

## Derin Öğrenmede Kullanılan Veri Setleri ve Yazılım Kütüphaneleri

Özkan İNİK<sup>1\*</sup>, Erkan ÜLKER<sup>2</sup>

<sup>1</sup> Bilgisayar Mühendisliği Bölümü, Gaziosmanpaşa Üniversitesi, Tokat, Türkiye

<sup>2</sup> Bilgisayar Mühendisliği Bölümü, Selçuk Üniversitesi, Konya, Türkiye

\*ozkan.inik@gop.edu.tr

**Özet** – : Bu çalışmadaki amaç Derin Öğrenme mimarileri için kullanılan yazılım kütüphaneleri ve veri setlerinin incelenmesidir. Yapay zeka alanına farklı bir bakış açısı getiren Derin Öğrenme son yıllarda inanılmaz bir şekilde geniş bir alanda kullanılmaya başlanmıştır. Bilgisayarlı görüde Derin Öğrenme modelleri zengin yüksek çözünürlüklü fotoğraflar işlemektedir. Fotoğraf üzerindeki nesnelerin tanımlanması için geleneksel yöntemlerde olduğu gibi fotoğrafta kırılma veya özellik çıkarma gibi ön işlem aşaması yoktur. Benzer şekilde eski ağlar yalnızca iki tür nesneyi (veya bazı durumlarda tek nesnenin varlığı ile yokluğu) tanımlayabilirken, bu modern ağlar birçok farklı nesne kategorisini tanımlayabilir. Derin Öğrenmenin özellikle son yıllarda ortaya çıkmasının en büyük iki sebebi vardır. Bunların ilki günümüzde yeteri kadar verinin olması ve ikinci olarak bu verileri işleyecek donanımsal alt yapıların olmasıdır. Bu kapsamda kullanım amacına göre yazılım kütüphaneleri geliştirilmiş ve veri setleri oluşturulmuştur. Bu çalışmada toplam 10 farklı veri seti ve 6 farklı yazılım kütüphanesi incelenmiştir. Veri setleri sırasıyla MNIST Veri Seti, CIFAR10 veri seti, CIFAR 100 veri seti, STL-10 veri seti, The Street View House Numbers (SVHN) veri seti, Large Scale Visual Recognition Challenge(LSVRC) veri seti, Caltech 101 veri seti, Caltech 256, Labelled Faces in the Wild veri seti ve Pascal VOC veri setidir. Her bir veri setindeki görüntü sayısı sınıf sayısı v.b gibi bilgiler ayrıntılı bir şekilde anlatılmıştır. Yazılım kütüphaneleri ise sırasıyla Theano, Caffe, Torch, TensorFlow, Keras ve MatConvNet'tir. Bu yazılım kütüphanelerin birbirinden üstünlükleri ve dezavantajları ve kullandıkları platform ve performans değerleri detaylı verilmiştir. Yapılan çalışmada birden fazla Derin öğrenme kütüphanesi ve veri setleri hakkında bilgiler verilerle özellikle Derin Öğrenme alanında çalışmalar yapmak isteyen araştırmacılar için gerekli alt yapı bilgileri sunulmuştur.

**Anahtar Kelimeler** – : Caffe, CIFAR10, CNN, Derin Öğrenme, Keras, MNIST, Pascal VOC, Sınıflandırma, TensorFlow, Theano, Torch

## Data Sets and Software Libraries Used for Deep Learning

Özkan İNİK<sup>1\*</sup>, Erkan ÜLKER<sup>2</sup>

<sup>1</sup>Department of Computer Engineering, Gaziosmanpaşa University, Turkey

<sup>2</sup>Department of Computer Engineering, Selçuk University, Turkey

\*ozkan.inik@gop.edu.tr

**Abstract** – : The purpose of this study is to examine the software libraries and data sets used for Deep Learning architectures. Deep Learning brings a different perspective to the field of artificial intelligence. It has begun to be used an incredibly wide field in recent years. Deep Learning models process high resolution photos in computer vision. Unlike traditional machine learning method, there is no pre-processing phase, such as cropping or extracting features for identifying objects on a photo. Similarly, while old networks can only define two types of objects (or, in some cases, the absence and presence of a single object), these modern networks can describe many different categories of objects. There are two main reasons why Deep Learning has emerged, especially in recent years. The first of these is the training data as much as today. Secondly there is hardware to process this data. In this context according to its purpose many software libraries have been developed and data sets have been created. A total of 10 different data sets and 6 different software libraries were examined in this study. The data sets are the MNIST data Set, CIFAR10 data set, CIFAR 100 data set, STL-10 data set, Street View House Numbers (SVHN) data set, Large Scale Visual Recognition Challenge (LSVRC) data set, Caltech 101 data set, Caltech 256, Labeled Faces in the Wild data set and Pascal VOC data set respectively. The number of images in each data set, the number of classes, etc. is explained in detail. The software libraries are Theano, Caffe, Torch, TensorFlow, Keras and MatConvNet respectively. The advantages and disadvantages of these software libraries are explained. In addition, platform and performance values are given in detail. Necessary information (Deep Learning Libraries and data sets) are presented especially for researchers who want to work in the field of Deep Learning.

**Keywords** – : Caffe, Classification, CIFAR10, CNN, Deep Learning, Keras, MNIST, Pascal VOC, TensorFlow, Theano, Torch

---

## I. GİRİŞ

Derin Öğrenme son zamanlarda yapay zeka alanında çok iyi gelişmeler elde ederek birçok alanda kullanılmaya başlanmıştır. Özellikle bilgisayarlı görü[1-9] alanında yüksek başarılar elde etmiştir. Son yıllarda ortaya çıkmasının en büyük iki sebebi vardır. Bunların ilki günümüzde yeteri kadar verinin olması ve ikinci olarak bu verileri işleyecek donanımsal alt yapıların olmasıdır.

Günümüzde yeteri kadar verinin olmasında en büyük katkıyı kuşkusuz sosyal medya ve bazı büyük kuruluşlar(Google, Amazon v.s) yapmıştır. Dijital ortama alınan görüntülerin yapay zeka yöntemlerinde kullanılması için belli bir ön işleme tabi tutulup düzenlenmesi gerekmektedir. Bu amaçla Geçmişte gerek makine öğrenmesi yöntemleri için gerek günümüzde derin öğrenme modelleri için geniş yelpazede veri setleri oluşturulmuştur. Her probleme özgü farklı veri setleri tasarlanmıştır. Örneğin yüz tanımlama için binlerce insanın yüz görüntüsünden oluşturulan veri setleri veya nesne tanımlama için milyonlarca nesne görüntüsünden oluşan veri setleri bunlara örnek olur. Yapılan bu çalışmanın bir kısmında, temel olarak derin öğrenme modellerinde kullanılan veri setleri verilmiştir.

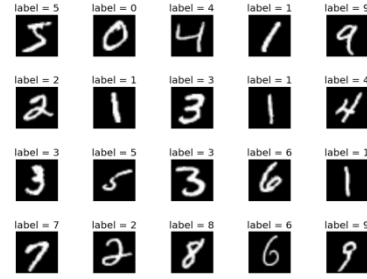
Derin öğrenme modellerinin donanımsal alt yapısını sağlayan Graphic Processing Unit(GPU) ve paralel çalışan Central Processing Unit(CPU) kullanılmaya başlanmasıyla bu donanımlarda modellerin eğitim ve test işlemini yapacak yazılım kütüphaneleri geliştirilmiştir. Derin öğrenme ile ilgili bütün yazılım kütüphaneleri ücretsiz ve açık kod kaynak olarak sunulmuştur. Hatta bir çok kütüphane geliştirilmeye açık ve sürekli farklı araştırmacılar vasıtasıyla güncellenmektedir. Oluşturulan bu yazılım kütüphanelerinin her birisinin kendine özgü üstün özellikleri ve bazen diğer kütüphanelere nazaran dezavantajları bulunmaktadır. Yapılan bu çalışmada diğer bir kısmında derin öğrenmede sıklıkla kullanılan yazılım kütüphaneleri verilmiştir. Bu kütüphanelerin belli çerçevede birbirlerine göre üstünlükleri sunulmuştur.

Bu çalışmanın yapısı şöyledir: Bölüm II’de derin öğrenme modellerinde kullanılan veri setleri anlatılmıştır. Bölüm III’te derin öğrenme modellerinde kullanılan yazılım kütüphaneleri aktarılmıştır. Son olarak Bölüm IV’te sonuç kısmı anlatılmıştır.

## II. DERİN ÖĞRENME MODELLERİNDE KULLANILAN VERİ SETLERİ

### A. MNIST Veri Seti

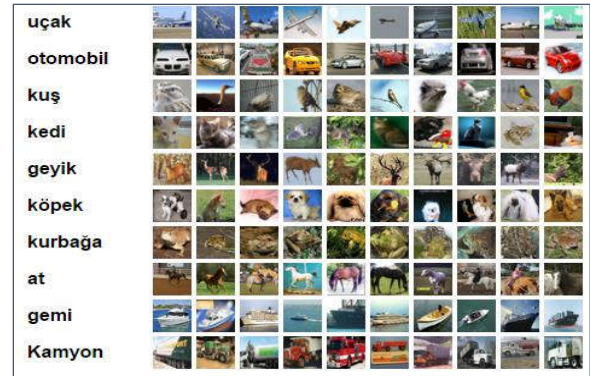
El yazısı rakamlarından oluşan bu veri seti[10] Yann LeCun, Corinna Cortes ve Christopher JC Burges tarafından oluşturulmuştur. Derin Öğrenmede temel veri seti olarak kabul edilir. Bir çok başlangıç seviyesindeki uygulamalar bu veri seti üzerinden yapılmaktadır. Veritabanı 60000 eğitim ve 10000 test görüntüsü içermektedir. Rakamlar, normal boyuta getirilmiş ve 28x28 büyüklükte siyah beyaz resimler olarak kaydedilmiştir. Veri setindeki her bir görüntü sayısal değeriyle etiketlenmiştir(Şekil 1).



Şekil 1. MNIST veri seti

### B. CIFAR10 Veri Seti

Bu veri seti[11] Alex Krizhevsky, Vinod Nair ve Geoffrey Hinton tarafından oluşturulmuştur. Veri setinde toplam 10 farklı sınıf bulunmakta ve her bir sınıf için 32x32 renkli 6000 adet resimden oluşmaktadır. Toplam görüntü âdeti 60000’dir. Bu resimlerin 50000 tanesi eğitim 10000 tanesi test görüntülerinden oluşmaktadır. Her bir sınıfa ait rastgele veri kümesi Şekil 2’de verilmiştir.



Şekil 2. CIFAR 10 veri setinde her bir sınıfa ait örnek resimler

### C. CIFAR100 Veri Seti

Bu veri kümesi[12], CIFAR 10 veri setinin geliştirilmiş halidir. Veri setinde 100 sınıf mevcut ve her sınıf 600 resim içermektedir. Sınıf başına 500 eğitim resmi ve 100 test görüntüsü vardır. CIFAR 100’deki 100 sınıf, 20 üst sınıflara ayrılmıştır. Her resim ait olduğu sınıf ve ait olduğu üst sınıf etiketi ile birlikte gelir. Veri setindeki Üst Sınıflar ve Üst sınıfları oluşturan alt sınıfların isimleri Tablo 1’de verilmiştir.

Tablo 1: CIFAR 100 veri setine ait sınıflar

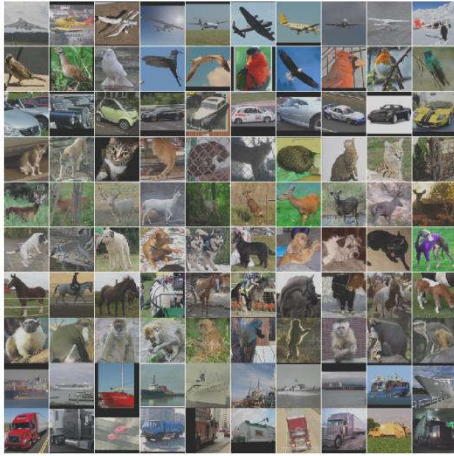
Üst Sınıf	Sınıf
aquatic mammals	beaver, dolphin, otter, seal, whale
fish	aquarium fish, flatfish, ray, shark, trout
flowers	orchids, poppies, roses, sunflowers, tulips
food containers	bottles, bowls, cans, cups, plates
fruit and vegetables	apples, mushrooms, oranges, pears, sweet peppers
household electrical devices	clock, computer keyboard, lamp, telephone, television
household furniture	bed, chair, couch, table, wardrobe
insects	bee, beetle, butterfly, caterpillar, cockroach
large carnivores	bear, leopard, lion, tiger, wolf
large man-made outdoor things	bridge, castle, house, road, skyscraper
large natural outdoor scenes	cloud, forest, mountain, plain, sea

large omnivores and herbivores	camel, cattle, chimpanzee, elephant, kangaroo
medium-sized mammals	fox, porcupine, possum, raccoon, skunk
non-insect invertebrates	crab, lobster, snail, spider, worm
people	baby, boy, girl, man, woman
reptiles	crocodile, dinosaur, lizard, snake, turtle
small mammals	hamster, mouse, rabbit, shrew, squirrel
trees	maple, oak, palm, pine, willow
vehicles 1	bicycle, bus, motorcycle, pickup truck, train
vehicles 2	lawn-mower, rocket, streetcar, tank, tractor

#### D. STL-10 veri seti

Stanford Üniversitesi tarafından oluşturulan bu veri seti[13] eğitimsiz özellik öğrenme yöntemleri, derin öğrenme ve kendi kendini öğreten öğrenme algoritmaları geliştirmek için bir resim tanıma veri kümesidir. Bazı farklılıkları yanında Cifar-10 veri setine benzemektedir. Özellikle, Eğitici öğrenmeden önce görüntü modellerini öğrenmek için çok geniş bir etiketsiz örnek seti sunulmaktadır. Veri setindeki resimlerin örnek bir çıktısı Şekil 3'te verilmiş olup genel özellikleri aşağıda verilmiştir.

- 10 sınıftan oluşmaktadır. Sınıflar: uçak, kuş, araba, kedi, geyik, köpek, at, maymun, gemi, kamyon.
- Resimler renkli ve 96x96 pikseldir.
- 500 eğitim resmi ve her sınıf başına 800 test görüntüsü mevcut.
- Eğitimsiz öğrenme için 100000 etiketlenmemiş görüntü vardır. Bu örnekler, benzer ama daha geniş bir görüntü dağılımı ile elde edilmiştir. Örneğin, etiketlenmiş kümesindeki araçlara ek olarak diğer tür hayvanlar (ayılar, tavşanlar vb.) ve araçlar (tren, otobüs vb.) içerir.



Şekil 3. STL-10 veri setine ait örnek resimler

#### E. The Street View House Numbers (SVHN) Veri Seti

SVHN, Stanford Üniversitesi tarafından Google Street View görüntülerinde ev numaralarından elde edilmiştir[14]. SVHN, veri ön işleme ve biçimlendirme üzerinde minimum gereksinim ile makine öğrenimi ve nesne tanıma algoritmaları geliştirmek için gerçek dünya görüntüleri veri kümesidir. MNIST'e benzeri olarak görülebilir (örneğin, resimler kırılmış küçük basamaklardan oluşur), ancak daha büyük etiketli verilere (600000'den fazla sayı görüntüleri) sahiptir. Veri seti 2 farklı formatta sunulmaktadır. Birinci

formatta karakter seviyesi sınırlayıcı kutuları bulunan orijinal görüntüleri(Şekil 4 sol) ve ikinci formatta(Şekil 4 sağ) ise MNIST veri seti benzeri 32x32 görüntüler tek bir karakter etrafında ortalanmıştır. SVHN veri setinin genel özellikleri aşağıdaki gibidir.

- 10 sınıf mevcuttur. 1 rakamı için "label 1", '9' için "label 9" ve '0' için "label10" etiketi kullanılmıştır.
- Eğitim için 73257, test için 26032 ve ekstra eğitim verisi olarak kullanılmak üzere biraz daha zor 531131 örnek verilmiştir.



Şekil 4. SVHN veri setine ait örnek resimler(Sol taraftaki format 1, sağ taraftaki format 2)

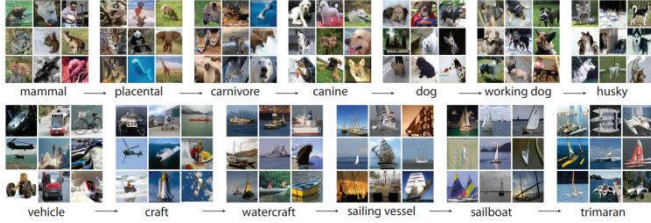
#### F. Large Scale Visual Recognition Challenge(LSVRC) Veri Seti

ImageNet yarışması kapsamında 2010 yılından yayınlanan veri setleridir. ImageNet veri seti olarak anılmaktadır[15]. Bu veri seti nesne tanımlama, sınıflandırma ve birçok yarışma kategorilerinde kullanılmaktadır. Ortalama 1.2 milyon resim ve 1000 tane sınıf mevcuttur. Bu veri seti hiyerarşik bir yapıda düzenlenmiştir. Her bir sınıfın bir üst sınıfı ve alt sınıfı mevcuttur. Örneğin köpek bir sınıf ise üst sınıfı hayvan ve alt sınıfı ise farklı köpek cinslerini içermektedir. Bu sebeple hiyerarşik derinliği diğer veri setlerine göre çok yüksektir. Örneğin diğer hiçbir veri setinde bulunmayan 147 farklı köpek cinsi katogorisi bu veri setinde mevcuttur. ImageNet veri setinde hiyerarjik yapısına göre kategoriler ve içerdikleri görüntü sayıları Tablo 2'de verilmiştir. Üst sınıf hayvanlar ve araçlar olmak üzere Şekil 5'te örnek bir gösterim sunulmuştur.

Tablo 2. ImageNet veri seti kategori sayıları

En üst sınıf	Alt sınıf sayısı	Her alt sınıf için ortalama görüntü sayısı	Toplam görüntü sayısı
amphibian	94	591	56K
Animal	3822	732	2799K
appliance	51	1164	59K
Bird	856	949	812K
Covering	946	819	774K
Device	2385	675	1610K
Fabric	262	690	181K
Fish	566	494	280K
Flower	462	735	339K
Food	1495	670	1001K
Fruit	309	607	188K
Fungus	303	453	137K
Furniture	187	1043	195K
Geological formation	151	838	127K
Invertebrate	728	573	417K
Mammal	1138	821	934K
Musical instrument	157	891	140K
Plant	1666	600	999K

Reptile	268	707	190K
Sport	166	1207	200K
Structure	1239	763	946K
Tool	316	551	174K
Tree	993	568	564K
Utensil	86	912	78K
Vegetable	176	764	135K
Vehicle	481	778	374K
Person	2035	468	952K



Şekil 5. ImageNet veri seti örnek hiyerarşik görüntüler

### G. Caltech 101 Veri Seti

Fei-Fei Li, Marco Andreetto ve Marc 'Aurelio Ranzato tarafından hazırlanan veri setidir[16]. 101 kategoriye ait nesnelerin resimlerinden oluşmaktadır. Her bir kategoride ortalama 40 ile 800 resim vardır. Çoğu kategoride yaklaşık 50 resim bulunur. Her resmin çözünürlüğü yaklaşık 300 x 200'dür.

### H. Caltech 256 Veri Seti

Bu veri seti Caltech 101 veri setinin daha gelişmiş halidir[17]. 256 kategoriye ait nesnelerin resimlerinden oluşmaktadır. Toplamda 3068 resim bulunmaktadır. Tablo 3'te Caltech 101 veri seti ile karşılaştırılmış hali verilmiştir.

Tablo 3. Caltech 101 ve Caltech 256 karşılaştırılması

Veri Seti	Kategori Sayısı	Toplam Resim Sayısı	Her kategori için			
			Min	Med	Mean	Max
Caltech 101	102	9144	31	59	90	800
Caltech 256	257	30608	80	100	119	827

### İ. Labelled Faces in the Wild Veri Seti

Bu veri seti yüz tanıma problemini incelemek için hazırlanmıştır. Veri seti internetten toplanan 13233 yüz görüntüsünden oluşmaktadır. Her görüntü resimdeki kişinin adıyla etiketlenmiştir. Yaklaşık 5749 kişinin yüz resmini içermektedir. Resimlerin boyutu 150x150'dir[18].

### J. Pascal VOC Veri Seti

2005 ile 2012 yılları arasında yapılan Pascal VOC Challenges yarışması için hazırlanan veri setidir[19]. Yıllara göre farklı veri setleri hazırlanmıştır. Bu veri setleri oluşturan kategoriler Tablo 3'te verilmiştir.

Tablo 3. Pascal VOC veri setine ait yıllara göre bilgiler

Yıl	Veri Seti Bilgileri
2005	4 sınıf: bisiklet, araba, motosiklet ve insan. Toplam 1578 görüntü ve bu görüntülerden 2209 etiketlenmiş obje içermektedir.
2006	10 sınıf: bisiklet, otobüs, araba, kedi, inek, köpek, at, motosiklet, insan, koyun. Toplam 2618 görüntü ve bu görüntülerden etiketlenmiş 4754 obje içermektedir.
2007	20 sınıf: ➤ İnsan: İnsan ➤ Hayvan: kuş, kedi, inek, köpek, at, koyun.

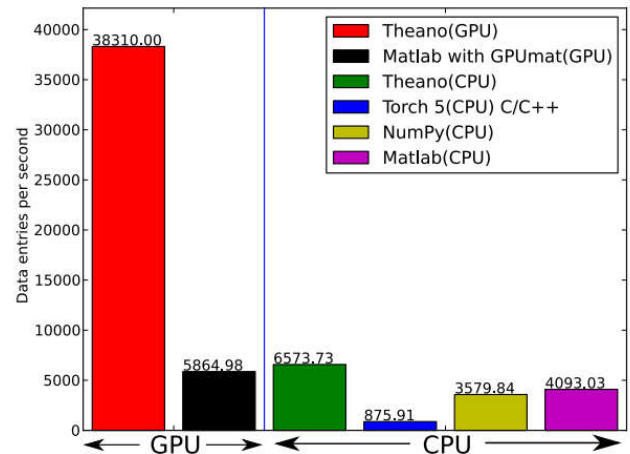
	➤ Araç: Uçak, bisiklet, bot, otobüs, araba, motosiklet, tren ➤ Ev içi: Şişe, sandalye, yemek masası, saksı, sofa, tv/monitor Toplam 9963 görüntü ve bu görüntülerden etiketlenmiş 24640 obje içermektedir.
2008	20 sınıf. Verilerin %50'si eğitim, %50'si test için bölünmüştür. Toplam 4340 görüntü ve bu görüntülerden etiketlenmiş 10363 obje içermektedir.
2009	20 sınıf. Toplam 7054 görüntü ve bu görüntülerden etiketlenmiş 17218 objeden oluşmaktadır. Ayrıca segmentasyon yapılmış 3211 görüntü içermektedir.
2010	20 sınıf. Toplam 10103 görüntü ve bu görüntülerden etiketlenmiş 23374 objeden oluşmaktadır. Ayrıca segmentasyon yapılmış 4203 görüntü içermektedir.
2011	20 sınıf. Toplam 11530 görüntü ve bu görüntülerden etiketlenmiş 27450 objeden oluşmaktadır. Ayrıca segmentasyon yapılmış 5034 görüntü içermektedir.
2012	20 sınıf. Toplam 11530 görüntü ve bu görüntülerden etiketlenmiş 27450 objeden oluşmaktadır. Ayrıca segmentasyon yapılmış 6929 görüntü içermektedir.

## III. DERİN ÖĞRENME MODELLERİNDE KULLANILAN YAZILIM KÜTÜPHANELERİ

Derin öğrenme modellerinin eğitimi için birçok açık kaynak kodlu yazılımlar geliştirilmiştir. Bu yazılımların bazıları geliştirilmeye açık ve sürekli farklı kullanıcılar tarafından geliştirilmektedir. Özellikle modellerin eğitimi esnasında geriye yayılımda türev hesaplarında ve tasarlanan modellerin mobil olarak çalıştırılması amacıyla yazılımlar geliştirilmiştir. Güncel olarak mevcut yazılım kütüphanelerinin bazıları aşağıda verilmiştir.

### A. Theano

Theano[20] Python dilinde yazılmış matematiksel ifadeler için bir derleyicidir. Montreal Üniversitesi LISA laboratuvarında geliştirilmiştir. Derin öğrenme mimarilerinde eğitim algoritmalarında türev alma işlemlerinde çok etkili bir şekilde kullanılır. Theano ile uygulanan yaygın makine öğrenme algoritmaları, CPU için derlendiğinde rakip alternatiflerden (C / C ++, NumPy / SciPy ve MATLAB) 1.6x ile 7.5x kat daha hızlı, GPU ile derlendiğinde ise 6.5x ile 44x daha hızlıdır[21]. Diğer kütüphanelerle karşılaştırılması Şekil 6'da verilmiştir. Birden fazla GPU desteklememesi dezavantajlarındandır.



Şekil 6. Theano kütüphanesinin diğer kütüphanelerle karşılaştırılması[21]

### B. Caffe

Caffe Berkeley Vision and Learning Center (BVLC) ve kullanıcı topluluğu tarafından geliştirilmiş ve sürekli güncellenmektedir. Yangqing JIA'nin[22] doktora sırasında ortaya koymuş olduğu bir projedir. Derin öğrenme modellerinin eğitimi için hem CPU hemde GPU ile eğitim yapılmasına olanak sağlamaktadır. C++ programlama dilinde geliştirilmiştir. Python ve Matlab arabirimlerini içerir. Araştırma deneyleri ve endüstride kullanım için Caffe ileri düzeyde kullanılmaktadır. Caffe, tek bir NVIDIA K40 GPU ile günlük 60M'tan fazla görüntü işleyebilir. Çıkarım hızı 1 ms / görüntü, iken eğitim için 4 ms / görüntü hızında çalışmaktadır. Caffe geliştiricileri mevcut en hızlı derin öğrenme modellerinin uygulaması Caffe ile yapıldığına inanmaktadır. Caffe yazılım kütüphanesine [23] online erişim adresinden erişilebilir.

### C. Torch

Torch, Lua kullanılarak oluşturulmuş ve Lua (JIT) derleyicisinde çalışan bilimsel bir yazılım kütüphanesidir [24]. Torch'un yazılım kütüphanesinin temel felsefesi, yapılacak işlemleri son derece basit hale getirmek ve bilimsel algoritmaların oluşturma konusunda maksimum esnekliğe ve hıza sahip olmaktır. Torch, makine öğrenimi, bilgisayarlı görme, sinyal işleme ve paralel işleme konularında geniş bir paket sunmaktadır. Torch'ta kullanımı kolay popüler sinir ağları ve optimizasyon kütüphaneleri bulunmaktadır. Derin Ağlar için esnek grafikler oluşturabilir ve CPU'lar ve GPU'lar üzerinde etkili bir şekilde paralel hale getirilebilir. Torch'un temel özellikleri aşağıda sunulmuştur.

- Modüler bir yapıda oluşturulmuş ve bu modüler parçaların birleştirilmesi çok kolay.
- GPU üzerinde çalıştırılacak katmanların yazılmasını kullanıcı tarafından kolaylıkla yapılabilir.
- Lua ile geliştirildiğinde kod okuması kolay
- Önceden eğitilmiş birçok derin öğrenme modelini barındırmaktadır.

Torch; Facebook, Google, Twitter gibi birçok şirket ve araştırma laboratuvarında kullanılmaktadır. Online erişim adresi[25]'dir.

### D. Tensor Flow

Google'ın makine öğrenimi araştırma grubu tarafından geliştirilen TensorFlow, veri akış grafikleri kullanarak sayısal hesaplama için kullanılan açık kaynaklı bir yazılım kütüphanesidir. Esnek mimari ile hesaplamayı tek bir API ile bir masaüstü, sunucu veya mobil cihazdaki bir veya daha fazla CPU'ya veya GPU'ya dağıtımına olanak tanır. TensorFlow başlangıçta, makine öğrenimi ve derin sinir ağları araştırması yürütmek amacıyla geliştirilmişti fakat şuan sistem diğer alanlarda geniş bir yelpazede kullanılmaktadır. Online erişim adresi [26]'dir.

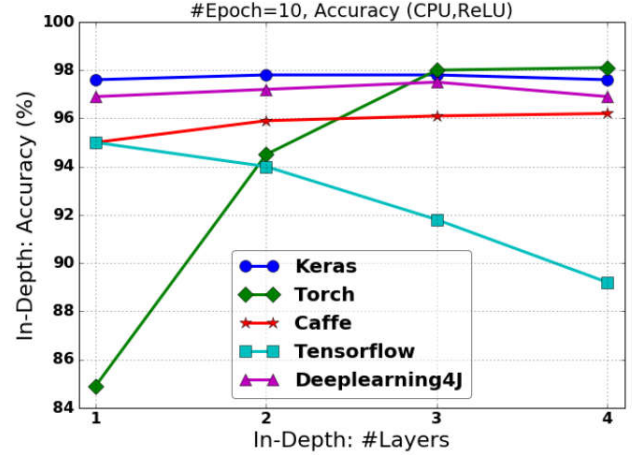
### E. Keras

Keras[27], üst düzey sinir ağları kütüphanesidir. Theano veya Tensorflow kütüphaneleri tabanda Keras kullanmıştır. Python dili ile yazılmıştır. Hızlı sonuçlar vermeyi odak noktası kabul ederek geliştirilmiştir. Temel anahtar kelimesi fikirden sonuca mümkün olan en az gecikmeyle gidebilmektir. Keras'ın temel üstünlükleri aşağıda belirtilmiştir.

- Kolay ve hızlı prototipleme sağlar.

- Hem konvolüsyonel ağları hem de tekrar eden ağları ve ikisinin kombinasyonlarını destekler.
- Raslantısal bağlantı planlarını destekler (çoklu giriş ve çoklu çıkış eğitimi dahil).
- CPU ve GPU üzerinde kusursuz çalışır.

Keras'ın sınıflandırma doğruluğunda diğer kütüphanelerle karşılaştırılması Şekil 7'de verilmiştir. Bu karşılaştırmada veri seti olarak MNIST ve ReLu fonksiyonu kullanılmıştır.



Şekil 7. Keras'ın sınıflandırma doğruluğunun diğer kütüphanelerle karşılaştırılması[28]

### F. MatConvNet

MatConvNet[29] derin öğrenme modellerinin eğitilmesi için Matlab tabanlı bir yazılım kütüphanesidir. Bu kütüphane derin ağların tasarımında kullanılan katmanların blok halinde tasarlanmasına olanak sunar. Örneğin konvolüsyon katmanı için `vl_nconv()`, pooling katmanı için `vl_nnpool()`, aktivasyon katman için `vl_nnsigmoid()`, sınıflandırma katmanı için `vl_nnsoftmax()` gibi fonksiyonlar tasarlanmıştır. Blok halindeki tasarlanan bu fonksiyonlar sayesinde, yeni bir derin öğrenme modelinin tasarlanması daha esnek, anlaşılır ve kolay olmaktadır. Hem GPU hem de CPU üzerinde verimli hesaplama yapılmasına olanak verir. Bu sayede oluşturulan bir ağın çok büyük veri setleri üzerinde çalışmasına olanak sunar. MatConvNet şu an her ne kadar MATLAB'a dayanıyor olsa da, bu kütüphane MATLAB kodu ile C++ ve CUDA çekirdeği arasında belli bir ayrımla geliştirilmektedir. Bu nedenle gelecekte kütüphane, MATLAB'dan bağımsız olarak derin ağların işlenmesine izin verecek şekilde genişletilebilir. MatConvNet kütüphanesinin ImageNet veri seti üzerinde eğitildiğinde aldığı süreler tek GPU ile Şekil 8 ve birden fazla GPU ile Şekil 9'da verilmiştir.

model	batch sz.	CPU	GPU	CuDNN
AlexNet	256	22.1	192.4	264.1
VGG-F	256	21.4	211.4	289.7
VGG-M	128	7.8	116.5	136.6
VGG-S	128	7.4	96.2	110.1
VGG-VD-16	24	1.7	18.4	20.0
VGG-VD-19	24	1.5	15.7	16.5

Şekil 8. ImageNet veri seti üzerinde eğitimde görüntü başına aldığı saniye.

num GPUs	1	2	3	4
VGG-VD-16 speed	20.0	22.20	38.18	44.8

Şekil 9. Birden fazla GPU ile eğitim süresi

## SONUÇ

Derin öğrenmede kullanılan veri setleri hakkında bilgiler sunulmuştur. Her bir veri setindeki görüntü sayısı sınıf sayısı v.b gibi bilgiler ayrıntılı bir şekilde anlatılmıştır. Ayrıca derin öğrenmede kullanılan yazılım kütüphaneleri hakkında bilgiler verilmiştir. Bu yazılım kütüphanelerin birbirinden üstünlükleri ve dezavantajları ve kullandıkları platform ve performans değerleri verilmiştir. Yapılan çalışma ile özellikle derin öğrenme alanında çalışmalar yapmak isteyen araştırmacılar için gerekli bilgiler sunulmuştur.

## REFERANSLAR

- [1] S. Ji, W. Xu, M. Yang, and K. Yu, "3D convolutional neural networks for human action recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, pp. 221-231, 2013.
- [2] K. Simonyan and A. Zisserman, "Two-stream convolutional networks for action recognition in videos," in *Advances in neural information processing systems*, 2014, pp. 568-576.
- [3] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei, "Large-scale video classification with convolutional neural networks," in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, 2014, pp. 1725-1732.
- [4] Y. Taigman, M. Yang, M. A. Ranzato, and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1701-1708.
- [5] S. Antol, A. Agrawal, J. Lu, M. Mitchell, D. Batra, C. Lawrence Zitnick, et al., "Vqa: Visual question answering," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 2425-2433.
- [6] A. Karpathy and L. Fei-Fei, "Deep visual-semantic alignments for generating image descriptions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3128-3137.
- [7] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3431-3440.
- [8] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [9] Y. Sun, X. Wang, and X. Tang, "Deep learning face representation from predicting 10,000 classes," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1891-1898.
- [10] Y. LeCun, C. Cortes, and C. J. Burges, "Mnist handwritten digit database. AT&T Labs," ed, 2010.
- [11] A. Krizhevsky, V. Nair, and G. Hinton, "The CIFAR-10 dataset," online: <http://www.cs.toronto.edu/kriz/cifar.html>, 2014.
- [12] A. Krizhevsky, V. Nair, and G. Hinton, "The CIFAR-100 dataset," online: <http://www.cs.toronto.edu/kriz/cifar.html>, 2014.
- [13] A. Coates, A. Ng, and H. Lee, "An analysis of single-layer networks in unsupervised feature learning," in *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, 2011, pp. 215-223 URL:<http://cs.stanford.edu/~acoates/stl10>
- [14] Y. Netzer, T. Wang, A. Coates, A. Bissacco, B. Wu, and A. Y. Ng, "Reading digits in natural images with unsupervised feature learning," in *NIPS workshop on deep learning and unsupervised feature learning*, 2011, p. 5 URL:<http://ufldl.stanford.edu/housenumbers>.
- [15] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, 2009, pp. 248-255.
- [16] L. Fei-Fei, R. Fergus, and P. Perona, "Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories," *Computer vision and Image understanding*, vol. 106, pp. 59-70 URL:[http://www.vision.caltech.edu/Image\\_Datasets/Caltech101/](http://www.vision.caltech.edu/Image_Datasets/Caltech101/), 2007.
- [17] G. H. Griffin and A. Perona, "P. the caltech-256," *Caltech Technical Report, Tech. Rep.*, 2012 URL:[http://www.vision.caltech.edu/Image\\_Datasets/Caltech256/](http://www.vision.caltech.edu/Image_Datasets/Caltech256/).
- [18] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," Technical Report 07-49, University of Massachusetts, Amherst 2007 URL <http://vis-www.cs.umass.edu/lfw/index.html>.
- [19] "Pascal VOC Challenges Datasets(2005-2010)," URL <http://host.robots.ox.ac.uk/pascal/VOC/>.
- [20] J. Bergstra, O. Breuleux, P. Lamblin, R. Pascanu, O. Delalleau, G. Desjardins, et al., "Theano: Deep learning on gpus with python," 2011 URL: <http://deeplearning.net/software/theano/index.html>.
- [21] J. Bergstra, O. Breuleux, F. Bastien, P. Lamblin, R. Pascanu, G. Desjardins, et al., "Theano: A CPU and GPU math compiler in Python," in *Proc. 9th Python in Science Conf*, 2010, pp. 1-7.
- [22] Y. JiA, "<http://daggerfs.com/>."
- [23] Caffe, "<http://caffe.berkeleyvision.org/>".
- [24] R. Collobert, K. Kavukcuoglu, and C. Farabet, "Torch7: A matlab-like environment for machine learning," in *BigLearn, NIPS Workshop*, 2011.
- [25] Torch, "<http://torch.ch/>".
- [26] T. Flow, <https://www.tensorflow.org/>.
- [27] F. Chollet, "Keras (2015)," URL <http://keras.io>, 2017.
- [28] V. Kovalev, A. Kalinovskiy, and S. Kovalev, "Deep Learning with Theano, Torch, Caffe, Tensorflow, and Deeplearning4J: Which One is the Best in Speed and Accuracy?," 2016.
- [29] A. Vedaldi and K. Lenc, "Matconvnet: Convolutional neural networks for matlab," in *Proceedings of the 23rd ACM international conference on Multimedia*, 2015 URL:<http://www.vlfeat.org/matconvnet/>, pp. 689-692.