Ταξινόμηση σκίτσων με τη χρήση Βαθιών Νευρωνικών Δικτύων και της TensorFlow

Ευθυμία Καφαλή

Περιεχόμενα

- Εργαλεία
- Δεδομένα
- Μοντέλο
- Αποτελέσματα

1 Εργαλεία

Εργαλεία

python™

TensorFlow

Keras

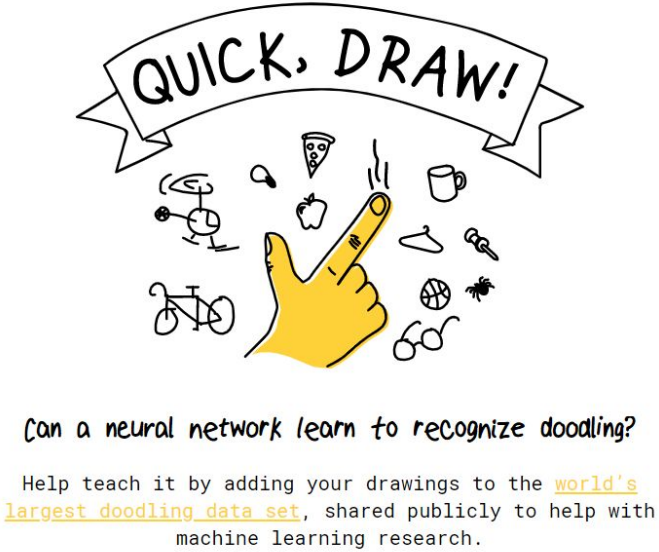
NumPy

scikit learn

matplotlib

2 Δεδομένα

Δεδομένα (1/4)



"Quick, Draw!" [Online]. Available: https://quickdraw.withgoogle.com/

Δεδομένα *(2/4)*

Σύνολο δεδομένων "Quick, Draw!" της Google

- Ασπρόμαυρα σκίτσα διαστάσεων 28 x 28
- Ζωγραφισμένα από χρήστες του online παιχνιδιού
- 345 κλάσεις
- ~ 2.5 από τα 15 εκατομμύρια διαθέσιμα σκίτσα

D. Ha and D. Eck, "A Neural Representation of Sketch Drawings," Apr. 2017

Δεδομένα (3/4)

Παραδείγματα σκίτσων

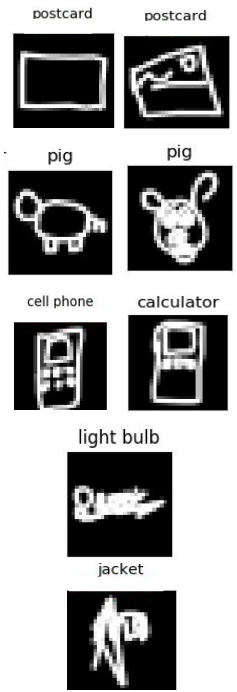
- Ζώα
- Μουσικά όργανα
- Εργαλεία
- Καθημερινά αντικείμενα
- Είδη ρουχισμού
- Οχήματα
- Απλά σχήματα
- Landmarks (Eiffel Tower, Great Wall of China) κ.ά.

Δεδομένα *(4/4)*

Challenges:

Τα δεδομένα δημιουργήθηκαν από την αλληλεπίδραση χρήστη - Η/Υ

- Μη ολοκληρωμένα σκίτσα
- Οπτική
- Ομοιότητα κλάσεων
- Κατανόηση των ζητούμενων αντικειμένων
- Εσκεμμένος αποπροσανατολισμός του ΝΔ



3 Μοντέλο

Μοντέλο *(1/3)*

Στόχος. Υλοποίηση ενός classifier για την ταξινόμηση των σκίτσων στις 345 κλάσεις

Μοντέλο. Convolutional Neural Network

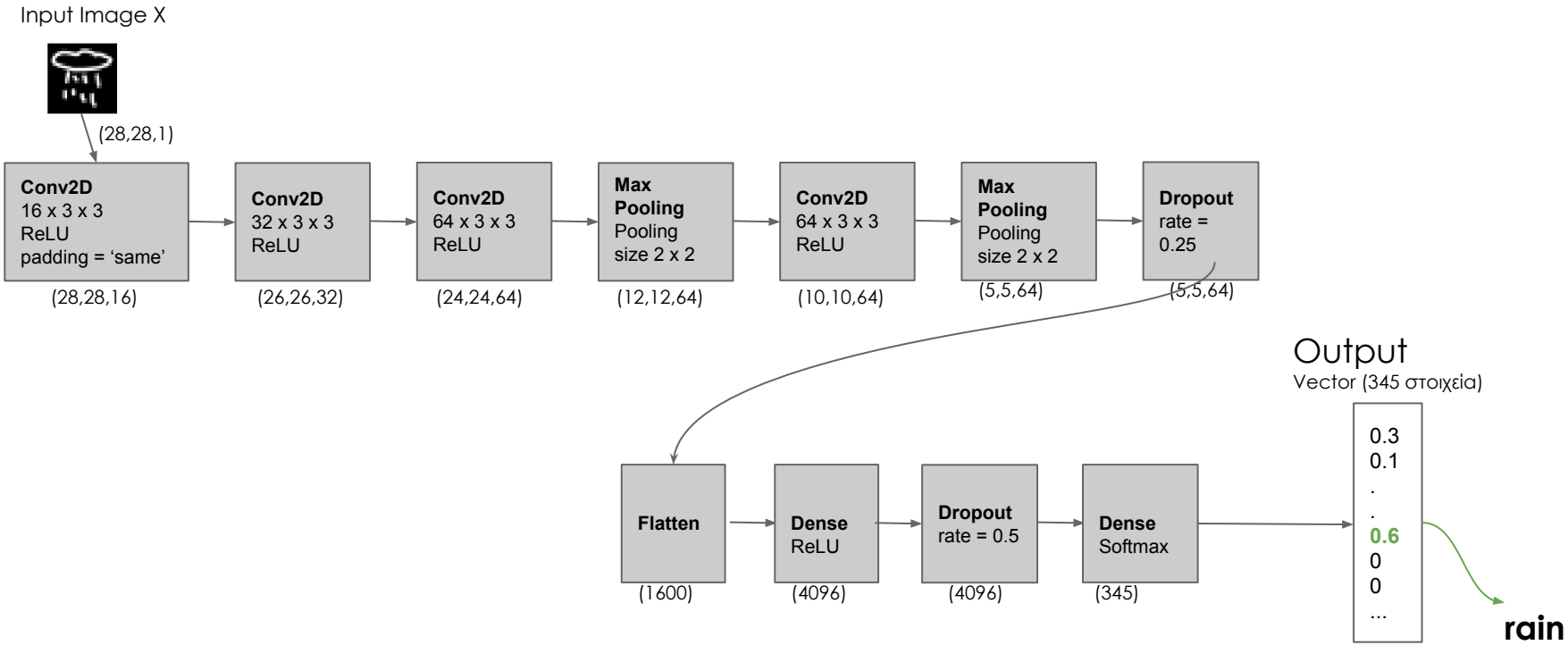
Loss Function. Categorical Cross Entropy

Optimizer. Adam

Training Epochs. 30

Batch size. 1024

Hardware: GPU NVIDIA Titan Xp



Συνολικός αριθμός παραμέτρων: 8,031,385

Μοντέλο *(3/3)*

Έλεγχος Overfitting

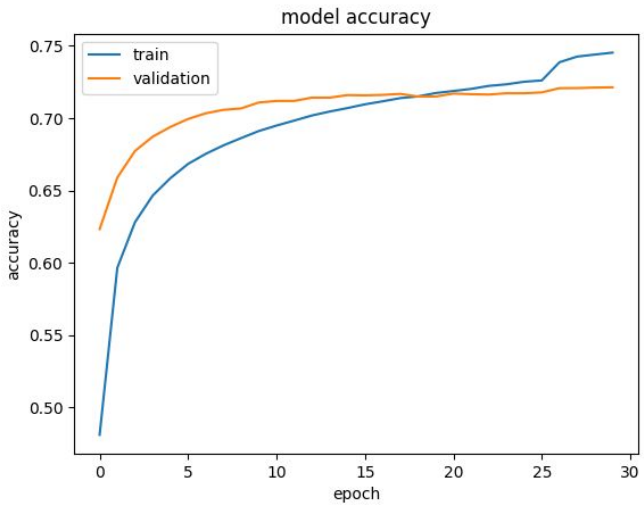
- Keras Callbacks
  - Model Checkpoint
  - Reduce learning rate on plateau (patience = 5)
  - Early Stopping (patience = 6)
- Data shuffling

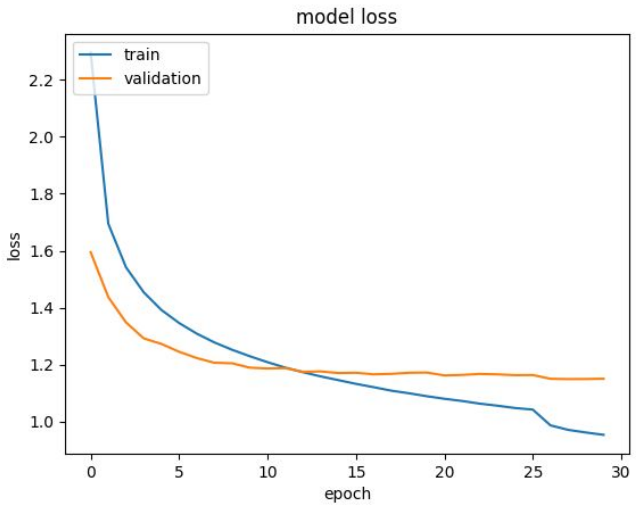
Data split

- Sklearn train_test_split 80-20 με ισορροπημένη συνεισφορά από κάθε κλάση στα training και test sets

4 Αποτελέσματα

Αποτελέσματα *(1/3)*





Αποτελέσματα (2/3)

Κλάσεις	345
Σκίτσα / κλάση	6450
Συνολικός Αριθμός σκίτσων	2.225.250
Training samples	1.602.180
Validation samples	178.020
Test samples	445.050

Metric	Classification Scores
Test Loss	1.156
Test Accuracy (top - 1)	0.719
Test Accuracy (top - 3)	0.869

Αποτελέσματα *(3/3)*

Class #	Class Name	Precision	Recall	F1 - score	Support
1	**rifle**	0.65	0.70	0.67	1274
2	**hedgehog**	0.72	0.79	0.75	1320
3	**axe**	0.79	0.78	0.78	1269
4	**floor lamp**	0.83	0.72	0.77	1306
5	**bottlecap**	0.67	0.54	0.60	1277
...	**...**	...	...	...	...
345	**monkey**	0.54	0.65	0.59	1262
	weighted avg	**0.72**	**0.72**	**0.72**	**445050**

Ερμηνεία

recall(rifle): *Από τα 1274 test samples που ανήκουν στην κλάση rifle, το μοντέλο ταξινόμησε σωστά το 70%.*

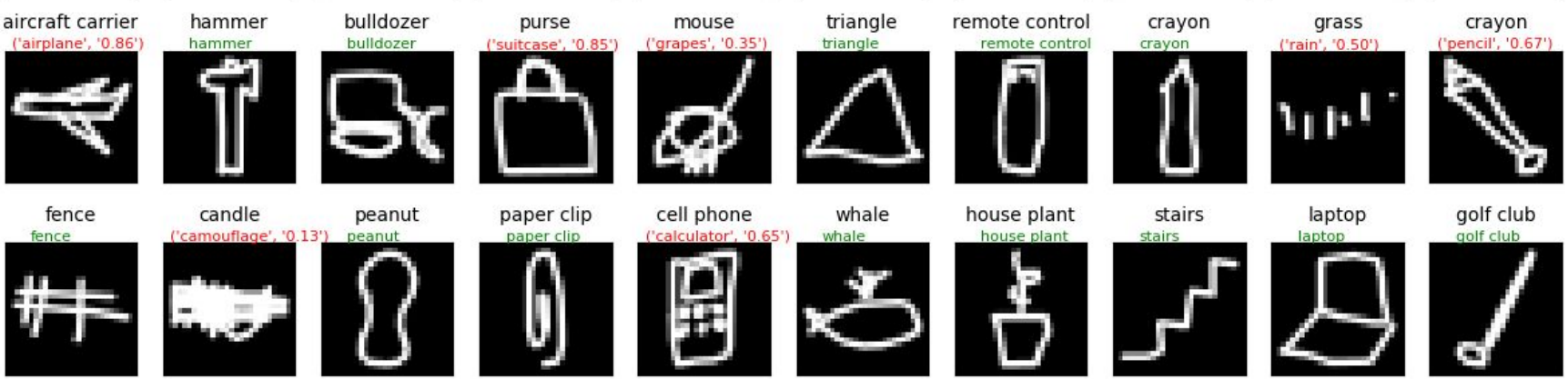
precision(rifle): *Το 65% του συνολικού αριθμού των δειγμάτων που το μοντέλο ταξινόμησε ως rifle ανήκουν όντως στην κλάση rifle.*

f1 - score(rifle) =
2 * precision(rifle) * recall(rifle) / precision(rifle) + recall (rifle)
= 2 * (0.65 * 0.7) / (0.65 + 0.7) = 0.67

Precision weighted avg =
precision(rifle)*support(rifle) +
precision(hedgehog)*support(hedgehog) +...+
precision(monkey)*support(monkey) / 345 = 0.72

https://quickdraw.withgoogle.com/data/aircraft_carrier

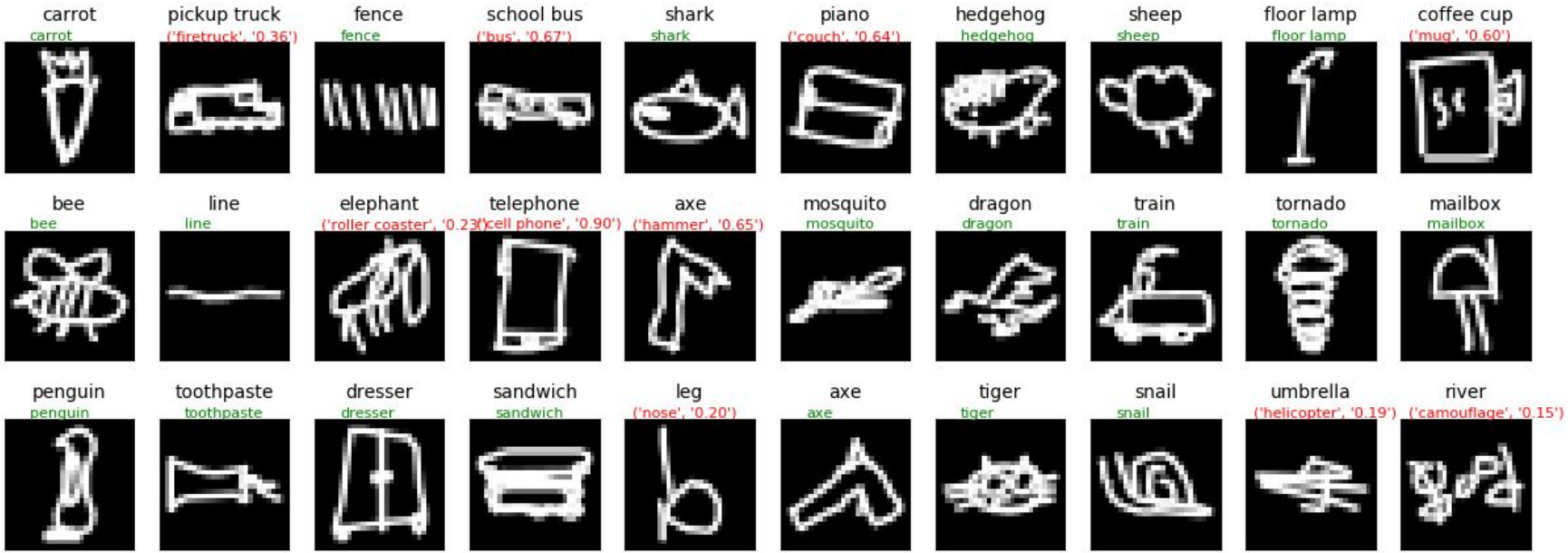
Visualizations *(1/2)*



Προβλέψεις του μοντέλου σε τυχαία δείγματα του test set (Metric: top - 1 Accuracy)

https://quickdraw.withgoogle.com/data/mug

Visualizations (2/2)



Προβλέψεις του μοντέλου σε τυχαία δείγματα του test set (Metric: top - 1 Accuracy)

Thanks!