

## REDES NEURONALES ARTIFICIALES – DEEP LEARNING

### CALSIFICADOR CON DEEP LEARNING – REDES CONVOLUCIONALES

LAURA DIAZ DÁVILA

REDES NEURONALES  
APRENDIZAJE SUPERVISADO – UNIDIRECCIONALES –  
MULTICAPA

EL PERCEPTRON GENERALIZADO

- 1. MLP BACK PROPAGATION**
- 2. DEEP LEARNING**

---

## Principales tipos de DNN

- Modelos predictivos "discriminatorios" para la clasificación (o regresión) con entrenamiento principalmente supervisado:

CNN - Convolutional Neural Network (o ConvNet)

FC DNN - Fully Connected DNN (MLP con al menos dos niveles ocultos)

HTM - Hierarchical Temporal Memory

- Entrenamiento no sobredesarrollado (modelos "generativos" entrenados en reconstruir la entrada, útil para pre-entrenamiento de otros modelos y para producir características destacadas):

Stacked (de-noising) Auto-Encoders

RBM - Restricted Boltzmann Machine

DBN - Deep Belief Networks

- Patrones recurrentes (utilizados para secuencias, reconocimiento de voz, análisis de sentimiento, procesamiento del lenguaje natural, ...):

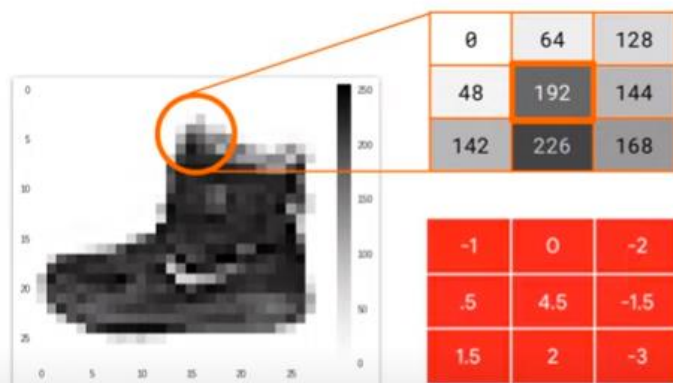
RNN - Recurrent Neural Network

LSTM - Long Short-Term Memory

- Aprendizaje de refuerzo (para aprender comportamientos):

Deep Q-Learning

# CNN FILTROS



Current pixel value is 192

Consider neighbor values

Filter definition

CURRENT\_PIXEL\_VALUE = 192

NEW\_PIXEL\_VALUE =  $(-1 * 0) + (0 * 64) + (-2 * 128) +$   
 $(.5 * 48) + (4.5 * 192) + (-1.5 * 144) +$   
 $(1.5 * 42) + (2 * 226) + (-3 * 168)$



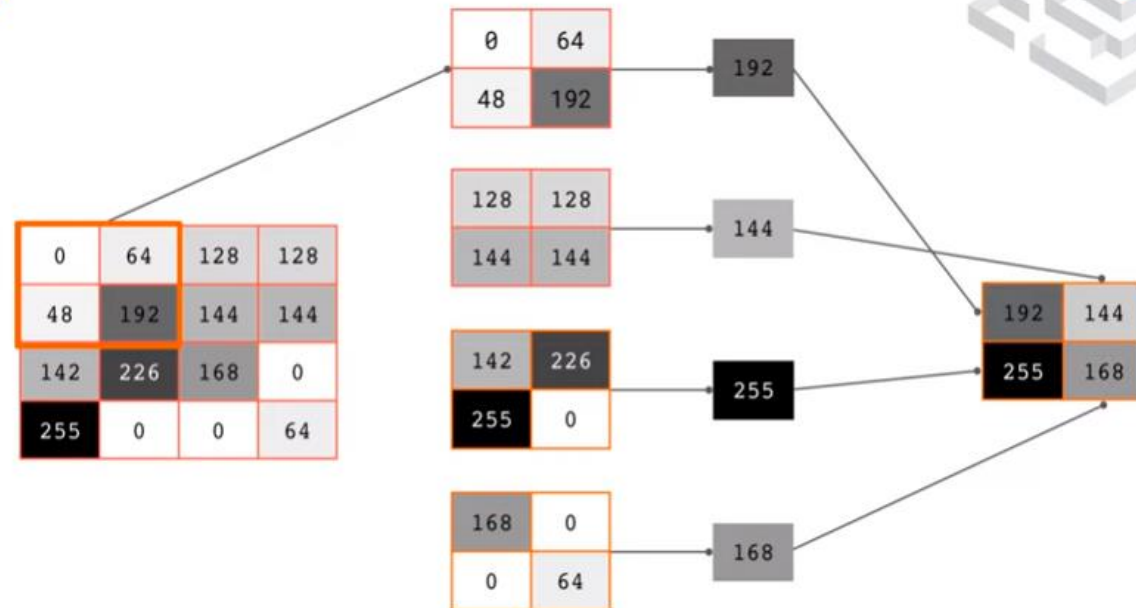
# CNN FILTROS



-1	-2	-1
0	0	0
-1	2	1



# MAX POOLING



# MAX POOLING



Max pooling 2x2



# CONVOLUTIONAL NEURAL NETWORKS



**FILTROS O  
MULTIPLICADORES**

**“SE VAN  
QUEDANDO  
CON LO QUE LES  
INTERESA”**

**RECONOCIMIENTO DE OBJETOS, DE IMÁGENES EN MOVIMIENTO**

**Fuente:** Redes Neuronales Artificiales - VIII Escuela de Verano (2020) –SEMF

<https://www.youtube.com/watch?v=z7Im7GSSQPI>



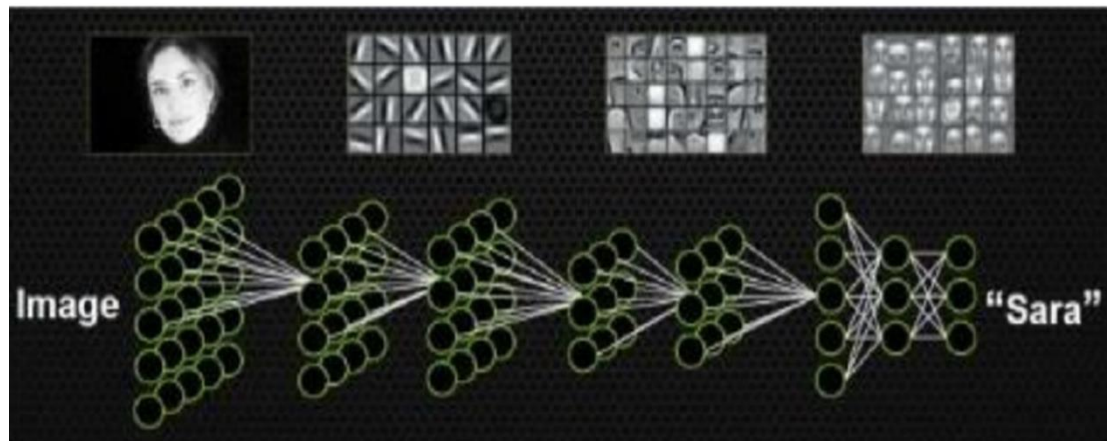
# CONVOLUTION NEURAL NETWORKS

**Procesamiento local:** las neuronas solo están conectadas localmente a neuronas del nivel anterior. Cada neurona luego realiza procesamiento local implicando un fuerte número de reducción de conexiones.

**Pesos compartidos:** los pesos se comparten en grupos. Diferentes neuronas en el mismo nivel, realizan el mismo tipo de procesamiento en diferentes porciones de la entrada. Fuerte reducción en el número de parámetros

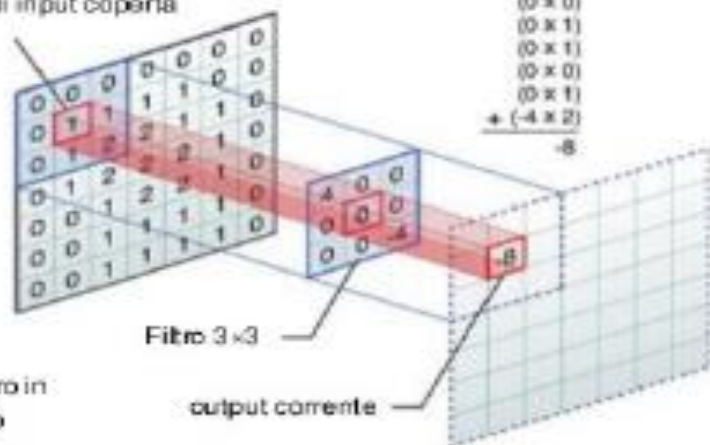
Explícitamente diseñado para procesar imágenes, para lo cual el procesamiento local, los pesos compartidos y la agrupación no solo simplifican el modelo, sino que es más efectivo que los modelos completamente conectados. Se aplican en otros dominios.

La arquitectura de un CNN consiste en una jerarquía de niveles. La capa de entrada está directamente conectada a los píxeles de la imagen, las últimas capas generalmente están completamente conectadas y funcionan como un clasificador MLP, mientras en las capas ocultas intermedias se usan conexiones locales y pesos compartidos



- La convolución es una de las operaciones más importantes de procesamiento de imágenes, a través de la cual se aplican los filtros digitales.
- Un filtro digital (una pequeña máscara 2D de pesos) se desliza sobre las diferentes posiciones de entrada; para cada posición se genera un valor de salida, ejecutando el producto escalar entre la máscara y la porción de la entrada cubierta (ambos tratados como vectores)

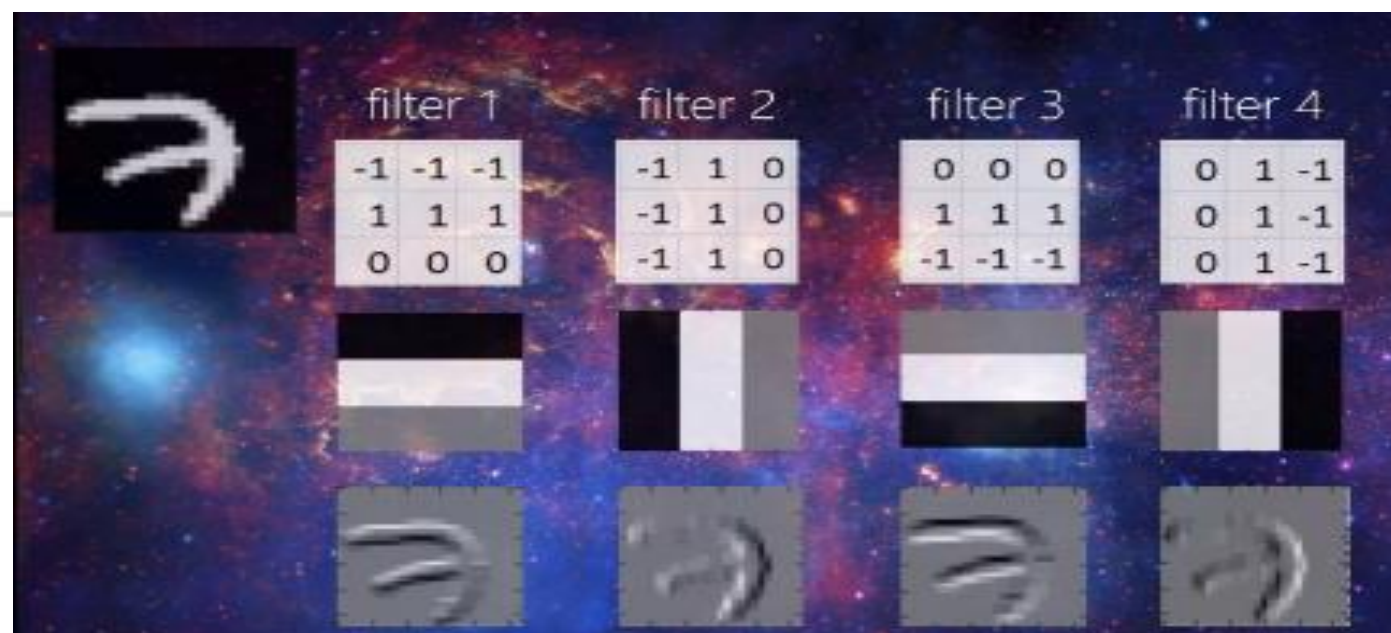
posizione corrente e  
porzione di input coperta  
dal filtro



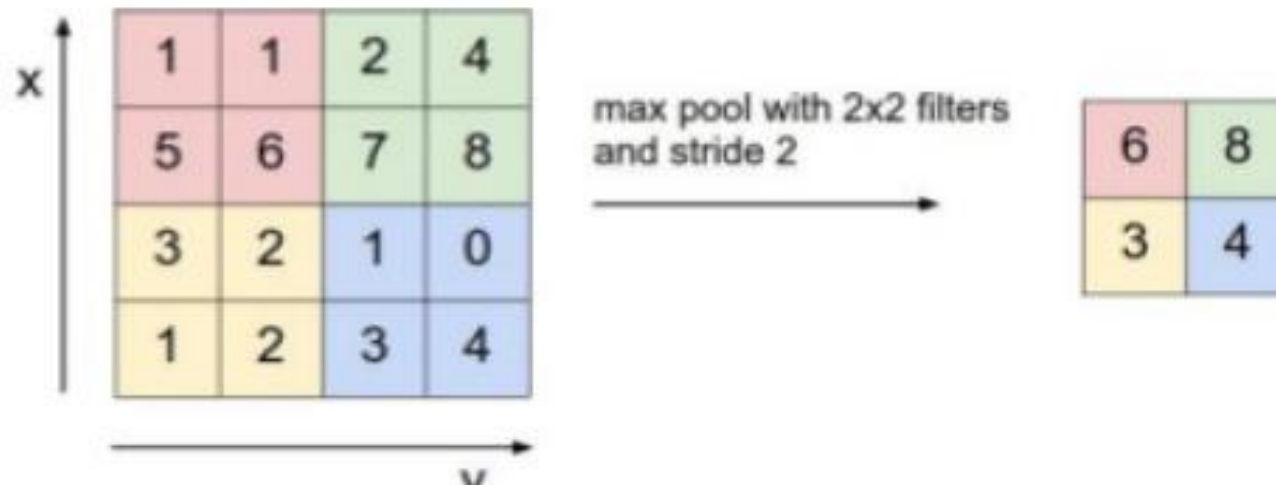
Esempio (animato): pesi del filtro in  
rosso, porzione coperta in giallo

1 <sub>x1</sub>	1 <sub>x0</sub>	1 <sub>x1</sub>	0	0
0 <sub>x0</sub>	1 <sub>x1</sub>	1 <sub>x0</sub>	1	0
0 <sub>x1</sub>	0 <sub>x0</sub>	1 <sub>x1</sub>	1	1
0	0	1	1	0
0	1	1	0	0

4		

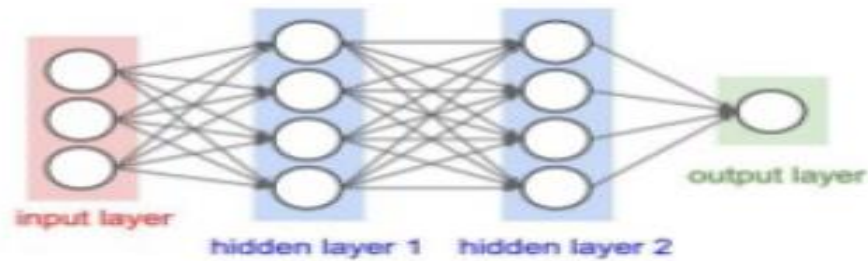


Los operadores de agrupamiento más utilizados son el promedio (Avg) y el máximo (Máx): ambos son "bastante" invariables para traducciones pequeñas. Este tipo de proceso no tiene "parámetros / pesos" para aprender. En el ejemplo, un máximo de agrupación con  $2 \times 2$  filtros y Stride = 2

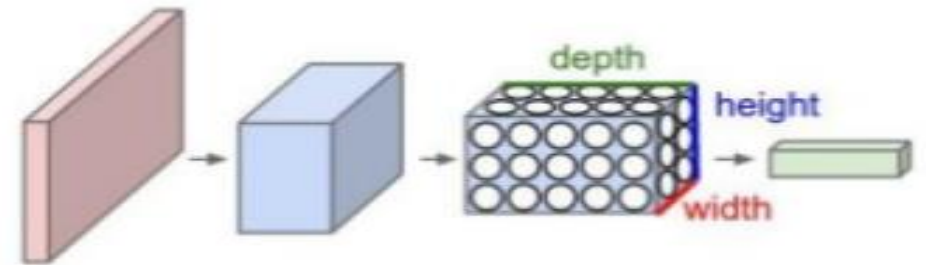


## ***Volúmenes:***

Las neuronas de cada nivel están organizadas en grillas o volúmenes 3D, a diferencia de MLP que tiene una organización lineal de neuronas en niveles

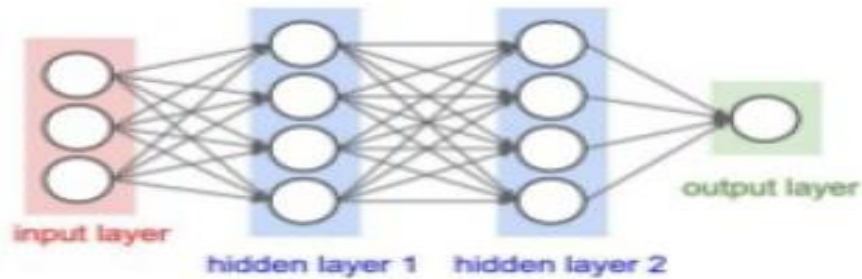


MLP

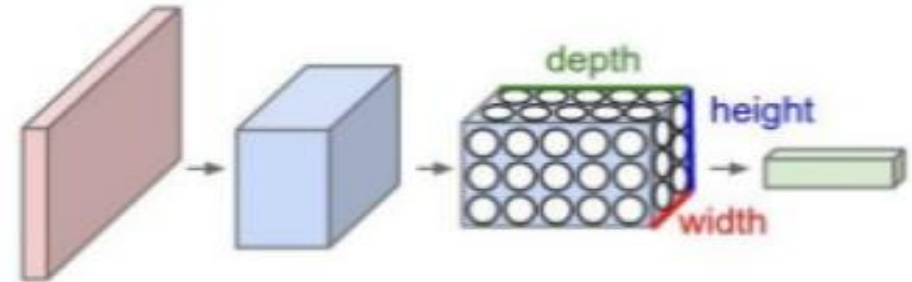


CNN

## ***Volúmenes:***



MLP

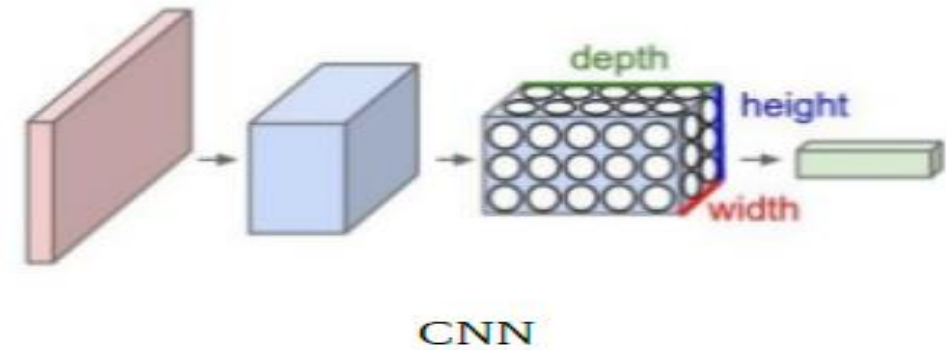
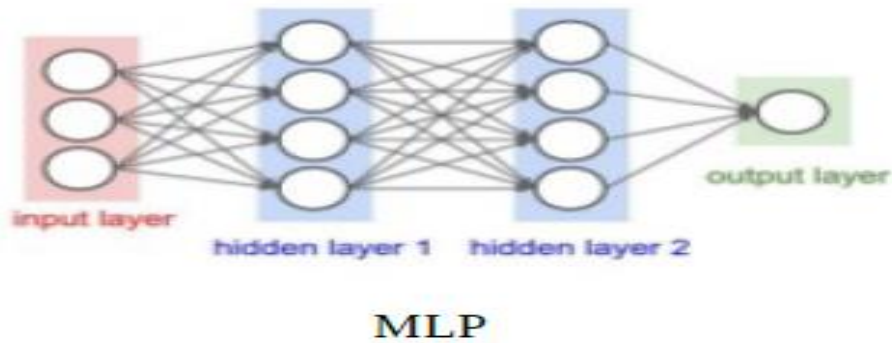


CNN

En el CNN height -width conservan la organización espacial "retinotípica" de la imagen de entrada. La profundidad (depth) de la tercera dimensión identifica las 'Features map'



## ***Volúmenes:***

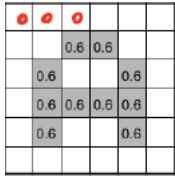
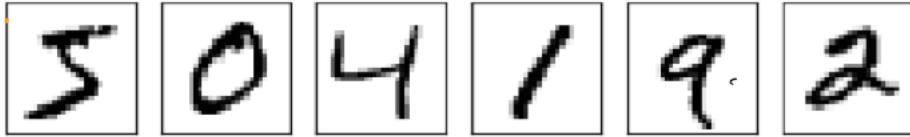


Los pesos se comparten en el nivel del Features map. Las neuronas de la misma Features map procesan diferentes porciones del volumen de entrada de la misma manera. Cada Features map puede verse como el resultado de un filtrado de entrada específico (filtro fijo)



# CNN

CAPA DE ENTRADA O CAMPO RECEPTIVO



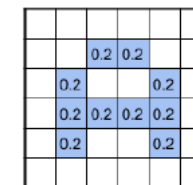
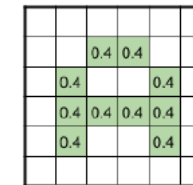
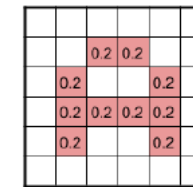
SE PRESERVA LA TOPOLOGÍA

SE NORMALIZA – ESCALA DE GRISES ENTRE 0 Y 255

CANTIDAD DE NEURONAS DE LA CAPA DE ENTRADA:  
 $28 * 28 = 784$  PÍXELES O ELEMENTOS



IMAGEN A COLOR COMPUESTA  
POR TRES CANALES (RGB)



CANTIDAD DE NEURONAS DE LA ENTRADA  
TRIDIMENSIONAL:  
 $(28 * 28) * 3 = 2352$  PÍXELES O ELEMENTOS

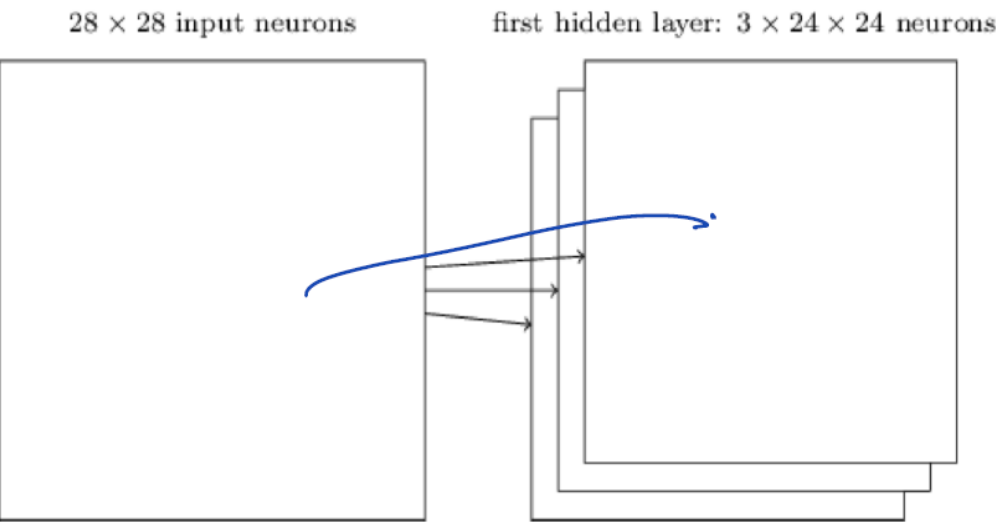


# CNN

## CAPA CONVOLUCIONAL

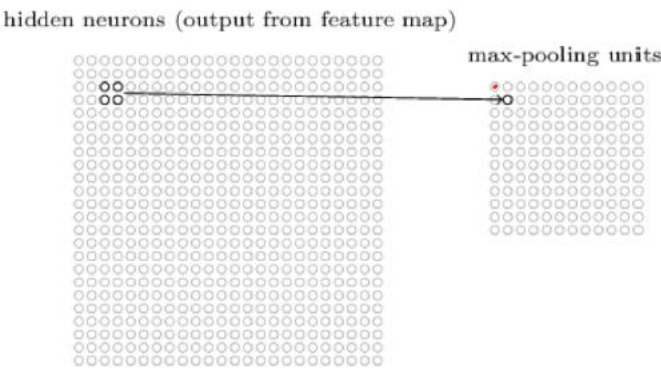
UNA CONVOLUCIÓN ES UN FILTRO: A PARTIR DE DOS FUNCIONES, SE GENERA OTRA “CONVOLUCIONADA”

LA RED APRENDE EL ÚNICO PESO SINÁPTICO QUE REPRESENTA A ESE GRUPO DE NEURONAS, “HIDDEN”

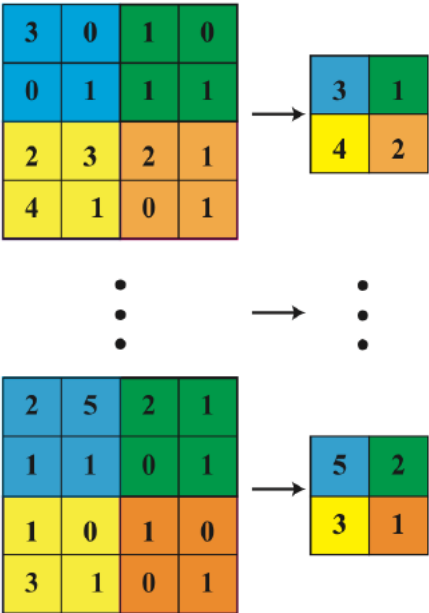


## CAPA POOLING

TOMA LA SALIDA DE LA CONVOLUCIÓN Y LA REDUCE, AGRUPANDO NEURONAS

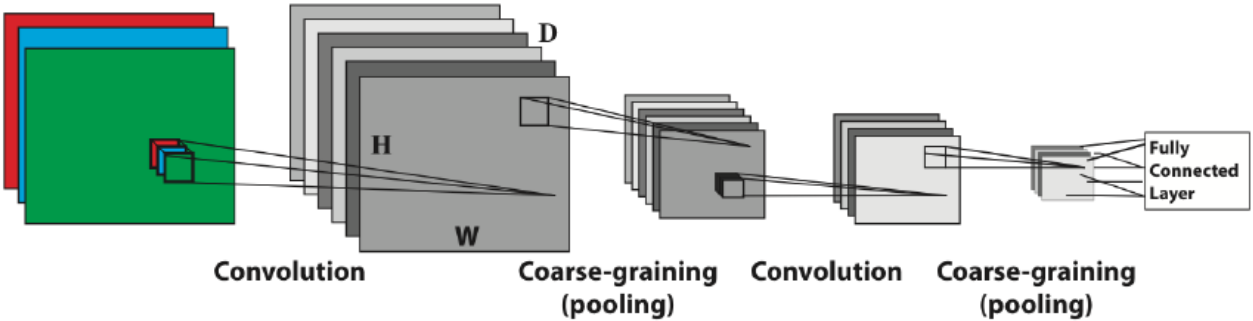
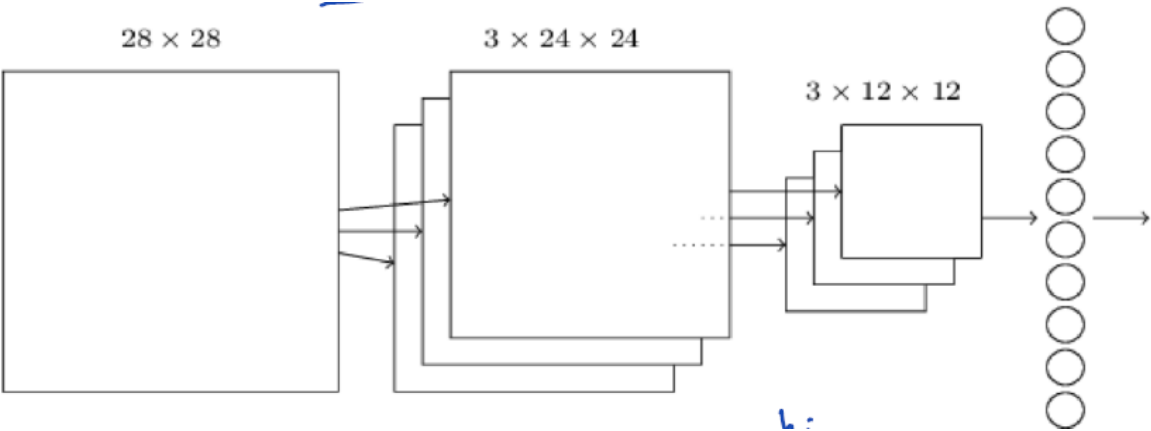


## EJEMPLO DE MAX POOLING



5 0 4 1 9 2

# CNN



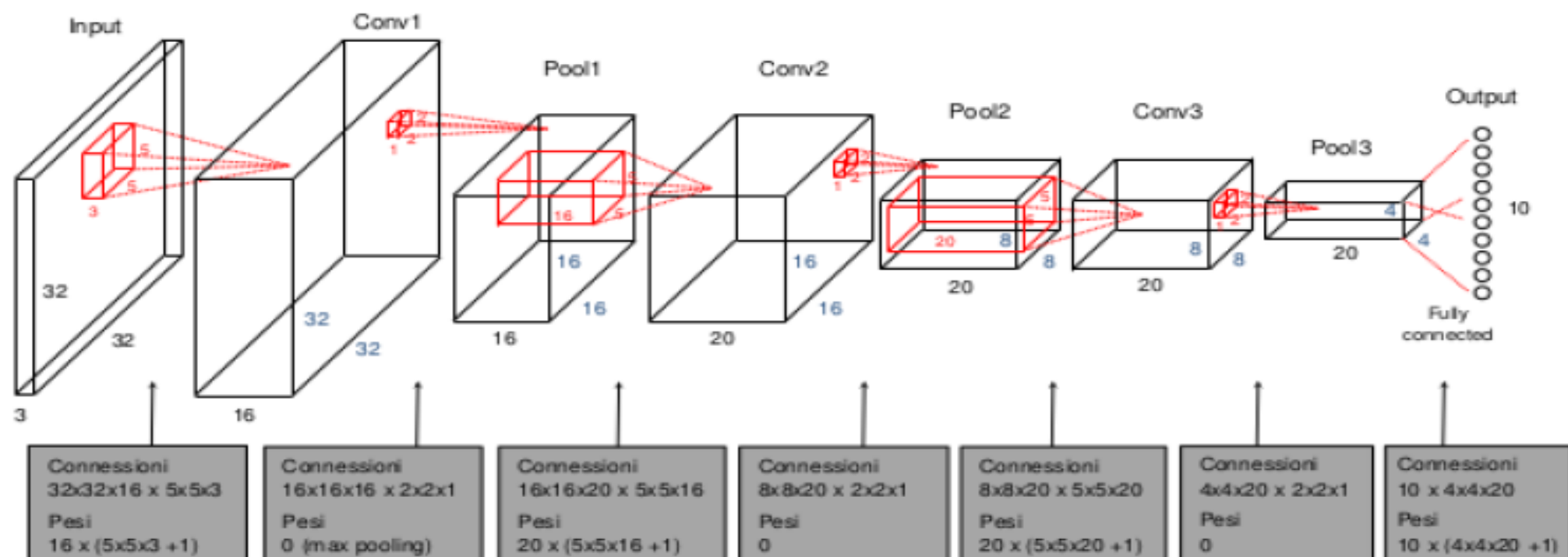
Neuronas totales: 31.562 (incluido el capa de entrada)  
Conexiones totales: 3,942,784  
Pesos totales: 22.466 (incluido bias)

## Ejemplo de arquitectura

### Architettura

- **Input**: Immagini RGB 32x32x3;
- **Conv1**: Filtri:5x5, FeatureMaps:16, stride:1, pad:2, attivazione: Relu
- **Pool1**: Tipo: Max, Filtri 2x2, stride:2
- **Conv2**: Filtri:5x5, FeatureMaps:20, stride:1, pad:2, attivazione: Relu
- **Pool2**: Tipo: Max, Filtri 2x2, stride:2
- **Conv3**: Filtri:5x5, FeatureMaps:20, stride:1, pad:2, attivazione: Relu
- **Pool3**: Tipo: Max, Filtri 2x2, stride:2
- **Output**: Softmax, NumClassi: 10

Neuronas totales: 31.562 (incluido el capa de entrada)  
 Conexiones totales: 3,942,784  
 Pesos totales: 22.466 (incluido bias)



¡GRACIAS!