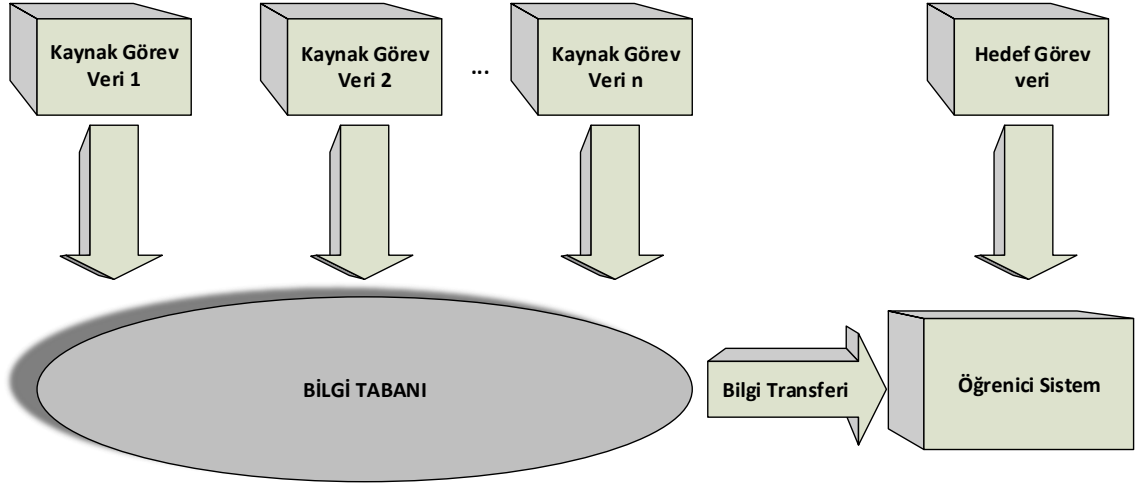
$$y = \max(0, x) \quad (2)$$

Evrişim katmanındaki son aşamada havuzlama (pooling) işlemi gerçekleştirilir. Bu işlem genellikle 2x2 lokal imge matrisi üzerinde maksimum, minimum, toplam ve ya Öznitelikler çıkarıldıktan sonra elde edilen düşük boyutlu imgeler sınıflandırma katmanına (tam bağlı yapay sinir ağı) aktarılır. Sinir ağının çıkışında hesaplanan hata değeri geri yayılım yöntemi kullanarak konvolüsyon filtre katsayılarının ve tam bağlı yapının katmanlar ağırlıklarının güncellenmesinde kullanılır.

## 3. Transfer Öğrenme

Transfer öğrenme yapay öğrenme sistemlerinin eğitim aşamasıyla öğrendiği bilgiyi farklı ve ya benzer problem çözümlerinde kullanılmasını inceleyen öğrenme yaklaşımıdır. İnsan tabiatının temelinde, transfer öğrenme yaklaşımı bulunmaktadır. Yani insanoğlu bilerek ya da bilmeyerek geçmiş tecrübelerinden yararlanarak karşısına çıkan farklı bir olaya çözüm bulmaktadır [11].

Gerçek hayatta insanın ilk kez karşılaştığı ve ani karar vermek zorunda olduğu (eğitim bilgisi edinmeye zamanı olmadığı) anlar olabilir. Bu anlarda insanın vereceği kararlar, hayatı boyunca öğrendiği sistemin yeni duruma tepkisi olacaktır. Şekil 2'de transfer öğrenmenin temel şeması gösterilmektedir



Şekil 2. Transfer öğrenme çalışma prensibi [11].

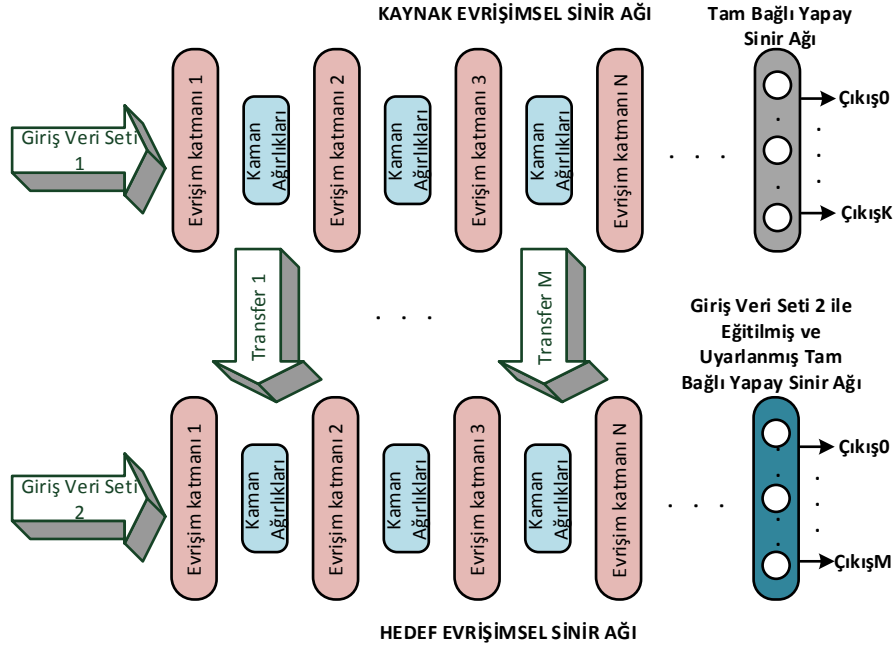
Krizhevsky ve diğerlerinin önerdiği 5 evrişim katmanlı 3 tam bağlı sınıflandırıcı bir ESA (60 milyon ağırlık ve 650 000 hücre) eğitildikten sonra öğrenilen bilgiler (ağırlıklar) farklı problem çözümlerinde kullanılmış ve yüksek başarımın elde edildiği gösterilmiştir [16]. Ayrıca[17] nolu çalışmada etiketli verilerde kullanılarak yüksek sınıflandırma başarısının elde edildiği gösterilmiştir.

#### 4. Evrişimsel Sinir Ağları İçin Kullanılan Transfer Öğrenme Yaklaşımları

##### 4.1. Özellik Çıkarıcı Metot

Yeni veri kümesine göre tasarlanmış bir ESA'nın evrişim katman ağırlık değerlerinin eğitilmiş bir ESA'dan kopyalanması ve eğitim işleminin sadece sınıflandırma katmanında yapılması şeklindeki kullanımdır. Sınıflandırma katmanı yapay sinir ağı olacağı gibi destek vektör makinesi gibi lineer bir sınıflandırıcıda olabilir. Böylece eğitilmiş ESA'nın özellik çıkarma kabiliyeti yeni ESA'ya aktarılmaktadır. Bu kullanım Şekil 3'de görselleştirilmiştir.

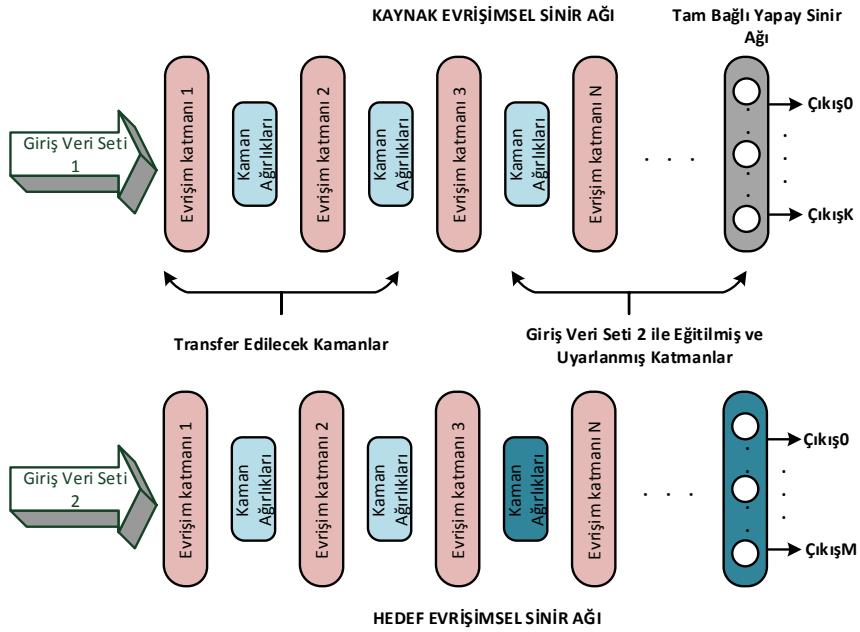
Piramitleme aşamasında, imge çözünürlüğü kademeli olarak azaltılır (downsampling) ve artırılır (upsampling). Üretilen imgeler üst üste dizildiğinde piramit yapısı oluştuğu için bu ismi almıştır. Çözünürlük azaltma işlemi kabalaştırma, arttırma işlemi keskinleştirme olarak ifade edilebilir. Piramitleştirme işlemi giriş ve gürültü imgelerinin her ikisine birden uygulanır ve aynı seviyede olan imgelerin histogramları karşılaştırılır. Histogram farkı minimum olması için gürültü imgesi iteratif bir şekilde güncellenir [13],[14]. Şekil 4'te piramit tabanlı desen yayma yönteminin örnek bir imge üzerindeki sonucu gösterilmektedir.



Şekil 3. Özellik Çıkarıcı Metod.

### 3. Kısmi Özellik Çıkarıcı Metot

Eğitilmiş ESA'nın evrişim katmanındaki ağırlıkların bütünü yerine belirli bir kısmının yeni ESA'ya kopyalanması ve yeni ESA'daki kalan evrişim katmanı ile birlikte sınıflandırıcı katmanın yeni eğitim kümesi kullanılarak güncellenmesi şeklindeki kullanımdır. Bu kullanımda eğitilmiş ESA'nın kaba özellik çıkarma kısmının transfer edildiği görülmektedir. Bu kullanım Şekil 4.'de görselleştirilmiştir.



Şekil 4. ESA'nın kısmi özellik çıkarıcı olarak kullanılması

### 4.3. Öneğitim Modeli

Öneğitim stratejisi daha önce bir eğitim setiyle eğitilip başarıyı kanıtlanmış bir ağın ağırlıklarının kullanılmasıdır. Günümüz Alexnet gibi çoklu katmana sahip ESA'lar Gpu ile ImageNet üzerinde bir ağ eğitimleri günler hatta haftalar (kullanılan donanım özelliklerine göre değişkenlik göstermektedir) sürmesinden dolayı, son ESA kontrol noktalarını

ince ayar veya özellik çıkarmak için ağırları kullanabilecek araştırmacılara, kontrol noktalarının erişime açık bıraktığını görmek yaygın bir durumdur. Örneğin, Caffè [18] kütüphanesinin insanların ağırlıklarını paylaştığı bir Zoo modeli vardır.

Öğretim stratejisi literatürde ince ayarlama ve özellik çıkarıcı yöntemler için vazgeçilmez bir önışlemdir. Literatürde derin ağlarda rastgele ağırlıklara göre önceden eğitilmiş bir ağırlıklarının kullanılmasının daha iyi sınıflandırma sonuçları verdiği gösterilmiştir [13].

## 5. Deneysel Sonuçlar

ESA'lar için kısmi özellik çıkarıcı ve özellik çıkarıcı transfer öğrenme yaklaşımları Tablo 1' de özellikleri verilen farklı veri kümeleri uygulanmış ve sınıflama sonuçları bu bölümde detaylandırılmıştır. Deneysel sonuçlar Matlab2018b ortamında ve 20 çekirdeğe sahip 2 işlemciden oluşan bir Cpu, Nvidia quard 4000 gpu hızlandırıcı kartta ve 64 gb ram sahip hesaplama sunucusunda yürütülmüştür. Transfer edilecek ağırlıklar, ImageNet ile eğitilmiş 8 katmandan oluşan ve AlexNet olarak bilinen ESA'dan öğretim olarak alınmıştır.

**Tablo 1.** Veri kümeleri

	<b>Mnist[19]</b>	<b>Cifar10[20]</b>	<b>Caltech101[21]</b>
Veri Kümesi Sayısı	60000	50000	~10000
Veri Kümesi Sınıf Sayısı	10	10	101
Renk Uzayı	Binary	RGB	Gri ve RGB
ImageNet Benzerliği	Yok	Az	Orta
Veri Türü	El yazısı ile rakamlar	Nesneler	Nesneler

Uygulamalarda kullanılan kısmi özellik çıkarıcı yaklaşımı, Alexnet' in son tam bağlı katmanındaki sınıflama için kullanılan tam bağlı katmanları kaldırarak veri kümesi sınıf sayısına göre tekrardan tam bağlı yapı oluşturulmuştur. Veri kümeleri farklı sayıda eğitim veri sayısı ve 227x227 çözünürlüğe dönüştürülen verilerle ESA'nın bütün katmanları sgd (single gradient descent), 256 veri kümeleri ile 0,001 öğrenme katsayısı ile yeniden eğitilmiştir. ESA'larda transfer öğrenme ile yeniden eğitilmesi aşamasında düşük öğrenme katsayısı kullanılması literatürde aktarılmıştır [22]. Özellik çıkarıcı yöntemde ise Alexnet'in ilk 7 katmanın çıkış sonucu 4096 tane değeri olan özellik vektörü Cifar, Mnist, Caltech veri kümesine uygulanmıştır. Uygulama sonucunda çıkan özellikler destek vektör makinesiyle sınıflandırılmıştır.

Deneysel sonuçlardaki doğruluk oranı, eğitim verisi dışından kalan test verilerinin sisteme uygulanarak elde edilen doğru tahminlerin yüzde oranıdır. Verinin tamamı kullanılmadan eğitilen sistemler yüksek tahmin başarımlarıyla transfer öğrenmenin derin öğrenme yöntemlerinde ne kadar etkili sonuçlar ortaya koyduğu önemli bir gelişme olarak not edilmiştir. Tablo 2' de farklı veri kümelerinin Alexnet ile kısmi özellik çıkarılması metodunun ve Tablo 3' de özellik çıkarıcı metodun sonuçları verilmiştir.

**Tablo 2.** Bilgi transferi sonrası sınıflama sonuçları. Eğitilmiş ESA'nın özellik çıkarma kabiliyeti kısmi olarak transfer edilmiştir.

Veri Kümeleri	Eğitim Sonuçları		
	Veri boyutu	Doğruluk (%)	Yürütülme Zamanı(dk)
Cifar10	5000	85,63	14,29
	25000	91,41	74
	45000	92	127
Mnist	6000	97	17,30
	30000	99,23	76
	54000	99,8	163
Caltech101	1000	74	7,46
	5000	89	11
	9000	91	14,14

**Tablo 3.** Bilgi transferi sonrası sınıflama sonuçları. Eğitilmiş ESA'nın özellik çıkarma kabiliyeti bütünüyle transfer edilmiştir.

Veri Kümeleri	Eğitim Sonuçları		
	Veri boyutu	Doğruluk (%)	Yürütülme Zamanı(dk)
Cifar10	5000	78,33	5,14
	25000	81,56	11,10
	35000	82,13	28,13
Mnist	6000	78	5,46
	30000	98,85	6,52
	42000	99,02	7,93
Caltech101	1000	62,69	21,59
	5000	78	32,84
	9000	83	43,86

Tablo 2 ve Tablo 3' de görüldüğü gibi kısmi ve özellik çıkarıcı yaklaşım, Cifar, Mnist ve Caltech eğitim verilerinin sayısı artması genel olarak sınıflama doğruluk oranını da artırmaktadır. Ayrıca doğruluk oranı ile özelliği transfer edilen ESA'nın eğitim kümesi ile yeni veri kümesinin benzerliğine bağlı olduğu da gözlemlenmiştir. Mnist veri kümesi basit el yazması rakamları benzerliği olmamasına rağmen yüksek doğruluk oranıyla sınıflandırmaktadır. Bu durum ImageNet ile eğitilmiş ESA'nın veri kümesinin özelliklerini çok iyi öğrendiğini göstermektedir. Cifar ve Caltech veri kümesi, ImageNet eğitim verisi ile kısmen benzerlik göstermektedir. Buna rağmen sınıflama tahmini Mnist veri setine göre düşüktür. Bunun başlıca nedeni ilgili veri setlerinin Mnist veri kümesine göre, imgelerin renk ve boyutsa özellikleri olarak söylenebilir. Tablo 3'de özellik çıkarıcı yöntemin düşük eğitim verisiyle son katmanındaki liner sınıflandırıcının düşük bir sınıflama skoru elde ettiği görülmüştür. Fakat eğitim verisinin belirli bir eşikte, yüksek bir sınıflama skoru elde ettiği ve eşik değerinden fazla eğitim verisi kullanıldığında sınıflandırıcı tarafından aşırı uyum sağladığı Mnist veri kümesi için gözlemlenmektedir. Bunun yanında kullandığımız kısmi özellik çıkarıcı yöntemin, özellik çıkarıcı yöntemine göre eğitim sırasından 2 ila 3 kat arası daha fazla zaman harcadığı da göz önüne alınması gereken önemli bir kriterdir.

Caltech101 veri kümesi, sınıf sayısının fazlalığı ve eğitim veri setinin kısıtlı olmasına rağmen ImageNet ile benzerliğinden dolayı doğruluk oranlarının Cifar10 veri kümesi ile paralel olduğu görülmüştür.

## 5. Sonuçlar

ESA' lar kullanılarak resimler üzerinde sınıflama yapmak literatürde yaygın kullanılan bir yöntemdir. ESA' larda kullanılan özellik çıkarıcı ve kısmi özellik çıkarıcı transfer öğrenme yöntemleri secimi hedef veri setindeki veri sayısı, kaynak veri setine olan benzerliğine göre değişmektedir. Öneğitim modeli hem özellik çıkarıcı hem de kısmi özellik çıkarıcı yaklaşımında standart olarak kullanılmaktadır. Alexnet' in ağırlıklarını araştırmacılar için paylaşması eğitimi günler süren önemli bir ESA'nın problemler için genelleştirmesini sağlamaktadır.

Bu makale çalışmasında eğitim verisinin durumuna göre hangi transfer yönteminin seçileceği ve aktarılan verilerin ESA için neler olacağı detaylı bir şekilde verilmiştir. Kısaca kaynak veri seti bilgilerle aynı özelliğe sahip bir hedef veri setine sahip ve elimizdeki hedef veri sayısı az ise kaynak ağı ilk katmanları özellik çıkarıcı olarak kullanılarak son katmanına liner bir sınıflandırıcı koymak gerektiği; veri sayısı fazla ise kaynak ağı hedef veri setiyle kısmi özellik çıkarıcının uygun olacağı sonucuna varılmıştır. Eğer kaynak veri seti ile hedef veri seti aynı özelliğe sahip değilse hedef veri seti sayısı az ise kaynak ağı sadece sınıflandırma katmanı destek vektör makinalarıyla yer değiştirmek gerekir. Ayrıca az sayıda veriyi sadece destek vektör makinası ile eğitmek yeterli olmaktadır. Eğer hedef veri kümesi fazla ise kaynak ağı bütün katmanlarının kısmi özellik çıkarma yaklaşımı ile tekrardan eğitilmesi, sınıflama doğruluk oranının yükselmesine neden olmaktadır.

## Kaynakça

- [1] Ethem Alpaydın, "Introduction to Machine Learning (Adaptive Computation and Machine Learning).", MIT Press, 2004.
- [2] C. M. Bishop, "Pattern Recognition and Machine Learning. Springer Science + Business Media", LLC, 2006.
- [3] C. Szegedy et al., "Going Deeper with Convolutions," in Computer Vision and Pattern Recognition (CVPR), 2014.

- [4] İnternet, “Yahoo Flickr.” [Online]. Available: <https://www.flickr.com/photos/tags/yahoo/>. [Accessed: 30-Dec-2017].
- [5] E. Cengil, A. Çınar, Z. Güler “A GPU-Based Convolutional Neural Network Approach for Image Classification,” in Artificial Intelligence and Data Processing Symposium (IDAP), 2017.
- [6] Şengör, Abdülkadir; Yaman, Akbulut; Yanhui, Guo; Varun, “Classification of amyotrophic lateral sclerosis disease based on convolutional neural network and reinforcement sample learning algorithm,” *Heal. Inf. Sci. Syst.*, 2017.
- [7] R. Girshick, J. Donahue, T. Darrell, J. Malik, and U. C. Berkeley, “Rich feature hierarchies for accurate object detection and semantic segmentation,” *Comput. Vis. Pattern Recognition. IEEE*, pp. 580–587, 2014.
- [8] İnternet, “Image.net,” 2018. [Online]. Available: <http://www.image-net.org>. [Accessed: 20-Dec-2017].
- [9] A. Krizhevsky and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” *Adv. Neural Inf. Process. Syst.*, vol. 1907–1105, pp. 1–9, 2012.
- [10] S. J. Pan and Q. Yang, “A Survey on Transfer Learning,” *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, 2010.
- [11] Barış Koçer, “Transfer Öğrenmede Yeni Yaklaşımlar”, Selçuk Üniversitesi, Doktora Tezi, 2012.
- [12] G. Konidaris and A. G. Barto, “Transfer in Reinforcement Learning via Shared Features,” *J. of Machine Learn. Res.*, vol. 13, pp. 1333–1371, 2012.
- [13] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, “How transferable are features in deep neural networks?,” *Adv. Neural Inf. Process. Syst.*, vol. 27, 2014.
- [14] Esra Mahsereci Karabulut, “Investigation Of Deep Learning Approaches For Biomedical Data Classification,” Çukurova University, Doktora Tezi, 2016.
- [15] Britz Denny, “Understanding Convolutional Neural Networks for NLP,” 2015. [Online]. Available: 20-Dec-2017 <http://www.wildml.com/2015/11/understanding-convolutional-neural-networks-for-nlp/>. [Accessed: 30-Dec-2017].
- [16] M. E. Akçay, Samet; Kundegorski, M. Devereux, and T. P. Breckon, “Transfer Learning Using Convolutional Neural Networks For Object Classification Within X-Ray Baggage Security Imagery,” in *Image Processing (ICIP)*, 2016.
- [17] Z. Huang, Z. Pan, and B. Lei, “Transfer Learning with Deep Convolutional Neural Network for SAR Target Classification with Limited Labeled Data,” pp. 1–21, 2017.
- [18] T. D. Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, “Caffe: Convolutional architecture for fast feature embedding,” in *ACMMM*, 2014, pp. 675–678.
- [19] Y. Bengio and P. Haffner, “Gradient-Based Learning Applied to Document Recognition,” vol. 86, no. 11, 1998.
- [20] A. Krizhevsky, “Learning multiple layers of features from tiny images.,” University of Toronto, 2009.
- [21] P. P. L. Fei-Fei, R. Fergus, “Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories,” *Comput. Vis. Image Underst.*, no. 106(1), pp. 59–70, 2007.
- [22] İnternet, “Transfer Learning.” [Online]. Available: <http://cs231n.github.io/transfer-learning/#tf>. [Accessed: 20-Dec-2018].