

# 2- Convolutional Neural Networks (CNN)

By: Ruba  
eng. Esraa Madhi

---

## ماهي (Convolutional Neural Networks (CNN) ؟

هي خوارزميات ذكاء اصطناعي تعتمد على الشبكات العصبية (artificial neural networks (ANNs متعددة الطبقات و تستخرج الميزات features من الصور و الفيديو , قدرة على أداء العديد من المهام مثل تصنيف (classification) و اكتشاف (detection) الكائنات و تجزئتها (segmentation).

### Object detection vs. image segmentation vs. image classification

Object detection



Image segmentation

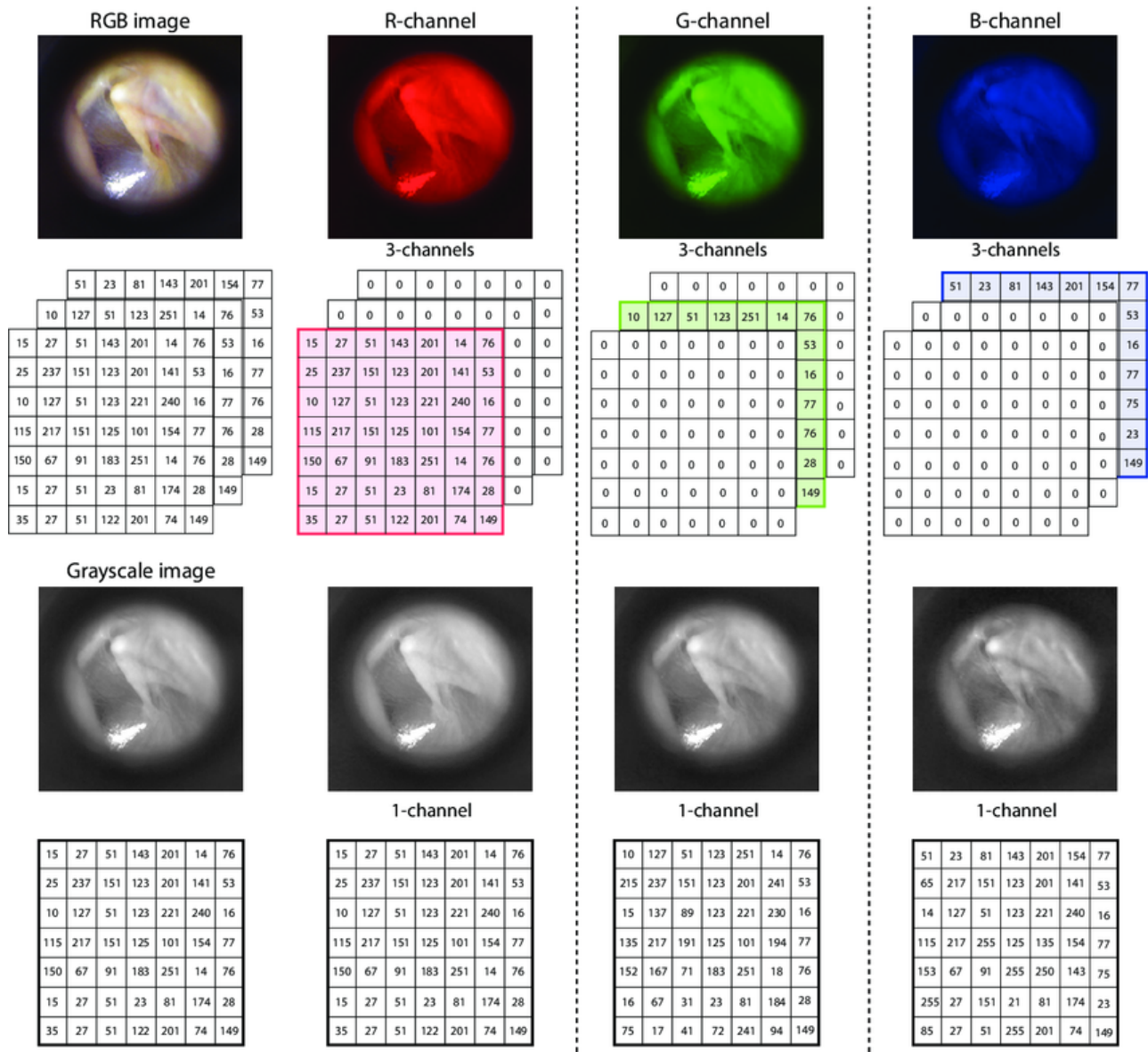


Image classification

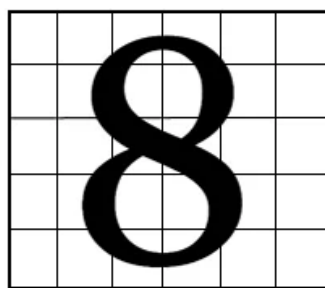


## Image Structure:

في CNN ، يتم تمثيل كل صورة في شكل مجموعة من قيم البكسل.



Real Image of the digit 8



Represented in the form of an array



0	0	1	1	0	0
0	1	0	0	1	0
0	0	1	1	0	0
0	1	0	0	1	0
0	0	1	1	0	0

Digit 8 represented in the form of pixels of 0's and 1's

# بنية (Convolutional Neural Networks (CNN):

تنقسم شبكة CNN إلى قسمين:

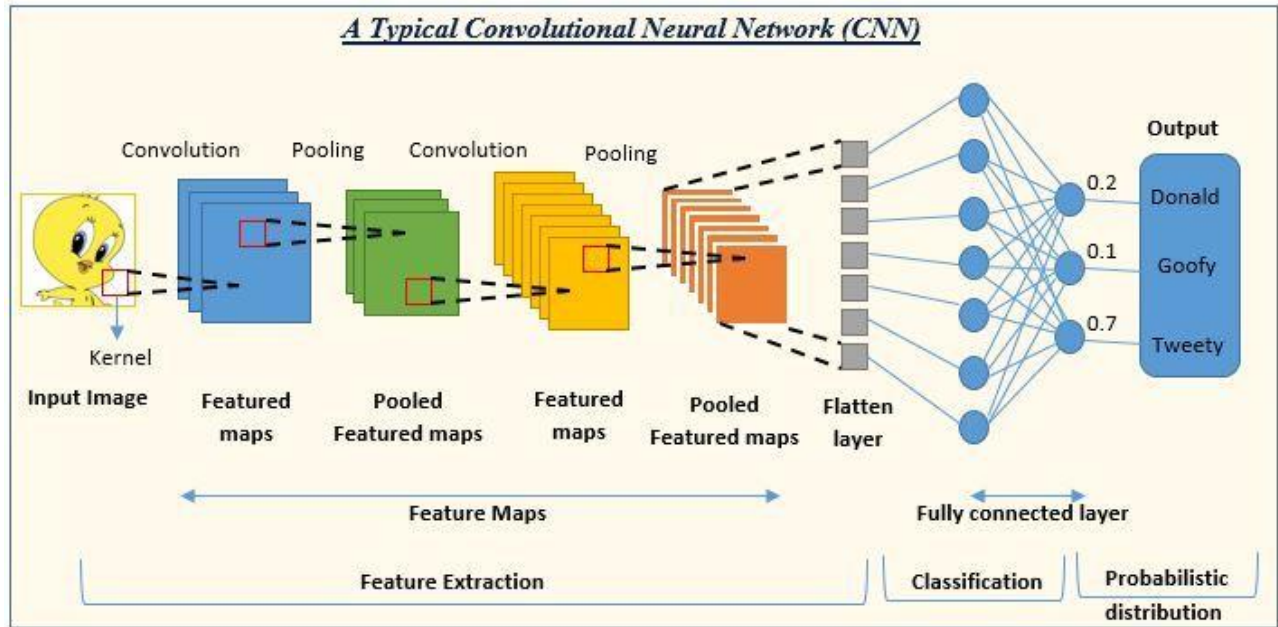
يتعلم أفضل الميزات (features) لاستخراجها من الصور	Convolutional steps:
يتعلم كيفية تصنيف الميزات (features) في فئات مختلفة.	Dense steps

تحتوي convolutional steps عادةً على ثلاث أنواع من الطبقات (Layers) طبقات مخفية متعددة تساعد في استخراج المعلومات من الصورة:

- Convolutional
- Pooling
- Relu

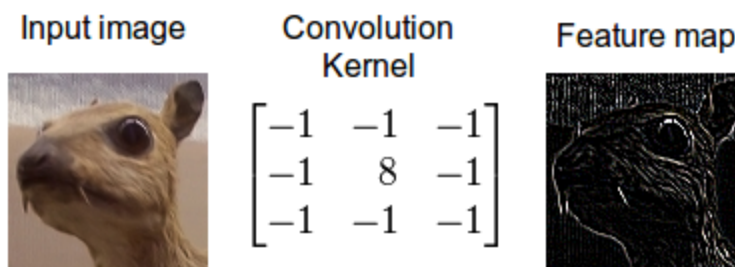
وتحتوي dense steps على طبقتان تساعدان في تصنيف المميزات التي تم استخراجها:

- Flatten
- Fully Connected






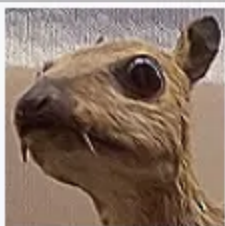
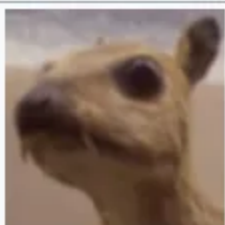
### طبقة الالتفاف (Convolution Layer):

- هي أساس بناء شبكة CNN, هي تتكون من مجموعة من المرشحات (filters) تسمى kernels وعادة ماتكون هذه المرشحات أصغر من الصورة الفعلية.



- An image kernel is a **small matrix** used to **apply effects** like the ones you might find in Photoshop or Gimp, such as blurring, sharpening, outlining or embossing. They're also used in machine learning for **'feature extraction'**, a technique for determining the most important portions of an image.

<https://youtu.be/CXOGvCMLrkA>

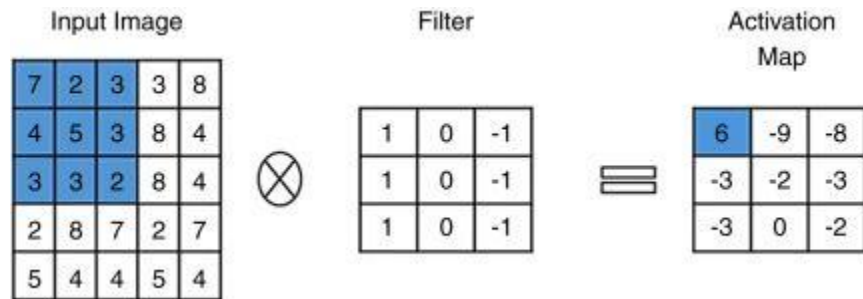
Edge detection	$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$	
	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$	
	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	
Sharpen	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	
Box blur (normalized)	$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	

- Kernels features:**

- كل kernel لها عرض و ارتفاع , العرض × الارتفاع ينتج لنا الوزن (weight).
- تستخدم الأوزان (weights) لإستخراج الميزات من بيانات الإدخال.
- في خطوة التدريب ، تبدأ الأوزان (weights) في kernel بقيم عشوائية .

- **How Convolution Layer works?**

في هذه الطبقة هي حاصل ضرب نقطي (dot product) بين مصفوفتين هما مصفوفة kernel و مصفوفة البيكسل للصورة , يلتف كل مرشح مع الصورة و يقوم بإنشاء خريطة ال-activation, و يتم تكرار هذه العملية لكل عنصر من عناصر الصورة المدخلة.



في المثال التوضيحي نرا أن لدينا صورة بمقاس 5x5 و التي تكون قيم البيكسل فيها إما 0 أو 1, و هناك أيضاً مصفوفة المرشح بأبعاد 3x3.

نقوم بتمرير مصفوفة ال-filter فوق الصورة المدخلة ويتم هنا حساب dot product للحصول على مصفوفة الميزات وهي activation map.

ال-activation map هي نتيجة تطبيق المرشحات على الصورة المدخلة و تسمى أيضاً feature map تسمح بفهم الميزات (feature) التي تم اكتشافها.

$$\text{Activation map} = \text{Input} * \text{Filter}$$

$$= \sum_{y=0}^{\text{columns}} \left( \sum_{x=0}^{\text{rows}} \right) \text{Input} (x - p, y - q) \text{Filter} (x, y)$$

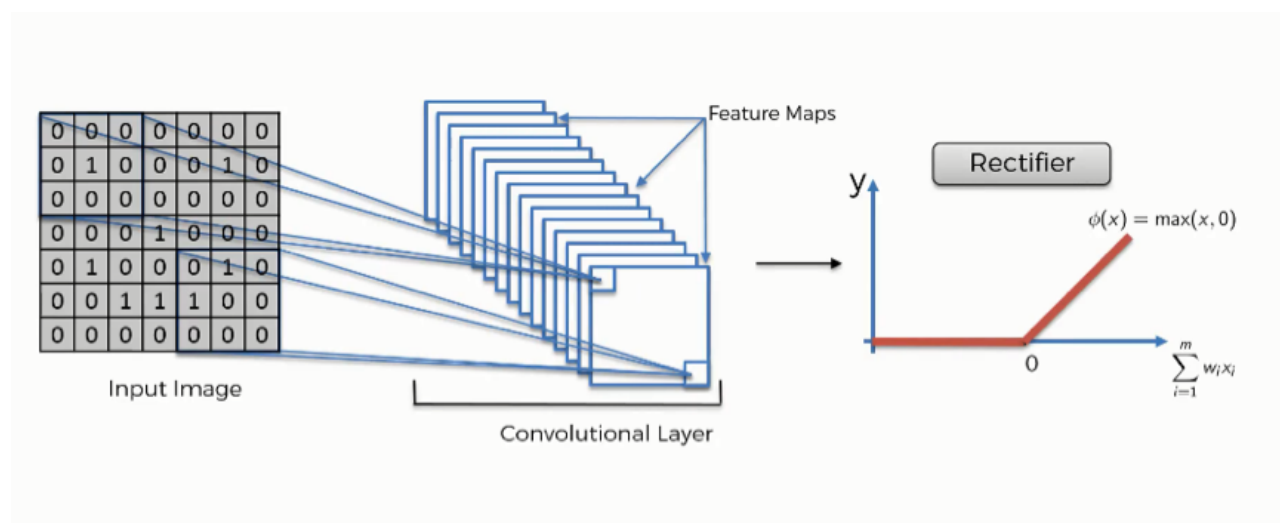
*Activation map algorithm*

- **Demo:**

- <https://setosa.io/ev/image-kernels/>

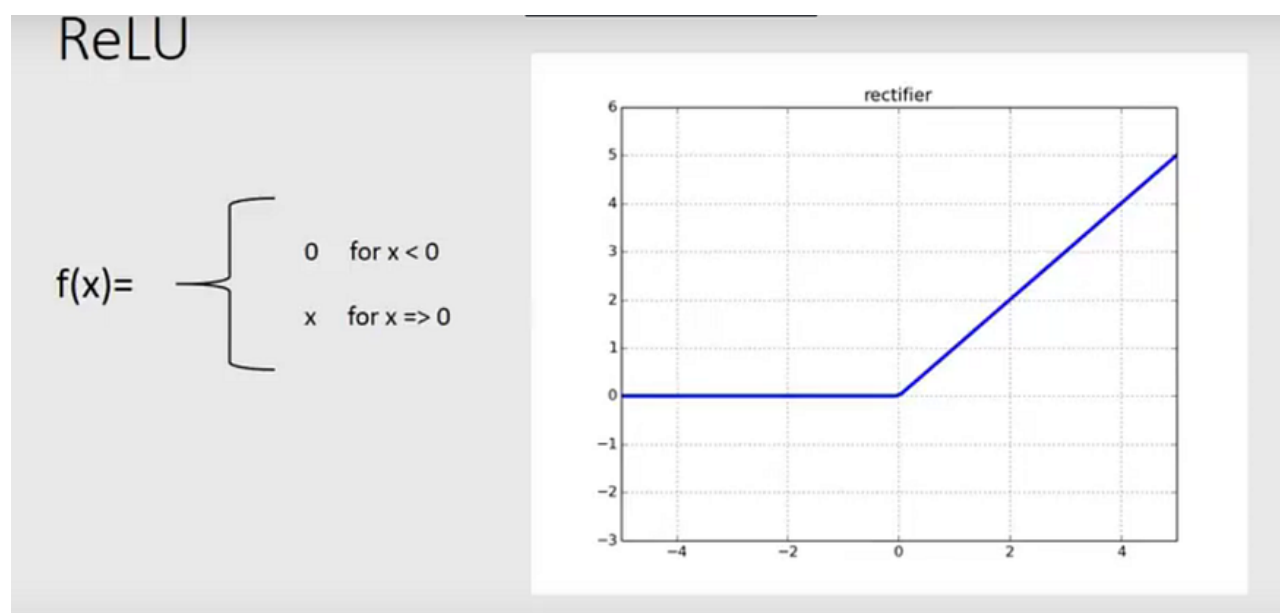
**طبقة (ReLU Layer):**

ReLU الوحدة الخطية المصححة (rectified linear unit). بمجرد استخراج الـ activation maps تكون الخطوة التالية هي نقلها إلى طبقة ReLU , أي أنها وظيفة التنشيط (activation function).

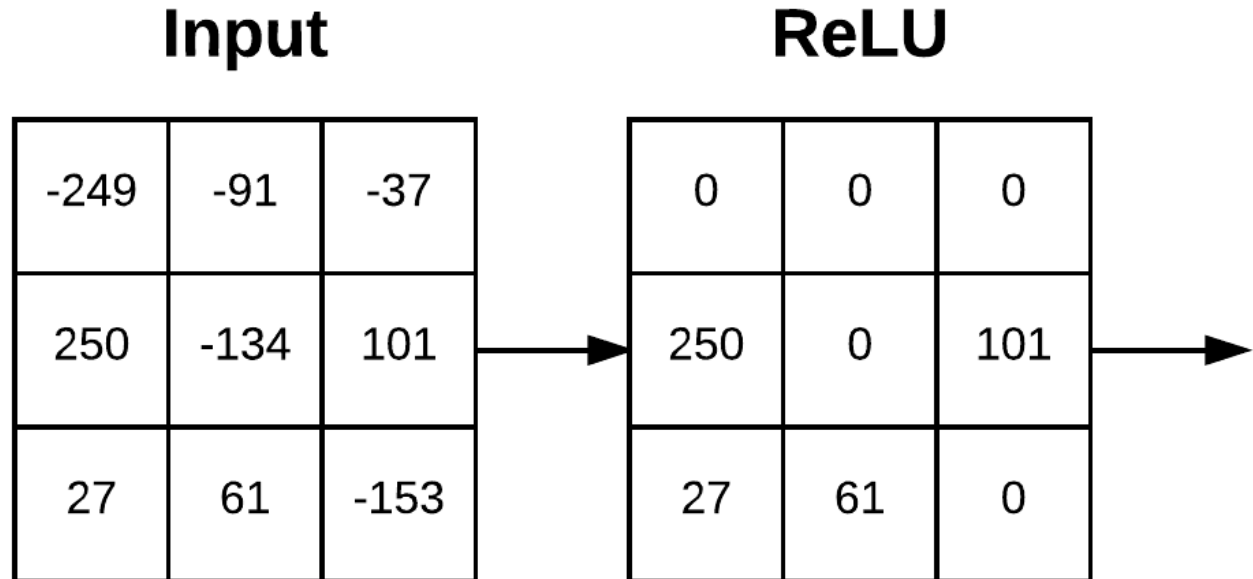


وهي وظيفة (function) تسمح بتعلم العلاقات المعقدة في البيانات, تساعد ReLU الشبكة على تعلم وظائف اتخاذ القرار الأصعب وتقليل overfitting, وهي عملية حسابية بسيطة :

$$y = \max(x, 0)$$



هي مرتبطة بـconvolution function فإذا كان ناتجها 0 أو قيمة سالبة فهذا يعني أن الميزة (feature) غير موجودة فـ نشير إليها بـ0, إما غير ذلك فـا نحتفظ بالقيمة نفسها.

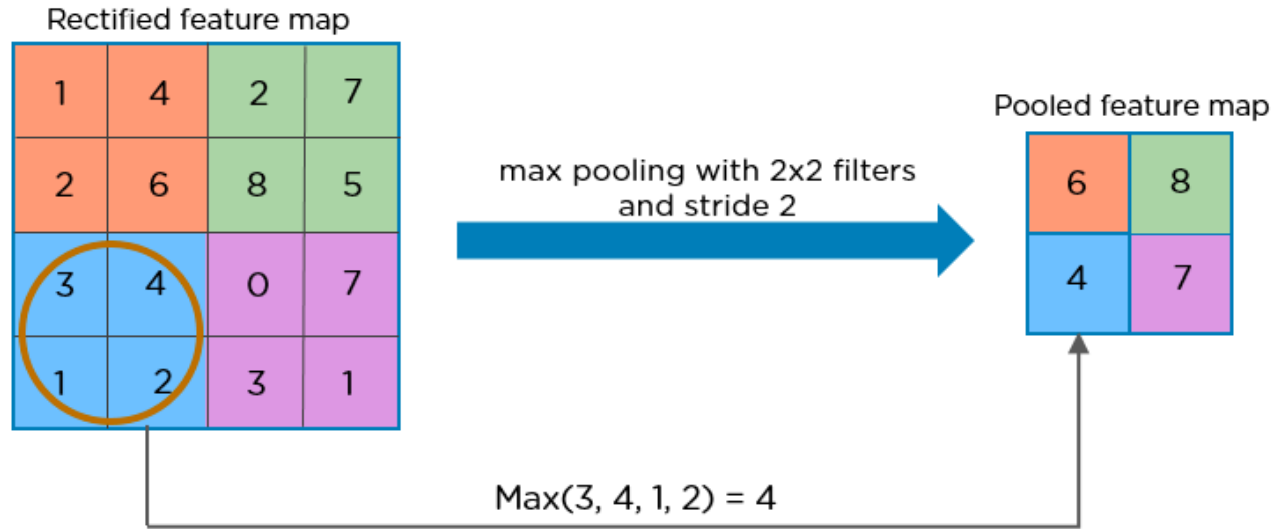


إذا الوظيفة خطية للقيم الأكبر من الصفر ، مما يعني أن لديها الكثير من الخصائص المرغوبة لوظيفة التنشيط الخطي عند تدريب شبكة عصبية.

### طبقة التجميع (Pooling Layer):

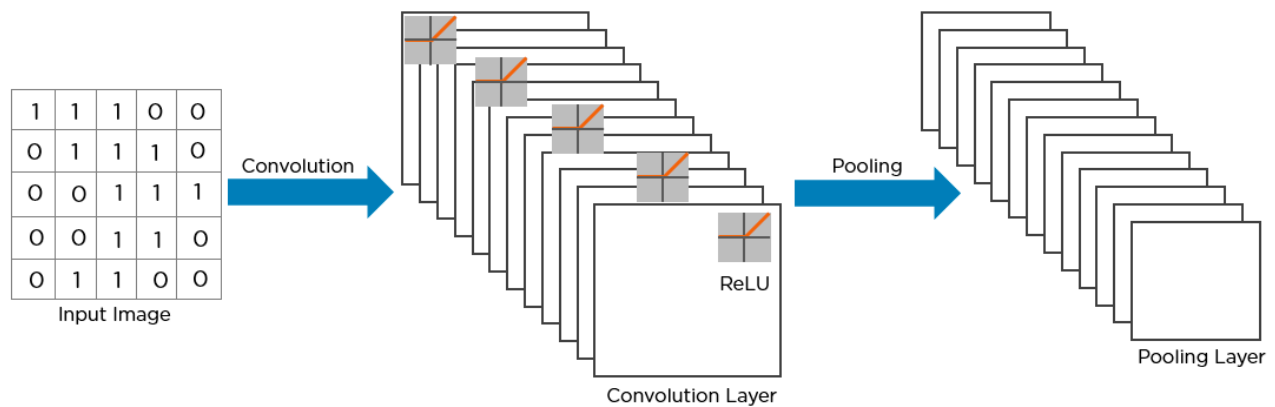
تسمى أيضاً طبقة العينات السفلية (down-sampling layer), تستخدم لتقليل أبعاد activation maps. في طبقة التجميع ، يطبق المرشح فوق بيانات الإدخال ثم بعدها ويطبق عملية التجميع (max, min, avg) لإنشاء pooled feature map, وعادة ما تكون القيمة القصوى (max) أو متوسط القيمة لجميع القيم (avg).





تستخدم طبقة التجميع مرشحات مختلفة لتحديد أجزاء مختلفة من الصورة, أي بمعنى اننا نأخذ في الواقع القيم التي تدل على وجود ميزة (feature) في هذا القسم من الصورة, بهذه الطريقة نتخلص من المعلومات غير المرغوب فيها. فيما يتعلق بوجود ميزة في جزء معين من الصورة ونأخذ في الاعتبار فقط ما هو مطلوب معرفته.

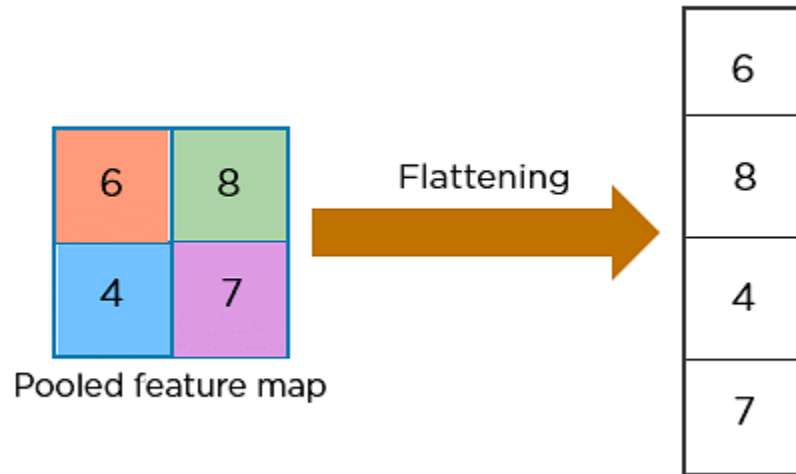
نرا كيف تبدو بنية convolution neural network حتى الآن:



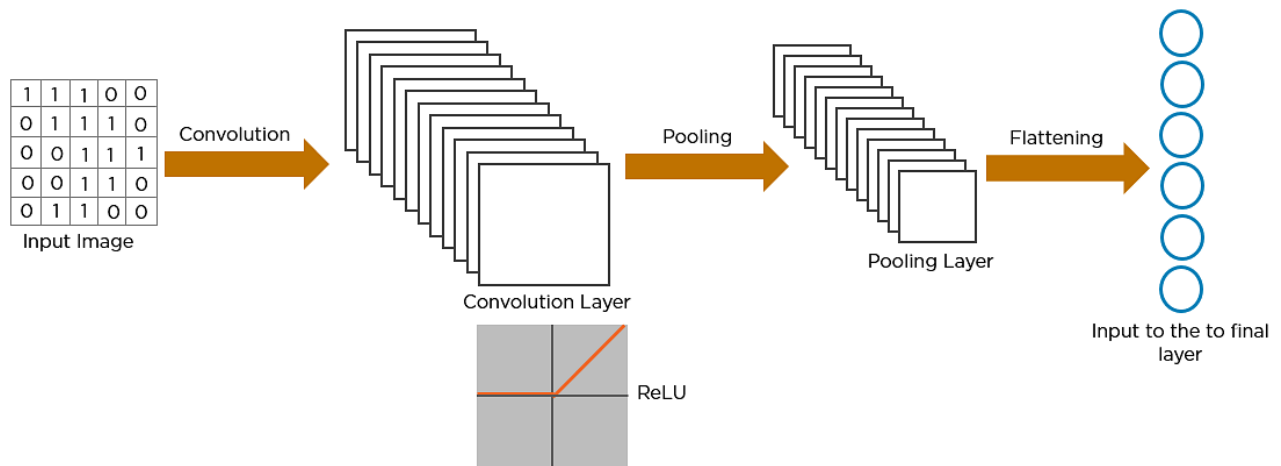
<https://youtu.be/hGOTo5ujH2w>

## التسطيح Flattening:

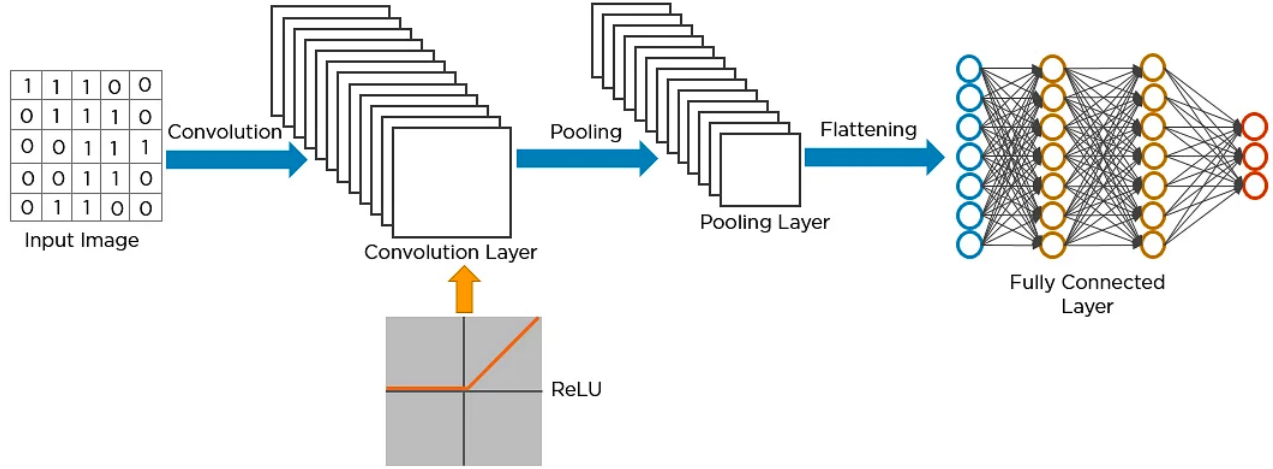
يتم استخدامها لتحويل كل مصفوفة ثنائية الأبعاد الناتجة من pooled feature maps إلى متجه خطي.



ويتم تعيين المصفوفة المسطحة كمدخل للطبقة المتصلة fully connected layer لتصنيف الصورة .

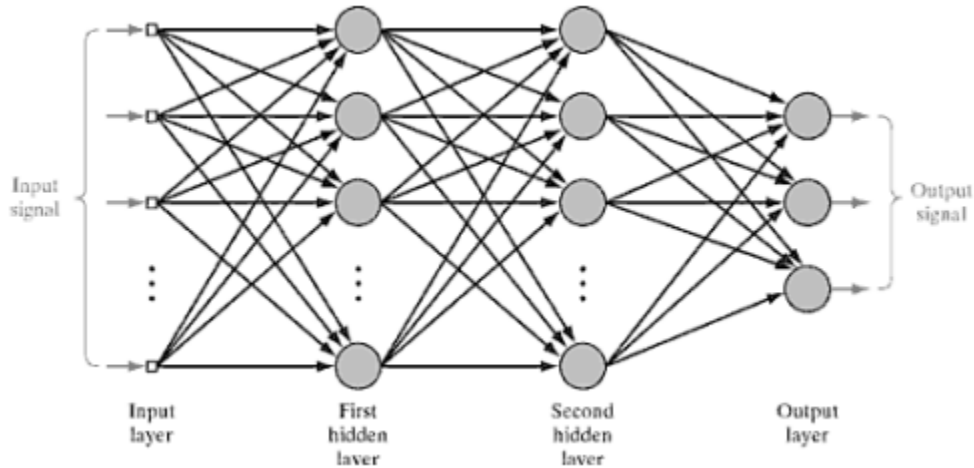


## الطبقة المتصلة بالكامل (Fully Connected):

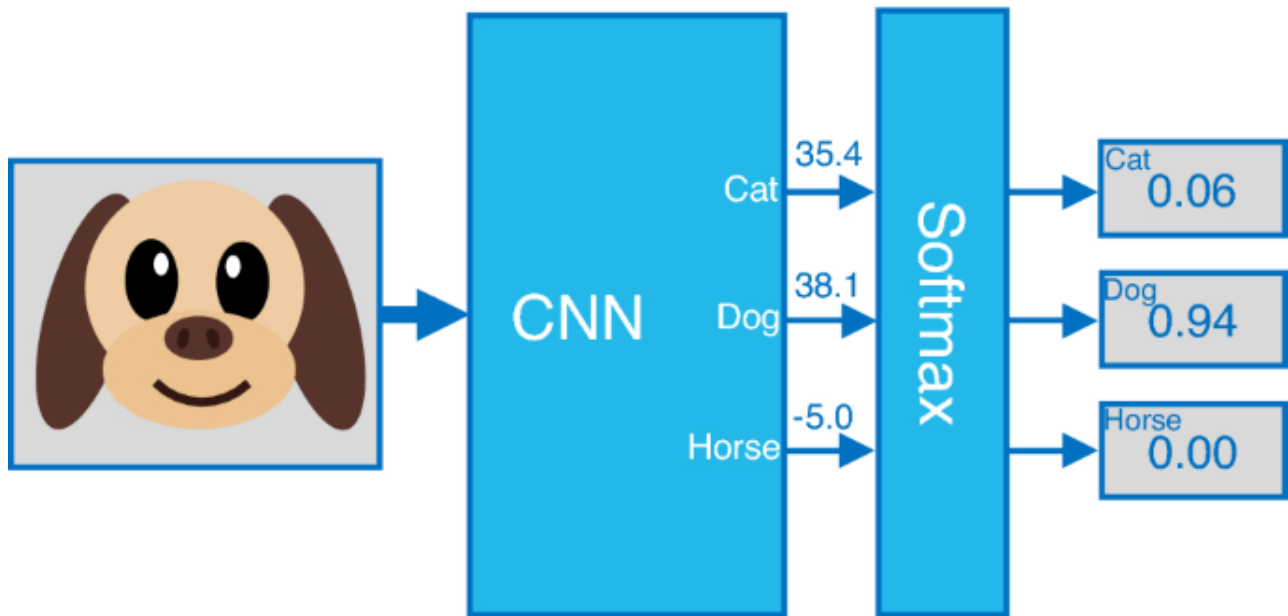


الطبقة المتصلة بالكامل هي ( MultiLayer Perceptron MLP ) ، يتكون من ثلاثة أنواع من الطبقات:

- الإدخال (input layer):
  - تتلقى طبقة الإدخال الميزات (features) التي تم إنشاؤها بواسطة CNN.
- المخفية (hidden layer):
  - الطبقة المخفية عبارة عن سلسلة من الخلايا العصبية ذات الأوزان التي سوف تتعلمها في خطوة التدريب.
  - ويمكن أن يتكون MLP من طبقة مخفية واحدة أو أكثر.
- المخرجة (output layer):
  - أيضاً طبقة الإخراج هي عبارة عن سلسلة من الخلايا العصبية , لكن تحتوي على وظيفة تنشيطية (activation function) مختلفة.



o عادة ما تستخدم وظيفة softmax لتوليد احتمالات كل فئة.



<https://youtu.be/Aop4rGjMskI>

<https://youtu.be/JboZfxUjLSk>

<https://youtu.be/JboZfxUjLSk>

<https://youtu.be/JboZfxUjLSk>

<https://youtu.be/JboZfxUjLSk>

---

---

## : Different types of CNN Architectures - CNN البنى الشائعة في

تنقسم إلى نوعين

● الشبكات الكلاسيكية (Classic network):

- LeNet-5
- AlexNet
- VGG 16

● الشبكات الحديثة (Modern network):

- Inception
- ResNet
- ResNeXt
- DenseNet

---

---

## Resources

- [https://www.researchgate.net/publication/333242381\\_Convolutional\\_Neural\\_Network\\_Layers\\_and\\_Architectures](https://www.researchgate.net/publication/333242381_Convolutional_Neural_Network_Layers_and_Architectures)
- <https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939>
- <https://www.sciencedirect.com/topics/engineering/convolutional-layer>
- <https://medium.com/skylar-salernos-tech-blog/image-convolution-filters-explained-c878f1056e78>
- <https://www.kaggle.com/discussions/general/171197>
- <https://youtu.be/yb2tPt0QVPY?feature=shared>
-