

A decorative graphic on the left side of the slide consisting of two overlapping parallelograms. The front one is blue and the back one is a light green color. They are positioned diagonally, with the blue one partially covering the green one.

Convolutional Neural Networks

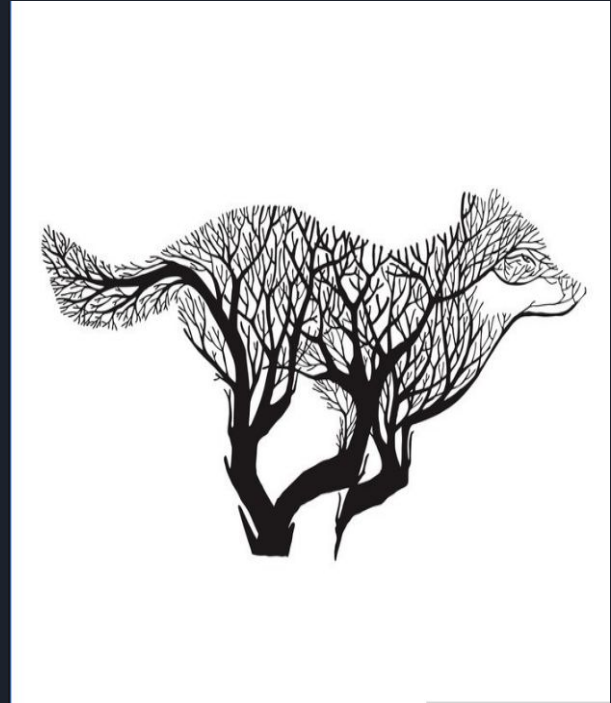
Καλτσόπουλος Ερρίκος,
Θεοφανίδης Ηρακλής



Τι είναι το CNN;

Το CNN είναι τεχνητό νευρωνικό δίκτυο το οποίο χρησιμοποιείται κυρίως για την ταξινόμηση (classification) εικόνων ,ομαδοποίηση(clustering), και επίσης μπορεί να πραγματοποιείται αναγνώριση αντικειμένων. Είναι αλγόριθμος που μπορεί να εντοπίσει πρόσωπα, άτομα, ταμπελες, όγκους, και πολλές άλλες πτυχές οπτικών δεδομένων.

Πώς ο εγκέφαλος μας αντιλαμβάνεται τις εικόνες





Αναγνώριση εικόνων από τον υπολογιστή

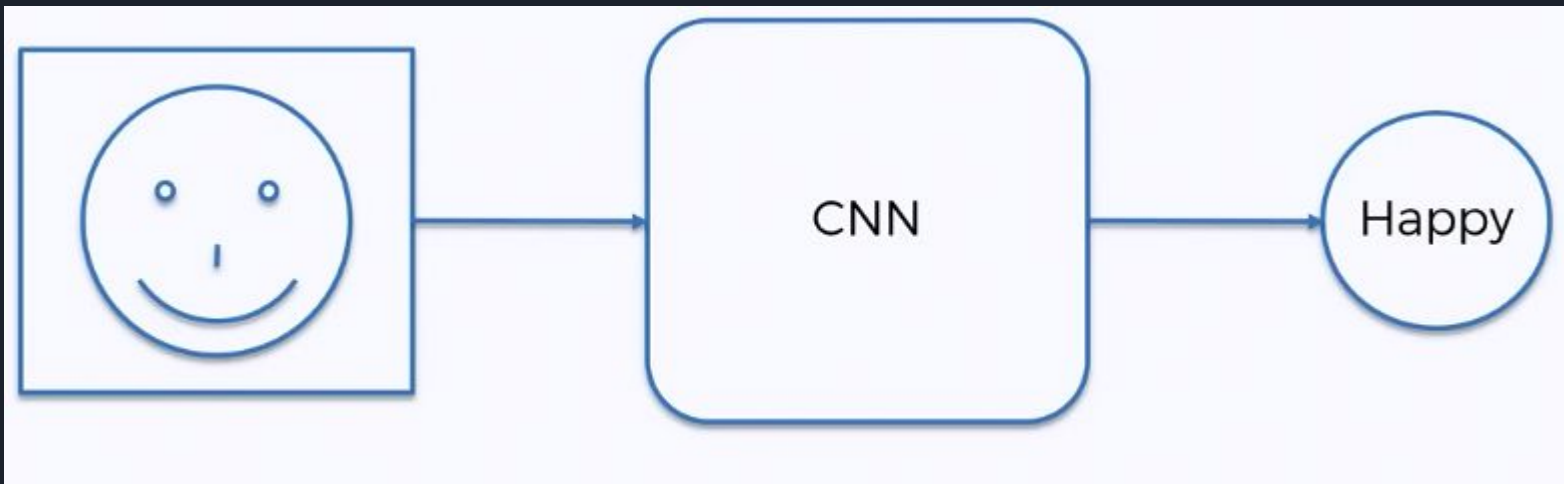
Ασπρόμαυρη εικόνα

- $2 \times 2 = 4 \text{ Pixel}$
- Το κάθε pixel μπορεί να πάρει τιμή από 0 έως και 255
- Αν είναι 0 είναι τελείως μαύρο, ενώ αν είναι 255 είναι τελείως άσπρο

Έγχρωμη εικόνα

- $2 \times 2 = 4 \text{ Pixel}$
- Έχει τρία στρώματα το Red, Green, Blue, το λεγόμενο μοντέλο RGB, οπότε το κάθε Pixel μπορεί να πάρει 3 τιμές από 0 έως 255 .

Απλό παράδειγμα νευρωνικού





Το CNN Αποτελείται από 4 βήματα

- Convolution
- Max Pooling
- Flattering
- Full Connection



Convolution

Το βήμα αυτό αποτελείται απο το

- Input Imag
 - εισαγόμενη εικόνα. ε:εισαγόμενη εικόνα
- Feature detector
 - βρίσκει τα κύρια χαρακτηριστικά μίας εικόνας.:βρίσκει τα κύρια χαρακτηρίστηκα μιάμιση εικόνας
- Feature Map
 - ουσιαστικά μας δείχνει που ακριβώς βρίσκονται αυτά τα κύρια χαρακτηριστικά.

:ουσιαστικά μας δείχνει που ακριβώς βρίσκονται αυτά τα κύρια χαρακτηριστικα. Αν θέλετε να δείτε τι κρύβεται πίσω από τα μαθηματικά στην μέθοδο convolution μπείτε στο Link :

<http://cs.niu.edu.cn/wuix/paper/CNN.pdf>

$$(f * g)(t) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(\tau) g(t - \tau) d\tau$$

Convolution

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

Input Image



0	0	1
1	0	0
0	1	1

Feature
Detector



0	1	0	0	0
0	1	1	1	0
1	0	1	2	1
1	4	2	1	0
0	0	1	2	1

Feature Map



Max Pooling($\frac{1}{2}$)

Το max pooling το χρησιμοποιούμε όταν θέλουμε να αναγνωρίσουμε ένα χαρακτηριστικό σε ένα αντικείμενο.(π.χ. Ένα ζώο όπως γάτα).

Όταν ψάχνουμε σε πολλές εικόνες ένα κοινό χαρακτηριστικό για το αντικείμενο στο οποίο μας ενδιαφέρει να αναγνωρίσουμε , το αντικείμενο μπορεί να βρίσκεται σε διαφορετική θέση,με διαφορετική φωτεινότητα , ή το αντικείμενο να μην είναι ακριβώς το ίδιο(π.χ. Πολλές γάτες με διαφορετικά χρώματα).

Επίσης με το pooling:

1. Μειώνεται το μέγεθος της εικόνας
2. Βοηθάει στην αποφυγή του Overfeeding

Max Pooling(2/2)

0	1	0	0	0
0	1	1	1	0
1	0	1	2	1
1	4	2	1	0
0	0	1	2	1

Feature Map

Max Pooling

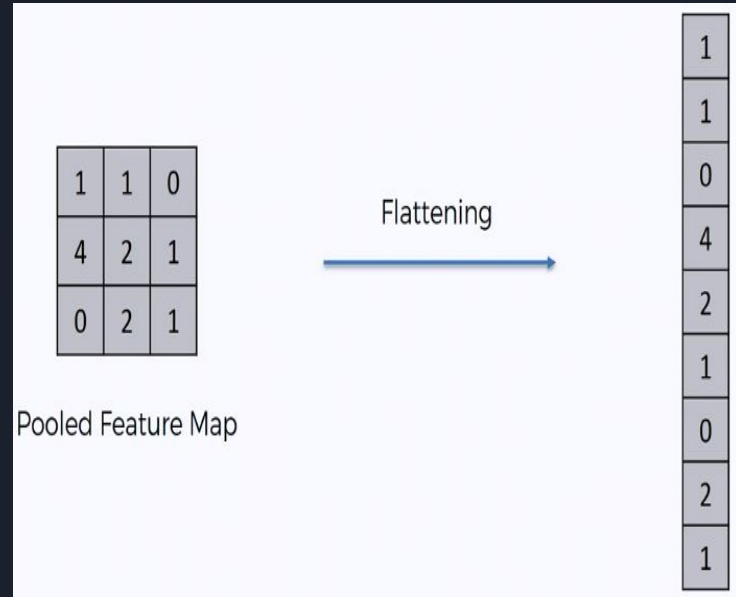


1	1	0
4	2	1
0	2	1

Pooled Feature Map

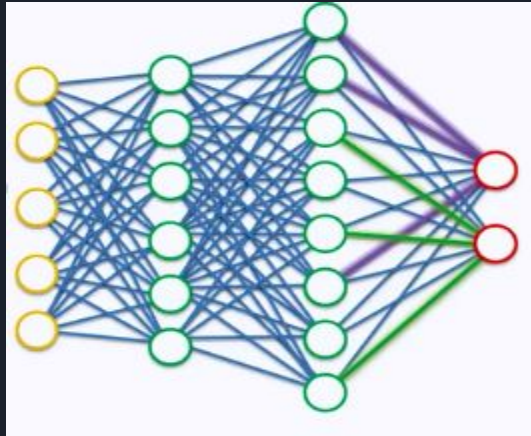
Flattening

Ουσιαστικά παίρνουμε τα Pooled feature maps από το pooling layer και κάνουμε το κάθε ένα από μία στήλη ώστε αυτές η στήλες να είναι η είσοδοι στο νευρωνικό μας.



Full Connection

Μετά το flattening εισάγονται τα δεδομένα στο νευρωνικό δίκτυο το οποίο τα επεξεργάζεται. Στο final fully connected layer γίνεται η ψηφοφορία για την κλάση στην οποία ανήκει. Βεβαίως προηγείται το training όπου γίνονται πολλές εποχές, η διορθώσεις και βρίσκουμε ποιοί νευρώνες είναι σημαντικοί για κάθε κλάση. Επίσης δεν εκπαιδεύεται μόνο το νευρωνικό αλλά και οι feature detectors.



Draw your number here



Downsampled drawing:

First guess:

Second guess:

Layer visibility

Input layer

Convolution layer 1

Downsampling layer 1

Convolution layer 2

Downsampling layer 2

Fully-connected layer 1

Fully-connected layer 2

Output layer

0123456789

Output Layer

Fully Connected Layer 2

Fully Connected Layer 1

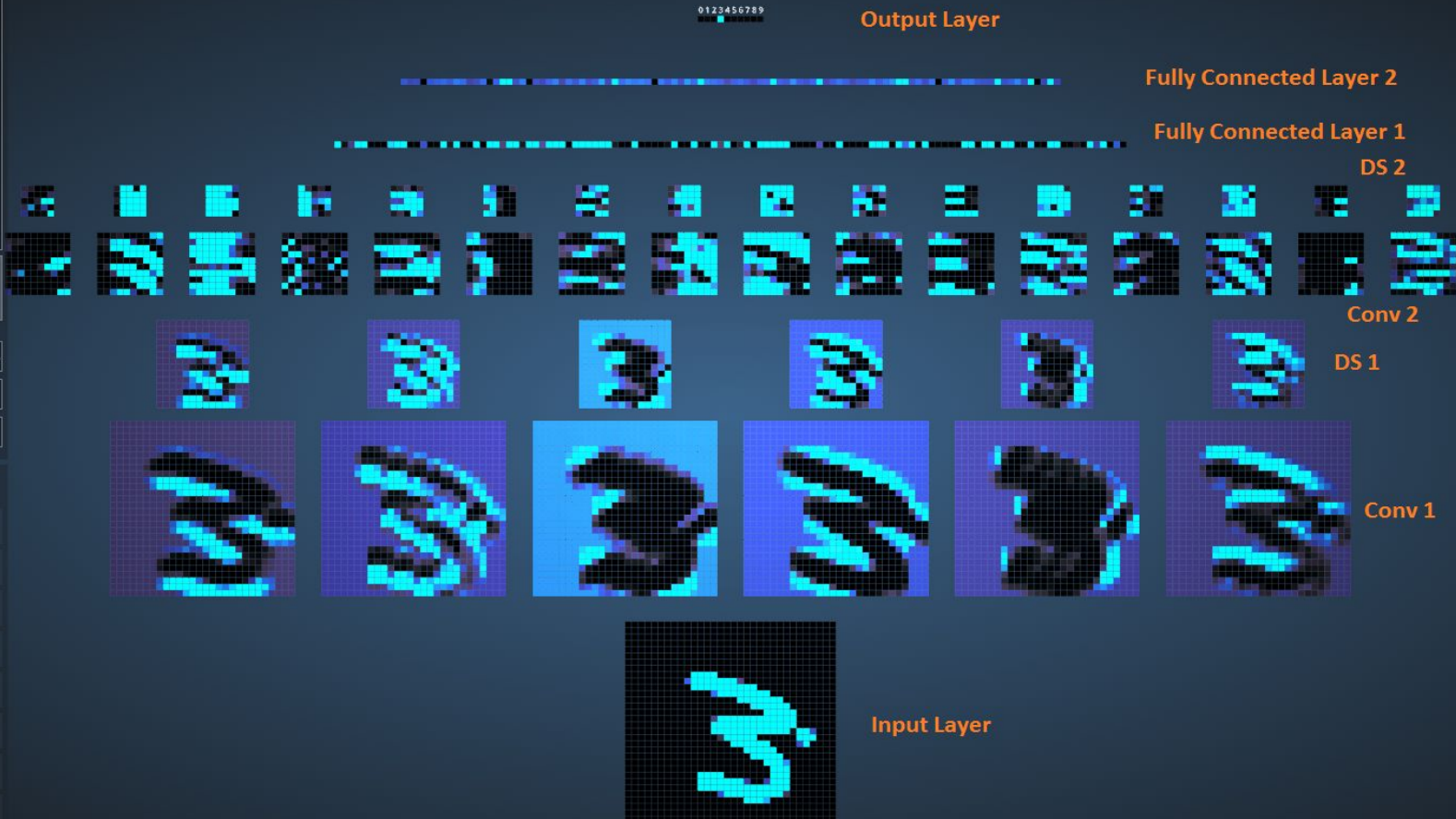
DS 2

Conv 2

DS 1

Conv 1

Input Layer





ΤΕΛΟΣ