

# Υπολογιστική Όραση (ΜΥΕ046)

## Τελική άσκηση

Ακαδημαϊκό έτος: 2019-2020  
Διδάσκων: Γιώργος Σφήκας

### Άσκηση 1

Σε αυτή την άσκηση πρέπει να χρησιμοποιήσετε ένα συνελκτικό νευρωνικό δίκτυο που έχει εκπαιδευτεί για να εκτελεί σημασιολογική κατάτμηση. Δεν χρειάζεται ούτε και πρέπει να μπειτε στη διαδικασία να εκπαιδεύσετε εκ νέου το δίκτυο (δηλαδή, να βρείτε με κάποιο τρόπο 'καλές' παραμέτρους για το δίκτυο). Η αρχιτεκτονική του δικτύου, όπως και οι παράμετροι του δικτύου θα θεωρηθούν σταθερές και γνωστές εκ των προτέρων.

Ο τρόπος που θα κληθείτε να χρησιμοποιήσετε το δίκτυο θα είναι κατ' αρχάς αποτιμώντας το αποτέλεσμα του στρώματος εξόδου του (output layer για συγκεκριμένη είσοδο-εικόνα). Επιπλέον, θα σας ζητηθεί να υπολογίσετε το αποτέλεσμα κάποιου ενδιάμεσου στρώματος, δηλαδή στρώματος που δεν αντιστοιχεί στην είσοδο ούτε στην έξοδο. Στην πρώτη περίπτωση, το αποτέλεσμα είναι μια σημασιολογική κατάτμηση της εικόνας εισόδου, οπότε θα πρέπει να πάρετε σαν αποτέλεσμα την αντιστοίχιση κάθε εικονοστοιχείου σε μια σημασιολογική κατηγορία. Στην δεύτερη περίπτωση, το αποτέλεσμα δεν είναι άμεσα χρηστικό, όμως το αποτέλεσμα μπορεί να χρησιμοποιηθεί σαν πολύ ισχυρά χαρακτηριστικά εικόνας<sup>1</sup>.

Το δίκτυο που θα χρησιμοποιήσετε είναι το DeepLab [1] και κώδικα που το υλοποιεί θα βρείτε στο <https://github.com/tensorflow/models/tree/master/research/deeplab>. Όπως θα δείτε και στο documentation του DeepLab, υπάρχει ένας αριθμός από διαφορετικές παραλλαγές-εκδοχές ανάλογα με ποιο σετ παραμέτρων και ποια συγκεκριμένη αρχιτεκτονική θα επιλέξετε να χρησιμοποιήσετε (με ονόματα τέτοια όπως Mobilenetv3, Xception71, κ.ο.κ.).

Τα ζητούμενα είναι τα παρακάτω:

**(50% του συνολικού βαθμού):** Να υπολογίσετε σημασιολογική κατάτμηση για ένα αριθμό από εικόνες που θα δώσετε στο δίκτυο σαν είσοδο (την καθεμία ξεχωριστά βεβαίως). Μπορείτε να χρησιμοποιήσετε όποια παραλλαγή της αρχιτεκτονικής επιθυμείτε. Οι εικόνες που θα πρέπει να δώσετε σαν είσοδο να είναι 5 έγχρωμες εικόνες της επιλογής σας. Να οπτικοποιήσετε την κατάτμηση-αποτέλεσμα σαν μια ψευδοχρωματισμένη εικόνα, όπου κάθε κατηγορία να αντιστοιχεί σε ξεχωριστό χρώμα της επιλογής σας (μπορείτε να το επιλέξετε τυχαία). Επίσης, να φαίνεται η αντιστοίχιση του κάθε χρώματος με την σημασιολογική κατηγορία (δηλαδή, π.χ. πράσινο = σκύλος, κόκκινο = γάτα, πορτοκαλί = άνθρωπος, μπλε = ποδήλατο, και ούτω καθεξής).

**(20% του συνολικού βαθμού):** Να υπολογίσετε για τις ίδιες εικόνες το αποτέλεσμα κάποιου ενδιάμεσου στρώματος της επιλογής σας. Αφού εκτελέσετε Principal Component Analysis (PCA) στα βαθιά χαρακτηριστικά ώστε να 'ρίξετε' την διάσταση τους σε  $d = 3$ , να οπτικοποιήσετε το αποτέλεσμα μεταχειριζοντάς το σαν να ήταν μια έγχρωμη εικόνα. Με άλλα λόγια, το κάθε ένα από τα τρία κανάλια του αποτελέσματος θα αντιστοιχεί σε ένα από τα γνώριμα Red-Green-Blue κανάλια. Το αποτέλεσμα θα είναι μια ψευδοχρωματισμένη εικόνα, όπου αντικείμενα παρόμοιας σημασιολογίας γενικά θα αντιστοιχούν σε παρόμοιο χρώμα (βλ. για παράδειγμα το αποτέλεσμα στο [2]).

**(30% του συνολικού βαθμού):** Αφού εκτελέσετε Principal Component Analysis (PCA) στα βαθιά χαρακτηριστικά ώστε να 'ρίξετε' την διάσταση τους σε  $d = 8$ , να εκτελέσετε κατάτμηση με αυτά πλέον σαν είσοδο χρησιμοποιώντας τον αλγόριθμο  $k - means$ . Χρησιμοποιήστε αριθμό τμημάτων  $k = 2$  για να δυαδικοποιήσετε την εικόνα.

Για τον PCA όπως και για τον  $k$ -means μπορείτε να χρησιμοποιήσετε έτοιμες υλοποιήσεις (π.χ. από το scikit-learn).

### Αναφορές

- [1] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 4, pp. 834–848, 2017.
- [2] Giorgos Sfikas, João Patacas, Charalampos Psarros, Antigoni Noula, Dimosthenis Ioannidis, and Dimitrios Tzovaras, "A deep neural network-based method for the detection and accurate thermography statistics estimation of aerially surveyed structures," in *19th International Conference on Construction Applications of Virtual Reality*, 2019.

<sup>1</sup>Βλ. 'Ψηφιακή Επεξεργασία Εικόνας 10b - Χαρακτηριστικά εικόνας'