



GLOBAL AI HUB  
AYGAZ YAPAY ZEKAYA GİRİŞ BOOTCAMP  
PROJE ÖDEVİ

Involutional Neural Network

SÜLEYMAN SAMET KAYA  
suleymansametskaya@gmail.com

İBRAHİM GÜLDEMİR  
İbrahimguldemir123@gmail.com

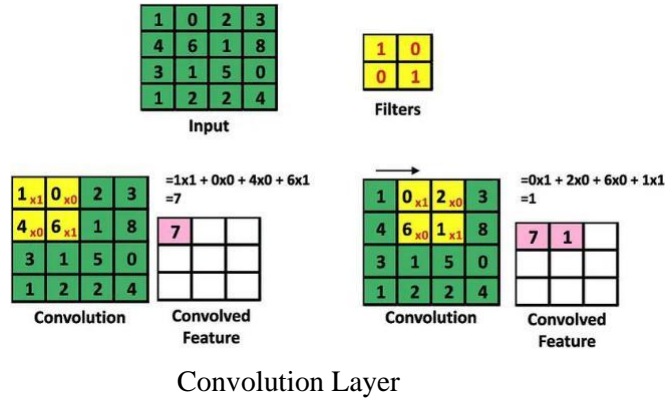
## Convolutional Neural Network:

Convolutional Neural Network (CNN) veya Evrişimsel Sinir Ağları, görüntü tanıma ve sınıflandırma gibi problemlerde sıklıkla kullanılan bir derin sinir ağı türüdür.

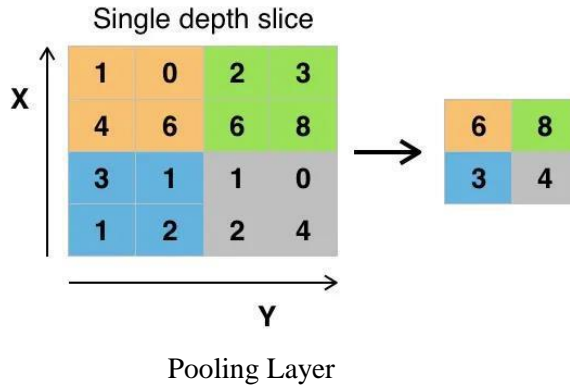
CNN'ler, insan beyindeki sinirsel bağlantıların davranışını taklit ederek görsel özellikleri öğrenme yeteneğine sahiptir.

- **Temel Katmanları:**

- **Convolution Layer:** Bu katman, giriş görüntüsünden özellikler çıkarır ve bu özellikleri bir sonraki katmana aktarır.



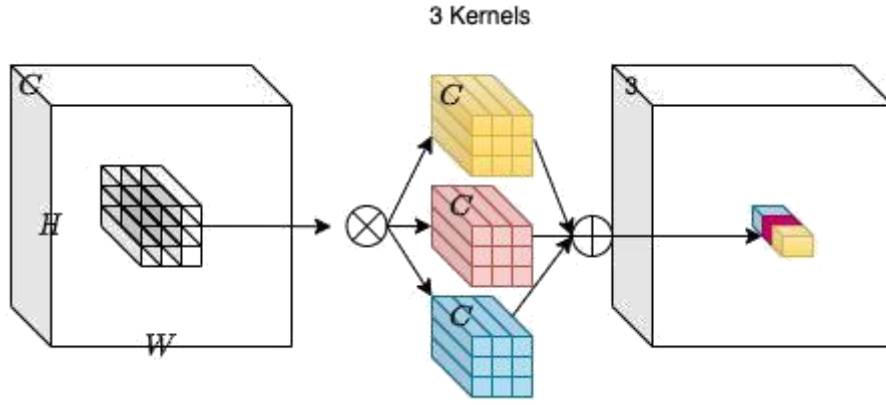
- **Pooling Layer:** Bu katman, özellik haritalarını azaltarak boyut küçültme ve aşırı uyumayı önleme işlevini gerçekleştirir.



- **Fully Connected Layer:** Bu katman, önceki katmanlardan elde edilen özellikleri sınıflandırma veya tahmin için kullanır.

- **Çalışma Prensipleri:**

- CNN'ler, giriş görüntüsünü işleyerek önemli özellikleri öğrenir ve bu özellikleri kullanarak sınıflandırma veya tahmin işlemi gerçekleştirir.
- Convolution ve Pooling katmanları, görüntüdeki yerel özellikler ve düzenli örüntüler hakkında bilgi edinir.
- Fully Connected katmanlar ise, önceki katmanlardan elde edilen özellikleri kullanarak nihai çıktıyı oluşturur.



Boyutları  $H$ ,  $W$  ve  $C_{in}$  olan girdi tensörü  $X$ 'e,  $K \times K \times C_{in}$  boyutunda bir koleksiyon konvolüsyon çekirdeği uygularız.

Çıktı tensörü  $Y$ 'nin boyutları  $H$ ,  $W$  ve  $C_{out}$ 'tur.

Konvolüsyon çekirdekleri mekansal konumdan bağımsızdır (girdinin mekansal konumuna dayanmazlar) ve kanal özgürlüğüne sahiptir (çıktı tensöründeki her kanal belirli bir konvolüsyon filtresine karşılık gelir).

- $K$ , konvolüsyon veya involution çekirdeklerinin boyutunu temsil eder. Örneğin, bir çekirdek boyutu  $3 \times 3$  ise,  $K = 3$  olur.
- $C_{in}$ , girdi kanal sayısını ifade eder. Bu, girdi verisinin her bir özelliğinin (örneğin, renk kanalları gibi) sayısını belirtir.
- Bu terim, bir çekirdeğin boyutunu ve girdi verisinin kanal sayısını birleştirir. Örneğin,  $3 \times 3 \times 64$  bir çekirdek, 64 kanallı bir girdi tensörü üzerinde işlem yapar.

### ***Involutional Neural Network:***

Involutional Neural Network (INV-NN), geleneksel Convolutional Neural Network (CNN) mimarisine dayanan ancak farklı bir tür evrişim katmanı kullanan bir yapay sinir ağıdır. Bu ağ türü, özellikle 3 boyutlu verilerle çalışırken, örneğin video veya vokal verileri gibi, daha etkili olabilir.

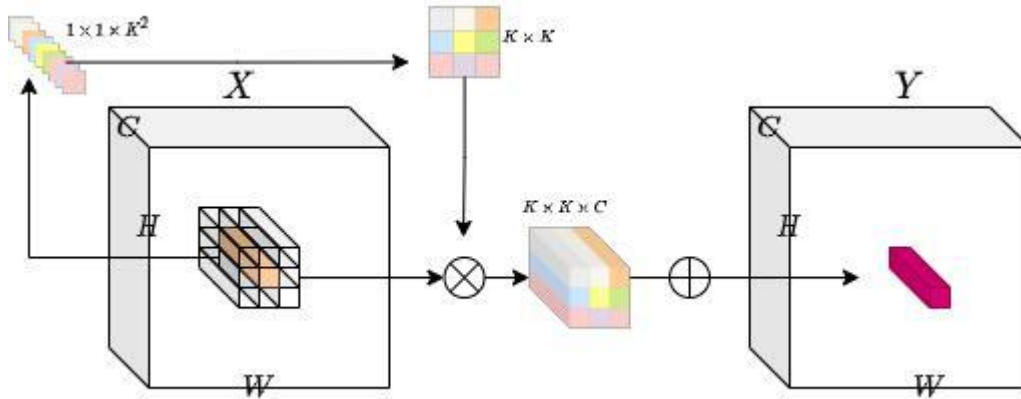
INV-NN'nin temel bileşeni olan "involution layer" veya involution katmanı, CNN'deki evrişim katmanına benzer bir şekilde girdi verisinden özellik haritaları çıkarmak için kullanılır. Ancak, evrişim işleminden farklı olarak, involution katmanı adı verilen ve genellikle daha küçük boyutlu filtreler kullanılan bir işleme dayanır. INV-NN, bu involution katmanlarını bir araya getirerek derin bir ağ yapısı oluşturur.

- **Temel Katmanları:**
  - **Involution Layer (Involution Katmanı):** INV-NN'nin temel yapı taşıdır. Involution katmanı, girdi verisinden özellik haritalarını çıkarmak için kullanılır. Geleneksel evrişim katmanlarına benzer şekilde işlev görür, ancak evrişim yerine daha küçük boyutlu filtrelerle işlem yapar.
  - **Aktivasyon Katmanları:** Bu katmanlar, özellik haritalarına bir aktivasyon fonksiyonu uygular. En yaygın olarak kullanılan aktivasyon fonksiyonları ReLU (Rectified Linear Unit), sigmoid ve tanh'dır. Bu katmanlar, modelin öğrenme yeteneğini artırarak ve ağı doğrusallığı kırarak non-lineerlik ekler.
  - **Toplama (Pooling) Katmanları:** Bu katmanlar, özellik haritalarını küçültmek ve böylece hesaplama maliyetini azaltmak için kullanılır. Max pooling ve average pooling gibi farklı pooling stratejileri mevcuttur. Bu katmanlar ayrıca ağı translasyon invariansını artırabilir.

- **Tam Bağlantılı (Fully Connected) Katmanlar:** Bu katmanlar, özellik haritalarını düzleştirir ve ardından sınıflandırma veya regresyon gibi son işlemleri gerçekleştirmek için kullanılır. Tipik olarak, çok katmanlı bir perceptron (MLP) olarak da adlandırılan bir yapay sinir ağı içerirler.
- **Normalleştirme Katmanları:** Batch normalizasyonu, layer normalizasyonu ve instance normalizasyonu gibi farklı normalizasyon teknikleri, modelin daha hızlı ve daha istikrarlı bir şekilde eğitilmesine yardımcı olabilir.
- **Düşürme (Dropout) Katmanları:** Bu katmanlar, aşırı uyum (overfitting) riskini azaltmak için ağıın öğrenme kapasitesini kontrol etmek için kullanılır. Belirli bir olasılıkla, belirli bir katmandaki nöronların çıkarılmasını sağlarlar.

#### • Çalışma Prensipleri:

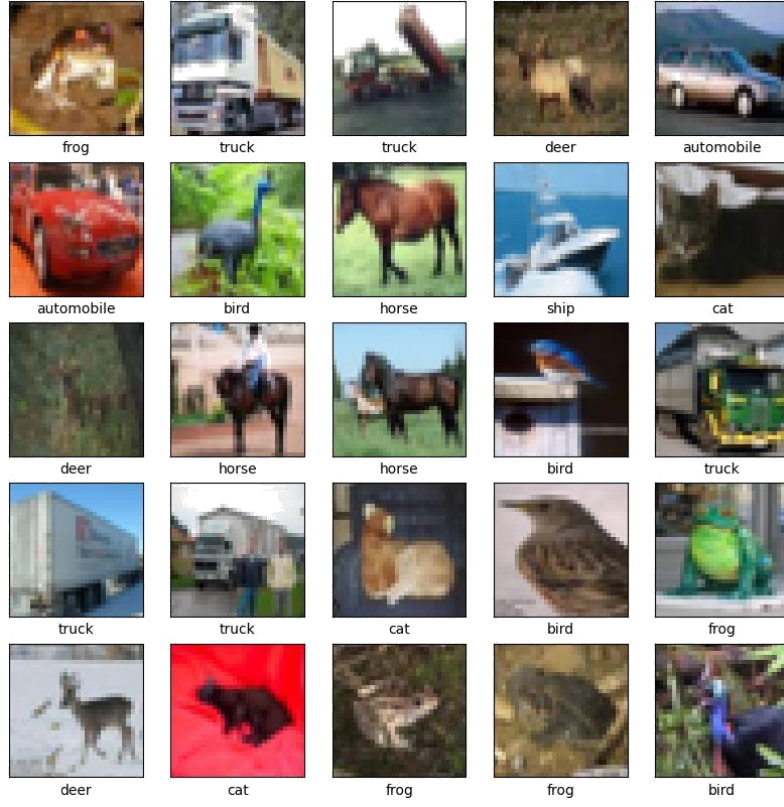
- Involution, hem mekansal konuma özgü hem de kanal bağımsız olmayı amaçlar. Zorluk, her mekansal konum için sabit sayıda involution çekirdeği oluşturmak ve aynı zamanda değişken çözünürlüklü girdi tensörlerini ele alabilmektir. Bu sorunu çözmek için yazarlar, her çekirdeği belirli mekansal konumlara bağlı olarak oluşturmayı önerir. Çekirdek oluşturma süreci, involution'un değişken çözünürlüklü girdileri işleyebilmesini sağlar.
- Giriş verisinden, her piksel için involution kanalları (IC) üretilir. Bu IC'ler, pikselin konumuna ve özelliklerine göre dinamik filtreler oluşturmak için kullanılır.
- Giriş verisi, IC'ler kullanılarak dinamik olarak konvolüsyon işlemine tabi tutulur. Bu sayede her piksel için özelleştirilmiş filtreler uygulanır.
- Elde edilen çıktı, daha sonraki katmanlara aktarılır.



- Böylece Involution, görüntü işleme görevlerinde daha etkili özellik öğrenmesi sağlar ve CNN tabanlı modellerin performansını artırır. Özellikle sıkışık veya karmaşık yapıdaki görüntüler üzerinde başarılı sonuçlar elde edilmiştir.

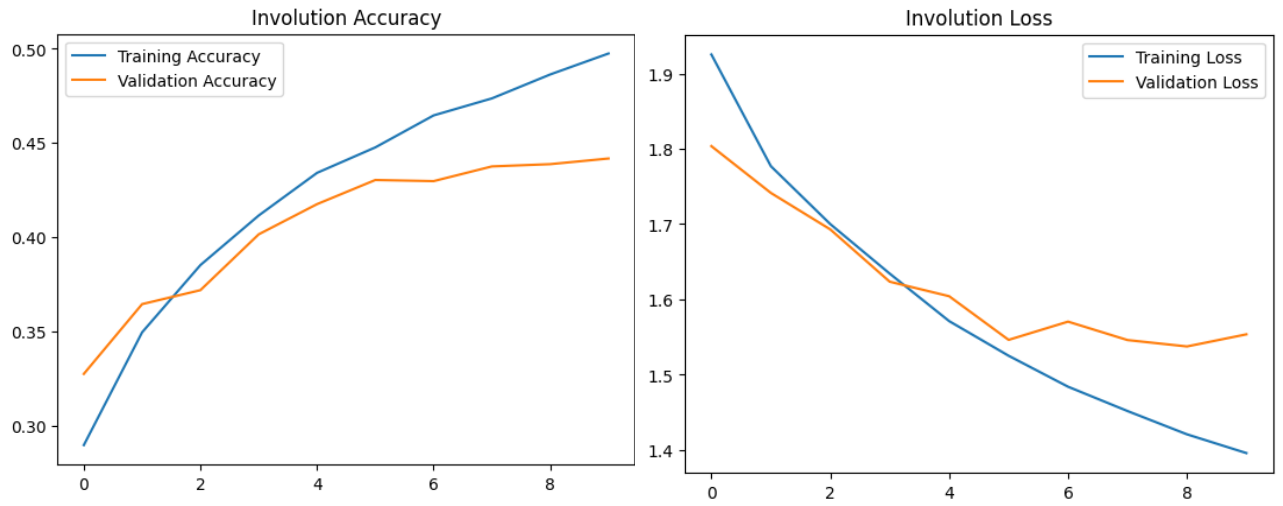
#### Dataset CIFAR-10:

- CIFAR-10, bilgisayarla görme ve derin öğrenme algoritmalarının performansını değerlendirmek için sıkça kullanılan bir veri setidir. Bu veri seti, 10 farklı sınıfa ait toplam 60.000 renkli (RGB) görüntüden oluşur. Her sınıf, farklı nesnelerin görüntülerini içerir: uçaklar, arabalar, kuşlar, kediler, geyikler, köpekler, kurbağalar, atlar, gemiler ve kamyonlar.
- Her sınıf, eşit sayıda örneğe sahip olacak şekilde dengelenmiştir, yani her sınıfın 6.000 görüntüsü vardır. Görüntüler, 32x32 piksel boyutunda küçük ve düşük çözünürlüklüdür.
- CIFAR-10, derin öğrenme modellerinin eğitilmesi, doğrulanması ve test edilmesi için bir standart haline gelmiştir çünkü gerçek dünya verilerine oldukça benzeyen çeşitli nesneler içerir. Bu veri seti üzerinde yapılan çalışmalar, makine öğrenimi ve yapay zekâ alanlarında önemli ilerlemeler sağlamış ve yeni algoritmaların geliştirilmesine katkıda bulunmuştur.



CIFAR-10 Örnek Resimler

*Datasetin Model Eğitimi Sürecinde Doğruluk ve Kayıp Oranları:*



*CIFAR-10 Oranları*

## **KAYNAKÇA**

- <https://www.mdpi.com/2072-4292/13/16/3055>
- <https://keras.io/examples/vision/involution/>
- <https://arxiv.org/abs/2103.06255v2>
- <https://keras.io/api/datasets/cifar10/>
- <https://ieeexplore.ieee.org/abstract/document/9451544>