

[6632] Διαχείριση και Επεξεργασία Μεγάλων Δεδομένων Παρατήρησης Γης

ΔΠΜΣ Γεωπληροφορική
ΔΠΜΣ ΕΔΕΜΜ
ΔΠΜΣ Μαθ. Προτυποποίηση

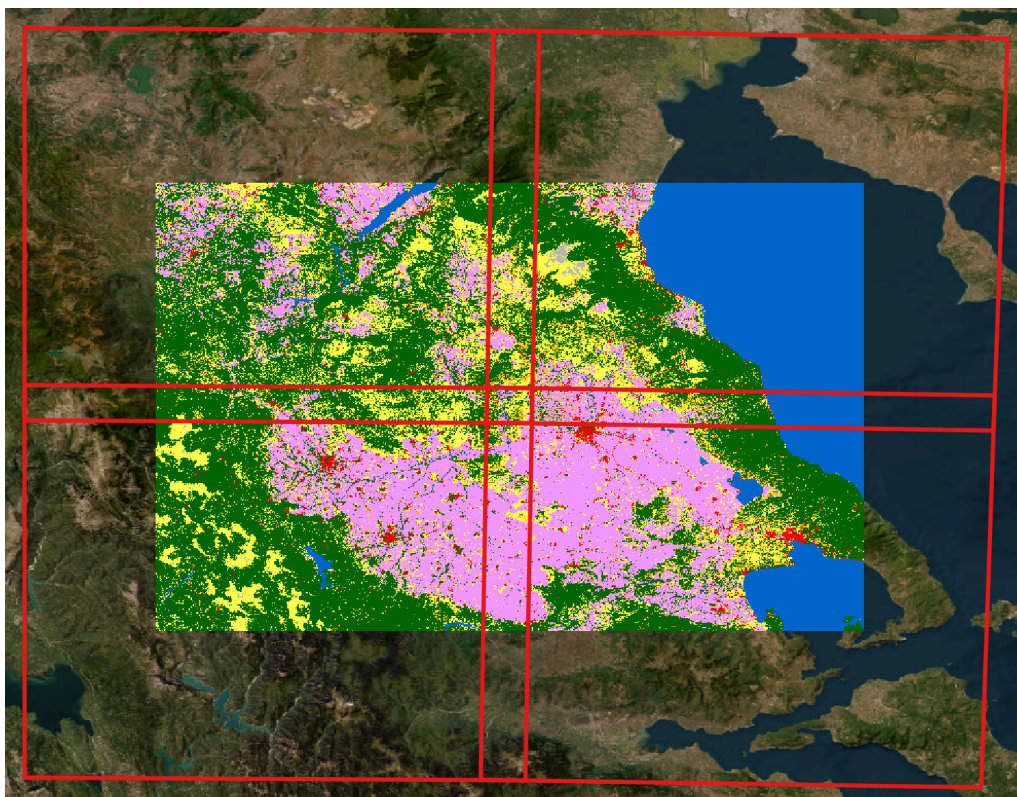
Άσκηση 2 - Χαρτογράφηση κάλυψης γης μεγάλης κλίμακας με δορυφορικές εικόνες και συνελκτικά νευρωνικά δίκτυα

10 Απριλίου, 2024

Περίληψη

Στην άσκηση 2 καλείστε να υλοποιήσετε μία ενιαία διαδικασία εκπαίδευσης και αξιολόγησης μοντέλων βαθιάς μηχανικής μάθησης με στόχο την διενέργεια αυτόματης χαρτογράφησης χρήσεων γης για μια μεγάλη περιοχή ενδιαφέροντος. Δημιουργώντας ένα δικό σας σετ δεδομένων από δορυφορικές απεικονίσεις της δορυφορικής πλατφόρμας Sentinel-2 και σε συνδυασμό με τα παρεχόμενα δεδομένα αληθείας καλείστε να εκπαιδεύσετε και να αξιολογήσετε ένα μοντέλο πλήρως συνελκτικών τεχνητών νευρωνικών δικτύων τύπου U-Net αξιοποιώντας μερικώς και τεχνικές μεταφοράς μάθησης.

Ως τελικό βήμα καλείστε να χρησιμοποιήσετε το μοντέλο σας ώστε να χαρτογραφήσετε μια διαφορετική περιοχή ενδιαφέροντος, αναπτύσσοντας μία ολοκληρωμένη λύση χαρτογράφησης.



Στόχοι

1. **Απόκτηση και προ-επεξεργασία γεωχωρικών δεδομένων τύπου raster:** Εξοικείωση με τη γεωχωρική διάσταση των τηλεπισκοπικών δεδομένων. Pansharpening, επαναπροβολή και διαχείριση συστημάτων αναφοράς. Γεωχωρικά ερωτήματα και κοπή εικόνας.
2. **Προετοιμασία δεδομένων για εκπαίδευση:** Εξοικείωση με την τυπική διαδικασία που ακολουθείται στην περίπτωση της επιβλεπόμενης μάθησης. Δημιουργία σετ δεδομένων. Χωρισμός σε υποσύνολα. Κανονικοποίηση δεδομένων. Τεχνικές επαύξησης δεδομένων για προβλήματα κατάτμησης. Ιδιαιτερότητες δορυφορικών απεικονίσεων. Ανάπτυξη αλγορίθμων τροφοδότησης δεδομένων.
3. **Σχεδιασμός μοντέλου U-Net:** Εξοικείωση με μοντέλα B.M.M. για κατάτμηση εικόνας. Μεταφορά μάθησης. Προ-εκπαιδευμένα μοντέλα για πολυφασματικά δεδομένα. Σχεδιασμός αρχιτεκτονικής.
4. **Εκπαίδευση και αξιολόγηση:** Εξοικείωση με τη διαδικασία εκπαίδευσης ενός ΤΝΔ. Επιλογή μεθόδου βελτιστοποίησης. Παρακολούθηση επιδόσεων μοντέλου. Μετρικές αξιολόγησης. Ρύθμιση υπερπαραμέτρων.
5. **Πρόβλεψη σε νέα δεδομένα:** Εξοικείωση με την εφαρμογή του μοντέλου για πρόβλεψη σε νέα δεδομένα. Πρόβλεψη σε υποπεριοχές. Τελική σύνθεση προϊόντος. Ποιοτική αξιολόγηση. Εκτίμηση σφάλματος γενίκευσης.

➤Οδηγίες / ★Προδιαγραφές/•Tips εκτέλεσης

Απόκτηση και προ-επεξεργασία γεωχωρικών δεδομένων τύπου raster:

- Λήψη δεδομένων αληθείας. [link](#)
 - Ονόματα / Τιμές κατηγοριών στα μεταδεδομένα του αρχείου
 - Χρονολογία αναφοράς 2021
- Λήψη προϊόντων L1C (Top-of-atmosphere reflectance) του Sentinel-2. Χρήσιμο [link](#)
 - ★ Το σύνολο των δεδομένων πρέπει να **καλύπτει πλήρως** την περιοχή για την οποία είναι διαθέσιμα τα δεδομένα αληθείας. Χρήσιμο [link](#)
 - ★ Η συνολική **νεφοκάλυψη** να μην υπερβαίνει το 10% ανά πλακίδιο (tile)
- Προεπεξεργασία δεδομένων
 - ★ Εφαρμογή τεχνικής **pansharpening** (έστω και απλοϊκότατης) ώστε να έχουν όλα τα κανάλια 10μ χωρική ανάλυση.
 - Προσοχή στη σειρά των καναλιών
 - ★ **Ευθυγράμμιση** δορυφορικών δεδομένων και δεδομένων αληθείας. Να λάβετε σοβαρά υπόψιν σας ότι τα δεδομένα έχουν γεωχωρική υπόσταση
- Δημιουργία σετ δεδομένων
 - ★ Δημιουργήστε ζευγάρια εικόνων δορυφορικών δεδομένων - δεδομένων αληθείας όπου υπάρχει 1-1 αντιστοιχία σε επίπεδο εικονοστοιχείου.

- Συνοδευτικά μπορείτε να δημιουργήσετε και μάσκες εγκυρότητας δεδομένων (valid masks)
- Η διαδικασία μπορεί να γίνει είτε προγραμματιστικά είτε σε περιβάλλον GIS. Στη δεύτερη περίπτωση να συμπεριλάβετε στην τεχνική σας έκθεση screenshots “βήμα-βήμα” της διαδικασίας που ακολουθήσατε καθώς και σχετική περιγραφή.

Προετοιμασία δεδομένων για εκπαίδευση:

- Χωρισμός σε υποσύνολα
- Κανονικοποίηση δεδομένων
 - Λάβετε υπόψη ότι τα δεδομένα τροφοδοτούνται σε προεκπαιδευμένο τμήμα δικτύου
- Επαύξηση δεδομένων
 - ★ Τουλάχιστον 1 μέθοδος γεωμετρικής και 1 μέθοδος ραδιομετρικής επαύξησης δεδομένων. Χρήσιμο [link](#)
- Προετοιμασία αλγορίθμων τροφοδότησης δεδομένων

Σχεδιασμός μοντέλου U-Net:

- Σχεδιασμός αρχιτεκτονικής για σημασιολογική κατάτμηση εικόνας τύπου U-Net
 - ★ 13 κανάλια εισόδου με 10μ χωρική ανάλυση ανά εικονοστοιχείο
 - ★ (Μεταφορά μάθησης) Το encoder μέρος του U-Net θα αποτελείται **αποκλειστικά** από στρώσεις **προ-εκπαιδευμένου** μοντέλου τύπου ResNet 50 ή 18. Χρήσιμο [link](#)
 - ★ (Σχεδιασμός αρχιτεκτονικής) Τουλάχιστον δύο (2) skip connections στο U-Net για μεταφορά χαρακτηριστικών συγκεκριμένης κλίμακας. Τα skip connections του ResNet δεν προσμετρώνται!

Εκπαίδευση και αξιολόγηση:

- Εκπαίδευση μοντέλου
 - Χρησιμοποιήστε την pytorch lightning ή παρόμοιο
 - ★ Παρακολούθηση τουλάχιστον δύο σχετικών (2) **μετρικών** αξιολόγησης κατά τη διάρκεια της εκπαίδευσης στο σετ ελέγχου (validation). Χρήσιμο [link](#)
 - ★ Δημιουργία τουλ. 2 + 1 **γραφημάτων** τιμών μετρικών / συνάρτησης κόστους ανά εποχή εκπαίδευσης
 - ★ Διενέργεια πολλαπλών πειραμάτων για τη ρύθμιση τουλάχιστον μίας (1) **υπερπαραμέτρου** (πλην του πλήθους εποχών εκπαίδευσης)
 - ★ Αποθήκευση του **καλύτερου** μοντέλου

Πρόβλεψη σε νέα δεδομένα:

- Λήψη προϊόντων L2A (surface reflectance) του Sentinel-2
 - ★ Λήψη ενός (1) πλακιδίου (χρον. 2021) με κωδικό **T34SEH (Αχαΐα)**
- Δημιουργία χάρτη προβλέψεων

- ★ Υλοποίηση μεθοδολογίας πρόβλεψης σε υπο-περιοχές και σύνθεση τελικού χάρτη προβλέψεων βάσει **καλύτερου** μοντέλου
- ★ Εφαρμόστε κατάλληλη **επικάλυψη** υπο-περιοχών ώστε να είναι “**συνεχείς**” οι προβλέψεις.
 - Εκτιμήστε το receptive field της αρχιτεκτονικής σας
- ★ **Ποιοτική αξιολόγηση** μέσω φωτοερμηνείας
- ★ **Ποσοτική αξιολόγηση** με τουλ. 2 σχετικές μετρικές
 - Τα δεδομένα αληθείας για την περιοχή αυτή βρίσκονται στο [link](#)

Σημειώσεις

1. Οι οδηγίες και τα tips είναι ενδεικτικά. Το παραδοτέο οφείλει να εκπληρώνει τις δοθείσες προδιαγραφές (★)
2. Ως **παραδοτέο** θεωρούνται τα εξής:
 - i. **Τεχνική έκθεση** με επαρκή περιγραφή της διαδικασίας ανά βήμα. Συμπεριλάβετε εικόνες με αποτελέσματα και γραφήματα. Εκτενής ποιοτική αξιολόγηση αποτελεσμάτων στην περιοχή των νέων δεδομένων
 - ii. Αρχεία πηγαίου **κώδικα** σε μορφή .py ή .ipynb
 - iii. Μία (1) **εικόνα-χάρτης κάλυψης γης** για την περιοχή “πρόβλεψης”
3. Δεκτή κάθε λύση σε pytorch/tensorflow/jax/keras