

NESNE ALGILAMA VE SINIFLANDIRM

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

Algoritma

Ön Hazırlık

Sistemi Tanı
(CNN Ağ Yapısı)

Yapılan Çalışmaları
İncele

Veri Setlerini Tanı

Mimarileri Tanı (Ağ
Modelleri)

Kütüphaneleri Tanı

Uygulama

Problem Belirle

Veri Topla

Model Belirle

Alt Yapı Oluştur

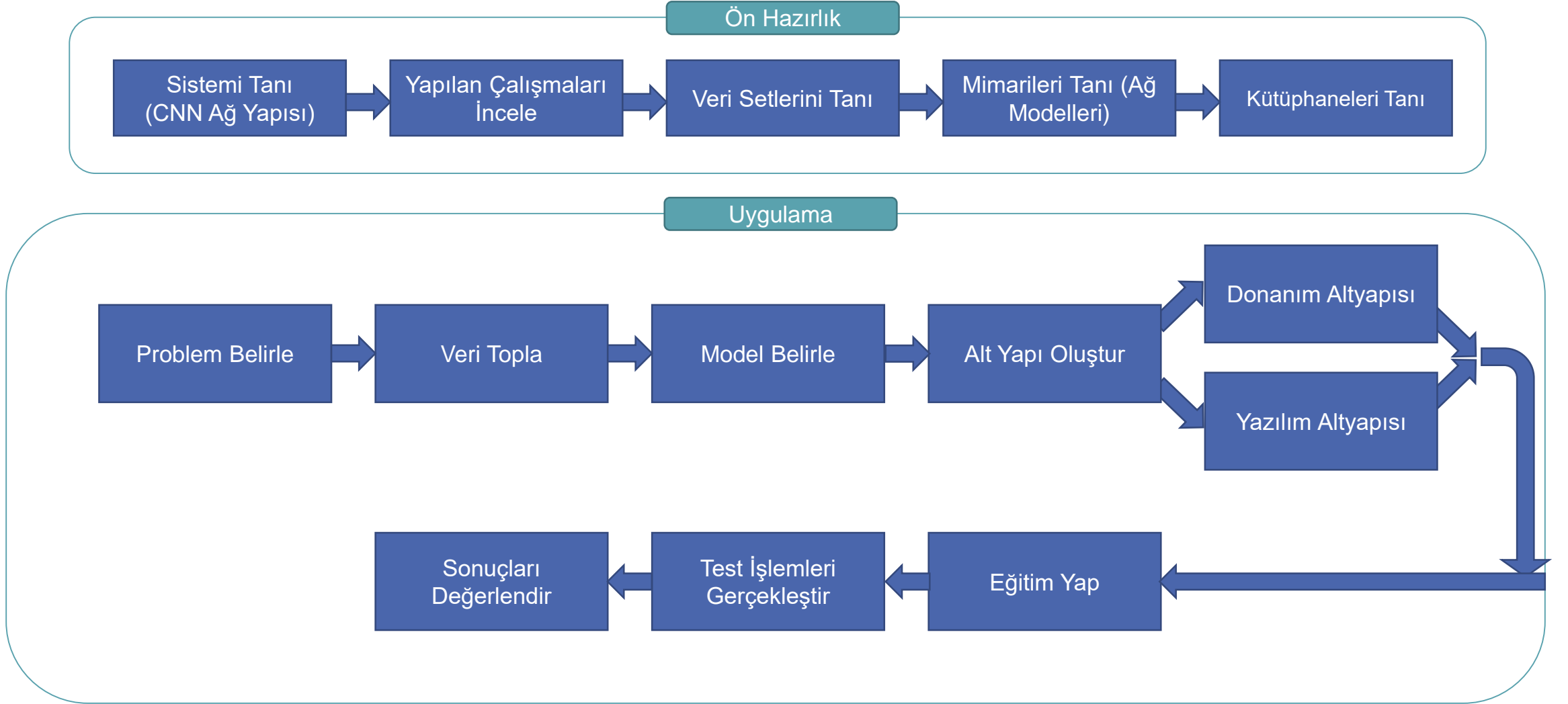
Donanım Altyapısı

Yazılım Altyapısı

Sonuçları
Değerlendir

Test İşlemleri
Gerçekleştir

Eğitim Yap

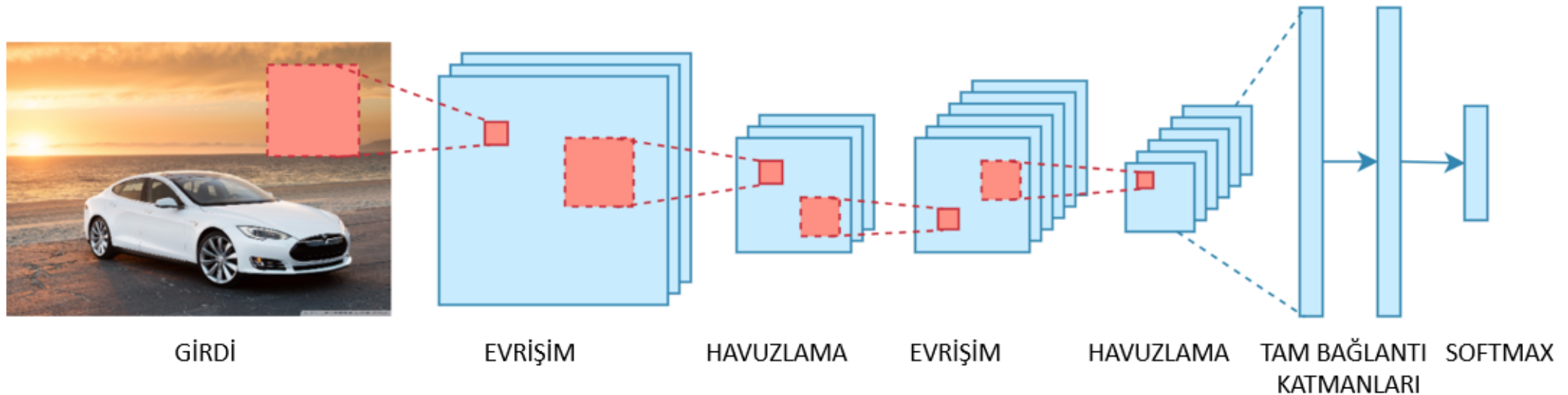


Evrişimli Sinir Ağları

Derin öğrenme ya da derin sinir ağı kavramı çok katmanlı yapay sinir ağlarını ifade etmektedir. En popüler derin sinir ağlarından biri evrişimli sinir ağları olarak görülmektedir.

Evrişimli sinir ağları ile yapay sinir ağları arasındaki dikkate değer tek fark evrişimli sinir ağlarının öncelikle görüntü işlemede örüntü tanıma alanında kullanılmasıdır.

Evrişimli sinir ağları evrişim, havuzlama ve tam bağlantılı olmak üzere üç tür katmandan oluşmaktadır.



Evrişimli Katmanlar

Evrişimli Katman, CNN algoritmalarında görüntü üzerinde işlem yapan ilk katmandır. Bu katman için CNN'nin temel yapı taşı denilmesi yanlış olmaz. Görüntüler, içlerinde belirli değerler taşıyan piksellerden oluşan matrislerdir. Bu katmanda gerçekleşen olay görüntü matrisleri üzerinde dolaşan, görüntü matrislerinden daha küçük boyutlu filtre matrisleri ile görüntülerin belirli özelliklerini elde etmektir. Bu özellikler düşük veya yüksek seviyeli olabilirler. Örnek vermek gerekirse kullandığımız filtre görüntü üzerindeki kenarları algılayacak bir filtre olabilir.

7	2	3	3	8
4	5	3	8	4
3	3	2	8	4
2	8	7	2	7
5	4	4	5	4

*

1	0	-1
1	0	-1
1	0	-1

=

6		

$$7 \times 1 + 4 \times 1 + 3 \times 1 + 2 \times 0 + 5 \times 0 + 3 \times 0 + 3 \times -1 + 3 \times -1 + 2 \times -1 = 6$$

Evrişimli Katmanlar

Görüntülerimiz tek bir özellikten oluşmadığı için genel olarak birden fazla filtre uygulanması gerekmektedir. Buradan CNN algoritmalarında çok sayıda evrişimli katman bulunduğu sonucuna ulaşılabilir. Ek olarak filtre işlemleri uygularken bilinmesi gereken iki terim daha vardır. Bu terimler Adım (Stride) ve Dolgu (Padding) terimleridir. Bunları kısaca açıklayacak olursak:

```
[convolutional]
batch_normalize=1
filters=64
size=1
stride=1
pad=1
activation=leaky
```

```
[convolutional]
batch_normalize=1
filters=128
size=3
stride=1
pad=1
activation=leaky
```

Evrişimli Katmanlar

Adım değeri CNN modellerinde parametre olarak değiştirilebilen bir değerdir. Bu değer filtrenin ana görsel üzerinde kaç piksel boyunca kayacağını belirler. Eğer Adım değeri büyük bir değer olursa filtre görüntü üzerinde adım değerine göre kayacağı için ortaya çıkacak olan öznelik haritası daha küçük boyutlu olacaktır.

Bir görsele filtre uygulandığında boyutlardan dolayı çıktı orijinal görselden daha küçük olur. Bunu önlemek için kullanabileceğimiz yöntem ise padding yani dolgulamadır. Dalgulama işleminde görsele adete bir çerçeve olacakmış gibi dört taraftan da sıfırlar eklenir. Filtrenin boyutuna göre bu sıfır eklenen katmanlar artırılabilir.

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	Görüntü						0	0
0	0	Görüntü						0	0
0	0	Görüntü						0	0
0	0	Görüntü						0	0
0	0	Görüntü						0	0
0	0	Görüntü						0	0
0	0	Görüntü						0	0
0	0	Görüntü						0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

Padding

Havuzlama Katmanı (Pooling Layers)

Havuzlama katmanının amacı, görüntüler çok büyük olduğunda parametre sayısını azaltarak görüntüyü küçültmektir ve aşırı öğrenmeyi kontrol etmektir. Bu işlem, her bir haritanın boyutsallığını azaltır ama önemli bilgileri korur. 3 tür havuzlama yöntemi vardır. Maksimum, ortalama ve minimum havuzlama yöntemidir. CNN ağlarında genellikle maksimum havuzlama kullanılır.

1	1	2	4
5	6	7	8
3	2	1	0
1	2	3	4

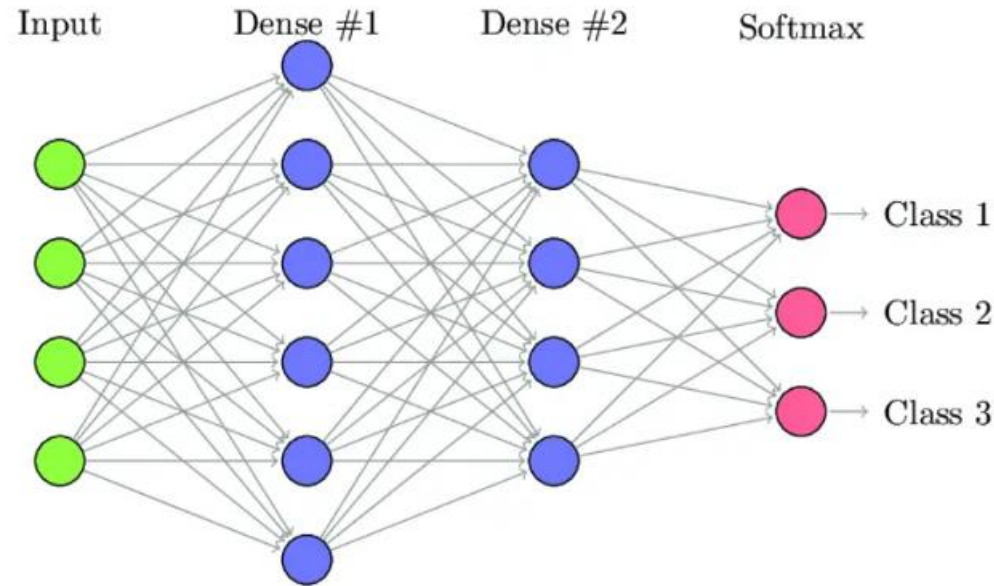
Feature map



Pooled
Feature map

Tam Baęlantı Katmanları (Fully Connected Layers)

Evriřimli Sinir Aęlarının son adımı bu katmanda gerekleřir. Tam Baęlantı Katmanları, mimarinin bir katmandaki bütn dęm ve nronlarının bir sonraki katmanlara baęlandığı bir yapay sinir aę trdr. Bu aę tr hesaplama aısından karmařık olmasının yanı sıra ařırı yklenmeye meyillidir.



Yapılan Çalışmalar

Ref No	Problem Description	Used DataSet	Input Data	DL Model	Architecture	Algorithm	Experimental Results	Year
[XX]		Images of 500	CNN	Darknet	YOLO		2021
[XY]		The study region	CNN	ResNet	FasterCNN		2020

Veri Setleri

IMAGENET

ImageNet veri seti 27 üst kategoriden ve bu kategorilerle ilişkili olarak en az 20bin alt kategoriden oluşan yaklaşık 14 milyon tam çözünürlüklü görüntü barındıran görüntü veri kümesidir.

Uçak



Otomobil



Kuş



Kedi



Geyik



Köpek



Kaplumbağa



At



Gemi



Kamyon



Veri Setleri



Visual Object Classes Challenge 2009 (VOC2009)



[click on an image to see the annotation]

Pascal VOC veri seti 2005 –2012 yılları arasında geliştirilerek yayınlanmış her yıl aynı isimle düzenlenen yarışmada kullanılan ve en son haliyle toplam 20 nesne sınıfı olmak üzere 11.530 görüntüden ve 27.450 nesne etiketi verisinden oluşan görsel veri setidir.

Veri Setleri



MS COCO veri kümesi, 82 nesne kategorisi 5 binden fazla etiketli içeriğe sahip toplam 91 ortak nesne kategorisi içerir. Toplamda veri kümesinde 328 bin görüntüde 2 milyon 500 bin etiketli örneği bulunmaktadır. ImageNet veri setinin aksine, COCO daha az kategoriye ve kategori başına daha fazla örneğe sahiptir.

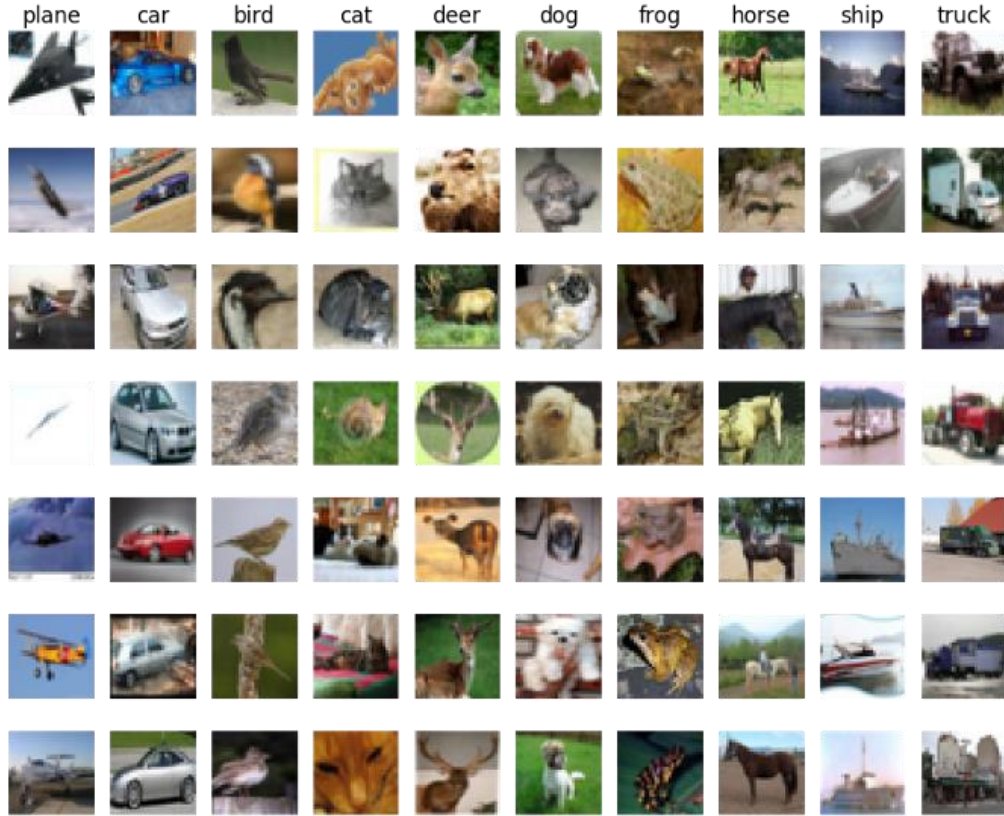
Veri Setleri

MNIST

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6
7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7
8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9

MNIST, el yazısıyla yazılmış rakamlardan oluşan ve 60 bin eğitim, 10 bin test olmak üzere toplamda 70 bin görüntüden oluşan bir veri setidir. Görüntüler 28x28 piksel boyutlarında olup gri seviyelidir.

Veri Setleri

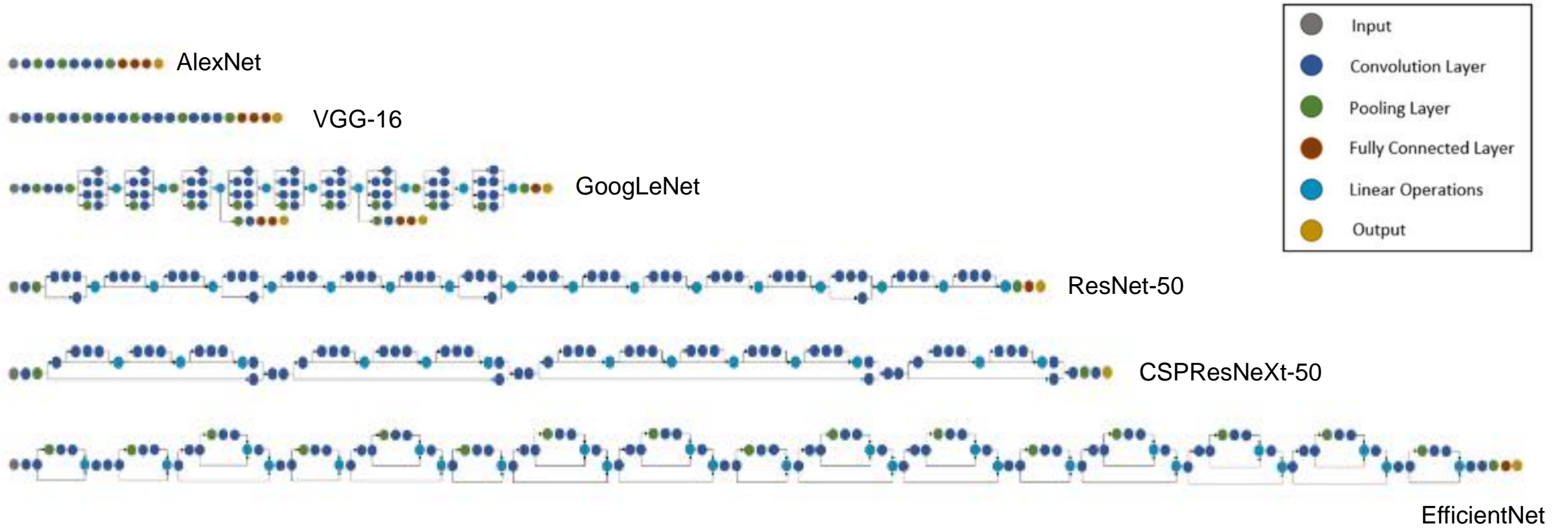


CIFAR, 80 milyon küçük boyutlu görüntü içeren veri setidir. İki farklı versiyonu olan veri setinde;

CIFAR-10 veri seti her sınıfta 6000 görüntü bulunan toplam 10 sınıftan oluşmaktadır. Görüntü boyutları 32x32 piksel boyutlarında olup 50 bin görüntü eğitim, 10 bin görüntü test kümesi olarak ayrılmıştır

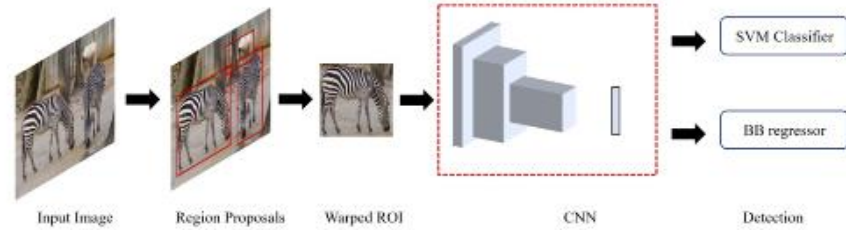
CIFAR-100 veri seti ise her sınıfta 600 görüntü bulunan toplam 100 sınıftan oluşmaktadır

Nesne Algılama Mimarileri

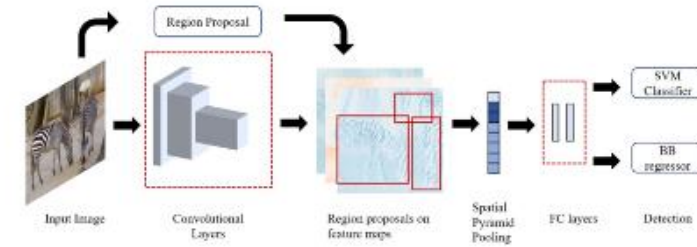


Nesne Algılama Algoritmaları: Two-stage detectors

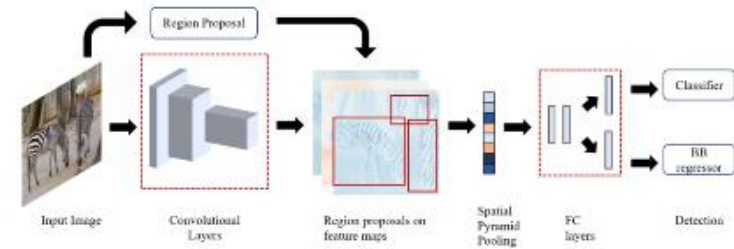
RCNN



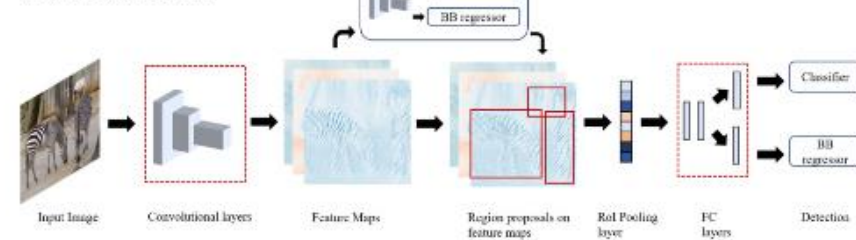
SPP-Net



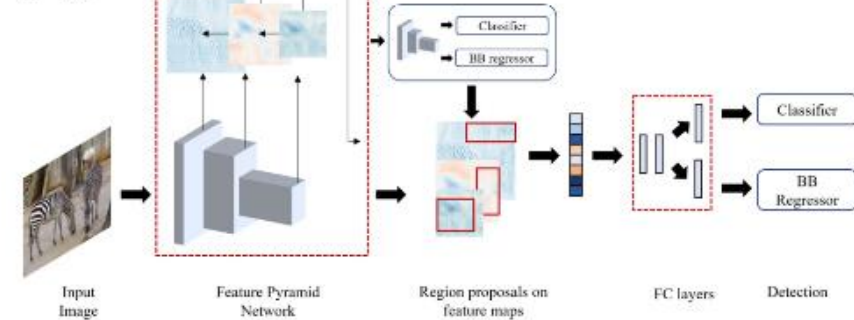
Fast RCNN



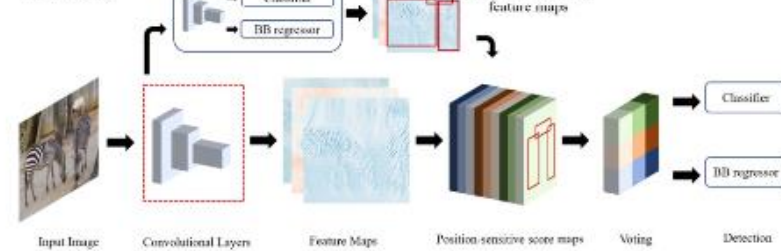
Faster RCNN



FPN

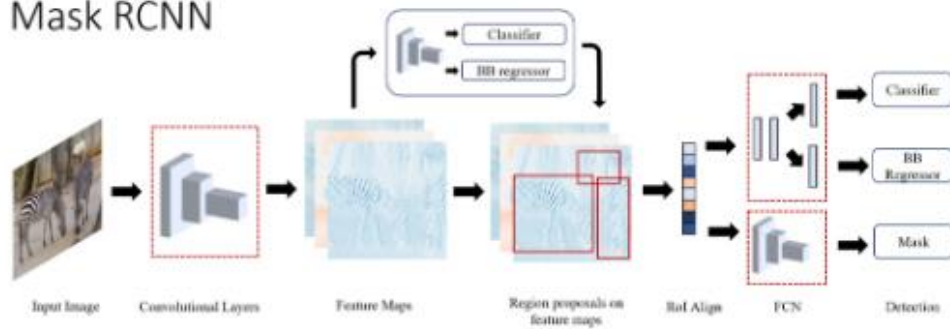


R-FCN

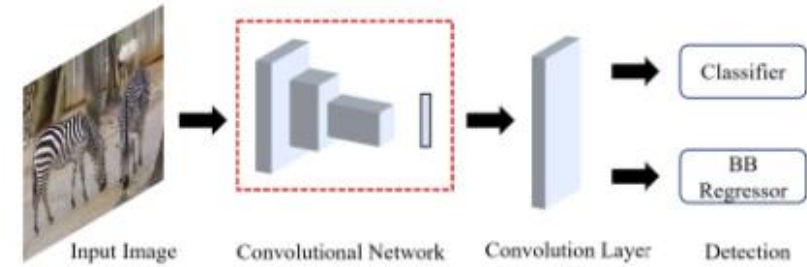


Nesne Algılama Algoritmaları: Single stage detectors

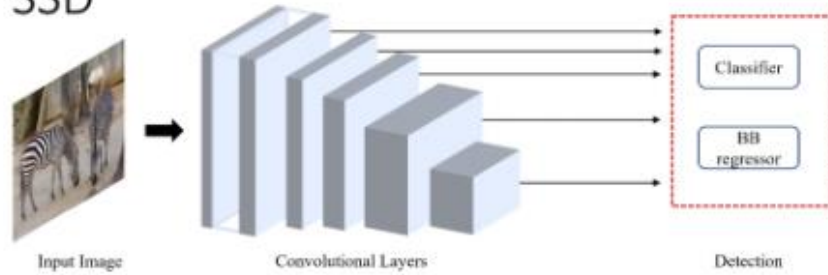
Mask RCNN



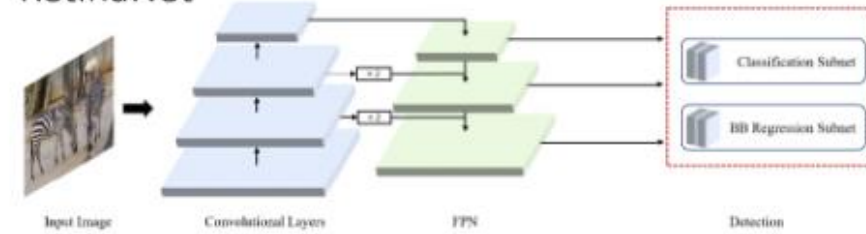
YOLO



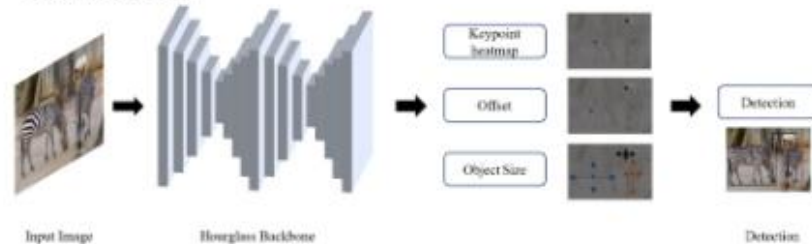
SSD



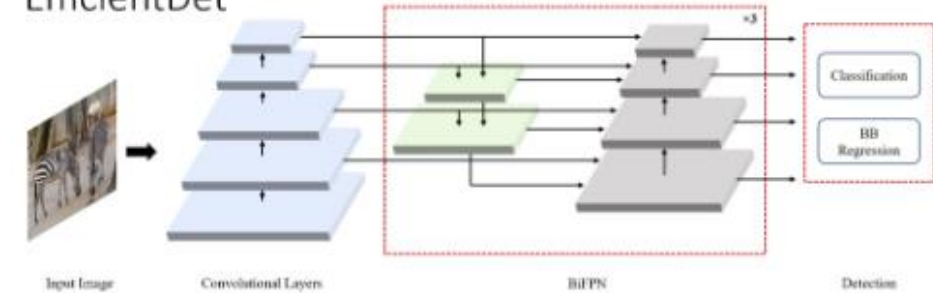
RetinaNet



CenterNet

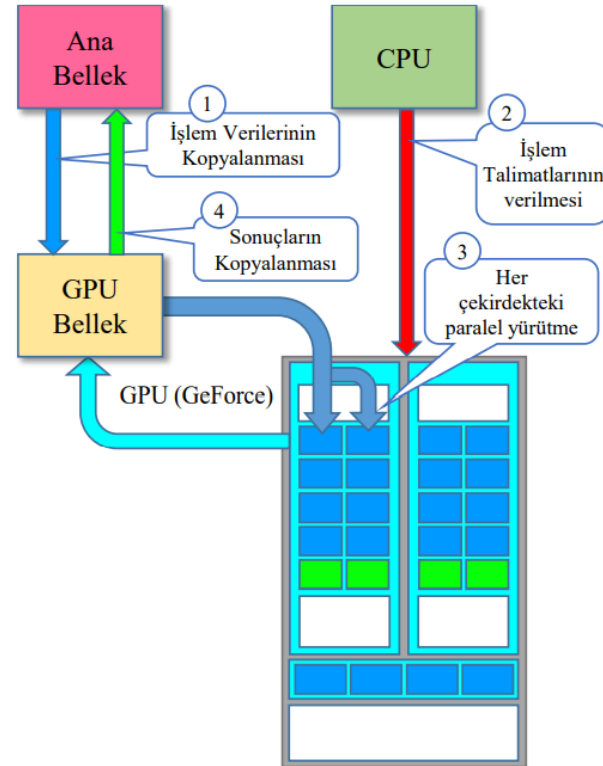


EfficientDet



Kütüphaneler

CUDA : CUDA (Compute Unified Device Architecture), Nvidia tarafından geliştirilen bir paralel hesaplama platformu ve uygulama programlama arayüzü (API) modelidir. C, C ++ ve Python gibi programlama dilleriyle çalışacak şekilde tasarlanmıştır.



Kütüphaneler

cuDNN: NVIDIA CUDA® Derin Sinir Ağı kütüphanesi (cuDNN), derin sinir ağlarının temelleri için GPU hızlandırmalı bir kütüphanedir. cuDNN, ileri ve geri konvülsyon, havuzlama, normalleştirme ve aktivasyon katmanları gibi standart görevler için yüksek düzeyde ayarlanmış uygulamalar sağlar.

Tensorflow: Başlangıçta Google tarafından Google Beyin projesinin bir parçası olarak geliştirilen Tensorflow, 2015'in sonlarında açık kaynak haline getirilmiştir. Tensorflow bir Python kütüphanesidir.

OpenCV: OpenCV (Açık Kaynak Kodlu Bilgisayarla Görme Kütüphanesi), açık kaynak kodlu bir bilgisayarla görme ve makine öğrenmesi yazılımı kütüphanesidir. OpenCV, bilgisayarla görme uygulamaları için ortak bir altyapı sağlamak ve ticari ürünlerde makine algısının kullanımını hızlandırmak için oluşturulmuştur.

Kütüphaneler

PyTorch: Torch kütüphanesine dayanan açık kaynaklı bir makine öğrenme kütüphanesidir, bilgisayarla görme ve doğal dil işleme gibi uygulamalar için kullanılır. Öncelikle Facebook'un AI Araştırma laboratuvarı (FAIR) tarafından geliştirilmiştir.

NumPy: NumPy Python programlama dili için oluşturulan bir kütüphanedir. Büyük, çok boyutlu diziler ve matrislerin yanı sıra geniş bir yüksek seviye matematiksel fonksiyon koleksiyonuna bu diziler üzerinde çalışabilmeleri için ek destek verir.

Matplotlib: Matplotlib, Python programlama dili ve sayısal matematik uzantısı NumPy için bir grafik kütüphanesidir. Grafikleri Tkinter, wxPython, Qt veya GTK + gibi genel amaçlı GUI araç takımlarını kullanarak uygulamalara gömmek için nesne yönelimli bir API sağlar.