

Μάθημα: Διαχείριση και Επεξεργασία Μεγάλων Δεδομένων Παρατήρησης Γης

ΔΠΜΣ: ΕΔΕΜΜ

Άσκηση: 1η - Χαρτογράφηση κάλυψης γης μεγάλης κλίμακας με δορυφορικές εικόνες και συνελκτικά νευρωνικά δίκτυα

Ονοματεπώνυμο: Στέφανος Πέτκοβιτς

Αριθμός Μητρώου: 03400281

Ημερομηνία Παράδοσης: 4/5/2025

Περίληψη Άσκησης

Στην παρούσα άσκηση, κατόπιν της λήψης 4 πολυφασματικών εικόνων Sentinel-2 που απεικονίζουν την περιοχή της Θεσσαλίας και επικαλύπτονται με έναν χάρτη “αληθείας” (ground truth) ο οποίος απεικονίζει σε επίπεδο εικονοστοιχείου διακριτές κλάσεις εδαφικής κάλυψης (αστικός ιστός, δασικές εκτάσεις κ.ο.κ.), εφαρμόστηκαν τα παρακάτω βήματα για την εκπαίδευση ενός συνελκτικού νευρωνικού δικτύου (CNN) τύπου Unet το οποίο θα έχει την δυνατότητα να υλοποιεί μια εργασία σημασιολογικής κατάτμησης όπου δοθέντος τουλάχιστον μιας πολυφασματικής απεικόνισης Sentinel-2, αναθέτει σε κάθε εικονοστοιχείο της απεικόνισης μια κλάση:

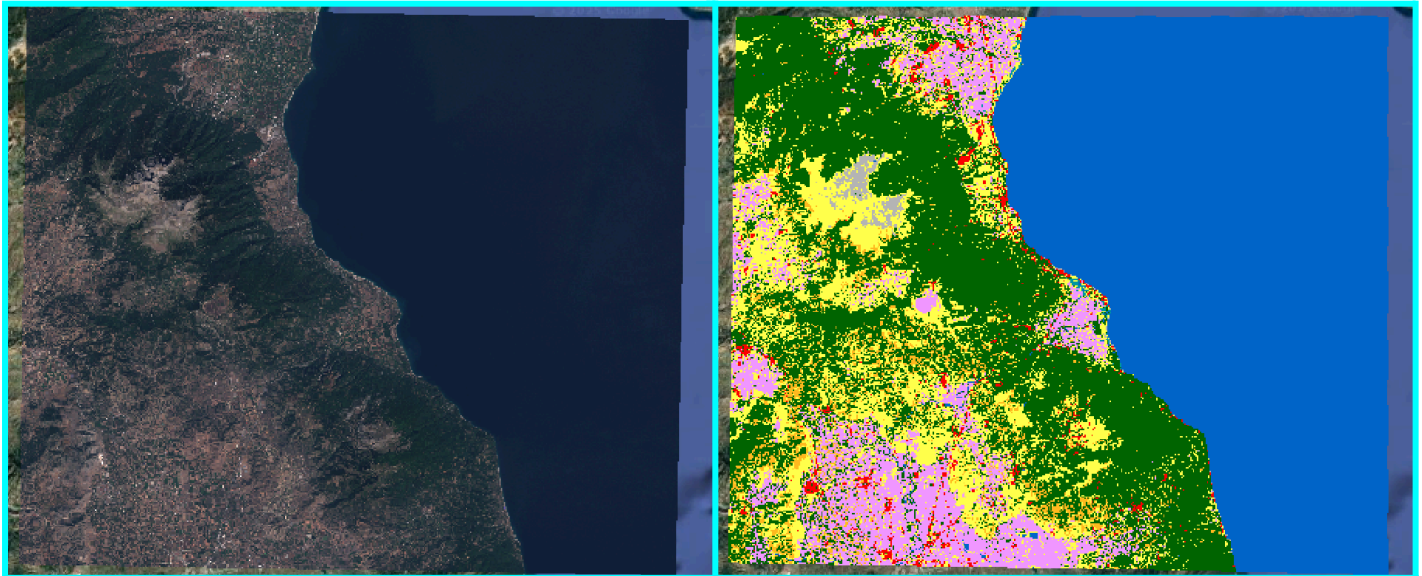
- Προεπεξεργασία πολυφασματικών εικόνων και χάρτη “αληθείας” έτσι ώστε να υπάρχει ταύτιση 1-1 σε επίπεδο εικονοστοιχείου.
- Κατασκευή συνόλων δεδομένων τα οποία θα εισαχθούν στο Unet CNN για εκπαίδευση ή πρόβλεψη.
- Ορισμός αρχιτεκτονικής Unet CNN, πραγματοποίηση δοκιμών εκπαίδευσης και αποθήκευση μοντέλου με την καλύτερη απόδοση βάσει συγκεκριμένων μετρικών.
- Πρόβλεψη κλάσεων σε εικόνα Sentinel-2 που απεικονίζει διαφορετική περιοχή (T34SEH - Αχαΐα) και ανάλυση απόδοσης βάσει συγκεκριμένων μετρικών.

Απόκτηση και Προ-Επεξεργασία Γεωχωρικών Δεδομένων Τύπου Raster

Κατόπιν της λήψης των τεσσάρων πολυφασματικών απεικονίσεων Sentinel-2 (ανήκουν σε 2 διαδοχικά και επικαλυπτόμενα στα άκρα τους tiles) οι οποίες αλληλοεπικαλύπτονται με τα δεδομένα αλήθειας (θεματικός χάρτης με 8 κλάσεις κάλυψης γης ανά εικονοστοιχείο), πραγματοποιήθηκαν τα ακόλουθα βήματα προ-επεξεργασίας των απεικονίσεων μέσω των μεθόδων της κλάσης *PreprocessRasters*:

1. Αποσυμπίεση αρχικού φακέλου και επαναπροσαρμογή (resampling) καναλιών των Sentinel-2 απεικονίσεων των οποίων η χωρική ανάλυση είναι μεγαλύτερη των 10 μέτρων. Η μέθοδος επαναπροσαρμογής που επιλέχθηκε ήταν η διγραμμική (bilinear) για την ομαλή “σμίκρυνση” των εικονοστοιχείων.
2. Αλλαγή συστήματος αναφοράς χάρτη “αληθείας” από το γεωγραφικό WGS 84’ (EPSG:4326) στο προβολικό σύστημα αναφοράς με κωδικό (EPSG: 32634) το οποίο είναι και το σύστημα αναφοράς των Sentinel-2 απεικονίσεων. Η συγκεκριμένη “επαναπροβολή/reprojection” επιλέχθηκε για την επίτευξη όσο το δυνατόν καλύτερης ταύτισης μεταξύ των απεικονίσεων Sentinel-2 και του χάρτη “αληθείας” σε επίπεδο εικονοστοιχείου. Επισημαίνεται ότι κατά την “επαναπροβολή” του χάρτη “αληθείας” εφαρμόστηκε και η μέθοδος επαναπροσαρμογής του εγγύτερου γείτονα έτσι ώστε να μην υπάρχει αλλοίωση των κατηγορικών τιμών των εικονοστοιχείων του.
3. Το τελευταίο βήμα προ-επεξεργασίας είναι η δημιουργία τεσσάρων συνόλων Sentinel-2 απεικόνισης και χάρτη “αληθείας”. Η δημιουργία των

συγκεκριμένων συνόλων επιτεύχθηκε μέσω 2 διαδοχικών “κοψιμάτων” των απεικονίσεων και των χαρτών έτσι ώστε να υπάρχει πλήρης ταύτιση μεταξύ των αντικειμένων του συνόλου αναφορικά με τα όρια τους.



Σχήμα 1. Παράδειγμα συνόλου τμήματος απεικόνισης Sentinel-2 και αντίστοιχου τμήματος χάρτη “αληθείας” με αντιστοιχία ορίων

Προετοιμασία Δεδομένων για Εκπαίδευση

Η προετοιμασία των συνόλων για να εισαχθούν σαν δεδομένα εκπαίδευσης στο συνελικτικό νευρωνικό δίκτυο Unet (του οποίου η αρχιτεκτονική καθώς και η διαδικασία εκπαίδευσης θα αναλυθούν στην επόμενη ενότητα) πραγματοποιήθηκε μέσω των μεθόδων των κλάσεων *CreateDataset* και *PatchDataset*. Αρχικά μέσω της μεθόδου *create_patch_index* της κλάσης *CreateDataset*, για όλα τα σύνολα δημιουργούνται τετραγωνικά “τεμάχια/patches” (στην εφαρμογή δημιουργήθηκαν patches μεγέθους 256x256). Έπειτα, μέσω της μεθόδου *create_splits* χωρίζονται τα patches των συνόλων απεικόνισης - χάρτη “αληθείας” στα σετ εκπαίδευσης (training) και επικύρωσης (validation) με ποσοστά 30% για επικύρωση και 70% για εκπαίδευση. Τέλος, μέσω της *CreateDataset* αρχικοποιείται και η σειρά των επαυξήσεων που εφαρμόζονται στα δεδομένα εκπαίδευσης και αληθείας. Ειδικότερα, εφαρμόζονται 2 γεωμετρικές (οριζόντια και κατακόρυφη περιστροφή) και 2 ραδιομετρικές (μεταβολή αντίθεσης/φωτεινότητας και κανονικοποίηση/blur) επαυξήσεις, με τις ραδιομετρικές επαυξήσεις να εφαρμόζονται μόνο στα patches των απεικονίσεων για να μην μεταβληθούν οι κατηγορικές τιμές των εικονοστοιχείων στα patches από τον χάρτη “αληθείας”. Όλα τα παραπάνω εφαρμόζονται μέσω της κλάσης *PatchDataset*.

Σχεδιασμός Μοντέλου U-net

Ο “κορμός” της αρχιτεκτονικής του μοντέλου U-net είναι οι μικρότερες encoder και decoder αρχιτεκτονικές οι οποίες δίνουν την δυνατότητα παραγωγής χάρτη κλάσεων ο οποίος έχει τις ίδιες χωρικές ιδιότητες με την είσοδο. Επομένως, δοθέντος ενός patch απεικόνισης Sentinel-2 256x256 η έξοδος θα είναι ένας χάρτης 256x256 με ένα ορισμένο πλήθος κλάσεων. Παρακάτω δίνεται μια συνοπτική περιγραφή για τα επιμέρους στοιχεία της αρχιτεκτονικής U-net (ορίζονται και σαν ξεχωριστές κλάσεις στο script *C_Unet_Arch.py*):

1. **Block Συνελίξεων**: Η πιο θεμελιώδης συνιστώσα της αρχιτεκτονικής είναι ένα block το οποίο αποτελείται από 2 επίπεδα συνελίξης τα οποία συνδέονται με μια ReLU

συνάρτηση ενεργοποίησης. Επισημαίνεται ότι στον decoder τα block εκτελούν ανάστροφη συνέλιξη για να μειώσουν τα επίπεδα που έχουν δημιουργηθεί από τις συνέλιξεις των encoder blocks.

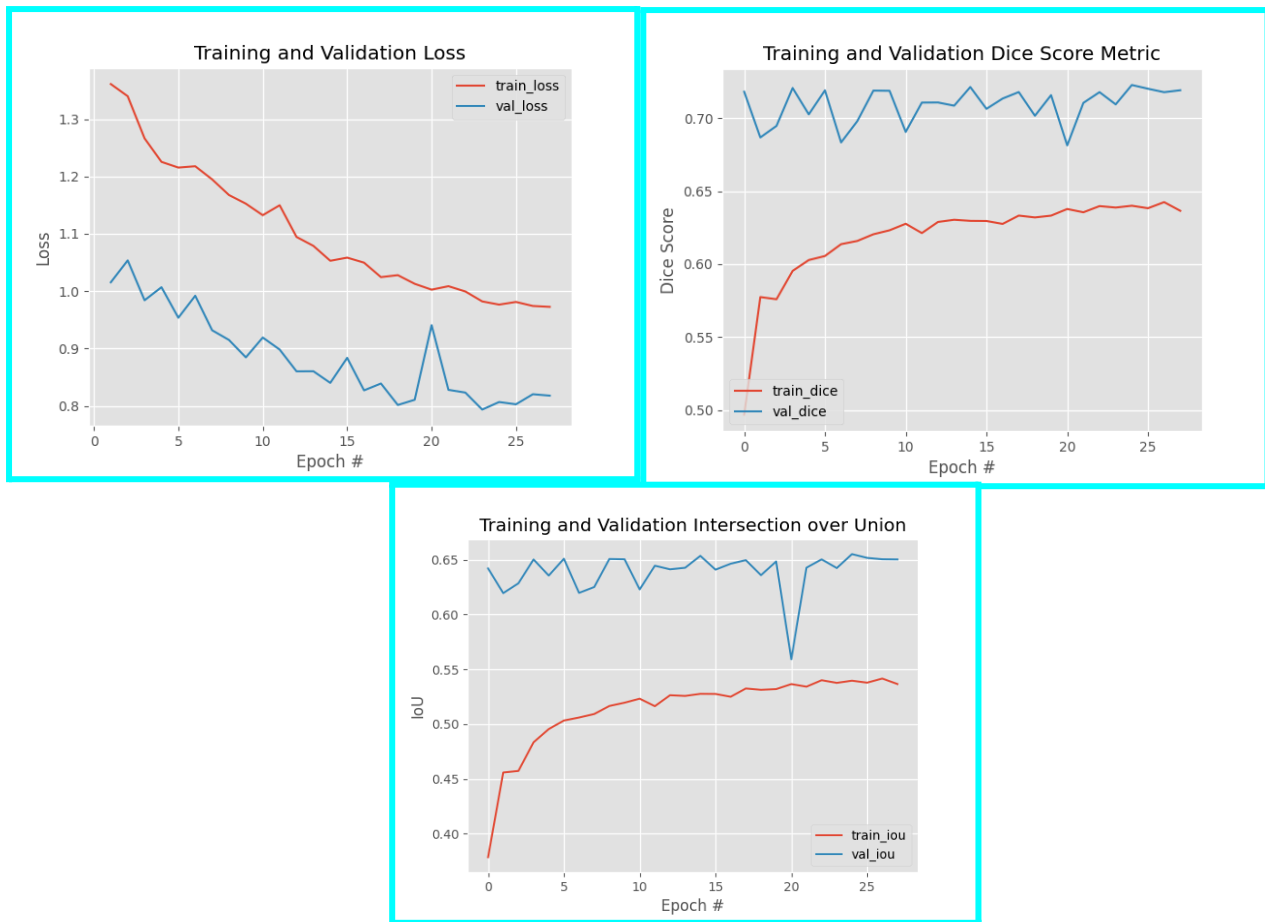
2. **Encoder:** Ο encoder ουσιαστικά αποτελεί μια συστοιχία από συνελικτικά blocks τα οποία συνδέονται μεταξύ τους με max pooling operators. Αυτή η διαδικασία έχει σαν αποτέλεσμα την αύξηση των επιπέδων μέχρις ότου κάθε εικονοστοιχείο της εισόδου να αποτελεί από μόνο του έναν “νευρώνα” στο πλήρως συνδεδεμένο επίπεδο. Επίσης, για να μπορέσουν να εισαχθούν τα βάρη του προ-εκπαιδευμένου δικτύου MoCo ResNet 50, κληρονομήθηκε η αρχιτεκτονική (αριθμός και τύπος επιπέδων, παράμετροι φίλτρου συνέλιξης κ.ο.κ.) ενός residual network.
3. **Decoder:** Ο decoder περιέχει τον ίδιο αριθμό block με τον encoder αλλά όπως αναφέρθηκε και παραπάνω οι επιμέρους τελεστές εκτελούν ανάστροφες συνέλιξεις για την σταδιακή αύξηση της διαστατικότητας των διαδοχικών επιπέδων μέχρις ότου να δοθεί σαν έξοδος ο χάρτης προβλέψεων με τον καθορισμένο αριθμό κλάσεων. Επίσης, στην *forward* μέθοδο της σχετικής κλάσης υλοποιούνται και τα skip connections με τα blocks του encoder μέσω του τελεστή *center_crop*, με την έξοδο του να συνενώνεται με το αντίστοιχο block του decoder.

Εκπαίδευση και Αξιολόγηση

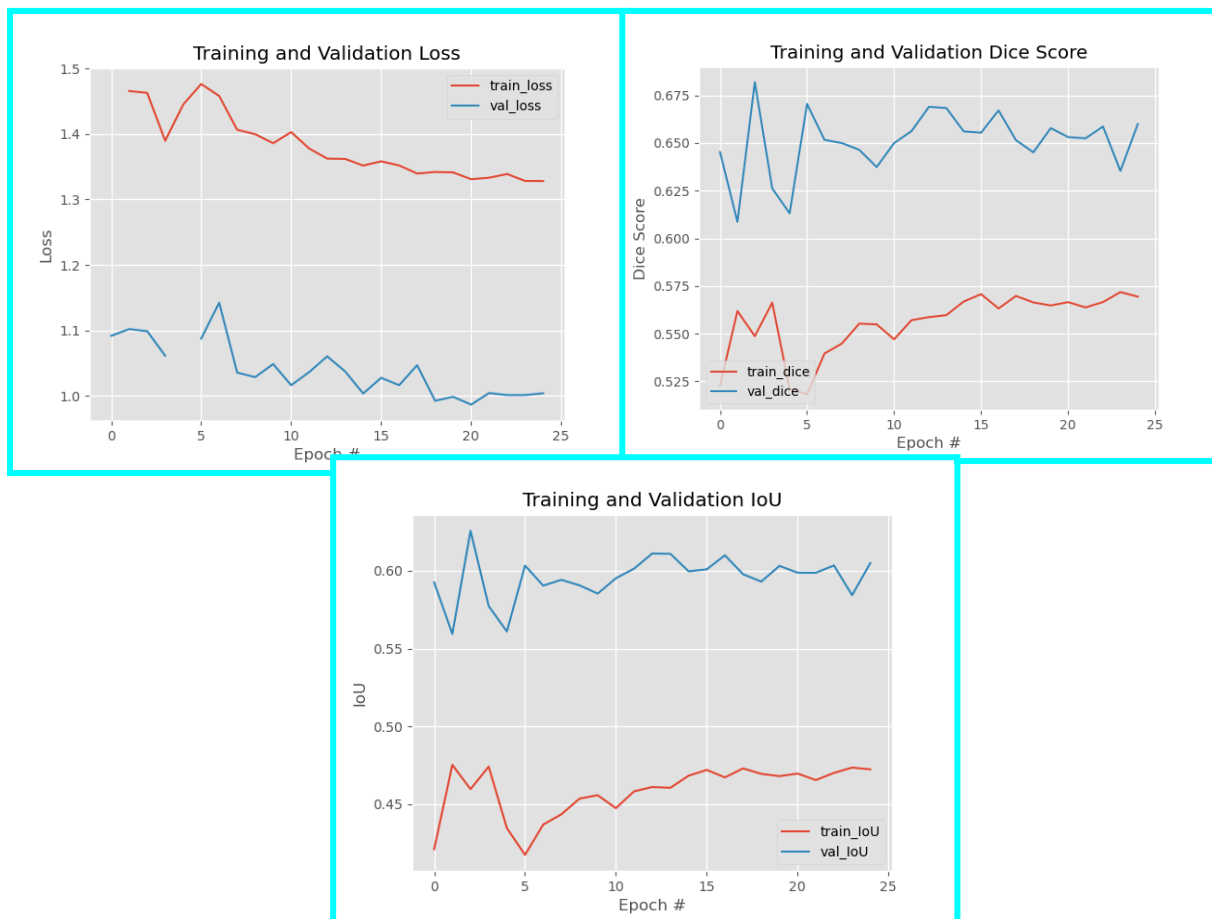
Κατόπιν εκτέλεσης των παραπάνω βημάτων μέσω των σχετικών κλάσεων και μεθόδων που δημιουργήθηκαν, έλαβε χώρα η εκπαίδευση του μοντέλου με 3 διαφορετικούς συνδυασμούς των υπερπαραμέτρων του batch size, του πλήθους των εποχών και του ρυθμού μάθησης. Επιπλέον, πριν την δημιουργία του επαναληπτικού βρόχου εκπαίδευσης (training loop) προσαρμόστηκε το πρώτο συνελικτικό επίπεδο των blocks του encoder έτσι ώστε να έχει 13 κανάλια και να μπορεί να χρησιμοποιεί τα προ-εκπαιδευμένα βάρη ενώ ορίστηκαν και οι παρακάτω σημαντικές παράμετροι εκπαίδευσης:

- **Συνάρτηση Κόστους Διασταυρούμενης Εντροπίας (Cross Entropy):** Η συγκεκριμένη συνάρτηση κόστους ενδείκνυται σε προβλήματα πολλαπλής ταξινόμησης.
- **Βελτιστοποιητής Adam:** Ο συγκεκριμένος βελτιστοποιητής (optimizer) βοηθά στην εύρεση του καλύτερου δυνατού ελαχίστου της συνάρτησης κόστους συνδυάζοντας τις τεχνικές της ορμής και του RMSprop.
- **Scheduler Ρυθμού Μάθησης (Reduce on Plateau):** Ο scheduler έχει σαν εργασία την μείωση του ρυθμού μάθησης όταν παύει να μειώνεται η συνάρτηση κόστους ή άλλη σχετική μετρική.

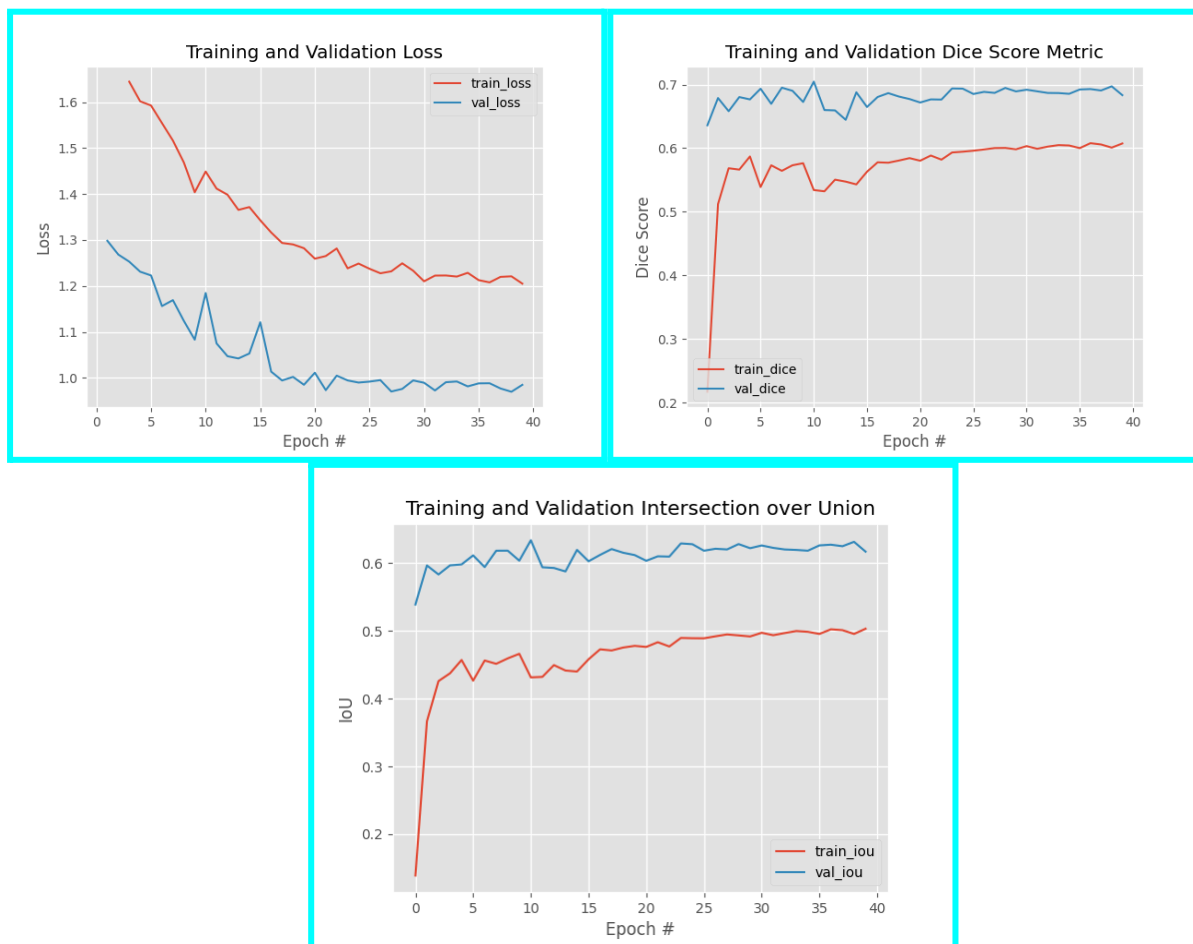
Προχωρώντας στην διαδικασία της εκπαίδευσης, για κάθε εποχή υπολογίζεται (σαν τον μέσο όρο μεταξύ των batches) η απώλεια/κόστος στα σύνολα εκπαίδευσης και επικύρωσης καθώς και οι μετρικές dice score και intersection over union (IoU). Αναφέρεται επιγραμματικά ότι η μετρική dice μεταξύ του χάρτη κατάτμησης “αληθείας” A και του χάρτη κατάτμησης που προκύπτει από την διαδικασία της εκπαίδευσης B, είναι το πηλίκο $2 * |A \cap B| / (|A| + |B|)$ το οποίο δείχνει πόσο όμοιοι είναι οι 2 χάρτες ενώ η μετρική IoU ορίζεται από την σχέση $|A \cap B| / |A \cup B|$ και ουσιαστικά αποτελεί μια καλή εκτίμηση για το ποσοστό χωρικής ταύτισης μεταξύ των απεικονιζόμενων κλάσεων. Τέλος, εφαρμόζεται και η τεχνική early stopping έτσι ώστε η εκπαίδευση να διακόπτεται όταν δεν υπάρχει βελτίωση της απώλειας/κόστους στο σύνολο επικύρωσης μετά από έναν αριθμό διαδοχικών εποχών. Παρακάτω παρουσιάζονται τα διαγράμματα απεικόνισης των μετρικών καθώς και της απώλειας/κόστους συναρτήσεις των εποχών εκπαίδευσης για κάθε έναν από τους τρεις συνδυασμούς υπερπαραμέτρων.



Σχήμα 2. Διάγραμμα απεικόνισης απώλειας/κόστους (πάνω αριστερά εικόνα), dice score (πάνω δεξιά εικόνα) και intersection over union (κάτω εικόνα) συναρτήσεως των εποχών εκπαίδευσης για batch size = 8 και ρυθμό εκπαίδευσης = 0.0005.



Σχήμα 3. Διάγραμμα απεικόνισης απώλειας/κόστους (πάνω αριστερά εικόνα), dice score (πάνω δεξιά εικόνα) και intersection over union (κάτω εικόνα) συναρτήσεως των εποχών εκπαίδευσης για batch size = 16 και ρυθμό εκπαίδευσης = 0.0005.



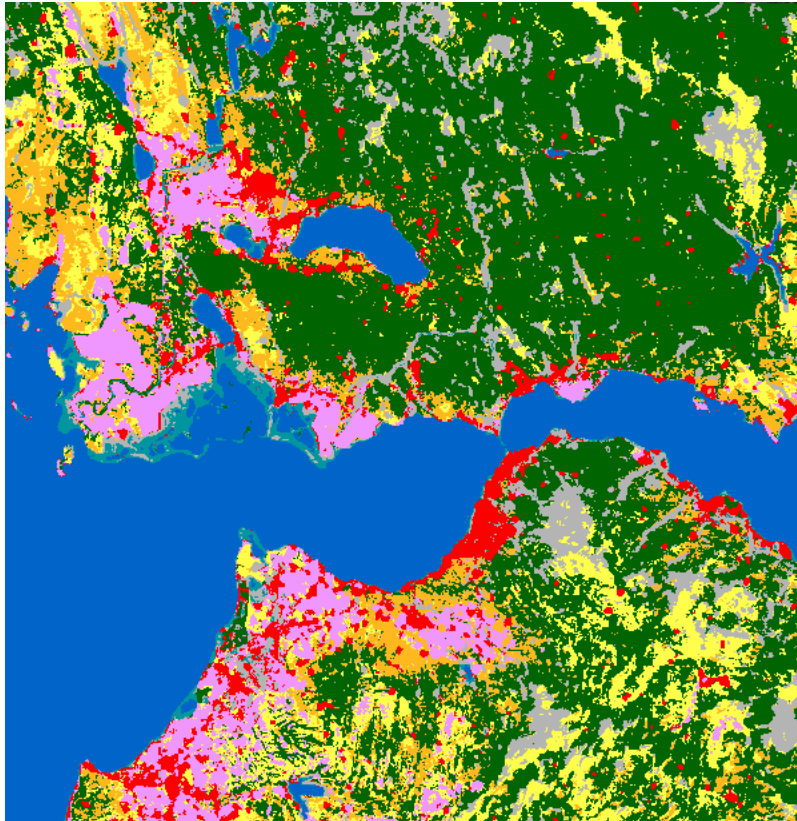
Σχήμα 4. Διάγραμμα απεικόνισης απώλειας/κόστους (πάνω αριστερά εικόνα), dice score (πάνω δεξιά εικόνα) και intersection over union (κάτω εικόνα) συναρτήσεως των εποχών εκπαίδευσης για batch size = 32 και ρυθμό εκπαίδευσης = 0.001.

Από τα παραπάνω διαγράμματα διακρίνεται ότι η επίδοση του μοντέλου για batch size = 8 είναι η καλύτερη όσον αφορά τις σχετικές μετρικές (μέγιστες τιμές dice score και intersection over union) καθώς και την ελαχιστοποίηση της απώλειας/κόστους και στα 2 σύνολα που αναλύονται (επομένως αυτό θα επιλεγεί). Μια σημαντική παρατήρηση που μπορεί να γίνει σε σχέση με την διαδικασία της εκπαίδευσης αφορά την σταθερή απόκλιση στις υπολογιζόμενες ποσότητες μεταξύ των συνόλων εκπαίδευσης και επικύρωσης, κάτι το οποίο κατά πάσα πιθανότητα δεν οφείλεται σε υπο-προσαρμογή (underfitting) αλλά στην επίδραση των επαυξήσεων στο σύνολο εκπαίδευσης (δεν εφαρμόζονται στο σύνολο επικύρωσης) καθώς και στην χρήση προ-εκπαιδευμένων βαρών και της εφαρμογής batch normalization στον encoder του Unet.

Πρόβλεψη σε Νέα Δεδομένα

Η ανάλυση της προβλεπτικής ικανότητας του μοντέλου με τον συνδυασμό υπερπαραμέτρων που επιτυγχάνει την βέλτιστη επίδοση υλοποιήθηκε με την παραγωγή ενός χάρτη

προβλέψεων από μια διαφορετική περιοχή σε σχέση με αυτή που έλαβε χώρα η εκπαίδευση. Επομένως, ελήφθη και πέρασε από τις απαιτούμενες επεξεργασίες (χωρισμός σε 25 patches χωρικών διαστάσεων 2096x2096 μέτρα) μια απεικόνιση Sentinel-2 της περιοχής της Αχαΐας. Έπειτα, το προκύπτον σύνολο δεδομένων δίνεται σαν είσοδο στο Unet του οποίου οι παράμετροι εισάγονται από το καλύτερο μοντέλο της προηγούμενης ενότητας και παράγεται ο χάρτης προβλέψεων που φαίνεται παρακάτω.

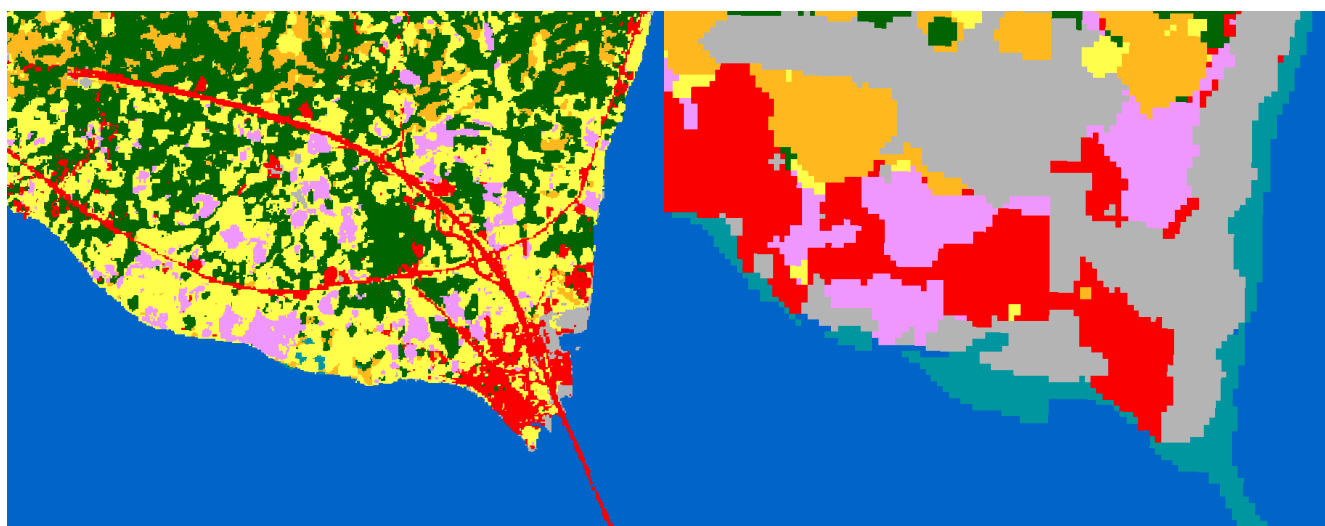


Σχήμα 5. Απόσπασμα χάρτη πρόβλεψης παραγόμενου από το καλύτερο μοντέλο Unet

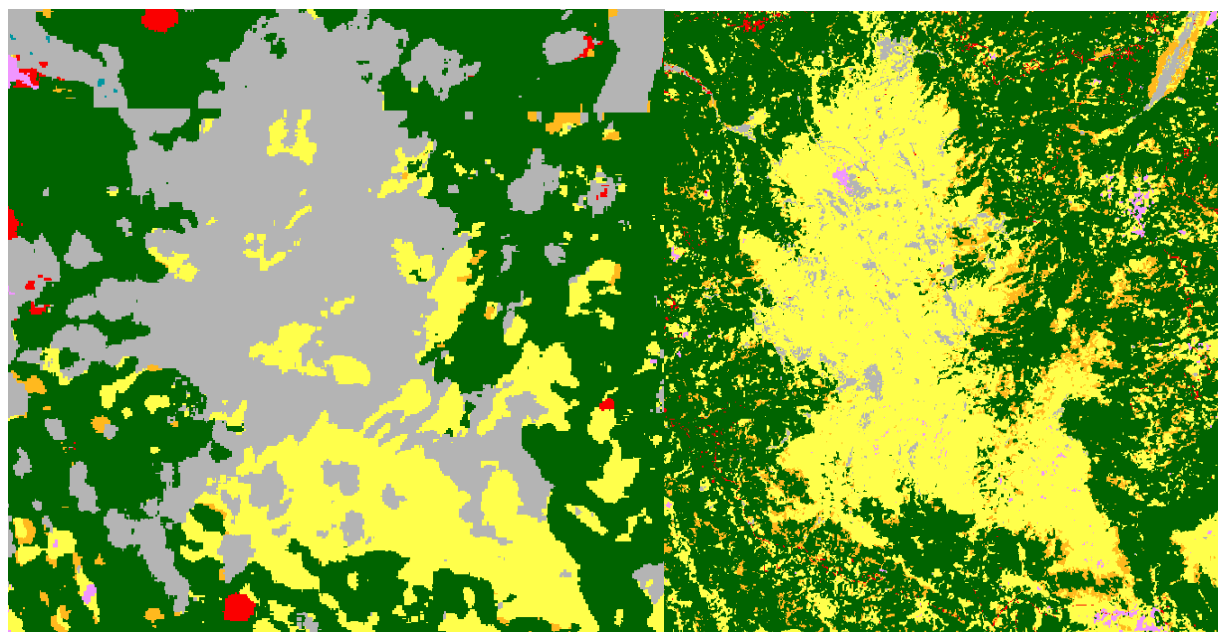
Μια γενική παρατήρηση που μπορεί να γίνει συγκρίνοντας τον παραγόμενο χάρτη πρόβλεψης με τον διαθέσιμο χάρτη “αληθείας” είναι το γεγονός ότι αυτός μπορεί μόνο “χονδρικά” να προβλέπει τις κλάσεις σε επίπεδο εικονοστοιχείου. Παραδείγματος χάρη, οι δομημένες και οι αγροτικές εκτάσεις απεικονίζονται αρκετά μεγαλύτερες από όσο είναι στην πραγματικότητα και το μοντέλο δεν μπορεί να προβλέψει με επιτυχία σχετιζόμενες οντότητες που καταλαμβάνουν λίγα εικονοστοιχεία όπως είναι το οδικό δίκτυο. Επίσης, παρατηρείται ότι δημιουργείται και αρκετή σύγχυση μεταξύ μεταξύ μερικών κλάσεων όπως είναι οι θαμνότοποι και οι περιοχές με αραιή/χαμηλή ή καθόλου βλάστηση καθώς και οι υδάτινες επιφάνειες και οι υγροβιότοποι κοντά στην ακτογραμμή. Η συγκεκριμένη συμπεριφορά του μοντέλου μπορεί να ερμηνευθεί από τον πίνακα που φαίνεται παρακάτω όπου οι περισσότερες κλάσεις (πλην των δασικών και υδατινών επιφανειών που κατανέμονται πιο “ομοιόμορφα” χωρικά) εντοπίζονται με αρκετά χαμηλή ακρίβεια βάσει των μετρικών dice score και intersection over union καθώς το μοντέλο αποτυγχάνει να αναγνωρίσει τα αληθή τους χωρικά πρότυπα.

<u>Κλάση</u>	<u>Dice Score</u>	<u>IoU</u>
<u>Δασική Έκταση</u>	0.72	0.62
<u>Θαμνότοπος</u>	0.21	0.13
<u>Χορτολιβαδική Έκταση</u>	0.36	0.32
<u>Αγροτική Έκταση</u>	0.41	0.35
<u>Δομημένη Έκταση</u>	0.19	0.13
<u>Αραιή/Χαμηλή Βλάστηση</u>	0.09	0.05
<u>Μόνιμες Υδάτινες Επιφάνειες</u>	0.91	0.94
<u>Υγροβιότοποι</u>	0.09	0.19

Πίνακας 1. Μετρικές κατάτμησης ανά κατηγορία/κλάση κάλυψης γης



Σχήμα 6. Παράδειγμα οδικού δικτύου (χαρακτηρίζεται στον χάρτη αληθείας σαν αστική έκταση) το οποίο το μοντέλο αποτυγχάνει να εντοπίσει



Σχήμα 7. Παράδειγμα έκτασης όπου το μοντέλο “συγχύζεται” και υπερεκτιμά την κάλυψη των εκτάσεων με αραιή ή/και χαμηλή βλάστηση σε σχέση με τους θαμνότοπους.

Επιπλέον, υπολογίστηκαν και οι μέσοι όροι των μετρικών IoU και dice score μεταξύ χάρτη πρόβλεψης και αληθείας οι οποίες είναι ίσες με 70.23% και 59.24% αντίστοιχα. Ανατρέχοντας στα γραφήματα των ίδιων μετρικών κατά την εκπαίδευση μπορεί να εξαχθεί το συμπέρασμα ότι δεδομένων των δυνατοτήτων του μοντέλου η ακρίβεια που επιτυγχάνεται για την πρόβλεψη είναι η καλύτερη δυνατή.

Τέλος, αναφέρεται ότι βάσει του receptive field του μοντέλου το οποίο υπολογίζεται από την αναδρομική σχέση $R_t = R_{t-1} + (k_l - 1) * j_{i-1}$, όπου k_l το μέγεθος του πυρήνα στο επίπεδο l και j_{i-1} το spacing των εικονοστοιχείων εισόδου τα οποία αντιστοιχούν σε μια μονάδα του feature map στο επίπεδο $l - 1$, και το οποίο για το μοντέλο υπό εξέταση βρέθηκε ίσο με 7 υποδηλώνοντας ουσιαστικά πόσα εικονοστοιχεία χρησιμοποιεί ο κάθε νευρώνας του CNN σε ένα συνελκτικό επίπεδο για να εξάγει τα feature maps, έγινε προσπάθεια να εφαρμοστεί μια τεχνική έτσι ώστε η απεικόνιση του χάρτη πρόβλεψης να είναι όσο το δυνατόν γίνεται πιο συνεχής χωρίς να φαίνονται τα όρια των patches. Συγκεκριμένα, από τα άκρα κάθε patch εφαρμόστηκε σε ένα offset πλάτους 16 εικονοστοιχείων η τεχνική zero padding (αντικατάσταση τιμής που παραπέμπει σε κάθε κλάση με το 0) και έπειτα με παρεμβολή εγγύτερου γείτονα έγινε ξανά η ανάθεση των κλάσεων. Βέβαια παρατηρώντας τον χάρτη πρόβλεψης η επίδραση στην “συνέχεια” της απεικόνισης φαίνεται να είναι αρκετά μικρή.