



SAKARYA
ÜNİVERSİTESİ

BSM 310 YAPAY ZEKA

CEMİL ÖZ, İSMAİL ÖZTEL

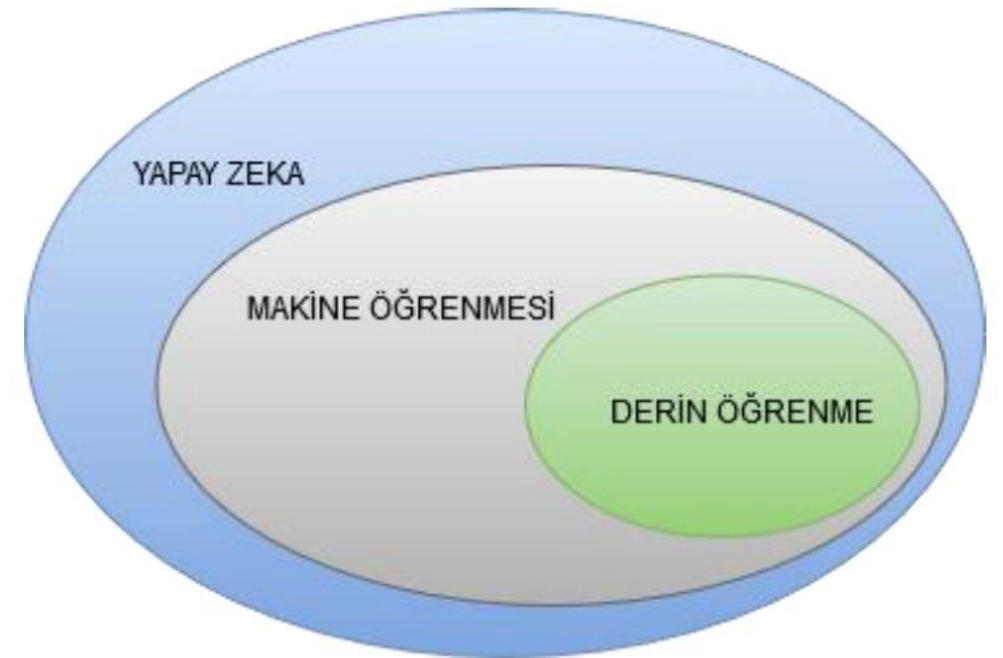
~ EVRİŞİMSEL SİNİR AĞLARI ~

KONULAR

- Derin öğrenme tanımı
- Sayısal görüntü işlemeye giriş
- Derin ağlarda veri kümesi
- Evrişimsel sinir ağları
- Konvolüsyon, Havuzlama, ReLU katmanları
- Derin ağ eğitiminde karşılaşılabilecek sorunlar
- Transfer öğrenme
- Veri çoklama

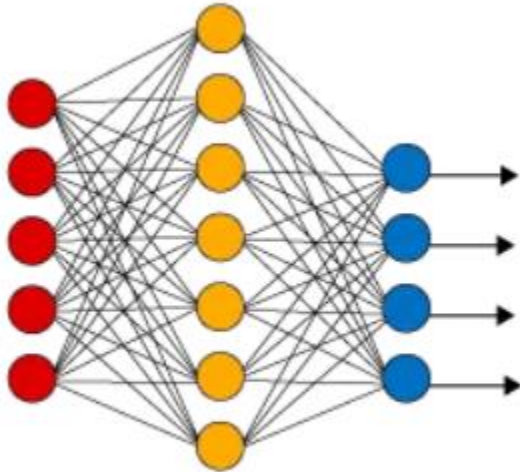
Derin Öğrenme

- Derin öğrenme, bünyesinde yapay sinir ağı (YSA / ANN) yapısal mimarisini kullanan bir makine öğrenmesi yöntemidir.

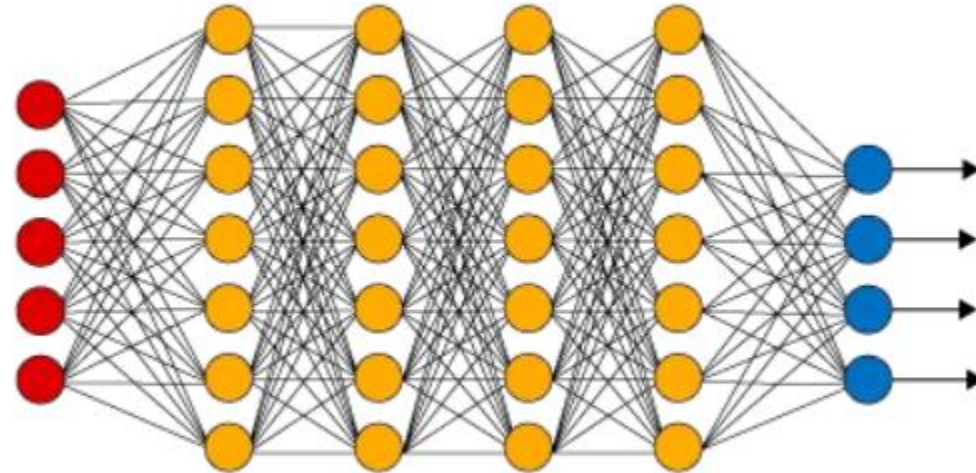


Derin Öğrenme

Simple Neural Network



Deep Learning Neural Network



● Input Layer ● Hidden Layer ● Output Layer

<https://yapayzeka.ai/python-icin-5-muhtesem-derin-ogrenme-kutuphanesi/>

Derin Öğrenme- Giriş

- Son yıllarda veri miktarının çok büyük boyutlara ulaşmış olması (Büyük Veri-Big Data)
- Grafiksel İşleme Birimi (GPU) ve Merkezi İşleme Birimi (CPU) gibi donanımların hızlanması / gelişmesi

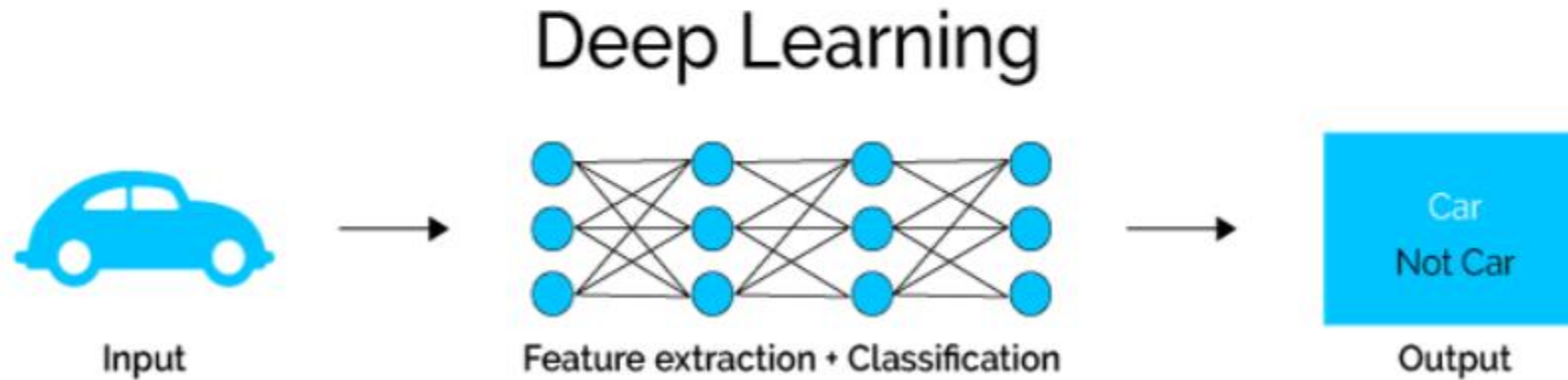
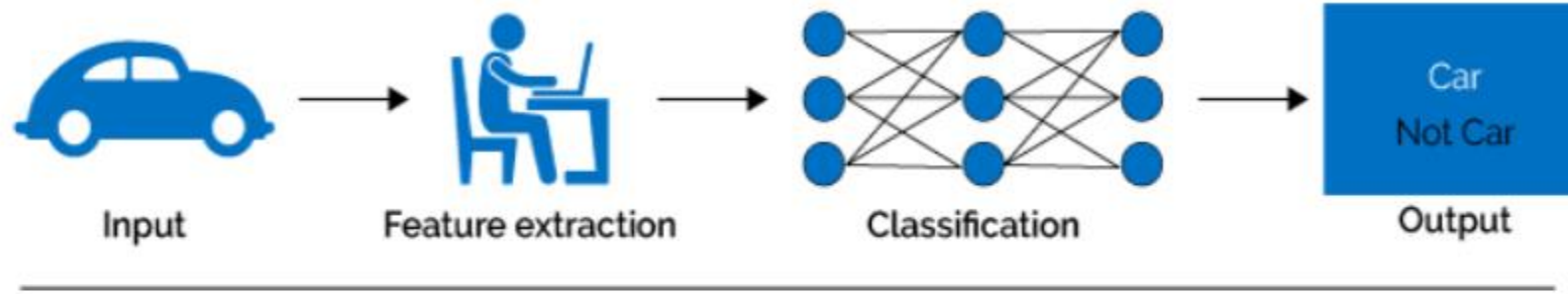
daha fazla sayıda sinir ve katmana sahip YSA'nın eğitilmesi ve çalıştırılmasını mümkün hale getirmiştir. Bu ağlar Derin Sinir Ağları, bu öğrenme de Derin Öğrenme olarak adlandırılır.

Derin Öğrenme-Giriş

- Derin öğrenme terimi ilk olarak 1986'da Rina Dechter tarafından kullanılmıştır.
- Derin ağların çok kullanılan türleri: Çok katmanlı algılayıcılar(Multi Layer Perceptron-MLP), Evrimsel Sinir ağlar (Convolutional Neural Network-CNN), Tekrarlayan Sinir ağları (Recurrent Neural Network-RNN).
- MLP işlemsel veriler üzerinde; CNN görüntü işlemede; ses, metin, zaman serisi gibi ardışık verilerin işlenmesinde ise RNN daha iyi sonuç vermektedir.

Derin Öğrenme-Giriş

- Derin öğrenme; ön işleme, boyutsal indirgeme, öznitelik çıkarma ve sınıflandırma aşamalarını tek ağda birleştirir.



<https://www.deeplearning-academy.com/p/ai-wiki-machine-learning-vs-deep-learning>

Derin Öğrenmede Popüler Etkinlik & Kuruluş

- The Neural Information Processing Systems (NIPS) Foundation (Sinir bilgisi işleme sistemleri vakfı):sinir bilgi işleme sistemleri üzerinde biyolojik teknolojik, matematiksel ve teorik açıdan araştırma ve bilgi alışverişi amacıyla kurulan kar gütmeyen bir kuruluştur.
- ImageNet Large Scale Visual Recognition Challenge –Büyük Ölçekli Görsel Tanıma Yarışması) 2010 yılından beri yapılmakta olan Google, Microsoft gibi kuruluşların yarıştığı ve desteklediği bir etkinliktir.

Derin Sinir Ağları Kullanım Alanları



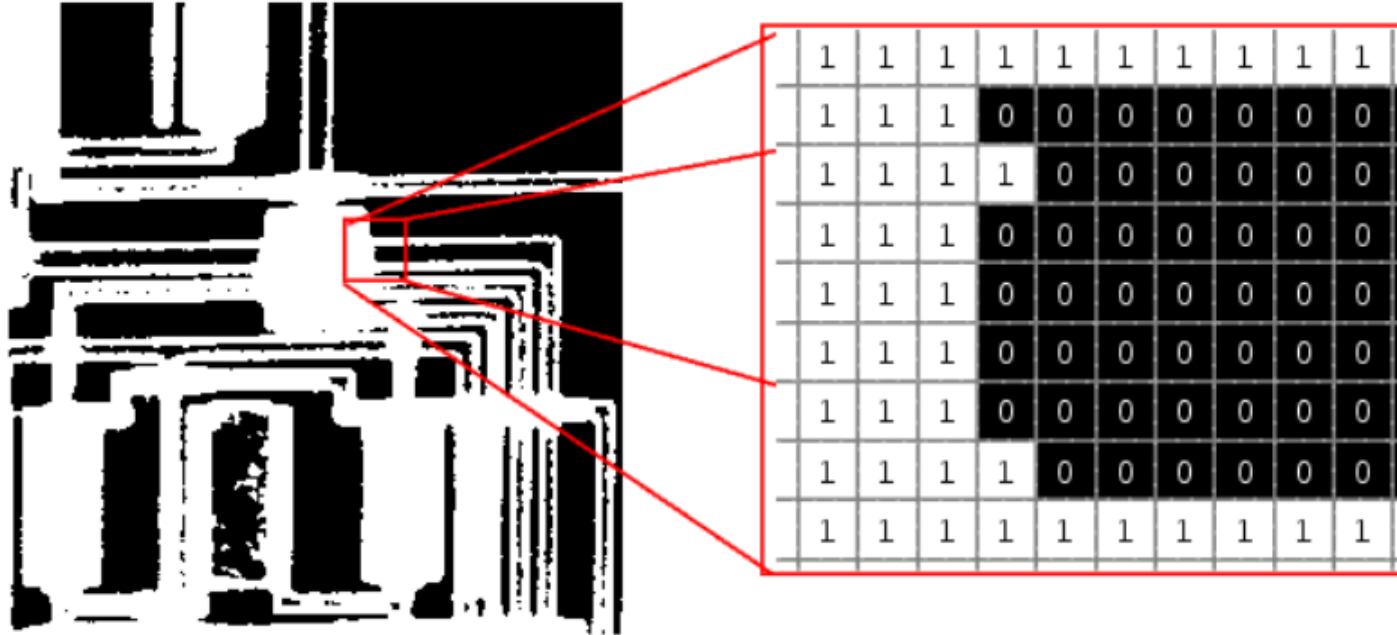
- Derin Sinir Ağları daha çok görüntü işleme görüntü sınıflandırma, görüntü bölütleme gibi amaçlarla kullanılmaktadır.

Sayısal Görüntü İşleme

- Sayısal görüntü, görüntünün (x,y) koordinatları ve bu koordinatlara ait renk değerinden oluşur (piksel). Piksellerin aldığı renk değerleri bakımından görüntü 3'e ayrılır.
 - Siyah beyaz görüntü
 - Gri seviyeli görüntü
 - Renkli görüntü
- Sayısal görüntü işleme, görüntünün onarılması, iyileştirilmesi, anlamlı bölgelere ayrıştırılması gibi amaçlarla gerçekleştirilen klasik veya yapay zekaya dayalı işlemlerdir.
- "Görüntü onarımı" çeşitli sebeplerle bozulan görüntünün tekrar orijinal haline dönüştürülmesi, "görüntü iyileştirme" ise görüntünün daha temiz ve anlaşılır olmasını sağlama amacıyla gürültü giderme kenar belirginleştirme vb. işlemlerdir.

Siyah-Beyaz Görüntü

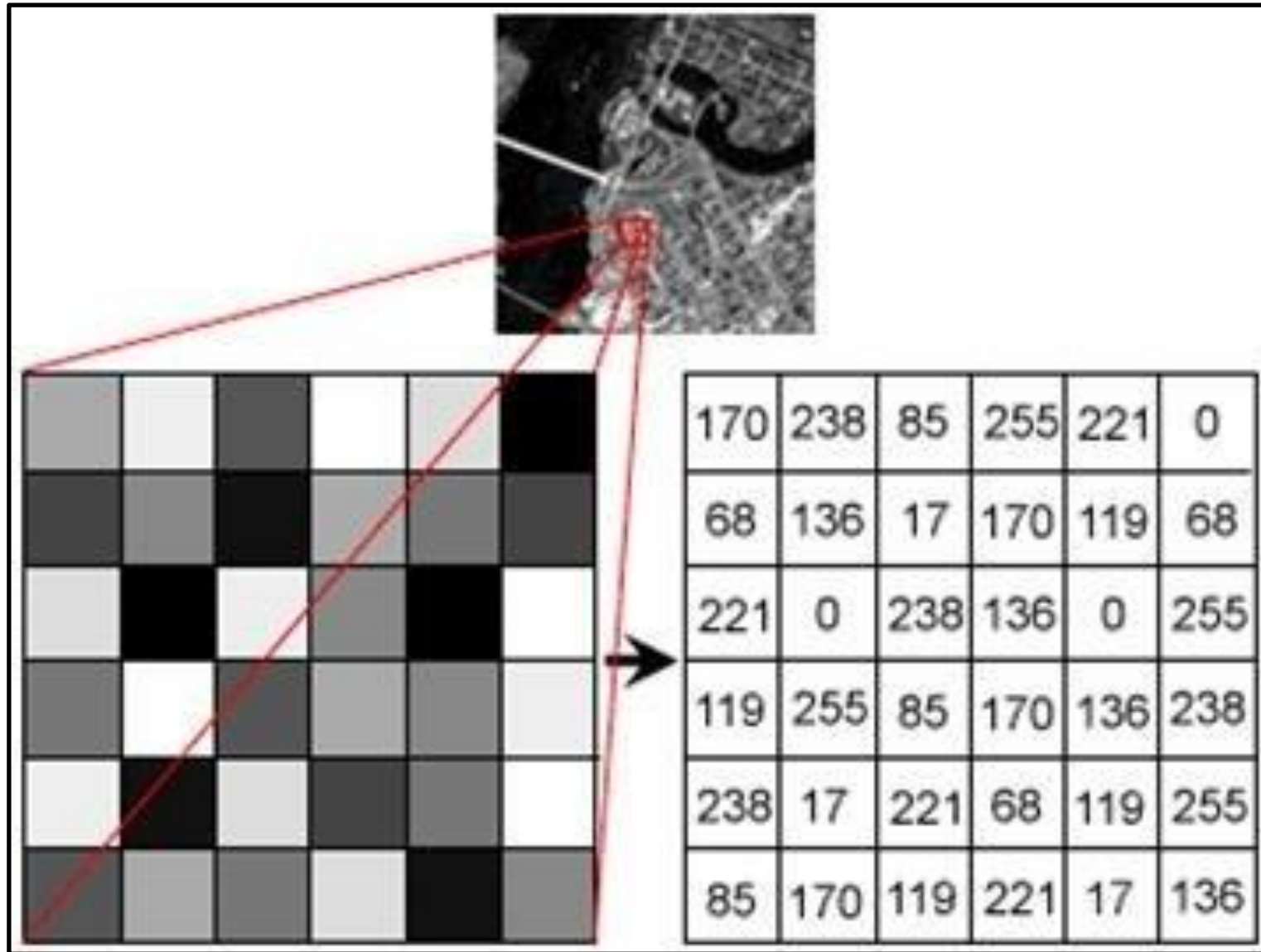
- Sadece siyah ve beyaz olmaz üzere iki piksel değerinden oluşan görüntülerdir



<https://www.mathworks.com/help/images/binary-images.html>

Gri Seviyeli Görüntü

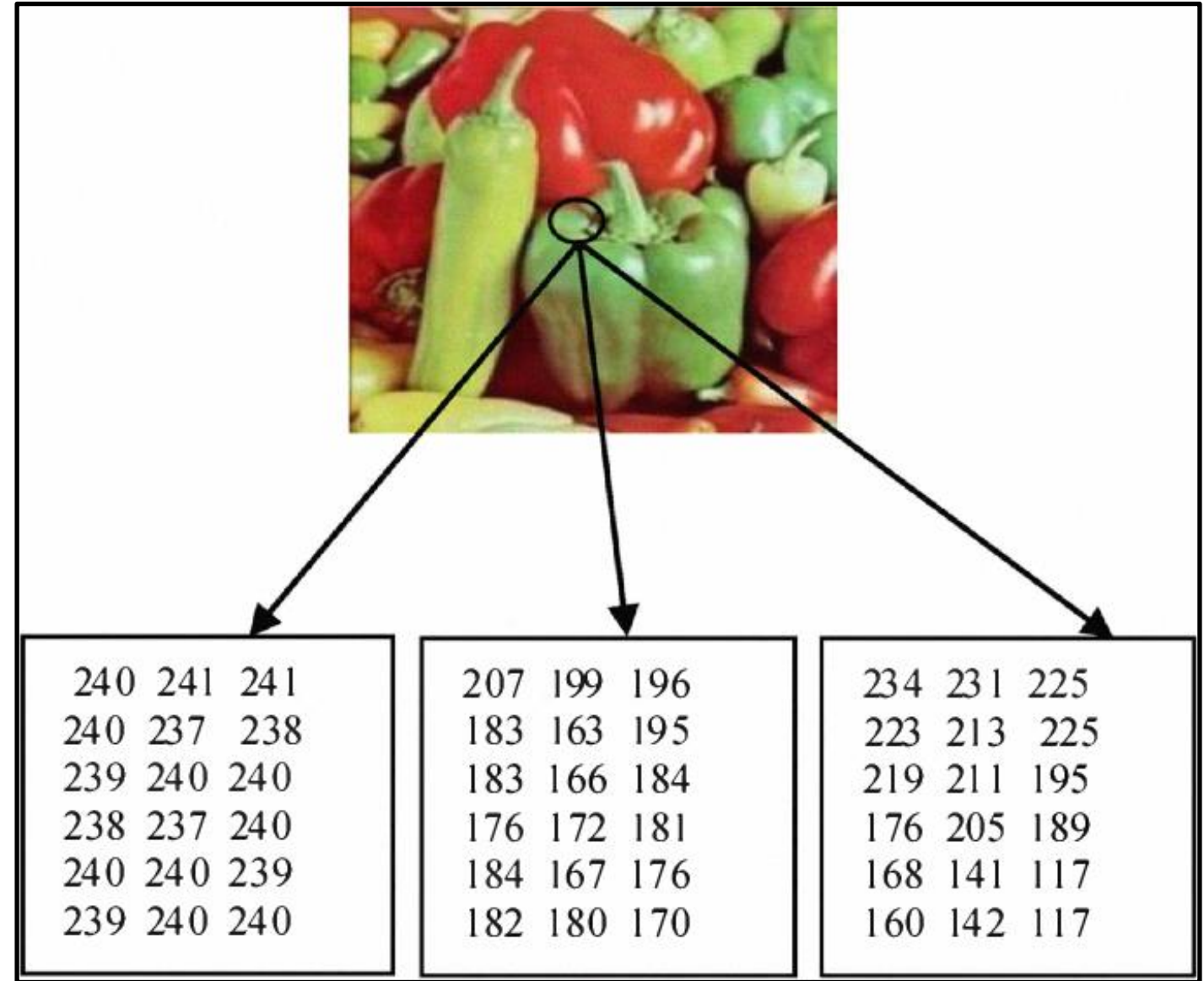
- Piksellerinin tümü gri seviye renklerden oluşur.



Neves et al., Analysis of Emotions From Body Postures Based on Digital Imaging, Third International Conference on Advances in Signal, Image and Video Processing, 2018

Renkli Görüntü

- Pikselleri kırmızı, yeşil ve mavi renk değerleri içerir.

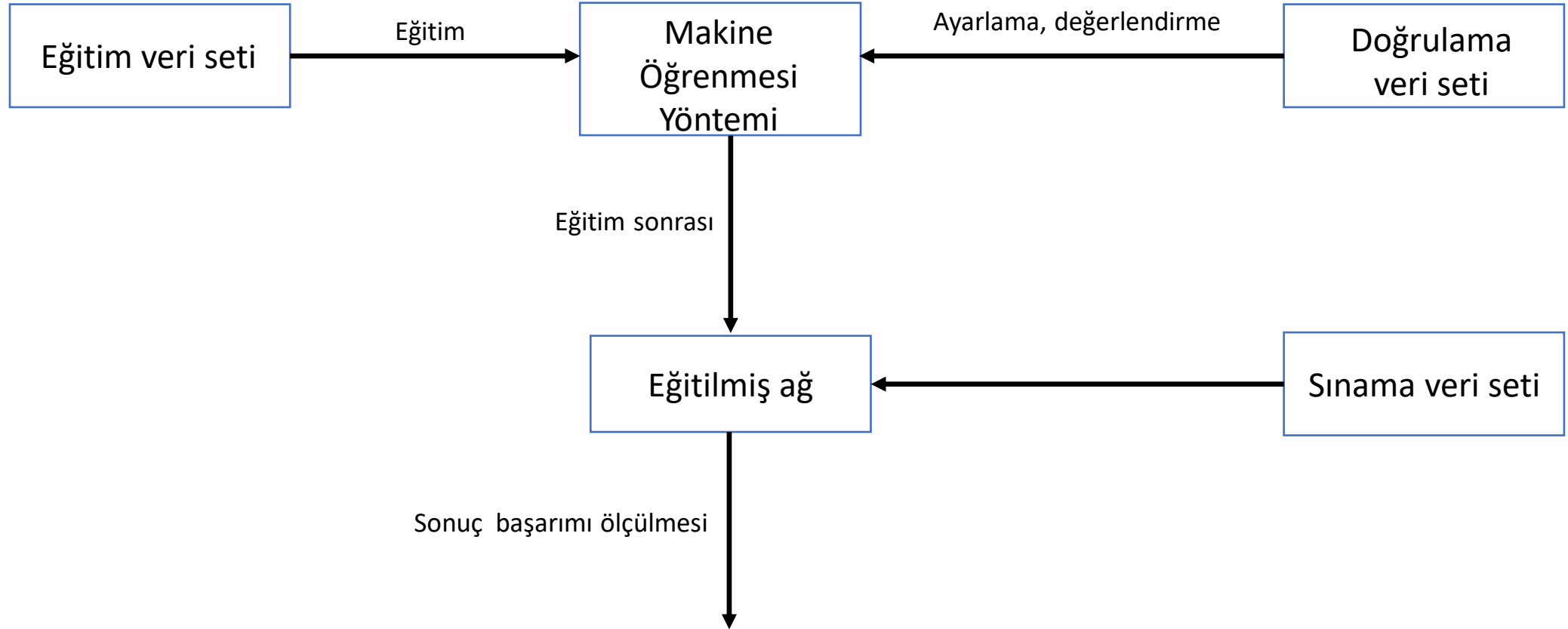


Karim et al., A new approach for LSB based image steganography using secret key, Computer and Information Technology (ICCIT), 2011

Derin Sinir Ağlarda veri kümesi

- YSA'da olduğu gibi derin sinir ağlarda da başarıımı etkileyen önemli faktörlerden biri ağa sunulacak veri kümesidir.
- Veri kümesi genel olarak 3'e bölünür.
 - Eğitim (training),
 - Test (testing)
 - Doğrulama (validation) olarak kullanılır.
- Eğitim kümesi ile ağ eğitilir. Doğrulama veri kümesi eğitim anında başarıımı kontrol etmek için kullanılır. Eğitim sonlandığında ise test veri kümesi ile ağın başarıımı hesaplanır.

Derin Sinir Ağlarda veri kümesi



Derin Sinir Ağlarda veri kümesi

- Derin ağların eğitiminde kullanılabilecek el yazısın tanıma, nesne tanıma vb. çok geniş yelpazede açık kaynak kodlu ücretsiz veri kümeleri bulunmaktadır. En bilinenleri aşağıda açıklanmıştır.
- MNIST (Mixed National Institute of Standards and Technology): El yazısı ile yazılmış ve uygun bir şekilde sınıflandırılmış 60.000 eğitim ve 10.000 test görüntüsünden oluşur.
- CIFAR (Canadian Institute for advanced Research): 80 milyon küçük görüntü kümesi içerir. Nesne tanıma amacıyla kullanılır. Uçak, kuş, kedi, vb. görüntüler içerir.
- ImageNet: 14 milyondan fazla görüntü ve 20.000'den fazla sınıf içerir. Sınıflar basit nesnelerden oluşur: araç, yiyecek, ağaç, vb.

Derin Öğrenme Kütüphaneleri

- Derin öğrenme için geliştirilmiş çok sayıda kütüphane vardır. En yaygın kullanılanları aşağıda açıklanmıştır.
 - TensorFlow: Sayısal hesaplama için kullanılan açık kaynak kodlu bir Python derin öğrenme kütüphanesidir.
 - Theano: Matematik ifadelerini etkili bir şekilde kullanmayı sağlayan Python derin öğrenme kütüphanesidir.
 - Keras: Hem Theano hem de Tensorflow üzerinde çalışan Python ile yazılmış bir derin öğrenme kütüphanesidir.

Derin Öğrenme Kütüphaneleri

- Torch: Algoritmaları oluşturma konusunda daha işlevsel, kullanıcı dostu arayüzlü bir derin öğrenme kütüphanesidir.
- DL4J (Deep Learning for Java): Keras, TensorFlow gibi açık kaynak kütüphaneleri üzerinde çalışabilen bir Java derin öğrenme kütüphanesidir.
- Caffe: Özellikle CNN’de çok hızlı çalışan modüler ve kullanıcı dostu arayüzlü derin öğrenme kütüphanesidir.

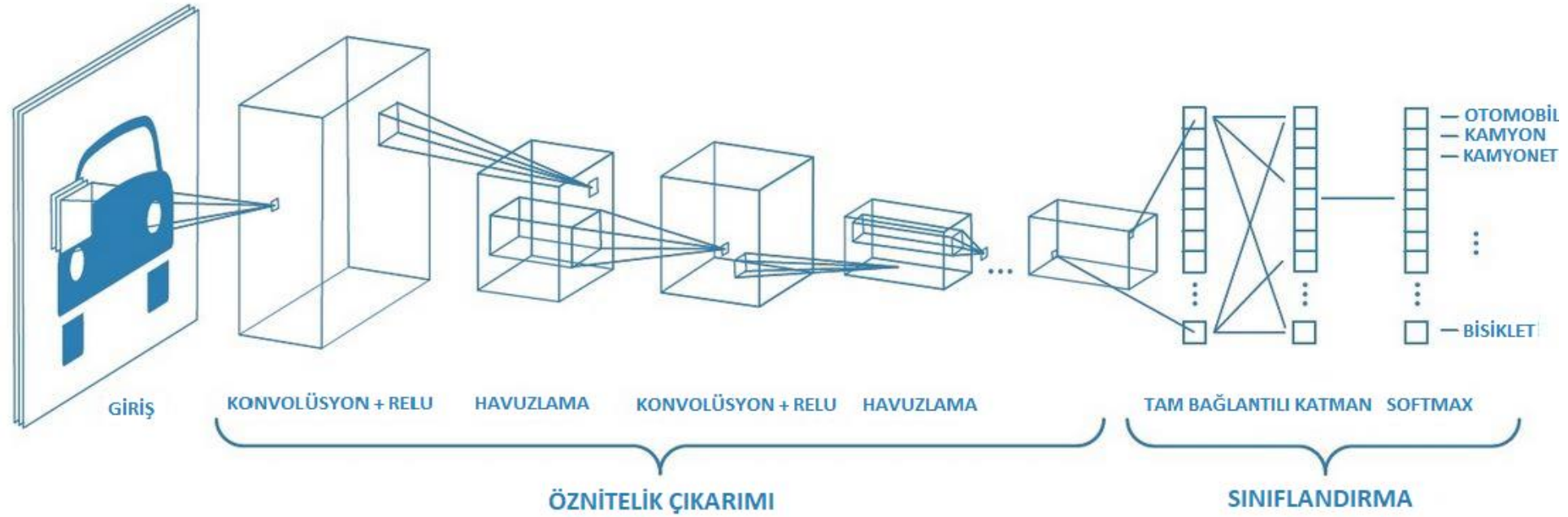
Derin Öğrenme için kullanılan programlama dilleri

- Derin öğrenme için genellikle kullanılan programlama dilleri aşağıda sıralanmıştır.
- Python: En esnek dillerden biridir. Basitliği sayesinde yapay zeka alanında en sık kullanılan dillerden biridir.
- C++: Donanım düzeyinde programlama yeteneği sayesinde derin öğrenme projelerine uygundur.
- Java: Algoritma kodlamada kolaylığı ile derin öğrenme projelerine uygundur.
- R dili: İstatistiksel modellerin üretilmesinde esneklik sağlar. Oldukça yeni bir dil olmasına rağmen akademik çalışmalarda ve sektörde yaygınlığı artmaktadır.

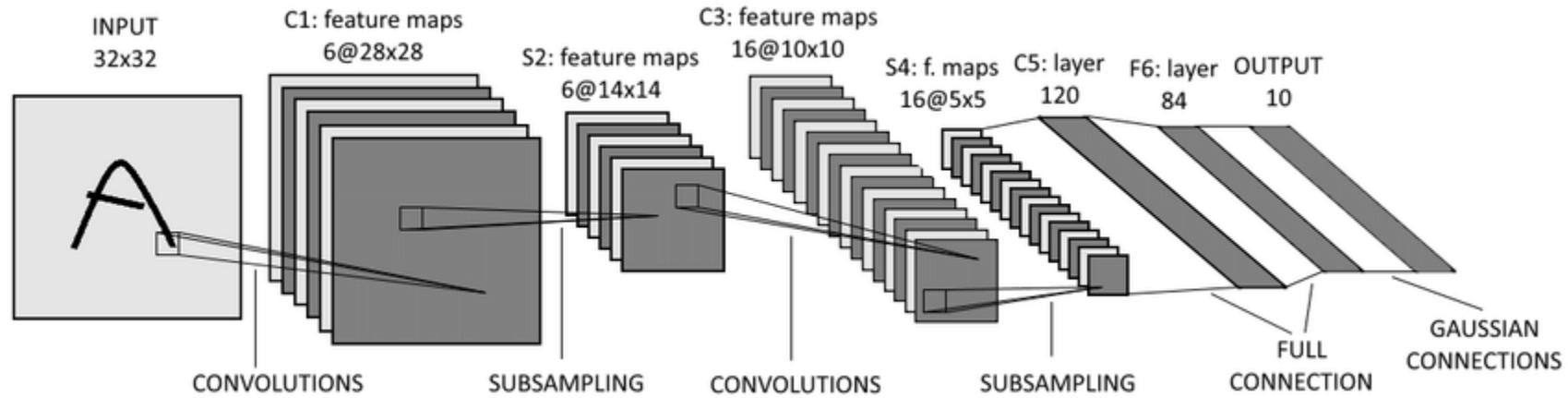
Evrişimsel Sinir Ağlar (CNN)

- YSA'lara göre daha fazla sinir ve katmana sahiptir.
- Çalışma şekilleri farklıdır. Örneğin YSA'lar şekil tanıma yaparken alt parçalar yerine şeklin tamamını tanımaya çalışır, CNN önce şeklin alt parçalarını bulur ardından bu şekiller birleştirilerek şeklin tamamı bulunur.
- CNN konvolüsyon (evrişim), havuzlama(pooling) ve tam bağlı (fully-connected) katmanlardan oluşur.
- CNN'e ait ilk konvolüsyon katmanı görüntüde kenar bulma, köşe bulma vb. genel özellikleri bulmaya yönelik çalışır. Sinir ağının sonuna doğru ise probleme özgü ayırt edici özellikler elde edilir.

Evrişimsel Sinir Ağlar (CNN)



Evrişimsel Sinir Ağlar (Convolutional Neural Networks - CNN)



- Yukarıdaki şekil 1998 yılında YannLeCun ve ekibi tarafından geliştirilen LeNet ağını gösterir. Bu ağ el yazımı rakam tanıma için geliştirilmiştir.
- Bu ağın sinir sayısı ve bağlantı sayısı 60bini bulmaktadır. Bu sayı YSA ile kıyaslandığında çok önemli ölçüde artmış bununla birlikte başarımlar da önemli ölçüde iyileşmiştir.

Evrişim (Konvolüsyon)

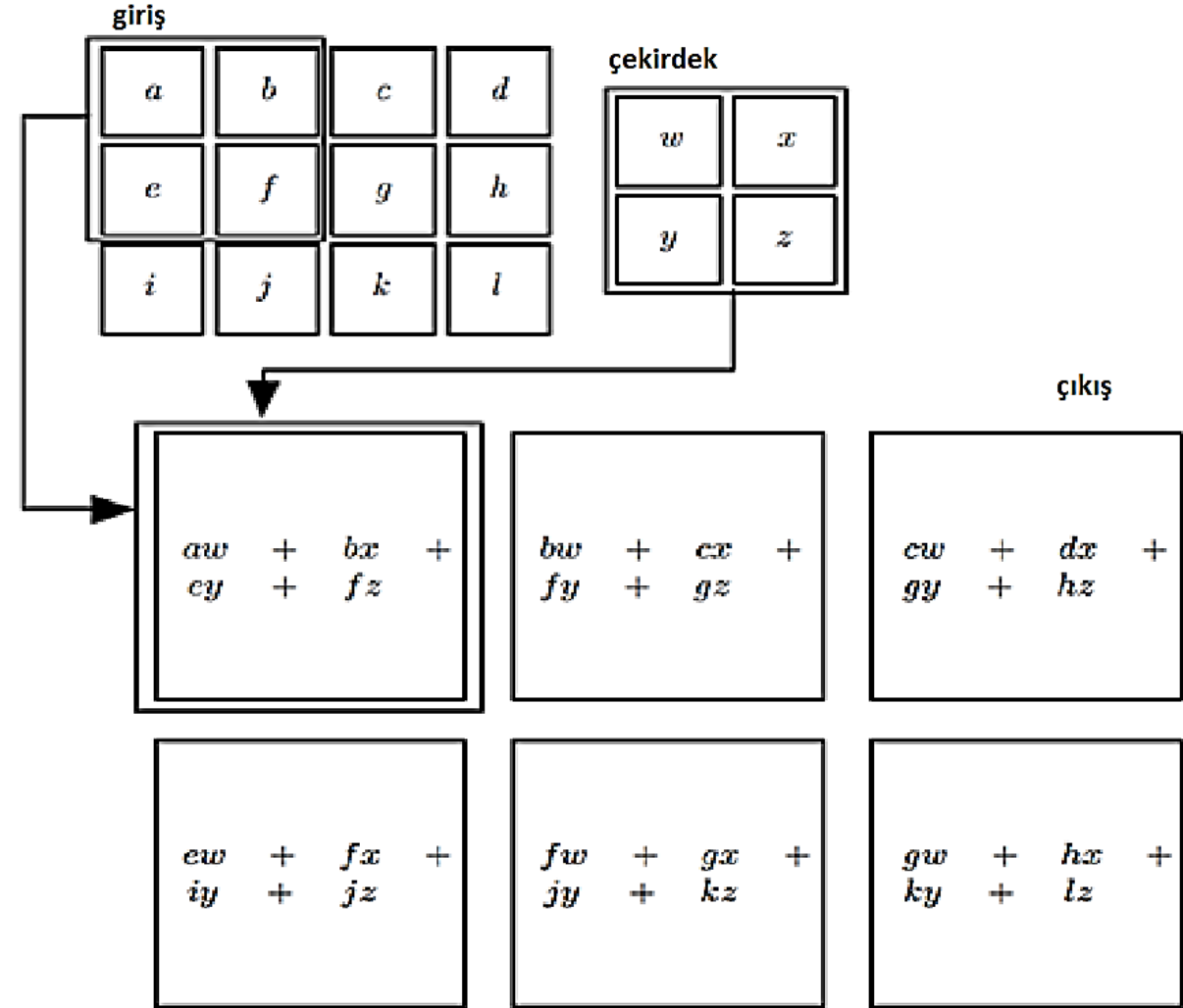
- Toplama, çıkarma, türev gibi matematiksel bir işlemdir.
- Birincil amacı giriş olarak verilen görüntülerin özellik haritalarını çıkarmaktır.
- Bu işlem girdi olarak genellikle iki boyutlu veri dizilerini ve bu diziler üzerinde dolaşabilecek daha küçük iki boyutlu çekirdek dizilerini parametre olarak alır.
- Konvolüsyon işlemi “ * ” ile gösterilir ve matematiksel ifadesi:

$$S(i, j) = (I * K)(i, j) = \sum_m \sum_n I(i + m, j + n) K(m, n)$$

- I , $m \times n$ boyutlu görüntü için girişi; K çekirdek verilerini, s ise i, j noktalarındaki işlem sonucunu temsil etmektedir.

Evrişim (Konvolüsyon)

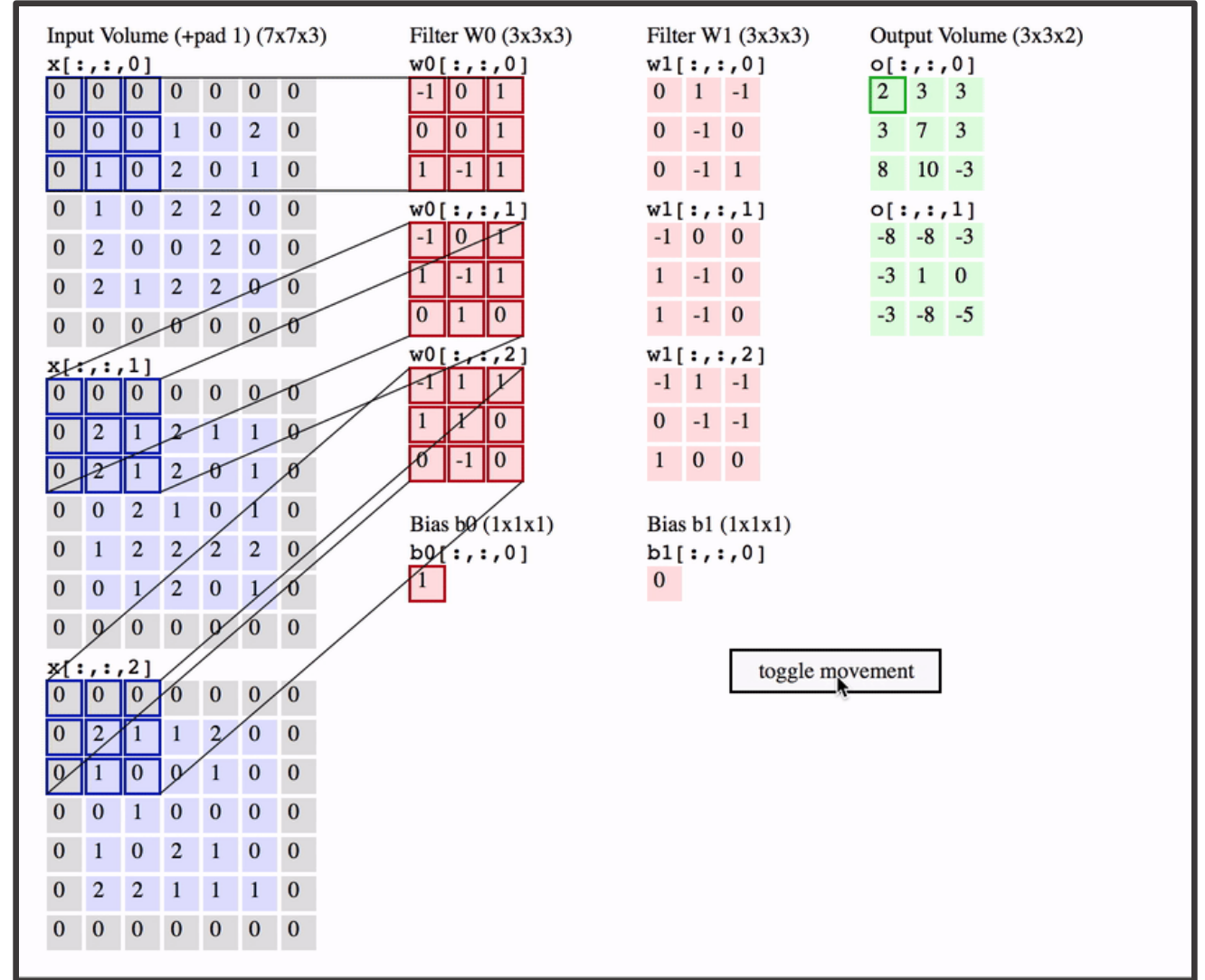
- kayma oranı (stride) 1
- görüntü kenarlarına ekleme yapma (zero padding)



Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, May 2015.

Evrişim (Konvolüsyon)

- stride 2
- padding 1



CS231n Course Materials, "CS231n Convolutional Neural Networks for Visual Recognition." <https://skymind.ai/wiki/convolutional-network>.

Sıfır Dolgu (padding)

- Giriş matrisinde çerçeve dışına sıfır ekleme işlemidir.
- Dolgu işlemi ile giriş boyutu ihtiyaca göre ayarlanabilir.
- Genellikle giriş boyutunun çıkışta ayarlanması gerektiği durumda kullanılır.
- Aşağıdaki şekilde 3x3 boyutlu bir matrise her kenardan sıfır dolgu uygulanarak matris 5x5 olarak büyütülmüştür.
- Sıfır dolgu sadece bir kenardan ya da iki kenardan da yapılabilir.

| | | |
|---|---|---|
| 1 | 2 | 1 |
| 4 | 5 | 1 |
| 6 | 8 | 1 |

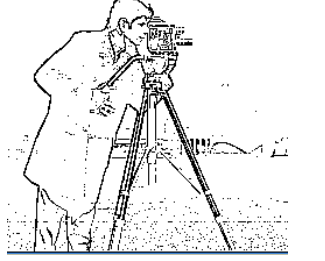




| | | | | |
|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 2 | 1 | 0 |
| 0 | 4 | 5 | 1 | 0 |
| 0 | 6 | 8 | 1 | 0 |
| 0 | 0 | 0 | 0 | 0 |

Farklı Filtrelerin Evrişime Etkileri

Özgün görüntü



| İşlem | Filtre | | | Evrişmiş Görüntü |
|---------------------|--------|-------|--------|--|
| Kenar Algılama | -1 | -1 | -1 |  |
| Keskinleştirme | 0 | -1 | 0 |  |
| Gaussian Bulanıklık | 0.0625 | 0.125 | 0.0625 |  |

Evriřim (Konvolüsyon)



Mathworks, "Visualize Activations of a Convolutional Neural Network. <https://ww2.mathworks.cn/help/nnet/examples/visualize-activations-of-a-convolutional-neural-network.html>.

Evriřim (Konvolüsyon)

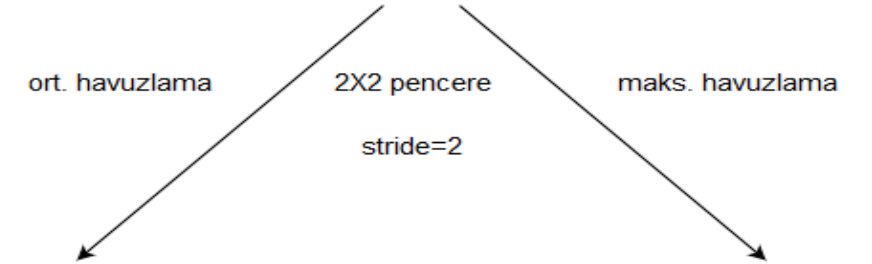


Mathworks, "Visualize Activations of a Convolutional Neural Network. <https://ww2.mathworks.cn/help/nnet/examples/visualize-activations-of-a-convolutional-neural-network.html>.

Havuzlama (Pooling) Katmanı

- Bir sonraki konvolüsyon katmanı için giriş boyutunu azaltır, bilgi kaybına yol açar.
- Hesaplama yükünü azaltır.
- Ezberlemeyi önler.

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



| | | | |
|---|---|---|---|
| 2 | 4 | 1 | 3 |
| 0 | 4 | 3 | 5 |
| 6 | 3 | 2 | 4 |
| 3 | 3 | 4 | 6 |

| | | | |
|---|---|---|---|
| 7 | 9 | 6 | 8 |
| 2 | 7 | 6 | 9 |
| 9 | 8 | 5 | 8 |
| 5 | 8 | 8 | 8 |

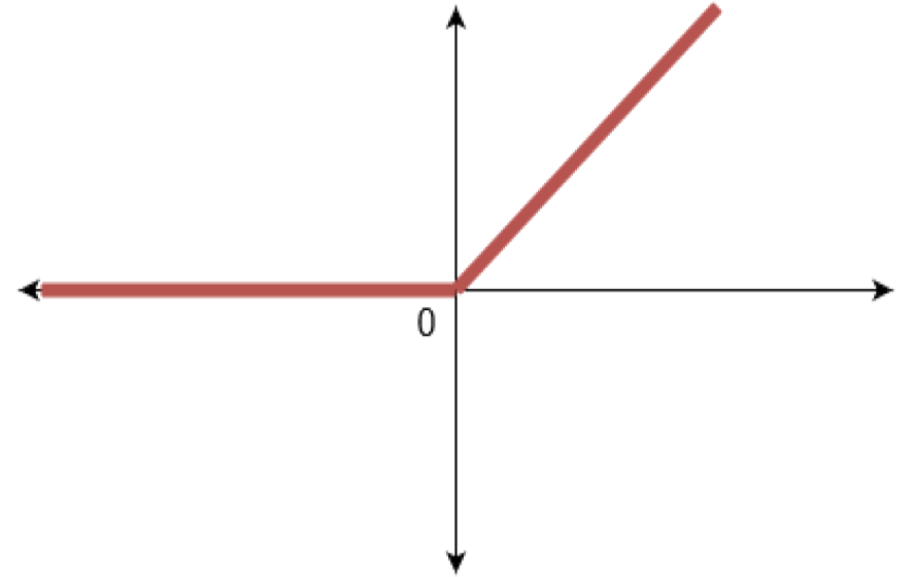
Katmanların çıkış boyutlarının hesaplanması

| Katman | Çıkış boyutu hesaplama formülü |
|-----------------------|--------------------------------|
| Havuzlama (Pooling) | $\frac{w - f}{s} + 1$ |
| Evrişim (Convolution) | $\frac{w - f + 2p}{s} + 1$ |

- w: giriş boyutu
- f: filtre boyutu
- s: adım sayısı
- p: padding sayısı

ReLU Katmanı

- $G(z) = \max\{0, z\}$
- ReLU'nun giriş parametresi olarak aldığı değer negatif veya sıfır ise sıfır, değilse kendisini döndürür.
- ReLU'yu diğer aktivasyon fonksiyonlarından ayıran en önemli özellik, diğerlerine göre daha hızlı çalışmasıdır.
- Bu katmanın kullanılması ile ağ daha hızlı bir eğitim gerçekler.



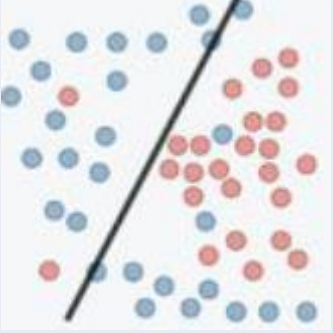
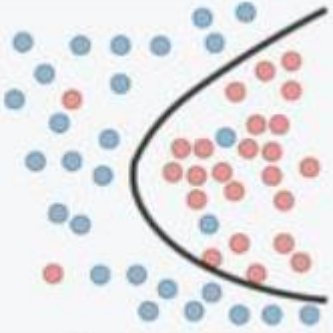
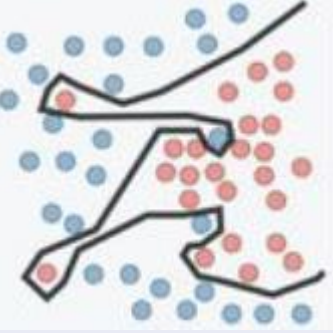



Ağın eğitiminde ortaya çıkabilecek sorunlar

- Ağ eğitilirken karşılaşılabilecek sorunlar
 - Ezberleme (Overfitting)
 - Yanlış öğrenme (Underfitting)
- Ezberleme ve yanlış öğrenme iki sebeple olabilir.
 - Veri kümesi ile ilgili nedenler,
 - Eğitim süresi ile ilgili nedenler

Ağın eğitiminde ortaya çıkabilecek sorunlar

- Ezberleme, ağın eğitim setine doğru cevap verip test verilerine genelleme yapamayıp düzgün sonuç üretememesidir.
 - Eğitim uzay yeterli genişliğe sahip değilse ezberleme ortaya çıkar.
 - Örnek: Bir ağın kuş cinslerini tanımak üzere eğitildiğini düşünelim. 50 bin görüntü ile ağ eğitilsin. Bu görüntülerin yarısı kanarya, diğer yarısı ise kanarya benzeri çeşitli kuşlara ait olsun. Bu durumda veri kümesi dengesiz dağıldığından ezberleme ihtimali olasıdır.
 - Ezberlemeyi önlemek için eğitim veri kümesi genişletilmelidir.
 - Aynı verilerle çok fazla eğitim yapıldığında da ezberleme sorunu ortaya çıkar.

Ağın eğitiminde ortaya çıkabilecek sorunlar

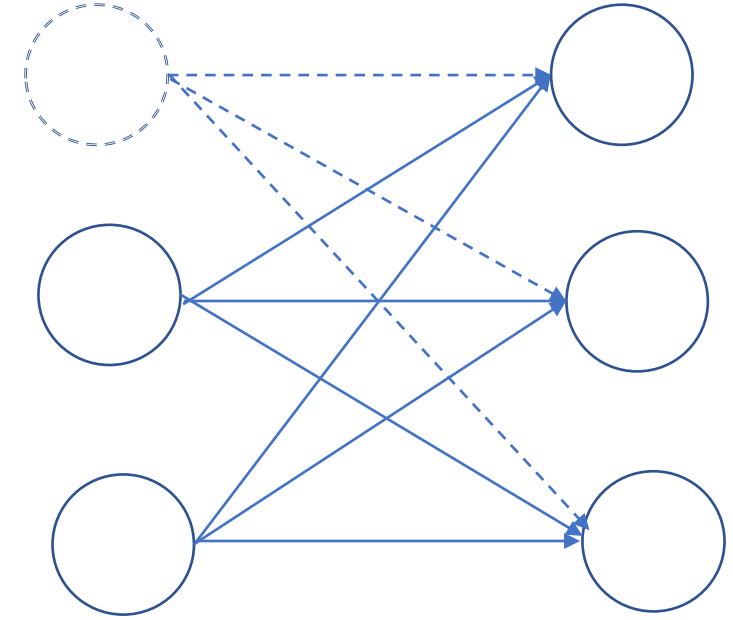
| | Underfitting (Yanlış öğrenme) | Öğrenme | Overfitting (Ezberleme) |
|------------------------------------|--|---|---|
| Belirtiler | <ul style="list-style-type: none">Yüksek eğitim hatasıEğitim hatası ile test hatası birbirine yakınYüksek bias | <ul style="list-style-type: none">Eğitim hatası test hatasından az miktarda düşüktür. | <ul style="list-style-type: none">Çok düşük eğitim hatasıEğitim hatası test hatasından çok çok düşüktür. |
| Sınıflandırma gösterimi |  |  |  |
| Derin Öğrenme gösterimi |  |  |  |
| Çözüm için yapılabilecekler | <ul style="list-style-type: none">Model karmaşılaştırılabilirDaha fazla özellik (feature) eklenebilirDaha uzun süre eğitim yapılabilir | | <ul style="list-style-type: none">Daha fazla veri eklenebilir |

Seyreltme

- Eğitimde ortaya çıkan ezberleme gibi sorunları önlemek için seyreltme başvurulabilecek bir başka yöntemdir.
- Seyreltme ağ boyutunu azaltarak eğitimin iyileşmesine yardımcı olur. Ağların daha genelleştirilmiş sonuçlar vermesine katkı sağlar.
- Seyreltme 3 şekilde yapılabilir:
 - Sinir Seyreltme (Dropout)
 - Bağlantı Seyreltme (Dropconnect)
 - Tümünü Seyreltme (DropAll)

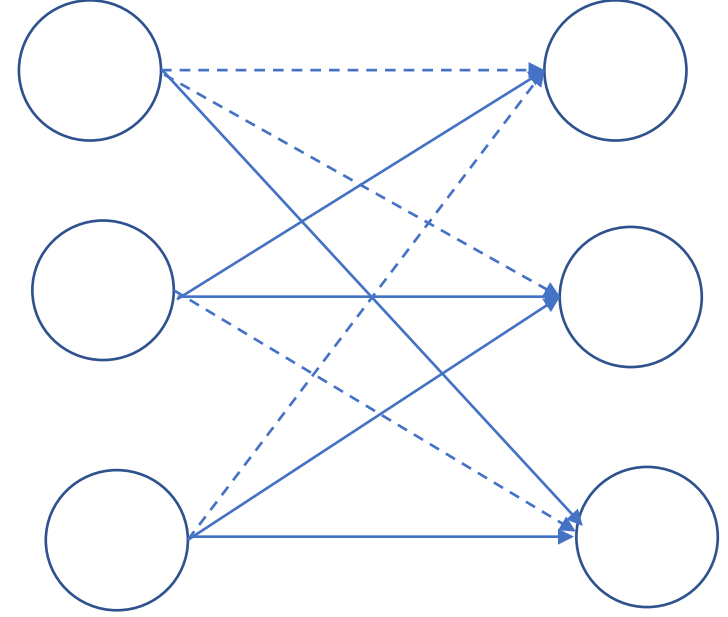
Dropout (Sinir seyreltme) katmanı:

- Dropout katmanı, ağlarda aşırı uydurma (overfitting) önüne geçmek için kullanılır.
- Her eğitim aşaması için belli bir oranda rastgele seçilen sinirler yok sayılır. O sinirden sonraki katmandaki sinirlere bağlantı yapılmaz.



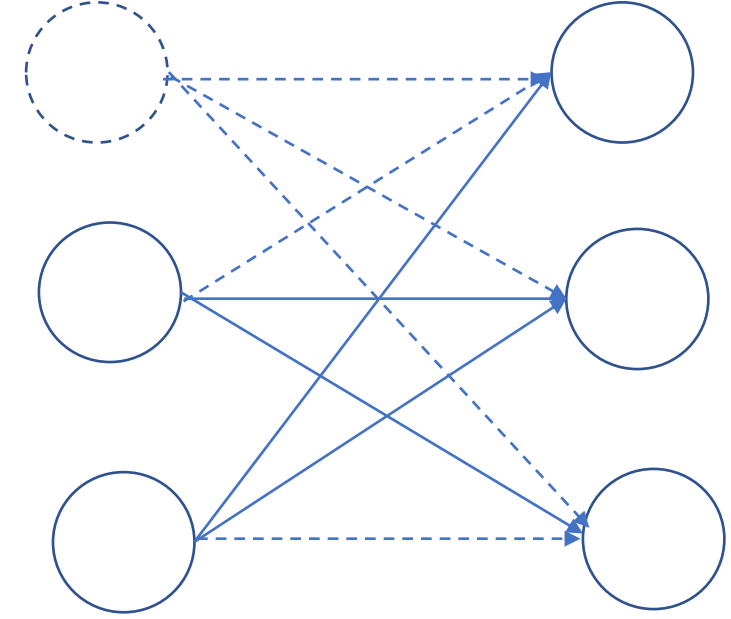
DropConnect (Bağlantı seyreltme) katmanı:

- Sinir seyreltme işlemindeki sinirlerin sonraki katman sinirlerine bağlanmaması yerine belli bir oranda rastgele seçilen ağırlıklar sıfır yapılarak sonraki katmandaki bazı sinirlere bağlantılar koparılır.



DropAll (Tümünü seyreltme) katmanı:

- Sinir seyreltme ve bağlantı seyreltmenin belli bir oranda rastgele seçilen sinir ve bağlantılara uygulanması işlemidir.



Tam Bağlantılı Katman (Fully Connected):

- Softmax ile birlikte sınıflandırma işlevlerini yerine getiren katmandır.
- Kendinden önceki katmanın tüm elemanlarıyla bağlantılıdır ve sınıflandırmadan önceki aşama için, sınıf etiketlerine dair olasılıksal değerler içerir.
- Bu katmandan elde edilen sayısal değerlerin hangi sınıf ile ilişkili olduğuna softmax katmanı karar verir ve bu olasılıksal değerleri $[0,1]$ aralığına çeker.
- Değerler incelendiğinde; en büyük değer hangi sınıfa ait ise, ağ o sınıf etiketini çıktı olarak üretir.

Geri Yayılım Algoritması (Backpropagation)

- Yapay sinir ağlarında kullanılan eğitim yöntemidir.
- Geri yayılım yaklaşımı ile güncellemeleri ağ üzerinde gerçekleştirir.
- Güncellemeler ağın çıktısı ile beklenen çıktı arasındaki hatayı minimize etmeyi amaçlar. Bu hata değerleri en küçük ortalamaların karesi (Least Mean Square - LMS) hatası ile hesaplanabilir.

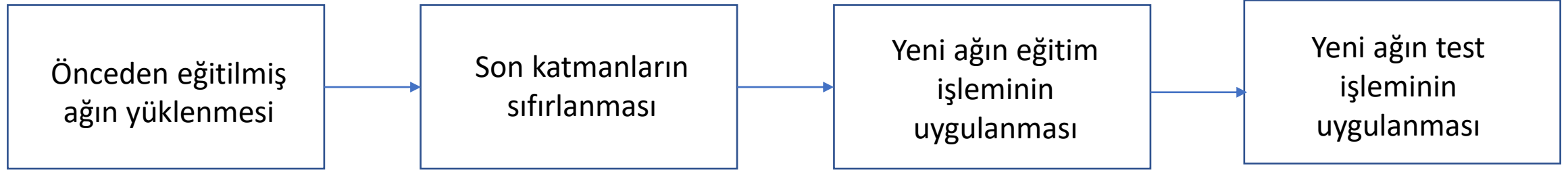
$$J(w) = \frac{1}{2} \sum_{k=1}^c (t_k - z_k)^2$$

- Denklemden z tahmin edilen sonuç vektörünü, t ise olması gereken sınıflandırma sonuç vektörünü göstermektedir. c vektör uzunluğunu, w ise ağdaki ağırlıkları sembolize etmektedir.

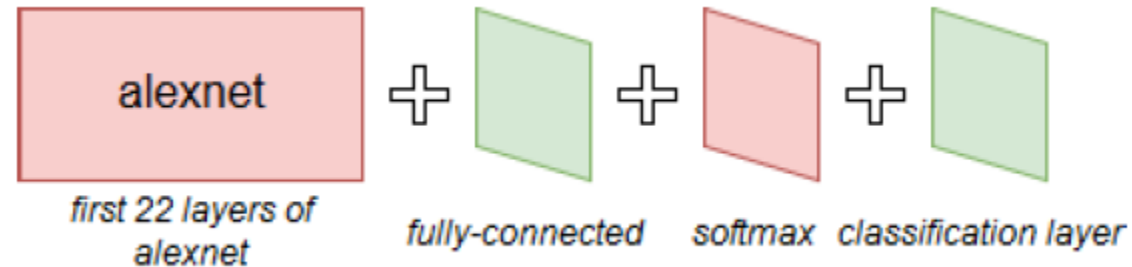
Transfer Öğrenme

- Problemler için yeni bir ağ tasarlanıp eğitim yapılabileceği gibi daha önce farklı bir problem için tasarlanan bir ağ da yeni bir problemin çözümünde kullanılabilir. Bu işlem transfer öğrenme olarak adlandırılabilir.
- Örneğin 1000 tane nesneyi sınıflandırmak için tasarlanan ve oldukça yüksek başarımlar veren AlexNet mimarisi yüz ifadesi tespiti için kullanılabilir ve yine başarılı sonuç üretebilir.
- Transfer öğrenme uygulamak için ağın ilk katmanları sabit bırakılıp son katmanları yeni problemi çözecek şekilde yeniden eğitim yapılarak güncellenir.

Transfer Öğrenme



Transfer öğrenme adımları



Alexnet mimarisinin transfer öğrenme için güncellenmesi

Transfer Öğrenme

- Genellikle, ağın ilk katmanları kenar bulma, renk belirleme, vb. gibi genel görüntü işleme işlemlerini yapar.
 - Yüz ifadesi tespiti için de, nesne tanıma için de, bölütleme işlemleri için de, diğer görüntü sınıflandırma işlemleri için de evrimsel sinir ağların ilk katmanlarında yapılan işlemler benzerdir.
 - Bu nedenle transfer öğrenme oldukça kullanışlı bir yaklaşımdır.

Transfer Öğrenme

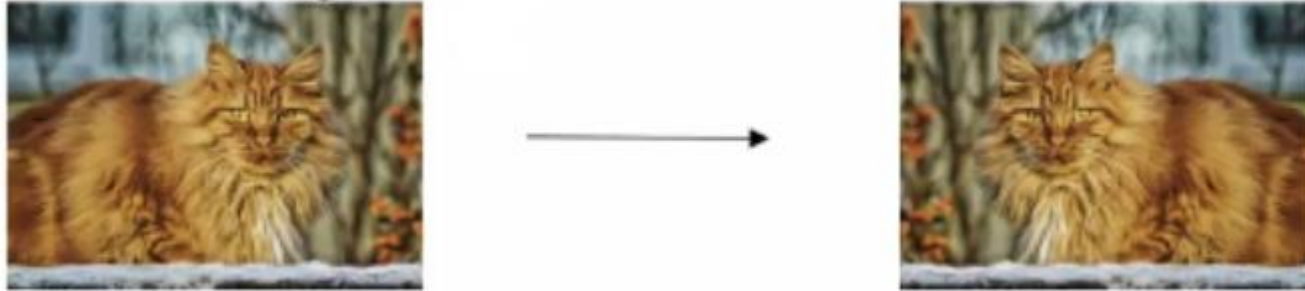
- Transfer öğrenmenin diğer faydaları aşağıda sıralanmıştır.
 - Bazı modellerin eğitimi, standart bilgisayar işlemcileri ile mümkün olmayıp grafik işlemcileriyle eğitmeye ihtiyaç duyar ve bazı eğitimler çok uzun sürer. Bu eğitim sonucunda oluşan ağ modeli farklı amaçlar için de kullanma açısından transfer öğrenme oldukça avantajlıdır.
 - Elimizde çok fazla veri yoksa dahi, transfer öğrenme sayesinde, daha önce milyonlarca veri ile eğitilen bir ağı kullanarak kendi problemimize başarılı sonuç üretebiliriz.
 - Transfer öğrenme sayesinde problemlere daha hızlı çözüm bulabiliriz. Zaman tasarrufu sağlar.

Veri Artırma (Data Augmentation)

- Derin öğrenmede veri sayısının çok olması ağın performansını önemli ölçüde artırır.
- Veri setinin küçük olduğu durumlarda başarıyı artırmak için veri çoklama uygulanabilir.
- Veri setindeki verilerin çeşitli işlemlerle çoklanarak veri sayısının artırılmasına veri çoklama denir.
- Veri azlığından dolayı sürekli aynı verilerle eğitim sonucunda ezberleme sorununun ortaya çıkabileceğini öğrenmiştik. Veri artırma ile ezberleme probleminin de önüne geçilmiş olur.

Veri Artırma (Data Augmentation)

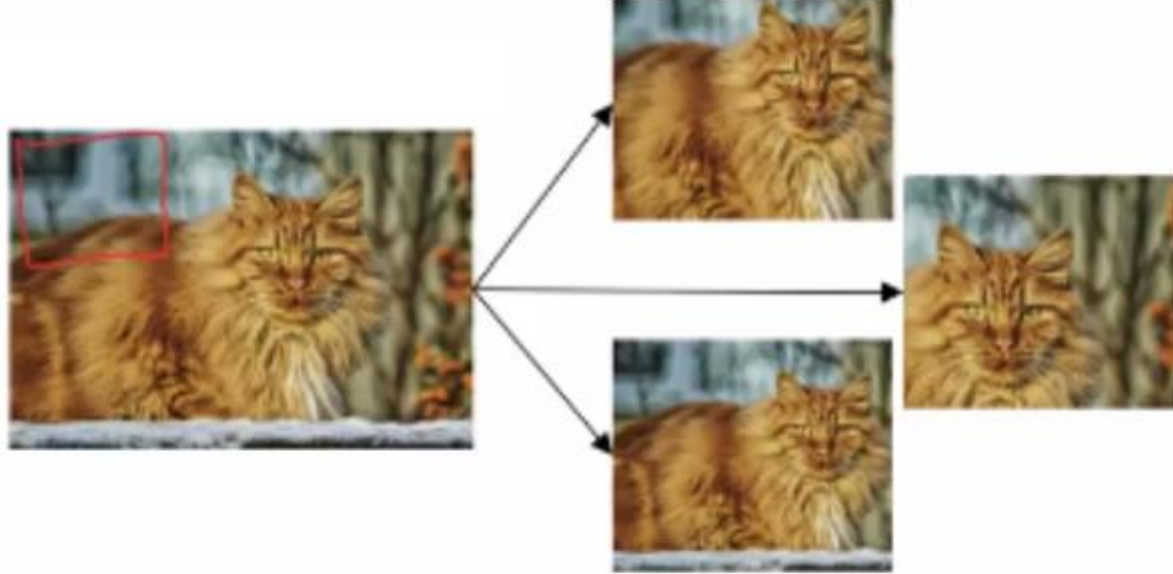
- Görüntünün eksenlere göre simetrilerinin alınması.



<https://medium.com/@ayyucekizrak/deri%CC%87ne-daha-deri%CC%87ne-evri%C5%9Fimli-sinir-a%C4%9Flar%C4%B1-2813a2c8b2a9>

Veri Artırma (Data Augmentation)

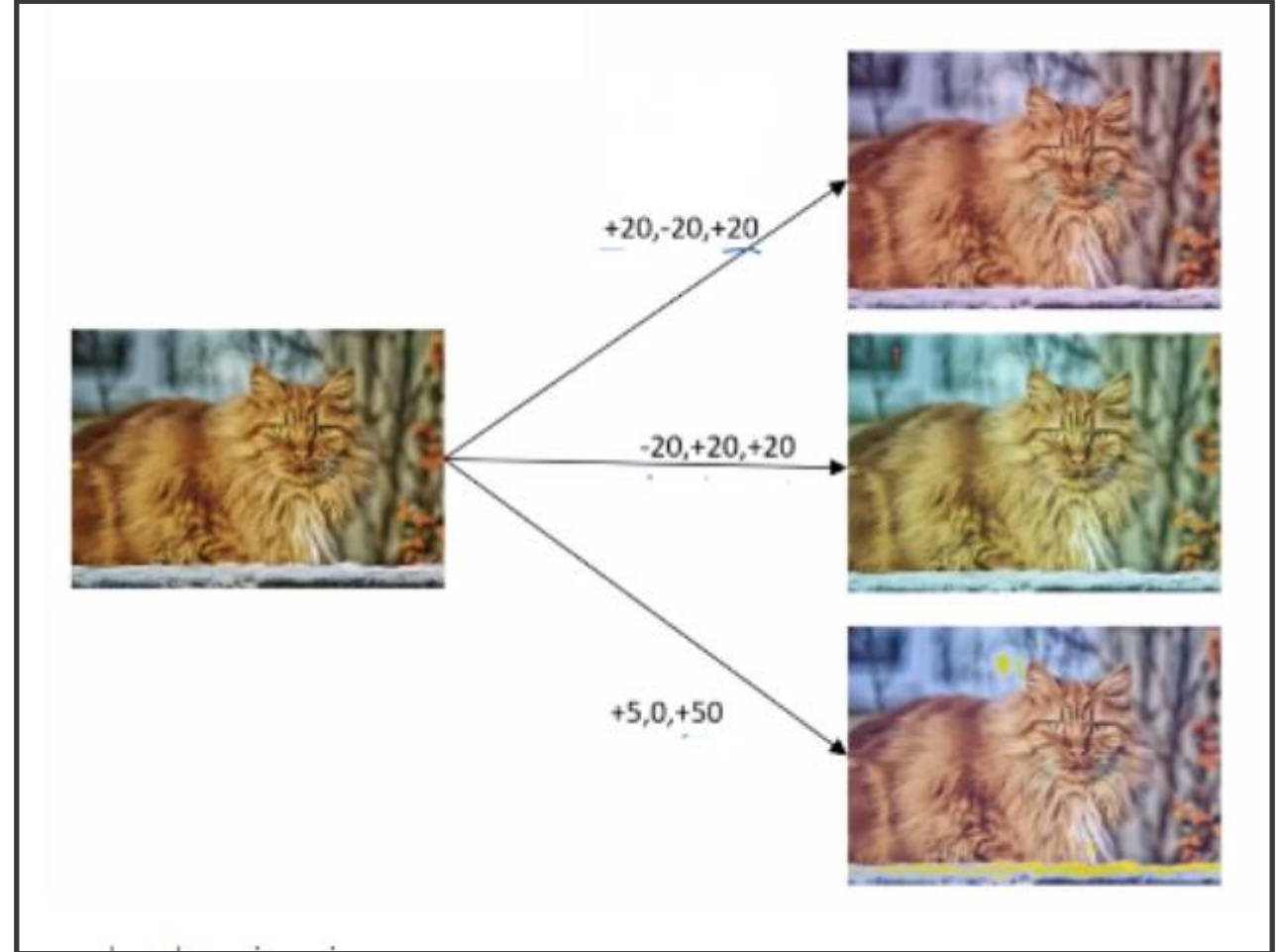
- Görüntüden bir parçanın kesilip alınması



<https://medium.com/@ayyucekizrak/deri%CC%87ne-daha-deri%CC%87ne-evri%C5%9Fimli-sinir-a%C4%9Flar%C4%B1-2813a2c8b2a9>

Veri Artırma (Data Augmentation)

- Görüntünün renk oranlarının değiştirilmesi



<https://medium.com/@ayyucekizrak/deri%CC%87ne-daha-deri%CC%87ne-evri%C5%9Fimli-sinir-a%C4%9Flar%C4%B1-2813a2c8b2a9>