

Convolutional Neural Networks

Dr. Mauricio Toledo-Acosta

Diplomado Ciencia de Datos con Python

Table of Contents

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

1 Deep Learning

2 Convolutional Neural Networks

- Convolution Kernels
- Pooling Layers

3 Further Applications

Deep Learning

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

El Aprendizaje Profundo (Deep Learning) es una nueva área de investigación de Machine Learning, cuyo objetivo es acercar el aprendizaje automático a la Inteligencia Artificial.

Deep Learning

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

El Aprendizaje Profundo (Deep Learning) es una nueva área de investigación de Machine Learning, cuyo objetivo es acercar el aprendizaje automático a la Inteligencia Artificial.

Resultados teóricos sugieren fuertemente que para aprender el tipo de funciones complicadas que pueden representar abstracciones de alto nivel, es necesario considerar arquitecturas profundas.

Deep Learning

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

El Aprendizaje Profundo (Deep Learning) es una nueva área de investigación de Machine Learning, cuyo objetivo es acercar el aprendizaje automático a la Inteligencia Artificial.

Resultados teóricos sugieren fuertemente que para aprender el tipo de funciones complicadas que pueden representar abstracciones de alto nivel, es necesario considerar arquitecturas profundas.

Las arquitecturas profundas se componen de múltiples niveles de operaciones no lineales, como en las redes neuronales con muchas capas ocultas.

Deep Learning

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

Las arquitecturas profundas son útiles para:

- Reducir la cantidad de elementos de cómputo en el cálculo de funciones complejas.

Deep Learning

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

Las arquitecturas profundas son útiles para:

- Reducir la cantidad de elementos de cómputo en el cálculo de funciones complejas.
- Generar representaciones distribuidas del espacio de entrada (embeddings).

Deep Learning

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

Las arquitecturas profundas son útiles para:

- Reducir la cantidad de elementos de cómputo en el cálculo de funciones complejas.
- Generar representaciones distribuidas del espacio de entrada (embeddings).
- Producir modelos que permitan generalizar funciones con entradas de alta variabilidad.

Usos

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

Algunos de los usos de estas arquitecturas profundas son:

- Reconocimiento automático del habla.

Usos

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

Algunos de los usos de estas arquitecturas profundas son:

- Reconocimiento automático del habla.
- Reconocimiento de imágenes.

Usos

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

Algunos de los usos de estas arquitecturas profundas son:

- Reconocimiento automático del habla.
- Reconocimiento de imágenes.
- Procesamiento de artes visuales.

Usos

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

Algunos de los usos de estas arquitecturas profundas son:

- Reconocimiento automático del habla.
- Reconocimiento de imágenes.
- Procesamiento de artes visuales.
- Procesamiento del lenguaje natural.

Usos

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

Algunos de los usos de estas arquitecturas profundas son:

- Reconocimiento automático del habla.
- Reconocimiento de imágenes.
- Procesamiento de artes visuales.
- Procesamiento del lenguaje natural.
- Sistemas de recomendación.

Usos

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

Algunos de los usos de estas arquitecturas profundas son:

- Reconocimiento automático del habla.
- Reconocimiento de imágenes.
- Procesamiento de artes visuales.
- Procesamiento del lenguaje natural.
- Sistemas de recomendación.
- Análisis de imágenes médicas.

Table of Contents

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

1 Deep Learning

2 Convolutional Neural Networks

- Convolution Kernels
- Pooling Layers

3 Further Applications

CNN

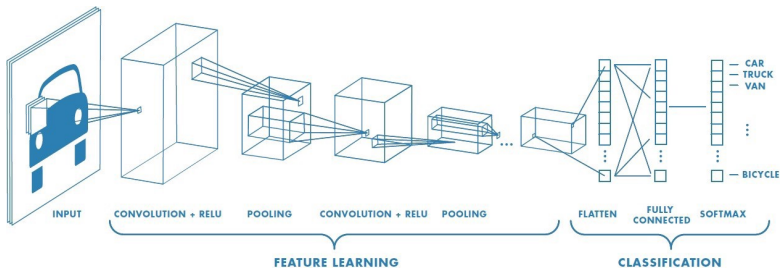
Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications



Source

Why CNNs?

- Una imagen no es más que una matriz de valores de píxeles.
¿Por qué no basta con aplanar la imagen y alimentar la MLP para tareas de clasificación?

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

Why CNNs?

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

- Una imagen no es más que una matriz de valores de píxeles.
¿Por qué no basta con aplanar la imagen y alimentar la MLP para tareas de clasificación?
- Para imágenes muy básicas, este enfoque puede exhibir un desempeño razonable al realizar tareas de clasificación, pero **tendría poca precisión cuando se trata de imágenes complejas que tienen dependencias entre píxeles.**

Why CNNs?

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

- Una imagen no es más que una matriz de valores de píxeles.
¿Por qué no basta con aplanar la imagen y alimentar la MLP para tareas de clasificación?
- Para imágenes muy básicas, este enfoque puede exhibir un desempeño razonable al realizar tareas de clasificación, pero **tendría poca precisión cuando se trata de imágenes complejas que tienen dependencias entre píxeles.**
- **Una CNN es capaz de capturar las dependencias espaciales de una imagen mediante la aplicación de filtros.** La arquitectura se ajusta mejor al conjunto de datos de la imagen gracias a la reducción del número de parámetros implicados.

CNN

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

Convolutional Neural Networks son muy similares a las redes neuronales ordinarias:

- Estan hechas de neuronas que aprenden pesos y sesgos.

CNN

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

Convolutional Neural Networks son muy similares a las redes neuronales ordinarias:

- Estan hechas de neuronas que aprenden pesos y sesgos.
- Cada neurona recibe una entrada, realiza un producto punto y tiene una activación no lineal.

CNN

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

Convolutional Neural Networks son muy similares a las redes neuronales ordinarias:

- Estan hechas de neuronas que aprenden pesos y sesgos.
- Cada neurona recibe una entrada, realiza un producto punto y tiene una activación no lineal.
- La red recibe las imágenes como conjuntos de píxels en un lado y produce scores de clases en el otro.

CNN

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

Convolutional Neural Networks son muy similares a las redes neuronales ordinarias:

- Estan hechas de neuronas que aprenden pesos y sesgos.
- Cada neurona recibe una entrada, realiza un producto punto y tiene una activación no lineal.
- La red recibe las imágenes como conjuntos de píxels en un lado y produce scores de clases en el otro.
- La red tiene una función de pérdida.

Convolutional Neural Networks son muy similares a las redes neuronales ordinarias:

- Estan hechas de neuronas que aprenden pesos y sesgos.
- Cada neurona recibe una entrada, realiza un producto punto y tiene una activación no lineal.
- La red recibe las imágenes como conjuntos de píxels en un lado y produce scores de clases en el otro.
- La red tiene una función de pérdida.
- Los pesos se actualizan con descenso de gradiente u otros optimizadores.

Layers of a CNN

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

Una red CNN se compone principalmente de tres tipos de capas:

- Capas convolucionales (convolutional).

Layers of a CNN

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

Una red CNN se compone principalmente de tres tipos de capas:

- Capas convolucionales (convolutional).
- Capas de pooling (pooling)

Layers of a CNN

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

Una red CNN se compone principalmente de tres tipos de capas:

- Capas convolucionales (convolutional).
- Capas de pooling (pooling)
- Capas totalmente conectadas (fully connected).

CNN

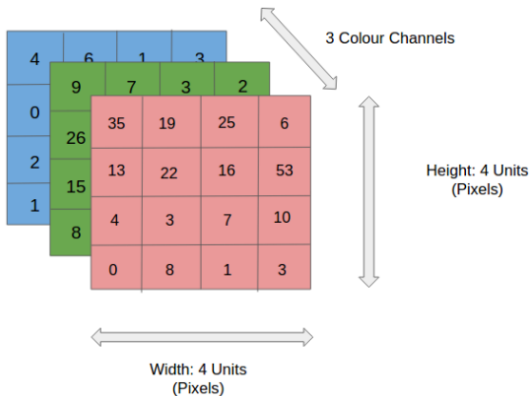
Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications



Source

CNN

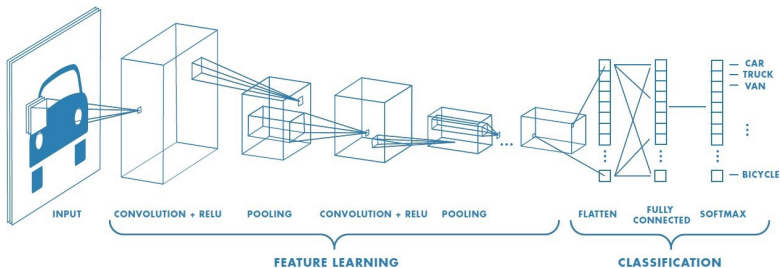
Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications



Convolution Kernels

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

- También llamada matriz de convolución o máscara.

Convolution Kernels

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

- También llamada matriz de convolución o máscara.
- Esta matriz es utilizada para transformar los valores de la imagen por medio de los valores del kernel.

Convolution Kernels

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

- También llamada matriz de convolución o máscara.
- Esta matriz es utilizada para transformar los valores de la imagen por medio de los valores del kernel.
 - Es cuadrada y pequeña (3×3 , 5×5).

Convolution Kernels

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

- También llamada matriz de convolución o máscara.
- Esta matriz es utilizada para transformar los valores de la imagen por medio de los valores del kernel.
 - Es cuadrada y pequeña (3×3 , 5×5).
 - Cuanto más grande es la matriz, más información local se pierde.

Convolution Kernels

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

- También llamada matriz de convolución o máscara.
- Esta matriz es utilizada para transformar los valores de la imagen por medio de los valores del kernel.
 - Es cuadrada y pequeña (3×3 , 5×5).
 - Cuanto más grande es la matriz, más información local se pierde.
- Permite efectos de *área* como desenfoque, nitidez y detección de bordes.

Convolution Kernels

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

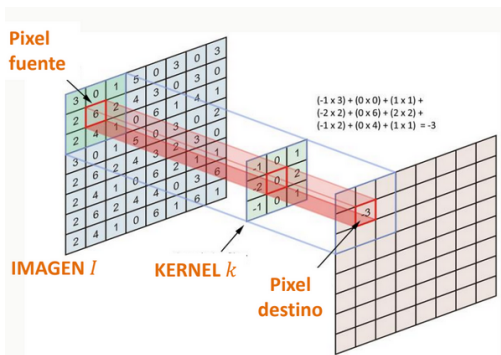
Further
Applications

- También llamada matriz de convolución o máscara.
- Esta matriz es utilizada para transformar los valores de la imagen por medio de los valores del kernel.
 - Es cuadrada y pequeña (3×3 , 5×5).
 - Cuanto más grande es la matriz, más información local se pierde.
- Permite efectos de *área* como desenfoque, nitidez y detección de bordes.
- No es una multiplicación de matrices.

Convolution Kernels

Al aplicar el kernel de convolución k a una entrada (i, j) de la imagen I , esta entrada se transforma en

$$I_{i,j} = \sum_{x,y=1}^n I_{x-i,y-j} k_{x,y}$$



Convolution

Convolutional
Neural
Networks

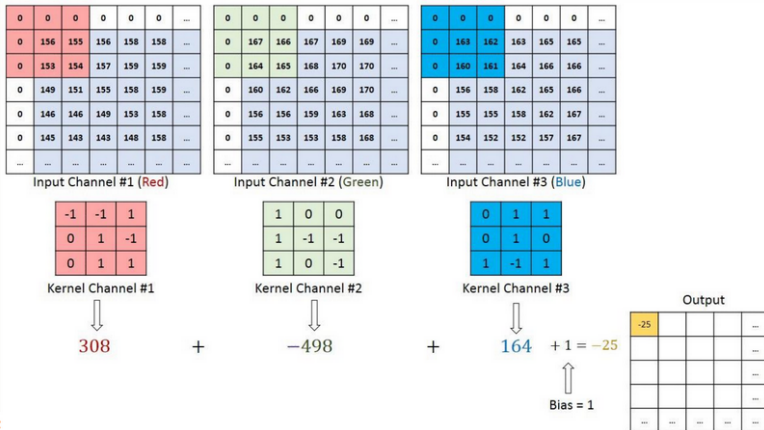
Deep Learning

Convolutional
Neural
Networks

Convolution Kernels

Pooling Layers

Further
Applications



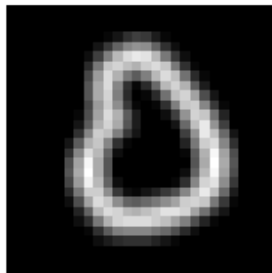
Convolution: Examples

$$K = \frac{1}{9} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Before



After



State of the Art: Edge Detection

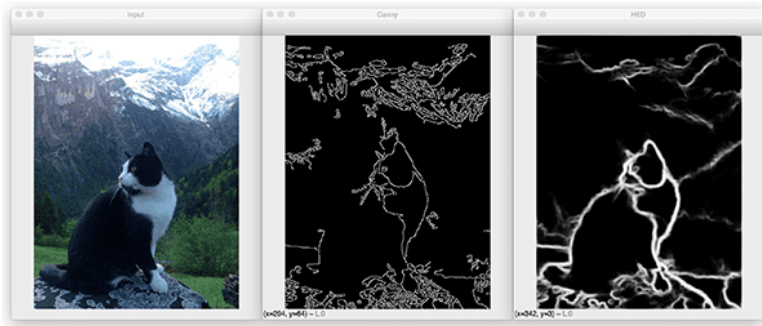
Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications



Holistically-Nested Edge Detection

Pooling Layers

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

- Uno de los principales objetivos de las CNN es aprender los filtros (kernels).

Pooling Layers

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

- Uno de los principales objetivos de las CNN es aprender los filtros (kernels).
- Los filtros sirven para detectar rasgos discriminativos en imágenes.

Pooling Layers

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

- Uno de los principales objetivos de las CNN es aprender los filtros (kernels).
- Los filtros sirven para detectar rasgos discriminativos en imágenes.
- El problema con este enfoque es que el proceso es que es sensible a la ubicación donde se encuentren estos rasgos.

Pooling Layers

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

- Uno de los principales objetivos de las CNN es aprender los filtros (kernels).
- Los filtros sirven para detectar rasgos discriminativos en imágenes.
- El problema con este enfoque es que el proceso es que es sensible a la ubicación donde se encuentren estos rasgos.
- Una solución a esto es subsamplear estas salidas para hacerlas más robustas al cambio de posición en la imagen.

Pooling Layers

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

- Uno de los principales objetivos de las CNN es aprender los filtros (kernels).
- Los filtros sirven para detectar rasgos discriminativos en imágenes.
- El problema con este enfoque es que el proceso es que es sensible a la ubicación donde se encuentren estos rasgos.
- Una solución a esto es subsamplear estas salidas para hacerlas más robustas al cambio de posición en la imagen.
- Aquí es donde entran las capas de **pooling**. El pooling se aplica después de una capa de convolución.

Pooling

La operación de pooling consiste en subsamplear la imagen de entrada. Esta operación se especifica, en vez de aprenderse.

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels

Pooling Layers

Further
Applications

Pooling

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

La operación de pooling consiste en subsamplear la imagen de entrada. Esta operación se especifica, en vez de aprenderse. Las dos maneras típicas que se usan son:

- Average Pooling: Calcula el valor promedio para cada porción de la imagen.
- Maximum Pooling (Max Pooling): Calcula el valor máximo para cada porción de la imagen.

Pooling

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

La operación de pooling consiste en subsamplear la imagen de entrada. Esta operación se especifica, en vez de aprenderse. Las dos maneras típicas que se usan son:

- Average Pooling: Calcula el valor promedio para cada porción de la imagen.
- Maximum Pooling (Max Pooling): Calcula el valor máximo para cada porción de la imagen.

El tamaño de la operación de pooling es más pequeña que el tamaño de la imagen; casi siempre es de 2×2 píxeles con un paso de 2 pixels. En este caso, se reduce el tamaño a la mitad. Por ejemplo, una capa de pooling aplicada a una imagen de 6×6 resultará en una salida de 3×3 .

Pooling

Max Pooling

29	15	28	184
0	100	70	38
12	12	7	2
12	12	45	6

2 x 2
pool size

100	184
12	45

Average Pooling

31	15	28	184
0	100	70	38
12	12	7	2
12	12	45	6

2 x 2
pool size

36	80
12	15

Example

Convolutional
Neural
Networks

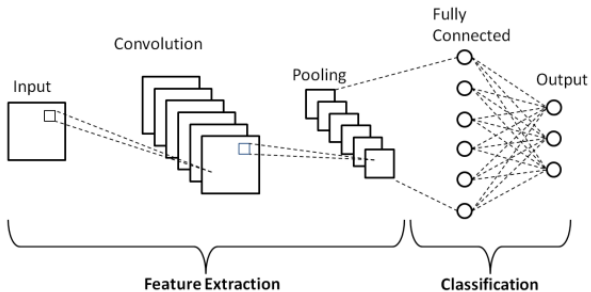
Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

Consideremos el siguiente ejemplo de CNN



Example

- La capa INPUT $[32 \times 32 \times 3]$ tendrá los valores de los pixeles de la imagen, en este caso la imagen es de tamaño 32×32 , con tres canales R,G,B.

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

Example

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

- La capa INPUT $[32 \times 32 \times 3]$ tendrá los valores de los pixeles de la imagen, en este caso la imagen es de tamaño 32×32 , con tres canales R,G,B.
- La capa CONV calculará la convolución con cada filtro. Esto resultará en un volumen de imágenes $[32 \times 32 \times 12]$ si es que usamos 12 filtros. A cada salida aplicamos la activación RELU, seguimos teniendo $([32 \times 32 \times 12])$.

Example

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

- La capa INPUT $[32 \times 32 \times 3]$ tendrá los valores de los pixeles de la imagen, en este caso la imagen es de tamaño 32×32 , con tres canales R,G,B.
- La capa CONV calculará la convolución con cada filtro. Esto resultará en un volumen de imágenes $[32 \times 32 \times 12]$ si es que usamos 12 filtros. A cada salida aplicamos la activación RELU, seguimos teniendo $([32 \times 32 \times 12])$.
- La capa POOL subsamplea a lo largo del ancho y largo de las imágenes, el resultado es de tamaño $[16 \times 16 \times 12]$.

Example

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

- La capa INPUT $[32 \times 32 \times 3]$ tendrá los valores de los pixeles de la imagen, en este caso la imagen es de tamaño 32×32 , con tres canales R, G, B.
- La capa CONV calculará la convolución con cada filtro. Esto resultará en un volumen de imágenes $[32 \times 32 \times 12]$ si es que usamos 12 filtros. A cada salida aplicamos la activación RELU, seguimos teniendo $([32 \times 32 \times 12])$.
- La capa POOL subsamplea a lo largo del ancho y largo de las imágenes, el resultado es de tamaño $[16 \times 16 \times 12]$.
- El volumen de datos anterior se aplanan y entra a la red FC, la cuál calculará los scores de clase resultando en un volumen de tamaño $[1 \times 1 \times 10]$.

Classification

Convolutional
Neural
Networks

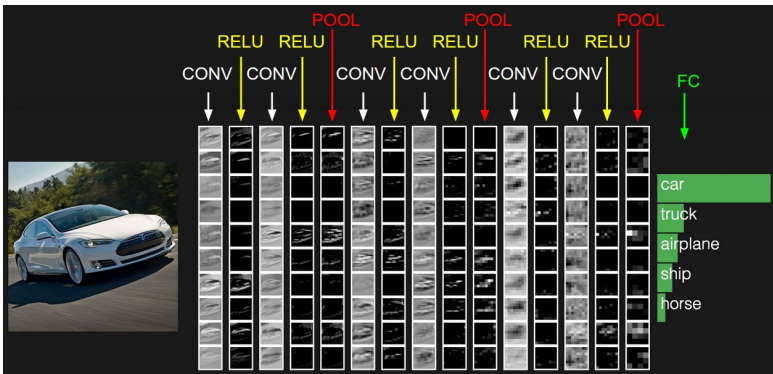
Deep Learning

Convolutional
Neural
Networks

Convolution Kernels

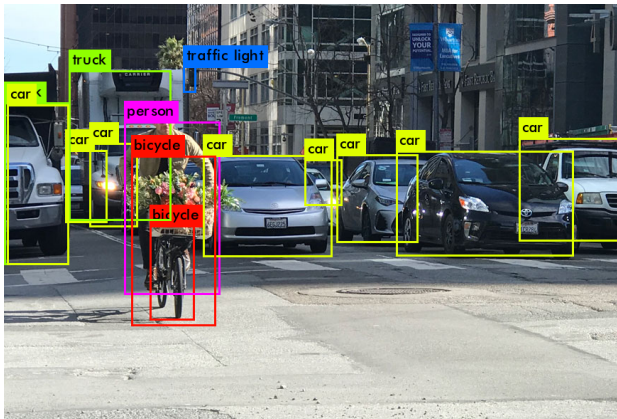
Pooling Layers

Further
Applications



State of the Art: Real-Time Object Detection

YOLO: You only look once



Video demo
Source

Table of Contents

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

1 Deep Learning

2 Convolutional Neural Networks

- Convolution Kernels
- Pooling Layers

3 Further Applications

Natural Language Processing

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

Yesterday I had a very good day with my son

Text message

Yesterday had very good day son

Text without stop-words

-0.5 -0.3 0.2 1 0.7 0.9

Word scores

$$\begin{pmatrix} 0.23 \\ -0.32 \\ \dots \\ 0.66 \end{pmatrix} \begin{pmatrix} 0.52 \\ 0.12 \\ \dots \\ -0.07 \end{pmatrix} \begin{pmatrix} -0.42 \\ 0.33 \\ \dots \\ -0.17 \end{pmatrix} \begin{pmatrix} -0.42 \\ -0.58 \\ \dots \\ -0.19 \end{pmatrix} \begin{pmatrix} -0.32 \\ 0.53 \\ \dots \\ 0.09 \end{pmatrix}$$

Word representations

$$\begin{pmatrix} 0.23 & 0.52 & -0.42 & -0.42 & -0.32 \\ -0.32 & 0.12 & 0.33 & -0.58 & 0.53 \\ \dots & \dots & \dots & \dots & \dots \\ 0.66 & -0.07 & -0.17 & -0.19 & 0.09 \end{pmatrix}$$

Matrix representation for the text

Natural Language Processing

Convolutional
Neural
Networks

Deep Learning

Convolutional
Neural
Networks

Convolution Kernels
Pooling Layers

Further
Applications

