



Τεχνητή Νοημοσύνη και Μηχανική Μάθηση

Κωνσταντίνος Καραμανής

The University of Texas at Austin & Archimedes/Athena RC

constantine@utexas.edu


<https://caramanis.github.io/>





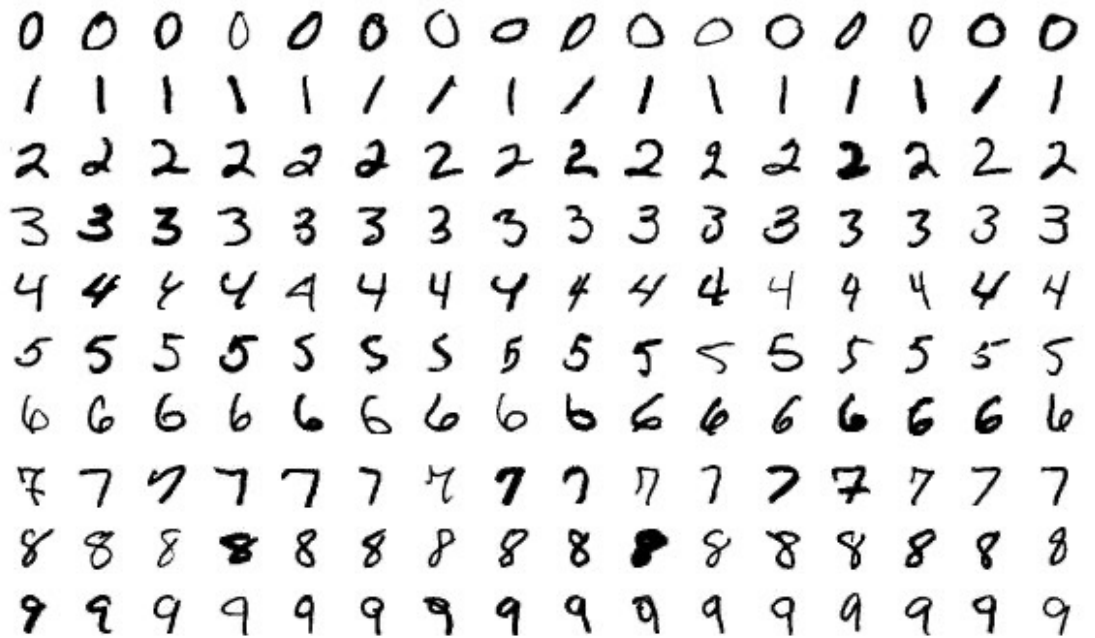
Μηχανική Όραση

Πως αποθηκεύεται μία
εικόνα στον υπολογιστή;



Μαυρόασπρες εικόνες

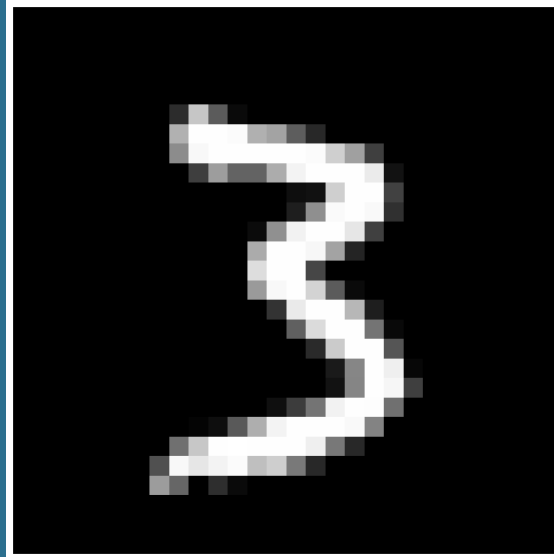
MNIST – σετ δεδομένων με
χειρόγραφα ψηφία 0-9



Μαύρα εικόνες

28 x 28 αριθμοί από
0 έως 255
κωδικοποιούν που
είναι τα μαύρα πίξελ
και που τα άσπρα

MNIST – σετ δεδομένων με
χειρόγραφα ψηφία 0-9

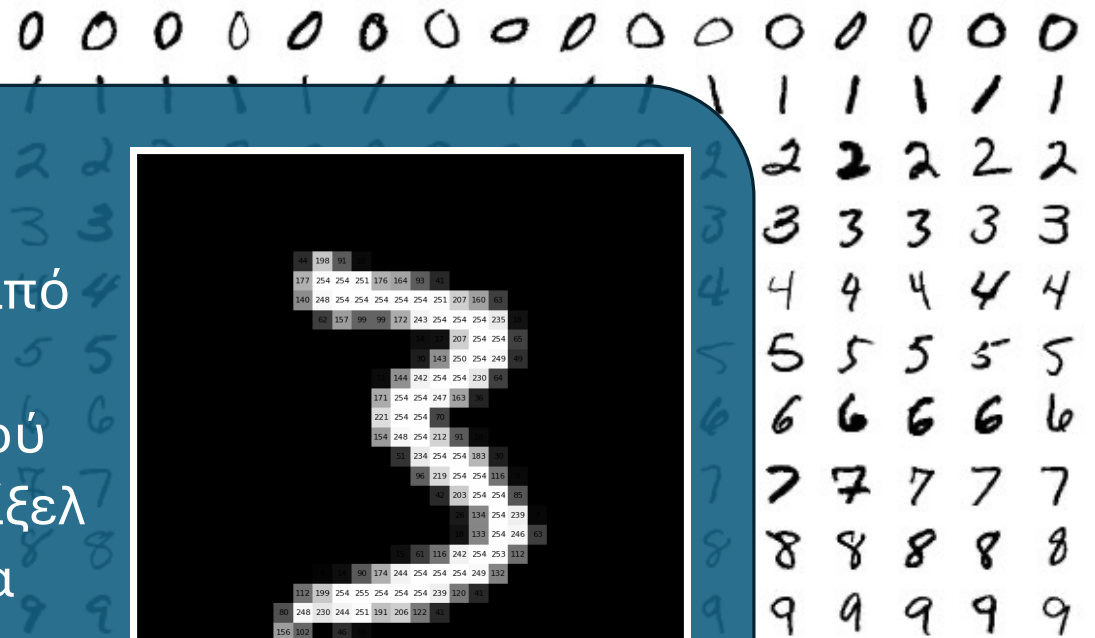


0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1 1 1 1 1 1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2 2 2 2 2 2
3 3 3 3 3 3 3 3 3 3 3 3 3 3
4 4 4 4 4 4 4 4 4 4 4 4 4 4
5 5 5 5 5 5 5 5 5 5 5 5 5 5
6 6 6 6 6 6 6 6 6 6 6 6 6 6
7 7 7 7 7 7 7 7 7 7 7 7 7 7
8 8 8 8 8 8 8 8 8 8 8 8 8 8
9 9 9 9 9 9 9 9 9 9 9 9 9 9

Μαύρες εικόνες

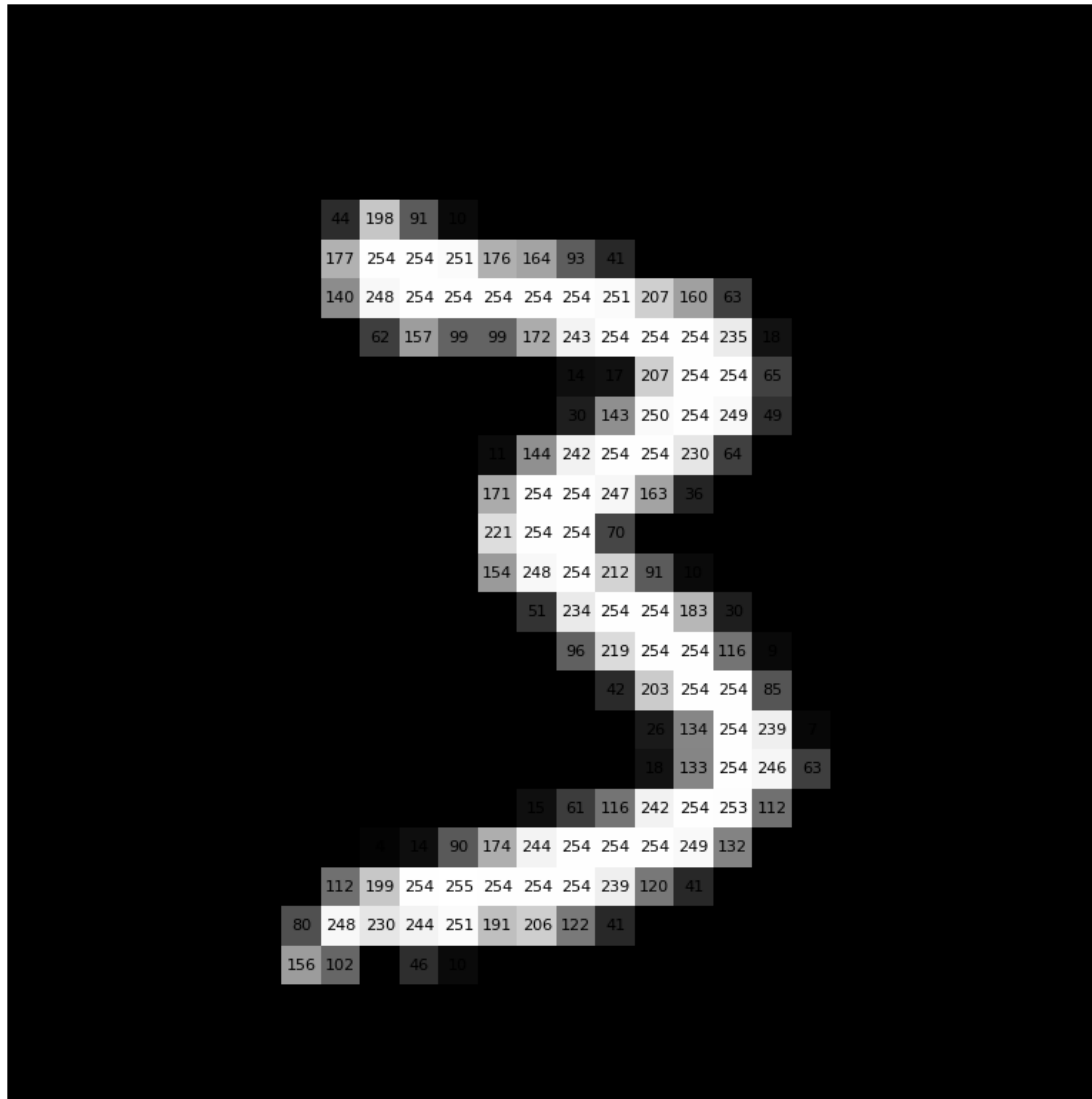
28 x 28 αριθμοί από
0 έως 255
κωδικοποιούν που
είναι τα μαύρα πίξελ
και που τα άσπρα

MNIST – σετ δεδομένων με
χειρόγραφα ψηφία 0-9



Μαυροί εικόνες

MNIST – σετ δεδομένων
χειρόγραφα ψηφία

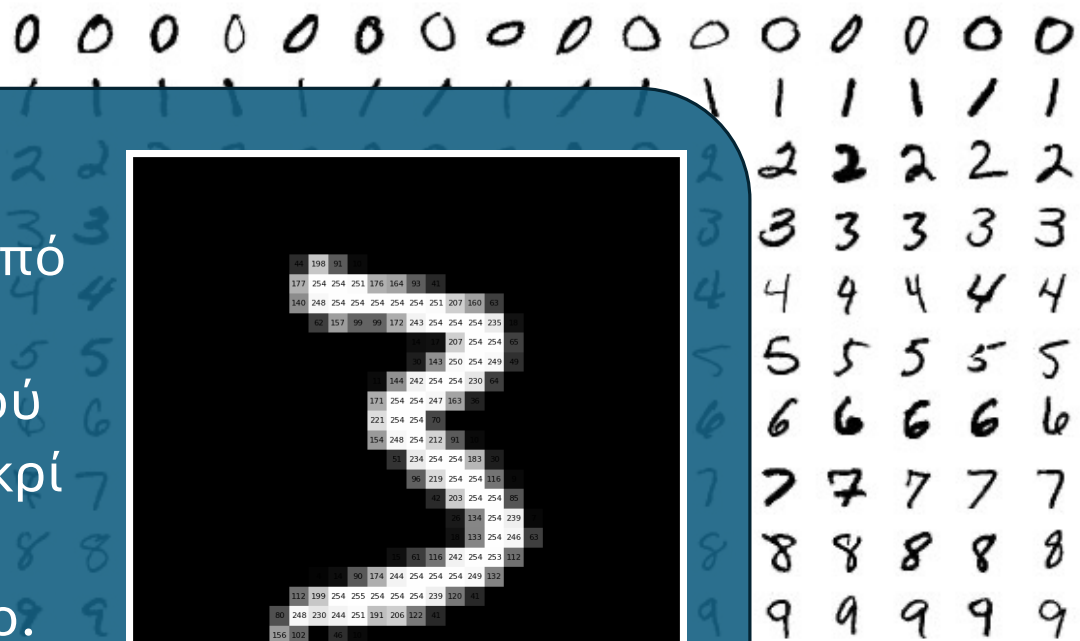
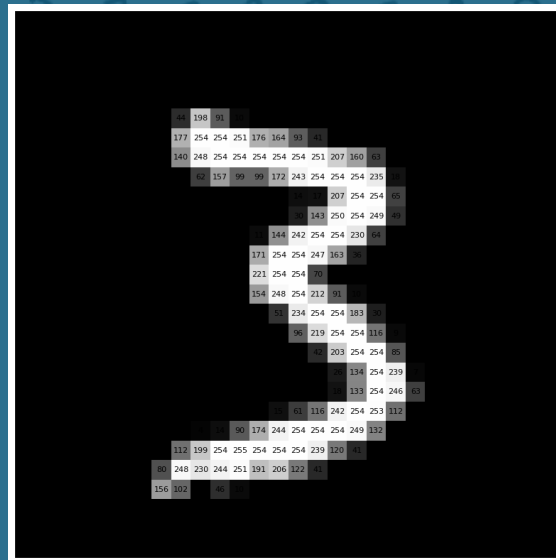


0 0 0 0 0 0
1 1 1 1 1 1
2 2 2 2 2 2
3 3 3 3 3 3
4 4 4 4 4 4
5 5 5 5 5 5
6 6 6 6 6 6
7 7 7 7 7 7
8 8 8 8 8 8
9 9 9 9 9 9

Μαυρά εικόνες

MNIST – σετ δεδομένων με
χειρόγραφα ψηφία 0-9

28 x 28 αριθμοί από
0 έως 255
κωδικοποιούν πού
είναι τα μαύρα/γκρί
ή άσπρα. 255 =
άσπρο, 0 = μαύρο.



Μαυρά εικόνες

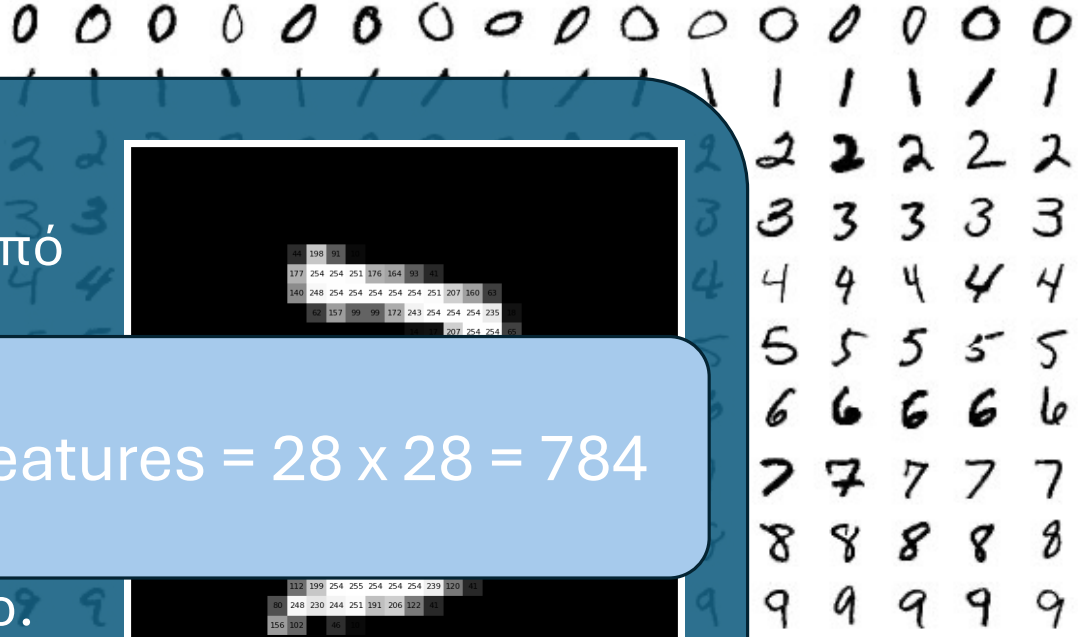
28 x 28 αριθμοί από
0 έως 255

κω
είν
ή α

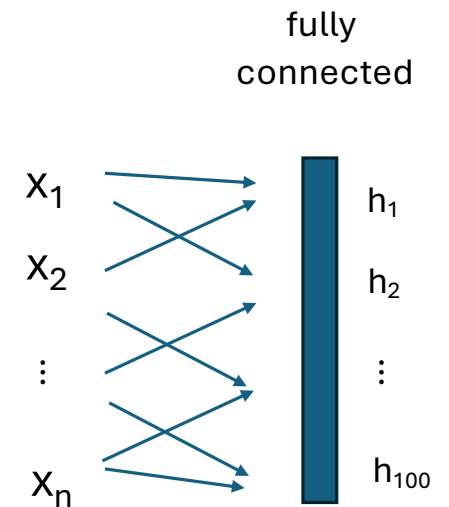
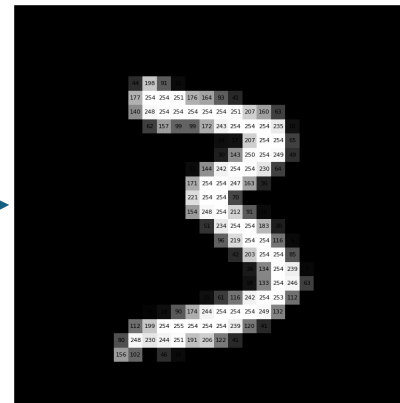
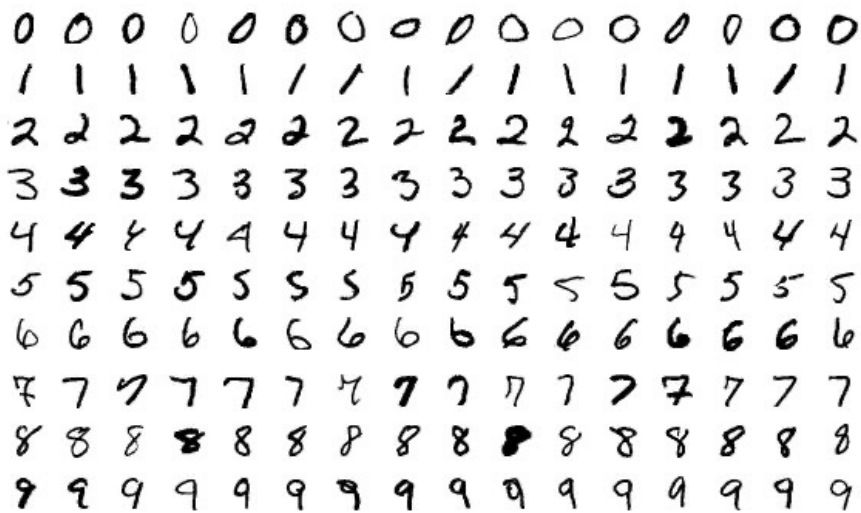
άσπρο, 0 = μαύρο.

MNIST – σετ δεδομένων με
χειρόγραφα ψηφία 0-9

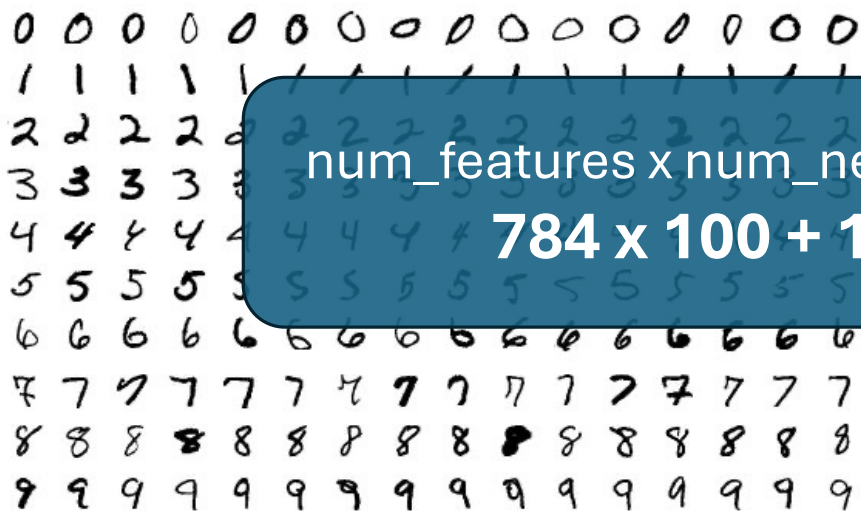
Οπότε: # features = 28 x 28 = 784



Πόσες παραμέτρους έχει ένα γραμμικό (fully connected) επίπεδο με 100 νευρώνες;

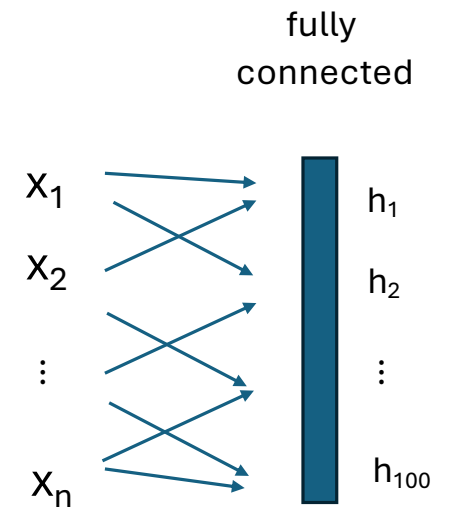
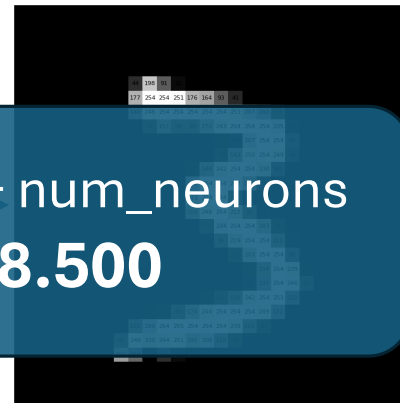


Πόσες παραμέτρους έχει ένα γραμμικό (fully connected) επίπεδο με 100 νευρώνες;



num_features x num_neurons + num_neurons

$$784 \times 100 + 100 = 78.500$$



Μηχανική Όραση

Πως αποθηκεύεται μία **έγχρωμη** εικόνα στον υπολογιστή;

Κάθε εικόνα 500 x 500, αντιπροσωπεύεται ως 500 επί 500 πίξελ, με κάθε πίξελ να αντιστοιχεί σε **τρεις** αριθμούς. Αυτοί οι τρεις αριθμοί μεταξύ 0 και 255, δηλώνουν την ένταση τριών χρωμάτων, R για Κόκκινο (Red), G για Πράσινο (Green), και B για Μπλε (Blue), εξού και το RGB. Συνδυάζοντας αυτά τα τρία χρώματα μπορούμε να πάρουμε τα περισσότερα χρώματα του φάσματος.

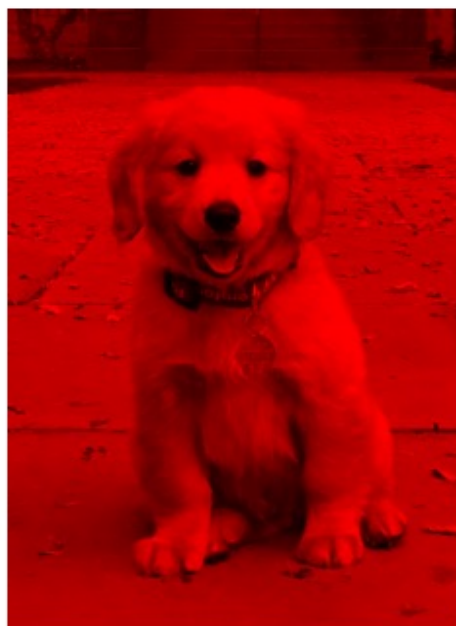


Μηχανική Όραση

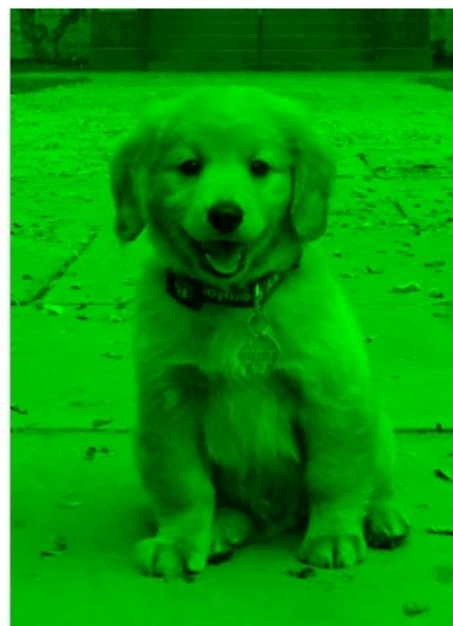
Original



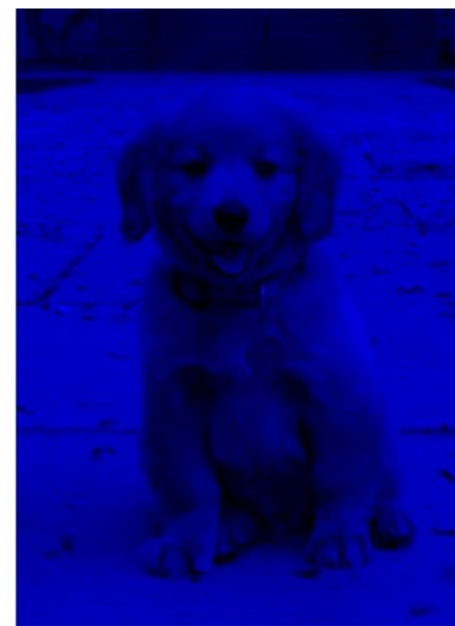
Red Channel



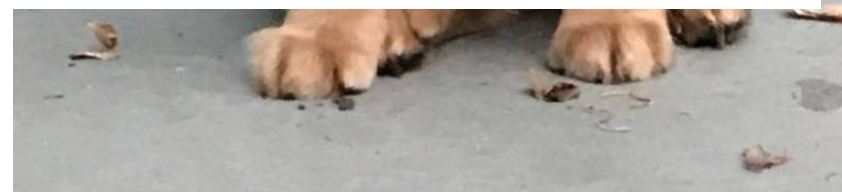
Green Channel



Blue Channel



τα τρία χρώματα μπορούμε να παρουμε τα περισσότερα χρώματα του φάσματος.



Μηχανική Όραση

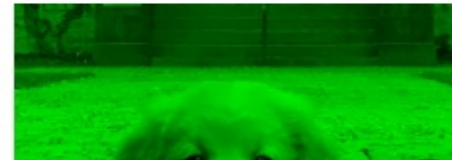


Original

Red Channel

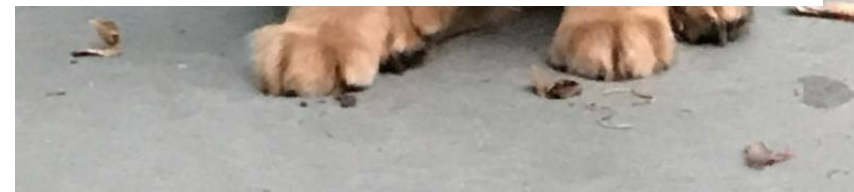
Green Channel

Blue Channel



Οπότε: # features = $500 \times 500 \times 3 = 750.000$

τα τρία χρώματα μπορούμε να παρουμε τα περισσότερα χρώματα του φάσματος.



CIFAR-10

50.000 έγχρωμες 32 x 32 εικόνες.

Οπότε πιο είναι το μέγεθος του πίνακα X που αποθηκεύει τα δεδομένα;

airplane



automobile



bird



cat



deer



dog



frog



horse



ship



truck

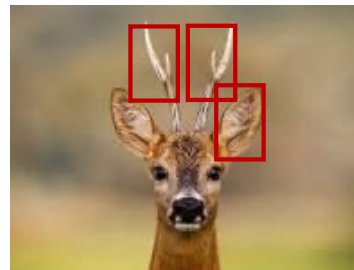
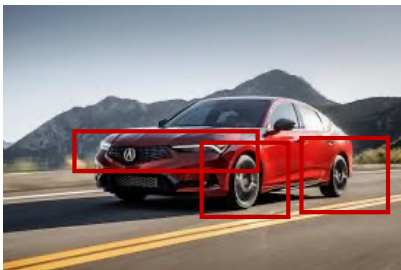
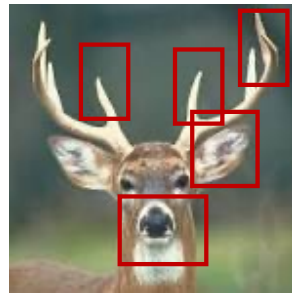
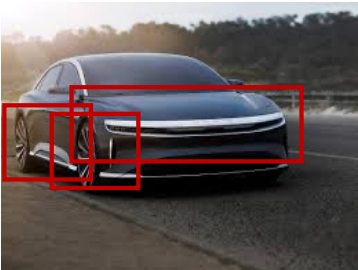
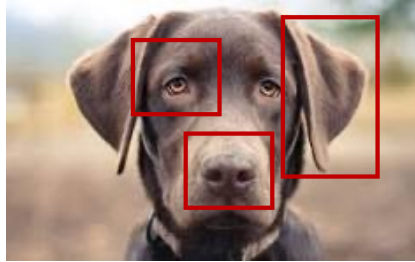
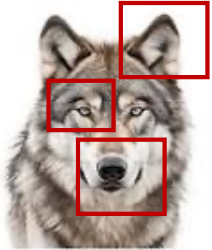


Νευρωνικά Δίκτυα

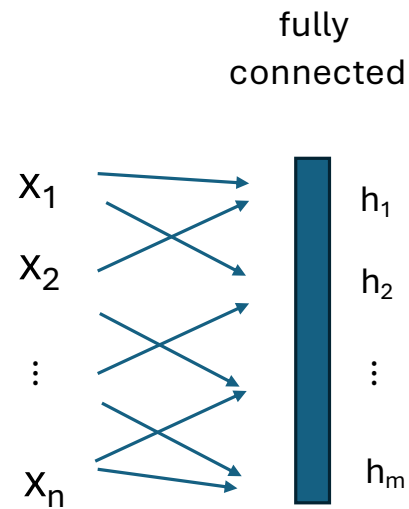
Μία απείρως πλούσια
οικογένεια
αλγορίθμων που εάν
σχεδιαστούν σωστά,
ταιριάζουν με πολλούς
τομείς



Πως
διαφέρουν
αυτές οι
εικόνες;

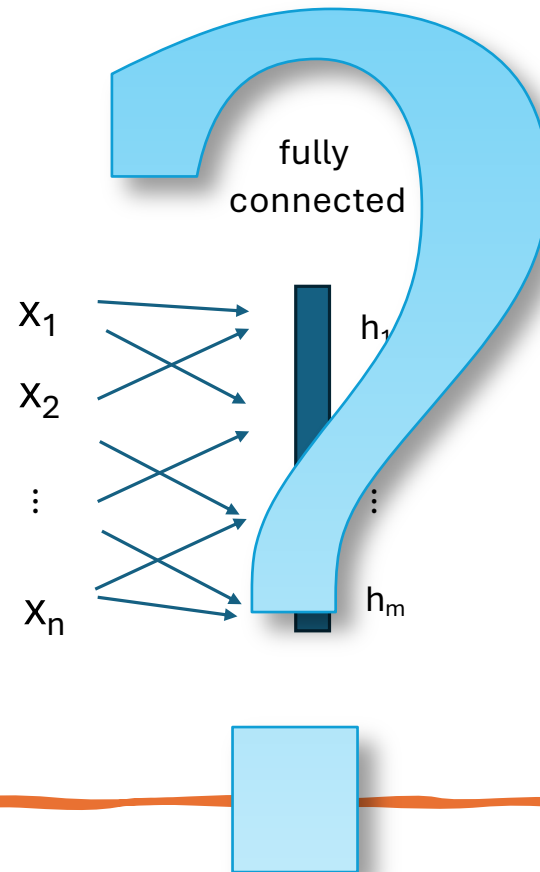


Πως
διαφέρουν
αυτές οι
εικόνες;



Νευρωνικά Δίκτυα

Μία απείρως πλούσια
οικογένεια
αλγορίθμων που εάν
σχεδιαστούν σωστά,
ταιριάζουν με πολλούς
τομείς



Νευρωνικά Δίκτυα

Μία απείρως πλούσια οικογένεια αλγορίθμων που εάν σχεδιαστούν σωστά, ταιριάζουν με πολλούς τομείς