



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Ανάπτυξη, εφαρμογή και αξιολόγηση μη επιβλεπόμενων, βαθιάς μάθησης, μεθοδολογιών αιχμής για τη βελτίωση της χωρικής ανάλυσης δορυφορικών εικόνων Sentinel-2.

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Σάββας Σιφναίος

Επιβλέπων : Π. Τσανάκας

Καθηγητής Ε.Μ.Π.

Αθήνα, Μάρτιος 2022



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ
ΥΠΟΛΟΓΙΣΤΩΝ

**Ανάπτυξη, εφαρμογή και αξιολόγηση μη επιβλεπόμενων,
βαθιάς μάθησης, μεθοδολογιών αιχμής για τη βελτίωση της
χωρικής ανάλυσης δορυφορικών εικόνων Sentinel-2.**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Σάββας Σιφναίος

Επιβλέπων : Π. Τσανάκας

Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 15^η Μαρτίου 2022.

.....
Π. Τσανάκας
Καθηγητής Ε.Μ.Π.

.....
Δ. Σούντρης
Καθηγητής Ε.Μ.Π.

.....
Β. Καντερέ
Επ. Καθηγήτρια Ε.Μ.Π.

Αθήνα, Μάρτιος 2022

.....

Σάββας Σιφναίος

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Σιφναίος Σάββας, 2022.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας Εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της Εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγκεκριμένο συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Η τηλεπισκόπηση, τα τελευταία χρόνια, χαίρει ραγδαίας ανάπτυξης και ενδιαφέροντος, ενώ το φάσμα των εφαρμογών που αξιοποιούν δορυφορικές εικόνες διευρύνεται συνεχώς, καλύπτοντας πλέον τομείς που εκτείνονται από την γεωργία, την οικολογία και την αρχαιολογία έως την πρόληψη φυσικών καταστροφών, την μεσιτική και τον πολεοδομικό σχεδιασμό. Η μεγάλη αυτή ζήτηση για δορυφορικές εικόνες, δημιουργεί συνεχώς αυξανόμενες απαιτήσεις ως προς τα χαρακτηριστικά, την ποιότητα του περιεχομένου και την ανάλυση των εικόνων αυτών. Ωστόσο, εξαιτίας περιορισμών που επιβάλλονται από την σημερινή τεχνολογία αισθητήρων, οι περισσότεροι από τους δορυφόρους που βρίσκονται υπό τροχιά περί τη Γη και χρησιμοποιούνται για εφαρμογές τηλεπισκόπησης αδυνατούν να παράξουν εικόνες με, ταυτόχρονα, υψηλή χωρική ανάλυση και πλούσιο φασματικό περιεχόμενο.

Ένας τρόπος για να ξεπεραστούν οι προαναφερθέντες περιορισμοί, είναι η επεξεργασία των δορυφορικών εικόνων, μετά την λήψη τους, με σκοπό την βελτίωση της χωρικής τους ανάλυσης. Οι αλγόριθμοι βελτίωσης της χωρικής ανάλυσης των δορυφορικών εικόνων, υπάγονται στην ευρύτερη κατηγορία του Super Resolution και σκοπός τους είναι η αύξηση της χωρικής ανάλυσης των πολυφασματικών αυτών εικόνων, διατηρώντας αμετάβλητο το φασματικό τους περιεχόμενο.

Η ανάπτυξη της βαθιάς μάθησης και ειδικότερα των βαθιών συνελικτικών δικτύων, έχει καταστήσει τα παραπάνω ένα ιδιαίτερα χρήσιμο εργαλείο για την επίλυση προβλημάτων σύνθεσης και ανακατασκευής εικόνων, επιτρέποντας την σημαντική βελτίωση της ποιότητας των εικόνων που προκύπτουν ως λύσεις του προβλήματος του Super Resolution. Στην πλειοψηφία τους, οι περισσότεροι αλγόριθμοι βαθιάς μάθησης για την επίλυση του Super Resolution προβλήματος ανήκουν στην οικογένεια της επιβλεπόμενης μάθησης, γεγονός το οποίο δημιουργεί προβλήματα στην περίπτωση που ένα μεγάλο σύνολο εικόνων εκπαίδευσης δεν είναι διαθέσιμο.

Στα πλαίσια της παρούσας διπλωματικής εργασίας εξετάζονται μεθοδολογίες μη επιβλεπόμενης βαθιάς μάθησης για την βελτίωση της χωρικής ανάλυσης εικόνων που έχουν ληφθεί από τον αστερισμό δορυφόρων Sentinel-2 του ευρωπαϊκού προγράμματος Copernicus. Αναλυτικότερα, αξιοποιούνται δύο διαφορετικά βαθιά συνελικτικά δίκτυα, τα οποία, μέσω αμιγώς μη επιβλεπόμενης μάθησης, δεχόμενα ως είσοδο μία δορυφορική εικόνα παράγουν ως έξοδο την βελτιωμένης χωρικής ανάλυσης εκδοχή της. Επιπλέον, χρησιμοποιείται ένα μοντέλο το οποίο παράγει, χωρίς καμία επίβλεψη, αναπαραστάσεις παραμορφώσεων που δίνονται, στην συνέχεια, ως είσοδοι μαζί με τις εικόνες χαμηλής ανάλυσης για την εκπαίδευση ενός νευρωνικού δικτύου ικανού να συνθέτει εικόνες βελτιωμένης χωρικής ανάλυσης.

Η αξιολόγηση των τριών μοντέλων έγινε μέσω ποιοτικής και ποσοτικής ανάλυσης. Για την ποσοτική αξιολόγηση των μοντέλων, χρησιμοποιούνται οι υποδειγματοληπτημένες εκδοχές των αποτελεσμάτων του εκάστοτε μοντέλου, δεδομένου ότι πραγματικές εικόνες υψηλής ανάλυσης δεν είναι διαθέσιμες. Επιπλέον, σημειώνεται πως για την αξιολόγηση αυτή χρησιμοποιήθηκαν εικόνες από πέντε διαφορετικές περιοχές μελέτης, οι οποίες είναι ευθυγραμμισμένες με τις ανάγκες του ερευνητικού έργου EIFFEL (European Union's Horizon 2020 research and innovation programme under Grant Agreement No 101003518), <https://www.eiffel4climate.eu>.

Λέξεις Κλειδιά : Super Resolution, μη επιβλεπόμενη μάθηση, βαθιά μάθηση, συνελικτικά νευρωνικά δίκτυα, Sentinel-2, τηλεπισκόπηση, δορυφορικές εικόνες.

Abstract

Remote sensing is a rapidly growing field which, in recent years, has drawn the interest of the scientific community, while the number of applications that utilize satellite images is constantly increasing, involving areas ranging from agriculture, ecology and archeology to disaster prevention, real estate and urban planning. This high demand for satellite imagery creates ever-increasing expectations regarding the features, the content quality and the resolution of these imagery. However, due to limitations imposed by current sensor technology, most of the satellites orbiting the Earth that are used for remote sensing applications are unable to produce images with high spatial resolution and rich spectral content at the same time.

To overcome the above-mentioned limitations one could process the satellite images, after receiving them, in order to improve their spatial resolution. Algorithms for improving the spatial resolution of satellite images (i.e. increasing the number of pixels they contain) fall into the broader category of Super Resolution and their purpose is to increase the spatial resolution of these multispectral images, without any loss of spectral information.

The fast development of deep learning, and in particular of deep convolutional networks, has rendered neural networks a particularly useful tool for solving inverse problems such as image generation and restoration, allowing significant improvements, in terms of quality, in the resulting images of Super Resolution algorithms. Most of the deep learning Super Resolution algorithms, however, are supervised, which creates problems when a large set of training images is not available.

In this work, we examine the use of unsupervised deep learning methodologies in order to improve the spatial resolution of images obtained from the Sentinel-2 satellite constellation, which is part of the European Copernicus program. More specifically, two different deep convolutional networks are utilized, which receive as input a satellite image and, through purely unsupervised learning, produce an image with improved spatial resolution. In addition, another model is used which learns, without any supervision, degradation representations which are then used, along with the low resolution images, as training inputs for the following Super Resolution neural network.

A qualitative and a quantitative evaluation of the three models was conducted. For the quantitative evaluation, the downsampled versions of the results of each model were used as ground truth, since the original high resolution images are not available. For this evaluation we used images from five different areas were, which are in line with the needs of the EIFFEL research project (European Union's Horizon 2020 research and innovation program under Grant Agreement No. 101003518), <https://www.eiffel4climate.eu>.

Keywords : *Super Resolution, unsupervised learning, deep learning, convolutional neural networks, Sentinel-2, remote sensing, satellite imagery.*

Ευχαριστίες

Θα ήθελα, πρωτίστως, να ευχαριστήσω τον Δρ. Παναγιώτη Τσανάκα, επιβλέποντα καθηγητή της παρούσας διπλωματικής εργασίας, για την ανάθεση του θέματος και την γενικότερη συμβολή του στην ολοκλήρωση της εργασίας αυτής. Παράλληλα, ευχαριστώ το ΕΠΙΣΕΥ, σε συνεργασία με το οποίο πραγματοποιήθηκε η διπλωματική αυτή εργασία, για την παραχώρηση των δεδομένων που χρησιμοποιήθηκαν για την αξιολόγηση των μεθοδολογιών που εξετάζονται στο παρόν έργο και τα οποία είναι ευθυγραμμισμένα με τις ανάγκες του ερευνητικού έργου EIFFEL (European Union's Horizon 2020 research and innovation programme under Grant Agreement No 101003518), <https://www.eiffel4climate.eu>.

Επίσης, ευχαριστώ θερμά τον κ. Δημήτρη Μπληζιώτη για την εξαιρετική συνεργασία μας, από την αρχή έως το τέλος, του οποίου η βοήθεια, οι συμβουλές και η καθοδήγηση διαδραμάτισαν καθοριστικό ρόλο στην τελική διαμόρφωση της παρούσας εργασίας. Ταυτόχρονα, θα ήθελα να ευχαριστήσω τους κ. Γεώργιο Τσιμικλή και κα. Μαρία Δεκαβάλλα για την συμβολή τους σε διαδικαστικά και προγραμματιστικά ζητήματα, καθώς και για τις ιδέες εμπλουτισμού του περιεχομένου της εργασίας.

Τέλος, ευχαριστώ την οικογένειά μου για την υπομονή, την υποστήριξη και την συνεχή εμπιστοσύνη τους σε εμένα, καθώς και τους φίλους μου που αποτέλεσαν, και συνεχίζουν να αποτελούν, πηγή έμπνευσης και έναυσμα προσωπικής βελτίωσης και προόδου.

Πίνακας Περιεχομένων

1. Εισαγωγή.....	20
1.1. Αντικείμενο της εργασίας.....	20
1.2. Διάρθρωση της εργασίας.....	21
2. Βιβλιογραφική Επισκόπηση.....	22
2.1. Δορυφορική Τηλεπισκόπηση.....	22
2.1.1 Εφαρμογές Δορυφορικής Τηλεπισκόπησης.....	23
2.2. Ευρωπαϊκά προγράμματα δορυφορικής τηλεπισκόπησης.....	25
2.2.1 Ευρωπαϊκός Οργανισμός Διαστήματος.....	25
2.2.2 Το πρόγραμμα Copernicus.....	25
2.2.3 Η αποστολή Sentinel 2.....	26
2.2.3.1 Γενικά Στοιχεία.....	26
2.2.3.2 Τεχνικές Προδιαγραφές.....	28
2.2.3.3 Τα προϊόντα.....	30
2.3. Βελτίωση της χωρικής ανάλυσης δορυφορικών εικόνων.....	32
2.3.1 Ορισμός.....	32
2.3.2 Αλγόριθμοι παρεμβολής.....	32
2.3.2.1 Μη προσαρμοστικοί αλγόριθμοι.....	33
2.3.2.1.1 Πλησιέστερος γείτονας.....	33
2.3.2.1.2 Διγραμμική παρεμβολή.....	34
2.3.2.1.3 Δικυβική παρεμβολή.....	34
2.3.2.1.4 Δικυβική B-Spline.....	35
2.3.2.2 Προσαρμοστικοί αλγόριθμοι.....	36
2.3.3 Αλγόριθμοι Pansharpening.....	36
2.3.3.1 Component Substitution.....	37
2.3.3.1.1 Η μέθοδος (G)IHS.....	39
2.3.3.1.2 Η μέθοδος PCA.....	39
2.3.3.2 Multiresolution Analysis.....	40
2.3.4 Σύγχρονες προσεγγίσεις.....	41
2.3.4.1 Multi-Image Super Resolution.....	41
2.3.4.2 Example-Based Super Resolution.....	42

2.3.4.3 Single-Image Super Resolution.....	42
2.4. Μηχανική μάθηση.....	43
2.4.1 Εφαρμογές της μηχανικής μάθησης.....	44
2.4.2 Επιβλεπόμενη Μάθηση.....	45
2.4.3 Μη Επιβλεπόμενη Μάθηση.....	45
2.4.4 Βαθιά Μάθηση.....	46
2.4.4.1 Το Perceptron.....	46
2.4.4.2 Feedforward Neural Networks.....	48
2.4.4.2.1 Η διαδικασία της εκπαίδευσης.....	51
2.4.4.3 Convolutional Neural Networks.....	52
2.4.4.3.1 Η διαδικασία της εκπαίδευσης.....	55
2.4.5 Εφαρμογές της βαθιάς μάθησης.....	56
2.4.5.1 Εφαρμογές στο Super Resolution.....	57
3. Βελτίωση της χωρικής ανάλυσης δορυφορικών εικόνων μέσω μη επιβλεπόμενης μάθησης.....	61
3.1. Τα μοντέλα.....	62
3.1.1 Deep Image Prior.....	62
3.1.1.1 Αρχιτεκτονική και βελτιστοποίηση του δικτύου.....	66
3.1.1.2 Τροποποιήσεις και συνεισφορά.....	67
3.1.2 Zero-Shot Super Resolution.....	68
3.1.2.1 Αρχιτεκτονική και βελτιστοποίηση του δικτύου.....	70
3.1.2.2 Τροποποιήσεις και συνεισφορά.....	71
3.1.3 Degradation Aware Super Resolution.....	72
3.1.3.1 Αρχιτεκτονική και βελτιστοποίηση των δικτύων.....	75
3.1.3.2 Τροποποιήσεις και συνεισφορά.....	78
3.2. Συγκεντρωτική περιγραφή μοντέλων.....	80
4. Αποτελέσματα.....	81
4.1. Το σύνολο ελέγχου.....	81
4.2. Αποτελέσματα του Deep Image Prior.....	82
4.2.1 Εικόνες με χωρική ανάλυση 10m GSD.....	83
4.2.2 Εικόνες με χωρική ανάλυση 20m GSD.....	85
4.2.3 Εικόνες με χωρική ανάλυση 60m GSD.....	88
4.2.4 Σχολιασμός Αποτελεσμάτων.....	90

4.3. Αποτελέσματα Zero-Shot Super Resolution.....	92
4.3.1 Εικόνες με χωρική ανάλυση 10m GSD.....	92
4.3.2 Εικόνες με χωρική ανάλυση 20m GSD.....	94
4.3.3 Εικόνες με χωρική ανάλυση 60m GSD.....	97
4.3.4 Σχολιασμός Αποτελεσμάτων.....	99
4.4. Αποτελέσματα Degradation Aware Super Resolution.....	100
4.4.1 Εικόνες χωρικής ανάλυσης 10m GSD.....	100
4.4.2 Εικόνες χωρικής ανάλυσης 20m GSD.....	103
4.4.3 Εικόνες χωρικής ανάλυσης 60m GSD.....	105
4.4.4 Σχολιασμός Αποτελεσμάτων.....	107
4.5. Ποιοτική αξιολόγηση αποτελεσμάτων μέσω φωτοερμηνείας.....	109
4.6. Ποσοτική αξιολόγηση αποτελεσμάτων.....	115
4.6.1 Αξιολόγηση βάσει μετρικών.....	115
4.6.2 Σύγκριση Ιστογραμμάτων.....	118
4.7. Συνολική αξιολόγηση.....	122
4.8. Τεχνικά χαρακτηριστικά των υπολογιστικών συστημάτων.....	122
5. Επίλογος.....	122
5.1. Μελλοντικές προοπτικές.....	124
6. Βιβλιογραφία.....	126

Κατάλογος Εικόνων

Εικόνα 1: Η κάλυψη της αποστολής Sentinel 2 και οι χρόνοι επανεπίσκεψης των περιοχών αυτών [14]	27
Εικόνα 2: Η γεωμετρική συχνότητα επανεπίσκεψης εξαιτίας της επικάλυψης ζωνών από γειτονικές τροχιές [14].....	28
Εικόνα 3: Σχηματική απεικόνιση της μεγέθυνσης εικόνας [21].....	33
Εικόνα 4: Οι παραγόμενες εικόνες από τις μεθόδους Nearest Neighbor, Bilinear και Bicubic.[23]..	35
Εικόνα 5: Σύγκριση μεταξύ βιολογικού και τεχνητού νευρώνα.....	47
Εικόνα 6: Οι γνωστότερες συναρτήσεις ενεργοποίησης.....	49
Εικόνα 7: Η συνάρτηση LeakyReLU.....	50
Εικόνα 8: Γραφική αναπαράσταση ενός Feedforward νευρωνικού δικτύου.....	50
Εικόνα 9: Εφαρμογή <i>zero padding</i> σε 6x6 εικόνα.....	54
Εικόνα 10: Η διαδικασία της συνέλιξης.....	54
Εικόνα 11: Παράδειγμα εφαρμογής <i>max pooling</i>	55
Εικόνα 12: Η αρχιτεκτονική του μοντέλου MSCARN [27].....	58
Εικόνα 13: Εσωτερική αρχιτεκτονική του MSRBCA block [27].....	58
Εικόνα 14: Εσωτερική αρχιτεκτονική του SE block (attention block) [27].....	59
Εικόνα 15: Η συνολική αρχιτεκτονική του DSen2 [46].....	60
Εικόνα 16: Η αρχιτεκτονική ενός ResBlock[46].....	60
Εικόνα 17: Σχήματα ελαχιστοποίησης της ενέργειας σφάλματος της εξόδου για διαφορετικές συναρτήσεις κανονικοποίησης [47].....	63
Εικόνα 18: Υπολογισμός του με χρήση της μεθόδου <i>gradient descent</i> [47].....	64
Εικόνα 19: Δείγματα <i>priors</i> που δημιουργούνται από διαφορετικές αρχιτεκτονικές νευρωνικών δικτύων [47].....	65
Εικόνα 20: Η αρχιτεκτονική του Deep Image Prior (DIP) δικτύου.....	66
Εικόνα 21: Η διαδικασία εκπαίδευσης του ZSSR [48].....	70
Εικόνα 22: Οπτικοποίηση της διαδικασίας εκμάθησης αναπαραστάσεων παραμόρφωσης [53]....	75
Εικόνα 23: Αρχιτεκτονική του κωδικοποιητή [53].....	76
Εικόνα 24: Η αρχιτεκτονική DASR [53].....	77
Εικόνα 25: Επεξήγηση των Blocks.....	77
Εικόνα 26: Η αρχιτεκτονική ενός DA Block [53].....	77
Εικόνα 27: Οι εικόνες του σύνολου ελέγχου.....	81
Εικόνα 28: Οι ευρύτερες περιοχές των υπό μελέτη εικόνων στο χάρτη.....	82

Εικόνα 48: Μεγέθυνση των 10m GSD καναλιών της 5ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.....	94
Εικόνα 49: Μεγέθυνση των 20m GSD καναλιών της 1ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.....	94
Εικόνα 50: Μεγέθυνση των 20m GSD καναλιών της 2ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.....	95
Εικόνα 51: Μεγέθυνση των 20m GSD καναλιών της 3ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.....	95
Εικόνα 52: Μεγέθυνση των 20m GSD καναλιών της 4ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.....	96
Εικόνα 53: Μεγέθυνση των 20m GSD καναλιών της 5ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.....	96
Εικόνα 54: Μεγέθυνση των 60m GSD καναλιών της 1ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.....	97
Εικόνα 55: Μεγέθυνση των 60m GSD καναλιών της 2ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.....	97
Εικόνα 56: Μεγέθυνση των 60m GSD καναλιών της 3ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.....	98
Εικόνα 57: Μεγέθυνση των 60m GSD καναλιών της 4ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.....	98
Εικόνα 58: Μεγέθυνση των 60m GSD καναλιών της 5ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.....	99
Εικόνα 59: Μεγέθυνση των 10m GSD καναλιών της 1ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.....	100
Εικόνα 60: Μεγέθυνση των 10m GSD καναλιών της 2ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.....	101
Εικόνα 61: Μεγέθυνση των 10m GSD καναλιών της 3ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.....	101
Εικόνα 62: Μεγέθυνση των 10m GSD καναλιών της 4ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.....	102
Εικόνα 63: Μεγέθυνση των 10m GSD καναλιών της 5ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.....	102
Εικόνα 64: Μεγέθυνση των 20m GSD καναλιών της 1ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.....	103
Εικόνα 65: Μεγέθυνση των 20m GSD καναλιών της 2ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.....	103
Εικόνα 66: Μεγέθυνση των 20m GSD καναλιών της 3ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.....	104

Εικόνα 67: Μεγέθυνση των 20m GSD καναλιών της 4ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.....	104
Εικόνα 68: Μεγέθυνση των 20m GSD καναλιών της 5ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.....	105
Εικόνα 69: Μεγέθυνση των 60m GSD καναλιών της 1ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.....	105
Εικόνα 70: Μεγέθυνση των 60m GSD καναλιών της 2ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.....	106
Εικόνα 71: Μεγέθυνση των 60m GSD καναλιών της 3ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.....	106
Εικόνα 72: Μεγέθυνση των 60m GSD καναλιών της 4ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.....	107
Εικόνα 73: Μεγέθυνση των 60m GSD καναλιών της 5ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.....	107
Εικόνα 74: DIP (10m).....	109
Εικόνα 75: ZSSR (10m).....	110
Εικόνα 76: DASR (10m).....	110
Εικόνα 77: DIP (20m).....	111
Εικόνα 78: ZSSR (20m).....	111
Εικόνα 79: DASR (20m).....	111
Εικόνα 80: DIP (60m).....	112
Εικόνα 81: ZSSR (60m).....	113
Εικόνα 82: DASR (60m).....	113
Εικόνα 83: Ιστόγραμμα 1ης εικόνας από το DIP (10m).....	118
Εικόνα 84: Ιστόγραμμα 2ης εικόνας από το DIP (10m).....	118
Εικόνα 85: Ιστόγραμμα 1ης εικόνας από το ZSSR (10m).....	118
Εικόνα 86: Ιστόγραμμα 2ης εικόνας από το ZSSR (10m).....	118
Εικόνα 87: Ιστόγραμμα 2ης εικόνας από το DASR (10m).....	119
Εικόνα 88: Ιστόγραμμα 1ης εικόνας από το DASR (10m).....	119
Εικόνα 89: Ιστόγραμμα 1ης εικόνας από το DIP (20m).....	119
Εικόνα 90: Ιστόγραμμα 2ης εικόνας από το DIP (20m).....	119
Εικόνα 91: Ιστόγραμμα 1ης εικόνας από το ZSSR (20m).....	119
Εικόνα 92: Ιστόγραμμα 2ης εικόνας από το ZSSR (20m).....	119
Εικόνα 93: Ιστόγραμμα 1ης εικόνας από το DASR (20m).....	119

Εικόνα 94: Ιστόγραμμα 2ης εικόνας από το DASR (20m).....	119
Εικόνα 95: Ιστόγραμμα 1ης εικόνας από το DIP (60m).....	120
Εικόνα 96: Ιστόγραμμα 2ης εικόνας από το DIP (60m).....	120
Εικόνα 97: Ιστόγραμμα 1ης εικόνας από το ZSSR (60m).....	120
Εικόνα 98: Ιστόγραμμα 2ης εικόνας από το ZSSR (60m).....	120
Εικόνα 99: Ιστόγραμμα 1ης εικόνας από το DASR (60m).....	120
Εικόνα 100: Ιστόγραμμα 2ης εικόνας από το DASR (60m).....	120

Κατάλογος Πινάκων

Πίνακας 1: Οι φασματικές ζώνες των αισθητήρων των Sentinel-2 δορυφόρων.....	29
Πίνακας 2: Επεξήγηση των συμβόλων.....	38
Πίνακας 3: Συνοπτική παρουσίαση μοντέλων και της συνεισφοράς της εργασίας.....	80
Πίνακας 4: Μέσοι χρόνοι εκτέλεσης για το DIP μοντέλο	91
Πίνακας 5: Μέσοι χρόνοι εκτέλεσης για το ZSSR μοντέλο.....	100
Πίνακας 6: Χρόνοι εκπαίδευσης του μοντέλου DASR, ανά συντελεστή μεγέθυνσης.....	108
Πίνακας 7: Μέσοι χρόνοι σύνθεσης εικόνων από το DASR.....	109
Πίνακας 8: Βαθμολόγηση των μοντέλων, ως προς τις κύριες παραμέτρους αξιολόγησης.....	115
Πίνακας 9: Μέσοι όροι μετρικών ανά μοντέλο και ανά χωρική ανάλυση εικόνων.....	117
Πίνακας 10: Συνολική αξιολόγηση των μοντέλων.....	121

1. Εισαγωγή

Στην σημερινή εποχή υπάρχει πληθώρα διαστημικών οργανισμών που διαθέτουν τεχνητούς δορυφόρους σε τροχιά γύρω από την Γη ή έχουν την ικανότητα εκτόξευσης και διαχείρισης τέτοιων συστημάτων. Αναφέρονται, ονομαστικά, οι European Space Agency (ESA) , National Aeronautics and Space Administration (NASA) , China National Space Administration (CNSA) ως παραδείγματα οργανισμών των οποίων το πεδίο δράσης περιέχει την διαχείριση δορυφορικών συστημάτων και κατ' επέκταση τον κλάδο της δορυφορικής τηλεπισκόπησης (Satellite Remote Sensing). Τα δορυφορικά συστήματα που αξιοποιούνται για εφαρμογές τηλεπισκόπησης είναι εξοπλισμένα με συστοιχίες αισθητήρων ικανές να λαμβάνουν εικόνες της Γης πέρα από το ορατό φάσμα της ηλεκτρομαγνητικής ακτινοβολίας. Συγκεκριμένα, οι δορυφορικές εικόνες αποτελούνται από πολλαπλά χρωματικά κανάλια, το πλήθος των οποίων ξεπερνά τα τρία που είναι το σύνηθες στην περίπτωση των συμβατικών εικόνων, καθένα από τα οποία περιέχει πληροφορίες σχετικά με μήκη κύματος της ηλεκτρομαγνητικής που δύναται να εκτείνονται από το υπεριώδες έως και το υπέρυθρο τμήμα του φάσματος. Τα πιο συνηθισμένα, ωστόσο, τμήματα της ηλεκτρομαγνητικής ακτινοβολίας που απεικονίζονται στις δορυφορικές εικόνες είναι το Short Wave Infrared (SWIR), το Visible and Near-Infrared (VNIR), καθώς και το ορατό τμήμα του φάσματος.

Η πολυφασματική φύση των δορυφορικών εικόνων, γίνεται εύκολα κατανοητό, πως είναι ιδιαίτερα επιθυμητή και χρήσιμη σε πληθώρα εφαρμογών το φάσμα των οποίων εκτείνεται από την γεωργία έως τον πολεοδομικό σχεδιασμό και τις στρατιωτικές δυνάμεις. Η ποιότητα των αποτελεσμάτων των παραπάνω εφαρμογών εξαρτάται, πέρα από το φασματικό περιεχόμενο, και από την χωρική ανάλυση των πολυφασματικών δορυφορικών εικόνων, η οποία πολλές φορές αποδεικνύεται ανεπαρκής στην περίπτωση σύγχρονων και πολύπλοκων εφαρμογών. Η προφανής λύση που θα μπορούσε κανείς να προτείνει είναι η βελτίωση της ποιότητας των ίδιων των αισθητήρων, προκειμένου αυτοί να είναι σε θέση να δημιουργούν εικόνες με ταυτόχρονα πλούσιο φασματικό περιεχόμενο και υψηλή χωρική ανάλυση. Ωστόσο, η πρόταση αυτή δεν είναι ιδιαίτερα εύκολο να υλοποιηθεί, καθώς η αντίστοιχη τεχνολογία ακόμα δεν είναι σε θέση να υποστηρίξει την κατασκευή τέτοιων αισθητήρων χωρίς την εκθετική αύξηση του κόστους των δορυφορικών συστημάτων στα οποία θα εγκατασταθούν.

Μία πιο εφικτή πρόταση για την αντιμετώπιση του περιορισμού της χαμηλής ευκρίνειας των δορυφορικών εικόνων είναι η χρήση λογισμικού για την εκ των υστέρων βελτίωση της χωρικής ανάλυσης των εικόνων αυτών. Αναλυτικότερα, στην βιβλιογραφία ανά τα χρόνια έχει προταθεί πληθώρα διαφορετικών προσεγγίσεων για την επίλυση του συγκεκριμένου προβλήματος. Οι αλγόριθμοι που εκτελούν την διαδικασία της βελτίωσης της χωρικής ανάλυσης μίας εικόνας καλούνται αλγόριθμοι Super Resolution και εκτείνονται από πολύ απλούς αλγορίθμους παρεμβολής, σε ειδικά σχεδιασμένους αλγορίθμους για δορυφορικές εικόνες που καλούνται αλγόριθμοι Pansharpening έως και σύνθετα μοντέλα μηχανικής και βαθιάς μάθησης.

1.1. Αντικείμενο της εργασίας

Η μηχανική, και κυρίως η βαθιά, μάθηση μέσα στις τελευταίες δεκαετίες έχει κερδίσει το ενδιαφέρον μιας πολύ μεγάλης μερίδας της επιστημονικής κοινότητας που ασχολείται με ζητήματα τεχνητής νοημοσύνης, με αποτέλεσμα την τεράστια και ταχύτατη ανάπτυξή της και την εισχώρησή της σε, πρακτικά, κάθε τομέα της επιστήμης των υπολογιστών. Οι αλγόριθμοι βαθιάς μάθησης έχουν καταφέρει να βελτιώσουν τα αποτελέσματα πληθώρας εφαρμογών, με

τις εφαρμογές στο τομέα της όρασης υπολογιστών να εκμεταλλεύονται σε πολύ μεγάλο βαθμό την ικανότητα των νευρωνικών δικτύων να μοντελοποιούν τις έντονες μη γραμμικές εξαρτήσεις που υπάρχουν μεταξύ γειτονικών εικονοστοιχείων μίας εικόνας, για να βελτιώσουν την ποιότητα των αποτελεσμάτων που λαμβάνονταν από τις αντίστοιχες κλασικές μεθόδους. Μία από τις εφαρμογές της όρασης υπολογιστών, στην οποία η βαθιά μάθηση αποδεικνύει την υπεροχή της, έναντι των κλασικών μεθοδολογιών, είναι αυτή του Super Resolution, σκοπός της οποίας είναι η αύξηση της χωρικής ανάλυσης μίας εικόνας, η αύξηση δηλαδή του πλήθους των εικονοστοιχείων που περιέχονται σε αυτή, με την εισαγωγή όσο το δυνατόν λιγότερων παραμορφώσεων και αλλοιώσεων. Στο μεγαλύτερο μέρος της βιβλιογραφίας, οι προτεινόμενες μέθοδοι βαθιάς μάθησης για την βελτίωση της χωρικής ανάλυσης μίας εικόνας ακολουθούν την προσέγγιση της επιβλεπόμενης μάθησης, κατά την οποία ένα εκτενές σύνολο με ζεύγη εικόνων υψηλής-χαμηλής ανάλυσης απαιτείται για την εκπαίδευση του εκάστοτε μοντέλου. Το κυριότερο πρόβλημα με αυτή την προσέγγιση στην περίπτωση των δορυφορικών εικόνων, που εξετάζονται στην παρούσα εργασία, είναι η έλλειψη τέτοιων συνόλων εκπαίδευσης, καθώς η απόκτηση εικόνων υψηλής χωρικής ανάλυσης, όπως προαναφέρθηκε, αποτελεί σχετικά δύσκολη διαδικασία.

Στα πλαίσια της παρούσας διπλωματικής εργασίας, εξετάζονται μεθοδολογίες βαθιάς, μη επιβλεπόμενης μάθησης για την εκτέλεση του Super Resolution σε εικόνες που έχουν ληφθεί από τον αστερισμό δορυφόρων Sentinel-2, ο οποίος αποτελεί μία από τις αποστολές του ευρωπαϊκού προγράμματος Copernicus. Το κύριο πλεονέκτημα των μη επιβλεπόμενων αλγορίθμων είναι το γεγονός ότι είναι σε θέση να συνθέσουν εικόνες βελτιωμένης χωρικής ανάλυσης, χωρίς την ανάγκη ύπαρξης ενός συνόλου δεδομένων της μορφής που περιγράφηκε στην παραπάνω παράγραφο. Επομένως, υπό την απουσία των πραγματικών εικόνων υψηλής χωρικής ανάλυσης, γίνεται κατανοητό πως τα μοντέλα μη επιβλεπόμενης μάθησης που θα χρησιμοποιηθούν στην εργασία αυτή έχουν στην διάθεσή τους μόνο την εικόνα χαμηλής ευκρίνειας, από την οποία πρέπει να εξαγάγουν τις απαραίτητες πληροφορίες και στατιστικές εξαρτήσεις για την σύνθεση του τελικού αποτελέσματος. Ο περιορισμός αυτός δεν πρέπει υποχρεωτικά να αντιμετωπιστεί ως πρόβλημα ή τροχοπέδη, καθώς κατά αυτό τον τρόπο το εκάστοτε μοντέλο “εκπαιδεύεται” αποκλειστικά και μόνο πάνω στην υπό εξέταση εικόνα, μαθαίνοντας εξαρτήσεις και εξαγοντας στατιστικά χαρακτηριστικά εξειδικευμένα για την εκάστοτε εικόνα, γεγονός το οποίο δεν συμβαίνει στο μοντέλα επιβλεπόμενης μάθησης και συχνά δημιουργεί προβλήματα κατά την σύνθεση της εικόνας βελτιωμένης ανάλυσης.

1.2. Διάρθρωση της εργασίας

Η παρούσα διπλωματική εργασία οργανώνεται σε έξι κεφάλαια, συμπεριλαμβανομένου και του παρόντος. Στο **δεύτερο κεφάλαιο**, παρατίθενται όλες οι πληροφορίες που χρειάζεται ο αναγνώστης της εργασίας αυτής για να είναι σε θέση να κατανοήσει πλήρως τις έννοιες και τις μεθοδολογίες που παρουσιάζονται στην συνέχεια. Αναλυτικότερα, στο συγκεκριμένο κεφάλαιο γίνεται μία εισαγωγή στην έννοια της τηλεπισκόπησης και των βασικότερων εφαρμογών αυτής, στην συνέχεια παρουσιάζεται ο Ευρωπαϊκός Διαστημικός Οργανισμός και το πρόγραμμα Copernicus. Η παρουσίαση αυτή ακολουθείται από μία αναλυτικότερη παρουσίαση της αποστολής Sentinel-2 της οποίας τα προϊόντα διαδραματίζουν κεντρικό ρόλο στην παρούσα εργασία. Επιπλέον, πραγματοποιείται μία ιστορική αναδρομή στις μεθοδολογίες που είχαν χρησιμοποιηθεί παλαιότερα για την βελτίωση της χωρικής ανάλυσης τόσο των τριχρωματικών εικόνων, γενικά, όσο και των δορυφορικών εικόνων ειδικότερα. Στο τέλος του συγκεκριμένου κεφαλαίου, παρατίθεται μία εκτενής ανάλυση της μηχανικής

μάθησης, η οποία καταλήγει στην επεξήγηση των μεθοδολογιών της βαθιάς μάθησης και την αναφορά ορισμένων ενδεικτικών εφαρμογών αυτής.

Το **τρίτο κεφάλαιο** αποτελεί το πυρήνα της εργασίας, καθώς εκεί παρουσιάζονται αναλυτικότερα οι μεθοδολογίες που ακολουθούνται από τα τρία μοντέλα που χρησιμοποιήθηκαν, καθώς και οι αρχιτεκτονικές αυτών. Επιπλέον, αναφέρονται, για κάθε ένα από τα μοντέλα, οι τροποποιήσεις που χρειάστηκαν να πραγματοποιηθούν, προκειμένου τα μοντέλα να λειτουργούν κατά τον αναμενόμενο τρόπο.

Στο **τέταρτο κεφάλαιο**, παρουσιάζονται οι εικόνες Sentinel-2 που χρησιμοποιήθηκαν για την αξιολόγηση των μοντέλων και παρατίθενται οι εικόνες βελτιωμένης χωρικής ανάλυσης που δημιουργήθηκαν από καθένα από τα τρία μοντέλα βαθιάς μάθησης. Η παράθεση των αποτελεσμάτων του εκάστοτε μοντέλου συνοδεύεται από σχολιασμό τους, προκειμένου στη συνέχεια να καταστεί εφικτή η ολική σύγκριση των μεθοδολογιών και να αναδειχθεί το βέλτιστο, μεταξύ των τριών, μοντέλο. Στο τέλος του συγκεκριμένου κεφαλαίου αναφέρονται τα τεχνικά χαρακτηριστικά των υπολογιστικών συστημάτων στα οποία πραγματοποιήθηκε η εκτέλεση των αλγορίθμων των μοντέλων.

Στο **πέμπτο κεφάλαιο**, συνοψίζονται τα συμπεράσματα που προκύπτουν σχετικά με τις μεθοδολογίες, ενώ παράλληλα παρουσιάζονται οι μελλοντικές προοπτικές επέκτασης και βελτίωσης της παρούσας εργασίας.

Το **έκτο κεφάλαιο**, περιέχει το σύνολο της βιβλιογραφίας που χρησιμοποιήθηκε για την εκπόνηση της παρούσας διπλωματικής εργασίας.

2. Βιβλιογραφική Επισκόπηση

Στο συγκεκριμένο κεφάλαιο γίνεται εισαγωγή ορισμένων από τις βασικές έννοιες οι οποίες θα μελετηθούν στην παρούσα εργασία. Αρχικά, δίνεται ο ορισμός της δορυφορικής τηλεπισκόπησης, οι λόγοι οι οποίοι οδήγησαν στην ανάπτυξη του συγκεκριμένου τομέα, καθώς και ορισμένες από τις κυριότερες εφαρμογές του πεδίου αυτού. Στην συνέχεια, γίνεται μια παρουσίαση του Ευρωπαϊκού Διαστημικού Οργανισμού, η οποία ακολουθείται από μία, σχετικά, εκτενή περιγραφή του προγράμματος Copernicus, εστιάζοντας, κυρίως, στην αποστολή Sentinel-2, καθώς τα προϊόντα της αποστολής αυτής κατέχουν κεντρική θέση στην παρούσα εργασία. Επιπρόσθετα, παρουσιάζονται ορισμένες από τις μεθόδους που έχουν χρησιμοποιηθεί παλαιότερα για την βελτίωση της χωρικής ανάλυσης δορυφορικών εικόνων. Στο τέλος του συγκεκριμένου κεφαλαίου, γίνεται αναφορά στην μηχανική μάθηση, αναλύοντας το θεωρητικό υπόβαθρο και τις εφαρμογές του πεδίου αυτού της τεχνητής νοημοσύνης.

2.1. Δορυφορική Τηλεπισκόπηση

Ως (δορυφορική) τηλεπισκόπηση ορίζεται η διαδικασία ανάκτησης πληροφοριών και δεδομένων που αφορούν τις χερσαίες και υδάτινες εκτάσεις της Γης, αξιοποιώντας την ηλεκτρομαγνητική ακτινοβολία ενός, ή και περισσοτέρων, τμημάτων του ηλεκτρομαγνητικού φάσματος, η οποία ανακλάται στην Γη ή εκπέμπεται από αυτή. Για την ανίχνευση της ακτινοβολίας αυτής, χρησιμοποιούνται συστοιχίες αισθητήρων οι οποίοι είναι ενσωματωμένοι στους υπό τροχιά, περί τη Γη, δορυφόρους. [1]

Η μεγάλη έκταση των τμημάτων της επιφάνειας της Γης, που δύναται να μελετηθούν, σε συνδυασμό με την πολυπλοκότητα των φαινομένων που λαμβάνουν χώρα εκεί, καθιστούν τις

συμβατικές (επίγειες) μεθόδους μελέτης, ανάλυσης και μοντελοποίησής τους ανεπαρκής. Συγκεκριμένα, αναφέρεται το παράδειγμα των φυσικών πόρων που εντοπίζονται σε υδροβιότοπους, η μελέτη των οποίων απαιτεί την αποθήκευση και παρακολούθηση δεδομένων που σχετίζονται τόσο με τους υπό εξέταση υδροβιότοπους όσο και τους γειτονικούς τους, αυξάνοντας σημαντικά το βαθμό πολυπλοκότητας μίας έρευνας που χρησιμοποιεί δεδομένα που προέκυψαν από επίγειες μετρήσεις και παρατηρήσεις [2]. Η δορυφορική τηλεπισκόπηση, αντιθέτως, προσφέρει προηγμένα μέσα και τεχνικές για την μελέτη των φυσικών πόρων, την παρακολούθηση τοπικών περιβαλλοντικών αλλαγών, καθώς και αλλαγών που σημειώνονται σε επίπεδο περιοχών ή ακόμη και σε παγκόσμια κλίμακα. Παράλληλα, η μορφή αυτή τηλεπισκόπησης φέρει τα πλεονεκτήματα πως παράγει με ικανοποιητική ταχύτητα και συχνότητα μακροσκοπικά δεδομένα, που αποτελούν μια ακριβή, περιεκτική και δυναμική αναπαράσταση των περιοχών ενδιαφέροντος. [3]

2.1.1 Εφαρμογές Δορυφορικής Τηλεπισκόπησης

Οι δορυφορικές εικόνες περιέχουν, συνήθως, σύνθετα και διαφορετικά πρότυπα, ενώ ταυτόχρονα είναι πλούσιες σε χωροχρονικές και φασματικές πληροφορίες σχετικά με το τοπίο/σκηνή που απεικονίζουν. Οι εικόνες αυτού του είδους έχουν αποδειχθεί ιδιαίτερα χρήσιμες σε εφαρμογές που αφορούν την παρακολούθηση της επιφάνειας, καθώς και της ατμόσφαιρας της Γης σε τοπικό, περιφερειακό και παγκόσμιο, ακόμη, επίπεδο [3], [4]. Στην συνέχεια παρατίθενται, επιγραμματικά, ορισμένες από τις πιο γνωστές περιπτώσεις αξιοποίησης δορυφορικών εικόνων σε εφαρμογές που συνδυάζουν τεχνικές και μεθόδους της όρασης υπολογιστών [5].

- **Γεωργία** : Στην συγκεκριμένη κατηγορία εντοπίζονται εφαρμογές όπως ο προσδιορισμός των επιπέδων υγρασίας του εδάφους, μέσω είτε ενεργών είτε παθητικών αισθητήρων. Οι ενεργοί αισθητήρες, για την λήψη μίας μέτρησης, εκπέμπουν ένα ηλεκτρομαγνητικό κύμα σε κάποια επιφάνεια της γης και μετράνε την ποσότητα αυτής που ανακλάται, δημιουργώντας έτσι εικόνες υψηλής χωρικής ευκρίνειας. Αντίθετα, οι παθητικοί ανιχνεύουν την ανακλώμενη “φυσική” ακτινοβολία από μία περιοχή με μεγάλη ακρίβεια, υστερώντας όμως σε χωρική ανάλυση. Παράλληλα, στην κατηγορία αυτή εντοπίζονται εφαρμογές όπως η ταξινόμηση του είδους του εδάφους και ο προσδιορισμός της υγείας των καλλιεργειών μέσω υπέρυθρης ακτινοβολίας.
- **Αρχαιολογία** : Ένα από τα βασικότερα προβλήματα στον κλάδο της αρχαιολογίας είναι ο εντοπισμός των περιοχών με την μεγαλύτερη πιθανότητα ύπαρξης απολιθωμάτων και, γενικότερα, αντικειμένων αρχαιολογικής αξίας. Συνεπώς, μέσω της χρήσης δορυφορικών εικόνων, καθίσταται εφικτός ο προσδιορισμός παραμέτρων, όπως παραδείγματος χάριν ο δείκτης βλάστησης και η κλίση του εδάφους, οι οποίοι είναι ενδεικτική για την ύπαρξη ή μη αντικειμένων αρχαιολογικής αξίας. Παράλληλα, στην αρχαιολογία, συχνά χρησιμοποιούνται υπέρυθρες και στερεοσκοπικές δορυφορικές εικόνες για την απεικόνιση αρχαιολογικών ευρημάτων που βρίσκονται κάτω από το έδαφος της γης, χωρίς να απαιτείται εκσκαφή.
- **Δασοπονία** : Η δορυφορική τηλεπισκόπηση έχει συμβάλει σε μεγάλο βαθμό στην ανάπτυξη του κλάδου της δασοπονίας. Αναλυτικότερα, μέσω των δορυφορικών εικόνων καθίσταται εφικτή η ανάλυση της μεταβολής της έκτασης των δασικών εκτάσεων. Επιπλέον, χρησιμοποιώντας δορυφορικές απεικόνισης των δασών, επιτυγχάνεται η έγκαιρη ανίχνευση ασθενειών σε δέντρα, κυρίως, μέσω της παρατήρησης των αλλαγών του χρώματός τους. Τέλος, μία ακόμα σημαντική εφαρμογή της τηλεπισκόπησης στον

τομέα αυτό είναι η ανίχνευση δασικών πυρκαγιών και ο ακριβής προσδιορισμός των μετώπων τους, μέσω πάλι υπέρυθρων και θερμικών απεικονίσεων των δασών.

- **Υδρολογία** : Στην κατηγορία αυτή εντοπίζονται εφαρμογές όπως ο εντοπισμός υδάτων στο υπέδαφος, μέσω υπέρυθρης ακτινοβολίας, η ανίχνευση μόλυνσης, μέσω της μέτρησης της ανακλώμενης από βακτήρια ακτινοβολίας, καθώς και η μελέτη υδροβιότοπων.
- **Περιβάλλον και οικολογία** : Με την χρήση δορυφορικών εικόνων, καθίσταται εφικτή η παρακολούθηση βιότοπων και, γενικότερα, εκτάσεων στην επιφάνεια της γης με σκοπό την μελέτη της βιοποικιλότητας εκεί, καθώς και τον δυναμικό τρόπο που αυτή μεταβάλλεται χρονικά. Επιπρόσθετα, η τηλεπισκόπηση επιτρέπει στους μελετητές να παρακολουθούν την άνοδο των επιπέδων της θάλασσας, καθώς και άλλων φαινομένων που οφείλονται στην κλιματική αλλαγή. Επίσης, προβλήματα που αφορούν την οικολογία, όπως η παρακολούθηση των πληθυσμών απειλούμενων προς εξαφάνιση ειδών μπορούν να επιλυθούν με την βοήθεια της τηλεπισκόπησης. Τέλος, εφαρμογές όπως η παρακολούθηση των τροχιών μεταναστευτικών πτηνών καθώς και η αξιολόγηση της καταλληλότητας περιοχών για την ανάπτυξη συγκεκριμένων ειδών επωφελούνται σε μεγάλο βαθμό από την ανάπτυξη της τηλεπισκόπησης.
- **Πρόληψη φυσικών καταστροφών** : Στην κατηγορία αυτή εντοπίζονται εφαρμογές όπως η παρακολούθηση της ηφαιστειακής δραστηριότητας, μέσω θερμικών αισθητήρων, η εκτίμηση της πιθανότητας κατολισθήσεων, μέσω στερεογραφικών απεικονίσεων για τον προσδιορισμό της κλίσης του εδάφους, η πρόβλεψη πλημμυρών και πυρκαγιών καθώς και η εκτίμηση ζημιών σε κτίρια μετά από σεισμούς.
- **Ωκεανογραφία** : Η συμβολή της τηλεπισκόπησης στην ωκεανογραφία εντοπίζεται σε εφαρμογές όπως η ανίχνευση πετρελαϊκών μολύνσεων, καθώς και η παρακολούθηση της εξέλιξης πετρελαϊκών μολύνσεων. Μία ακόμη εφαρμογή της τηλεπισκόπησης είναι η παρακολούθηση της ροής των κυμάτων, μέσω μετρήσεων που αφορούν τα υψομετρικά επίπεδα θαλάσσιων εκτάσεων, την ταχύτητα των επιφανειακών ανέμων και την θερμοκρασία στην επιφάνεια της θάλασσας. Επιπρόσθετα, η μελέτη της υποθαλάσσιας βλάστησης και των υφάλων επωφελείται από την τηλεπισκόπηση, η οποία επιτρέπει την απομακρυσμένη παρακολούθηση τέτοιων εκτάσεων (ακόμη και σε μεγάλα βάθη). Τέλος, αξίζει να σημειωθεί πως συχνά χρησιμοποιούνται υπέρυθρες απεικονίσεις των ωκεανών με σκοπό την χαρτογράφηση του πυθμένα τους και τον εντοπισμό αντικειμένων εκεί.
- **Πολοδομικό σχεδιασμό** : Οι δορυφορικές εικόνες συχνά χρησιμοποιούνται για την κατασκευή τρισδιάστατων μοντέλων πόλεων, καθώς επίσης και για τον προσδιορισμό της πυκνότητας πληθυσμού σε επιμέρους τμήματα πόλεων.

Στο σημείο αυτό, αξίζει να σημειωθεί πως οι εφαρμογές που αξιοποιούν δορυφορικές εικόνες και, γενικότερα, δορυφορικά δεδομένα δεν περιορίζονται στις παραπάνω αναφερθείσες. Τα δεδομένα που προκύπτουν μέσω της δορυφορικής τηλεπισκόπησης βρίσκονται στο επίκεντρο πολλών, ακόμα, διαφορετικών εφαρμογών το εύρος των οποίων εκτείνεται από τον τομέα της μεσιτικής, της ασφάλισης και της παρακολούθησης (surveillance) έως και εφαρμογές στις στρατιωτικές δυνάμεις και σε κρατικούς μηχανισμούς. [4]

2.2. Ευρωπαϊκά προγράμματα δορυφορικής τηλεπισκόπησης

2.2.1 Ευρωπαϊκός Οργανισμός Διαστήματος

Ο Ευρωπαϊκός Οργανισμός Διαστήματος (ΕΟΔ) είναι οργανισμός με έδρα το Παρίσι. Δημιουργήθηκε στις 30 Μαΐου του 1975 με σκοπό τον καλύτερο συντονισμό των ευρωπαϊκών διαστημικών δραστηριοτήτων, ως αποτέλεσμα της κατανόησης του αδυνάτου της διατήρησης ενός ξεχωριστού εθνικού διαστημικού προγράμματος από κάθε χώρα μέλος, προκειμένου να ανταγωνιστεί την τεχνολογική υπεροχή της Ένωσης Σοβιετικών Σοσιαλιστικών Δημοκρατιών (ΕΣΣΔ) και των Ηνωμένων Πολιτειών της Αμερικής (ΗΠΑ) στον συγκεκριμένο τομέα. Αυτή την στιγμή, ο ΕΟΔ απασχολεί περίπου 2.200 επιστήμονες και μηχανικούς (βάσει των στατιστικών του 2018) και συμπεριλαμβάνει 22 κράτη μέλη.

Το πρόγραμμα διαστημικών πτήσεων του Ευρωπαϊκού Οργανισμού Διαστήματος περιλαμβάνει ανθρώπινες διαστημικές πτήσεις (κυρίως μέσω της συμμετοχής στο πρόγραμμα του Διεθνούς Διαστημικού Σταθμού), την έναρξη και λειτουργία μη επανδρωμένων αποστολών εξερεύνησης σε άλλους πλανήτες και τη Σελήνη, την παρατήρηση της Γης (Earth Observation), τον σχεδιασμό οχημάτων και συστημάτων εκτόξευσης. Επιπλέον διατηρεί ένα μεγάλο διαστημικό αεροδρόμιο, το Διαστημικό Κέντρο της Γουιάνας στο Κουρού της Γαλλικής Γουιάνας. Το κύριο ευρωπαϊκό όχημα εκτόξευσης, Ariane 5, λειτουργεί μέσω της Arianespace με τον ΕΟΔ να μοιράζεται το κόστος της εκτόξευσης και της περαιτέρω ανάπτυξης αυτού του οχήματος εκτόξευσης. Ο ΕΟΔ συνεργάζεται επίσης με τη NASA για την κατασκευή της μονάδας υπηρεσίας Orion Spacecraft [6]–[8].

Στα πλαίσια των δράσεων που αφορούν την παρατήρηση της Γης (Earth Observation), ο Ευρωπαϊκός Οργανισμός Διαστήματος, διαθέτει ένα μεγάλο αριθμό ενεργών προγραμμάτων. Σκοπός των προγραμμάτων αυτών είναι να διατηρήσουν τις τεχνικές δορυφορικής τηλεπισκόπησης στο υψηλότερο, κατά το δυνατόν, επίπεδο τεχνολογικής προόδου, να διευρύνουν τους ορίζοντες της ανθρώπινης κατανόησης για τον πλανήτη μας, καθώς και να παράγουν αποτελέσματα και προϊόντα αιχμής. Επιπλέον, τα προγράμματα αυτά αποσκοπούν στη συγκέντρωση γνώσης και αξιοποίησης αυτής προς όφελος ολόκληρης της επιστημονικής κοινότητας [9].

Ένα από τα προγράμματα του ΕΟΔ είναι το “Living Planet” το οποίο περιλαμβάνει ένα επιστημονικό και ερευνητικό τμήμα που περιέχει την αποστολή “Earth Explorer”, καθώς και ένα τμήμα παρατήρησης της Γης, με την ονομασία “Earth Watch”. Το συγκεκριμένο τμήμα έχει σχεδιαστεί ώστε διευκολύνει την μετάδοση δορυφορικών εικόνων της Γης προς τους χρήστες των λειτουργικών υπηρεσιών. Τέλος, ο Ευρωπαϊκός Οργανισμός Διαστήματος αναπτύσσει την σειρά δορυφόρων Sentinel, σε συνεργασία με την Ευρωπαϊκή Επιτροπή, στα πλαίσια του προγράμματος Copernicus [9].

2.2.2 Το πρόγραμμα Copernicus

Το Copernicus αποτελεί το πιο αναπτυγμένο από τα προγράμματα του ΕΟΔ, στο τομέα του Earth Observation, και αποσκοπεί στην ανάπτυξη μιας μεθόδου δορυφορικής τηλεπισκόπησης η οποία θα είναι συνεχής, παγκόσμια, αυτόνομη, μεγάλης εμβέλειας και υψηλής ποιότητας. Μέσω αυτής της μεθόδου λαμβάνονται ακριβείς, έγκαιρες και εύκολα προσβάσιμες πληροφορίες συμβάλλοντας έτσι, μεταξύ άλλων, στην βελτίωση της διαχείρισης του περιβάλλοντος, στην κατανόηση και τον μετριασμό των επιπτώσεων της κλιματικής αλλαγής, καθώς και στην διασφάλιση της αστικής ασφάλειας [10].

Στα πλαίσια του συγκεκριμένου προγράμματος, ο ΕΟΔ αναπτύσσει μια νέα οικογένεια δορυφόρων, οι οποίοι ονομάζονται Sentinel, με σκοπό να καλύψει τις συγκεκριμένες και εξειδικευμένες λειτουργικές ανάγκες του προγράμματος αυτού. Μέχρι στιγμής, σε τροχιά περί τη Γη βρίσκονται τρεις αστερισμοί αποτελούμενοι από δύο δορυφόρους, καθώς επίσης και ένας μονήρης δορυφόρος, ο Sentinel 5P. Η οικογένεια των δορυφόρων αυτών, αποτελείται από τους εξής αστερισμούς :

- *Sentinel 1* : Προσφέρει ένα σύνολο παρατηρήσεων αποτελούμενο από εικόνες ραντάρ που έχουν ληφθεί τόσο την ημέρα όσο και την νύχτα και περιέχουν όλες τις πιθανές καιρικές συνθήκες.
- *Sentinel 2* : Έχει σχεδιαστεί για να προσφέρει υψηλής ευκρίνειας οπτικές εικόνες για υπηρεσίες που αφορούν την επιφάνεια της Γης.
- *Sentinel 3* : Παράγει δεδομένα χρήσιμα για υπηρεσίες των οποίων οι εφαρμογές αφορούν τόσο τους ωκεανούς όσο και τις χερσαίες εκτάσεις του πλανήτη μας.
- *Sentinel 4 – 5* : Σχεδιάστηκαν με σκοπό να παρέχουν δεδομένα παρακολούθησης της ατμοσφαιρικής σύνθεσης, μέσω γεωστατικής και πολιτικής τροχιάς αντίστοιχα.
- *Sentinel 6* : Ο αστερισμός αυτός σχεδιάζεται με στόχο την μέτρηση του επίπεδου της στάθμης της θάλασσας, μέσω ραντάρ, σε παγκόσμιο επίπεδο. Τα προκύπτοντα δεδομένα, θα προορίζονται κυρίως για εφαρμογές της ωκεανογραφίας καθώς και για κλιματικές μελέτες.

Συνοψίζοντας, το πρόγραμμα Copernicus παρέχει ένα ενοποιημένο σύστημα μέσω του οποίου τεράστιες ποσότητες δεδομένων τροφοδοτούνται σε μια σειρά θεματικών υπηρεσιών που έχουν σχεδιαστεί για να ωφελήσουν το περιβάλλον, τον ανθρώπινο τρόπο ζωής, τις ανθρωπιστικές ανάγκες και για να στηρίξουν ,παράλληλα, την χάραξη μίας αποτελεσματικής πολιτικής για ένα πιο βιώσιμο μέλλον.

Τέλος, αξίζει να αναφερθεί πως οι θεματικές υπηρεσίες στις οποίες απευθύνεται το πρόγραμμα μπορούν να χωριστούν στις εξής έξι κύριες κατηγορίες :

- Η διαχείριση της γης
- Το υδάτινο περιβάλλον
- Η κλιματική αλλαγή
- Η ασφάλεια
- Η ατμόσφαιρα
- Η ανταπόκριση σε έκτακτες καταστάσεις. [10]

2.2.3 Η αποστολή Sentinel 2

2.2.3.1 Γενικά Στοιχεία

Το Sentinel 2 είναι μια αποστολή πολυφασματικής απεικόνισης ευρείας έκτασης και υψηλής ανάλυσης, με σκοπό την υποστήριξη μελετών παρακολούθησης του εδάφους ,στα πλαίσια του προγράμματος Copernicus, συμπεριλαμβανομένης της παρακολούθησης της βλάστησης, της κάλυψης του εδάφους (Land Cover) και των υδάτων. Επιπλέον, η αποστολή Sentinel 2 παρέχει δεδομένα που ενισχύουν μελέτες σχετικές με την παρατήρηση των εσωτερικών πλωτών οδών και των παράκτιων περιοχών [11].

Συγκεκριμένα, η αποστολή παρέχει πληροφορίες που αφορούν γεωργικές και δασικές διαδικασίες καθώς και τη διαχείριση της επισιτιστικής ασφάλειας. Οι δορυφορικές αυτές εικόνες, μπορούν να χρησιμοποιηθούν για τον προσδιορισμό διαφόρων δεικτών των φυτών όπως παραδείγματος χάριν δείκτες χλωροφύλλης καθώς και δείκτες του περιεχομένου του νερού. Εκτός, όμως, από την παρακολούθηση της ανάπτυξης των φυτών, ο αστερισμός Sentinel-2 μπορεί να χρησιμοποιηθεί για τη χαρτογράφηση των αλλαγών στην κάλυψη γης (Land Cover) και για την παρακολούθηση των παγκόσμιων δασών. Παρέχει επίσης πληροφορίες για τη ρύπανση σε λίμνες και παράκτια ύδατα. Τέλος, οι εικόνες πλημμυρών, ηφαιστειακών εκρήξεων και κατολισθήσεων που προσφέρει, συμβάλλουν στη χαρτογράφηση καταστροφών και βοηθούν τις ανθρωπιστικές προσπάθειες ανακούφισης [12].

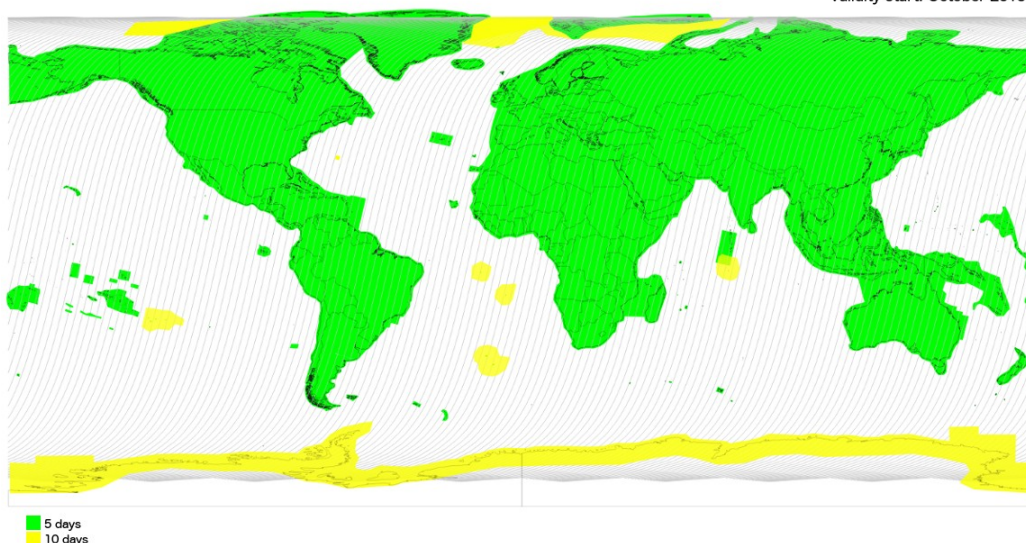
Ο αστερισμός των Sentinel-2 αποτελείται από δύο πανομοιότυπους δορυφόρους οι οποίοι βρίσκονται σε πολική και συγχρονισμένη με τον ήλιο τροχιά σε υψόμετρο 786 χιλιομέτρων. Οι δύο αυτοί δορυφόροι εκτοξεύτηκαν στις 23 Ιουνίου του 2015 (Sentinel-2A) και στις 7 Μαρτίου του 2017 (Sentinel-2B), εξασφαλίζοντας με αυτό τον τρόπο την ύπαρξη διαφοράς φάσης μεταξύ τους ίση με 180° . Η διαμόρφωση αυτή των δύο δορυφόρων στον αστερισμό του Sentinel-2 επιτυγχάνει βέλτιστη κάλυψη της επιφάνειας της Γης και μετάδοση των δεδομένων. Αναλυτικότερα, οι δύο αυτοί δορυφόροι μαζί καλύπτουν όλες τις χερσαίες επιφάνειες της Γης, τα μεγάλα νησιά, εσωτερικά (λίμνες, ποτάμια) και παράκτια ύδατα μεταξύ των γεωγραφικών μηκών 84° βόρεια και 84° νότια, με συχνότητα επανεπίσκεψης ίση με πέντε ημέρες [12], [13]. Επιπλέον, αξίζει να σημειωθεί πως, εξαιτίας των επικαλύψεων ορισμένων ζωνών από γειτονικές τροχιές, η συχνότητα επανεπίσκεψης ορισμένων περιοχών υπό διαφορετικές βέβαια, συνθήκες παρατήρησης (γωνία λήψης, κλίση της Γης κ.α.) δύναται να αυξηθεί [14].

Στις επόμενες δύο εικόνες φαίνονται οι περιοχές της επιφάνειας της Γης που έχουν συχνότητα επανεπίσκεψης ίση με πέντε ημέρες, καθώς και οι περιοχές οι οποίες υπό διαφορετικές συνθήκες παρακολούθησης γίνεται να επαναπαρατηρηθούν από τον αστερισμό των δορυφόρων σε λιγότερο από πέντε ημέρες.

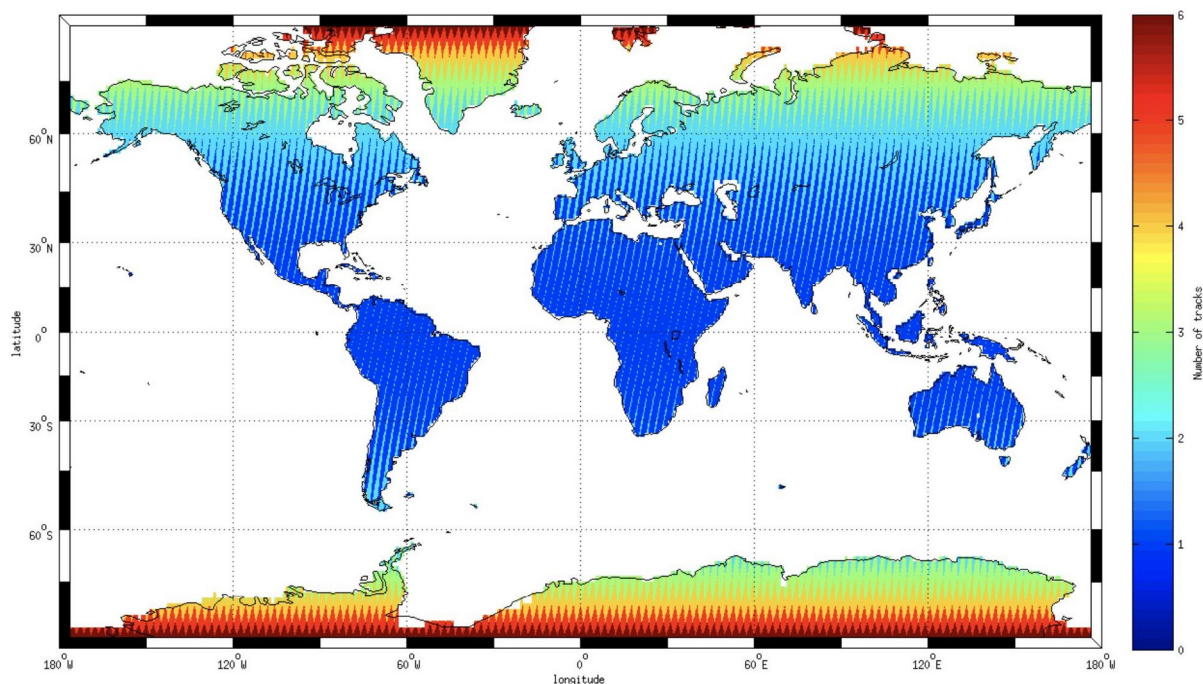
Sentinel-2 Constellation Observation Scenario: Revisit Frequency



Validity start: October 2019



Εικόνα 1: Η κάλυψη της αποστολής Sentinel 2 και οι χρόνοι επανεπίσκεψης των περιοχών αυτών [14].



Εικόνα 2: Η γεωμετρική συχνότητα επανεπίσκεψης εξαιτίας της επικάλυψης ζωνών από γειτονικές τροχιές [14].

2.2.3.2 Τεχνικές Προδιαγραφές

Στους δορυφόρους Sentinel-2 το βασικό όργανο λήψης εικόνων είναι το λεγόμενο “Multispectral Instrument” (MSI), το οποίο λαμβάνει μετρήσεις από 13 συχνотικές ζώνες της ανακλώμενης στην Γη ακτινοβολίας, το συνολικό φάσμα των οποίων εκτείνονται από τα 443 nm έως και τα 2190 nm του ηλεκτρομαγνητικού φάσματος, εντός (χωρικών) ζωνών εύρους 290 km. Πιο συγκεκριμένα, δειγματοληπτεί :

- Τρεις ορατές μπάντες και μία Visible and Near-Infrared (VNIR) μπάντα με χωρική ανάλυση ίση με 10m GSD (Ground Sampling Distance).
- Έξι υπέρυθρες μικροκυματικές (SWIR) μπάντες με χωρική ανάλυση ίση με 20m GSD.
- Τρεις μπάντες ατμοσφαιρικής διόρθωσης με χωρική ανάλυση ίση με 60m GSD. [11], [13]

Στον παρακάτω πίνακα παρουσιάζονται συνοπτικά τα κεντρικά μήκη κύματος και το αντίστοιχο bandwidth καθώς και η χωρική ανάλυση κάθε μίας από τις 13 φασματικές ζώνες που καταγράφει ο αστερισμός των Sentinel-2.

Πίνακας 1: Οι φασματικές ζώνες των αισθητήρων των Sentinel-2 δορυφόρων.

	S2A		S2B		
Band Number	Central wavelength (nm)	Bandwidth (nm)	Central wavelength (nm)	Bandwidth (nm)	Spatial resolution (m)
1	442.7	21	442.3	21	60
2	492.4	66	492.1	66	10
3	559.8	36	559.0	36	10
4	664.6	31	665.0	31	10
5	704.1	15	703.8	16	20
6	740.5	15	739.1	15	20
7	782.8	20	779.7	20	20
8	832.8	106	833.0	106	10
8a	864.7	21	864.0	22	20
9	945.1	20	943.2	21	60
10	1373.5	31	1376.9	30	60
11	1613.7	91	1610.4	94	20
12	2202.4	175	2185.7	185	20

Στο σημείο αυτό, αξίζει να αναφερθεί πως ο σχεδιασμός του MSI ,και η κατασκευή του, πραγματοποιήθηκε από την Airbus Defense and Space και καθοδηγήθηκε από την απαίτηση για μεγάλη γεωμετρική (χωρική) και φασματική απόδοση των μετρήσεων [15]. Συγκεκριμένα, οι αισθητήρες πρέπει να επιδεικνύουν ισχυρή ομοιομορφία, όσων αφορά την γεωμετρική απόδοση, μεταξύ των εστιακών επιπέδων ώστε να αποφεύγεται η παραμόρφωση της εικόνας. Παράλληλα, σχετικά με την ραδιομετρική απόδοση των δορυφόρων έχει υιοθετηθεί μια αυστηρή γεωμετρική διάταξη όλων των οπτικών και μηχανικών στοιχείων, προκειμένου να αποφευχθεί στον βέλτιστο βαθμό οποιαδήποτε, μη σχετική με την υπό παρατήρηση περιοχή, εξωτερική ηλιακή ακτινοβολία. Προκειμένου, λοιπόν, να ικανοποιηθούν οι παραπάνω απαιτήσεις το MSI των Sentinel-2 δορυφόρων διαθέτει τους εξής αισθητήρες και χαρακτηριστικά :

- * Ένα TMA (Three Mirror Anastigmat) τηλεσκόπιο με διάμετρο οπής ίση με 150 mm, ισοστατικά τοποθετημένη στην πλατφόρμα προκειμένου να ελαχιστοποιηθούν οι θερμοελαστικές παραμορφώσεις. Η σχεδίαση των φακών ,από οπτικής άποψης, έχει

βελτιστοποιηθεί έτσι ώστε να επιτευχθεί κορυφαία ποιότητα απεικόνισης σε ολόκληρο το οπτικό του φάσμα. Ο σκελετός του τηλεσκόπιο και οι ίδιοι οι φακοί είναι κατασκευασμένοι από καρβίδιο του πυριτίου, εξασφαλίζοντας με αυτό το τρόπο ιδιαίτερα υψηλό λόγο σταθερότητας- μάζας και ελαχιστοποιώντας τις θερμικές παραμορφώσεις.

- * Δύο εστιακά επίπεδα κατασκευασμένα από :
 - μονολιθικούς CMOS (Complementary metal-oxide-semiconductor) αισθητήρες για τον εντοπισμό των VNIR κυμάτων.
 - MCT (Mercury - Cadmium - Telluride) αισθητήρες υβριδοποιημένους πάνω σε ένα CMOS κύκλωμα για την ανίχνευση των SWIR κυμάτων.
- * Ένα διχρωμικό διαχωριστή δέσμης VNIR / SWIR. Ο φασματικός διαχωρισμός στις διάφορες συχνοτικές ζώνες παρέχεται από φίλτρα λωρίδων που είναι τοποθετημένα πάνω από τους ανιχνευτές.
- * Δύο ξεχωριστές συστοιχίες 12 ανιχνευτών τοποθετημένες σε κάθε εστιακό επίπεδο καλύπτοντας τα κανάλια VNIR και SWIR αντίστοιχα. Οι 12 ανιχνευτές σε κάθε εστιακό επίπεδο βρίσκονται σε κλιμακωτή διαμόρφωση ώστε να καλύπτουν ολόκληρο το οπτικό πεδίο.
- * Έναν ενσωματωμένο διαχύτη πλήρους πεδίου για την πραγματοποίηση ραδιομετρικής βαθμονόμησης ώστε να εξασφαλισθεί υψηλή ποιότητα ραδιομετρικής απόδοσης.
- * Ένα μηχανισμό κλείστρου που εμποδίζει την άμεση έκθεση του οργάνου στην ηλιακή ακτινοβολία. Αυτός ο μηχανισμός χρησιμοποιείται επίσης ως συσκευή βαθμονόμησης συλλέγοντας το φως του ήλιου αφού πρώτα ανακλαστεί σε ένα διαχύτη [15].

Τέλος, αξίζει να αναφερθεί πως κάθε δορυφόρος έχει ιδιαίτερα μεγάλη αυτονομία, έτσι ώστε να μπορεί να λειτουργήσει δίχως εξωτερική παρέμβαση από την Γη για έως και τα 15 ημέρες [16].

2.2.3.3 Τα προϊόντα

Ο αστερισμός των Sentinel-2 παράγει , καθημερινώς, 1 Terabyte (TB) όγκο δεδομένων. Ο τεράστιος αυτός όγκος δεδομένων οφείλεται στο γεγονός ότι οι παραγόμενες εικόνες χωρικού εύρους 290 km διαθέτουν 13 διαφορετικές συχνοτικές μπάντες σε τρεις διαφορετικές χωρικές αναλύσεις [16]. Γίνεται, επομένως, εύκολα κατανοητό πως η μετάδοση των παραγόμενων δεδομένων από τον αστερισμό προς την Γη είναι ένα εξίσου σημαντικό και απαιτητικό έργο.

Μόλις, λοιπόν, ληφθούν και κατασκευαστούν τα δεδομένα στο σύστημα των Sentinel-2 δορυφόρων, αυτά αποστέλλονται προς την Γη προκειμένου να ληφθούν από τους X-band κεντρικούς σταθμούς, οι οποίοι είναι τοποθετημένοι στην Ματέρα της Ιταλίας, την Μασπαλόμας της Ισπανίας και Σβάλμπαντ της Νορβηγίας. Επιπλέον, τα δεδομένα δύναται να μεταδοθούν με μια δέσμη λέιζερ, πρώτα, προς το Ευρωπαϊκό Σύστημα Αναμετάδοσης Δεδομένων, το οποίο αναμεταδίδει δεδομένα χρησιμοποιώντας γεωστατικούς δορυφόρους προς τους Ka-band σταθμούς των χρηστών. Η δεύτερη μέθοδος μετάδοσης της πληροφορίας, αποδεικνύεται ιδιαίτερα χρήσιμη όταν ο αστερισμός των Sentinel-2 δεν είναι εντός εμβέλειας με κάποιον από τους X-band σταθμούς στην Γη, επιτρέποντας έτσι γρηγορότερη διαθεσιμότητα των δεδομένων στους τελικούς χρήστες [17].

Τα δεδομένα, ή αλλιώς προϊόντα, που μεταδίδονται από τους Sentinel-2 δορυφόρους προς την Γη καλούνται προϊόντα βαθμού-0 (Level-0) και είναι στην ουσία τα δεδομένα δίχως

καμία επεξεργασία σε συμπιεσμένη μορφή. Τα βαθμού-0 προϊόντα περιέχουν όλη την απαραίτητη πληροφορία ώστε να παραχθούν, στην συνέχεια, τα προϊόντα βαθμού-1 και πάνω.

Στην συνέχεια, τα βαθμού-1A προϊόντα είναι τα αποσυμπιεσμένα δεδομένα με τις φασματικές μπάντες αδρά συγκαταχωρημένες (coregistered) σε συνδυασμό με ορισμένες βοηθητικές πληροφορίες. Τα προϊόντα βαθμού-1B περιέχουν επιπλέον πληροφορίες για την ραδιομετρική διόρθωση των δεδομένων ακτινοβολίας. Το φυσικό γεωμετρικό μοντέλο για την διόρθωση των δεδομένων προσαρμόζει τις παραμέτρους του χρησιμοποιώντας τα διαθέσιμα σημεία ελέγχου εδάφους και προσαρτάται στο προϊόν, χωρίς ωστόσο να εφαρμόζεται εφαρμόζεται σε αυτό [18].

Τα προϊόντα βαθμού-1C παρέχουν ορθό-ανακατασκευασμένες πληροφορίες ανακλαστικότητας για το υψηλότερο τμήμα της ατμόσφαιρας (Top-Of-Atmosphere), με πολυφασματικές καταχωρήσεις τύπου sub-pixel. Επιπλέον, στα προϊόντα αυτά συμπεριλαμβάνονται και μάσκες νέφωσης καθώς και μάσκες υδάτων/εδάφους. Τέλος, τα προϊόντα βαθμού-2A παρέχουν τις ίδιου είδους πληροφορίες με τα προϊόντα βαθμού-1C, μόνο που τώρα αφορούν το κατώτερο μέρος της ατμόσφαιρας, ενώ ταυτόχρονα περιέχουν και ένα “λεξικό” ταξινόμησης σκηνής (πχ νέφωση, σκίαση λόγω νέφωσης, καλλιέργειες, ύδατα, χιόνια και άλλα) [18].

Παρακάτω παρατίθεται μια λεπτομερέστερη παρουσίαση των προϊόντων βαθμού-1B έως και βαθμού-2A :

- Προϊόντα βαθμού-1B : Τα προϊόντα αυτού του βαθμού αποτελούνται από κόκκους (granules), ένας κόκκος αντιπροσωπεύει την υπο-εικόνα ενός από τους 12 ανιχνευτές κατά την κατεύθυνση της τροχιάς (25 km) και περιέχει έναν δεδομένο αριθμό γραμμών κατά μήκος της τροχιάς (περίπου 23 km). Κάθε κόκκος βαθμού-1B έχει όγκο δεδομένων περίπου ίσο με 27 MB. Τα δεδομένα αυτά αφορούν, επίσης, την ανακλαστικότητα στο ανώτατο στρώμα της ατμόσφαιρας.
- Προϊόντα βαθμού-1C : Περιέχουν πληροφορίες για την ανακλαστικότητα των ανώτερων στρωμάτων της ατμόσφαιρας υπό το πρίσμα μιας αυστηρής χαρτογραφικής γεωμετρίας (συνδυασμένη προβολή UTM - Universal Transverse Mercator - και ελλειψοειδές WGS84). Οι εικόνες βαθμού-1C είναι ένα σύνολο πλακιδίων (tiles) 100 τετραγωνικών χιλιομέτρων, καθένα από τα οποία είναι περίπου 500 Megabyte (MB). Στα προϊόντα αυτά έχουν εφαρμοστεί ραδιομετρικές και γεωμετρικές διορθώσεις.
- Προϊόντα βαθμού-2A : Περιέχουν πληροφορίες για την ανακλαστικότητα των κατώτερων στρωμάτων της ατμόσφαιρας, πάλι, υπό το πρίσμα μιας αυστηρής χαρτογραφικής γεωμετρίας [18].

Τα δεδομένα που λαμβάνονται από το MSI όργανο των Sentinel-2 δορυφόρων επεξεργάζονται συστηματικά, προκειμένου να φτάσουν στην μορφή προϊόντων βαθμού-1C, από το Payload Ground Segment (PDGS). Η επεξεργασία από τον βαθμό-0 μέχρι και τον βαθμό-1C πραγματοποιείται από την λειτουργικότητα Instrument Data Processing (IDP) του PDGS. Επιπλέον, σημειώνεται πως τα προϊόντα βαθμού-2A, αρχικά, επεξεργάζονταν από την πλευρά του χρήστη, χρησιμοποιώντας έναν επεξεργαστή που εκτελείται στην εργαλειοθήκη Sentinel-2 του ΕΟΔ, ενώ πλέον υπάρχει και η δυνατότητα παραγωγής τους και από το PDGS χρησιμοποιώντας τον επεξεργαστή Sen2Cor. Τέλος, κρίνεται σκόπιμο να σημειωθεί πως μόνο τα προϊόντα βαθμού-1C και βαθμού-2A είναι πλήρως διαθέσιμα στους τελικούς χρήστες μέσω της πλατφόρμας Copernicus Open Access Hub [19].

2.3. Βελτίωση της χωρικής ανάλυσης δορυφορικών εικόνων

Στο σύνολό τους, οι δορυφόροι που χρησιμοποιούνται για εφαρμογές επισκόπησης της γης, όπως λόγου χάριν οι IKONOS και QuickBird, παράγουν δύο είδη εικόνων, την πανχρωματική (PAN) εικόνα η οποία έχει υψηλή χωρική ανάλυση, αλλά χαμηλό φασματικό περιεχόμενο και τις πολυφασματικές (MS) εικόνες με υψηλή φασματική αλλά χαμηλή χωρική ανάλυση. Στην περίπτωση του αστερισμού Sentinel-2, δεν υπάρχει κάποιος αισθητήρας ικανός να παράξει την πανχρωματική εικόνα, ωστόσο στα πλαίσια της ανάλυσής μας, στην παρούσα εργασία, πραγματοποιούνται ορισμένες παραδοχές, προκειμένου να ενταχθούν τα προϊόντα των Sentinel-2 δορυφόρων στις παραπάνω δύο κατηγορίες. Συγκεκριμένα, η πανχρωματική εικόνα των Sentinel-2 δορυφόρων θεωρείται πως είναι αυτή που προκύπτει από τις ορατές και την VNIR μπάντα συχνοτήτων με χωρική ανάλυση 10m GSD, ενώ οι εικόνες που προκύπτουν από τους λοιπούς αισθητήρες ανήκουν στην κατηγορία των MS εικόνων του Sentinel-2.

Για πολλαπλές εφαρμογές στον τομέα της Παρατήρησης της Γης (Earth Observation), όπως λόγου χάριν ο εντοπισμός στόχου, η αναγνώριση προτύπων, η ανίχνευση μεταβολών, η ταξινόμηση κάλυψης εδάφους (land-cover) και άλλες, είναι ιδιαίτερα βοηθητική η χρήση εικόνων που συνδυάζουν υψηλή χωρική και φασματική ανάλυση, ωστόσο λόγω των τεχνολογικών περιορισμών των σύγχρονων δορυφορικών αισθητήρων η απευθείας λήψη τέτοιων εικόνων αποδεικνύεται αρκετά δύσκολη. Επομένως, μπορεί εύκολα κανείς να αντιληφθεί πως για την απόκτηση εικόνων πλούσιων τόσο σε φασματικές όσο και σε χωρικές πληροφορίες απαιτείται η πραγματοποίηση κάποιας μετά-επεξεργασίας των δορυφορικών εικόνων προκειμένου να λάβουμε τα επιθυμητά αποτελέσματα.

2.3.1 Ορισμός

Ένας από τους πιο διαδεδομένους τρόπους για την απόκτηση εικόνων με τα χαρακτηριστικά που αναφέραμε παραπάνω είναι η βελτίωση της χωρικής ανάλυσης των MS εικόνων που λαμβάνονται από τους δορυφόρους. Στην γενικότερη περίπτωση, η διαδικασία βελτίωσης της χωρικής ανάλυσης μίας εικόνας στην βιβλιογραφία, συχνά, καλείται *Super Resolution* (SR) και έχει ως γενικό στόχο την αύξηση της ανάλυσής της πέρα από τα όρια του αισθητήρα, την αύξηση δηλαδή του αριθμού των εικονοστοιχείων της εικόνας, παρέχοντας παράλληλα περισσότερες χωρικές λεπτομέρειες από αυτές που καταγράφονται από το αρχικό όργανο λήψης [20].

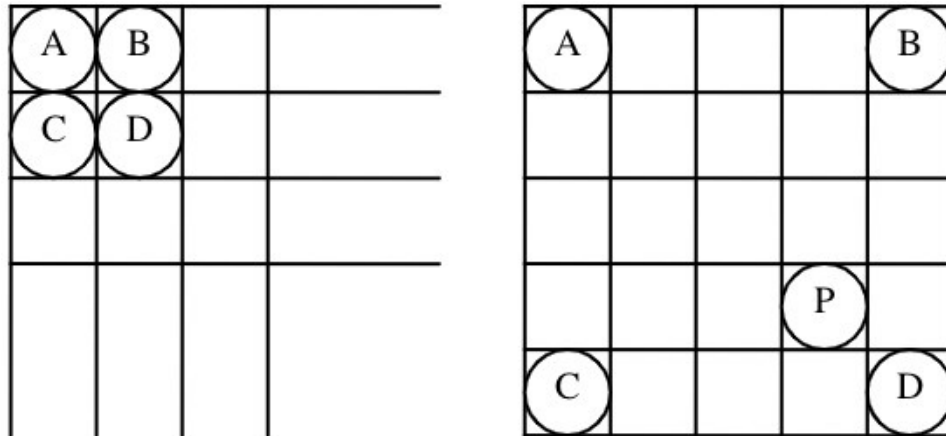
Στην προσπάθεια βελτίωσης της χωρικής ανάλυσης των εικόνων έχει υλοποιηθεί και χρησιμοποιηθεί μεγάλο πλήθος μεθόδων. Οι μέθοδοι αυτοί, εκτείνονται από απλούς αλγορίθμους παρεμβολής, σε σύνθετους αλγορίθμους σχεδιασμένους αποκλειστικά για δορυφορικές εικόνες και αλγορίθμους βαθιάς μάθησης.

2.3.2 Αλγόριθμοι παρεμβολής

Η μεγέθυνση ή γενικότερα η αλλαγή των διαστάσεων των εικόνων αποτελεί μία από τις βασικότερες λειτουργίες στον τομέα της επεξεργασίας εικόνων. Η διαδικασία της μεγέθυνσης μπορεί να θεωρηθεί ως η πρώτη, ιστορικά, προσπάθεια υλοποίησης ενός αλγορίθμου βελτίωσης της χωρικής ανάλυσης μίας εικόνας, καθώς έχοντας ως είσοδο μία εικόνα χαμηλής ανάλυσης αποσκοπεί στην δημιουργία ενός αντίγραφου, μεγαλύτερης ανάλυσης, το οποίο θα διατηρεί ταυτόχρονα το περιεχόμενο και τα χαρακτηριστικά της εικόνας εισόδου.

Ο βασικός στόχος της μεγέθυνσης μιας εικόνας, είναι η αύξηση του πλήθους των εικονοστοιχείων που περιέχονται σε αυτή, έτσι ώστε μία εικόνα χαμηλής ευκρίνειας να

μπορέσει να μετατραπεί σε εικόνα με υψηλότερη ανάλυση. Ωστόσο, όταν μια μικρή εικόνα μεγεθύνεται, παραδείγματος χάρη κατά παράγοντα 400% όπως φαίνεται στην εικόνα 3, οι τιμές των τεσσάρων, αρχικά, γειτονικών εικονοστοιχείων (που σημειώνονται με τα γράμματα A,B,C,D στο αριστερό μέρος) θα αποδοθούν στις νέες θέσεις των εικονοστοιχείων αυτών στην μεγεθυσμένη εικόνα, σύμφωνα με τον συντελεστή μεγέθυνσης. Το πρόβλημα, τώρα, είναι πως στην νέα εικόνα μεταξύ των εικονοστοιχείων A,B,C,D παρεμβάλλεται ένας αριθμός εικονοστοιχείων των οποίων η τιμή είναι άγνωστη, όπως παραδείγματος χάριν το εικονοστοιχείο P.



Εικόνα 3: Σχηματική απεικόνιση της μεγέθυνσης εικόνας [21].

Η παρεμβολή είναι ο τρόπος με τον οποίο υπολογίζονται οι άγνωστες τιμές των νέων εικονοστοιχείων που εισάγονται σε μία εικόνα και κατ' επέκταση ο τρόπος που οι εικόνες κλιμακώνονται. Υπάρχουν αρκετές και διαφορετικές μέθοδοι παρεμβολής, καθεμία από τις οποίες παράγει διαφορετικό τελικό αποτέλεσμα.

Οι αλγόριθμοι παρεμβολής σε εικόνες λειτουργούν προς δύο κατευθύνσεις, προκειμένου να υπολογίσουν την καλύτερη προσέγγιση του χρώματος και της έντασης ενός εικονοστοιχείου με βάση τις τιμές των γειτονικών του εικονοστοιχείων. Επιπλέον, ανάλογα με την μέθοδο υπολογισμού των τιμών των νέων εικονοστοιχείων, οι παραπάνω αλγόριθμοι διακρίνονται στις εξής δύο κατηγορίες :

- Μη προσαρμοστικοί (Non-Adaptive)
- Προσαρμοστικοί

2.3.2.1 Μη προσαρμοστικοί αλγόριθμοι

Οι τεχνικές μη προσαρμοστικής (Non-Adaptive) παρεμβολής βασίζονται σε άμεσο χειρισμό των εικονοστοιχείων αντί να εξετάζουν οποιοδήποτε χαρακτηριστικό ή περιεχόμενο μιας εικόνας. Αυτές οι τεχνικές ακολουθούν το ίδιο μοτίβο για όλα τα pixels και είναι εύκολες στην εκτέλεση και έχουν μικρότερο κόστος υπολογισμού. Ορισμένες μη προσαρμοστικές τεχνικές είναι η διγραμμική (Bilinear) και η δικυβική (Bicubic), η Lanczos, η Sinc, καθώς και οι προσέγγιση του πλησιέστερου γείτονα (Nearest Neighbor) [22].

2.3.2.1.1 Πλησιέστερος γείτονας

Αυτή είναι η απλούστερη προσέγγιση και απαιτεί τον λιγότερο χρόνο επεξεργασίας από όλους τους αλγόριθμους παρεμβολής. Ο αλγόριθμος κοντινότερου γείτονα (Nearest Neighbor)

επιλέγει την τιμή του πλησιέστερου pixel, στρογγυλοποιώντας τις συντεταγμένες του επιθυμητού σημείου παρεμβολής. Χρησιμοποιώντας αυτή τη μέθοδο βρίσκουμε το πλησιέστερο αντίστοιχο pixel στην εικόνα προέλευσης (αρχική) για κάθε pixel στην τελική εικόνα και στην συνέχεια τα κοντινότερα αυτά εικονοστοιχεία αντιγράφονται για τη δημιουργία νέων pixel καθώς μεγαλώνει η εικόνα. Αυτή η μορφή παρεμβολής, εξαιτίας της απλότητάς της, υποφέρει, συνήθως, από τα φαινόμενα του pixelation και της εμφάνισης εντόνων γωνιών που χωρίζουν τις καμπύλες σε σκαλοπάτια ή οδοντωτών ακμών, με αποτέλεσμα η τελική εικόνα να είναι ιδιαίτερα χαμηλής ποιότητας [21], [22].

2.3.2.1.2 Διγραμμική παρεμβολή

Η διγραμμική παρεμβολή παίρνει έναν σταθμισμένο μέσο όρο των τεσσάρων γειτονικών εικονοστοιχείων για να υπολογίσει την τελικά παρεμβαλλόμενη τιμή του νέου εικονοστοιχείου. Το αποτέλεσμα είναι μία πολύ πιο ομαλή εικόνα από την αρχική εικόνα. Στην περίπτωση που τα τέσσερα γειτονικά εικονοστοιχεία με γνώστες τιμές (από την αρχική εικόνα) ισαπέχουν από το νέα pixel, τότε η παρεμβαλλόμενη τιμή είναι ο απλός μέσος όρος των γνωστών τιμών. Αυτή η τεχνική εκτελεί παρεμβολή και στις δύο κατευθύνσεις, οριζόντια και κάθετη, με άμεση συνέπεια να δίνει καλύτερα αποτελέσματα από την παρεμβολή πλησιέστερου γείτονα. Ο πυρήνας παρεμβολής για τη διγραμμική παρεμβολή σε κάθε διάσταση της εικόνας είναι :

$$f(x) = \begin{cases} 0, & \text{if } |x| > 1 \\ 1 - |x|, & \text{if } |x| < 1 \end{cases}$$

, όπου x η απόσταση μεταξύ του νέου και του αρχικού pixel. [22], [23]

2.3.2.1.3 Δικυβική παρεμβολή

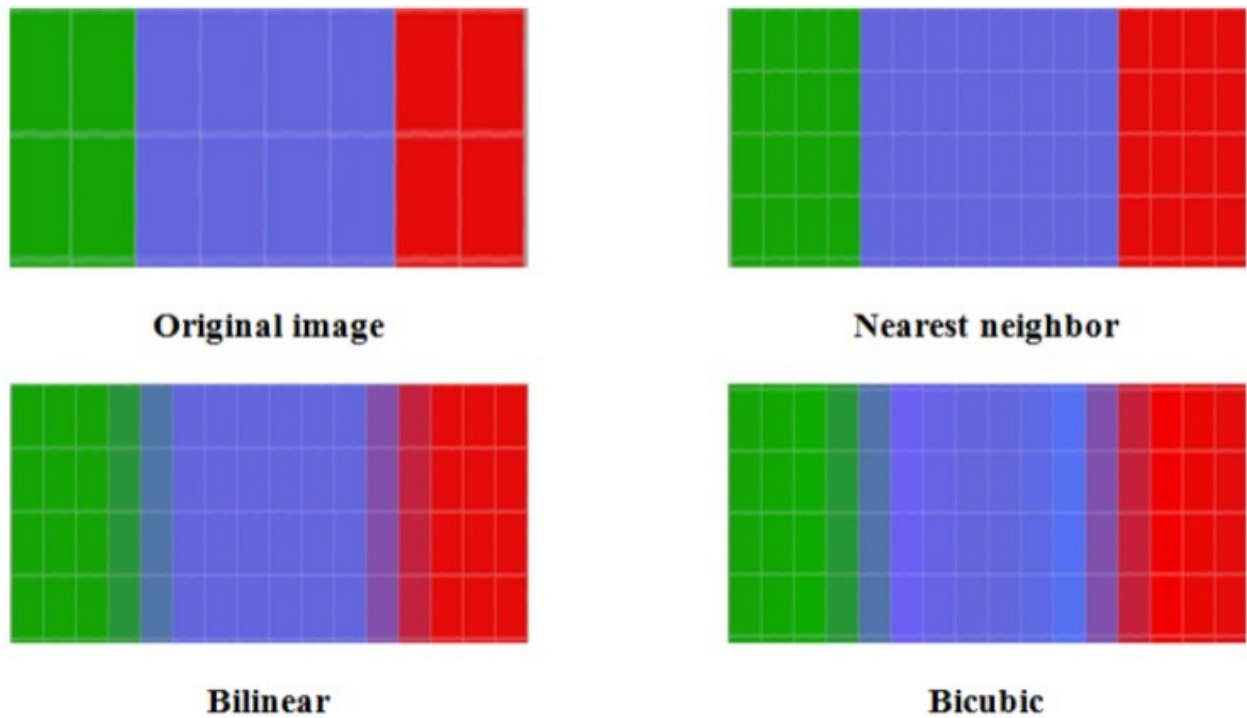
Η δικυβική παρεμβολή πηγαίνει ένα βήμα παρά πέρα από τη διγραμμική, λαμβάνοντας υπόψη την πλησιέστερη γειτονιά 4x4 γνωστών pixel για συνολικά δεκαέξι pixel. Εφόσον αυτά βρίσκονται σε διάφορες αποστάσεις από το άγνωστο εικονοστοιχείο, τα πιο κοντινά εικονοστοιχεία λαμβάνουν μεγαλύτερη στάθμιση στον υπολογισμό. Η μέθοδος αυτή παρεμβολής παράγει αισθητά πιο ευκρινείς εικόνες από τις δύο προηγούμενες μεθόδους και είναι ίσως ο ιδανικός συνδυασμός χρόνου επεξεργασίας και ποιότητας εξόδου. Για το λόγο χρησιμοποιείται ευρέως από πολλά προγράμματα επεξεργασίας εικόνων, όπως το Adobe Photoshop, και προγράμματα οδήγησης εκτυπωτών [21], [22].

Ο πυρήνας παρεμβολής της μεθόδους αυτής για κάθε διάσταση είναι :

$$f(x) = \begin{cases} \frac{3}{2}|x|^3 - \frac{5}{2}|x|^2 + 1, & \text{if } 0 \leq |x| < 1 \\ -\frac{1}{2}|x|^3 + \frac{5}{2}|x|^2 - 4|x| + 2, & \text{if } 1 \leq |x| < 2 \\ 0, & \text{if } 2 < |x| \end{cases}$$

, όπου x η απόσταση μεταξύ του νέου και του αρχικού pixel.[23]

Στην παρακάτω εικόνα, φαίνονται τα αποτελέσματα που παράγονται από τις τρεις προαναφερθείσες μεθόδους παρεμβολής.



Εικόνα 4: Οι παραγόμενες εικόνες από τις μεθόδους Nearest Neighbor, Bilinear και Bicubic.[23]

2.3.2.1.4 Δικυβική B-Spline

Όπως και η δικυβική παρεμβολή, ο αλγόριθμος παρεμβολής δικυβικών B-spline επίσης αξιολογεί τα πλησιέστερα δεκαέξι εικονοστοιχεία εικόνας εισόδου. Ωστόσο, αυτός ο αλγόριθμος χρησιμοποιεί συναρτήσεις παρεμβολής B-spline αντί για κυβικά splines, οι οποίες γενικά παράγουν αρκετά πιο ομαλά αποτελέσματα. Αναλυτικότερα, η μέθοδος των κυβικών B-splines εκτελεί μια συνέλιξη με ένα δισδιάστατο μη διαχωρίσιμο φίλτρο, και κατά συνέπεια έχει αυξημένη πολυπλοκότητα συγκριτικά με την δικυβική παρεμβολή. Αντίθετα, η δικυβική παρεμβολή χρησιμοποιεί μια συνέλιξη με ένα διαχωρίσιμο φίλτρο, και ως εκ τούτου η πολυπλοκότητά της είναι μικρότερη. Παρά τη διαφορά αυτή στην απόδοση, η μέθοδος των B-splines έχει ενδιαφέροντα χαρακτηριστικά ομαλότητας που την καθιστούν καλή επιλογή σε ορισμένες περιπτώσεις [22].

Ο πυρήνας παρεμβολής της μεθόδου αυτής είναι :

$$f(x) = \frac{1}{6} \begin{cases} 3|x|^3 - 6|x|^2 + 4, & \text{if } 0 \leq |x| < 1 \\ |x|^3 + 6|x|^2 - 12|x| + 8, & \text{if } 1 \leq |x| < 2 \\ 0, & \text{if } 2 \leq |x| \end{cases}$$

όπου x η απόσταση μεταξύ του νέου και του αρχικού pixel.

2.3.2.2 Προσαρμοστικοί αλγόριθμοι

Οι μη προσαρμοστικοί αλγόριθμοι παρεμβολής αποθηκεύουν μόνο τα στοιχεία χαμηλής συχνότητας της αρχικής εικόνας με αποτέλεσμα να αδυνατούν να διατηρήσουν καθαρές τις ακμές της εικόνας ή και να εισάγουν τεχνουργήματα (artifacts) γύρω από αυτές. Για την καλύτερη οπτική ποιότητα είναι αναγκαία η αξιοποίηση και των στοιχείων υψηλής συχνότητας, τα οποία είναι δυνατόν να διατηρηθούν από τους προσαρμοστικούς αλγορίθμους κλιμάκωσης των εικόνων.

Οι προσαρμοστικές (Adaptive) τεχνικές λαμβάνουν υπόψη τα χαρακτηριστικά της εικόνας εισόδου, όπως λόγου χάριν την τιμή της έντασης των pixel, τις ακμές, την υφή κ.α. Αυτές οι τεχνικές δίνουν καλύτερα αποτελέσματα από τις μη προσαρμοστικές τεχνικές, καθώς όπως προαναφέρθηκε διατηρούν τις υψίσυχνες λεπτομέρειες μίας εικόνας, αλλά απαιτούν περισσότερο υπολογιστικό χρόνο. Ορισμένα παραδείγματα προσαρμοστικών τεχνικών είναι η αλλαγή μεγέθους εικόνας με επίγνωση του περιβάλλοντος (Context aware image resizing), η χάραξη ραφής (Seam Carving), καθώς και οι τεχνικές που βασίζονται στη τμηματοποίηση της, υπό επεξεργασία, εικόνας. Στην συνέχεια, παρουσιάζονται επιγραμματικά ορισμένες από τις πιο γνώστες μεθόδους αλλαγής μεγέθους εικόνας με επίγνωση του περιβάλλοντος.

Υπάρχουν πολλές τεχνικές που προτείνονται για αλλαγή μεγέθους εικόνας με επίγνωση του περιβάλλοντος από διαφορετικούς ερευνητές. Σύμφωνα με τις τεχνικές αυτές, καθίσταται εφικτή η αλλαγή του μεγέθους της εικόνας για την περικοπή εικόνας, χωρίς ωστόσο να καθορίζεται το μέγεθος της εικόνας εξόδου. Στα πλαίσια λοιπόν της αλλαγής μεγέθους εικόνας με επίγνωση του περιβάλλοντος, οι Suh et al.[24] πρότειναν ότι τα αντικείμενα σε μια εικόνα είναι σημαντικά σε πολλές εργασίες ανάκτησης περιεχομένου. Ο κύριος στόχος της μεθόδου αυτής είναι, πριν από την κλιμάκωση μιας εικόνας, να περικόψει αυτόματα τις εικόνες προκειμένου να δημιουργήσει μία μικρογραφία αυτών (thumbnail). Στην συνέχεια, υπολογίζεται ο χάρτης προεξοχής και το παράθυρο περικοπής χρησιμοποιώντας άπληστη αναζήτηση (greedy search). Αυτός ο χάρτης προεξοχής λειτουργεί ως κριτήριο σπουδαιότητας για την αλλαγή του μεγέθους της εικόνας.

Παράλληλα, οι Zhang et al.[25] επιχείρησαν να μοντελοποιήσουν την περικοπή μιας εικόνας ως ένα πρόβλημα βελτιστοποίησης, στο οποίο η αντικειμενική συνάρτηση ορίζεται ως το άθροισμα τριών υπομοντέλων ενεργειών, το υπομοντέλο, δηλαδή, για την ενέργεια σύνθεσης, το υπομοντέλο για την ενέργεια διατήρησης και το υπομοντέλο για την ενέργεια ποιότητας. Στη συνέχεια, εφαρμόζεται η βελτιστοποίηση σμηνών σωματιδίων (Particle Swarm Optimization) για να ληφθεί η βέλτιστη λύση μεγιστοποιώντας την αντικειμενική συνάρτηση. Η υποψήφια λύση που μεγιστοποιεί την αντικειμενική συνάρτηση θα είναι το τελικό αποτέλεσμα, η περικεκομμένη δηλαδή εικόνα.

2.3.3 Αλγόριθμοι Pansharpening

Οι αλγόριθμοι παρεμβολής, που παρουσιάστηκαν στην παραπάνω ενότητα, παρά την ευκολία στην υλοποίησή τους, πολλές φορές αποδεικνύονται ανεπαρκείς για την επεξεργασία δορυφορικών εικόνων. Η αδυναμία των αλγορίθμων αυτών να δημιουργήσουν εικόνες μεγαλύτερων διαστάσεων, χωρίς να εισάγουν κάποιου είδους είτε χωρικής είτε φασματικής αλλοίωσης, οδήγησαν τους μελετητές του πεδίου στην ανάπτυξη νέων αλγορίθμων που στοχεύουν αποκλειστικά στην αύξηση των διαστάσεων δορυφορικών, πολυφασματικών εικόνων. Παραδοσιακά, η διαδικασία βελτίωσης της χωρικής ανάλυσης των πολυφασματικών εικόνων ονομάζεται Pansharpening και αποσκοπεί στην κατάλληλη ανάμιξη των PAN και MS εικόνων, που έχουν ληφθεί ταυτόχρονα και απεικονίζουν την ίδια περιοχή της επιφάνειας της

Γης, ώστε να κατασκευαστεί μια νέα εικόνα που να συνδυάζει την μεγάλη χωρική ευκρίνεια των πρώτων εικόνων με την υψηλή φασματική ανάλυση των δευτέρων. Το πρόβλημα αυτό μπορεί να ενταχθεί στην γενικότερη κατηγορία της συγχώνευσης δεδομένων, καθώς όπως προαναφέραμε σκοπός τους είναι ο συνδυασμός δύο διαφορετικών ειδών εικόνων για την κατασκευή μίας νέας εικόνας με βελτιωμένα χαρακτηριστικά. Επιπλέον, αξίζει να αναφερθεί πως το pansharpening, συγκριτικά με την γενικότερη συγχώνευση δεδομένων που έχουν προκύψει από πολλαπλούς αισθητήρες είναι ελαφρώς ευκολότερη καθώς δεν απαιτείται η ευθυγράμμιση των δύο εικόνων μίας και έχουν ληφθεί ταυτόχρονα από σχεδόν ίδια ακριβώς θέση [26].

Στα πλαίσια του pansharpening έχουν αναπτυχθεί πληθώρα αλγορίθμων που επιδιώκουν να υλοποιήσουν τις ιδέες της συγχώνευσης δεδομένων με σκοπό την κατασκευή μιας μοναδικής εικόνας με υψηλή χωρική και φασματική ανάλυση. Οι γνωστότερες κατηγορίες αλγορίθμων είναι οι *Component Substitution (CS)* και *Multiresolution Analysis (MRA)*, τις οποίες και παρουσιάζονται στις παρακάτω υποενότητες.

2.3.3.1 Component Substitution

Οι αλγόριθμοι *Component Substitution (CS)* αποτελούν την πιο κλασική προσέγγιση στο πρόβλημα του pansharpening. Οι μέθοδοι αυτού του είδους αφορούν την προβολή της MS εικόνας σε κάποιον άλλο χρωματικό χώρο, υπό την προϋπόθεση πως ο μετασχηματισμός αυτός διαχωρίζει ικανοποιητικά την χωρική δομή της εικόνας από τις φασματικές της πληροφορίες στις προκύπτουσες συνιστώσες. Στην συνέχεια, η χωρική ανάλυση της μετασχηματισμένης MS εικόνας ενισχύεται αντικαθιστώντας την συνιστώσα που περιέχει τις χωρικές δομές με την PAN εικόνα υψηλής ευκρίνειας. Στην συνέχεια, χρησιμοποιείται ο αντίστροφος χρωματικός μετασχηματισμός πάνω στην νέα πλέον εικόνα, αφού πρώτα υπερδειγματοληπτηθούν οι υπόλοιπες συνιστώσες στις διαστάσεις της PAN εικόνας, προκειμένου να ληφθεί το τελικό αποτέλεσμα. Από τον τρόπο που λειτουργεί η μέθοδος αυτή είναι εμφανές πως η ποιότητα του φασματικού περιεχομένου του τελικού αποτελέσματος εξαρτάται σε μεγάλο βαθμό από την ομοιότητα μεταξύ της PAN εικόνας και της αντικατασταθείσας συνιστώσας, ενώ παράλληλα η συγκεκριμένη κατηγορία αλγορίθμων συνδέεται με την εμφάνιση μεγάλων φασματικών παραμορφώσεων παρά την εγγυημένη βελτίωση της χωρικής ανάλυσης. Για τον λόγο αυτό, συχνά εφαρμόζεται ταίριασμα των ιστογραμμάτων της PAN εικόνας με την επιλεγείσα συνιστώσα, πριν την εκτέλεση της αντικατάστασης, προκειμένου να εξασφαλισθεί τουλάχιστον όμοια μέση τιμή και διασπορά μεταξύ της συνιστώσας και της ιστογραμμικά ταιριασμένης PAN εικόνας [26], [27].

Για παράδειγμα, στην περίπτωση του μετασχηματισμού της MS εικόνας στον χώρο Intensity Hue Saturation (IHS), η Intensity συνιστώσα θεωρείται πως περιέχει τις χωρικές δομές της MS εικόνας και συνεπώς είναι αυτή που θα αντικατασταθεί από την ιστογραμμικά ταιριασμένη αντίστοιχη PAN εικόνα. Στο σημείο αυτό αξίζει να αναφερθεί πως εκτός από τον χρωματικό μετασχηματισμό της MS εικόνας, ορισμένες υλοποιήσεις προτείνουν την χρήση PCA μετασχηματισμού και επιλογή της πρώτης συνιστώσας ως την περιέχουσα χωρικών πληροφοριών συνιστώσα. Τέλος, σημειώνεται πως στην κατηγορία αυτών των αλγορίθμων συμπεριλαμβάνεται και η μέθοδος *Gram-Schmidt spectral sharpening* [26].

Η συγκεκριμένη προσέγγιση στο πρόβλημα του pansharpening είναι καθολική, δηλαδή εφαρμόζεται και δρα με τον ίδιο τρόπο σε ολόκληρη την εικόνα, οδηγώντας έτσι σε ορισμένα πλεονεκτήματα, αλλά και μειονεκτήματα. Αναλυτικότερα, οι μέθοδοι αυτής της οικογένειας αλγορίθμων χαρακτηρίζονται, συνήθως, από αρκετά αξιόπιστη ικανότητα αύξησης της χωρικής

ανάλυσης του τελικού του προϊόντος, όντας ταυτόχρονα σχετικά γρήγορες και εύκολες ως προς την υλοποίησή τους. Στον αντίποδα, αδυνατούν να διαχειριστούν τοπικές ανομοιομορφίες μεταξύ της PAN και της MS εικόνας, οι οποίες οφείλονται στην όχι απόλυτη φασματική ταύτιση των δύο αυτών εικόνων, με αποτέλεσμα το τελικό προϊόν των μεθόδων της οικογένειας *Component Substitution* να περιέχει, συχνά, φασματικές παραμορφώσεις [26].

Σε μία προσπάθεια μοντελοποίησης της προσέγγισης των *Component Substitution* αλγορίθμων, οι *Tu et al* στο [28] πρότειναν πως, υπό την προϋπόθεση ύπαρξης ενός γραμμικού μετασχηματισμού της MS εικόνας που να καθιστά εφικτή την αντικατάσταση μίας μόνο κατάλληλης συνιστώσας όπως περιγράφεται παραπάνω, η διαδικασία σύντηξης των συνιστωσών μπορεί να επιτευχθεί χωρίς τον ρητό υπολογισμό του ευθέος και του αντιστρόφου μετασχηματισμού, αλλά μέσω ενός κατάλληλου σχήματος έγχυσης. Αυτή η παρατήρηση οδηγεί σε ταχύτερη εφαρμογή αυτών των μεθόδων και μπορεί να μοντελοποιηθεί μέσω της παρακάτω εξίσωσης :

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + g_k(\mathbf{P} - \mathbf{I}_L), \quad k = 1, \dots, N \quad (1)$$

,όπου ο δείκτης k αντιστοιχεί στην k -οστή φασματική μπάντα, N το πλήθος των φασματικών μπάντων που περιέχει η MS εικόνα και $\mathbf{g}=[g_1, \dots, g_k, \dots, g_N]$ είναι το διάνυσμα των συντελεστών κέρδους έγχυσης (*injection gains*).

Τα σύμβολα στην εξίσωση αυτή εξηγούνται στον παρακάτω πίνακα :

Πίνακας 2: Επεξήγηση των συμβόλων.

Symbol	Description
MS	MS image
$\widetilde{\mathbf{MS}}$	MS image interpolated at the scale of PAN
P	PAN image
$\widehat{\mathbf{MS}}$	Pansharpened image

Επιπλέον, σημειώνεται πως η εικόνα $\mathbf{I}_L$ δίνεται από την σχέση :

$$\mathbf{I}_L = \sum_{i=1}^N w_i \widetilde{\mathbf{MS}}_i \quad (2)$$

στην οποία το διάνυσμα βαρών $\mathbf{w} = [w_1, \dots, w_i, \dots, w_N]$ αντιστοιχεί στην πρώτη σειρά του πίνακα του ευθέος μετασχηματισμού που εφαρμόζεται στην εκάστοτε μέθοδο.

Παρακάτω παρατίθεται μία σύντομη ανάλυση των γνωστότερων μεθόδων που υπάγονται στην κατηγορία των *Component Substitution* αλγορίθμων.

2.3.3.1.1 Η μέθοδος (G)IHS

Η μέθοδος pansharpening IHS μετασχηματίζει την MS εικόνα στον ομώνυμο χρωματικό χώρο, ο οποίος μιμείται το ανθρώπινο οπτικό σύστημα κατά την επεξεργασία των πληροφοριών έντασης (I), απόχρωσης (H) και κορεσμού (S). Ο μετασχηματισμός IHS, υπό τον κλασικό του ορισμό, μπορεί να εφαρμοστεί μόνο σε έγχρωμες εικόνες RGB (Red-Green-Blue), οδηγώντας σε σημαντικούς περιορισμούς σχετικά με την επεξεργασία MS εικόνων. Στο [28], οι συγγραφείς έχουν γενικεύσει το φάσμα εφαρμογών του IHS μετασχηματισμού σε εικόνες με περισσότερα από τρία χρωματικά κανάλια, δημιουργώντας τον Generalized IHS μετασχηματισμό (GIHS), ενώ μεταγενέστερες μελέτες μοντελοποίησαν το πρόβλημα του *GIHS Component Substitution* για οποιοδήποτε αυθαίρετο μη αρνητικό σύνολο φασματικών βαρών w_i ως εξής :

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + \left(\sum_{i=1}^N w_i \right)^{-1} (\mathbf{P} - \mathbf{I}_L), \quad k = 1, \dots, N \quad (3)$$

όπου η εικόνα I_L δίνεται από την σχέση (2).

Στην γενικότερη περίπτωση τα βάρη $\{w_i\}_{i=1, \dots, N}$ τίθενται όλα στην τιμή $1/N$ [28], ωστόσο στο [29] προτείνεται μια βελτιστοποίηση των βαρών, προκειμένου αυτά να περιγράφουν καλύτερα τα φασματικά κανάλια των MS και PAN εικόνων. Επιπλέον, αξίζει να σημειωθεί πως παρόλο που τα φασματικά βάρη πρέπει να είναι μη αρνητικά, χωρίς ωστόσο να είναι υποχρεωτικό αυτά να αθροίζονται στη τιμή ένα.

Συνεχίζοντας την ανάλυση της (G)IHS μεθόδου, κρίνεται σκόπιμο να σημειωθεί πως αν στην εξίσωση (1) τεθούν οι συντελεστές κέρδους έγχυσης στις τιμές που δίνονται από την σχέση :

$$g_k = \frac{\widetilde{\mathbf{MS}}_k}{\mathbf{I}_L}, \quad k = 1, \dots, N \quad (4)$$

,τότε αυτή γίνεται :

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + \frac{\widetilde{\mathbf{MS}}_k}{\mathbf{I}_L} \cdot (\mathbf{P} - \mathbf{I}_L) = \widetilde{\mathbf{MS}}_k \cdot \frac{\mathbf{P}}{\mathbf{I}_L} \quad (5)$$

Η παραπάνω σχέση, μοντελοποιεί την ευρέως γνωστή μέθοδο Pansharpening μέσω του μετασχηματισμού Brovey (BT Pansharpening).

2.3.3.1.2 Η μέθοδος PCA

Το PCA, γνωστό και ως μετασχηματισμός Karhunen-Loève, είναι ένας άλλος μετασχηματισμός που έχει χρησιμοποιηθεί ευρέως για τη βελτίωση της χωρικής ανάλυσης φασματικών εικόνων. Ο μετασχηματισμός PCA επιτυγχάνεται μέσω μιας πολυδιάστατης περιστροφής του αρχικού συστήματος συντεταγμένων του N-διάστατου διανυσματικού χώρου. Αναλυτικότερα είναι ένας γραμμικός μετασχηματισμός των δεδομένων, έτσι ώστε η

προβολή των αρχικών φασματικών διανυσμάτων στους νέους άξονες, οι οποίοι είναι τα ιδιοδιανύσματα του πίνακα συνδιακύμανσης της φασματικής εικόνας, να παράγει ένα σύνολο βαθμωτών εικόνων, που ονομάζονται κύριες συνιστώσες (PCs), οι οποίες είναι στατιστικά ασυσχέτιστες μεταξύ τους. Οι συνιστώσες αυτές, στην συνέχεια, ταξινομούνται σε φθίνουσα σειρά ως προς την διακύμανσή τους, η οποία αποτελεί μέτρο των πληροφοριών που αυτές περιέχουν.

Όπως σε όλες τις μεθόδους της κατηγορίας *Component Substitution*, έτσι και τώρα γίνεται η υπόθεση ότι οι χωρικές πληροφορίες των MS εικόνων (οι οποίες είναι κοινές σε όλα τα κανάλια της εικόνας) συγκεντρώνονται, μετά τον μετασχηματισμό PCA, στην πρώτη συνιστώσα (PC_1), ενώ οι φασματικές πληροφορίες (συγκεκριμένες για κάθε μεμονωμένη μπάντα) επιμερίζονται στις υπόλοιπες $N-1$ συνιστώσες. Ωστόσο, η ισοδύναμη φασματική απόκριση της PC_1 , δηλαδή ο γραμμικός συνδυασμός των φασματικών αποκρίσεων του οργάνου λήψης της MS εικόνας σύμφωνα με τα βάρη του μετασχηματισμού PCA, ενδέχεται να μην ταιριάζει με τη φασματική απόκριση του οργάνου λήψης της PAN εικόνας. Στην πραγματικότητα, οι χωρικές πληροφορίες αντιστοιχίζονται στην πρώτη συνιστώσα, σε βαθμό ανάλογο με τη συσχέτιση μεταξύ των πολυφασματικών καναλιών. Η διαδικασία σύντηξης, όπως και στην περίπτωση του (G)IHS μετασχηματισμού, μπορεί να περιγραφεί από τη γενική διατύπωση της σχέσης (1), όπου τα διανύσματα συντελεστών $\mathbf{w}$ και $\mathbf{g}$ προέρχονται από τη διαδικασία PCA στην εικόνα MS. Αναλυτικότερα, όπως προαναφέρθηκε, το $\mathbf{w}$ είναι η πρώτη σειρά του πίνακα του ευθέως μετασχηματισμού, ενώ το διάνυσμα $\mathbf{g}$ στην περίπτωση του PCA είναι η πρώτη στήλη του πίνακα του αντίστροφου μετασχηματισμού [26].

2.3.3.2 Multiresolution Analysis

Η οικογένεια των *Multiresolution Analysis* (MRA) αλγορίθμων στοχεύει στην εξαγωγή λεπτομερειών υψηλής συχνότητας από την PAN εικόνα και στην έγχυση αυτών, στην συνέχεια, στην υπερδειγματοληπτημένη MS εικόνα. Αναλυτικότερα, η συμβολή της PAN εικόνας στην διαδικασία της βελτίωσης της χωρικής ανάλυσης της MS εικόνας επιτυγχάνεται μέσω του υπολογισμού της διαφοράς μεταξύ της πρώτης και μίας Low-Pass εκδοχής της. Η υψηλής ευκρίνειας MS εικόνα υπολογίζεται, στην γενικότερη περίπτωση, από την σχέση :

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + g_k(\mathbf{P} - \mathbf{P}_L), \quad k = 1, \dots, N.$$

Από την παραπάνω εξίσωση, επιβεβαιώνεται εύκολα πως σκοπός της συγκεκριμένης κατηγορίας μεθόδων είναι η διατήρηση του πλήρους φασματικού περιεχομένου της αρχικής MS εικόνας, ενώ ταυτόχρονα να προστεθεί επιπλέον πληροφορία η οποία έχει ληφθεί από την PAN εικόνα μέσω χωρικού φιλτραρίσματος [26], [27].

Η συγκεκριμένη κατηγορία αλγορίθμων έχει το πλεονέκτημα πως, εξαιτίας της διατήρησης του φασματικού περιεχομένου της MS εικόνας, είναι λιγότερη επιρρεπής στην εμφάνιση φασματικών παραμορφώσεων. Ωστόσο, οι μέθοδοι αυτοί βασίζονται σε μεγάλο όγκο πρότερης γνώσης και στην καλή ευθυγράμμιση των PAN και MS εικόνων.

Αναλυτικότερα σημειώνεται πως, ανεξάρτητα από τον τρόπο απόκτησης των χωρικών λεπτομερειών από την PAN εικόνα, πρέπει να πραγματοποιηθεί η έγχυση αυτών στην υπερδειγματοληπτημένη MS εικόνα. Προκειμένου, ωστόσο, η έγχυση αυτή να εισάγει την

λιγότερη δυνατή χωρική παραμόρφωση του εικονιζόμενου τοπίου είναι αναγκαία η κατασκευή ενός μοντέλου που βασίζεται στα δεδομένα που εμπλέκονται στην παραπάνω διαδικασία. Η υλοποίηση ενός τέτοιου μοντέλου μπορεί να πραγματοποιηθεί σε μία μικρότερη χωρική κλίμακα όπου τόσο η MS όσο και η PAN εικόνα είναι διαθέσιμες και στην συνέχεια να επεκταθεί στην ζητούμενη χωρική ανάλυση, υποθέτοντας πως αυτό θα ισχύει και στην χωρική αυτή κλίμακα. Η υπόθεση αυτή, όμως, πολλές φορές δεν ισχύει, ειδικά σε περιπτώσεις δεδομένων που διαθέτουν ήδη πολύ υψηλή χωρική ανάλυση, με αποτέλεσμα η τελική εικόνα, που παράγεται από τους MRA αλγορίθμους, να εμφανίζει αλλοιώσεις στις υπάρχουσες χωρικές δομές της ή και να περιέχει τεχνητές δομές (artifacts) οι οποίες δεν υπήρχαν στις αρχικές δύο εικόνες [30].

2.3.4 Σύγχρονες προσεγγίσεις

Οι παραπάνω δύο κατηγορίες αλγορίθμων βελτίωσης της χωρικής ανάλυσης των φασματικών δορυφορικών εικόνων, όπως παρουσιάσαμε και προηγουμένως, λειτουργούν υπό συγκεκριμένες, για την κατηγορία, προϋποθέσεις. Ωστόσο, πολλές φορές, στην πράξη, οι προϋποθέσεις αυτές αποδεικνύονται ιδιαίτερα απαιτητικές και δύσκολες να ικανοποιηθούν, με άμεση συνέπεια τα παραγόμενα, από τους αλγορίθμους αυτούς, αποτελέσματα να είναι συχνά χαμηλής ποιότητας και κατ' επέκταση όχι ιδιαίτερα χρήσιμα σε επόμενες εφαρμογές που θα τα αξιοποιούσαν.

2.3.4.1 Multi-Image Super Resolution

Μια πρώτη προσπάθεια στην βελτίωση της ποιότητας των εικόνων που παράγονται από τις παραπάνω κλασικές μεθόδους, είναι η οικογένεια των *Multi-Image Super Resolution* αλγορίθμων. Στην κατηγορία αυτή των αλγορίθμων, οι λεπτομέρειες υψηλής συχνότητας, που απαιτούνται για την σύνθεση της εικόνας υψηλής χωρικής ανάλυσης, θεωρείται πως βρίσκονται “διασκορπισμένες” σε ένα σύνολο εικόνων, της ίδιας σκηνής με μικρές (sub-pixel) μετατοπίσεις, χαμηλής ευκρίνειας. Κάθε μία, λοιπόν, από τις εικόνες του διαθέσιμου συνόλου δεδομένων αποτελεί έναν γραμμικό περιορισμό στο αντίστροφο πρόβλημα της ανακατασκευής (αποσυνέλιξης) της εικόνας υψηλής ευκρίνειας H [22], [23].

Αναλυτικότερα, θεωρούμε πως κάθε εικόνα χαμηλής ανάλυσης $L_j (j=1,2,...,n)$ έχει παραχθεί από την άγνωστη εικόνα υψηλής ανάλυσης H , μέσω μίας διαδικασίας θολώματος και υποδειγματοληψίας :

$$L_j = (H * B_j) \downarrow_{s_j}, \text{ όπου } B_j: \text{Blur Kernel},$$

$$\downarrow_{s_j}: \text{Υποδειγματοληψία κατά παράγοντα } s_j$$

Επομένως, κάθε pixel $p=(x,y)$ κάθε εικόνας χαμηλής ανάλυσης L_j εισάγει έναν γραμμικό περιορισμό στις άγνωστες τιμές των pixel της εικόνας υψηλής ανάλυσης εντός μιας γειτονίας γύρω από το αντίστοιχο pixel $q \in H$ (το μέγεθος της γειτονίας αυτής ορίζεται αποκλειστικά από το μέγεθος του πυρήνα θολώματος B_j). Το πρόβλημα που περιγράφηκε παραπάνω, μπορεί να μοντελοποιηθεί ως εξής :

$$L_j(p) = (H * B_j) \downarrow_{s_j}(q) = \sum_{q_i \in \text{support}(B_j)} H(q_i) B_j(q_i - q)$$

Συνεπώς, γίνεται κατανοητό πως αν υπάρχουν αρκετές εικόνες χαμηλής ευκρίνειας, τότε ο αριθμός των ανεξάρτητων εξισώσεων ξεπερνάει αυτό των αγνώστων και κατ' επέκταση ο

προσδιορισμός της εικόνας υψηλής ευκρίνειας καθίσταται εφικτός. Τέλος, αξίζει να σημειωθεί πως αυτού του είδους βελτίωση της χωρικής ανάλυσης εικόνων έχει αποδειχθεί πως παράγει ιδιαίτερα σταθερά και ικανοποιητικά αποτελέσματα για συντελεστές μεγέθυνσης που δεν ξεπερνούν την τιμή δύο [33].

2.3.4.2 Example-Based Super Resolution

Ορισμένες πιο σύγχρονες και σύνθετες μέθοδοι χρησιμοποιούν κάποιο εξωτερικό σύνολο δεδομένων, αποτελούμενο από ζεύγη εικόνων υψηλής και χαμηλής χωρικής ευκρίνειας, προκειμένου να μοντελοποιήσουν την σχέση-αντιστοιχία μεταξύ τμημάτων των εικόνων χαμηλής και υψηλής χωρικής ανάλυσης. Τελικά, εφαρμόζοντας την αντιστοιχία αυτή ,ως μετασχηματισμό, σε μία άγνωστη εικόνα χαμηλής ευκρίνειας επιδιώκουν να κατασκευάσουν την υψηλής ευκρίνειας εκδοχή της εικόνας εισόδου, η οποία αποτελεί και το ζητούμενο αποτέλεσμα [34]–[36]. Το βασικό πλεονέκτημα της μεθόδου αυτής, είναι πως δύναται να παράξει εικόνες για συντελεστές μεγέθυνσης μεγαλύτερους του δύο. Ωστόσο, σε αντίθεση με το Multi-Image SR, οι παραγόμενες εικόνες που ανακατασκευάζονται, δεν είναι εγγυημένο ότι περιέχουν τις πραγματικές (άγνωστες) λεπτομέρειες υψηλής ανάλυσης [33].

2.3.4.3 Single-Image Super Resolution

Οι αλγόριθμοι Super Resolution που βασίζονται στην ύπαρξη ενός συνόλου δεδομένων, είτε αυτό είναι σε μορφή πολλαπλών εικόνων χαμηλής χωρικής ανάλυσης της ίδιας σκηνής είτε σε μορφή ζευγών εικόνων υψηλής-χαμηλής ευκρίνειας, έχουν καταφέρει να βελτιώσουν σε κάποιο βαθμό την ποιότητα των αποτελεσμάτων συγκριτικά με τις κλασικές μεθόδους που παρουσιάστηκαν στην ενότητα 2.3.3 .

Στην πράξη, ωστόσο, και ειδικά στην περίπτωση των δορυφορικών εικόνων, η απόκτηση τέτοιων συνόλων δεδομένων μπορεί να αποδειχθεί ιδιαίτερα δύσκολη και απαιτητική διαδικασία. Σε μια προσπάθεια αντιμετώπισης του προβλήματος αυτού, της έλλειψης δηλαδή συνόλων δεδομένων αποτελούμενων από ζεύγη εικόνων υψηλής και χαμηλής ανάλυσης της ίδιας σκηνής, οι Glasner et al. στο [33] ,βασιζόμενοι στις προαναφερθείσες δύο τεχνικές, πρότειναν μια μέθοδο βελτίωσης της χωρικής ανάλυσης συμβατικών εικόνων, χρησιμοποιώντας αποκλειστικά και μόνο την υπό εξέταση εικόνα χαμηλής ευκρίνειας. Η μέθοδος αυτή, αν και προτάθηκε για συμβατικές RGB εικόνες, μπορεί εύκολα να επεκταθεί και στην κατηγορία των δορυφορικών εικόνων που εξετάζονται στην παρούσα εργασία. Για τον λόγο αυτό, σε συνδυασμό με την καινοτομία της υλοποίησης αυτής και την συμβολή της στην ανάπτυξη νεότερων αλγορίθμων που αξιοποιούνται για την εκπόνηση της παρούσας εργασίας, παρουσιάζεται συνοπτικά ,στην συνέχεια, η μέθοδος αυτή.

Η προσέγγιση των Glasner et al. , βασίζεται στην παρατήρηση πως τα επιμέρους τμήματα (patches) μιας ενιαίας φυσικής εικόνας τείνουν να επαναλαμβάνονται πολλαπλές φορές μέσα στην εικόνα, τόσο εντός της ίδιας κλίμακας όσο και σε διαφορετικές κλίμακες. Η επανάληψη των τμημάτων αυτών (με μετατοπίσεις σε sub-pixel επίπεδο) εντός της εικόνας ίδιας κλίμακας αποτελεί τη βάση για την εφαρμογή των περιορισμών στις εξισώσεις του κλασικού Multi-Image SR που παρουσιάσαμε στην ενότητα 2.3.4.1, από μια μεμονωμένη εικόνα. Αντίθετα, η επανεμφάνιση των επιμέρους τμημάτων της εικόνας σε διαφορετικές (πιο χονδροειδείς) κλίμακες της εικόνας παρέχει απλά παραδείγματα ζευγών εικόνων χαμηλής-υψηλής ανάλυσης, επιτρέποντας έτσι την εφαρμογή του Example-Based SR, της υποενότητας 2.3.4.2, αξιοποιώντας δεδομένα από μία και μόνο εικόνα (χωρίς την ανάγκη ύπαρξης κάποιας εξωτερικής βάσης δεδομένων ή προηγούμενων παραδειγμάτων). Αναλυτικότερα, σε περίπτωση εντοπισμού ενός όμοιου τμήματος της εικόνας σε κάποια, χαμηλότερης κλίμακας,

εκδοχή της μπορούμε να εξάγουμε το τμήμα-γονέα υψηλής ανάλυσης από την αρχική εικόνα και συνδυάζοντας τα δύο αυτά patches να κατασκευάσουμε ένα ζεύγος υψηλής-χαμηλής ανάλυσης εικόνων που θα αξιοποιηθούν στην συνέχεια από τον Example-Based αλγόριθμο [33].

Με βάση, λοιπόν, τις παραπάνω παρατηρήσεις, οι συγγραφείς πρότειναν μια ενοποιημένη προσέγγιση, στην οποία αξιοποιούν τις πολλαπλές εμφανίσεις ίδιων τμημάτων της εικόνας σε διάφορες κλίμακες προκειμένου να αναπτύξουν έναν προσαρμοστικό αλγόριθμο βελτίωσης της χωρικής ανάλυσης μιας εικόνας, χωρίς την ανάγκη ύπαρξης κάποιου εξωτερικού συνόλου δεδομένων. Αξίζει, τέλος, να σημειωθεί πως η προτεινόμενη μέθοδος εγγυάται πως τα παραγόμενα αποτελέσματα δεν θα είναι ποτέ χειρότερης ποιότητας από αυτά που θα προέκυπταν από έναν αλγόριθμο απλής παρεμβολής, καθώς σε περίπτωση που για κάποιο εικονοστοιχείο τα μόνο όμοια patches εντοπίζονται εντός της υπό εξέταση εικόνας τότε η μέθοδος επάγεται σε απλή εφαρμογή του Multi-Image SR αλγορίθμου, ενώ αν δεν βρεθεί κανένα όμοιο patch τότε η μέθοδος αντιστοιχεί σε απλή αποθορυβοποίηση της εικόνας εισόδου [33].

2.4. Μηχανική μάθηση

Η ιδέα των σκεπτόμενων μηχανών και της τεχνητής νοημοσύνης, παρά την ραγδαία ανάπτυξη που χαίρει τις τελευταίες δεκαετίες, άρχισε να σχηματίζεται σχεδόν ταυτόχρονα με την εμφάνιση των πρώτων ηλεκτρονικών υπολογιστών, όταν οι άνθρωποι ξεκίνησαν να αναρωτιούνται αν αυτά τα μηχανήματα θα μπορούσαν να αναπτύξουν κάποιου είδους νοημοσύνη η οποία, ενδεχομένως, με το πέρας των χρόνων θα ξεπερνούσε ακόμα και την ανθρώπινη.

Στα πρώτα χρόνια της ανάπτυξης του πεδίου της τεχνητής νοημοσύνης, ο κύριος στόχος των ερευνητών ήταν η υλοποίηση αλγορίθμων και τεχνικών που θα μπορούσαν να επιλύσουν προβλήματα αυξημένης (νοητικής) δυσκολίας για τους ανθρώπους. Οι αλγόριθμοι αυτοί κατάφεραν, πολύ σύντομα, να επιλύουν σύνθετες μαθηματικές εξισώσεις και να μοντελοποιούν περίπλοκες διεργασίες με εκπληκτική ταχύτητα και ακρίβεια, καθώς παρά την σύνθετη και δύσκολη στην κατανόηση, για τον άνθρωπο, φύση τους, τα προβλήματα αυτά μπορούσαν να περιγραφούν και να αναλυθούν από ένα σύνολο αυστηρά ορισμένων μαθηματικών κανόνων. Η πραγματική πρόκληση, παραδόξως, στον τομέα της τεχνητής νοημοσύνης, ακόμη και σήμερα, είναι η επίλυση των προβλημάτων εκείνων που για τους ανθρώπους φαίνονται εύκολα, τα οποία αντιμετωπίζονται με σχεδόν αυτόματο και διαισθητικό τρόπο, όπως παραδείγματος χάριν η ανίχνευση προσώπων σε μία εικόνα ή ο διαχωρισμός ενός σκύλου από μία γάτα. Ο λόγος που οι πρώτοι αλγόριθμοι τεχνητής νοημοσύνης απέτυχαν να επιλύσουν αυτή την φαινομενικά εύκολη κατηγορία προβλημάτων είναι η αδυναμία των ανθρώπων να θεμελιώσουν ένα, παρόμοιο, σύνολο αυστηρών κανόνων για να τα περιγράψουν και να τα αναλύσουν [37].

Προκειμένου, λοιπόν, να καταστεί εφικτή η αντιμετώπιση τέτοιου είδους προβλημάτων κρίνεται αναγκαίο οι υπολογιστές να μπορούν να εξάγουν πληροφορίες που βρίσκονται αποθηκευμένες σε δεδομένα και μέσω ενός αλγορίθμου μηχανικής μάθησης να αποκτήσουν εμπειρία που θα τους οδηγήσει, τελικά, στην επίλυση το προβλήματος. Σε μία προσπάθεια θεμελίωσης του πεδίου της μηχανικής μάθησης, ο *Mitchell* παρέχει στο βιβλίο του τον εξής ορισμό : “ Ένα υπολογιστικό πρόγραμμα θεωρούμε πως μαθαίνει από μία εμπειρία *E*, ως προς

μία κλάση διαδικασιών T και μια μετρική P , αν η επίδοσή του στην εργασία T βελτιώνεται μέσω της εμπειρίας E , ως προς την μετρική P "[38]. Ορισμένα παραδείγματα απλών αλγορίθμων μηχανικής μάθησης είναι το *Logistic Regression*, ο οποίος με βάση την είσοδό του μπορεί να υπολογίσει την πιθανότητα έκβασης ενός γεγονότος και ο *Naive Bayes ταξινομητής*, που κατατάσσει την είσοδό του σε κάποια πιθανή κατηγορία.

2.4.1 Εφαρμογές της μηχανικής μάθησης

Όπως αναφέρθηκε και προηγουμένως, οι αλγόριθμοι μηχανικής μάθησης αναπτύχθηκαν με σκοπό την διεύρυνση του φάσματος των προβλημάτων που μπορεί να διαχειριστεί ένας υπολογιστής, προβλημάτων δηλαδή τα οποία είναι δύσκολο να περιγραφούν από ένα σύνολο σχέσεων και κανόνων και που για την επίλυση τους χρειάζεται να αποκτηθεί γνώση από κάποιο εξωτερικό σύνολο δεδομένων. Οι κατηγορίες των προβλημάτων με τις οποίες σχετίζεται η μηχανική μάθηση είναι η εξής :

- **Ταξινόμηση** : Σκοπός της εργασίας αυτής είναι η κατάταξη ενός δεδομένου εισόδου σε μία από τις k πιθανές κατηγορίες. Τα προβλήματα ταξινόμησης εντοπίζονται πολύ συχνά στον τομέα της όρασης υπολογιστών, όπου συχνά ζητείται από μία φωτογραφία είτε να εντοπισθούν (και να ταξινομηθούν) τα εικονιζόμενα σε αυτή αντικείμενα ή και πρόσωπα , είτε να καταταχθεί η εικονιζόμενη σκηνή σε μία από τις πιθανές κατηγορίες τοπίων κ.α.
- **Παλινδρόμηση (Regression)** : Στην περίπτωση αυτών των προβλημάτων, ζητούμενο από τον αλγόριθμο είναι, με βάση την είσοδό του, να προβλέψει κάποια αριθμητική τιμή. Τα προβλήματα παλινδρόμησης είναι παρόμοια με εκείνα της ταξινόμησης μόνο που τώρα η έξοδος του αλγορίθμου έχει διαφορετική μορφή. Αναλυτικότερα για την επίλυση των προβλημάτων αυτών, ζητείται από τον αλγόριθμο μηχανικής μάθησης να μάθει μία συνάρτηση εξόδου της μορφής : $f: \mathbb{R}^n \rightarrow \mathbb{R}$. Ένα παράδειγμα προβλήματος παλινδρόμησης είναι η πρόβλεψη των μελλοντικών τιμών της αξίας κάποιου κρυπτονομίσματος, με βάση τις προηγούμενες γνωστές τιμές.
- **Μηχανική Μετάφραση (Machine Translation)** : Στην μηχανική μετάφραση, όπως προδιαθέτει και η ονομασία του προβλήματος, ζητούμενο είναι η μετατροπή μίας ακολουθίας συμβόλων σε κάποια άλλη γλώσσα.
- **Σύνθεση δεδομένων (Synthesis and Sampling)** : Σε αυτή την κατηγορία προβλημάτων, ο αλγόριθμος μηχανικής μάθησης καλείται να δημιουργήσει νέα δεδομένα, παρόμοια με εκείνα που είχε δει κατά την διαδικασία της εκπαίδευσής του. Η σύνθεση δεδομένων βρίσκει εφαρμογές σε ένα ευρύ φάσμα που εκτείνεται από τον τομέα το πολυμέσων έως και την ίδια την μηχανική μάθηση, σε περιπτώσεις που το σύνολο δεδομένων εκπαίδευσης δεν είναι επαρκώς μεγάλο για την επίλυση ενός προβλήματος.
- **Αποθορυβοποίηση (Denoising)** : Στα προβλήματα αποθορυβοποίησης, ο αλγόριθμος δέχεται ως είσοδο ένα αλλοιωμένο παράδειγμα $\tilde{x} \in \mathbb{R}^n$ το οποίο έχει προκύψει, μέσω μιας άγνωστης διεργασίας αλλοίωσης, από το παράδειγμα $x \in \mathbb{R}^n$. Στόχος είναι ο αλγόριθμος να καταφέρει να προβλέψει την καθαρή εκδοχή x από την αλλοιωμένη $\tilde{x}$, η γενικότερα να προβλέψει την υπό συνθήκη κατανομή πιθανότητας $p(x|\tilde{x})$.

Στο σημείο αυτό, είναι σημαντικό να επισημανθεί πως τα παραπάνω αποτελούν ορισμένα από τα πιο γνωστά πεδία, άλλα όχι ολόκληρο το φάσμα, εφαρμογών της μηχανικής μάθησης [37].

Τέλος, κρίνεται σκόπιμο να αναφερθεί ότι η ικανότητα των αλγορίθμων μηχανικής μάθησης να εκτελέσουν την εργασία για την οποία σχεδιάστηκαν, εξαρτάται σε μεγάλο βαθμό από την αναπαράσταση των δεδομένων εισόδου τους. Ως αναπαράσταση, αναγνωρίζεται ένα διάνυσμα που ορίζεται για κάθε ένα από τα δεδομένα εισόδου και το οποίο περιέχει τις σημαντικότερες πληροφορίες σχετικά με αυτό, καθιστώντας παράλληλα δυνατή την διάκρισή του από τα υπόλοιπα στοιχεία του συνόλου δεδομένων. Το διάνυσμα αναπαράστασης αποτελείται από μετρήσιμες ποσότητες, που ονομάζονται *χαρακτηριστικά (features)*. Επομένως, στην γενική περίπτωση που υπάρχουν l το πλήθος χαρακτηριστικά $x_i, i=1,2,\dots,l$, η αναπαράσταση ενός στοιχείου εισόδου είναι το διάνυσμα χαρακτηριστικών $x=[x_1, x_2, \dots, x_l]^T$, όπου ο τελεστής T υποδηλώνει το ανάστροφο διάνυσμα [37], [39].

2.4.2 Επιβλεπόμενη Μάθηση

Οι εφαρμογές της μηχανικής μάθησης για τις οποίες τα δεδομένα εκπαίδευσης έχουν την μορφή ζευγών διανύσματος χαρακτηριστικών - διανύσματος στόχου (ground truth) είναι γνωστές ως εφαρμογές *επιβλεπόμενης μάθησης (Supervised Learning)*. Ένα διάνυσμα στόχου αποτελεί την αναπαράσταση της αναμενόμενης, από το μοντέλο, εξόδου και χρησιμοποιείται για τον υπολογισμό των επιδόσεων του μοντέλου πάνω στην εκάστοτε εργασία που καλείται να εκτελέσει. Στην περίπτωση που διάνυσμα στόχου είναι ένα συγκεκριμένο σύνολο από κατηγορίες ή κλάσεις, τότε η εργασία που εκτελεί το μοντέλο υπάγεται στην κατηγορία της ταξινόμησης, που περιγράψαμε στην προηγούμενη υποενότητα. Διαφορετικά, αν το διάνυσμα στόχου περιέχει μία ή περισσότερες συνεχείς μεταβλητές, τότε το μοντέλο είναι σχεδιασμένο για να εκτελεί εργασίες παλινδρόμησης [40].

2.4.3 Μη Επιβλεπόμενη Μάθηση

Ένα διαφορετικό είδος μεθόδων μηχανικής μάθησης είναι η κατηγορία των αλγορίθμων *μη επιβλεπόμενης μάθησης (Unsupervised Learning)*. Σε αυτή την κατηγορία των αλγορίθμων, από το σύνολο εκπαίδευσης απουσιάζουν τα διανύσματα στόχου και, συνεπώς, τα μόνο διαθέσιμα δεδομένα είναι αυτά των διανυσμάτων χαρακτηριστικών.

Οι αλγόριθμοι μη επιβλεπόμενης μάθησης, πολύ συχνά, βρίσκουν εφαρμογές σε προβλήματα συσταδοποίησης (clustering), στα οποία ζητούμενο είναι η εύρεση ομοιοτήτων μεταξύ των δεδομένων εισόδου και η ομαδοποίησή τους, στην συνέχεια, σε συστάδες (clusters) δεδομένων με βάση τις ομοιότητες αυτές. Ορισμένα άλλα παραδείγματα προβλημάτων που επιλύονται μέσω αυτών των αλγορίθμων είναι η προβολή των δεδομένων εισόδου από κάποιο χώρο πολλαπλών διαστάσεων σε κάποιον άλλο χώρο λιγότερων διαστάσεων, με σκοπό την οπτικοποίηση των δεδομένων αυτών, καθώς και ο προσδιορισμός της κατανομής των στοιχείων του συνόλου δεδομένων, εντός του χώρου εισόδου [39], [40]. Ο προσδιορισμός της στατιστικής κατανομής που οδήγησε στο υπό εξέταση σύνολο δεδομένων, μπορεί στην συνέχεια να χρησιμοποιηθεί για την εκτέλεση εργασιών σχετικές με τη σύνθεση και την αποθρομβοποίηση δεδομένων [37].

Συνοψίζοντας, λοιπόν, θα μπορούσε κανείς να πει πως αλγόριθμοι μη επιβλεπόμενης μάθησης αφορούν την παρατήρηση ορισμένων δειγμάτων ενός τυχαίου διανύσματος

χαρακτηριστικών x , με σκοπό είτε τον προσδιορισμό της κατανομής πιθανότητας $p(x)$ αυτής καθ' αυτής, είτε την εξαγωγή ορισμένων χρήσιμων ιδιοτήτων αυτής. Αντίθετα, οι αλγόριθμοι επιβλεπόμενης μηχανικής μάθησης αφορούν την παρατήρηση πολλαπλών δειγμάτων ενός τυχαίου διανύσματος x μαζί με τα αντίστοιχα διανύσματα στόχου y , προκειμένου να μάθουν να προβλέπουν το y βλέποντας το διάνυσμα x . Για την πρόβλεψή τους, τα μοντέλα επιβλεπόμενης μάθησης, συχνά, υπολογίζουν μία εκτίμηση της υπό συνθήκη πιθανότητας $p(x|y)$ [37].

2.4.4 Βαθιά Μάθηση

Στην ενότητα 2.4.1, έγινε αναφορά στην επίδραση που έχουν τα διανύσματα χαρακτηριστικών, και πιο συγκεκριμένα τα ίδια τα χαρακτηριστικά, στις επιδόσεις ενός μοντέλου μηχανικής μάθησης. Αναλυτικότερα, παράγοντες όπως η επιλογή των κατάλληλων μεταβλητών που θα χρησιμοποιηθούν ως χαρακτηριστικά, το πλήθος των χαρακτηριστικών, ακόμη και η μορφή των χαρακτηριστικών, καθορίζουν σε πολύ μεγάλο βαθμό την ποιότητα των αποτελεσμάτων που παράγονται από έναν αλγόριθμο μηχανικής μάθησης. Ωστόσο, η επιλογή των κατάλληλων χαρακτηριστικών είναι διαδικασία που εξαρτάται άμεσα από το προς επίλυση πρόβλημα και το διαθέσιμο σύνολο, καθιστώντας έτσι την εύρεση της βέλτιστης αναπαράστασης των δεδομένων μία ιδιαίτερα απαιτητική διαδικασία.

Σε μία προσπάθεια αυτοματοποίησης της εύρεσης της βέλτιστης αναπαράστασης των δεδομένων εισόδου, προτάθηκε η επέκταση της χρήσης της μηχανικής μάθησης για την εύρεση, όχι απλά της αντιστοιχίας μεταξύ αναπαράστασης και εξόδου, αλλά και της ίδιας της αναπαράστασης. Οι αναπαραστάσεις αυτές, πολύ συχνά, αποδεικνύεται ότι οδηγούν σε καλύτερες και πιο γρήγορες επιδόσεις των μοντέλων, πυροδοτώντας έτσι την ανάπτυξη νέων τεχνολογιών και αλγορίθμων που ακολουθούν την συγκεκριμένη προσέγγιση. Προς την κατεύθυνση αυτή, της αυτοματοποιημένης δηλαδή εξαγωγής χαρακτηριστικών, η βαθιά μάθηση παρέχει την δυνατότητα εξαγωγής χαρακτηριστικών υψηλής σημασιολογικής αξίας απευθείας από τα δεδομένα εισόδου, εισάγοντας σύνθετες αναπαραστάσεις που προκύπτουν από άλλες, σημασιολογικά απλούστερες αναπαραστάσεις [37].

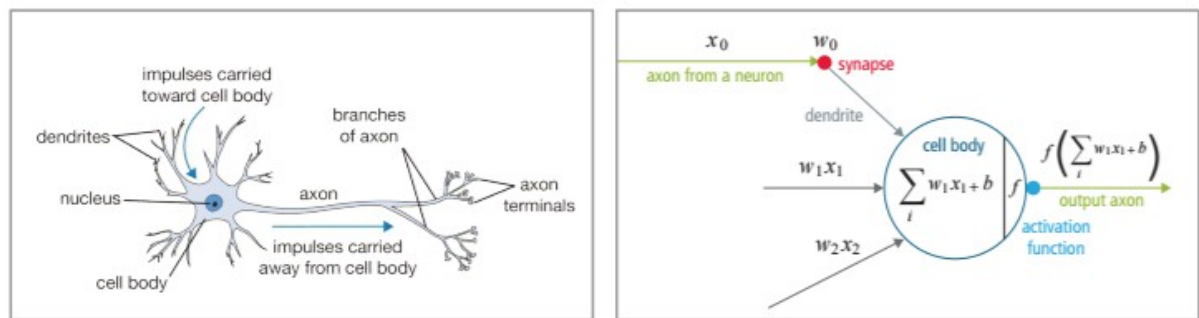
Η βαθιά μάθηση αποτελεί μία υποκατηγορία της μηχανικής μάθησης, η οποία αποσκοπεί στην αξιοποίηση πληθώρας θεμελιωδών μη γραμμικών υπολογιστικών μονάδων. Αναλυτικότερα, οι μονάδες αυτές υπολογίζουν τον γραμμικό συνδυασμό των εισόδων τους, τον οποίο στην συνέχεια περνάνε από μία μη γραμμική συνάρτηση, την αποκαλούμενη συνάρτηση ενεργοποίησης. Επιπρόσθετα, οι μονάδες αυτές ονομάζονται “τεχνητοί νευρώνες” και οργανώνονται σε δίκτυα, με τρόπο τέτοιο ώστε να προσομοιάζουν την δομή του ανθρώπινου εγκεφάλου. Τα προκύπτοντα δίκτυα είναι γνωστά ως νευρωνικά δίκτυα, των οποίων η εκπαίδευση επιτυγχάνεται μέσω μίας επαναληπτικής παρουσίασης των δειγμάτων εκπαίδευσης.

2.4.4.1 Το Perceptron

Τα νευρωνικά δίκτυα, παρά την σπουδαία ανάπτυξη που σημειώνεται στις μέρες μας πάνω στον συγκεκριμένο τομέα, έκαναν την πρώτη τους εμφάνιση στα τέλη της δεκαετίας του 1950, όταν ο Rosenblatt πρότεινε το μοντέλο Perceptron. Το Perceptron, αποτελεί την προσπάθεια του συγγραφέα να μοντελοποιήσει την λειτουργία ενός βιολογικού νευρώνα. Επιγραμματικά αναφέρεται πως οι βιολογικοί νευρώνες συνδέονται μεταξύ τους με συνάψεις. Κάθε νευρώνας λαμβάνει σήματα από τους δενδρίτες του και παράγει σήματα εξόδου κατά μήκος του νευράξονά του. Ο άξονας αυτός, τελικά, συνδέεται με τους δενδρίτες άλλων

νευρώνων και ούτω καθεξής. Στο υπολογιστικό μοντέλο του Perceptron τα σήματα που ταξιθεύουν στους νευράξονες αλληλεπιδρούν πολλαπλασιαστικά. Η ιδέα είναι ότι τα συναπτικά βάρη $\mathbf{W}$ μπορούν να μαθευτούν (με τη χρήση αλγορίθμων gradient descent) και ελέγχουν την επιρροή (και την κατεύθυνση της πληροφορίας για θετικά και αρνητικά βάρη) μεταξύ των νευρώνων. Σε ένα απλό μοντέλο, οι δονδρίτες μεταφέρουν το σήμα στο σώμα του κυττάρου όπου αθροίζονται όλα τα σήματα, υπολογίζεται δηλαδή ο γραμμικός συνδυασμός των σημάτων. Εάν το τελικό άθροισμα είναι θετικό, τότε η έξοδος του νευρώνα παίρνει την τιμή 1 και θεωρούμε πως ο νευρώνας ενεργοποιείται και στέλνει τις πληροφορίες κατά μήκος του άξονά του [41].

Στο παρακάτω σχήμα απεικονίζονται ένας βιολογικός νευρώνας (εικόνα 5α) και το μαθηματικό μοντέλο ενός τεχνητού νευρώνα (εικόνα 5β)



Εικόνα 5: Σύγκριση μεταξύ βιολογικού και τεχνητού νευρώνα
(α) (β)

Η λειτουργία του Perceptron, η οποία περιγράφηκε παραπάνω, συνοψίζεται από τις παρακάτω σχέσεις :

Έστω είσοδος $\mathbf{x} = [x_1, x_2, \dots, x_n]$, τότε ως γραμμικό συνδυασμό της εισόδου ορίζουμε την ποσότητα :

$$\sum_{i=1}^n w_i x_i = \mathbf{w}^T \mathbf{x} + b$$

,όπου $\mathbf{w} = [w_1, w_2, \dots, w_n]$, το διάνυσμα (συναπτικών) βαρών και b αριθμητική ποσότητα που καλείται bias.

Τελικά, η έξοδος του Perceptron δίνεται από την σχέση :

$$P_{out} = \begin{cases} 1, & \text{if } \mathbf{w}^T \mathbf{x} + b > 0 \\ -1, & \text{if } \mathbf{w}^T \mathbf{x} + b < 0 \end{cases}$$

Εξετάζοντας την παραπάνω σχέση, μπορεί κανείς εύκολα παρατηρήσει πως ο γραμμικός συνδυασμός της εισόδου του Perceptron αντιστοιχεί στο γραμμικό σύνορα $\mathbf{w}^T \mathbf{x} + b = 0$ στον n -διάστατο χώρο. Αυτό σημαίνει πως το Perceptron δύναται να χρησιμοποιηθεί για την επίλυση του προβλήματος της ταξινόμησης n -διάστατων γραμμικά διαχωρίσιμων διανυσμάτων $\mathbf{x}$ σε δύο κλάσεις. Αναλυτικότερα, είναι εφικτό να επιλυθεί το πρόβλημα της ταξινόμησης

δύο γραμμικά διαχωρίσιμων κλάσεων c_1, c_2 , υπολογίζοντας τις κατάλληλες τιμές w_i για το διάνυσμα βαρών και b , υπό την προϋπόθεση ότι :

$$\mathbf{w}^T \mathbf{x} + b = \begin{cases} > 0, & \text{if } \mathbf{x} \in c_1 \\ < 0, & \text{if } \mathbf{x} \in c_2 \end{cases}$$

Για υπολογισμό των άγνωστων τιμών χρησιμοποιείται ένας απλός επαναληπτικός αλγόριθμος εκπαίδευσης ο οποίος, σύμφωνα με το θεώρημα σύγκλισης του Perceptron [41], είναι εγγυημένο ότι θα οδηγήσει στην λύση του προβλήματος μετά από πεπερασμένο πλήθος επαναλήψεων. Ο αλγόριθμος εκπαίδευσης του Perceptron, όπως αποκαλείται, έχει την εξής μορφή :

Αρχικά, ορίζεται η παράμετρος η , η οποία ονομάζεται *learning rate* και καθορίζει το πόσο έντονη θα είναι η μεταβολή των τιμών του διανύσματος βαρών σε κάθε επανάληψη του αλγορίθμου. Επιπλέον, για την εκτέλεση της πρώτης επανάληψης, αποδίδονται τυχαίες τιμές στο διάνυσμα βαρών, το οποίο συμβολίζεται ως $\mathbf{w}(1)$.

Έστω ότι το σύνολο δεδομένων αποτελείται από N το πλήθος δείγματα $\mathbf{x}_j, j=1,2,\dots,N$, τότε για κάθε δείγμα $\mathbf{x}_j$ σε κάθε βήμα k , με $k=2,3,4,\dots$ εκτελούνται οι παρακάτω ενέργειες [41] :

- Αν $\mathbf{w}^T \mathbf{x} + b \leq 0$ και $\mathbf{x}_j \in c_1$

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \eta \mathbf{x}_j$$
- Αν $\mathbf{w}^T \mathbf{x} + b \geq 0$ και $\mathbf{x}_j \in c_2$

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \eta \mathbf{x}_j$$
- Αλλιώς
$$\mathbf{w}(k+1) = \mathbf{w}(k)$$

Ο αλγόριθμος αυτός, όπως γίνεται ξεκάθαρο παραπάνω, εξετάζει σε κάθε επανάληψή του όλα τα δείγματα εισόδου και στην περίπτωση που εντοπίσει κάποιο το οποίο (με βάση το $\mathbf{w}(k)$) ταξινομείται στην λάθος κλάση προχωράει σε ενημέρωση του διανύσματος βαρών προς την κατεύθυνση εκείνη που αυξάνει την πιθανότητα σωστής ταξινόμησης στο μέλλον. Σε περίπτωση που, μετά από K επαναλήψεις, όλα τα δείγματα ταξινομούνται σωστά, ο αλγόριθμος τερματίζει και επιστρέφει το διάνυσμα βαρών $\mathbf{w}(K)$ ως λύση του προβλήματος ταξινόμησης.

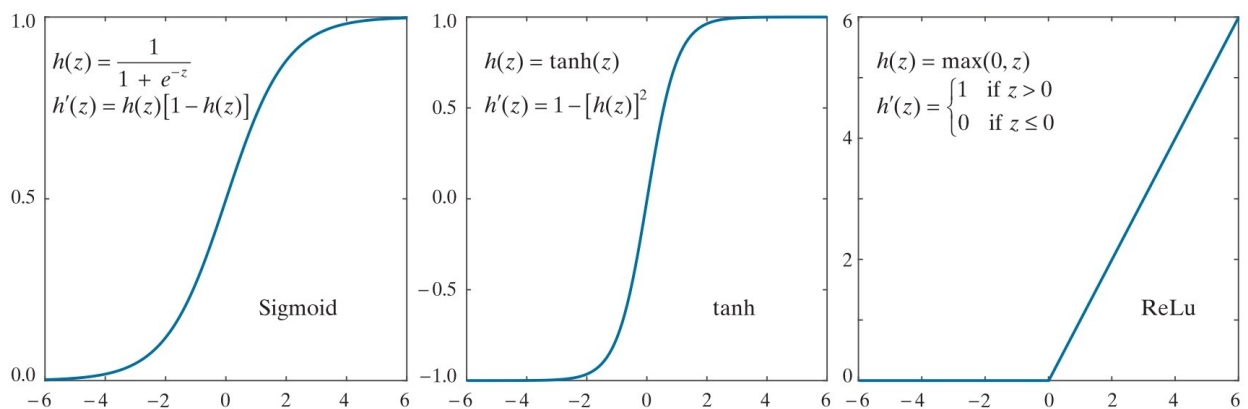
2.4.4.2 Feedforward Neural Networks

Το Perceptron, που παρουσιάστηκε παραπάνω, μπορεί να θεωρηθεί ως η πρώτη και απλούστερη υλοποίηση ενός νευρωνικού δικτύου. Το συγκεκριμένο μαθηματικό μοντέλο, είναι σε θέση να επιλύει, με μεγάλη επιτυχία, προβλήματα που αφορούν σύνολα δεδομένων με

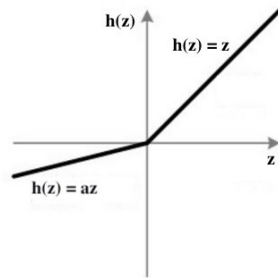
γραμμικά διαχωρίσιμα επιμέρους στοιχεία. Στην πράξη, ωστόσο, τα περισσότερα σύνολα δεδομένων αποτελούνται από στοιχεία τα οποία παρουσιάζουν μη γραμμικές εξαρτήσεις μεταξύ τους, καθιστώντας έτσι το Perceptron ανίκανο να τα διαχειριστεί. Σε μία προσπάθεια αντιμετώπισης του συγκεκριμένου εμποδίου, προτάθηκε ο συνδυασμός πολλαπλών Perceptron για την επίλυση προβλημάτων που αφορούν μη γραμμικά διαχωρίσιμα δεδομένα. Με βάση, λοιπόν, την παραπάνω ιδέα ξεκίνησε η υλοποίηση των σύγχρονων νευρωνικών δικτύων, τα οποία ονομάζονται επίσης και **βαθιά νευρωνικά δίκτυα**.

Τα βαθιά νευρωνικά δίκτυα, αποτελούνται από πολλαπλά επίπεδα (*layers*) νευρώνων που συνδέονται μεταξύ τους, δημιουργώντας την εικόνα ενός ακυκλικού κατευθυνόμενου γραφήματος (DAG). Κάθε επίπεδο του δικτύου, μπορεί να περιέχει διαφορετικό πλήθος νευρώνων, που για δική μας διευκόλυνση θεωρείται πως βρίσκονται στοιβαγμένοι σε μία στήλη. Οι νευρώνες του δικτύου παρουσιάζουν παρόμοια συμπεριφορά με αυτή του Perceptron, υπολογίζουν δηλαδή το σταθμισμένο άθροισμα των εισόδων τους, ωστόσο για τον υπολογισμό της εξόδου τους έχουν αντικαταστήσει την βηματική συνάρτηση, με άλλες πιο ομαλές μη γραμμικές συναρτήσεις ενεργοποίησης. Ο λόγος για τον οποίο αντικαταστάθηκε η βηματική συνάρτηση είναι η έντονη ευαισθησία της, ακόμη και σε απειροστά μικρά θετικά σήματα, γεγονός το οποίο δημιουργεί σημαντικά προβλήματα ευστάθειας στα σύγχρονα νευρωνικά δίκτυα. Επίσης, ένας ακόμη παράγοντας που οδήγησε στην αποφυγή χρήσης της βηματικής ως συνάρτηση ενεργοποίησης για τους νευρώνες, είναι η αδυναμία υπολογισμού της παραγώγου της, γεγονός το οποίο θα καθιστούσε αδύνατη την εφαρμογή του αλγορίθμου backpropagation που χρησιμοποιείται για την εκπαίδευση των δικτύων. Οι γνωστότερες συναρτήσεις που χρησιμοποιούνται σήμερα στα νευρωνικά δίκτυα ως συναρτήσεις ενεργοποίησης είναι εξής [42] :

- **Σιγμοειδής :** $h(z) = \frac{1}{1+e^{-z}}$
- **Υπερβολική εφαπτομένη :** $h(z) = \tanh(z)$
- **Rectified Linear Unit (ReLU) :** $h(z) = \max(0, z)$
- **Leaky Rectified Linear Unit (LeakyReLU) :** $h(z) = \begin{cases} a * z, & \text{if } z < 0 \\ z, & \text{if } z \geq 0 \end{cases}$, where a is a slope parameter for negative values of z

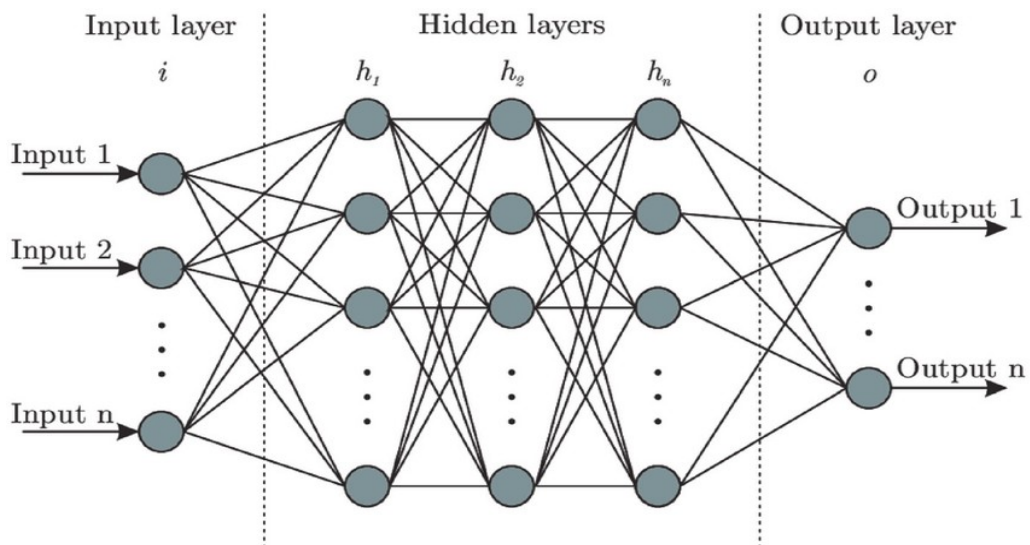


Εικόνα 6: Οι γνωστότερες συναρτήσεις ενεργοποίησης



Εικόνα 7: H
συνάρτηση LeakyReLU

Στην εικόνα 8 φαίνεται η αρχιτεκτονική ενός σύγχρονου νευρωνικού δικτύου. Αναλυτικότερα, όπως σημειώθηκε και προηγουμένως, το δίκτυο οργανώνεται σε επίπεδα και σε κάθε επίπεδο υπάρχουν πολλαπλοί νευρώνες, ή κόμβοι. Στο σημείο αυτό, κρίνεται σκόπιμο να σημειωθεί πως οι νευρώνες σε όλα τα επίπεδα του δικτύου είναι πανομοιότυποι και συμπεριφέρονται όπως περιγράφηκαν παραπάνω, εκτός από αυτούς που βρίσκονται στο πρώτο επίπεδο, ή αλλιώς στο επίπεδο της εισόδου (*input layer*). Κάθε ένας από τους νευρώνες του συγκεκριμένου επιπέδου, αντιστοιχεί σε ένα στοιχείο του διανύσματος που δίνεται ως είσοδος στο δίκτυο και συνεπώς οι εξόδοι, ή ενεργοποιήσεις, των νευρώνων αυτών είναι η τιμή του αντίστοιχου στοιχείου του διανύσματος εισόδου. Συνεχίζοντας την ανάλυση της προαναφερθείσας εικόνας, οι πολλαπλές γραμμές που ενώνουν τους κόμβους του δικτύου υποδηλώνουν ότι οι εξόδοι των νευρώνων ενός επιπέδου τροφοδοτούνται, ως είσοδοι, σε κάθε ένα από τους νευρώνες του επόμενου επιπέδου, σχηματίζοντας έτσι ένα πλήρως συνδεδεμένο δίκτυο. Τέλος, αξίζει να αναφερθεί πως οι τιμές των εξόδων των νευρώνων στα επίπεδα της εισόδου και της εξόδου είτε είναι γνωστές είτε μπορούν να παρατηρηθούν. Ωστόσο, σε όλα τα ενδιάμεσα επίπεδα του νευρωνικού δικτύου, οι τιμές των εξόδων των νευρώνων είναι αδύνατο να παρατηρηθούν και για τον λόγο αυτό, τα ενδιάμεσα αυτά επίπεδα ονομάζονται κρυφά επίπεδα [42].



Εικόνα 8: Γραφική αναπαράσταση ενός Feedforward νευρωνικού δικτύου.

2.4.4.2.1 Η διαδικασία της εκπαίδευσης

Τα βαθιά νευρωνικά δίκτυα μπορούν να περιγραφούν πλήρως από τα βάρη, τις τιμές των bias και τις συναρτήσεις ενεργοποίησης των νευρώνων τους. Επομένως, η εκπαίδευση ενός νευρωνικού δικτύου αφορά την αξιοποίηση ενός συνόλου δεδομένων εκπαίδευσης προκειμένου να καταστεί εφικτή η απόδοση κατάλληλων τιμών, για την επίλυση του εκάστοτε προβλήματος, στις παραμέτρους αυτές. Ο αλγόριθμος που χρησιμοποιείται για την εκπαίδευση των νευρωνικών δικτύων ονομάζεται *backpropagation* και μπορεί να συνοψιστεί στα παρακάτω τέσσερα βήματα :

- Τροφοδότηση του δικτύου με ένα δείγμα εκπαίδευσης.
- Ένα εμπρόσθιο πέρασμα του δείγματος από το δίκτυο και υπολογισμός της μετρικής σφάλματος.
- Ένα οπίσθιο πέρασμα του σφάλματος εξόδου, προκειμένου να υπολογιστούν οι αναγκαίες παράγωγοι για την ενημέρωση των τιμών των παραμέτρων.
- Ενημέρωση των βαρών και τιμών bias των νευρώνων.

Η παραπάνω διαδικασία επαναλαμβάνεται για όλα τα δείγματα στο σύνολο εκπαίδευσης και ολοκληρώνεται μόλις το σφάλμα της εξόδου του δικτύου γίνει αποδεκτά μικρό. Στο σημείο αυτό, αξίζει να σημειωθεί πως δεν υπάρχει κάποια μαθηματική απόδειξη, αντίστοιχη αυτής του Perceptron, ότι η συγκεκριμένη μέθοδος θα συγκλίνει σε κάποια λύση του, υπό μελέτη, προβλήματος, ωστόσο εμπειρικά φαίνεται πως στην πλειονότητα των περιπτώσεων οδηγεί σε εντυπωσιακά αποτελέσματα [42].

Ο σκοπός της εκπαίδευσης ενός νευρωνικού δικτύου είναι η ελαχιστοποίηση της συνάρτησης σφάλματος, η οποία προφανώς διαφέρει ανάλογα με την φύση του προβλήματος που καλείται το δίκτυο να επιλύσει. Προκειμένου, λοιπόν, να υπολογιστούν οι τιμές των παραμέτρων εκείνων που ελαχιστοποιούν την συνάρτηση σφάλματος, γίνεται χρήση της μεθόδου βελτιστοποίησης *gradient descent*. Σύμφωνα με την συγκεκριμένη μέθοδο, η ενημέρωση των βαρών και των biases θα γίνεται πάντοτε προς την κατεύθυνση εκείνη που υποδεικνύουν οι αντίστοιχες μερικές παράγωγοι :

$$w_{ij}(l) = w_{ij}(l) - \eta \frac{\partial E}{\partial w_{ij}(l)}$$

$$b_i(l) = b_i(l) - \eta \frac{\partial E}{\partial b_i(l)}$$

όπου η το learning rate, E η συνάρτηση σφάλματος,

$w_{ij}(l)$ το βάρος του νευρώνα i όταν δέχεται την έξοδο του νευρώνα j στο επίπεδο l

και $b_i(l)$ η τιμή bias του νευρώνα i στο επίπεδο l .

Ο υπολογισμός, ωστόσο, των παραπάνω μερικών παραγώγων της συνάρτησης σφάλματος, απευθείας, για τους νευρώνες των κρυφών επιπέδων δεν είναι εφικτός. Για τον λόγο αυτό, είναι αναγκαία η οπίσθια προώθηση του σφάλματος. Τέλος, δεδομένου ότι η συνάρτηση σφάλματος δεν εξαρτάται άμεσα από τα $w_{ij}(l), b_i(l)$, για τον υπολογισμό των

αντίστοιχων μερικών παραγώγων γίνεται χρήση του κανόνα της αλυσίδας, η ανάλυση του οποίου ξεφεύγει από τα πλαίσια της παρούσας εργασίας και για αυτό παραλείπεται.

Συνοψίζοντας, λοιπόν, τα σύγχρονα νευρωνικά δίκτυα, αποτελούνται από πολλαπλά επίπεδα τεχνητών νευρώνων, οι οποίοι παρουσιάζουν παρόμοια συμπεριφορά με το Perceptron με την διαφορά πως διευρύνουν το φάσμα των πιθανών συναρτήσεων ενεργοποίησης πέρα από την βηματική του Perceptron. Οι έξοδοι των νευρώνων ενός επιπέδου προωθούνται, ως είσοδοι, στους νευρώνες του αμέσως επόμενου επιπέδου του δικτύου και αντιστοιχούν, πρακτικά, σε απλά χαρακτηριστικά της εισόδου του δικτύου. Με αυτό το τρόπο, τα βαθιά νευρωνικά δίκτυα μαθαίνουν να συνδυάζουν, μέσω του αλγορίθμου backpropagation, σταδιακά, απλούστερα χαρακτηριστικά για να συνθέσουν νέα, υψηλότερης σημασιολογικής αξίας χαρακτηριστικά μέχρις ότου η πληροφορία να φτάσει στην έξοδο του δικτύου, η οποία αποτελεί έναν αρκετά σύνθετο μη γραμμικό μετασχηματισμό της εισόδου του δικτύου. Η βαθιά αυτή αρχιτεκτονική των σύγχρονων νευρωνικών δικτύων προσδίδει σε αυτά μία εντυπωσιακά ισχυρή ικανότητα αφαίρεσης και γενίκευσης.

2.4.4.3 Convolutional Neural Networks

Μέχρι το σημείο αυτό, η ανάλυση των σύγχρονων νευρωνικών δικτύων αφορούσε σύνολα δεδομένων των οποίων τα επιμέρους στοιχεία ήταν σε μορφή διανυσμάτων. Όταν τα δεδομένα εισόδου είναι εικόνες, δισδιάστατοι δηλαδή πίνακες, η παραπάνω υπόθεση συνεπάγεται την ανάγκη “χειροκίνητης” εξαγωγής χαρακτηριστικών για την δημιουργία ενός διανύσματος χαρακτηριστικών, το οποίο θα τροφοδοτηθεί, στην συνέχεια, ως είσοδος στο νευρωνικό δίκτυο. Η παραπάνω διαδικασία, ωστόσο, αναιρεί ένα από τα κυριότερα πλεονεκτήματα των βαθιών νευρωνικών δικτύων, την δυνατότητα δηλαδή αυτόματης εξαγωγής χαρακτηριστικών απευθείας από τα δεδομένα εκπαίδευσης. Ένας τρόπος για να γίνει αυτό θα ήταν να μετατρέψουμε τις εικόνες σε διανύσματα απευθείας, οργανώνοντας τα εικονοστοιχεία με βάση έναν γραμμικό δείκτη και στη συνέχεια να θεωρήσουμε πως κάθε στοιχείο (pixel) είναι ένα στοιχείο του διανύσματος. Ωστόσο, αυτή η προσέγγιση δεν λαμβάνει υπόψιν της καμία από τις χωρικές εξαρτήσεις που μπορεί να υπάρχουν μεταξύ των εικονοστοιχείων σε μια εικόνα, όπως παραδείγματος χάριν διατάξεις εικονοστοιχείων σε γωνίες, παρουσία ακμών και άλλα χαρακτηριστικά που μπορεί να βοηθήσουν στη διαφοροποίηση μιας εικόνας από την άλλη [42]. Προκειμένου, να καταστεί δυνατή η χρήση εικόνων απευθείας ως είσοδοι σε νευρωνικά δίκτυα οι LeCun et al. στο [43] πρότειναν μία νέα υλοποίηση νευρωνικών δικτύων, των *συνελικτικών νευρωνικών δικτύων (CNN)*.

Τα συνελικτικά νευρωνικά δίκτυα είναι παρόμοια με τα κοινά νευρωνικά δίκτυα που περιγράφηκαν προηγουμένως όσον αφορά την αρχιτεκτονική. Αποτελούνται από νευρώνες που έχουν βάρη και *biases*, τα οποία ενημερώνονται μέσω της διαδικασίας της εκπαίδευσης του δικτύου. Διαθέτουν, επίσης, τις ίδιες συναρτήσεις απώλειας στο επίπεδο της εξόδου και γενικά εκφράζουν μια συνάρτηση μεταφοράς που δέχεται ως είσοδο εικόνες, αυτούσιες, και παράγει ως έξοδο ορισμένα αποτελέσματα, τα οποία προφανώς σχετίζονται με το εκάστοτε υπό ανάλυση πρόβλημα.

Η κύρια διαφορά μεταξύ των συνελικτικών νευρωνικών δικτύων και των τυπικών νευρωνικών δικτύων, είναι ότι τα CNN κάνουν τη ρητή υπόθεση ότι τα δεδομένα εισόδου είναι αποκλειστικά δισδιάστατοι πίνακες δεδομένων (όπως παραδείγματος χάριν εικόνες), γεγονός που καθιστά δυνατή την κωδικοποίηση ορισμένων ιδιοτήτων στο εσωτερικό της αρχιτεκτονικής, μειώνοντας έτσι κατά πολύ τον αριθμό των παραμέτρων του δικτύου. Επιπλέον, στην περίπτωση των CNN η βασική πράξη που εκτελείται στους νευρώνες είναι η

συνέλιξη με ορισμένους πυρήνες, ή φίλτρα, που περιέχουν τα βάρη. Το αποτέλεσμα της συνέλιξης, στην συνέχεια, περνάει από μια συνάρτηση ενεργοποίησης η οποία εφαρμόζεται τώρα σε κάθε ένα από τα εικονοστοιχεία της παραγόμενης, από την συνέλιξη, εικόνας.

Η ιδέα των συνελικτικών νευρωνικών δικτύων βασίζεται, όπως και γενικότερα τα νευρωνικά δίκτυα, στον τρόπο λειτουργίας του εγκεφάλου και ακριβέστερα στον τρόπο λειτουργίας του οπτικού φλοιού. Συγκεκριμένο τμήμα νευρώνων, σε αυτό το τμήμα του εγκεφάλου, είναι υπεύθυνο για την επεξεργασία ενός περιορισμένου τμήματος του οπτικού πεδίου. Παραδείγματος χάριν, αν στο οπτικό πεδίο βρίσκονται μία ακμή ενός αντικειμένου και μία καμπύλη, διαφορετικοί νευρώνες διεγείρονται για τον εντοπισμό της ακμής, ενώ εντελώς διαφορετικοί νευρώνες πυροδοτούνται για την ανίχνευση της καμπύλης. Τα συνελικτικά δίκτυα χρησιμοποιούν ακριβώς αυτήν την ιδέα για να ανιχνεύσουν μοναδικά χαρακτηριστικά που περιέχονται, για παράδειγμα, σε μια εικόνα ενός αυτοκινήτου ή ενός αεροπλάνου. Ωστόσο, αν και ο ανθρώπινος εγκέφαλος υποσυνείδητα μπορεί να αναγνωρίσει ένα αεροπλάνο ανιχνεύοντας χαρακτηριστικά όπως δύο μεταλλικά φτερά, την ουρά του κ.λπ., τα συνελικτικά δίκτυα προσεγγίζουν την ίδια εργασία ανιχνεύοντας αρχικά χαρακτηριστικά χαμηλότερου επιπέδου, όπως ακμές και καμπύλες, και σταδιακά δημιουργούν πιο αφηρημένες έννοιες όπως αυτή ενός φτερού.

Για την εκτέλεση της παραπάνω λειτουργίας, τα CNN χρησιμοποιούν ένα σύνολο κατηγοριών επιπέδων, από τα οποία περνούν το δείγμα εισόδου. Οι κυριότερες κατηγορίες επιπέδων που χρησιμοποιούνται από τα συνελικτικά νευρωνικά δίκτυα είναι οι εξής :

- **Συνελικτικό επίπεδο** : Αυτό είναι το βασικότερο είδος επιπέδου στο συγκεκριμένο είδος δικτύων. Σε αυτό πραγματοποιείται η διαδικασία της συνέλιξης, μέσω της οποίας γίνεται η εξαγωγή των χαρακτηριστικών μίας εικόνας. Η εικόνες που δημιουργούνται στην έξοδο του συγκεκριμένου επιπέδου ονομάζονται *feature maps*.
- **Επίπεδο ενεργοποίησης** : Στο επίπεδο αυτό γίνεται εφαρμογή της συνάρτησης ενεργοποίησης.
- **Pooling επίπεδο** : Το συγκεκριμένο επίπεδο χρησιμοποιείται για την μείωση των χωρικών διαστάσεων των *feature maps* που έχουν προκύψει μετά το πέρασμα από το επίπεδο ενεργοποίησης.
- **Flattening επίπεδο** : Στο επίπεδο αυτό πραγματοποιείται η μετατροπή ενός *feature map* από δισδιάστατος πίνακας n -διάστατο διάνυσμα. Το επίπεδο αυτό είναι αναγκαίο σε περιπτώσεις εφαρμογών ταξινόμησης αντικειμένων, *regression* κ.α. , καθώς προετοιμάζει τα δεδομένα για να δοθούν ως είσοδος στο πλήρως συνδεδεμένο επίπεδο που ακολουθεί.
- **Πλήρως συνδεδεμένο επίπεδο** : Το επίπεδο αυτό έχει την μορφή που περιγράψαμε στην παραπάνω ενότητα και, όπως προαναφέρθηκε, χρησιμοποιείται για την εκτέλεση εργασιών όπως η ταξινόμηση και η παλινδρόμηση.

Στην συνέχεια, παρατίθεται μία σύντομη περιγραφή και ανάλυση των τριών πρώτων κατηγοριών επιπέδων που αναφέρθηκαν παραπάνω.

Συνελικτικό επίπεδο

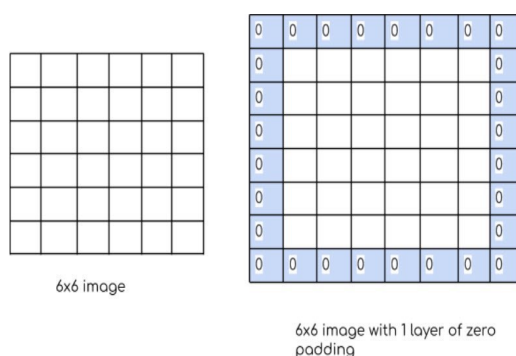
Αυτό είναι το κύριο επίπεδο που χρησιμοποιείται από τα συνελικτικά νευρωνικά δίκτυα, στο οποίο πραγματοποιείται το μεγαλύτερο μέρος της επεξεργασίας των, υπό ανάλυση,

εικόνων. Ο σκοπός του είναι να ανιχνεύσει διακριτά χαρακτηριστικά στην εικόνα. Αυτό γίνεται με την εφαρμογή ενός πυρήνα στην εικόνα που είναι υπεύθυνος για την εξαγωγή ορισμένων χαρακτηριστικών χαμηλότερου επιπέδου. Καθώς τα feature maps προωθούνται σε μεγαλύτερο βάθος του δικτύου, το σημασιολογικό περιεχόμενο των χαρακτηριστικών που ανιχνεύονται γίνεται όλο και υψηλότερο. Για παράδειγμα, ορισμένα φίλτρα, στην αρχή του δικτύου, μπορούν να ανιχνεύσουν ακμές. Ένας πυρήνας δεν είναι τίποτα άλλο παρά ένας πολυδιάστατος πίνακας τιμών $\text{pixel } n \times m \times c$, όπου n είναι το πλάτος, m είναι το ύψος και c το πλήθος των φίλτρων.

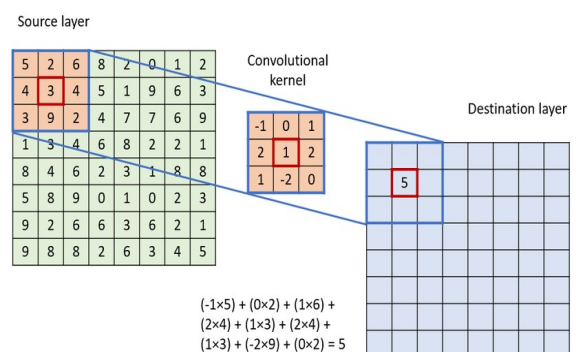
Η διαδικασία της συνέλιξης περιλαμβάνει την εφαρμογή του πυρήνα σε γειτονίες pixel της εικόνας (το μέγεθος της γειτονίας καθορίζεται από το μέγεθος του φίλτρου), μετατοπίζοντάς το κατά ένα, ή περισσότερα εικονοστοιχεία, κατά μήκος κάθε διάστασης της εικόνας (το πλήθος των εικονοστοιχείων που μετατοπίζεται το φίλτρο σε κάθε βήμα είναι γνωστό ως *stride*). Σε κάθε βήμα πραγματοποιείται ο πολλαπλασιασμός μεταξύ των αντίστοιχων τιμών των εικονοστοιχείων της εικόνας και των εικονοστοιχείων του φίλτρου. Τα προκύπτοντα γινόμενα αθροίζονται μεταξύ τους, προκειμένου να δημιουργήσουν την τιμή του νέου εικονοστοιχείου στο feature map που θα δημιουργηθεί. Στο σημείο αυτό, σημειώνεται ότι η έξοδος αυτής της διαδικασίας είναι μια μήτρα μικρότερων διαστάσεων από την αρχική εικόνα, ανάλογα με τον *stride*. Γι' αυτό χρησιμοποιείται, σε ορισμένες εφαρμογές, το *padding*, πριν την εκτέλεση της συνέλιξης, το οποίο προσθέτει εικονοστοιχεία περιμετρικά της εικόνας, ώστε η έξοδος του συνελικτικού επιπέδου να έχει τις ίδιες διαστάσεις με την αρχική εικόνα. Οι τιμές των προστιθέμενων εικονοστοιχείων εξαρτώνται από την μέθοδο *padding* που θα χρησιμοποιήσουμε [42].

Τέλος, αναφέρεται ότι στα σύγχρονα συνελικτικά δίκτυα χρησιμοποιούνται πολλαπλά επίπεδα αυτού του τύπου, επιτρέποντας την εξαγωγή διαφορετικών χαρακτηριστικών τα οποία προωθούνται εντός του δικτύου για την δημιουργία νέων και πιο σύνθετων χαρακτηριστικών.

Παρακάτω παρατίθενται ορισμένες εικόνες, με σκοπό την οπτικοποίηση των διαδικασιών που περιγράφηκαν.



Εικόνα 9: Εφαρμογή zero padding σε 6x6 εικόνα.



Εικόνα 10: Η διαδικασία της συνέλιξης

Επίπεδο ενεργοποίησης

Μετά από κάθε συνελικτικό επίπεδο ακολουθεί συνήθως ένα επίπεδο ενεργοποίησης ή μη γραμμικότητας. Αυτό διασφαλίζει ότι η έξοδος ενός συνελικτικού επιπέδου δεν μπορεί να υπολογιστεί ως γραμμικός συνδυασμός των εξόδων των προηγούμενων επιπέδων, καθώς διαφορετικά το άθροισμα όλων των επιπέδων θα εξακολουθεί να είναι μια γραμμική συνάρτηση. Σε αυτό το επίπεδο εφαρμόζεται μία από τις συναρτήσεις ενεργοποίησης, που

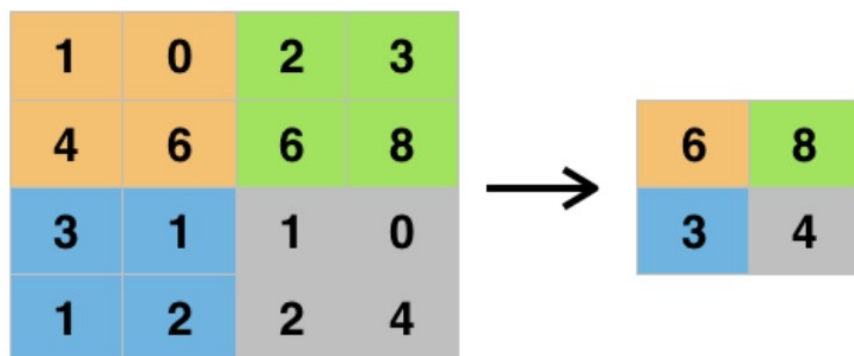
περιγράφηκαν προηγουμένως, σε κάθε ένα από τα στοιχεία της εικόνας-εξόδου του συνελικτικού επιπέδου. Σήμερα, η πιο συχνά χρησιμοποιούμενη συνάρτηση ενεργοποίησης είναι η συνάρτηση Rectified Linear Unit, καθώς εμπειρικά φαίνεται να δίνει τα καλύτερα αποτελέσματα λαμβάνοντας υπόψη την ταχύτητα της εκπαίδευσης σε αντίθεση με άλλες συναρτήσεις όπως η σιγμοειδής και η υπερβολική εφαιπτομένη που χρησιμοποιούνταν στο παρελθόν [42].

Pooling επίπεδο

Αυτό το επίπεδο εντοπίζεται, συνήθως, μεταξύ συνελικτικών επιπέδων του δικτύου και συμβάλει στην μείωση του πλήθους των τιμών των εικονοστοιχείων που περνούν δια μέσου του δικτύου και κατ' επέκταση τον αριθμό των υπερπαραμέτρων του δικτύου (καθώς το τελικό δίκτυο έχει λιγότερα βάρη για υπολογισμό). Ελέγχει επίσης το overfitting του δικτύου. Επιπρόσθετα, η διαδικασία αυτή είναι ιδιαίτερα χρήσιμη, επειδή επιτυγχάνει το translation invariance, που σημαίνει ότι εάν συμβεί μια μικρή μετατόπιση, οι περισσότερες από τις “συγκεντρωμένες” (pooled) τιμές δεν θα αλλάξουν. Αυτή η ιδιότητα είναι σημαντική, ειδικά αν μας ενδιαφέρει περισσότερο αν υπάρχει ένα συγκεκριμένο χαρακτηριστικό στην εικόνα, παρά η ακριβής τοποθεσία του στην εικόνα. Παρόλο που υπάρχουν άλλοι τύποι στρωμάτων συγκέντρωσης (pooling) που λειτουργούν με την ίδια αρχή, όπως average pooling ή L2-norm pooling, το max pooling είναι η πιο δημοφιλής και ευρέως χρησιμοποιούμενη εκδοχή.

Ο τρόπος με τον οποίο λειτουργεί το max pooling, είναι μέσω της εφαρμογής ενός φίλτρου στο feature map από το προηγούμενο επίπεδο, το οποίο απορρίπτει όλα τα εικονοστοιχεία εκτός από αυτό που διαθέτει τη μέγιστη τιμή εντός της γειτονίας που ορίζεται από το φίλτρο, καταλήγοντας σε μια μικρότερη εκδοχή της εικόνας, η οποία ωστόσο περιέχει αρκετές χρήσιμες, για το δίκτυο, πληροφορίες.

Αξίζει να αναφερθεί ότι σε πολλές εφαρμογές τα επίπεδα συγκέντρωσης αντικαθίστανται είτε από πιο συνελικτικά στρώματα με μεγαλύτερο stride (που μειώνουν επίσης το μέγεθος της αναπαράστασης) είτε ακόμη και απορρίπτονται εντελώς εάν η μείωση της απαιτούμενης υπολογιστικής ισχύος είναι επιτακτική ανάγκη.



Εικόνα 11: Παράδειγμα εφαρμογής max pooling.

2.4.4.3.1 Η διαδικασία της εκπαίδευσης

Όπως ήδη αναφέρθηκε, τα CNN μοιράζονται κάποιες ομοιότητες με τα νευρωνικά δίκτυα Feedforward. Υπό αυτή την έννοια, ένα CNN εκπαιδεύεται επίσης με τη μέθοδο backpropagation. Το σφάλμα που χρησιμοποιείται για την εκπαίδευση του δικτύου είναι μπορεί να είναι οποιοδήποτε από τα σφάλματα εξόδου που χρησιμοποιούνται και στα συμβατικά νευρωνικά δίκτυα. Ο κανόνας της αλυσίδας εφαρμόζεται επίσης εδώ, για να

ληφθούν οι μερικοί παράγωγοι του σφάλματος ως προς τα βάρη των πυρήνων και τα biases, αλλά επειδή ο απλός πολλαπλασιασμός των νευρωνικών δικτύων Feedforward αντικαθίσταται από τη συνέλιξη στα CNN, οι εξισώσεις είναι πιο σύνθετες. Είναι σημαντικό να σημειωθεί ότι η οπίσθια διάδοση επηρεάζει μόνο τις παραμέτρους του συνελκτικού επιπέδου.

2.4.5 Εφαρμογές της βαθιάς μάθησης

Το πεδίο της βαθιάς μάθησης στην σημερινή εποχή έχει προσελκύσει το ενδιαφέρον ενός πολύ μεγάλου τμήματος της κοινότητας που ασχολείται με την τεχνητή νοημοσύνη και την μηχανική μάθηση. Αποτέλεσμα της άνθισης του συγκεκριμένου πεδίου είναι εμφάνιση ενός ιδιαίτερα ευρέος φάσματος εφαρμογών, το οποίο εκτείνεται από εφαρμογές που χρησιμοποιούν οι άνθρωποι πλέον ακόμη και στην καθημερινότητά τους έως και σύνθετες εφαρμογές στο πεδίο της ιατρικής και της βιολογίας. Συγκεκριμένα αναφέρονται η περίπτωση της αναγνώρισης ομιλίας, η οποία έχει γίνει αναπόσπαστο κομμάτι της καθημερινότητας μας . Όταν οι άνθρωποι μιλάνε σε ένα έξυπνο τηλέφωνο, η σχεδόν άψογη αναγνώριση του λόγου εκτελείται από ένα νευρωνικό δίκτυο. Αυτό το είδος απόδοσης ήταν ανέφικτο μόλις πριν από λίγα χρόνια. Άλλες εφαρμογές από οι χρησιμοποιούνται καθημερινά, ίσως χωρίς να γίνεται αντιληπτό, είναι τα έξυπνα φίλτρα που μαθαίνουν τις προτιμήσεις των χρηστών, στο διαδίκτυο, για την αναδρομολόγηση ανεπιθύμητης αλληλογραφίας και άλλων ανεπιθύμητων μηνυμάτων από λογαριασμούς email, καθώς και τα συστήματα που διαβάζουν ταχυδρομικούς κώδικες στους φακέλους της συμβατικής αλληλογραφίας. Η βαθιά μάθηση, επιπλέον, επεκτείνεται και πέρα από τις απλές, καθημερινές δραστηριότητες. Τα τελευταία χρόνια έχουν αναπτυχθεί βαθιά νευρωνικά δίκτυα ικανά να χειρίζονται οχήματα (αυτόνομα οχήματα) και ρομπότ, τα οποία βρίσκονται στο πραγματικό ανθρώπινο περιβάλλον και αλληλεπιδρούν με αυτό. Τέλος, ορισμένες από τις λιγότερο γνωστές εφαρμογές περιλαμβάνουν την αυτοματοποιημένη ανακάλυψη νέων φαρμάκων, την πρόβλεψη γονιδιακών μεταλλάξεων στην έρευνα του DNA και την πρόοδο στην κατανόηση της φυσικής γλώσσας [42].

Οι εφαρμογές της βαθιάς μάθησης, ωστόσο, δεν περιορίζονται στους προαναφερθέντες τομείς. Ιδιαίτερα σημαντική είναι η συμβολή των νευρικών δικτύων στην όραση υπολογιστών και γενικότερα στην επεξεργασία εικόνων και βίντεο. Τα τελευταία χρόνια, έχει αναπτυχθεί πληθώρα συνελκτικών νευρωνικών δικτύων με σκοπό την ανάλυση και βελτίωση συμβατικών (RGB) και μη εικόνων. Παρακάτω αναφέρονται, επιγραμματικά, ορισμένα από τα πεδία όπου η χρήση CNN έχει προσφέρει, πραγματικά, εντυπωσιακά αποτελέσματα.

- **Ανίχνευση ασθενειών** : Στις συγκεκριμένες εφαρμογές χρησιμοποιούνται ιατρικές εικόνες, όπως αξονικές και μαγνητικές τομογραφίες, υπερηχογραφικές απεικονίσεις, ακτινογραφίες κ.α. , με σκοπό την εκπαίδευση νευρωνικών δικτύων τα όποια θα είναι σε θέση να ανιχνεύουν την ύπαρξη ή μη κάποιας ασθένειας. Χαρακτηριστικό παράδειγμα αποτελεί η ανίχνευση της νόσου Alzheimer's με ακρίβεια 98.8%.
- **Εκτίμηση δράσης** : Στην συγκεκριμένη κατηγορία, τα συνελκτικά δίκτυα χρησιμοποιούν πολλαπλά καρέ από ένα βίντεο με σκοπό την εξαγωγή της πόζας ενός ανθρώπου και του τρόπου που αυτή μεταβάλλεται χρονικά, προκειμένου να παράξουν μία εκτίμηση της κίνησης που εκτελεί.
- **Παρακολούθηση αντικειμένου.**
- **Εκτίμηση της στάσης του σώματος σε ανθρώπους.**
- **Ανίχνευση και αναγνώριση κειμένου** [44].

Ένα ακόμη σημαντικό πεδίο το οποίο χαίρει της προσφοράς των σύγχρονων βαθιών νευρωνικών δικτύων, είναι αυτό της τηλεπισκόπησης. Συγκεκριμένα, έχουν χρησιμοποιηθεί, με επιτυχία, νευρωνικά δίκτυα με σκοπό την ταξινόμηση των απεικονιζόμενων, σε μία δορυφορική εικόνα, τοπίων σε κατηγορίες, όπως αστικό, αγροτικό ή φυσικό περιβάλλον. Επιπλέον, στις εικόνες που περιέχουν πάνω από ένα είδος τοπίου, οι οποίες είναι και οι πιο συνηθισμένες στο τομέα της τηλεπισκόπησης, τα CNN επιτυγχάνουν να εκτελέσουν την λειτουργία της τμηματοποίησης (segmentation), του διαχωρισμού δηλαδή των επιμέρους τοπίων που απεικονίζονται στην εικόνα και την ένταξή τους σε μία από τις πιθανές κατηγορίες. Μία ακόμη χρήσιμη και αρκετά εντυπωσιακή εφαρμογή των νευρωνικών δικτύων, είναι ο εντοπισμός ζημιών που έχουν προκληθεί σε κτίρια, χρησιμοποιώντας δορυφορικές εικόνες, με ακρίβεια που φτάνει έως και το 83.02% [44]. Τέλος, δεν να γίνεται να παραληφθεί η συμβολή των νευρωνικών δικτύων σε εφαρμογές πάνω στο αντικείμενο που μελετάται στην παρούσα εργασία, στην βελτίωση, δηλαδή, της χωρικής ανάλυσης πολυφασματικών, δορυφορικών εικόνων.

2.4.5.1 Εφαρμογές στο Super Resolution

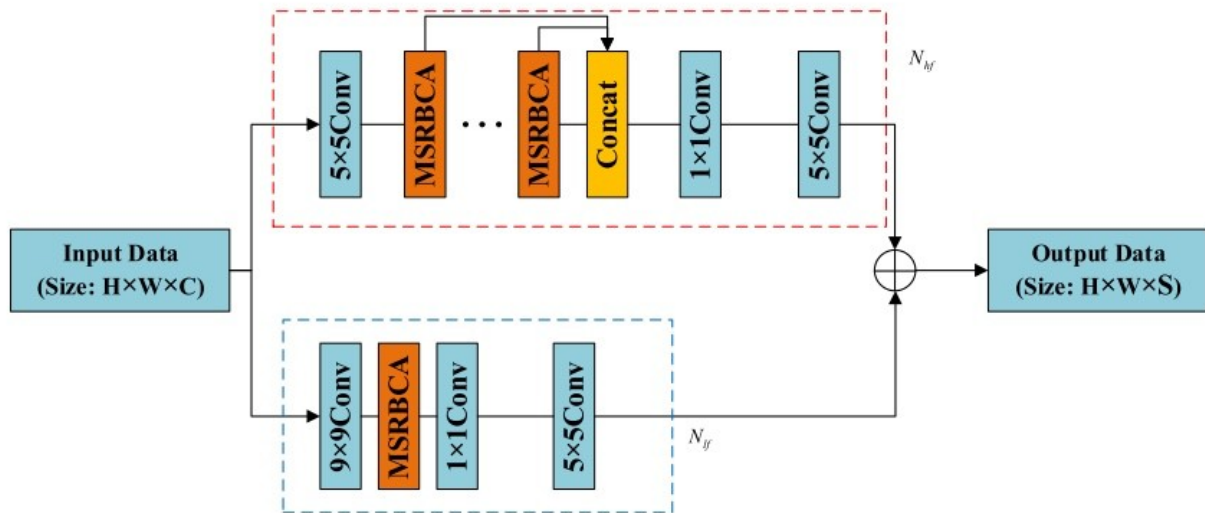
Οι κλασικές μέθοδοι βελτίωσης της χωρικής ανάλυσης των δορυφορικών εικόνων, όπως φάνηκε στο κεφάλαιο 2.3, συχνά αποτυγχάνουν να κατασκευάσουν μια τελική εικόνα πλούσια σε χωρικές και φασματικές πληροφορίες χωρίς να εισάγουν κάποιο είδους παραμόρφωση ή αλλοίωση. Ένας από τους βασικούς λόγους για τους οποίους οι κλασικοί αλγόριθμοι αδυνατούν να κατασκευάσουν μία εικόνα που να διατηρεί όλες τις πληροφορίες τόσο από την PAN όσο και από την MS εικόνα είναι η μεγάλη πολυπλοκότητα των ίδιων των δεδομένων, καθώς και οι έντονες μη γραμμικές συσχετίσεις των pixel από κάθε κανάλι των PAN και MS εικόνων. Τα μοντέλα βαθιάς μάθησης, όπως έχει ήδη αναφερθεί, κατασκευάζονται με πολλαπλά επίπεδα μετασχηματισμών, όπου σε κάθε επίπεδο, η είσοδος φιλτράρεται γραμμικά και πολλαπλά επίπεδα στοιβάζονται για τον σχηματισμό ενός έντονα μη γραμμικού μετασχηματισμού. Τα βαθιά νευρωνικά δίκτυα, ειδικά τα CNN, παρέχουν τη δυνατότητα μιας καλύτερης μοντελοποίησης των μη γραμμικών σχέσεων μεταξύ των pixel των δορυφορικών εικόνων, η οποία θα μπορούσε να βοηθήσει στην παραγωγή καλύτερων αποτελεσμάτων στο πρόβλημα της βελτίωσης της χωρικής ανάλυσης των εικόνων αυτών.

Χάρη, λοιπόν, των δυνατοτήτων και των εντυπωσιακών αποτελεσμάτων των συνελικτικών δικτύων στην επεξεργασία συμβατικών εικόνων (RGB), τα τελευταία χρόνια έχει αναπτυχθεί πληθώρα βαθιών μοντέλων που εκτελούν την διαδικασία του Pansharpening. Παρακάτω επιλέγονται προς παρουσίαση δύο μοντέλα τα οποία προσεγγίζουν με διαφορετικό τρόπο την διαδικασία επεξεργασίας των εικόνων χαμηλής ανάλυσης και σύνθεσης, στην συνέχεια, της ζητούμενης εικόνας υψηλής χωρικής ευκρίνειας. Επίσης, κρίνεται σκόπιμο να αναφερθεί πως και οι δύο μέθοδοι υπάγονται στην ευρύτερη κατηγορία της επιβλεπόμενης μάθησης, προϋποθέτουν δηλαδή την ύπαρξη ενός συνόλου δεδομένων στην μορφή ζευγών που περιγράφηκε στην ενότητα 2.3.

Multi-scale Channel Attention Residual Network

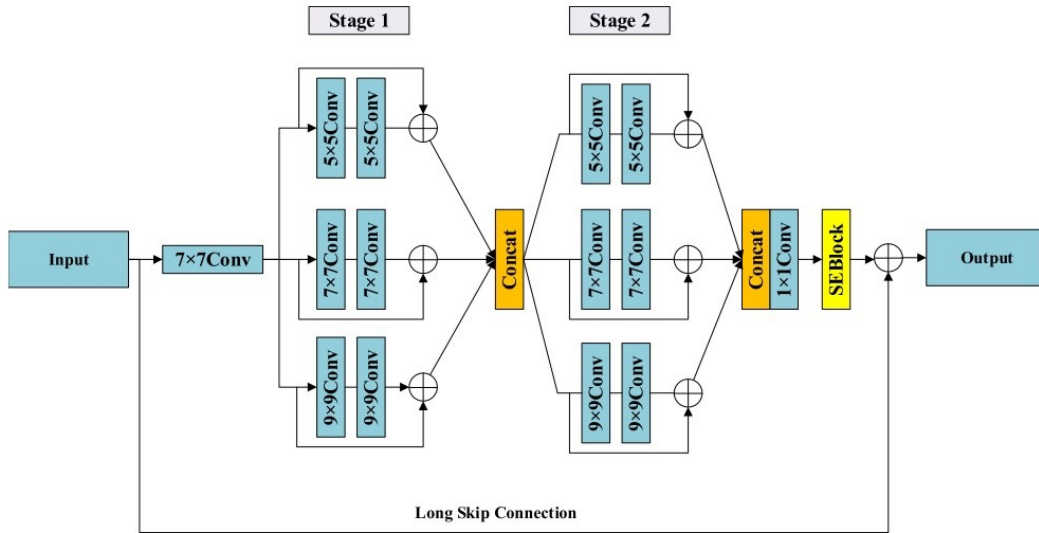
Το συγκεκριμένο μοντέλο προτάθηκε από τους Li et al. Στο [27]. Οι συγγραφείς στο συγκεκριμένο άρθρο πρότειναν ένα βαθύ συνελικτικό νευρωνικό δίκτυο, το οποίο βασίζεται στην αρχιτεκτονική του ResNet [45], περιέχει δηλαδή πέρα των κλασικών συνδέσεων που εντοπίζονται σε ένα νευρωνικό δίκτυο και συνδέσεις μεταξύ μη διαδοχικών επιπέδων (skip connections) του δικτύου. Χρησιμοποιώντας αυτή την τεχνική, οι συγγραφείς είναι σε θέση να

δημιουργήσουν ένα δίκτυο ικανό να εξάγει χαρακτηριστικά των εικόνων από διαφορετικές κλίμακες. Επιπλέον, η αξιοποίηση των skip connections επιτρέπει στο CNN να συγκλίνει ταχύτερα και με μικρότερο σφάλμα στο ζητούμενο αποτέλεσμα. Επιπρόσθετα, στο [27] πραγματοποιείται η παρατήρηση πως τα χαρακτηριστικά που εξάγονται από το δίκτυο στις διαφορετικές κλίμακες, δεν συμβάλουν όλα με την ίδια βαρύτητα στην σύνθεση της τελικής εικόνας. Συνεπώς, οι συγγραφείς προχώρησαν στην εισαγωγή ενός επιπέδου προσοχής (attention block), το οποίο είναι υπεύθυνο για την αξιολόγηση των εξαγόμενων χαρακτηριστικών και, τελικά, τον υπολογισμό ενός συντελεστή για τον προσδιορισμό του βαθμού σημαντικότητας των χαρακτηριστικών στην σύνθεση της εικόνας βελτιωμένης ανάλυσης. Η αρχιτεκτονική του δικτύου που μόλις περιγράφηκε φαίνεται στην εικόνα 12.

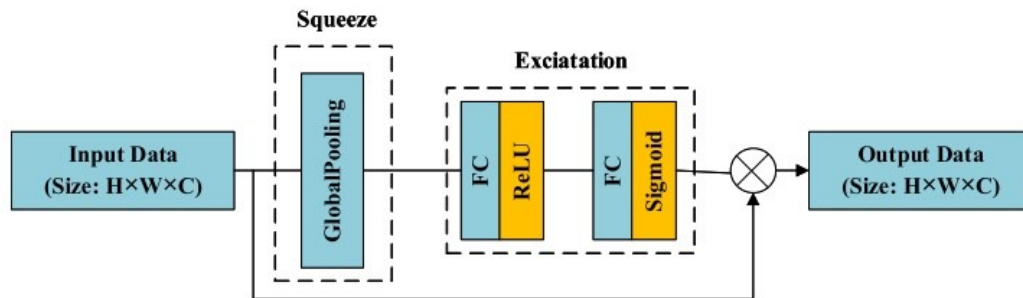


Εικόνα 12: Η αρχιτεκτονική του μοντέλου MSCARN [27].

Από την παραπάνω αρχιτεκτονική φαίνεται πως το συνολικό δίκτυο αποτελείται από δύο επιμέρους υποδίκτυα, ή κλάδους. Ο επάνω κλάδος του δικτύου είναι υπεύθυνος για την εξαγωγή των υψίσυχων χαρακτηριστικών της εικόνας, όπως ακμές, γωνίες, υφές και περιοχές φασματικού ενδιαφέροντος (blob), ενώ ο κάτω κλάδος είναι επιφορτισμένος με την ανίχνευση των χαρακτηριστικών χαμηλότερης συχνότητας, τις δομές δηλαδή που απεικονίζονται. Μετά την εξαγωγή τους, τα προαναφερθέντα χαρακτηριστικά συνδυάζονται γραμμικά, μέσω ενός τελεστή άθροισης, για την σύνθεση της τελικής εικόνας που διαθέτει υψηλή χωρική και φασματική ανάλυση.



Εικόνα 13: Εσωτερική αρχιτεκτονική του MSRBCA block [27].

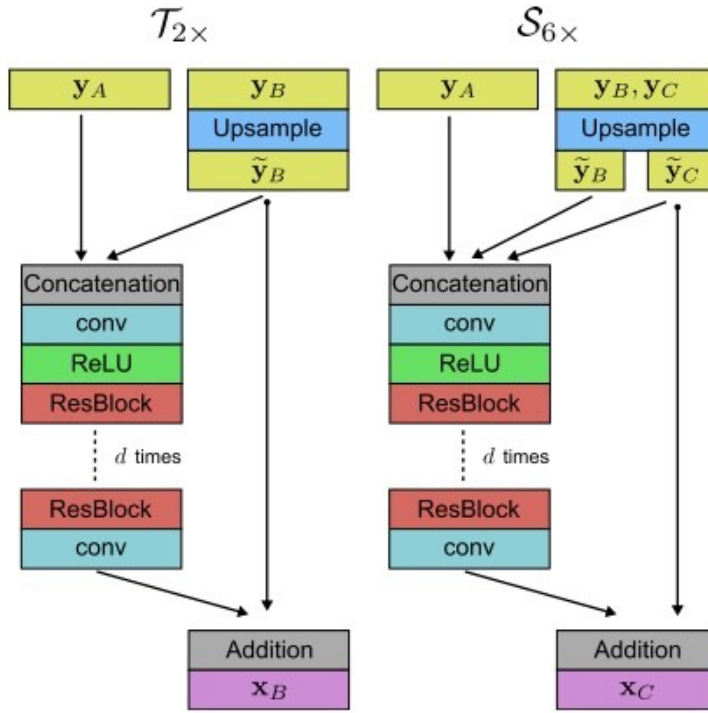


Εικόνα 14: Εσωτερική αρχιτεκτονική του SE block (attention block) [27].

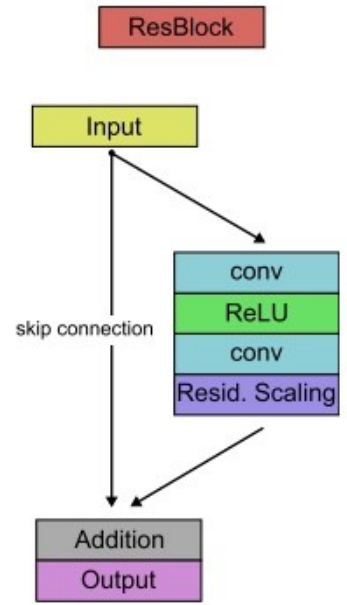
Τέλος, είναι σημαντικό να αναφερθεί πως για την εκπαίδευση του συγκεκριμένου μοντέλου, οι εικόνες PAN και MS υποδειγματοληπτούνται για την κατασκευή των δειγμάτων χαμηλής ανάλυσης, ενώ οι αρχικές MS εικόνες αντιμετωπίζονται ως εικόνες υψηλής ανάλυσης και χρησιμοποιούνται για τον υπολογισμό του σφάλματος εξόδου του δικτύου και για την συνολική του αξιολόγηση τελικά.

DSen2

Μία εναλλακτική προσέγγιση στο πρόβλημα του Pansharpening, μέσω νευρωνικών δικτύων, προτείνεται από τους Lanaras et al. Στο [46]. Οι συγγραφείς στην συγκεκριμένη προσέγγιση επέλεξαν να εργαστούν με εικόνες που έχουν ληφθεί από τον αστερισμό Sentinel-2, καθιστώντας τα αποτελέσματα της μεθόδου αυτής ιδιαίτερα σημαντικά, καθώς δημιουργούν μία βάση σύγκρισης για τα αποτελέσματα που θα παραχθούν από τα υπό μελέτη δίκτυα στην παρούσα εργασία. Η αρχιτεκτονική του δικτύου που προτείνεται στο [46] βασίζεται, επίσης στο ResNet και φαίνεται στην παρακάτω εικόνα.



Εικόνα 15: Η συνολική αρχιτεκτονική του DS2 [46].



Εικόνα 16: Η αρχιτεκτονική ενός ResBlock[46].

Στο συγκεκριμένο άρθρο, προτείνονται δύο νευρωνικά δίκτυα, ένα για την βελτίωση της χωρικής ανάλυσης των MS εικόνων ανάλυσης 20m κατά παράγοντα μεγέθυνσης 2 και ένα για την βελτίωση της χωρικής ανάλυσης των MS εικόνων ανάλυσης 60m κατά παράγοντα μεγέθυνσης 6. Η λογική, ωστόσο, πίσω και από τα δύο μοντέλα παραμένει ακριβώς η ίδια και η μόνη διαφοροποίηση τους εντοπίζεται στα δεδομένα που δίνονται ως είσοδοι κάθε φορά. Αναλυτικότερα, η προσέγγιση αυτή βασίζεται στην υπόθεση του αμετάβλητου του κλίμακας, υποθέτει δηλαδή πως τα αποτελέσματα που παράγονται από το δίκτυο εξαρτώνται από τον συντελεστή μεγέθυνσης και όχι από τις αρχικές και τελικές αναλύσεις των εικόνων. Ως παράδειγμα αναφέρεται η περίπτωση που το δίκτυο είναι εκπαιδευμένο για την αύξηση της ανάλυσης μίας εικόνας κατά συντελεστή 2, έχοντας ως δεδομένα εκπαίδευσης εικόνες ανάλυσης 40m και 20m αντίστοιχα. Το δίκτυο αυτό, σύμφωνα με το αναλλοίωτο της κλίμακας, είναι σε θέση να λάβει ως είσοδο εικόνες ανάλυσης 10m ή 60m και να παράξει ως αποτέλεσμα εικόνες ανάλυσης 5m ή 30m αντίστοιχα.

Βασιζόμενοι, λοιπόν, στην παραδοχή αυτή, οι συγγραφείς πρότειναν ένα pipeline υποδειγματοληψίας των αρχικών εικόνων PAN και MS, το οποίο αποτελείται από ένα αρχικό φιλτράρισμα της εκάστοτε εικόνας με γκαουσιανό πυρήνα τυπικής απόκλισης $\sigma=1/s$, όπου s ο συντελεστής μεγέθυνσης της εικόνας, ακολουθούμενου από ένα φιλτράρισμα μέσου όρου με κυλιόμενο παράθυρο μεγέθους $s \times s$ για την παραγωγή της τελικής υποδειγματοληπτικής εικόνας που θα χρησιμοποιηθεί για την εκπαίδευση του αντίστοιχου δικτύου.

Έχοντας κατασκευάσει το σύνολο εκπαίδευσης, η διαδικασία που ακολουθείται στο εσωτερικό του δικτύου συνοψίζεται στα εξής βήματα :

- Δίνονται ως είσοδοι στο δίκτυο οι υποδειγματοληπτημένες εκδοχές των εικόνων $10m$ (y_a) και $20m$ (y_b) για την περίπτωση του δικτύου με συντελεστή μεγέθυνσης 2. Για την περίπτωση του δικτύου με συντελεστή μεγέθυνσης 6 δίνεται επίσης και η υποδειγματοληπτημένη εκδοχή της εικόνας $60m$ (y_c).
- Πραγματοποίηση υπερδειγματοληψίας των εικόνων των οποίων η αρχική ανάλυση είναι $20m$ για την παραγωγή της εικόνας $\tilde{y}_b$ (και $60m$ ($\tilde{y}_c$) στην περίπτωση του δεύτερου δικτύου).
- Συνένωση των εικόνων $y_a, \tilde{y}_b$ στην περίπτωση του πρώτου δικτύου, ή τις εικόνες $y_a, \tilde{y}_b, \tilde{y}_c$ στην περίπτωση του δεύτερου δικτύου, για την παραγωγή της εικόνας x_0 .
- Πέρασμα της εικόνας x_0 από το πρώτο συνελικτικό επίπεδο και λήψη του x_1 .
- Πέρασμα του x_1 από τα d ResBlocks και λήψη του x_{d+1} .
- Πέρασμα του x_{d+1} από ένα τελευταίο συνελικτικό επίπεδο για την παραγωγή υπολειμματικής εικόνας, κατάλληλων διαστάσεων ανάλογα με το εκάστοτε δίκτυο.
- Μέσω του Skip Connection, αθροίζονται η υπολειμματική εικόνα με την εικόνα $\tilde{y}_b$, στην περίπτωση του πρώτου δικτύου, ή με την εικόνα $\tilde{y}_c$ στην περίπτωση του δεύτερου δικτύου.
- Επιστροφή του τελικού αποτελέσματος.

Από την ανάλυση των διαδικασιών που εκτελούνται στο εσωτερικό των δικτύων που προτείνονται στο [46], μπορεί εύκολα κανείς να οδηγηθεί στο συμπέρασμα πως τα δίκτυα αυτά δεν συνθέτουν, εκ νέου, την εικόνα υψηλής ανάλυσης. Αντιθέτως, στόχος τους είναι να προσεγγίσουν την κατάλληλη υπολειμματική εικόνα, η οποία όταν προστεθεί στην υπερδειγματοληπτημένη εκδοχή της εικόνας εισόδου χαμηλής ανάλυσης θα διορθώσει τις παραμορφώσεις που, όπως προαναφέρθηκε, εισάγει η απλοϊκή μέθοδος παρεμβολής. Με αυτό το τρόπο τα συγκεκριμένα δίκτυα επιτυγχάνουν να διατηρήσουν περισσότερα ραδιομετρικά χαρακτηριστικά των εικόνων MS.

3. Βελτίωση της χωρικής ανάλυσης δορυφορικών εικόνων μέσω μη επιβλεπόμενης μάθησης

Στο τέλος του προηγούμενου κεφαλαίου έγινε αναφορά σε ορισμένες περιπτώσεις αξιοποίησης βαθιών συνελικτικών δικτύων για την βελτίωση της χωρικής ανάλυσης δορυφορικών εικόνων. Οι δύο αυτές προσεγγίσεις βασίζονται στην προϋπόθεση της ύπαρξης ενός, αρκετά μεγάλου, συνόλου δεδομένων υπό την μορφή ζευγών που περιγράφηκε στην ενότητα 2.3 και συνεπώς υπάγονται στην κατηγορία της επιβλεπόμενης μάθησης. Στην πράξη, ωστόσο, πολλές φορές δεν υπάρχουν διαθέσιμα σύνολα δεδομένων με ζεύγη δορυφορικών εικόνων υψηλής-χαμηλής ανάλυσης, ενώ ταυτόχρονα η κατασκευή ενός τέτοιου συνόλου δεδομένων είναι ιδιαίτερα περίπλοκη διαδικασία, διότι πέρα από την προφανή δυσκολία της εύρεσης δορυφορικών εικόνων υψηλής ευκρίνειας πρέπει να ληφθεί υπόψη και η ευθυγράμμιση των εικόνων έτσι ώστε το, υπό εξέταση, τοπίο να απεικονίζεται και στις δύο εικόνες (που αποτελούν το ζεύγος) με τον ίδιο προσανατολισμό και από παρόμοια γωνία

λήψης. Εξαιτίας, λοιπόν, των παραπάνω δυσκολιών, έχουν ξεκινήσει, τα τελευταία χρόνια, να προτείνονται αλγόριθμοι μη επιβλεπόμενης μάθησης με σκοπό την επίλυση του προβλήματος του Super Resolution στις δορυφορικές εικόνες.

3.1. Τα μοντέλα

Στο συγκεκριμένο κεφάλαιο ακολουθεί μία αναλυτική παρουσίαση των μοντέλων βαθιάς, μη επιβλεπόμενης, μάθησης που χρησιμοποιήθηκαν κατά την εκπόνηση της παρούσας εργασίας, καθώς και των τροποποιήσεων που χρειάστηκαν να γίνουν ανά μοντέλο προκειμένου αυτά να είναι σε θέση να διαχειριστούν δορυφορικές εικόνες. Στο σημείο αυτό, αξίζει να αναφερθεί πως και τα τρία μοντέλα, που παρουσιάζονται παρακάτω, παρά τις διαφορές που επιδεικνύουν ως προς τον τρόπο επίλυσης του προβλήματος της βελτίωσης της χωρικής ανάλυσης των δορυφορικών εικόνων, μοιράζονται ως κοινό χαρακτηριστικό το γεγονός ότι για την δημιουργία του τελικού αποτελέσματος (της εικόνας υψηλής χωρικής ανάλυσης) χρησιμοποιείται αποκλειστικά και μόνο η, υπό εξέταση, εικόνα χαμηλής ευκρίνειας.

3.1.1 Deep Image Prior

Στο [47] εξετάζεται, από τους Ulyanov et al. , η ικανότητα συγκεκριμένων αρχιτεκτονικών συνελκτικών δικτύων να μοντελοποιούν και να αποθηκεύουν στις εσωτερικές τους παραμέτρους στατιστικά χαρακτηριστικά των υπό εξέταση εικόνων, χωρίς την ανάγκη προηγούμενης εκπαίδευσης τους πάνω σε κάποιο σύνολο δεδομένων. Αναλυτικότερα, οι Ulyanov et al. πρότειναν πως ορισμένες στατιστικές εξαρτήσεις μεταξύ των pixel μίας εικόνας, που απαιτούνται για την επίλυση προβλημάτων ανακατασκευής και αποκατάστασης όπως είναι και το Super Resolution, μπορούν να ανιχνευθούν από βαθιά δίκτυα γεννητόρων (deep generator networks) και στην συνέχεια να αποθηκευτούν στις εσωτερικές τους παραμέτρους. Με βάση την συγκεκριμένη προσέγγιση, μπορεί κανείς να θεωρήσει πως το πρόβλημα του Super Resolution εξαρτάται αποκλειστικά και μόνο από την εικόνα χαμηλής χωρικής ανάλυσης και την αρχιτεκτονική του συνελκτικού δικτύου που θα αναλάβει την διαδικασία σύνθεσης της εικόνας υψηλής ευκρίνειας, με τα βάρη και τα biases του δικτύου να αποτελούν μία παραμετροποίηση του προβλήματος ανακατασκευής.

Ένα βαθύ δίκτυο γεννήτορα μπορεί να μοντελοποιηθεί ως μία παραμετρική συνάρτηση $x = f_{\theta}(z)$ η οποία αντιστοιχίζει ένα διάνυσμα z στην εικόνα x . Για το πρόβλημα του Super Resolution γίνεται η θεώρηση πως η έξοδος του νευρωνικού δικτύου είναι της μορφής $x \in \mathbb{R}^{C \times H \times W}$ και η είσοδος του δικτύου είναι ένας σταθερός τυχαίος τανυστής $z \in \mathbb{R}^{C \times H \times W}$, όπου C, W, H είναι το πλήθος των καναλιών, το πλάτος και το ύψος της εικόνας και του τανυστή, ενώ οι παράμετροι θ αποτελούν το σύνολο των βαρών και των biases του συνελκτικού δικτύου. Επιπλέον, σημειώνεται πως το δίκτυο, που χρησιμοποιείται σε κάθε περίπτωση, έχει σταθερή εσωτερική δομή και εκτελεί διαδοχικά φιλτραρίσματα στον τανυστή εισόδου, όπως γραμμική συνέλιξη με τους πυρήνες των νευρώνων, υπερδειγματοληψία, pooling και εφαρμογή των μη γραμμικών συναρτήσεων ενεργοποίησης.

Στο σημείο αυτό, κρίνεται σκόπιμο να αναφερθεί πως με την συγκεκριμένη προσέγγιση οι πρότερες γνώσεις (priors) που θα αποθηκευτούν στα βάρη του δικτύου δεν είναι δυνατόν να αφορούν συγκεκριμένες έννοιες, όπως παραδείγματος χάριν η ύπαρξη ενός αντικειμένου στην εικόνα ή ενός σημείου ενδιαφέροντος (blob). Αντιθέτως, το δίκτυο χάρη στη συνεχή εκτέλεση συνελίξεων είναι σε θέση να ανιχνεύει τις εξαρτήσεις μεταξύ pixel μίας γειτονιάς σε πολλαπλές κλίμακες και στην συνέχεια, με βάση τις εξαρτήσεις αυτές, να μοντελοποιήσει την υπό συνθήκη κατανομή της πιθανότητας $p(x|x_o)$, όπου x, x_o η ζητούμενη εικόνα και η

χαμηλότερης ανάλυση εκδοχή της αντίστοιχα. Η κατανομή αυτή πιθανότητας, χρησιμοποιείται γενικά σε προβλήματα ανακατασκευής και αποκατάστασης εικόνων, επιβεβαιώνοντας έτσι την υπόθεση των αρθρογράφων περί μη αναγκαιότητας εκπαίδευσης των δικτύων, σε κάποιο εξωτερικό σύνολο δεδομένων, για την επίλυση προβλημάτων όπως το Super Resolution.

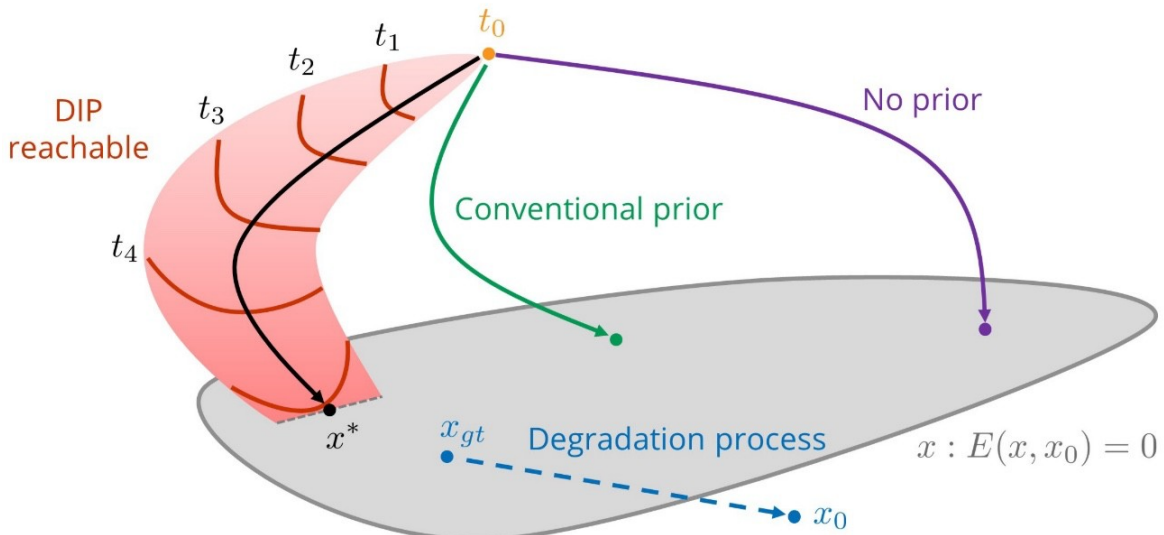
Το πρόβλημα του Super Resolution, όπως και πολλαπλά άλλα προβλήματα σύνθεσης και ανακατασκευής εικόνων, μπορεί να λυθεί μέσω μίας επαναληπτικής διαδικασίας ελαχιστοποίησης της ενέργειας σφάλματος της εξόδου του νευρωνικού δικτύου, μπορεί δηλαδή να περιγραφεί ως :

$$x^* = \operatorname{argmin}_x E(x; x_o) + R(x) \quad (1)$$

,όπου x^* η βέλτιστη λύση ,

$E(x; x_o)$ η ενέργεια σφάλματος της εξόδου,

$R(x)$ μία συνάρτηση κανονικοποίησης.



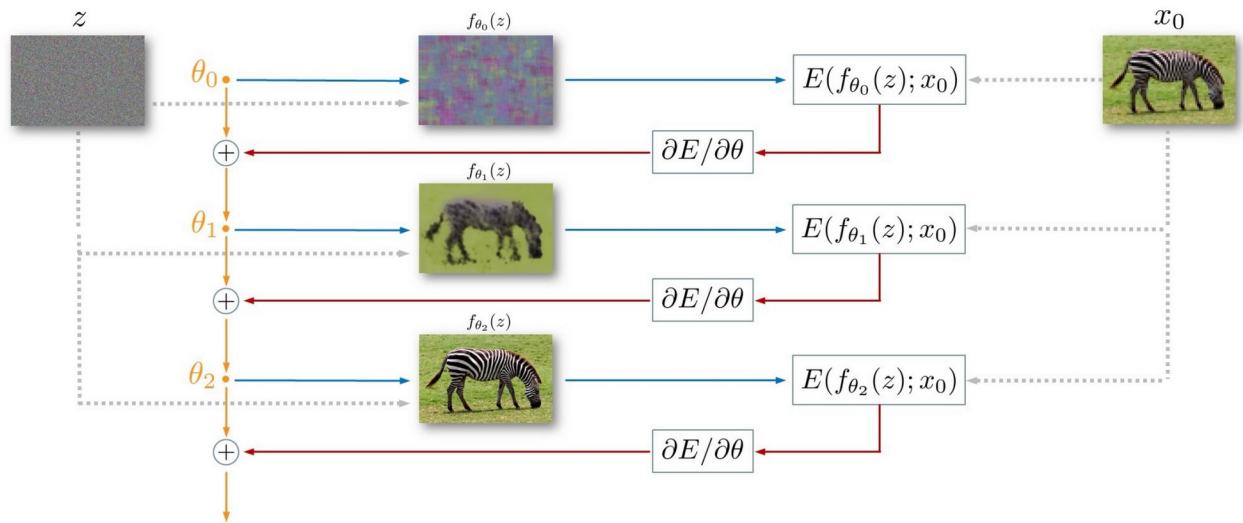
Εικόνα 17: Σχήματα ελαχιστοποίησης της ενέργειας σφάλματος της εξόδου για διαφορετικές συναρτήσεις κανονικοποίησης [47].

Στην παρούσα εργασία, η συνάρτηση $R(x)$ τίθεται στην σταθερή τιμή μηδέν, καθώς σκοπός είναι η αξιολόγηση των δυνατοτήτων του μοντέλου να ανιχνεύσει τις απαραίτητες εξαρτήσεις μεταξύ των pixel της υπό μελέτη εικόνας προκειμένου να παράξει το επιθυμητό αποτέλεσμα, χωρίς την εξωτερική παρέμβαση μίας συνάρτησης κανονικοποίησης (όπως παραδείγματος χάριν είναι το Total Variation Loss). Επιπλέον, χρησιμοποιώντας την παραμετροποίηση $x = f_\theta(z)$, γίνεται εύκολα κατανοητό πως το πρόβλημα της βελτίωσης της χωρικής ανάλυσης των δορυφορικών εικόνων μπορεί να μετασχηματιστεί σε ένα πρόβλημα προσδιορισμού των βέλτιστων τιμών των εσωτερικών παραμέτρων θ του νευρωνικού δικτύου. Συνεπώς, η (1) μπορεί να γραφεί στην μορφή :

$$\theta^* = \operatorname{argmin}_\theta E(f_\theta(z); x_o) \quad (2)$$

Για τον υπολογισμό του θ^* , ακολουθείται ένα επαναληπτικό σχήμα ελαχιστοποίησης του $E(f_\theta(z); x_o)$, όπως αυτό του gradient descent, ξεκινώντας από μία τυχαία

αρχικοποίηση των εσωτερικών παραμέτρων του δικτύου. Μία οπτικοποίηση της προαναφερθείσας διαδικασίας φαίνεται στο παρακάτω σχήμα.

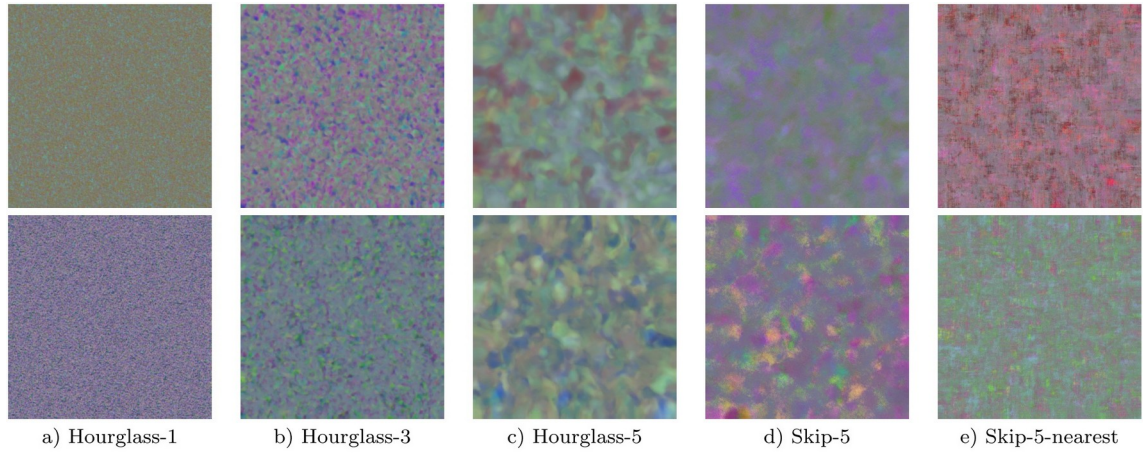


Εικόνα 18: Υπολογισμός του με χρήση της μεθόδου gradient descent [47].

Η τελική λύση στο πρόβλημα του Super Resolution δίνεται από την σχέση $x^* = f_{\theta^*}(z)$, οδηγώντας στο συμπέρασμα πως η αρχιτεκτονική του δικτύου που εκτελεί την ζητούμενη εργασία διαδραματίζει καθοριστικό ρόλο στην ποιότητα του τελικού αποτελέσματος. Στο σημείο αυτό αξίζει να αναφερθεί πως η προκύπτουσα πρότερη γνώση που αποθηκεύεται στις παραμέτρους του εκάστοτε δικτύου μπορεί να ερμηνευθεί και ως η προβολή στο χώρο, μειωμένων διαστάσεων, όλων των εικόνων που μπορούν να παραχθούν από το συνελκτικό δίκτυο με παραμέτρους θ . Συνεπώς, η επιλογή της αρχιτεκτονικής του μοντέλου είναι αυτή που θα καθορίσει τον τρόπο αναζήτησης του τελικού αποτελέσματος στο προαναφερθέντα χώρο.

Η παραμετροποίηση της εξίσωσης (2), αν και δεν προσδιορίζει μία συνάρτηση πυκνότητας πιθανότητας, δύναται να δειγματοληπτηθεί (κατά ένα πιο αφηρημένο τρόπο) μέσω της τυχαίας επιλογής τιμών για τις παραμέτρου θ και εξετάζοντας τα αποτελέσματα $f_{\theta}(z)$ που παράγονται από το νευρωνικό δίκτυο. Με αυτό τον τρόπο, καθίσταται εφικτή η οπτικοποίηση του σημείου έναρξης και των ενδιάμεσων βημάτων της διαδικασίας ελαχιστοποίησης της ενέργειας σφάλματος της εξόδου.

Στην παρακάτω εικόνα παρουσιάζονται ορισμένα δείγματα από τα priors που δημιουργούνται από διαφορετικές αρχιτεκτονικές δικτύων τύπου κλειψύδρας, όταν ως είσοδο δεχτούν τυχαίο ομοιόμορφο θόρυβο.



Εικόνα 19: Δείγματα *priors* που δημιουργούνται από διαφορετικές αρχιτεκτονικές νευρωνικών δικτύων [47].

Στην εικόνα 19, κάθε στήλη αντιστοιχεί στις εξόδους του εκάστοτε συνελκτικού δικτύου, όταν αυτό δέχεται ως είσοδο τον ίδιο τανυστή θορύβου z για δύο διαφορετικές αποδόσεις τιμών στις παραμέτρους θ . Οι αρχιτεκτονικές, τα αποτελέσματα των οποίων επιλέχθηκαν να οπτικοποιηθούν, είναι α) μία αρχιτεκτονική κλεψύδρας με ένα επίπεδο υποδειγματοληψίας και ένα επίπεδο διγραμμικής παρεμβολής, β) μία βαθύτερη αρχιτεκτονική κλεψύδρας με τρία επίπεδα υποδειγματοληψίας και τρία επίπεδα διγραμμικής παρεμβολής, γ) μια ακόμη βαθύτερη αρχιτεκτονική κλεψύδρας με πέντε και πέντε επίπεδα αντίστοιχα, δ) μία αρχιτεκτονική ίδια με την (γ) στην οποία έχουν προστεθεί και skip connections μεταξύ των αντίστοιχων επιπέδων υποδειγματοληψίας και παρεμβολής, ε) μία αρχιτεκτονική όμοια με αυτή που περιγράφηκε στο (δ) μόνο που στα αντίστοιχα επίπεδα εκτελείται παρεμβολή κοντινότερου γείτονα αντί διγραμμικής. Από την εικόνα 19 φαίνεται πως τα δείγματα παρουσιάζουν χωρικές δομές, η κλίμακα των οποίων εξαρτάται άμεσα από το βάθος (πλήθος επιπέδων) του δικτύου. Επιπρόσθετα, αξίζει να σημειωθεί πως η αξιοποίηση των skip connections, στα παραπάνω μοντέλα, έχει ως αποτέλεσμα την μεταφορά χωρικών δομών από διαφορετικές κλίμακες στο τελικό αποτέλεσμα που παράγεται από το μοντέλο, χαρακτηριστικό το οποίο εντοπίζεται και στις φυσικές και δορυφορικές εικόνες, καθιστώντας έτσι την χρήση συνελκτικών δικτύων αυτής της αρχιτεκτονικής προτιμητέα σε εφαρμογές Super Resolution.

Μέχρι στιγμής η ανάλυση της μεθοδολογίας που προτείνεται στο [47] διατηρούσε γενικό χαρακτήρα με δυνατότητα εφαρμογής σε διάφορα προβλήματα που σχετίζονται με την ανακατασκευή και την αποκατάσταση εικόνων. Για το πρόβλημα του Super Resolution που μελετάται στην παρούσα εργασία, κατά το οποίο ζητείται η αύξηση της χωρικής ανάλυσης της εικόνας κατά παράγοντα t , πρέπει να οριστεί μία συνάρτηση ενέργειας σφάλματος της εξόδου του μοντέλου ειδική για το πρόβλημα αυτό. Η συνάρτηση, λοιπόν, που επιλέγεται είναι η εξής :

$$E(x; x_o) = \|d(x) - x_o\|^2$$

, όπου ο τελεστής $d(\cdot): \mathbb{R}^{C \times tH \times tW} \rightarrow \mathbb{R}^{C \times H \times W}$ είναι τελεστής υποδειγματοληψίας κατά παράγοντα t ,

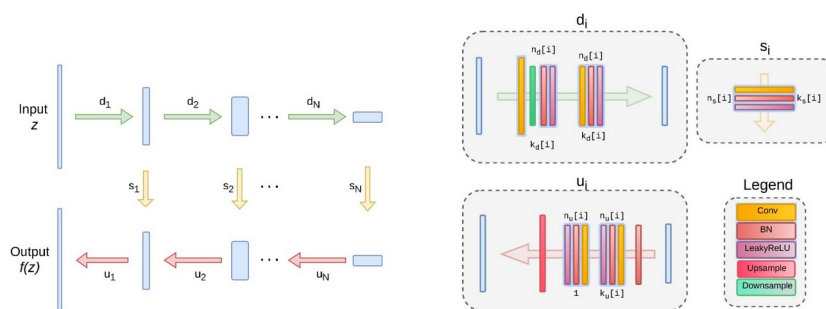
x_o η εικόνα χαμηλής χωρικής ανάλυσης,

$x = f_\theta(z)$ η έξοδος του δικτύου.

Επομένως, χρησιμοποιώντας την συγκεκριμένη συνάρτηση ενέργειας, το πρόβλημα μετασχηματίζεται στην εύρεση της εικόνας υψηλής ανάλυσης x , η οποία όταν υποδειγματοληπτηθεί είναι ίδια με την γνωστή εικόνα χαμηλής ανάλυσης x_0 . Ωστόσο, το πρόβλημα του Super Resolution, όπως και τα υπόλοιπα προβλήματα της ανακατασκευής εικόνων, δεν είναι καλά ορισμένο (ill-posed), υπό την αυστηρή μαθηματική έννοια, καθώς υπάρχει απειρία εικόνων x που με υποδειγματοληψία παράγουν την x_0 . Εξαιτίας, λοιπόν, του ill-posed χαρακτήρα του συγκεκριμένου προβλήματος, είναι αναγκαία η χρήση των εσωτερικών παραμέτρων του δικτύου ως παράγοντα κανονικοποίησης προκειμένου να περιοριστεί η απειρία των πιθανών λύσεων σε ένα πεπερασμένο σύνολο εικόνων που βρίσκονται το κοντινότερο δυνατό στην πραγματική εικόνα υψηλής ευκρίνειας, όπως δηλαδή φαίνεται στην εικόνα 17.

3.1.1.1 Αρχιτεκτονική και βελτιστοποίηση του δικτύου

Στα πλαίσια της συγκεκριμένης εργασίας, επιλέχθηκε να χρησιμοποιηθεί το προτεινόμενο, από το συγγραφείς του άρθρου, δίκτυο για την εκτέλεση της διαδικασίας του Super Resolution των δορυφορικών εικόνων. Το επιλεγθέν μοντέλο ακολουθεί την αρχιτεκτονική κλειψύδρας (U-Net) με skip connections, όπως ήταν αναμενόμενο από τις παρατηρήσεις που προηγήθηκαν. Αναλυτικότερα περιέχει πέντε στρώματα υποδειγματοληψίας, κάθε ένα από τα οποία παράγει 128 feature maps με χωρικές διαστάσεις ίσες με το μισό των feature maps που δέχονται ως είσοδο, και πέντε στρώματα υπερδειγματοληψίας μέσω διγραμμικής παρεμβολής. Επιπλέον, σημειώνεται πως τα feature maps που προωθούνται σε επόμενα επίπεδα του δικτύου, κατά την διέλευσή τους από τα Skip Connection εφαρμόζεται σε αυτά Batch Normalization, μία γραμμική συνέλιξη με stride ίσο με 1 και μία μη γραμμική ενεργοποίηση με την συνάρτηση LeakyReLU. Στο σημείο αυτό σημειώνεται πως η συνάρτηση ενεργοποίησης που χρησιμοποιείται σε οποιοδήποτε κρυφό επίπεδο του δικτύου είναι η LeakyReLU. Η έξοδος του δικτύου σχηματίζεται από ένα τελικό συνελικτικό επίπεδο, το οποίο δέχεται ως είσοδο τα feature maps που έχουν δημιουργηθεί κατά την διέλευση της εισόδου από τα κρυφά επίπεδα του δικτύου και, χρησιμοποιώντας ως συνάρτηση ενεργοποίησης την σιγμοειδή που περιορίζει τις τιμές των pixel μεταξύ του [0,1], παράγει ως έξοδο την εικόνα υψηλής ανάλυσης. Τέλος κρίνεται σκόπιμο να αναφερθεί πως για την παραγωγή του τελικού αποτελέσματος, στο νευρωνικό δίκτυο δίνεται ως είσοδος ένας τανυστής $z \in \mathbb{R}^{32 \times tH \times tW}$, όπου t ο παράγοντας μεγέθυνσης της εικόνας και H, W οι χωρικές διαστάσεις της εικόνας χαμηλής ευκρίνειας.



Εικόνα 20: Η αρχιτεκτονική του Deep Image Prior (DIP) δικτύου.

Τέλος, αξίζει να αναφερθεί πως για την εκπαίδευση του παραπάνω μοντέλου χρησιμοποιήθηκε ο βελτιστοποιητής (optimizer) ADAM, του οποίου το learning rate είχε τεθεί σε σταθερή τιμή ίση με 0.001.

3.1.1.2 Τροποποιήσεις και συνεισφορά

Η μεθοδολογία και το μοντέλο που παρουσιάστηκαν στις προηγούμενες ενότητες, προτάθηκαν αρχικά για συμβατικές RGB εικόνες, οι οποίες αποτελούνται από τρία χρωματικά κανάλια. Ωστόσο, οι δορυφορικές εικόνες της αποστολής Sentinel-2 που εξετάζονται στην παρούσα εργασία, όπως έχει αναφερθεί στην ενότητα 2.2.3.3, δύναται να αποτελούνται είτε από τέσσερα κανάλια (εικόνες με 10m GSD ανάλυση), είτε από έξι κανάλια (εικόνες με 20m GSD ανάλυση) , είτε από δύο κανάλια (εικόνες με 60m GSD ανάλυση). Επομένως, γίνεται άμεσα κατανοητό πως προκειμένου να λειτουργήσει το δίκτυο με τις συγκεκριμένες εικόνες είναι αναγκαία η αλλαγή του συνελικτικού επιπέδου που βρίσκεται στο τέλος του νευρωνικού δικτύου, με τρόπο τέτοιο ώστε να παράγει κάθε φορά εικόνες με τόσα κανάλια όσα και η εικόνα χαμηλής ανάλυσης, η οποία χρησιμοποιείται στην συνέχεια για τον υπολογισμό της ενέργειας σφάλματος και κατ' επέκταση για την προσαρμογή των παραμέτρων του δικτύου. Επιπλέον, επειδή το δίκτυο ήταν σχεδιασμένο για την επεξεργασία RGB εικόνων, είχε επιλεγεί από τους Ulyanov et al. να χρησιμοποιούνται εικόνες με 8bit. Η επιλογή αυτή, για ακόμη μία φορά αποτελούσε εμπόδιο στην επεξεργασία των δορυφορικών εικόνων, των οποίων ο τύπος δεδομένων είναι int16, καθώς αυτές χρειάζεται να αποθηκεύουν περισσότερες πληροφορίες από τις συμβατικές εικόνες. Για την αντιμετώπιση του συγκεκριμένου προβλήματος, αποφασίστηκε το νευρωνικό δίκτυο να αναμένει ως είσοδο ένα τανυστή του οποίου ο τύπος δεδομένων θα είναι float32, προκειμένου το τελικό αποτέλεσμα που παράγεται να είναι επίσης σε μορφή float32 και συνεπώς να μην υπάρχει καμία απώλεια πληροφοριών εξαιτίας της μετατροπής σε ένα τύπο δεδομένων χαμηλότερης ακρίβειας (όπως είναι το uint8).

Μία ακόμη τροποποίηση που χρειάστηκε να γίνει δεν αφορά άμεσα το δίκτυο, παρόλα αυτά συνέβαλε σημαντικά στην βελτίωση της ποιότητας των αποτελεσμάτων και για αυτό κρίνεται αναγκαίο να αναφερθεί. Τα ιστογράμματα των δορυφορικών εικόνων, συχνά, παρουσιάζουν μία μεγάλη συγκέντρωση γύρω από τις χαμηλότερες τιμές των pixel. Αντίθετα, όπως προαναφέρθηκε στην ενότητα 3.2.1.1, η είσοδος του συνελικτικού δικτύου είναι ομοιόμορφος θόρυβος, γεγονός που οδηγεί στο συμπέρασμα πως το ιστόγραμμα του είναι ομοιόμορφα κατανομημένο σε όλες τις τιμές στο διάστημα $[0,1]$. Η μεγάλη αυτή διαφορά στα ιστογράμματα των δύο “εικόνων” που εμπλέκονται στην διαδικασία προσαρμογής των βαρών του δικτύου αποδείχθηκε, εμπειρικά, πως οδηγούσε στην δημιουργία μαύρων στιγμάτων στα τελικά αποτελέσματα. Για την αντιμετώπιση του συγκεκριμένου προβλήματος, επιλέχθηκε να εφαρμόζεται μια προεπεξεργασία στις δορυφορικές εικόνες πριν δοθούν στο συνελικτικό δίκτυο. Αναλυτικότερα, εφαρμόζεται “τέντωμα” στο ιστόγραμμα κάθε καναλιού της υπό μελέτη δορυφορικής εικόνας (histogram stretching) , προκειμένου αυτά να επεκταθούν και να καταλαμβάνουν όλες τιμές στο διάστημα $[0,1]$.

Στο σημείο αυτό, κρίνεται σκόπιμο να αναφερθεί πως κατά την εκτέλεση των πειραμάτων με τις εικόνες χωρικής ανάλυσης 10m GSD, αλλά με τις εικόνες 20m GSD για συντελεστές μεγέθυνσης μεγαλύτερους του δύο, αντιμετωπίστηκαν προβλήματα σχετικά με το hardware του συστήματος στο οποίο αυτά εκτελούνταν. Αναλυτικότερα, όταν έγινε προσπάθεια επεξεργασίας των εικόνων με τους προαναφερθέντες συνδυασμούς χωρικής ανάλυσης και συντελεστή μεγέθυνσης, γέμιζε η μνήμη της GPU του συστήματος με άμεσο αποτέλεσμα τον πρόωρο, ανεπιτυχή τερματισμό του αλγορίθμου. Για την αντιμετώπιση του συγκεκριμένου προβλήματος, προτάθηκε ο διαχωρισμός των εικόνων σε επιμέρους μη επικαλυπτόμενα τμήματα, κατάλληλου μεγέθους, και η ανεξάρτητη επεξεργασία των τμημάτων αυτών από το μοντέλο DIP. Τα μεγεθυμένα επιμέρους τμήματα της εκάστοτε εικόνας, στην συνέχεια, ενώνονται ξανά, δημιουργώντας κατά αυτό το τρόπο την τελική εικόνα αυξημένης χωρικής

ανάλυσης. Παρακάτω φαίνονται οι διαστάσεις και το πλήθος των υπο-εικόνων που χρησιμοποιήθηκαν για κάθε συνδυασμό χωρικής ανάλυσης και συντελεστή μεγέθυνσης.

- **10m, x2** : 4 υπο-εικόνες διαστάσεων 210x210.
- **10m, x4** : 9 υπο-εικόνες διαστάσεων 140x140
- **10m, x6** : 25 υπο-εικόνες διαστάσεων 84x84
- **20m, x4** : 4 υπο-εικόνες διαστάσεων 105x105
- **20m, x6** : 9 υπο-εικόνες διαστάσεων 70x70

Τέλος κατά την διαδικασία των πειραματισμών με τις υπερπαραμέτρους του μοντέλου, παρατηρήθηκε, για ορισμένες εικόνες του συνόλου δοκιμών, αλλοίωση της ποιότητας της τελικής εικόνας όταν η εκπαίδευση του δικτύου ξεπερνούσε κάποιο αριθμό εποχών. Ωστόσο, ο αριθμός αυτός δεν είναι ίδιος για όλες τις εικόνες του συνόλου δοκιμών. Επομένως, για την αντιμετώπιση του συγκεκριμένου προβλήματος, υλοποιήθηκε μία μέθοδος Early Stopping η οποία ήταν υπεύθυνη για τον πρόωρο τερματισμό της εκπαίδευσης του νευρωνικού δικτύου όταν η ενέργεια του σφάλματος εξόδου δεν μειωνόταν για 100 συνεχόμενες εποχές. Στο σημείο αυτό, κρίνεται σκόπιμο να σημειωθεί πως σε κάθε εποχή εκπαίδευσης που το μοντέλο πετύχαινε το ελάχιστο (μέχρι εκείνη τη στιγμή) σφάλμα εξόδου δημιουργούταν ένα στιγμιότυπο του μοντέλου, προκειμένου να επιστραφεί ως βέλτιστο σε περίπτωση πρόωρου τερματισμού της εκπαίδευσης.

3.1.2 Zero-Shot Super Resolution

Στην ενότητα 2.4.5.1 παρουσιάστηκαν δύο βαθιά συνελκτικά δίκτυα επιβλεπόμενης μάθησης για την επίλυση του προβλήματος του Super Resolution (SR), τα οποία επιτυγχάνουν να βελτιώσουν σε σημαντικό βαθμό την ποιότητα των παραγόμενων αποτελεσμάτων συγκριτικά με τις συμβατικές μεθόδους που αναλύθηκαν στο κεφάλαιο 2.3 . Ωστόσο, τα μοντέλα αυτά, όντας μοντέλα επιβλεπόμενης μάθησης, περιορίζονται σε συγκεκριμένα δεδομένα εκπαίδευσης, όπου η απόκτηση των εικόνων χαμηλής ανάλυσης (LR) από τις αντίστοιχες εκδοχές τους υψηλής ανάλυσης (HR) γίνεται με προκαθορισμένο, ομοιόμορφο τρόπο κάθε φορά (π.χ. μέσω δικυβικής υποδειγματοληψίας) και συνεπώς δεν περιέχουν παραμορφώσεις που πιθανώς να οφείλονται σε παράγοντες όπως ο θόρυβος του αισθητήρα, συμπίεση της εικόνας ή μία μη ιδανική Point Spread Function (PSF). Στην πράξη, όμως, οι πραγματικές εικόνες LR σπάνια ικανοποιούν τις παραπάνω προϋποθέσεις, με αποτέλεσμα τα παραγόμενα προϊόντα, από τα νευρωνικά δίκτυα, να μην πετυχαίνουν την μέγιστη δυνατή ποιότητα στην περίπτωση εικόνων αυτών.

Οι Shocher et al. στο [48] προτείνουν μία προσέγγιση μη επιβλεπόμενης μάθησης, κατά την οποία το νευρωνικό δίκτυο εκπαιδεύεται αποκλειστικά και μόνο από την πραγματική εικόνα χαμηλής χωρικής ανάλυσης. Το μοντέλο ονομάζεται Zero-Shot SR ή ZSSR και αξιοποιεί την παρατήρηση των Glasner et al. [33] περί επανεμφάνισης μικρών τμημάτων στις φυσικές εικόνες, τόσο εντός της ίδιας κλίμακας όσο και σε διαφορετικές κλίμακες, για την εκπαίδευση ενός νευρωνικού δικτύου με παραδείγματα τα οποία έχουν εξαχθεί αποκλειστικά από την υπό εξέταση εικόνα χαμηλής ευκρίνειας. Στην ενότητα 2.3.4.3 αναλύθηκε η μεθοδολογία των Glasner et al. για την βελτίωση της χωρικής ανάλυσης εικόνων χρησιμοποιώντας ως “σύνολο” εκπαίδευσης την ίδια την εικόνα χαμηλής χωρικής ανάλυσης. Από την συγκεκριμένη ανάλυση, γίνεται κατανοητό πως για τον εντοπισμό των όμοιων τμημάτων εντός της ίδιας κλίμακας και σε διαφορετικές κλίμακες γίνεται χρήση της απλής Ευκλείδειας ομοιότητας για την εκτέλεση

μίας γενικευμένης αναζήτησης K-κοντινότερων γειτόνων. Η χρήση, ωστόσο, της Ευκλείδειας ομοιότητας καθιστά το μοντέλο της ενότητας 2.3.4.3 ανίκανο να γενικεύσει την εφαρμογή του τόσο σε περιπτώσεις που το ένα τμήμα εικόνας δεν περιέχεται στην υπό εξέταση LR εικόνα, όσο και σε περιπτώσεις τμημάτων με μη ομοιόμορφα μεγέθη εντός της εικόνας. Αντιθέτως, η προτεινόμενη μέθοδος, χάρη στην εξαιρετική ικανότητα γενίκευσης των βαθιών συνελικτικών δικτύων, δεν υπόκειται στους προαναφερθέντες περιορισμούς και είναι σε θέση να εξάγει σύνθετες συσχετίσεις του τύπου HR-LR (ειδικές για κάθε εικόνα) από την εικόνα χαμηλής χωρικής ανάλυσης που εξετάζεται και τις υποδειγματοληπτημένες εκδοχές της. Τέλος, αξίζει να αναφερθεί πως η “εσωτερική” εντροπία των τμημάτων που εμφανίζονται σε μία εικόνα (χαμηλής) ανάλυσης είναι αρκετά μικρότερη από την εντροπία των εικόνων ενός εξωτερικού συνόλου εκπαίδευσης, γεγονός που οδηγεί στο συμπέρασμα πως οι εξαρτήσεις εντός της εκάστοτε εικόνας είναι συχνά πιο χρήσιμες για το πρόβλημα του Super Resolution, συγκριτικά με τις στατιστικές πληροφορίες που εξάγονται από ένα εξωτερικό σύνολο εκπαίδευσης.

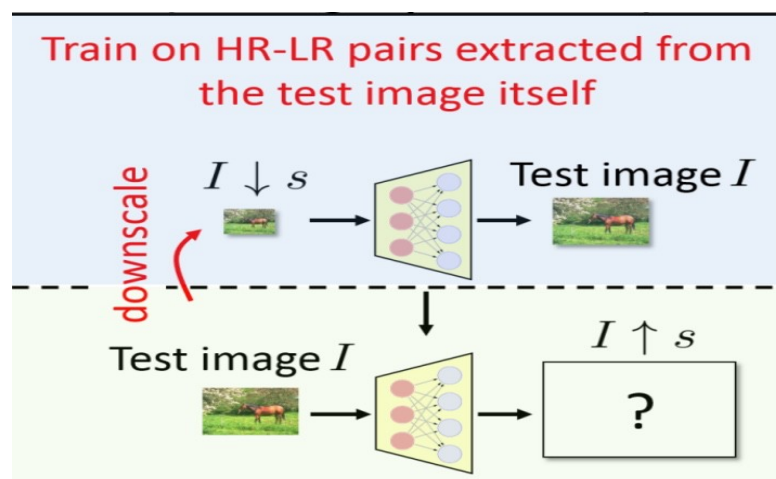
Στο σημείο αυτό, κρίνεται σκόπιμο να σημειωθεί πως, αν και ο όρος Zero-Shot προέρχεται από το πεδίο της ανίχνευσης αντικειμένων και της ταξινόμησης, το μοντέλο ZSSR δεν απαιτεί την ύπαρξη πρότερης γνώσης, υπό την μορφή εξωτερικών πληροφοριών σχετικές με την υπό ανάλυση εικόνα, όπως συμβαίνει στην περίπτωση του Zero-Shot Learning [49]. Παρόλα αυτά, στην περίπτωση που επιπρόσθετες πληροφορίες σχετικές με την εικόνα χαμηλής ανάλυσης είναι διαθέσιμες, το συγκεκριμένο μοντέλο είναι σε θέση να της αξιοποιήσει κατά την διαδικασία σύνθεσης του ζητούμενο αποτελέσματος προκειμένου να βελτιώσει περαιτέρω την ποιότητα των αποτελεσμάτων. Η δυνατότητα αυτή, προφανώς, δεν είναι διαθέσιμη στην περίπτωση μοντέλων επιβλεπόμενης μάθησης, τα οποία για να αξιοποιήσουν επιπρόσθετη πληροφορία απαιτούν την εκ νέου εκπαίδευσή τους, διαδικασία η οποία μπορεί να διαρκέσει έως και βδομάδες.

Η προτεινόμενη, από τους συγγραφείς, μεθοδολογία εκμεταλλεύεται τις παρατηρήσεις που έγιναν παραπάνω σχετικά με την αξία των επανεμφανιζόμενων τμημάτων στο εσωτερικό μίας εικόνας και προσπαθεί να τις συνδυάσει με την ισχυρή ικανότητα γενίκευσης των βαθιών νευρωνικών δικτύων. Αναλυτικότερα, δεδομένης μίας εικόνας χαμηλής ανάλυσης I ορίζεται ένα συνελικτικό δίκτυο, το οποίο στην συνέχεια θα προσαρμοστεί για την επεξεργασία αποκλειστικά της εικόνας I . Για την εκπαίδευση του δικτύου, όπως έχει προαναφερθεί, εξάγονται μικρότερα τμήματα της εικόνας I προκειμένου να χρησιμοποιηθούν ως δείγματα εκπαίδευσης. Η απόκτηση των δειγμάτων αυτών επιτυγχάνεται μέσω υποδειγματοληψίας της εικόνας I κατά παράγοντα s , όπου s είναι ο επιθυμητός συντελεστής μεγέθυνσης της εικόνας, και στην συνέχεια χρησιμοποιείται ένα σχετικά ρηχό συνελικτικό δίκτυο, η αρχιτεκτονική του οποίου θα αναλυθεί στην επόμενη ενότητα, για την ανακατασκευή της εικόνας I από την υποδειγματοληπτημένη εκδοχή της. Στην συνέχεια, το “εκπαιδευμένο” πλέον δίκτυο δέχεται ως είσοδο την εικόνα χαμηλής ανάλυσης I προκειμένου να παράξει την εικόνα υψηλότερης, κατά παράγοντα s , ανάλυσης.

Δεδομένου ότι στο σύνολο εκπαίδευσης αποτελείται αποκλειστικά και μόνο από την εικόνα χαμηλής ανάλυσης, οι Shocher et al. πρότειναν την εφαρμογή προσαύξησης δεδομένων (data augmentation) στην εικόνα I με σκοπό την παραγωγή περισσότερων ζευγών υψηλής-χαμηλής ανάλυσης. Η προσαύξηση αυτή επιτυγχάνεται μέσω υποδειγματοληψίας της εικόνας I κατά διαφορετικούς παράγοντες, δημιουργώντας έτσι πολλαπλές μικρότερες εκδοχές της συγκεκριμένης εικόνας ($I_o = I, I_1, I_2, \dots, I_n$). Οι εικόνες I_j , με $j=0,1,2,\dots,n$, που στο άρθρο ονομάζονται “HR fathers”, υποδειγματοληπτούνται

ξανά, αυτή την φορά, κατά τον επιθυμητό παράγοντα μεγέθυνσης s , δημιουργώντας τις αποκαλούμενες “LR sons” εικόνες οι οποίες δρουν ως δείγματα εισόδου του νευρωνικού δικτύου. Στην συνέχεια, στα προκύπτοντα νέα ζεύγη εφαρμόζονται επιπλέον μετασχηματισμοί, προκειμένου να αυξηθεί ακόμα παραπάνω το μέγεθος του συνόλου εκπαίδευσης. Αναλυτικότερα, σε καθένα από τα νέα HR-LR ζεύγη εικόνων εφαρμόζονται τέσσερις μετασχηματισμοί περιστροφής κατά γωνίες $0^\circ, 90^\circ, 180^\circ$ και 270° . Επιπλέον, για καθένα από τα ζεύγη λαμβάνονται και τα κατοπτρικά τους είδωλα ως προς την οριζόντια και κατακόρυφη κατεύθυνση, δημιουργώντας έτσι οκτώ νέα ζεύγη εικόνων υψηλής-χαμηλής ανάλυσης από κάθε αρχικό ζεύγος.

Ολοκληρώνοντας την ανάλυση της μεθοδολογίας πίσω από το μοντέλο ZSSR, κρίνεται σκόπιμο να γίνει αναφορά στο τρόπο εκπαίδευσης του μοντέλου. Συγκεκριμένα, για λόγους ευστάθειας του μοντέλου, καθώς και για την επίτευξη βελτίωσης της χωρικής ανάλυσης πολύ μικρών LR εικόνων κατά σχετικά μεγάλους παράγοντες μεγέθυνσης s , αλγόριθμος του [48] εφαρμόζεται σε πολλαπλά στάδια για ενδιάμεσους παράγοντες μεγέθυνσης ($s_1, s_2, \dots, s_m = s$). Σε κάθε ενδιάμεση κλίμακα s_i προστίθεται στο σύνολο εκπαίδευσης η παραγόμενη από νευρωνικό δίκτυο εικόνα μαζί με τις υποδειγματοληπτημένες και περιστρεφόμενες εκδοχές της, εμπλουτίζοντας ακόμα παραπάνω το σύνολο εκπαίδευσης. Στην συνέχεια, τα δείγματα του συνόλου εκπαίδευσης υποδειγματοληπτούνται κατά παράγοντα s_{i+1} για την παράγωγή των νέων ζευγών εκπαίδευσης. Η διαδικασία αυτή επαναλαμβάνεται μέχρι ο αλγόριθμος να φτάσει στον ζητούμενο παράγοντα μεγέθυνσης s .



Εικόνα 21: Η διαδικασία εκπαίδευσης του ZSSR [48].

3.1.2.1 Αρχιτεκτονική και βελτιστοποίηση του δικτύου

Στην προηγούμενη ενότητα έγινε αναφορά στην πολυπλοκότητα των εξαρτήσεων που εμφανίζονται μεταξύ των ζευγών HR-LR της συγκεκριμένης μεθοδολογίας. Εξαιτίας, λοιπόν, της σχετικά μικρής πολυπλοκότητας, οι συγγραφείς πρότειναν ένα σχετικά ρηχό συνελκτικό δίκτυο για την επίλυση του προβλήματος του Super Resolution. Το συνελκτικό δίκτυο που χρησιμοποιήθηκε στην συγκεκριμένη προσέγγιση και κατ' επέκταση στην παρούσα εργασία, αποτελείται από ένα συνελκτικό επίπεδο εισόδου και έξι (6) κρυφά συνελκτικά επίπεδα καθένα από τα οποία διαθέτει εξηντατέσσερα (64) κανάλια (θεωρείται πως το επίπεδο ενεργοποίησης βρίσκεται ενσωματωμένο στο τέλος κάθε συνελκτικού επιπέδου). Ως συνάρτηση ενεργοποίησης των νευρώνων των κρυφών επιπέδων του δικτύου επιλέχθηκε να χρησιμοποιηθεί η ReLU. Το επίπεδο της εξόδου του δικτύου, είναι επίσης ένα συνελκτικό

επίπεδο στην έξοδο του οποίου, ωστόσο, δεν εφαρμόζεται καμία συνάρτηση ενεργοποίησης. Η είσοδος του συνελκτικού δικτύου είναι μια υπερδειγματοληπτημένη εκδοχή της εικόνας χαμηλής ανάλυσης, έστω I_{1s} . Επιπρόσθετα, αξίζει να σημειωθεί πως, όπως και στην προσέγγιση των Lanaras et al., το δίκτυο περιέχει Skip Connection από το επίπεδο της εισόδου προς το επίπεδο της εξόδου, προκειμένου αυτό μετά το πέρας της εκπαίδευσής του να έχει μάθει την υπολειμματική εικόνα που όταν προστεθεί στην I_{1s} θα παράξει το επιθυμητό αποτέλεσμα.

Για την εκπαίδευση του συνελκτικού δικτύου, χρησιμοποιείται ως συνάρτηση σφάλματος της εξόδου η L_1 απόσταση μεταξύ της εξόδου του δικτύου και του εκάστοτε HR father. Επίσης, αξιοποιείται ο βελτιστοποιητής (optimizer) Adam του οποίου το learning rate, στην συγκεκριμένη προσέγγιση, δεν είναι σταθερό. Αντιθέτως, στην αρχή της εκπαίδευσης η παράμετρος learning rate τίθεται στην τιμή 0.001 και μειώνεται καθώς προχωράει η εκπαίδευση του δικτύου. Για την επιλογή των εποχών που θα μειωθεί η τιμή της συγκεκριμένης παραμέτρου, προτάθηκε η πραγματοποίηση μίας εκτίμησης των μελλοντικών τιμών του σφάλματος ανακατασκευής, μέσω γραμμικής παρεμβολής, και αν η τυπική απόκλιση του σφάλματος είναι μεγαλύτερη από την κλίση της προσεγγιστικής ευθείας τότε η τιμή του learning rate υποδεκαπλασιάζεται. Στο σημείο αυτό, αξίζει να αναφερθεί πως η εκπαίδευση του δικτύου ολοκληρώνεται είτε όταν το learning rate γίνει μικρότερο του 10^{-6} είτε σε κάποιο μέγιστο αριθμό εποχών που ορίζεται από τον χρήστη (στην παρούσα εργασία η μέγιστη τιμή επιλέχθηκε ίση με 5000 εποχές).

Τέλος, για την σύνθεση της εξόδου του νευρωνικού δικτύου εφαρμόζεται μία μέθοδος παρόμοια με το γεωμετρικό self-ensemble που προτείνεται στο [50], η οποία δημιουργεί οκτώ διαφορετικές εξόδους, μία για κάθε περιστροφή και κατοπτρισμό της υπό εξέταση εικόνας I . Επιπλέον, εφαρμόζεται η τεχνική της οπισθοπροβολής (back-projection) των [51],[33] σε κάθε μία από τις οκτώ εξόδους. Στην συνέχεια, από τις τροποποιημένες εξόδους επιλέγεται ο διάμεσος όρος, αντί το μέσου όρου που επιλέγεται στην υλοποίηση του [50].

3.1.2.2 Τροποποιήσεις και συνεισφορά

Η υλοποίηση του ZSSR, όπως και αυτή του Deep Image Prior, είχε προταθεί αρχικά για την επεξεργασία συμβατικών τριχρωματικών (RGB) εικόνων. Επομένως, γίνεται κατανοητό πως απαιτούνται ορισμένες τροποποιήσεις στην δομή του δικτύου προκειμένου αυτό να είναι σε θέση να διαχειριστεί δορυφορικές εικόνες. Αρχικά, ήταν αναγκαία η αλλαγή του πρώτου συνελκτικού επιπέδου, ώστε αυτό να περιμένει ως είσοδό του ένα τανυστή με πλήθος καναλιών ίδιο με αυτό της εικόνας I . Με την αλλαγή αυτή, επιτυγχάνεται η ικανότητα του δικτύου να διαχειρίζεται εικόνες οποιουδήποτε μεγέθους και με οποιοδήποτε πλήθος καναλιών, χωρίς την ανάγκη εξωτερικής εξωτερικής παρέμβασης από τον χρήστη. Επιπλέον, επειδή το δίκτυο ήταν σχεδιασμένο για την επεξεργασία RGB εικόνων, είχε επιλεγεί από τους Shocher et al. να χρησιμοποιούνται PNG εικόνες με 8bit ακρίβεια. Η επιλογή αυτή, όπως και στο προηγούμενο μοντέλο, αποτελεί εμπόδιο στην επεξεργασία των δορυφορικών εικόνων, των οποίων ο τύπος δεδομένων είναι, όπως έχει ήδη αναφερθεί, είναι int16. Για την αντιμετώπιση του συγκεκριμένου προβλήματος, αποφασίστηκε το νευρωνικό δίκτυο να αναμένει ως είσοδο ένα τανυστή του οποίου ο τύπος δεδομένων θα είναι float32, προκειμένου το τελικό αποτέλεσμα που παράγεται να είναι επίσης σε μορφή float32 και συνεπώς να μην υπάρχει καμία απώλεια πληροφοριών εξαιτίας της μετατροπής σε ένα τύπο δεδομένων χαμηλότερης ακρίβειας (όπως είναι το uint8). Επιπρόσθετα, για την εκτέλεση της προσαύξησης των δεδομένων εισόδου και συγκεκριμένα για την εκτέλεση των

μετασχηματισμών στροφής και κατοπτρισμού, οι συγγραφείς είχαν χρησιμοποιήσει την μέθοδο *wrapPerspective* της βιβλιοθήκης *OpenCV*, η οποία ωστόσο μπορεί να διαχειριστεί εικόνες με έως και τέσσερα χρωματικά κανάλια. Για την αντιμετώπιση του συγκεκριμένου προβλήματος, επιλέχθηκε στην περίπτωση των εικόνων ανάλυσης 20m GSD (οι οποίες έχουν έξι χρωματικά κανάλια) να πραγματοποιηθεί, σε πρώτη φάση, μία διάσπαση της αρχικής εικόνας σε δύο επιμέρους εικόνες με τρία χρωματικά κανάλια η κάθε μία στις οποίες θα εφαρμοστούν οι προαναφερθέντες μετασχηματισμοί. Μετά την εφαρμογή των μετασχηματισμών, οι αντίστοιχες εικόνες που προκύπτουν συνενώνονται ξανά για την δημιουργία της, τελικά, μετασχηματισμένης εικόνας με ανάλυση 20m GSD.

Στο σημείο αυτό, κρίνεται σκόπιμο να σημειωθεί πως στην περίπτωση του ZSSR μοντέλου δεν χρειάστηκε η εφαρμογή *histogram stretching* στις εικόνες εισόδου, καθώς όλες οι εικόνες που εμπλέκονται στην εκπαίδευση του νευρωνικού δικτύου προέρχονται από την υπό μελέτη εικόνα χαμηλής ανάλυσης και συνεπώς τα ιστογράμμά τους μοιράζονται παρόμοια μορφή και χαρακτηριστικά.

3.1.3 Degradation Aware Super Resolution

Στην ενότητα 2.3.4.1, υιοθετήθηκε η σχέση $I^{LR} = (I^{HR} \otimes k) \downarrow s + n$, όπου I^{LR}, I^{HR}, k, n είναι η εικόνα χαμηλής χωρικής ανάλυσης, η εικόνα υψηλής χωρικής ανάλυσης, ο πυρήνας παραμόρφωσης και ο προσθετικός θόρυβος αντίστοιχα, ενώ το $\otimes$ υποδηλώνει την πράξη της συνέλιξης και το $\downarrow s$ την πράξη της υποδειγματοληψίας κατά παράγοντα s , για την μοντελοποίηση του προβλήματος του Super Resolution, ως ένα αντίστροφο πρόβλημα αποσυνέλιξης για τον προσδιορισμό της εικόνας υψηλής χωρικής ανάλυσης I^{HR} . Από την παραπάνω μοντελοποίηση, γίνεται εύκολα κατανοητό πως το πρόβλημα του Super Resolution σχετίζεται σε μεγάλο βαθμό από τον πυρήνα παραμόρφωσης. Τα μοντέλα βαθιάς μάθησης που παρουσιάστηκαν στην ενότητα 2.4.5.1, όπως και η πλειοψηφία των μοντέλων που χρησιμοποιούνται για την επίλυση του προβλήματος του Super Resolution, ακολουθούν την υπόθεση ότι ο πυρήνας παραμόρφωσης είναι γνωστός και σταθερός (π.χ. πυρήνας δικυβικής ή διαγραμμικής παρεμβολής) σε όλες τις εικόνες που πρόκειται να αναλυθούν. Δυστυχώς, στην πράξη πολλές φορές αυτή η υπόθεση παραβιάζεται, συνεπάγοντας άμεσα την σημαντική μείωση των επιδόσεων των μοντέλων αυτών σε πραγματικές εικόνες χαμηλής χωρικής ανάλυσης.

Σε μία προσπάθεια βελτίωσης των αποτελεσμάτων των μοντέλων που επιλύουν το πρόβλημα του Super Resolution, προτάθηκε η υλοποίηση αλγορίθμων για την ανάλυση της εκάστοτε εικόνας με σκοπό τον προσδιορισμό του πιθανότερου πυρήνα παραμόρφωσης που οδήγησε στην δημιουργία αυτής. Τα μοντέλα, ωστόσο, που χρησιμοποιούσαν την επιπρόσθετη αυτή πληροφορία, αποδείχθηκαν ιδιαίτερα ευαίσθητα στην προσέγγιση της παραμόρφωσης, καθώς ακόμη και ένα μικρό σφάλμα στην προσέγγιση του πυρήνα αυτού ενισχύεται πολλαπλασιαστικά από το μοντέλο επίλυσης του Super Resolution, οδηγώντας έτσι στην εμφάνιση τεχνουργημάτων (*artifacts*) στο παραγόμενο αποτέλεσμα. Για την επίλυση του συγκεκριμένου προβλήματος προτάθηκε από τους Gu et al. στο [52] μία μεθοδολογία επαναληπτικής διόρθωσης της προσέγγισης του πυρήνα παραμόρφωσης (IKC), μέσω της παρατήρησης των προηγούμενων αποτελεσμάτων υψηλής ανάλυσης που παράγονται από ένα μοντέλο. Η μεθοδολογία αυτή, αν και πράγματι μετά από πεπερασμένο αριθμό επαναλήψεων παράγει εικόνες υψηλής ανάλυσης χωρίς την παρουσία *artifacts*, είναι ιδιαίτερα χρονοβόρα και συνεπώς σχετικά ακριβή σε υπολογιστικούς πόρους.

Σε αντίθεση με τις μεθοδολογίες που αναλύθηκαν προηγουμένως, οι Wang et al. στο [53] πρότειναν μία διαφορετική προσέγγιση κατά την οποία μαθαίνεται μία αναπαράσταση της παραμόρφωσης, και όχι ο ακριβής πυρήνας, που είναι επαρκής για το διαχωρισμό δύο διαφορετικών ειδών παραμόρφωσης. Για την εξαγωγή της προαναφερθείσας αναπαράστασης χρησιμοποιείται ένας διαφορετικός τρόπος μάθησης, το λεγόμενο *contrastive learning*, ο οποίος συγκρίνει, χωρίς την ανάγκη επίβλεψης, θετικά και αρνητικά ζεύγη δειγμάτων εντός του χώρου αναπαράστασης.

Το *contrastive learning* έχει αποδειχθεί ιδιαίτερα αποτελεσματική μεθοδολογία για την εκμάθηση μίας αναπαράστασης χωρίς επίβλεψη. Σε προηγούμενες εφαρμογές, η εκμάθηση μίας αναπαράστασης συνήθως γινόταν με τρόπο παρόμοιο με αυτόν της επιβλεπόμενης εκπαίδευσης ενός νευρωνικού δικτύου, όπως περιγράφηκε στην ενότητα 2.4.4.2.1, ελαχιστοποιώντας δηλαδή τη διαφορά μεταξύ της εξόδου και ενός σταθερού στόχου (π.χ. της ίδια της εισόδου για τους κωδικοποιητές - *autoencoders*). Στην περίπτωση του *contrastive learning*, αντί της ελαχιστοποίησης της διαφοράς μεταξύ εξόδου και ενός προκαθορισμένου, σταθερού στόχου, επιχειρείται η μεγιστοποίηση της αμοιβαίας πληροφορίας (*mutual information*) στο χώρο αναπαράστασης. Συγκεκριμένα, η αναπαράσταση ενός υπό εξέταση δείγματος θα πρέπει να προσελκύει τα θετικά (παρόμοια) δείγματα και να απωθεί τα αρνητικά. Τα θετικά παραδείγματα δύναται να είναι μετασχηματισμένες εκδοχές της εισόδου, πολλαπλές εμφανίσεις της εισόδου, καθώς και γειτονικά τμήματα της ίδιας εικόνα.

Τα πλεονεκτήματα της εκμάθησης μίας αναπαράστασης της παραμόρφωσης, αντί ολόκληρου του πυρήνα, είναι δίπτυχα. Αρχικά, συγκριτικά με την εξαγωγή ολόκληρης της προσέγγισης του πυρήνα παραμόρφωσης, είναι πολύ ευκολότερο να μαθευτεί μία γενικότερη και αφηρημένη αναπαράσταση επαρκής για τον διαχωρισμό δύο διαφορετικών ειδών παραμόρφωσης. Συνεπώς, καθίσταται δυνατή η απόκτηση διακριτικών αναπαραστάσεων ικανές να παρέχουν επαρκείς και ακριβείς πληροφορίες σχετικά με το είδος της παραμόρφωσης, χωρίς την ανάγκη εφαρμογής διορθωτικών επαναληπτικών αλγορίθμων όπως το IKC. Επιπλέον, χάρη στο *contrastive learning*, η εκμάθηση μίας αναπαράστασης παραμόρφωσης γίνεται χωρίς επίβλεψη, χωρίς δηλαδή να είναι διαθέσιμες οι πληροφορίες σχετικά με τις πραγματικές παραμορφώσεις των εικόνων που χρησιμοποιούνται κατά την εκπαίδευση. Επομένως, η χρήση των αναπαραστάσεων αυτών αποδεικνύεται ιδιαίτερα χρήσιμη σε εφαρμογές που αναλύουν πραγματικές εικόνες στις οποίες οι υπάρχουσες παραμορφώσεις είναι άγνωστες.

Βασιζόμενοι, λοιπόν, στις παραπάνω παρατηρήσεις οι Wang et al. πρότειναν μία ενοποιημένη μεθοδολογία που χρησιμοποιεί ένα συνελικτικό νευρωνικό δίκτυο (κωδικοποιητής) για την εξαγωγή μίας αναπαράστασης της παραμόρφωσης, η οποία στην συνέχεια δίνεται ως είσοδος μαζί με την υπό εξέταση εικόνα σε ένα δεύτερο συνελικτικό δίκτυο με σκοπό την σύνθεση, τελικά, της ζητούμενης εικόνας υψηλής χωρικής ανάλυσης. Στη μεθοδολογία αυτή, πραγματοποιείται η υπόθεση πως η παραμόρφωση είναι παντού ίδια στο εσωτερικό μίας εικόνας και διαφορετική μεταξύ δύο ανεξάρτητων εικόνων, γεγονός το οποίο ισχύει στην πλειοψηφία των πραγματικών εικόνων. Σύμφωνα με την παραπάνω υπόθεση, όλα τα τμήματα μίας εικόνας πρέπει να είναι όμοια μεταξύ τους, ως προς την παραμόρφωση, στον χώρο αναπαραστάσεων και να διαφέρουν από τα επιμέρους τμήματα μίας δεύτερης διαφορετικής εικόνας. Επομένως, ως θετικά δείγματα αναγνωρίζονται τμήματα εικόνων που προέρχονται από την ίδια εικόνα και ως αρνητικά αναγνωρίζονται τα τμήματα εκείνα που προέρχονται από διαφορετικές εικόνες, ενώ στόχος του *contrastive learning* είναι η εξαγωγή

μίας αναπαράστασης των παραμορφώσεων η οποία θα είναι ανεξάρτητη του περιεχομένου της εκάστοτε εικόνας.

Η διαδικασία εκπαίδευσης του κωδικοποιητή ακολουθεί τις ίδιες αρχές με αυτή της επιβλεπόμενης εκπαίδευσης ενός κλασικού νευρωνικού δικτύου, με την διαφορά πως στην συγκεκριμένη περίπτωση αλλάζει η συνάρτηση υπολογισμού του σφάλματος της εξόδου. Αναλυτικότερα, δεδομένου ενός υπό εξέταση τμήματος εικόνας, εξάγονται επιπλέον επιμέρους τμήματα από την ίδια εικόνα για την σύνθεση του συνόλου θετικών δειγμάτων και επιμέρους τμήματα από διαφορετικές εικόνες για την κατασκευή του αντίστοιχου συνόλου αρνητικών δειγμάτων. Όλα τα προαναφερθέντα τμήματα περνάνε, σε πρώτη φάση, το δίκτυο του κωδικοποιητή και στην συνέχεια, όπως προτείνεται στα [54], [55] , η έξοδος του κωδικοποιητή διέρχεται και από ένα feed forward νευρωνικό δίκτυο δύο επιπέδων για την απόκτηση των αναπαραστάσεων x, x^+, x^- , που είναι οι αναπαραστάσεις του υπό εξέταση τμήματος της εικόνας, των θετικών και αρνητικών δειγμάτων αντίστοιχα. Για τον υπολογισμό της ομοιότητας στον χώρο των αναπαραστάσεων χρησιμοποιείται το λεγόμενο InfoNCE Loss που υπολογίζεται από την σχέση :

$$L_x = -\log \frac{\exp(x \cdot x^+ / \tau)}{\sum_{n=1}^N \exp(x \cdot x_n^- / \tau)} \quad (1)$$

, όπου N το πλήθος των αρνητικών δειγμάτων,

τ μία υπερπαράμετρος θερμοκρασίας ,

“•” ο τελεστής του διανυσματικού εσωτερικού γινομένου.

Στο σημείο αυτό, κρίνεται σκόπιμο να αναφερθεί πως για την απόκτηση “καλών” αναπαραστάσεων παραμόρφωσης, που χαρακτηρίζονται από ανεξαρτησία από το χωρικό περιεχόμενο των εικόνων από όπου προκύπτουν, γίνεται χρήση μίας ουράς στην οποία περιέχονται δείγματα από διάφορες εικόνες με διαφορετικό περιεχόμενο και διαφορετικές παραμορφώσεις. Συγκεκριμένα, κατά την διάρκεια της εκπαίδευσης του μοντέλου, επιλέγονται

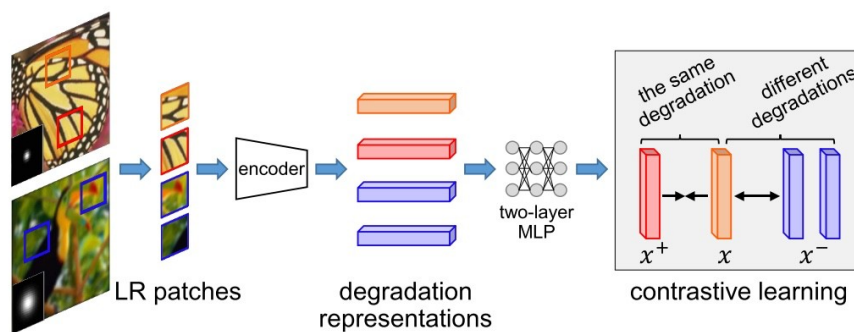
B τυχαίες εικόνες χαμηλής ευκρίνειας από τις οποίες εξάγονται δύο επιμέρους τμήματα των εικόνων μέσω τυχαίας περικοπής. Στην συνέχεια, αυτά τα $2B$ τμήματα εικόνων κωδικοποιούνται από τον κωδικοποιητή στις αναπαραστάσεις $p_i^1, p_i^2 \in \mathbb{R}^{256}$, όπου p_i^1 αντιστοιχεί στο πρώτο επιμέρους τμήμα που εξάχθηκε από την i -οστή εικόνα. Για κάθε εικόνα i , τα p_i^1, p_i^2 αναγνωρίζονται ως το υπό εξέταση τμήμα και το θετικό δείγμα αντίστοιχα. Με βάση, λοιπόν, τις παραπάνω τροποποιήσεις, το συνολικό σφάλμα για το contrastive learning ορίζεται ως :

$$L_{degrad} = \sum_{i=1}^B -\log \frac{\exp(p_i^1 \cdot p_i^2 / \tau)}{\sum_{j=1}^{N_{queue}} \exp(p_i^1 \cdot p_{queue}^j / \tau)} \quad (2)$$

, όπου N_{queue} ο αριθμός δειγμάτων στην ουρά,

p_{queue}^j το j -οστό αρνητικό δείγμα.

Στην εικόνα 22 φαίνεται οπτικοποίηση της διαδικασίας εκμάθησης αναπαραστάσεων παραμόρφωσης. Συγκεκριμένα, το πορτοκαλί τμήμα της επάνω εικόνας θεωρείται πως είναι το υπό εξέταση τμήμα, ενώ το κόκκινο τμήμα από την ίδια εικόνα αναγνωρίζεται ως θετικό δείγμα. Αντιθέτως, τα δύο επιμέρους τμήματα, μπλε χρώματος, που εξάγονται από την κάτω εικόνα αναγνωρίζονται ως αρνητικά δείγματα.



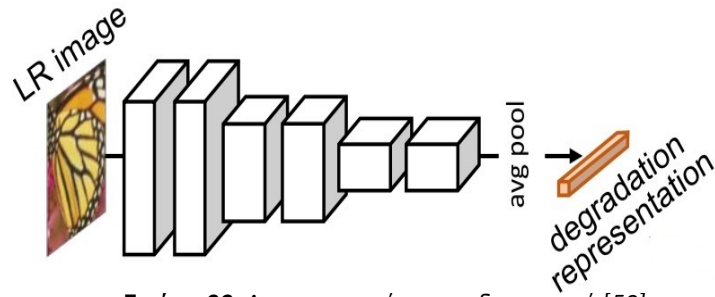
Εικόνα 22: Οπτικοποίηση της διαδικασίας εκμάθησης αναπαραστάσεων παραμόρφωσης [53].

3.1.3.1 Αρχιτεκτονική και βελτιστοποίηση των δικτύων

Στην προηγούμενη ενότητα έγινε μια σύντομη αναφορά στην δομή του συνολικού μοντέλου που χρησιμοποιείται στο [53] για την επίλυση του προβλήματος του Super Resolution. Συγκεκριμένα, σημειώθηκε πως η συγκεκριμένη μεθοδολογία αξιοποιεί δύο επιμέρους δίκτυα, ένα δίκτυο κωδικοποιητή για την παραγωγή των αναπαραστάσεων παραμόρφωσης η έξοδος του οποίου δίνεται ως είσοδος στο επόμενο συνελικτικό δίκτυο που είναι υπεύθυνο για την σύνθεση των εικόνων υψηλής χωρικής ανάλυσης. Στην συνέχεια, παρουσιάζονται αναλυτικά οι αρχιτεκτονικές των δύο αυτών δικτύων, καθώς και υπερπαραμέτροι που χρησιμοποιήθηκαν για την εκπαίδευσή τους.

Αρχιτεκτονική κωδικοποιητή

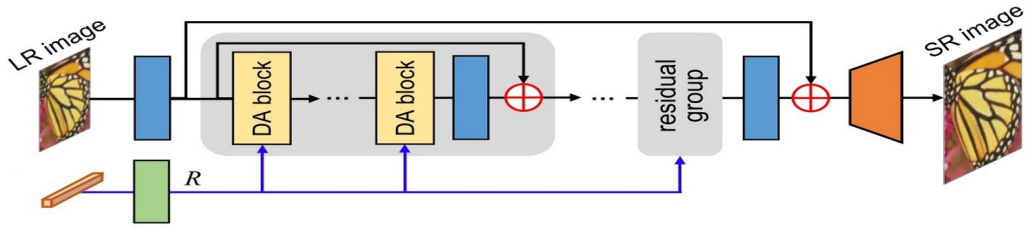
Το δίκτυο του κωδικοποιητή αποτελείται από έξι συνελικτικά επίπεδα, των οποίων ο πυρήνας συνέλιξης έχει διαστάσεις 3x3, ακολουθούμενα το καθένα από ένα επίπεδο Batch Normalization και στην συνέχεια ένα επίπεδο ενεργοποίησης μέσω της μη γραμμικής συνάρτησης LeakyReLU, με κλίση ίση με 0.1. Τα έξι συνελικτικά στρώματα με τα δύο επίπεδα που ακολουθούν το καθένα, δύναται να χωριστούν σε 3 υποομάδες με την κάθε μία να περιέχει δύο συνελικτικά στρώματα και τα αντίστοιχα επακόλουθα επίπεδα. Αναλυτικότερα, τα συνελικτικά επίπεδα της πρώτης ομάδας διαθέτουν 64 κανάλια και η τιμή του stride έχει τεθεί ίση με την μονάδα, τα συνελικτικά επίπεδα της δεύτερης ομάδας έχουν 128 κανάλια και η τιμή του stride είναι ίση με την μονάδα στο πρώτο επίπεδο και ίση με δύο στο δεύτερο επίπεδο. Τέλος, η τρίτη υποομάδα περιέχει συνελικτικά επίπεδα με 256 κανάλια, των οποίων οι τιμές του stride ακολουθεί την ίδια λογική με την προηγούμενη υποομάδα, και ακολουθείται από ένα επίπεδο Average Pooling. Η έξοδος του συνελικτικού αυτού τμήματος, προωθείται έπειτα σε ένα πλήρως συνδεδεμένο ρηχό δίκτυο δύο επιπέδων, από τα οποία το πρώτο έχει ως συνάρτηση ενεργοποίησης πάλι την LeakyReLU (με κλίση 0.1) ενώ το δεύτερο δεν διαθέτει κάποια ενεργοποίηση. Το τμήμα αυτό του δικτύου το κωδικοποιητή παράγει την ζητούμενη αναπαραστάση της παραμόρφωσης, η οποία είναι ένα διάνυσμα με 256 χαρακτηριστικά (features).



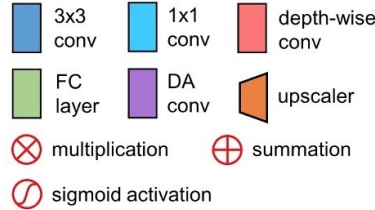
Εικόνα 23: Αρχιτεκτονική του κωδικοποιητή [53].

Αρχιτεκτονική του Degradation Aware Super Resolution (DASR) δικτύου

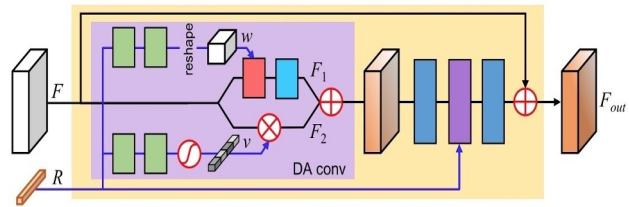
Η αρχιτεκτονική του δικτύου που εκτελεί την σύνθεση της ζητούμενης εικόνας υψηλής ανάλυσης είναι αρκετά πιο σύνθετη από του κωδικοποιητή και συνεπώς θα διασπαστεί σε επιμέρους αφηρημένα τμήματα για λόγους ευκολίας στην περιγραφή. Αναλυτικότερα, το δίκτυο για την επεξεργασία των εισόδων του, την έξοδο δηλαδή του κωδικοποιητή και την εκάστοτε εικόνα χαμηλής χωρικής ανάλυσης, διαθέτει αρχικά ένα συνελκτικό επίπεδο με πυρήνα διαστάσεων 3×3 και 64 κανάλια καθώς και ένα πλήρως συνδεδεμένο επίπεδο με σκοπό την συμπίεση της αναπαράστασης της παραμόρφωσης από τα 256 features στα 64. Στην συνέχεια, διαθέτει πέντε υπολειμματικές ομάδες (residual groups), κάθε μία από τις οποίες αποτελείται πέντε Degradation Aware (DA) blocks ακολουθούμενα από ένα συνελκτικό επίπεδο με πυρήνα 3×3 και 64 κανάλια. Επιπλέον, μετά το τελευταίο residual group υπάρχει ένα ακόμη συνελκτικό επίπεδο με πυρήνα 3×3 και 64 κανάλια το οποίο δέχεται ως είσοδο, προφανώς, την έξοδο του τελευταίου residual group. Επιπρόσθετα, αξιοποιείται ένα short Skip Connection σε καθένα από τα residual groups, προκειμένου να επιτευχθεί η άθροιση της εισόδου του εκάστοτε group με την έξοδο του τελευταίου του συνελκτικού επιπέδου. Το αποτέλεσμα της άθροισης αυτής αποτελεί την είσοδο του επόμενου residual group. Επίσης, χρησιμοποιείται και ένα long Skip Connection το οποίο προωθεί τα 64 feature maps, που παράγονται από το συνελκτικό επίπεδο στην είσοδο του δικτύου, προς το επίπεδο μετά το τελευταίο συνελκτικό επίπεδο. Η έξοδος του τελευταίου συνελκτικού επιπέδου αθροίζεται με τα feature maps που προωθούνται μέσω του long Skip Connection και το αποτέλεσμα της άθροισης περνά τελικά από ένα στρώμα υπερδειγματοληψίας με σκοπό την παραγωγή της ζητούμενης εικόνας χαμηλής ανάλυσης. Η αρχιτεκτονική που μόλις περιγράφηκε παρουσιάζεται στην εικόνα 24.



Εικόνα 24: Η αρχιτεκτονική DASR [53].



Εικόνα 25: Επεξήγηση των Blocks.



Εικόνα 26: Η αρχιτεκτονική ενός DA Block [53].

Στην εικόνα 26 παρουσιάζεται η αρχιτεκτονική ενός DA block, όπου φαίνεται πως κάθε DA block αποτελείται από δύο DA συνελικτικά επίπεδα, ακολουθούμενα το καθένα από ένα συνελικτικό επίπεδο με 64 κανάλια και πυρήνα διαστάσεων 3x3. Η ενεργοποίηση τόσο των συμβατικών όσο και των γενικευμένων (DA) συνελικτικών επιπέδων γίνεται με την συνάρτηση LeakyReLU της οποίας η παράμετρος κλίσης τίθεται στην τιμή 0.1. Τέλος, αξίζει να σημειωθεί πως για ακόμη μία φορά, χρησιμοποιείται Skip Connection με σκοπό την προώθηση της εισόδου του DA block προκειμένου αυτή να αθροιστεί με την εξόδο του τελευταίου συνελικτικού επιπέδου και να δημιουργηθεί η εξόδος του DA block.

Στο εσωτερικό ενός DA συνελικτικού επιπέδου υπάρχουν δύο επιμέρους κλάδοι. Στο πρώτο κλάδο, η συμπιεσμένη αναπαράσταση της παραμόρφωσης R περνά από δύο πλήρως συνδεδεμένα επίπεδα, από τα οποία μόνο το πρώτο διαθέτει συνάρτηση ενεργοποίησης την LeakyReLU με τιμή της παραμέτρου κλίσης ίση με 0.1. Η εξόδος των δύο αυτών επιπέδων, που διαθέτει 64 features, μετασχηματίζεται σε ένα τανυστή $w \in \mathbb{R}^{C \times 1 \times 3 \times 3}$. Στην συνέχεια, τα feature maps F που δίνονται ως είσοδοι στο εκάστοτε DA συνελικτικό επίπεδο, περνάνε από ένα συνελικτικό επίπεδο, όπου η συνέλιξη γίνεται κατά μήκος του άξονα των καναλιών (αντί των χωρικών διαστάσεων που είναι το συνηθισμένο) και πυρήνας του οποίου είναι ο τανυστής w που προσδιορίστηκε προηγουμένως. Η εξόδος του συνελικτικού αυτού επιπέδου περνά, έπειτα, από ένα ακόμη συμβατικό συνελικτικό επίπεδο 64 καναλιών με πυρήνα διαστάσεων 1x1, με σκοπό την δημιουργία της εξόδου F_1 του πρώτου κλάδου.

Στο δεύτερο κλάδο ενός DA συνελικτικού επιπέδου, η συμπιεσμένη αναπαράσταση R περνά για ακόμη μία φορά από δύο πλήρως συνδεδεμένα επίπεδα, τα οποία ακολουθούν την ίδια ακριβώς λογική με αυτά του πρώτου κλάδου. Ωστόσο, αυτή την φορά το δεύτερο πλήρως συνδεδεμένο επίπεδο διαθέτει συνάρτηση ενεργοποίησης την σιγμοειδή, παράγοντας έτσι τους αποκαλούμενους από τους συγγραφείς συντελεστές διαμόρφωσης ως προς την διάσταση των καναλιών v . Οι συντελεστές v πολλαπλασιάζονται, στην συνέχεια, με την τα feature maps εισόδου F κατά μήκος της διάστασης των καναλιών, προκειμένου να παραχθεί η εξόδος F_2 του δεύτερου κλάδου. Τελικά, οι εξόδοι F_1, F_2 των δύο κλάδων αθροίζονται με σκοπό την δημιουργία των feature maps εξόδου F_{out} του DA συνελικτικού επιπέδου, το οποίο δίνεται ως είσοδος στα επόμενα επίπεδα του DA block.

Βελτιστοποίηση

Στο σημείο αυτό κρίνεται σκόπιμο να αναφερθούν ορισμένες λεπτομέρειες σχετικά με τις υπερπαραμέτρους που χρησιμοποιήθηκαν κατά την διαδικασία της εκπαίδευσης του μοντέλου, καθώς και για την ίδια την εκπαίδευση του μοντέλου. Αρχικά, σημειώνεται πως για την εκπαίδευση του δικτύου χρησιμοποιήθηκαν 800 εικόνες από το σύνολο DIV2K και 2650 από το Flickr2K. Ο συντελεστής B που εμφανίζεται στην εξίσωση (2) της προηγούμενης ενότητας τέθηκε στην τιμή 32. Συγκεκριμένα, για την απόκτηση των 32 εικόνων χαμηλής ανάλυσης, που θα χρησιμοποιηθούν στην συνέχεια για την εξαγωγή των επιμέρους τμημάτων, επιλέχθηκαν 32 τυχαίες εικόνες από το σύνολο εκπαίδευσης στις οποίες εφαρμόστηκαν 32 διαφορετικοί, τυχαία επιλεγμένοι, γκαουσιανοί πυρήνες παραμόρφωσης, των οποίων η τυπική απόκλιση σ κυμαίνεται μεταξύ του $[0.2, 2.0]$ και $[0.2, 4.0]$ για παράγοντες μεγέθυνσης 2 και 4 αντίστοιχα. Επιπλέον, η παράμετρος θερμοκρασίας τ τέθηκε στην τιμή 0.07, ενώ το μέγεθος της ουράς N_{queue} θεωρήθηκε ίσο με 8192 δείγματα.

Για την εκπαίδευση του μοντέλου, χρησιμοποιήθηκε ο βελτιστοποιητής (optimizer) Adam με τιμές υπερπαραμέτρων $\beta_1=0.9, \beta_2=0.999$. Επιπρόσθετα, αξίζει να σημειωθεί πως για τις πρώτες 100 εποχές εκπαίδευσης του μοντέλου, το τμήμα DASR, παρέμεινε “παγωμένο” προκειμένου να βελτιστοποιηθεί αποκλειστικά και μόνο ο κωδικοποιητής. Το αρχικό learning rate τέθηκε ίσο με 10^{-3} και μετά από 60 εποχές μειώθηκε στην τιμή 10^{-4} . Με την ολοκλήρωση της εκατοστής εποχής, ξεκινάει η εκπαίδευση ολόκληρου του μοντέλου για άλλες 500 εποχές, με αρχικό learning rate ίσο με 10^{-4} το οποίο υποδιπλασιαζόταν κάθε 125 εποχές. Στην φάση εκπαίδευσης ολόκληρου το δικτύου, ως συνάρτηση σφάλματος της εξόδου θεωρήθηκε η :

$$L = L_{SR} + L_{degrad}, \text{ όπου } L_{SR} \text{ η } L_1 \text{ απόσταση μεταξύ της εξόδου του DASR και της εικόνας υψηλής ανάλυσης και}$$

$$L_{degrad} \text{ η συνάρτηση σφάλματος που ορίζεται από την (2).}$$

3.1.3.2 Τροποποιήσεις και συνεισφορά

Στην παραπάνω ενότητα αναλύθηκε ο τρόπος εκπαίδευσης και βελτιστοποίησης του μοντέλου που προτάθηκε από τους Wang et al. . Από την ανάλυση αυτή, γίνεται κατανοητό πως το συγκεκριμένο μοντέλο, αν και είναι σε θέση να μάθει τις αναπαραστάσεις των παραμορφώσεων που εμφανίζονται σε ένα σύνολο εικόνων χωρίς την ανάγκη καμίας επίβλεψης, μέσω του δικτύου του κωδικοποιητή και του contrastive learning, για την εκπαίδευση του συνολικού δικτύου ακολουθεί την μεθοδολογία της επιβλεπόμενης μάθησης. Η επιβλεπόμενη φύση της εκπαίδευσης του DASR τμήματος του μοντέλου, ωστόσο, δεν επηρεάζει άμεσα το γενικότερο στόχο της παρούσας εργασίας (που είναι η επίλυση του Super Resolution προβλήματος μέσω μη επιβλεπόμενης μάθησης), καθώς η σημαντικότερη συμβολή του [53] είναι η παρουσίαση μίας νέας μη επιβλεπόμενης μεθοδολογίας για την παραγωγή αναπαραστάσεων παραμόρφωσης οι οποίες στην συνέχεια μπορούν να αξιοποιηθούν από βαθιά νευρωνικά δίκτυα, ανεξάρτητα του τρόπου εκπαίδευσής τους, για την σύνθεση εικόνων αυξημένης χωρικής ανάλυσης. Δεδομένης, λοιπόν, της επιβλεπόμενης φύσης του DASR κομματιού, μπορεί εύκολα κανείς να αντιληφθεί την αναγκαιότητα ύπαρξης ενός συνόλου εκπαίδευσης. Το σύνολο εκπαίδευσης, όπως αναφέρθηκε και προηγουμένως, πρόκειται για μία ανάμιξη μεταξύ των συνόλων DIV2K και Flickr2K. Αναλυτικότερα, χρησιμοποιήθηκαν οι

800 εικόνες υψηλής ανάλυσης του υποσυνόλου εκπαίδευσης του DIV2K και οι αντίστοιχες 2650 εικόνες από το Flickr2K, ενώ το σύνολο ελέγχου αποτελείται από τις δορυφορικές εικόνες οι οποίες εξετάζονται στην παρούσα εργασία.

Η εκπαίδευση του DASR πάνω σε συμβατικές RGB εικόνες με ακρίβεια 8bit, δημιουργεί ορισμένους περιορισμούς στην διαδικασία απόκτησης των δορυφορικών εικόνων υψηλής ανάλυσης, κυρίως ως προς την μορφή και τον τύπο δεδομένων των δορυφορικών εικόνων που δίνονται ως είσοδοι στο δίκτυο αυτό. Αναλυτικότερα, επειδή το DASR αναμένει ως είσοδο τριχρωματικές εικόνες, κρίθηκε αναγκαία η διάσπαση των εικόνων ανάλυσης 10m GSD σε μία τριχρωματική εικόνα που περιέχει τις μπάντες B2,B3,B4 και μία μονοχρωματική εικόνα με την απεικόνιση της B8 μπάντας. Η ίδια διαδικασία ακολουθήθηκε για τις εικόνες ανάλυσης 20m GSD, οι οποίες διασπάστηκαν σε δύο επιμέρους τριχρωματικές εικόνες από τις οποίες η πρώτη περιείχε τις μπάντες B5,B6,B7 και η δεύτερη τις μπάντες B8a, B11,B12. Επίσης, σημειώνεται πως κάθε εικόνα με ανάλυση 60m GSD αντιμετωπίστηκε, στην παρούσα εργασία, ως δύο μονοχρωματικές εικόνες μία με την απεικόνιση της μπάντας B9 και μία για αυτή της B10. Τέλος, σημειώνεται πως όλες οι επιμέρους εικόνες που δημιουργήθηκαν από τις αρχικές δορυφορικές, αποθηκεύτηκαν ως 8bit jpeg αρχεία.

Στο σημείο αυτό κρίνεται σκόπιμο να αναφερθεί πως η διαθέσιμη υλοποίηση περιείχε μοντέλα εκπαιδευμένα για την μεγέθυνση εικόνων κατά παράγοντες 2 και 4, ωστόσο στην παρούσα εργασία εξετάζεται και η μεγέθυνση κατά παράγοντα 6. Στις προηγούμενες υλοποιήσεις, που δεν χρειαζόταν εκπαίδευση του δικτύου υπό την έννοια της επιβλεπόμενης μάθησης, η επιλογή του παράγοντα μεγέθυνσης γινόταν την στιγμή της ανάλυσης της εκάστοτε εικόνας. Στην περίπτωση, όμως, του μοντέλου που προτείνεται από τους Wang et al. αυτή η δυνατότητα δεν είναι διαθέσιμη. Επομένως, γίνεται εύκολα κατανοητό πως για την απόκτηση εικόνων μεγεθυμένων κατά παράγοντα 6, είναι αναγκαία η εκ νέου εκπαίδευση του δικτύου. Η εκ νέου εκπαίδευση του δικτύου συνεπάγεται με την σειρά ορισμένες αλλαγές στο τρόπο σύνθεσης των 32 εικόνων χαμηλής ανάλυσης, καθώς και στην δομή του επιπέδου υπερδειγματοληψίας στο τέλος του DASR τμήματος. Αναλυτικότερα, για την σύνθεση των 32 εικόνων χαμηλής ανάλυσης από τις αντίστοιχες εικόνες υψηλής ανάλυσης χρησιμοποιούνται γκαουσιανοί πυρήνες, η τυπική απόκλιση των οποίων αυτή την φορά κυμαίνεται στο διάστημα [0.2,6.0], ενώ το υπερδειγματοληπτικό επίπεδο ορίζεται κατάλληλα προκειμένου η έξοδός του να αντιστοιχεί σε μεγέθυνση κατά παράγοντα 6.

3.2. Συγκεντρωτική περιγραφή μοντέλων

Στον παρακάτω πίνακα παρουσιάζονται συνοπτικά τα χαρακτηριστικά των μοντέλων που περιγράφηκαν αναλυτικά στις προηγούμενες ενότητες του κεφαλαίου αυτού, καθώς οι προσωπικές τροποποιήσεις και η συνεισφορά του συγγραφέα της παρούσας εργασίας στις υλοποιήσεις των μοντέλων αυτών.

Πίνακας 3: Συνοπτική παρουσίαση μοντέλων και της συνεισφοράς της εργασίας.

Μοντέλο	Χαρακτηριστικά	Συνεισφορά
Deep Image Prior	- Πλήρως Unsupervised. - Παραγωγή SR εικόνας στην φάση ελέγχου. - Κωδικοποίηση στατιστικών εξαρτήσεων των pixel στα βάρη του δικτύου. - Δημιουργία SR εικόνας από θόρυβο. - Αρχιτεκτονική U-Net με Skip Connections. - Αξιοποίηση Skip Connection για μεταφορά χαρακτηριστικών από διαφορετικές κλίμακες.	- Γενίκευση σε εικόνες ακρίβειας 16bit. - Επεξεργασία εικόνων αυθαίρετου πλήθους καναλιών. - Μεταβλητές δικτύου έχουν τύπο δεδομένων float32 . - Υλοποίηση Early Stopping μεθόδου. - Εφαρμογή histogram stretching σε περιπτώσεις μεγάλων ανομοιομορφιών μεταξύ θορύβου εισόδου και υπό μελέτη εικόνας. - Βελτίωση της χωρικής ανάλυσης, σειριακά σε μη επικαλυπτόμενες υποεικόνες, σε περιπτώσεις υπερφόρτωσης της GPU VRAM
Zero-Shot Super Resolution	- Πλήρως Unsupervised. - Αξιοποίηση επανεμφάνσεων τμημάτων στο εσωτερικό μίας εικόνας. - Δημιουργία ζευγών HR-LR εικόνων αποκλειστικά από την υπό εξέταση εικόνα με εφαρμογή του data augmentation. - Σχετικά ρηχό συνελκτικό δίκτυο. - Αξιοποίηση Skip Connections, για επιτάχυνση εκπαίδευσης. - Μεταβλητό(μειούμενο) learning rate.	- Γενίκευση σε εικόνες ακρίβειας 16bit. - Γενίκευση σε εικόνες αυθαίρετου πλήθους καναλιών. - Τροποποίηση εσωτερικών συναρτήσεων για την εκτέλεση data augmentation και σε εικόνες με 6 κανάλια. - Μεταβλητές δικτύου έχουν τύπο δεδομένων float32 .
Degradation Aware Super Resolution	- Σύνθεση δύο επιμέρους δικτύων. - Unsupervised εξαγωγή αναπαραστάσεων παραμόρφωσης, μέσω contrastive learning. - Supervised εκπαίδευση SR υποδικτύου.	- Υλοποίηση και εκπαίδευση μοντέλου για μεγέθυνση κατά παράγοντα 6. - Σύνθεση συνόλου εκπαίδευσης του δικτύου, συνδυάζοντας
	- Αξιοποίηση αναπαραστάσεων παραμόρφωσης για βελτίωση ποιότητας SR αποτελεσμάτων. - Σύνθετες συναρτήσεις σφάλματος. - Σύνθετη αρχιτεκτονική, πολλαπλών blocks. Ακολουθεί την λογική της αρχιτεκτονικής του RCAN. - Ανεξάρτητη εκπαίδευση κωδικοποιητή σε πρώτη φάση. - Μεταβλητό (μειούμενο) learning rate.	- τα σύνολα DIV2K και Flickr2K. - Γενίκευση σε εικόνες αυθαίρετου πλήθους καναλιών, μέσω διάσπασης σε επιμέρους τριχρωματικές και μονοχρωματικές εικόνες. - Συμβιβασμός ως προς την ακρίβεια των αποτελεσμάτων, τα οποία είναι 8bit εικόνες.

4. Αποτελέσματα

Στο κεφάλαιο αυτό παρουσιάζονται οι δορυφορικές εικόνες που χρησιμοποιήθηκαν για την αξιολόγηση των μοντέλων του κεφαλαίου 3, καθώς επίσης και οι παραγόμενες από κάθε μέθοδο δορυφορικές εικόνες υψηλής χωρικής ανάλυσης. Αναλυτικότερα, για κάθε εικόνα του συνόλου ελέγχου παρατίθενται τα αποτελέσματα των τριών μοντέλα για συντελεστές μεγέθυνσης ίσους με 2, 4 και 6, με σκοπό τον εντοπισμό των καλύτερων αποτελεσμάτων για το εκάστοτε μοντέλο μεταξύ των τριών διαφορετικών συντελεστών. Στην συνέχεια, ακολουθεί μία ποιοτική σύγκριση των αποτελεσμάτων ανά συντελεστή μεγέθυνσης για όλα τα μοντέλα, με σκοπό την ανίχνευση του μοντέλου με την καλύτερη συμπεριφορά για τον εκάστοτε συντελεστή μεγέθυνσης. Στο τέλος του συγκεκριμένου κεφαλαίου, παρατίθεται μία ποσοτική αξιολόγηση των αποτελεσμάτων που παράγονται από το εκάστοτε μοντέλο, καθώς και μία συνολική σύγκριση όλων των μοντέλων με σκοπό την ανάδειξη του βέλτιστου μεταξύ των τριών.

4.1. Το σύνολο ελέγχου

Στα πλαίσια της παρούσας εργασίας, για την αξιολόγηση των μοντέλων που παρουσιάστηκαν στο κεφάλαιο 3 χρησιμοποιήθηκαν πέντε δορυφορικές εικόνες που λήφθηκαν από τον αστερισμό δορυφόρων Sentinel-2 οι οποίες απεικονίζουν πέντε διαφορετικές γεωγραφικές περιοχές της Ευρώπης. Αναλυτικότερα, οι εικόνες που χρησιμοποιήθηκαν αποτελούν επιμέρους τμήματα από προϊόντα τύπου 2A, που σημαίνει ότι περιέχουν πληροφορίες σχετικές με την ανακλαστικότητα της ακτινοβολίας στο κάτω μέρος της ατμόσφαιρας της Γης. Οι περιοχές μελέτης που απεικονίζονται στις εικόνες του συνόλου, καθώς και η ημερομηνία λήψης του παρουσιάζονται παρακάτω :

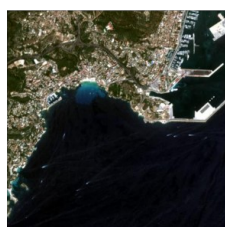
- **Εικόνα 1** : Kylaniemi, Φινλανδία, 10/09/2021
- **Εικόνα 2** : Mallorca, Ισπανία, 15/08/2021
- **Εικόνα 3** : Breda, Ολλανδία, 07/09/2021
- **Εικόνα 4** : Uryna, Λιθουανία, 26/09/2021
- **Εικόνα 5** : Αθήνα, Ελλάδα 27/09/2021

Στο σημείο αυτό τονίζεται πως, για λόγους ευκολίας οπτικοποίησης των προϊόντων του Sentinel-2, παρακάτω παρατίθενται οι εικόνες που αποτελούνται από τις μπάντες B2,B3,B4 που έχουν χωρική ανάλυση 10m GSD και αντιστοιχούν στο ορατό φάσμα της ηλεκτρομαγνητικής ακτινοβολίας.

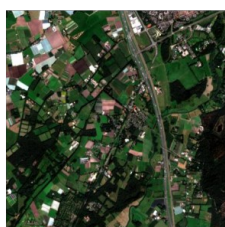
Εικόνα 1 : Φινλανδία



Εικόνα 2 : Ισπανία



Εικόνα 3 : Ολλανδία



Εικόνα 4 : Λιθουανία

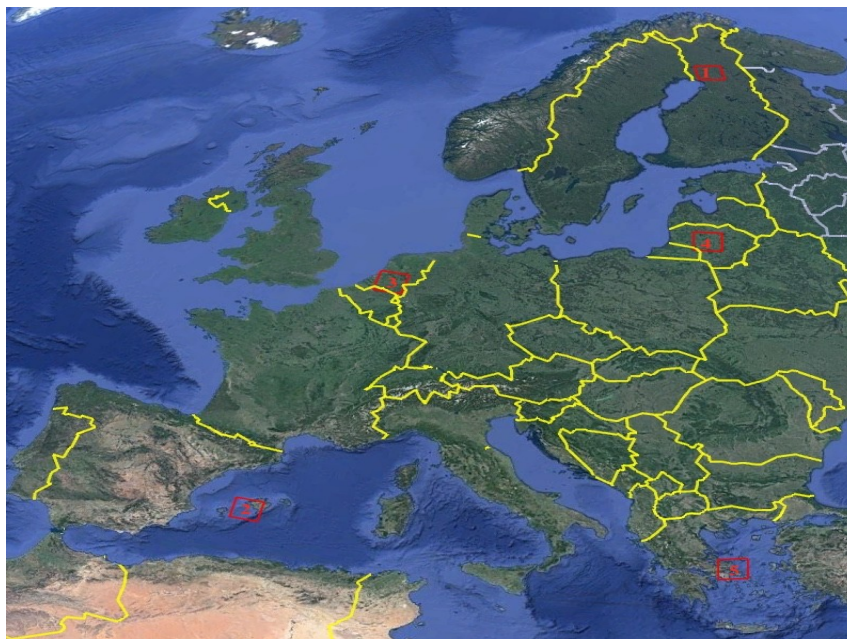


Εικόνα 5 : Ελλάδα



Εικόνα 27: Οι εικόνες του συνόλου ελέγχου.

Επιπλέον, στην εικόνα 28 σημειώνονται με κόκκινα αριθμημένα τετράγωνα οι ευρύτερες περιοχές που απεικονίζονται στα προϊόντα του αστερισμού Sentinel-2 των οποίων επιμέρους τμήματα είναι οι εικόνες που εξετάζονται στην παρούσα εργασία.



Εικόνα 28: Οι ευρύτερες περιοχές των υπό μελέτη εικόνων στο χάρτη.

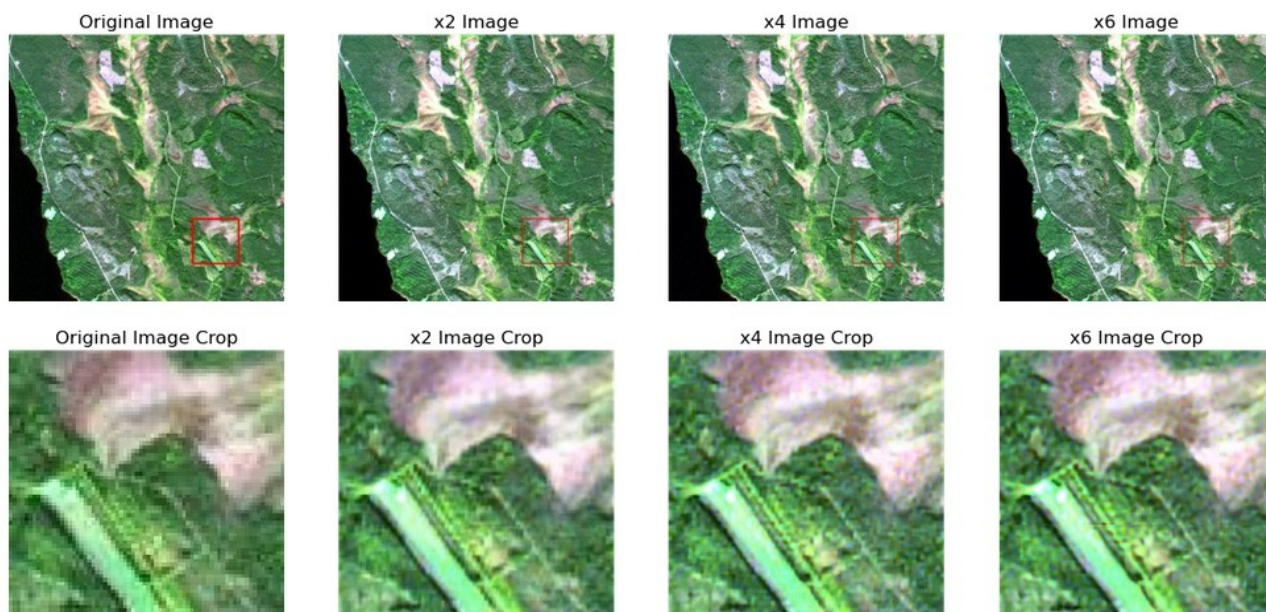
Τέλος, είναι σημαντικό να σημειωθεί πως οι περιοχές μελέτης που επιλέχθηκαν είναι ευθυγραμμισμένες με τις ανάγκες του ερευνητικού έργου EIFFEL (European Union's Horizon 2020 research and innovation programme under Grant Agreement No 101003518), <https://www.eiffel4climate.eu/>, και τα αποτελέσματα της συγκεκριμένης διπλωματικής εργασίας συμβάλλουν στις δραστηριότητες του έργου.

4.2. Αποτελέσματα του Deep Image Prior

Στην ενότητα αυτή πραγματοποιείται μία παρουσίαση των αποτελεσμάτων που παράχθηκαν από το μοντέλο Deep Image Prior (DIP) των Ulyanov et al. . Αναλυτικότερα, για κάθε μία από τις εικόνες που παρουσιάστηκαν παραπάνω, εφαρμόστηκε ο αλγόριθμος Super Resolution της συγκεκριμένης μεθόδου, για την παραγωγή εικόνων μεγεθυμένων κατά συντελεστές $\times 2$, $\times 4$, $\times 6$. Στην συνέχεια, για κάθε μία από τις εικόνες ελέγχου παρουσιάζονται οι αρχική εικόνα χαμηλότερης χωρικής ανάλυσης, οι μεγεθυμένες εικόνες κατά τους προαναφερθέντες παράγοντες, καθώς και μία περικοπή (σημειωμένη με κόκκινο τετράγωνο) τόσο της αρχικής όσο και των μεγεθυμένων εικόνων για διευκόλυνση της ποιοτικής σύγκρισης. Στο σημείο αυτό σημειώνεται πως για την περίπτωση των εικόνων με χωρική ανάλυση 10m GSD η περικεκομμένη εικόνα έχει διαστάσεις 70x70, ενώ για την περίπτωση των εικόνων με χωρική ανάλυση 20m και 60m GSD η περικεκομμένη εικόνα έχει διαστάσεις 100x100 και 120x120 αντίστοιχα. Επιπλέον, αναφέρεται πως για τις εικόνες με χωρική ανάλυση 10m και 20m GSD παρατίθενται μόνο τα τρία πρώτα χρωματικά τους κανάλια, ενώ για τις εικόνες 60m GSD παρατίθεται μόνο το δεύτερο από τα δύο χρωματικά κανάλια ως grayscale εικόνα.

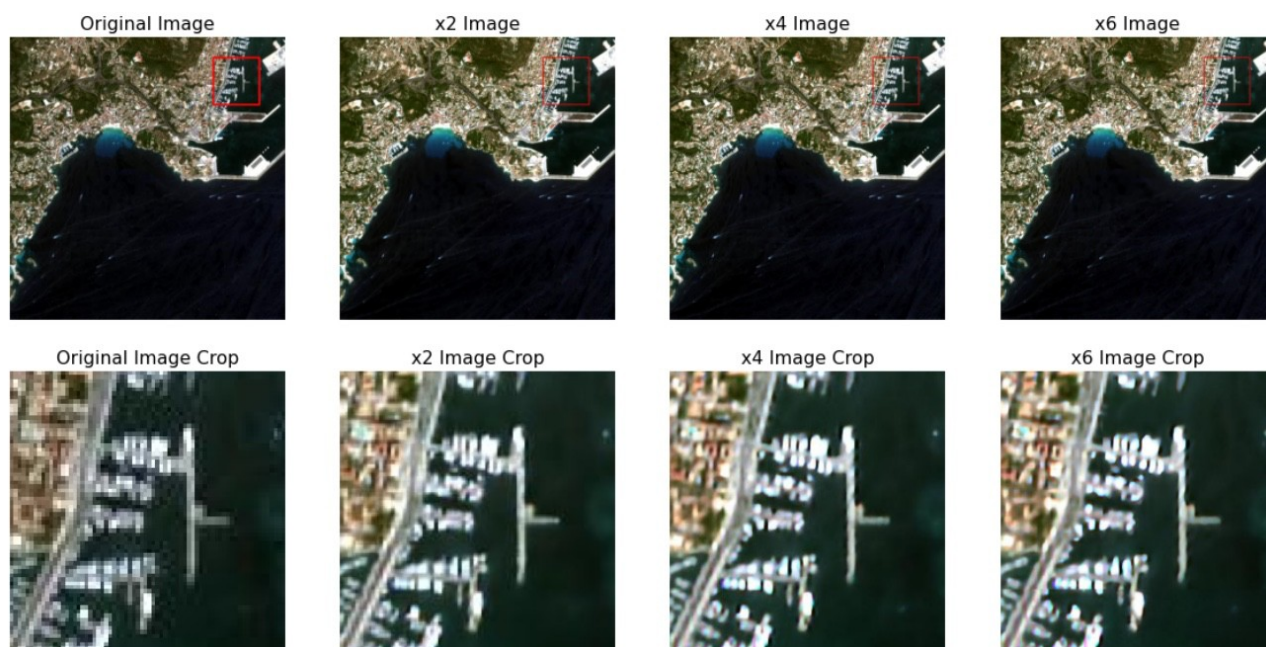
4.2.1 Εικόνες με χωρική ανάλυση 10m GSD

Εικόνα 1



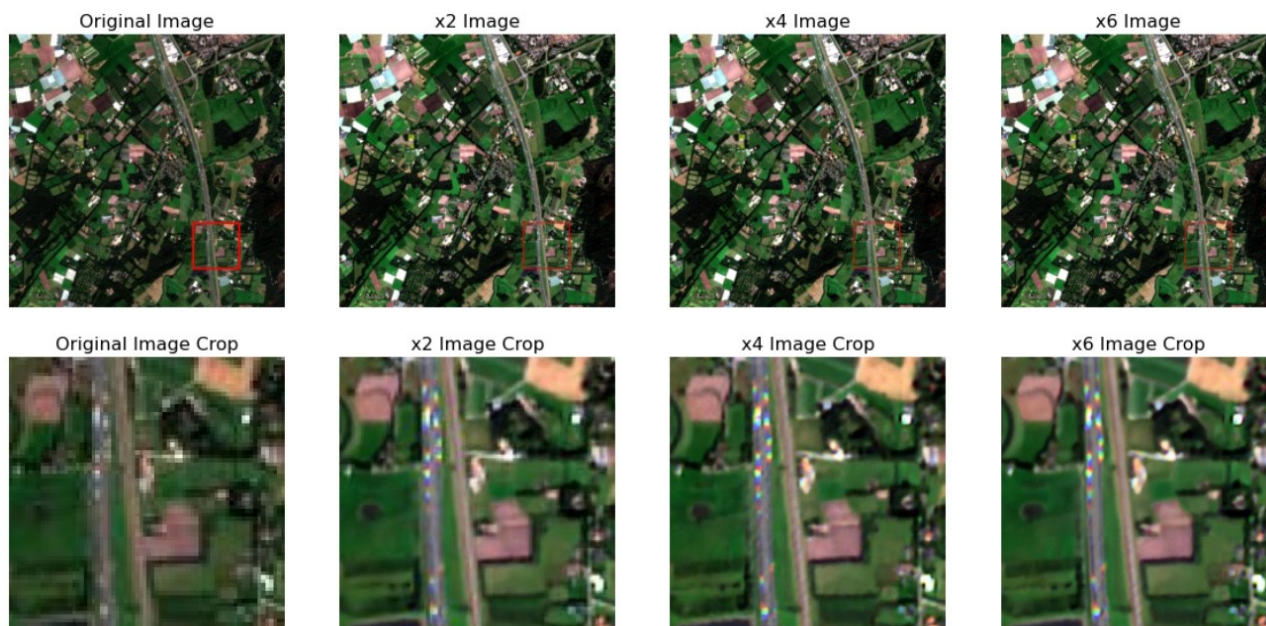
Εικόνα 29: Μεγέθυνση των 10m GSD καναλιών της 1ης εικόνας από το μοντέλο DIP κατά παράγοντες 2,4,6.

Εικόνα 2



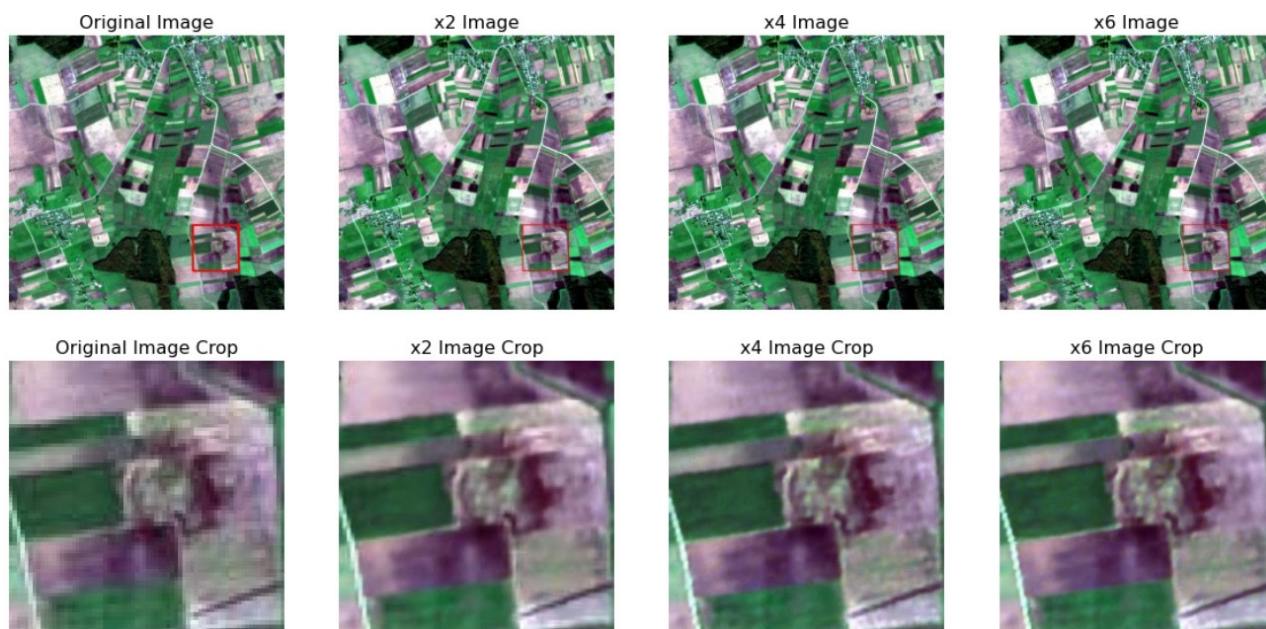
Εικόνα 30: Μεγέθυνση των 10m GSD καναλιών της 2ης εικόνας από το μοντέλο DIP κατά παράγοντες 2,4,6.

Εικόνα 3



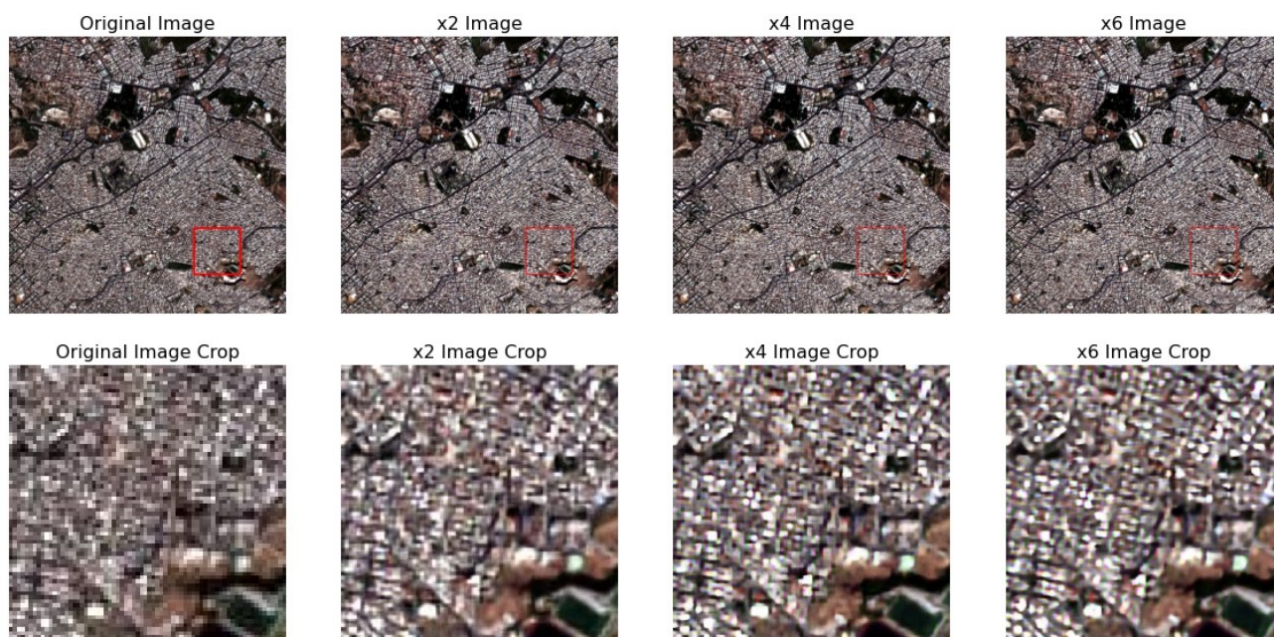
Εικόνα 31: Μεγέθυνση των 10m GSD καναλιών της 3ης εικόνας από το μοντέλο DIP κατά παράγοντες 2,4,6.

Εικόνα 4



Εικόνα 32: Μεγέθυνση των 10m GSD καναλιών της 4ης εικόνας από το μοντέλο DIP κατά παράγοντες 2,4,6.

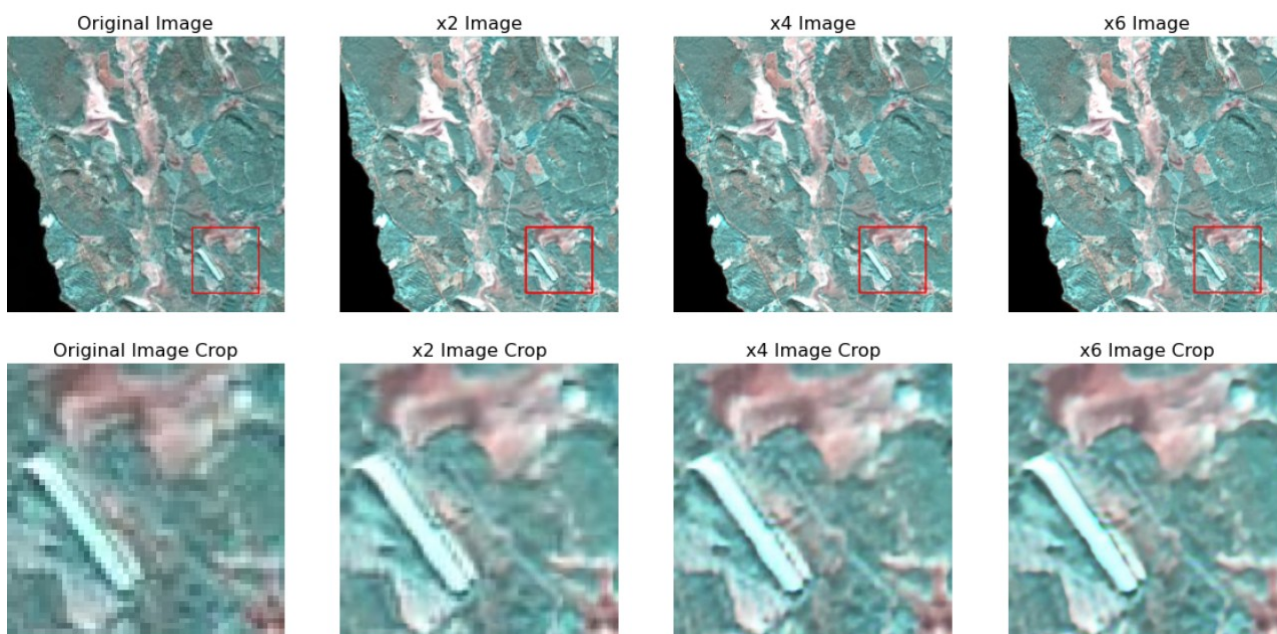
Εικόνα 5



Εικόνα 33: Μεγέθυνση των 10m GSD καναλιών της 5ης εικόνας από το μοντέλο DIP κατά παράγοντες 2,4,6.

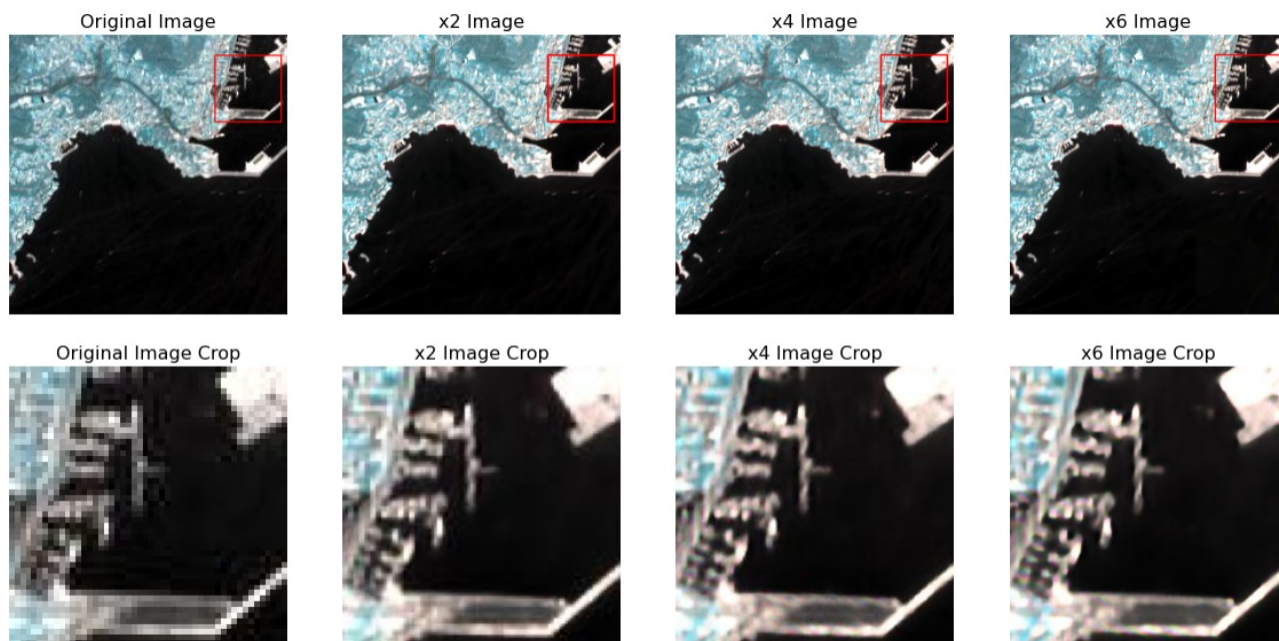
4.2.2 Εικόνες με χωρική ανάλυση 20m GSD

Εικόνα 1



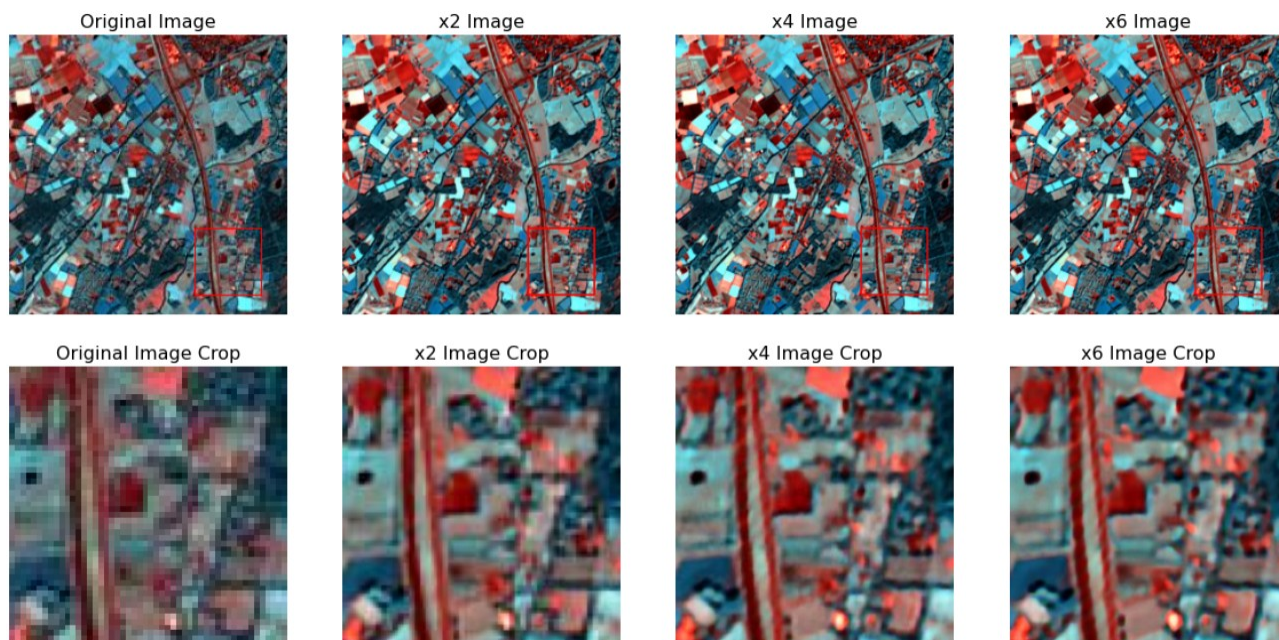
Εικόνα 34: Μεγέθυνση των 20m GSD καναλιών της 1ης εικόνας από το μοντέλο DIP κατά παράγοντες 2,4,6.

Εικόνα 2



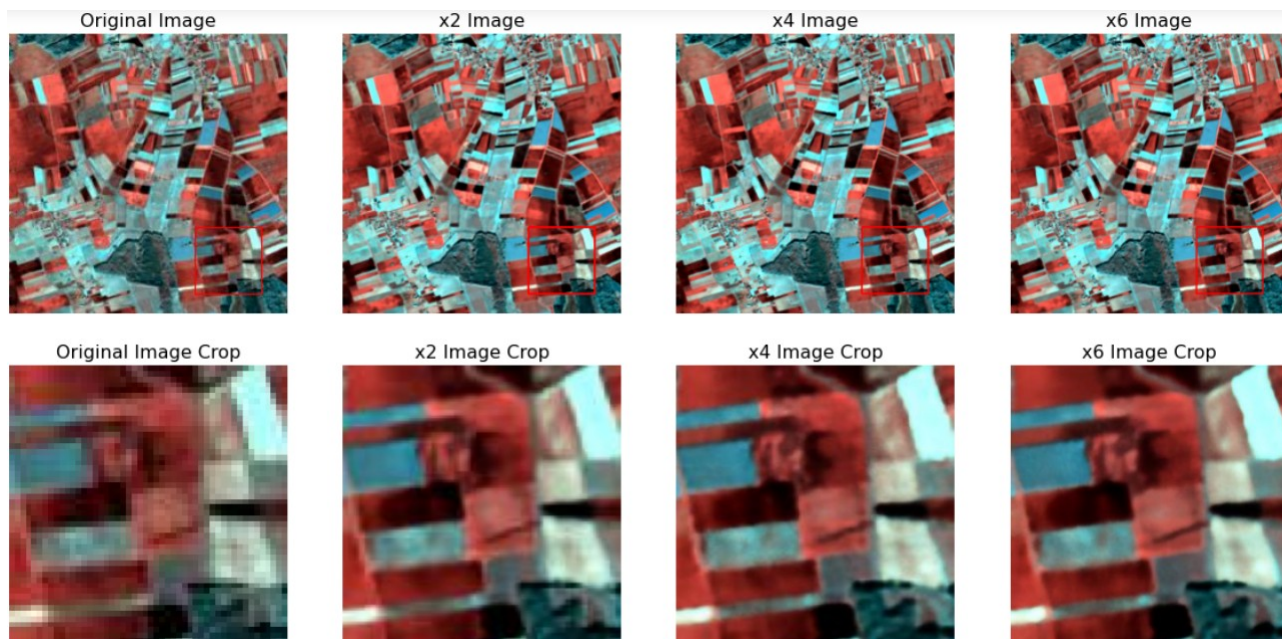
Εικόνα 35: Μεγέθυνση των 20m GSD καναλιών της 2ης εικόνας από το μοντέλο DIP κατά παράγοντες 2,4,6.

Εικόνα 3



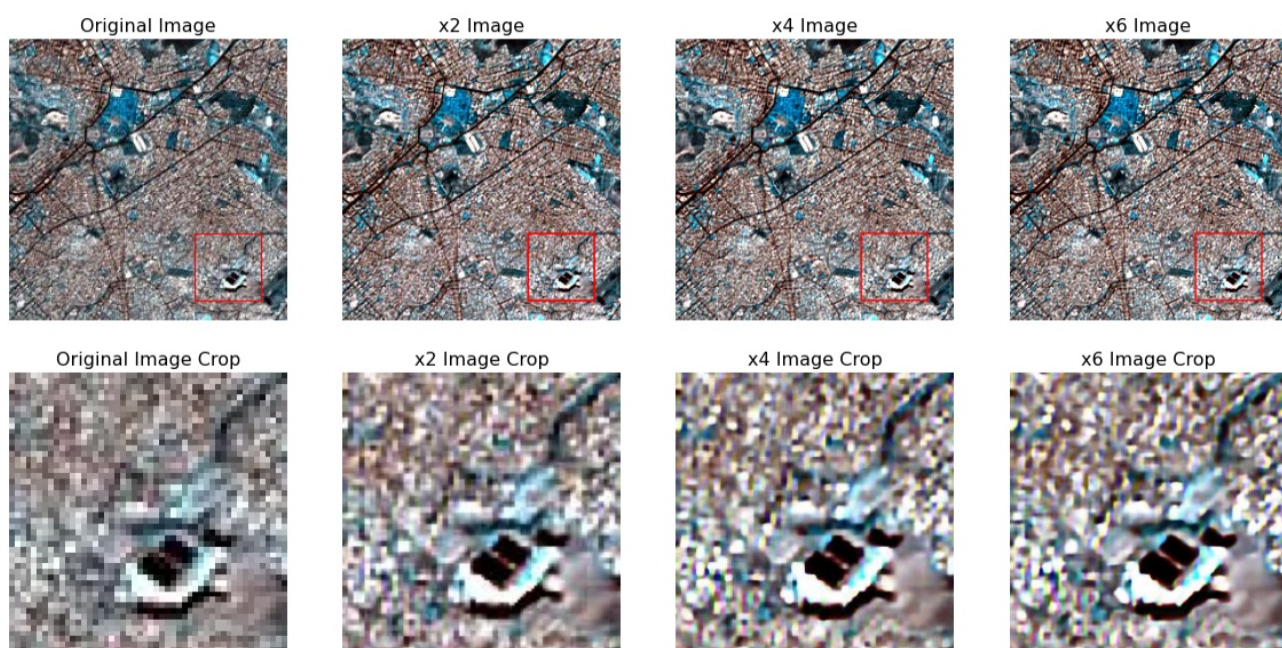
Εικόνα 36: Μεγέθυνση των 20m GSD καναλιών της 3ης εικόνας από το μοντέλο DIP κατά παράγοντες 2,4,6.

Εικόνα 4



Εικόνα 37: Μεγέθυνση των 20m GSD καναλιών της 4ης εικόνας από το μοντέλο DIP κατά παράγοντες 2,4,6.

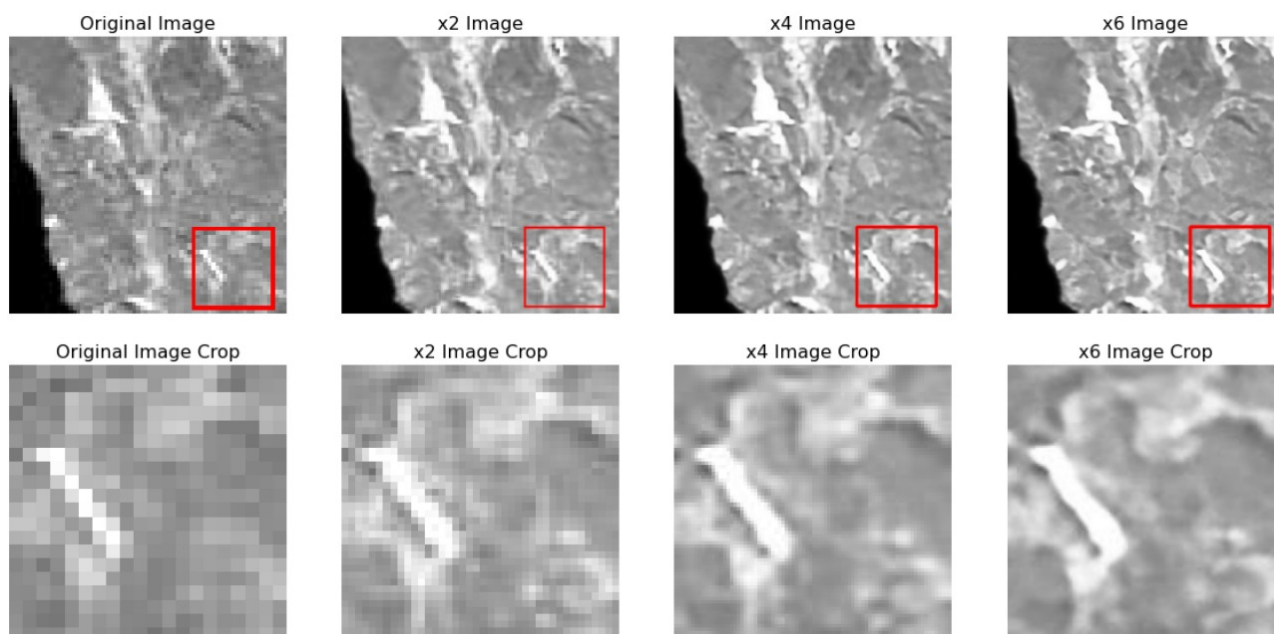
Εικόνα 5



Εικόνα 38: Μεγέθυνση των 20m GSD καναλιών της 5ης εικόνας από το μοντέλο DIP κατά παράγοντες 2,4,6.

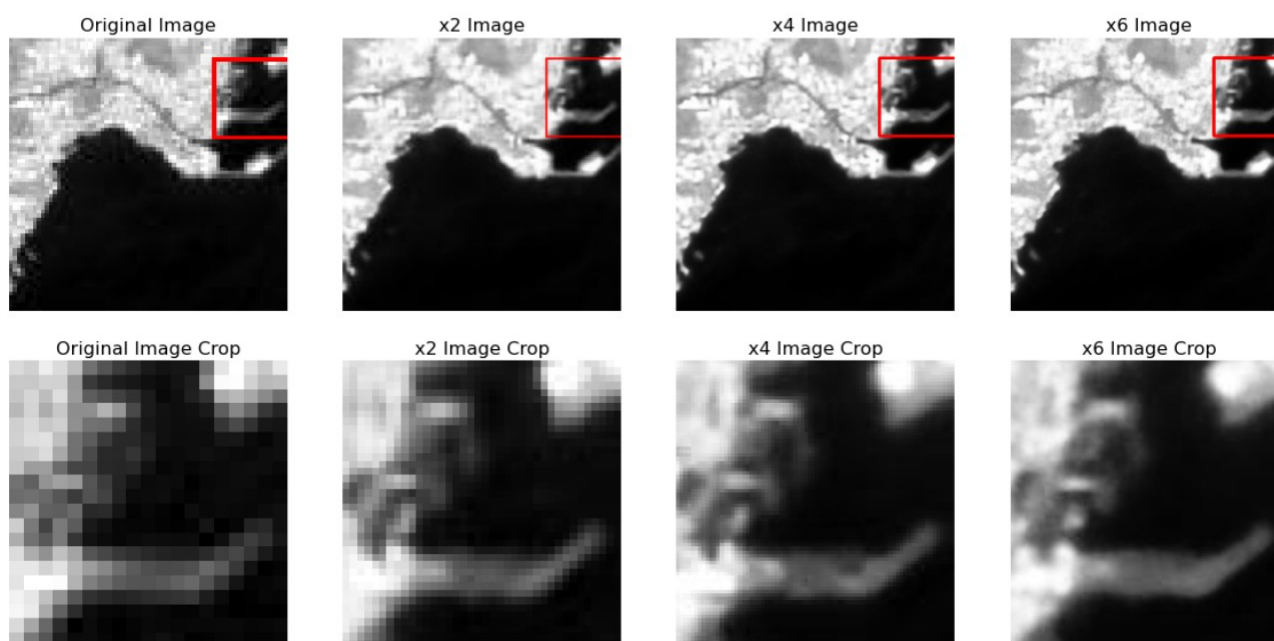
4.2.3 Εικόνες με χωρική ανάλυση 60m GSD

Εικόνα 1



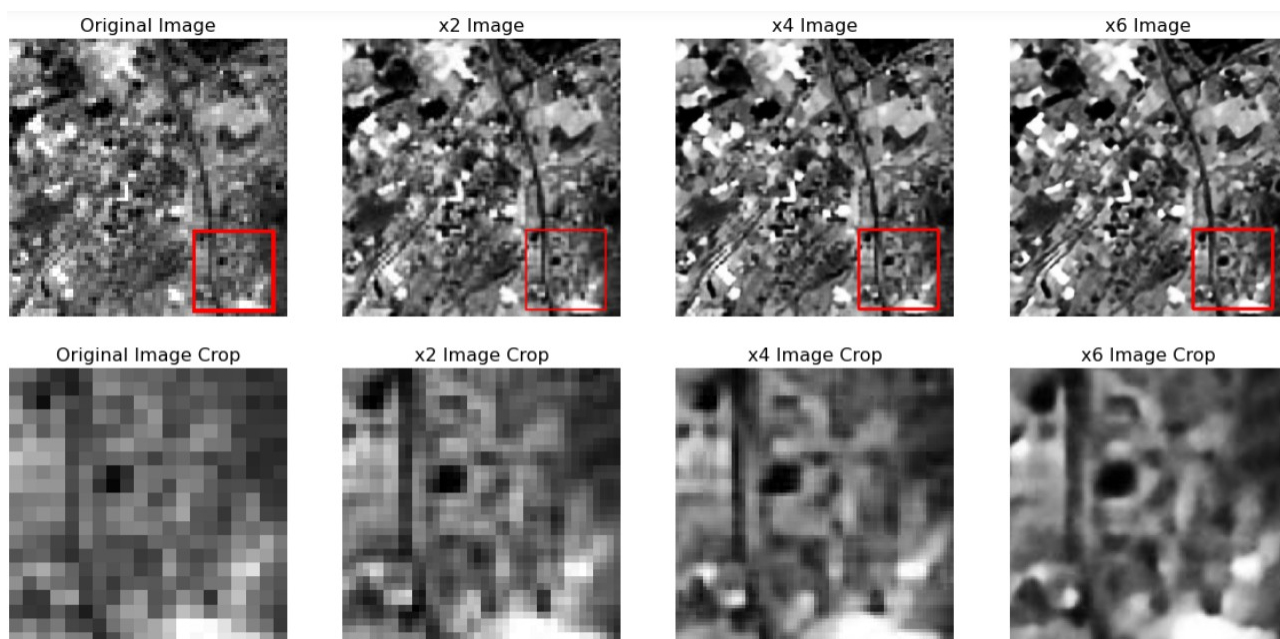
Εικόνα 39: Μεγέθυνση των 60m GSD καναλιών της 1ης εικόνας από το μοντέλο DIP κατά παράγοντες 2,4,6.

Εικόνα 2



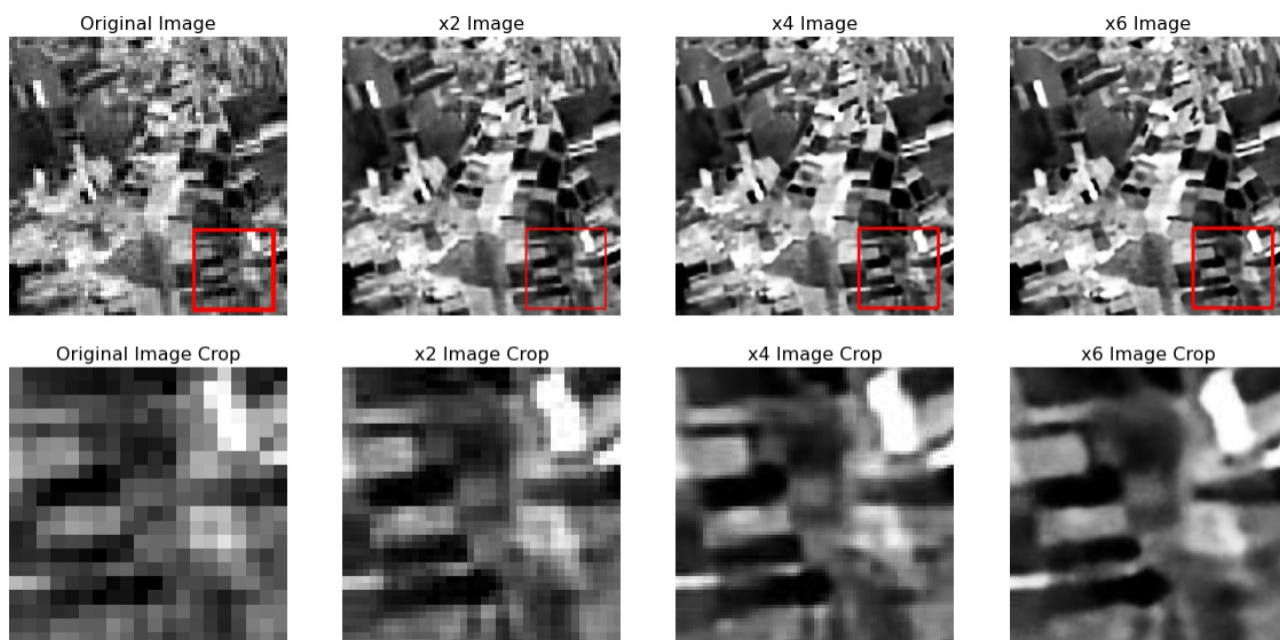
Εικόνα 40: Μεγέθυνση των 60m GSD καναλιών της 2ης εικόνας από το μοντέλο DIP κατά παράγοντες 2,4,6.

Εικόνα 3



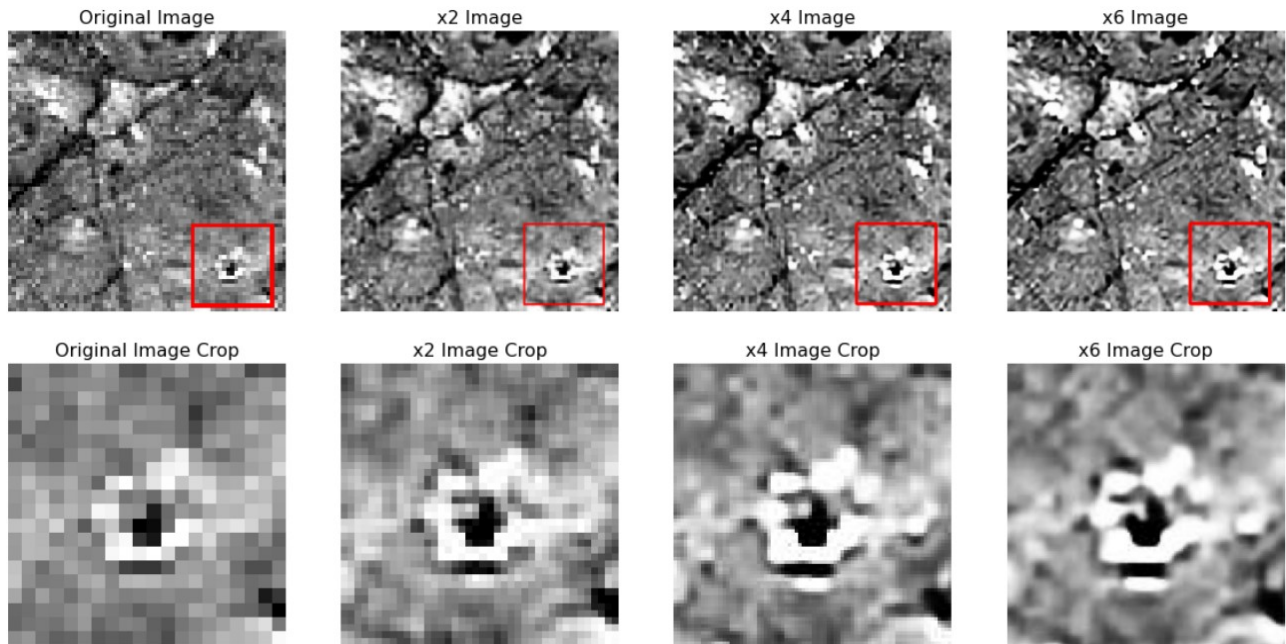
Εικόνα 41: Μεγέθυνση των 60m GSD καναλιών της 3ης εικόνας από το μοντέλο DIP κατά παράγοντες 2,4,6.

Εικόνα 4



Εικόνα 42: Μεγέθυνση των 60m GSD καναλιών της 4ης εικόνας από το μοντέλο DIP κατά παράγοντες 2,4,6.

Εικόνα 5



Εικόνα 43: Μεγέθυνση των 60m GSD καναλιών της 5ης εικόνας από το μοντέλο DIP κατά παράγοντες 2,4,6.

4.2.4 Σχολιασμός Αποτελεσμάτων

Αναπόσπαστο μέρος της παρουσίασης των αποτελεσμάτων, μίας μεθοδολογίας ή ενός αλγορίθμου, αποτελεί η ανάλυση και ο σχολιασμός τους. Αρχικά, λοιπόν, με μία γρήγορη επισκόπηση των εικόνων υψηλής ανάλυσης, και κυρίως των περικοπών αυτών, που παράγονται από το Deep Image Prior (DIP), μπορεί κανείς να επιβεβαιώσει την ικανότητα του μοντέλου να εντοπίζει τις εξαρτήσεις μεταξύ γειτονικών pixels στο εσωτερικό μία εικόνας και να δημιουργεί αποτελέσματα τα οποία είναι, οπτικά, βελτιωμένης ποιότητας. Αναλυτικότερα, σημειώνεται πως, ήδη για συντελεστή μεγέθυνσης ίσο με δύο, το φαινόμενο του pixelation μειώνεται σε σημαντικό βαθμό, ενώ καθώς αυξάνεται η τιμή αυτού το συγκεκριμένο φαινόμενο τείνει να εξαφανιστεί πλήρως. Επιπρόσθετα, στις μεγεθυμένες εικόνες, οι χωρικές δομές, που περιέχονται στην αντίστοιχη εικόνα χαμηλής ανάλυσης, απεικονίζονται με μεγαλύτερη ακρίβεια, με τα περιγράμματα και τα όρια αυτών να φαίνονται αρκετά πιο ξεκάθαρα και ομαλά. Μία ακόμα σημαντική παρατήρηση, είναι το γεγονός πως το μοντέλο είναι σε θέση να εξάγει πληροφορίες σχετικές με την υφή των υποπεριοχών της εκάστοτε εικόνας και να τις χρησιμοποιεί στην συνέχεια για την σύνθεση των νέων εικόνων, οι οποίες εκτός από πιο ξεκάθαρα περιγράμματα περιέχουν και περισσότερες λεπτομέρειες σχετικά με την υφή των απεικονιζόμενων χωρικών δομών. Τέλος, είναι ιδιαίτερα σημαντικό να αναφερθεί πως, παρόλο που το μοντέλο δημιουργεί τις εικόνες υψηλής ανάλυσης από τυχαίο θόρυβο, σε κανένα από τελικά αποτελέσματα δεν εντοπίζονται παραμορφώσεις των χωρικών δομών ή τεχνουργήματα (artifacts).

Στην περίπτωση των εικόνων χωρικής ανάλυσης 10m GSD αξίζει να σημειωθεί πως, αν και υπάρχει βελτίωση της ποιότητας των αποτελεσμάτων μεταξύ των συντελεστών μεγέθυνσης 4 και 6, η διαφορά αυτή δεν γίνεται εύκολα αισθητή με μία γρήγορη επισκόπηση των εκάστοτε δύο εικόνων. Ωστόσο, η παρατήρηση αυτή παύει να ισχύει καθώς μειώνεται η χωρική ανάλυση της υπό εξέταση εικόνας χαμηλής ανάλυσης. Αναλυτικότερα, στις περιπτώσεις των εικόνων με 20m, 60m GSD χωρική ανάλυση, καθώς αυξάνεται ο συντελεστής μεγέθυνσης τα

αποτελέσματα του DIP μοντέλου γίνονται συνεχώς, οπτικά, καλύτερα, ακολουθώντας τις γενικές παρατηρήσεις που πραγματοποιήθηκαν στην προηγούμενη παράγραφο. Εντυπωσιακά είναι τα αποτελέσματα στην περίπτωση των εικόνων χωρικής ανάλυσης 60m GSD, στα οποία εξετάζοντας κανείς την περικεκομμένη εικόνα είναι σε θέση να ανιχνεύσει μόνο διάσπαρτα ρίχει τα οποία αμυδρά συνθέτουν κάποια χωρική δομή. Αντιθέτως, εξετάζοντας τις υπο-εικόνες που εξάγονται από τις εικόνες αυξημένης χωρικής ανάλυσης, καθίσταται εφικτή, με την αύξηση του συντελεστή μεγέθυνσης, η ανίχνευση όλο και περισσότερων χωρικών δομών που προηγουμένως ήταν πρακτικά αδύνατο να αναγνωριστούν.

Στο σημείο αυτό, κρίνεται σκόπιμο να αναφερθεί πως το συγκεκριμένο μοντέλο αποδίδει πολύ καλά σε εικόνες όπου υπάρχουν μεγάλες επιφάνειες και οι ακμές ή οι διαχωριστές καμπύλες μεταξύ των χωρικών δομών που απεικονίζονται είναι σχετικά αραιές, όπως παραδείγματος χάριν συμβαίνει σε όλο το εύρος των εικόνων 1,3 και 4. Σε περιπτώσεις εικόνων που περιέχουν συχνές και πολλαπλές χωρικές δομές σε σχετικά μικρή απόσταση μεταξύ τους, όπως συμβαίνει στις εικόνες 2 και 5, η ποιότητα των αποτελεσμάτων του DIP μοντέλου μειώνεται σε κάποιο βαθμό, χωρίς ωστόσο να καθιστά τις παραγόμενες εικόνες κακής ποιότητας ή ακατάλληλες για χρήση σε λειψές εφαρμογές.

Ολοκληρώνοντας τον σχολιασμό των εικόνων που δημιουργήθηκαν από το DIP μοντέλο, αναφέρονται λεπτομέρειες σχετικά με το χρόνο που απαιτείται για την δημιουργία των εικόνων υψηλής χωρικής ανάλυσης ανά κατηγορία εικόνων χαμηλής ανάλυσης (10m, 20m, 60m GSD) και ανά συντελεστή μεγέθυνσης. Οι πληροφορίες αυτές παρατίθενται υπό την μορφή του παρακάτω πίνακα.

Πίνακας 4: Μέσοι χρόνοι εκτέλεσης για το DIP μοντέλο .

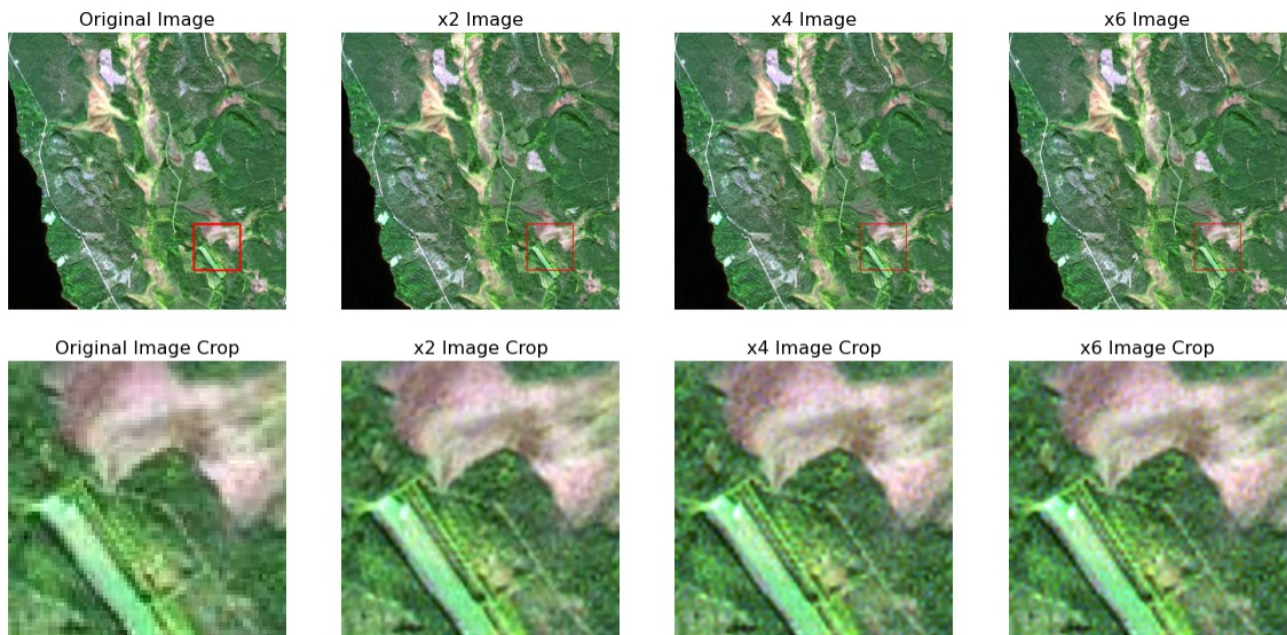
Συντελεστής μεγέθυνσης / Ανάλυση εικόνας	10m	20m	60m
x2	16.82 min	4.49 min	0.78 min
x4	60.89 min	12.43 min	2.03 min
x6	113.85 min	25.21 min	4.02 min

4.3. Αποτελέσματα Zero-Shot Super Resolution

Στην ενότητα αυτή παρουσιάζονται οι εικόνες υψηλής ανάλυσης που δημιουργήθηκαν από το ZSSR μοντέλο. Ο τρόπος παράθεσης των αποτελεσμάτων είναι ακριβώς ίδιος με αυτόν της προηγούμενης ενότητας. Επιπλέον, στο τέλος της ενότητας αυτής παρατίθεται ο σχολιασμός των αποτελεσμάτων.

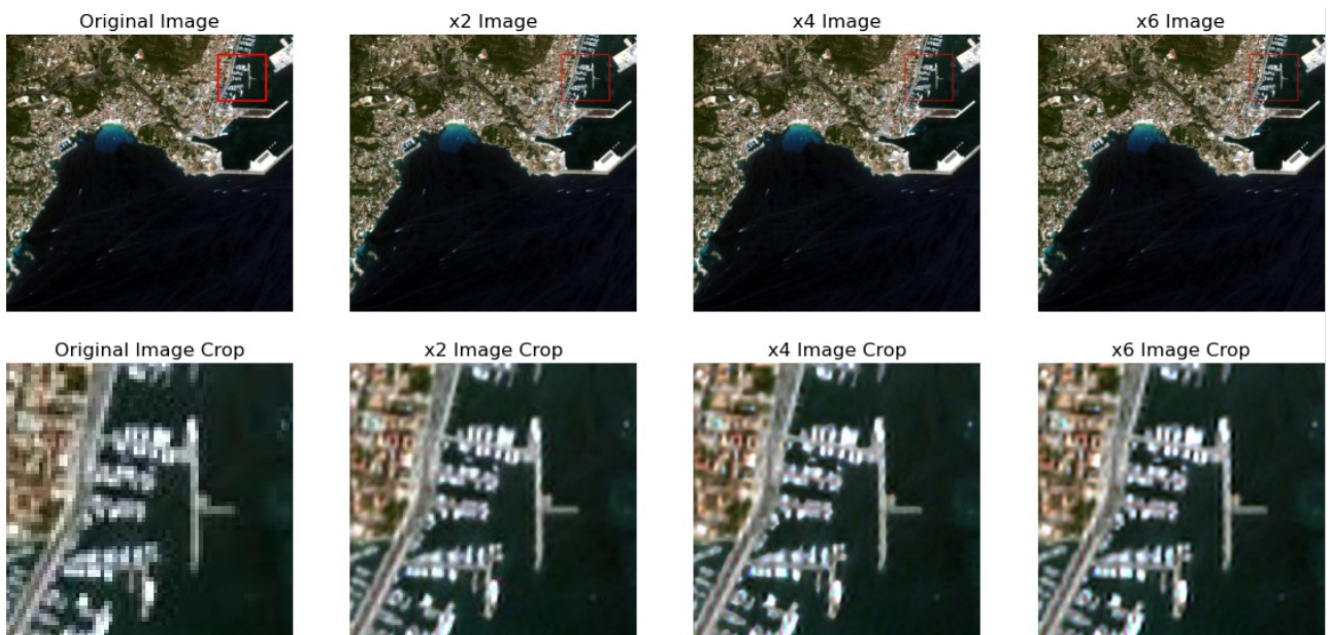
4.3.1 Εικόνες με χωρική ανάλυση 10m GSD

Εικόνα 1



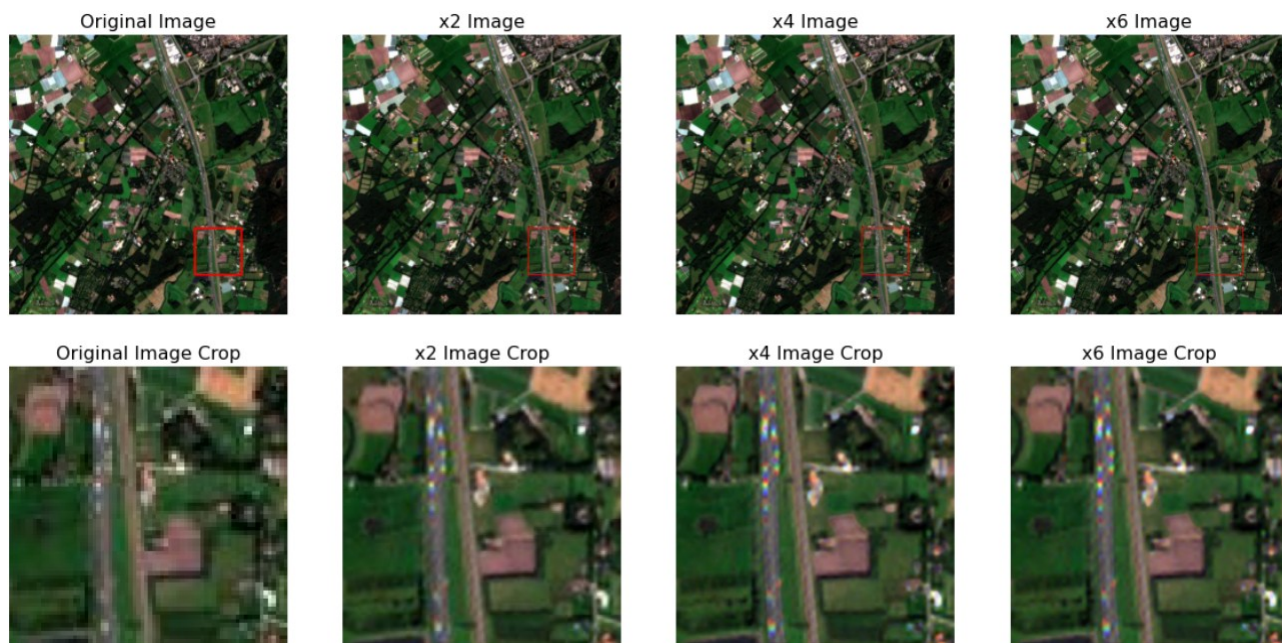
Εικόνα 44: Μεγέθυνση των 10m GSD καναλιών της 1ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.

Εικόνα 2



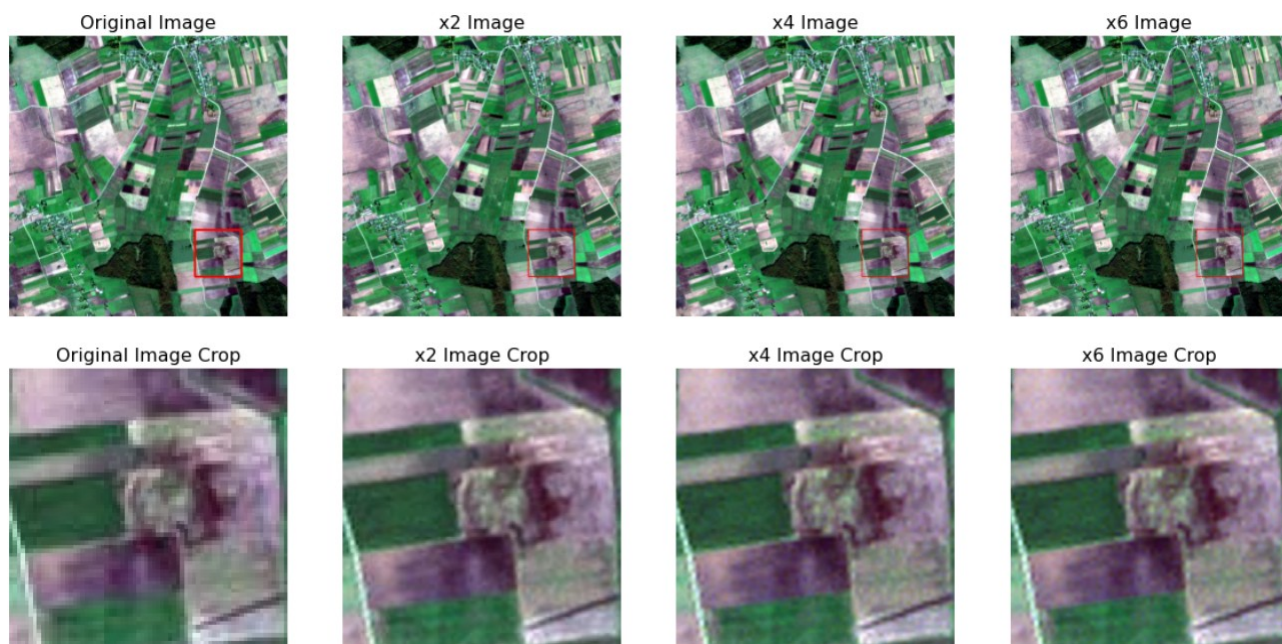
Εικόνα 45: Μεγέθυνση των 10m GSD καναλιών της 2ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.

Εικόνα 3



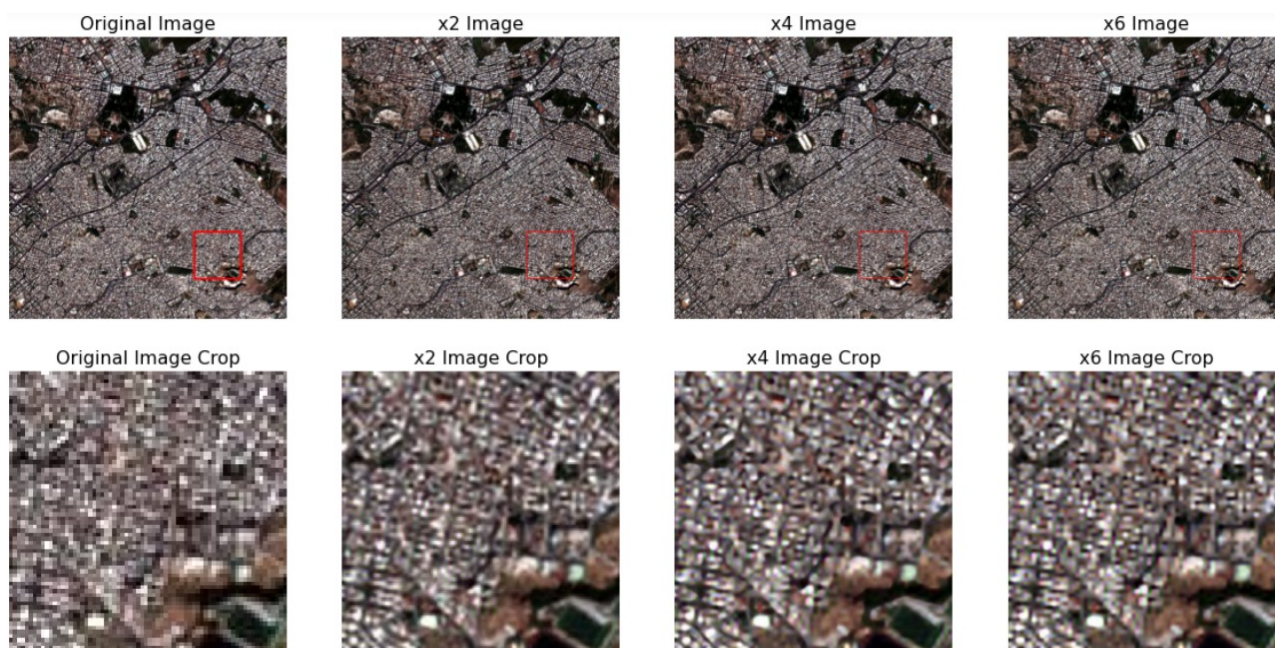
Εικόνα 46: Μεγέθυνση των 10m GSD καναλιών της 3ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.

Εικόνα 4



Εικόνα 47: Μεγέθυνση των 10m GSD καναλιών της 4ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.

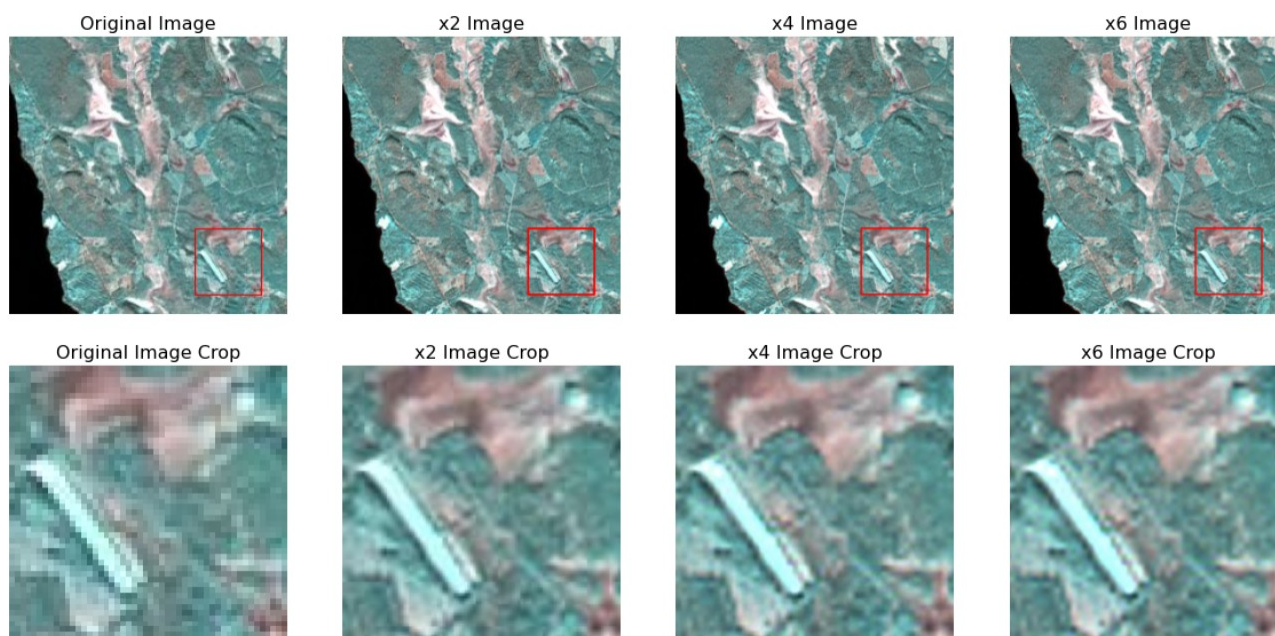
Εικόνα 5



Εικόνα 48: Μεγέθυνση των 10m GSD καναλιών της 5ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.

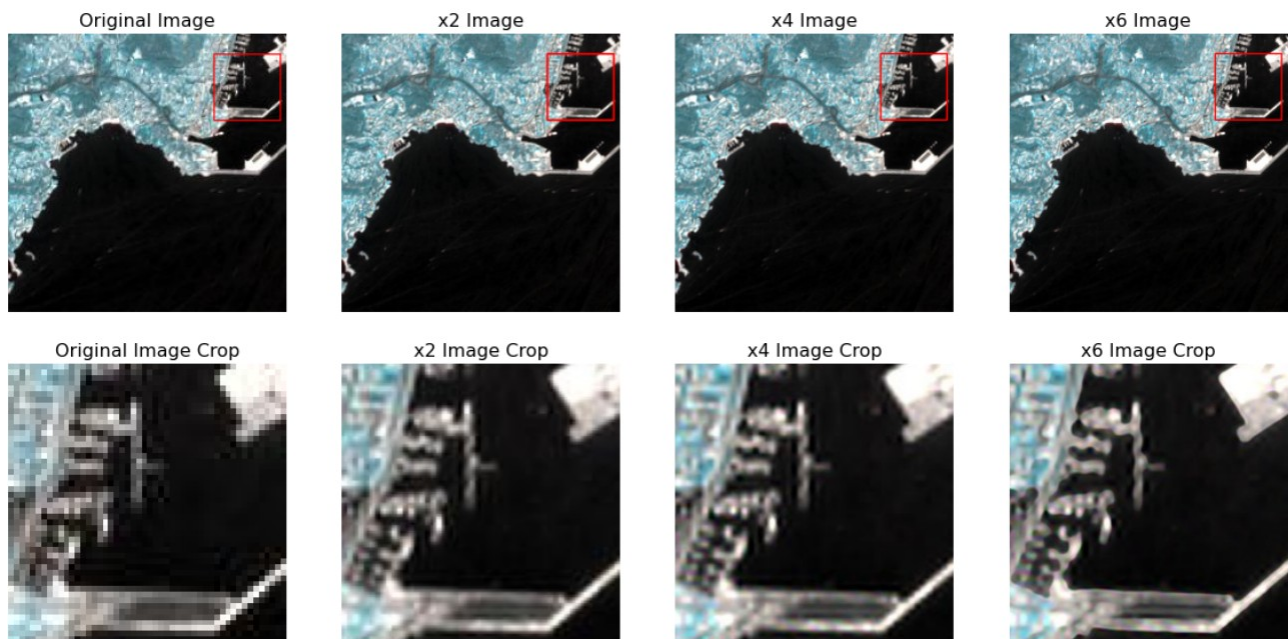
4.3.2 Εικόνες με χωρική ανάλυση 20m GSD

Εικόνα 1



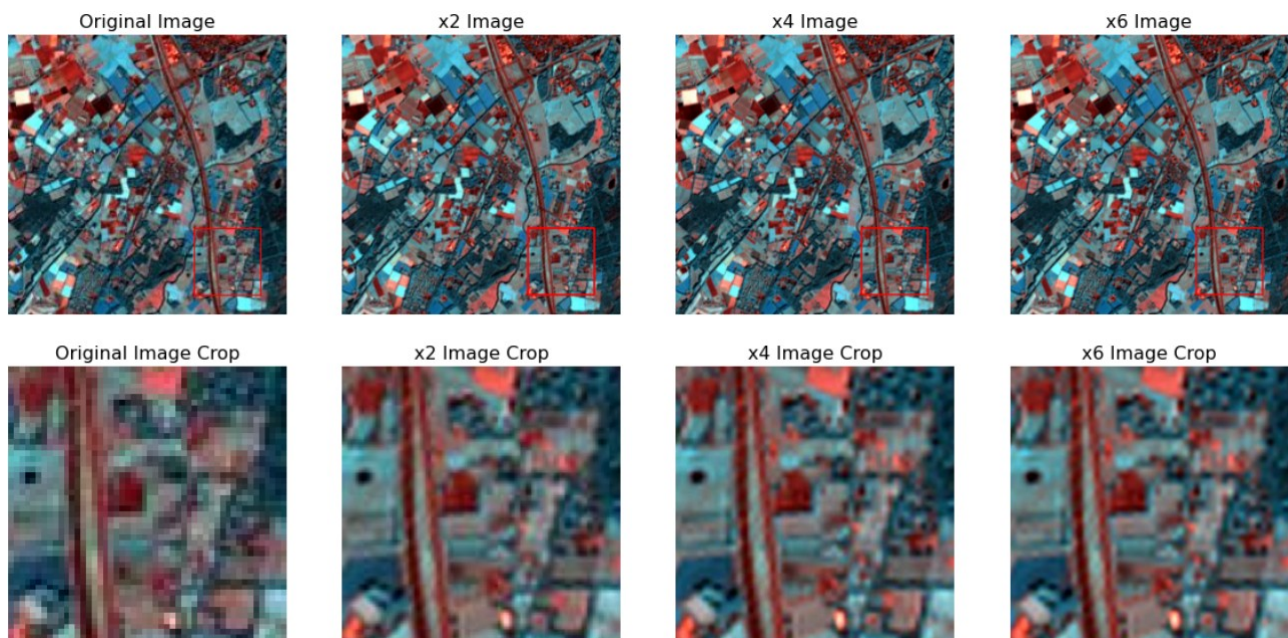
Εικόνα 49: Μεγέθυνση των 20m GSD καναλιών της 1ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.

Εικόνα 2



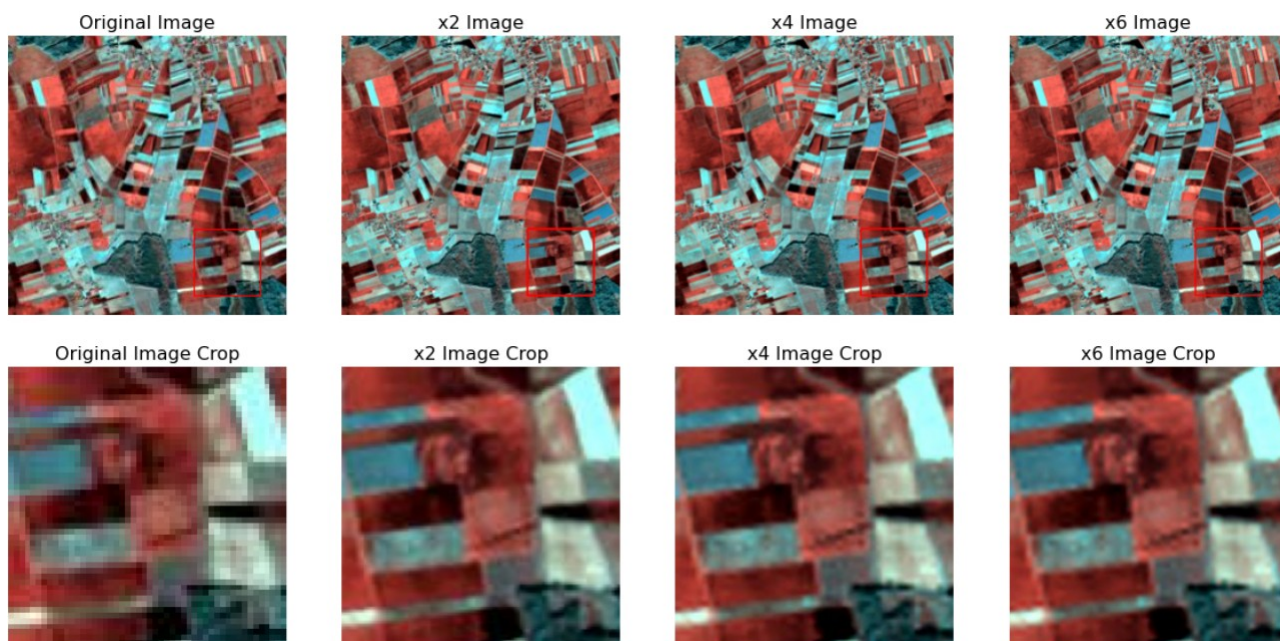
Εικόνα 50: Μεγέθυνση των 20m GSD καναλιών της 2ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.

Εικόνα 3



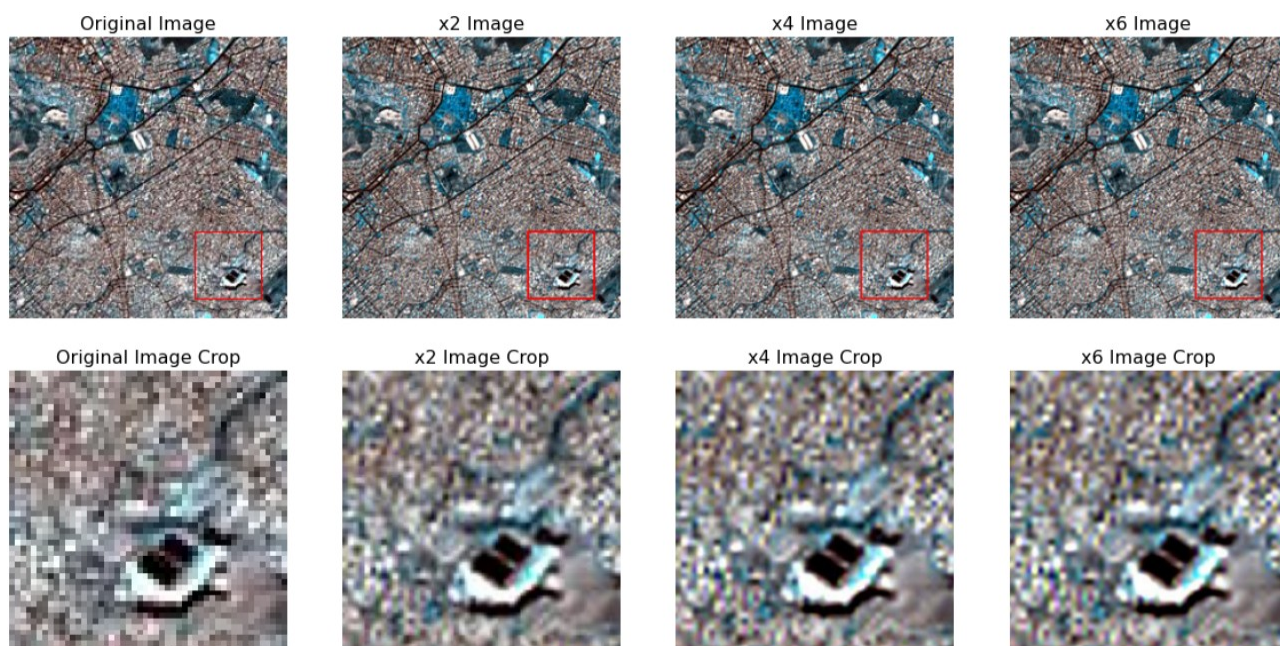
Εικόνα 51: Μεγέθυνση των 20m GSD καναλιών της 3ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.

Εικόνα 4



Εικόνα 52: Μεγέθυνση των 20m GSD καναλιών της 4ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.

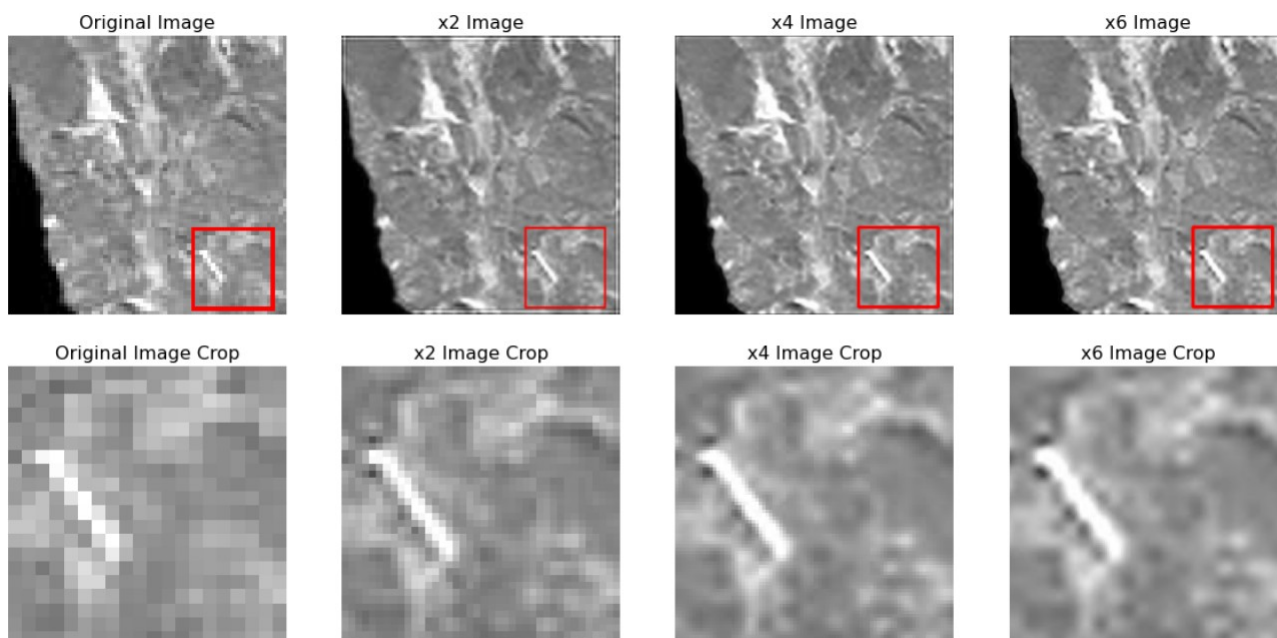
Εικόνα 5



Εικόνα 53: Μεγέθυνση των 20m GSD καναλιών της 5ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.

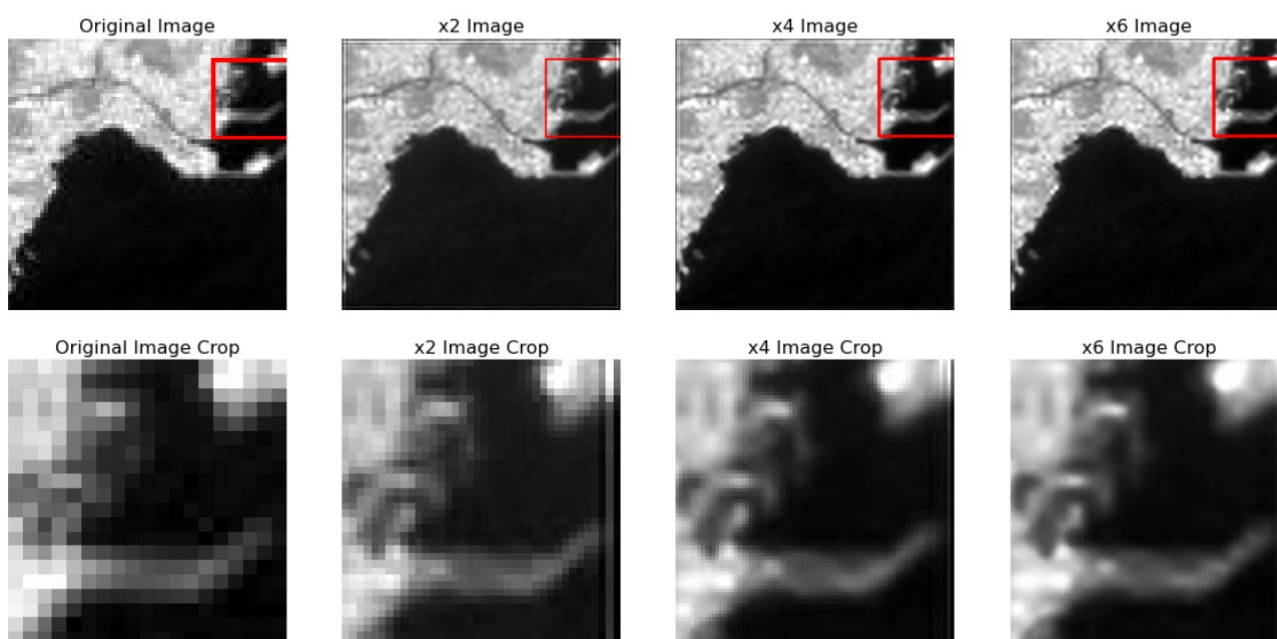
4.3.3 Εικόνες με χωρική ανάλυση 60m GSD

Εικόνα 1



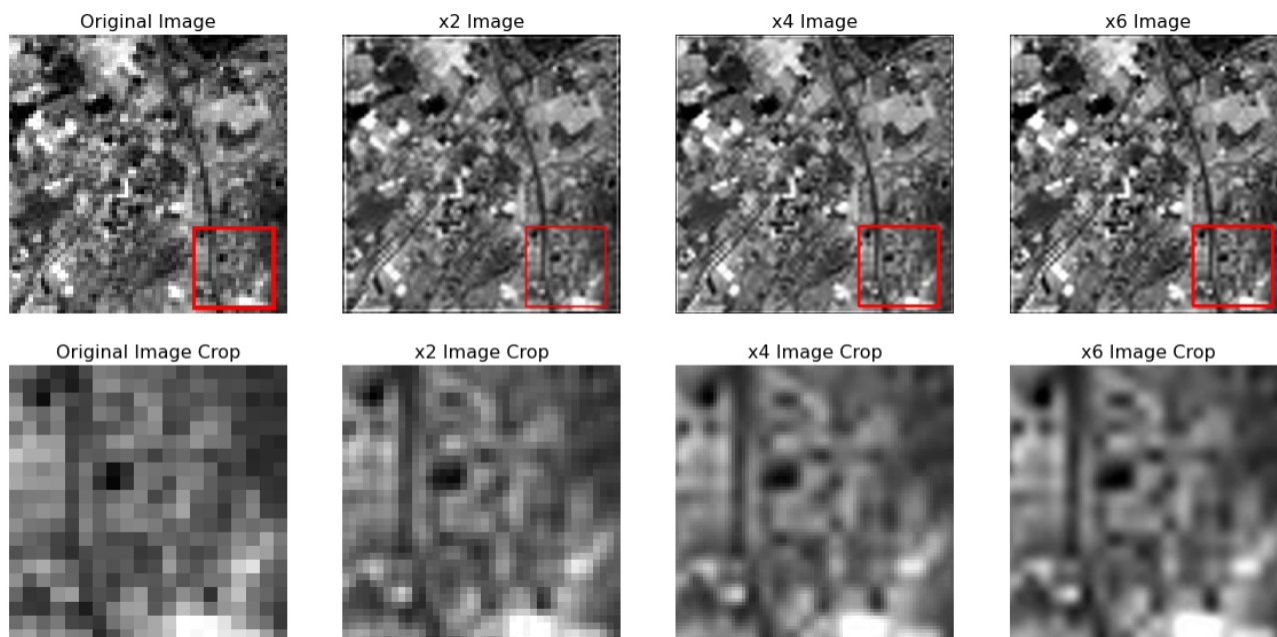
Εικόνα 54: Μεγέθυνση των 60m GSD καναλιών της 1ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.

Εικόνα 2



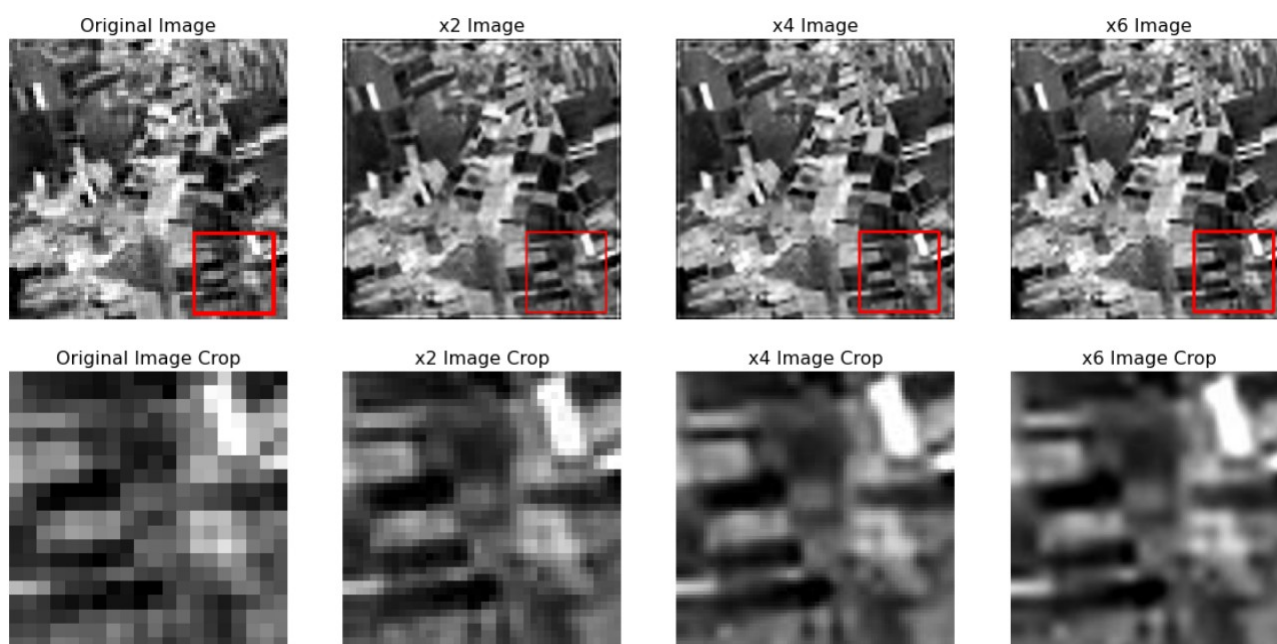
Εικόνα 55: Μεγέθυνση των 60m GSD καναλιών της 2ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.

Εικόνα 3



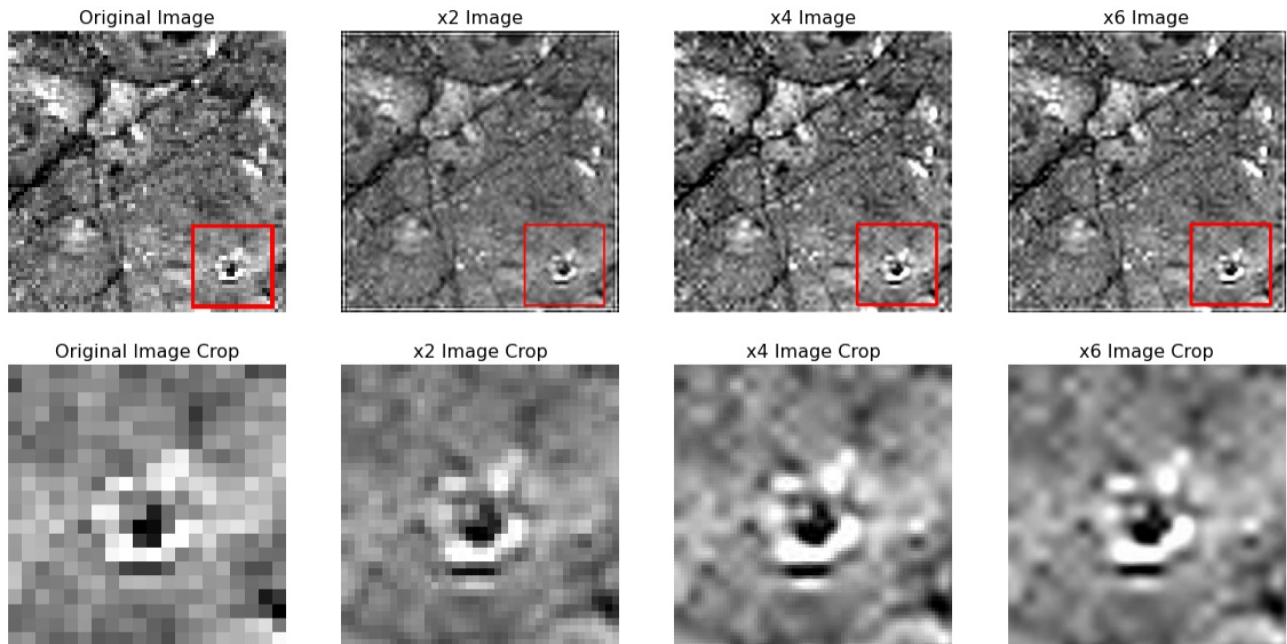
Εικόνα 56: Μεγέθυνση των 60m GSD καναλιών της 3ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.

Εικόνα 4



Εικόνα 57: Μεγέθυνση των 60m GSD καναλιών της 4ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.

Εικόνα 5



Εικόνα 58: Μεγέθυνση των 60m GSD καναλιών της 5ης εικόνας από το μοντέλο ZSSR κατά παράγοντες 2,4,6.

4.3.4 Σχολιασμός Αποτελεσμάτων

Εξετάζοντας τα αποτελέσματα που προκύπτουν από το μοντέλο Zero-Shot Super Resolution μπορεί κανείς να οδηγηθεί σε παρόμοια συμπεράσματα με εκείνα που παρουσιάστηκαν στην ενότητα του Deep Image Prior μοντέλου. Συγκεκριμένα, με μία γρήγορη ανάλυση των εικόνων που συντίθενται από το ZSSR μπορεί κανείς να επιβεβαιώσει την ικανότητα και του συγκεκριμένου μοντέλου να επεξεργάζεται και να βελτιώνει την χωρική ανάλυση δορυφορικών εικόνων. Αναλυτικότερα, οι εικόνες υψηλής ανάλυσης που παράγονται από το μοντέλο περιέχουν πιο ξεκάθαρες ακμές μεταξύ των χωρικών δομών που απεικονίζονται στην αντίστοιχη αρχική εικόνα χαμηλής χωρικής ανάλυσης, με την ποιότητα των χαρακτηριστικών αυτών να βελτιώνεται καθώς αυξάνεται ο συντελεστής μεγέθυνσης. Ωστόσο, το συγκεκριμένο μοντέλο, παρουσιάζει σχετική χαμηλή ικανότητα εξαγωγής χαρακτηριστικών υφής, γεγονός το οποίο επιβεβαιώνεται στις εικόνες που παράγονται για συντελεστές μεγέθυνσης 4 ή και 6 στις οποίες, παρά τις βελτιωμένες ακμές και τα καλύτερα περιγράμματα, το μοντέλο αδυνατεί να αποδώσει με μεγάλη ακρίβεια την υφή των συγκεκριμένων επιφανειών. Τέλος, αξίζει να σημειωθεί πως οι τα περικεκομμένα τμήματα από τα αποτελέσματα του ZSSR φαίνεται να είναι ελαφρώς θολά, γεγονός το οποίο μπορεί να αποδοθεί μερικώς στην αδυναμία του μοντέλου να εξάγει χαρακτηριστικά υφής των χωρικών δομών που υπάρχουν στην εκάστοτε εικόνα. Ένας ακόμη πιθανός λόγος είναι η σχετικά ρηχή αρχιτεκτονική του συγκεκριμένου μοντέλου, η οποία δεν επαρκεί για την μοντελοποίηση όλων των στατιστικών εξαρτήσεων που υπάρχουν σε μία εικόνας δορυφορικής φύσης.

Αναφορικά με το τους χρόνους εκτέλεσης του αλγορίθμου του συγκεκριμένου μοντέλου, παρατίθεται παρακάτω ο αντίστοιχος πίνακας των χρόνων εκτέλεσης του μοντέλου ανά ανάλυση εικόνας και ανά συντελεστή μεγέθυνσης.

Πίνακας 5: Μέσοι χρόνοι εκτέλεσης για το ZSSR μοντέλο.

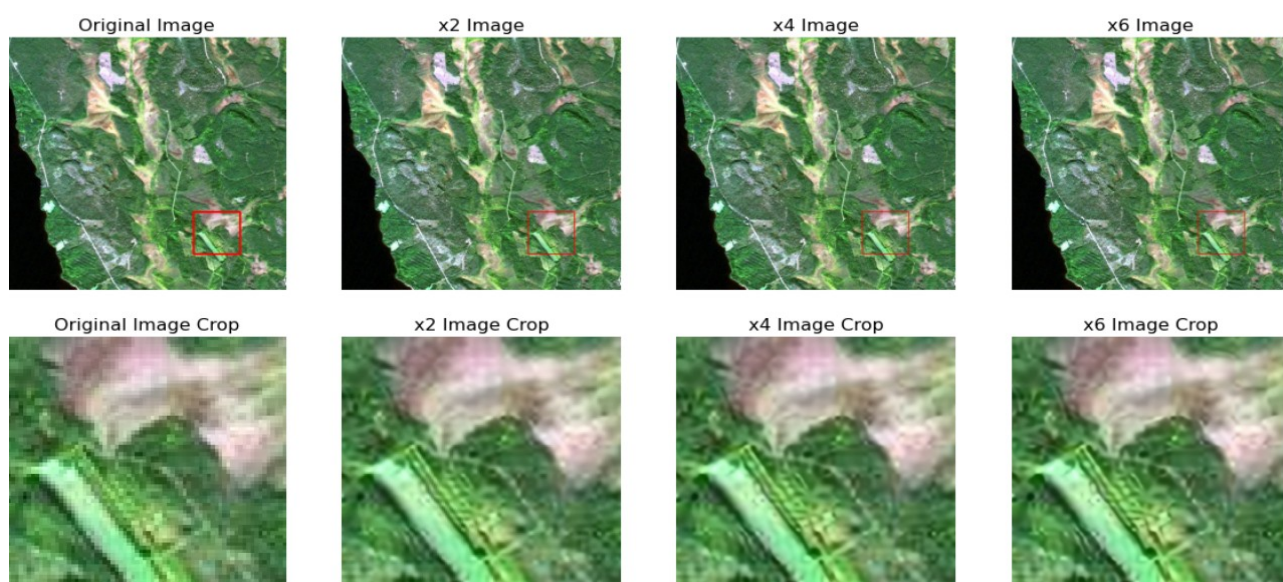
Συντελεστής μεγέθυνσης / Ανάλυση εικόνας	10m	20m	60m
x2	2.46 min	2.46 min	0.62 min
x4	5.1 min	3.18 min	0.56 min
x6	19 min	11 min	1.5 min

4.4. Αποτελέσματα Degradation Aware Super Resolution

Στην ενότητα αυτή παρουσιάζονται τα αποτελέσματα του τρίτου και τελευταίου μοντέλου που εξετάστηκε στην παρούσα εργασία. Σημειώνεται, για ακόμη μία φορά, ότι η παράθεση των αποτελεσμάτων του DASR μοντέλου ακολουθεί ακριβώς την ίδια λογική με προηγούμενως.

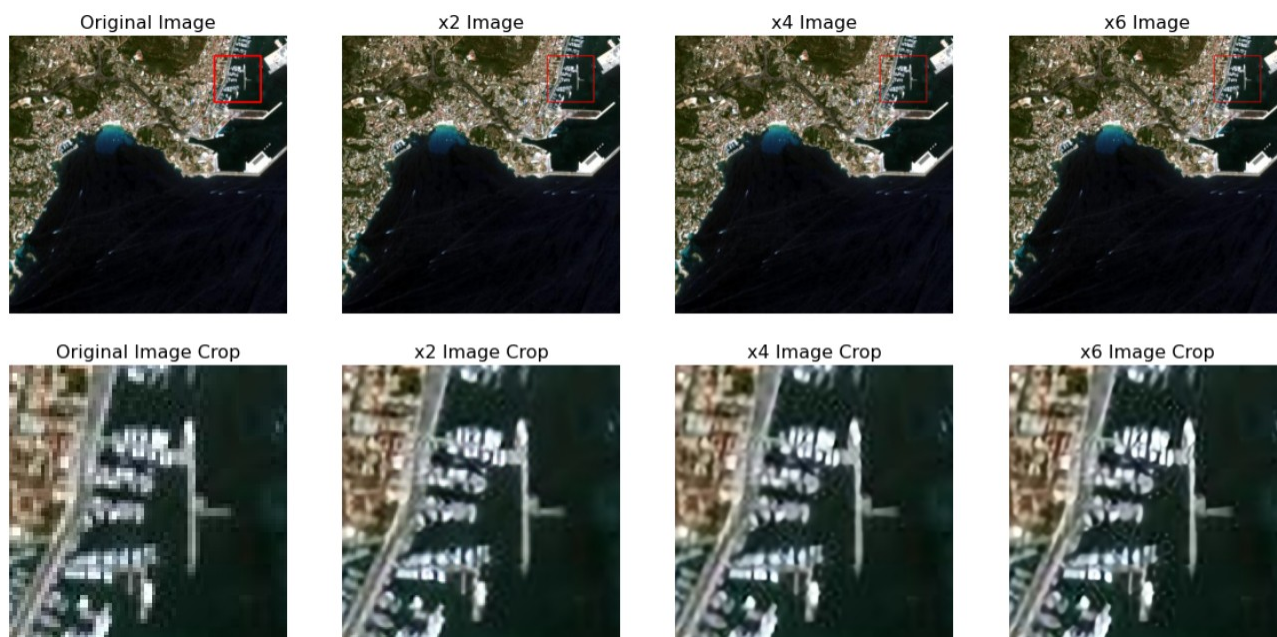
4.4.1 Εικόνες χωρικής ανάλυσης 10m GSD

Εικόνα 1



Εικόνα 59: Μεγέθυνση των 10m GSD καναλιών της 1ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.

Εικόνα 2



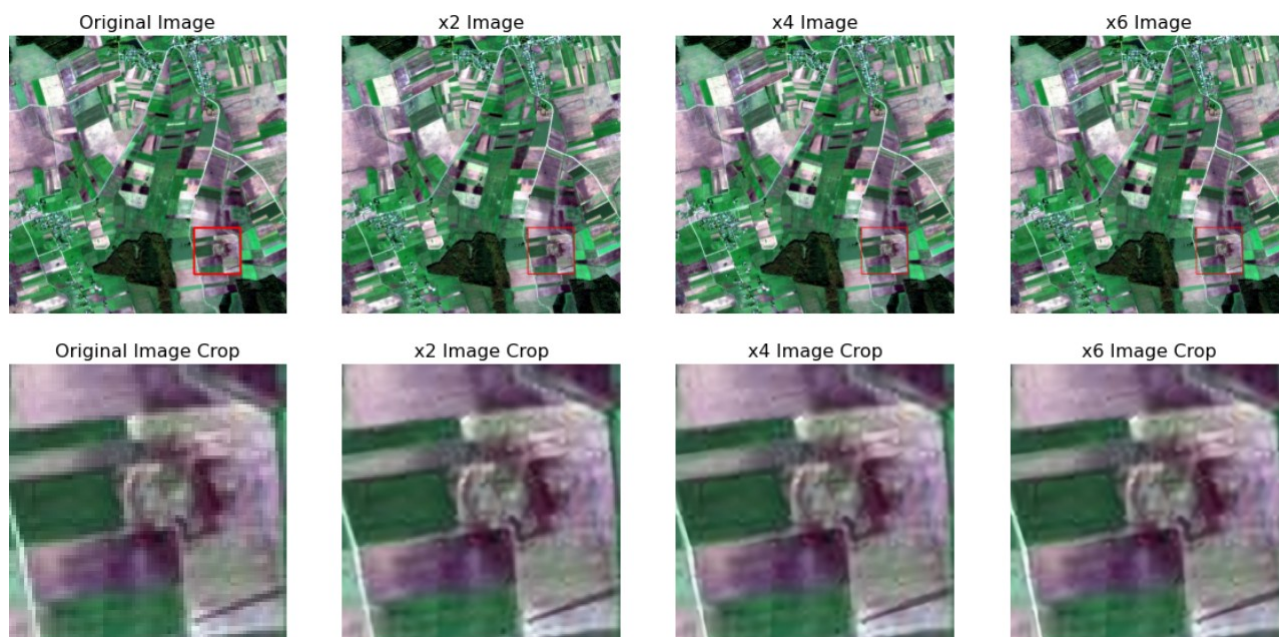
Εικόνα 60: Μεγέθυνση των 10m GSD καναλιών της 2ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.

Εικόνα 3



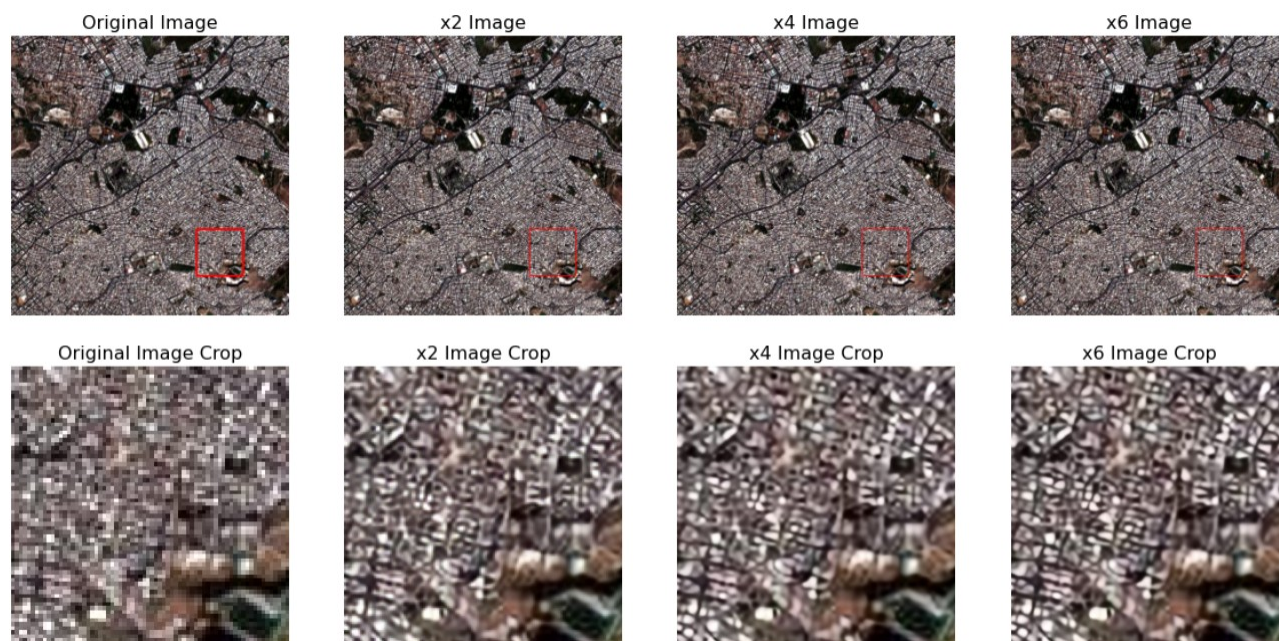
Εικόνα 61: Μεγέθυνση των 10m GSD καναλιών της 3ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.

Εικόνα 4



Εικόνα 62: Μεγέθυνση των 10m GSD καναλιών της 4ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.

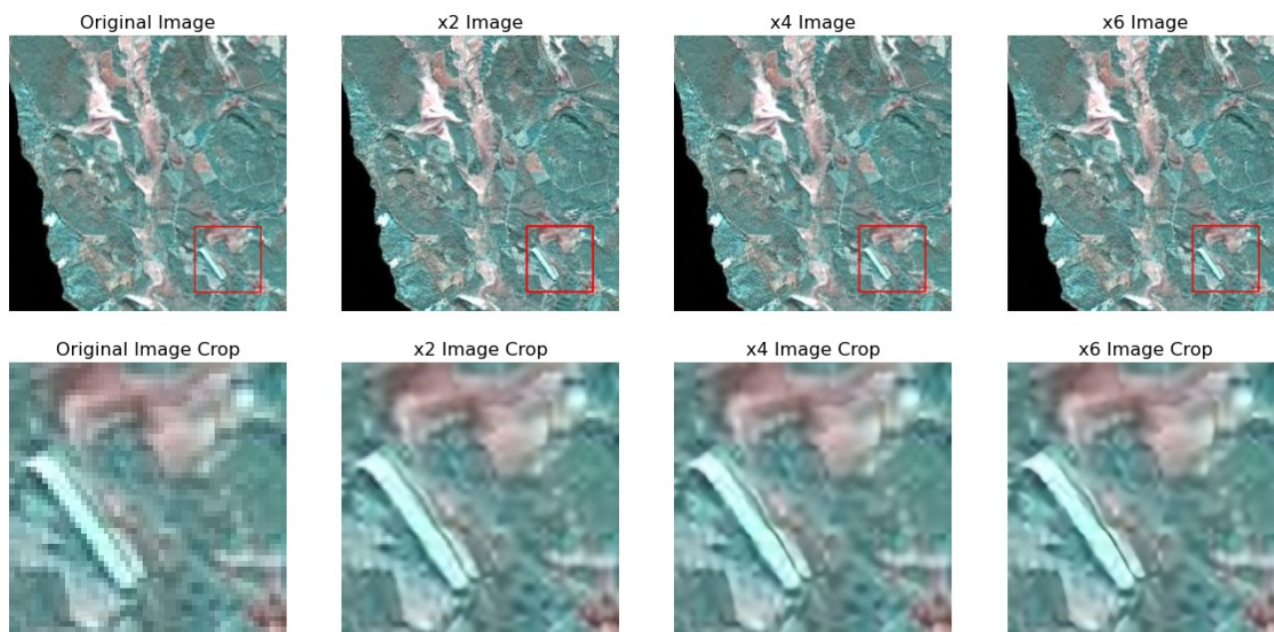
Εικόνα 5



Εικόνα 63: Μεγέθυνση των 10m GSD καναλιών της 5ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.

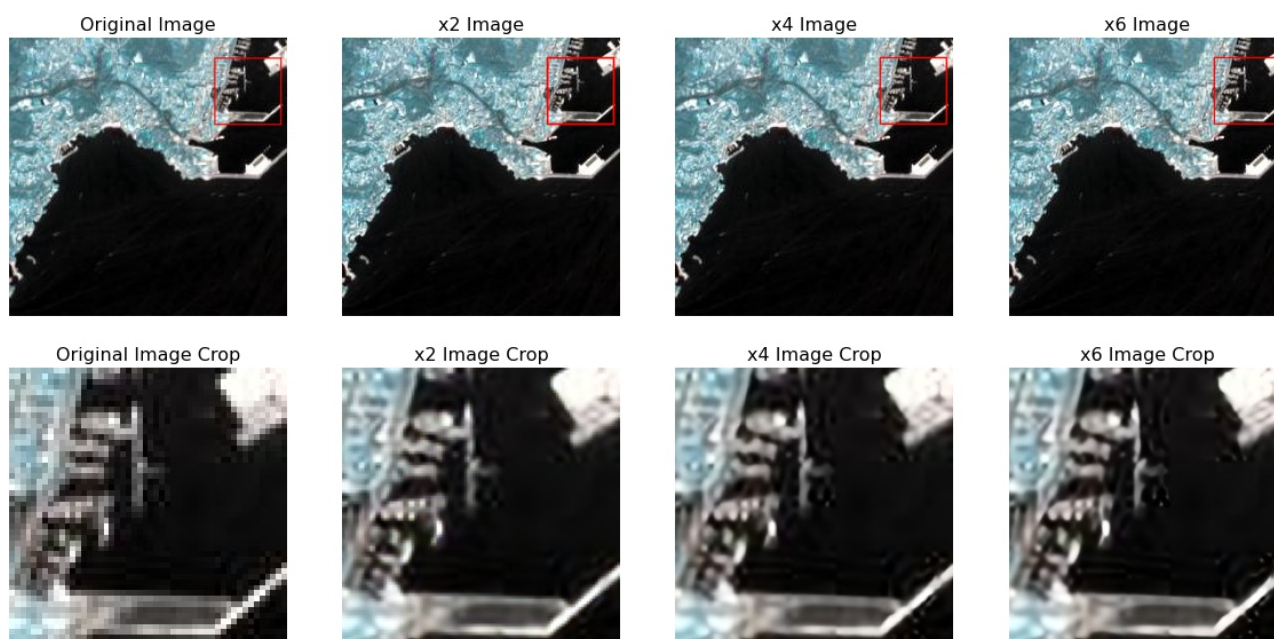
4.4.2 Εικόνες χωρικής ανάλυσης 20m GSD

Εικόνα 1



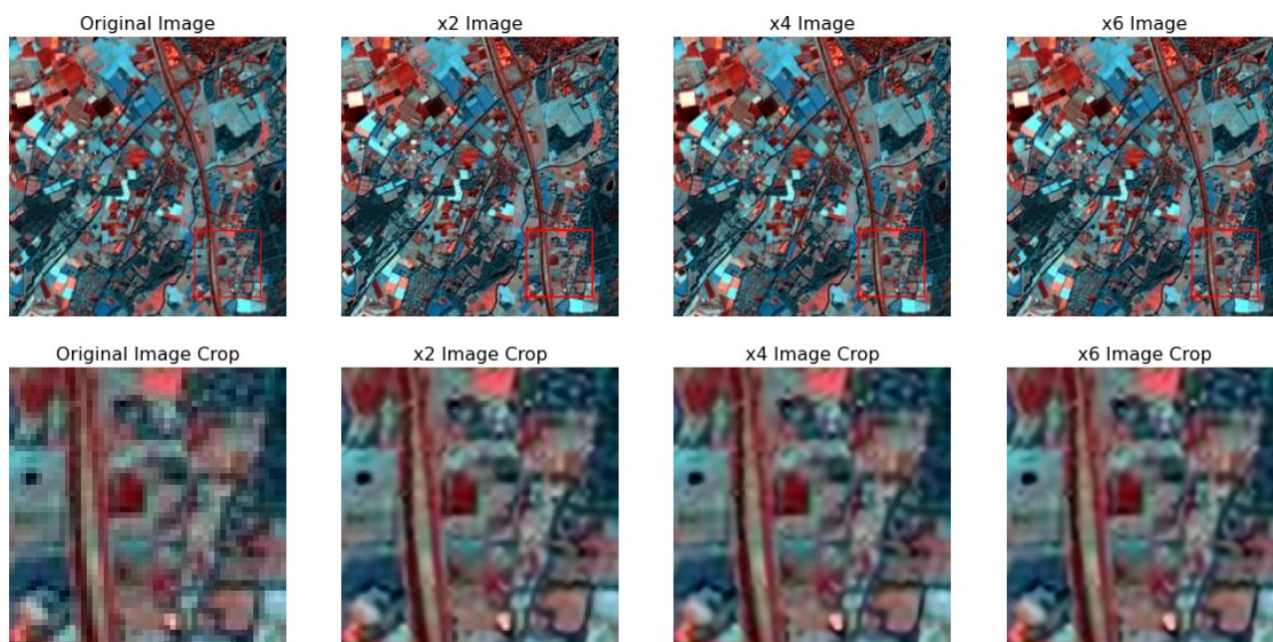
Εικόνα 64: Μεγέθυνση των 20m GSD καναλιών της 1ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.

Εικόνα 2



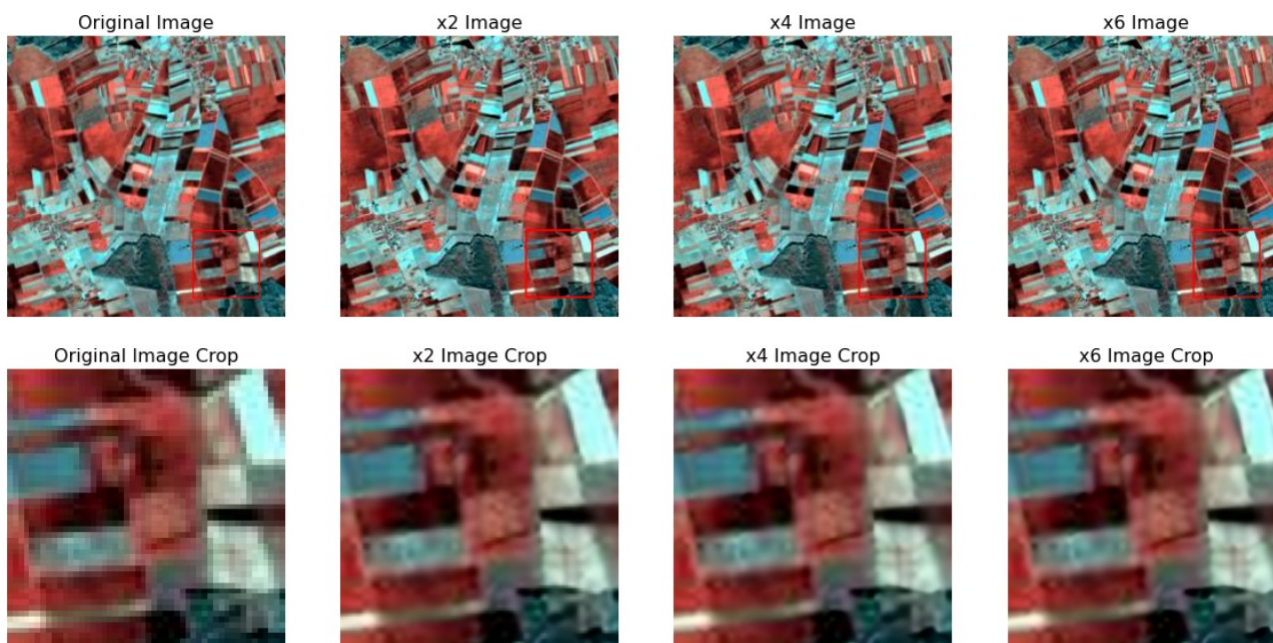
Εικόνα 65: Μεγέθυνση των 20m GSD καναλιών της 2ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.

Εικόνα 3



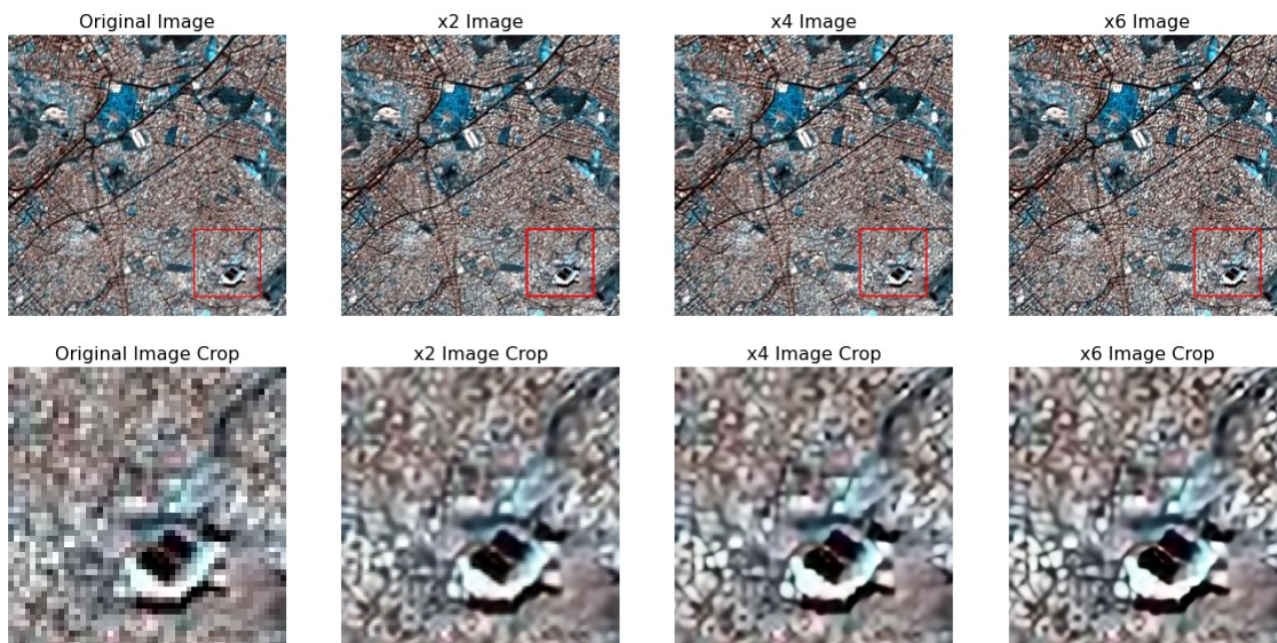
Εικόνα 66: Μεγέθυνση των 20m GSD καναλιών της 3ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.

Εικόνα 4



Εικόνα 67: Μεγέθυνση των 20m GSD καναλιών της 4ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.

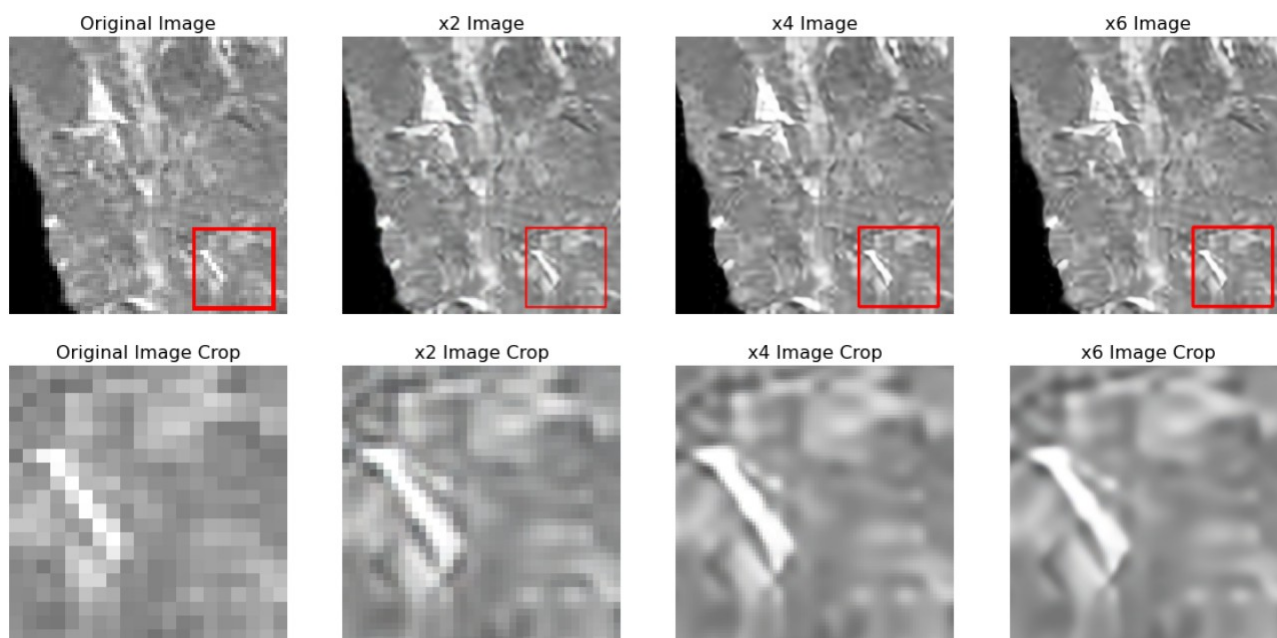
Εικόνα 5



Εικόνα 68: Μεγέθυνση των 20m GSD καναλιών της 5ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.

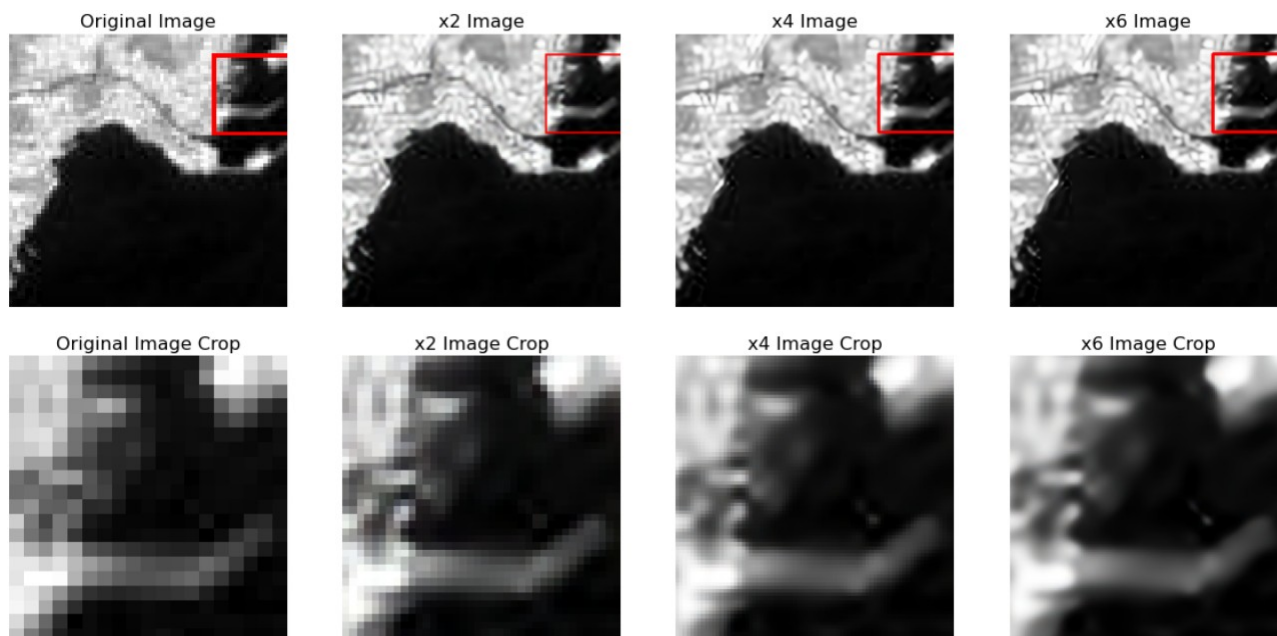
4.4.3 Εικόνες χωρικής ανάλυσης 60m GSD

Εικόνα 1



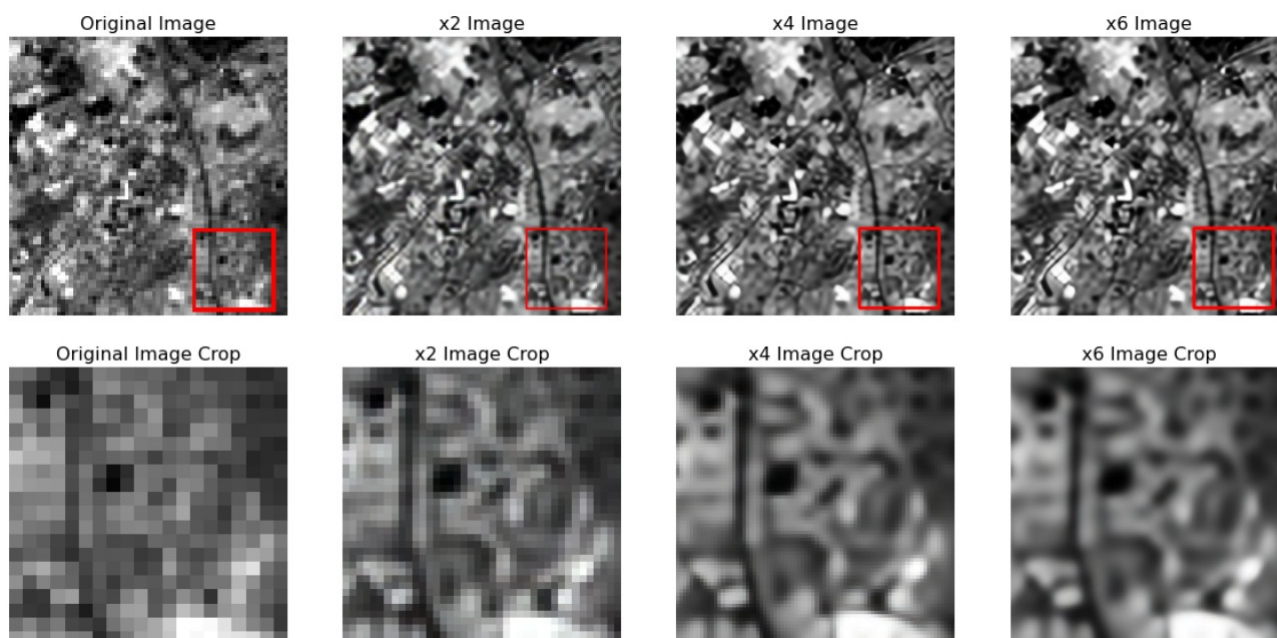
Εικόνα 69: Μεγέθυνση των 60m GSD καναλιών της 1ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.

Εικόνα 2



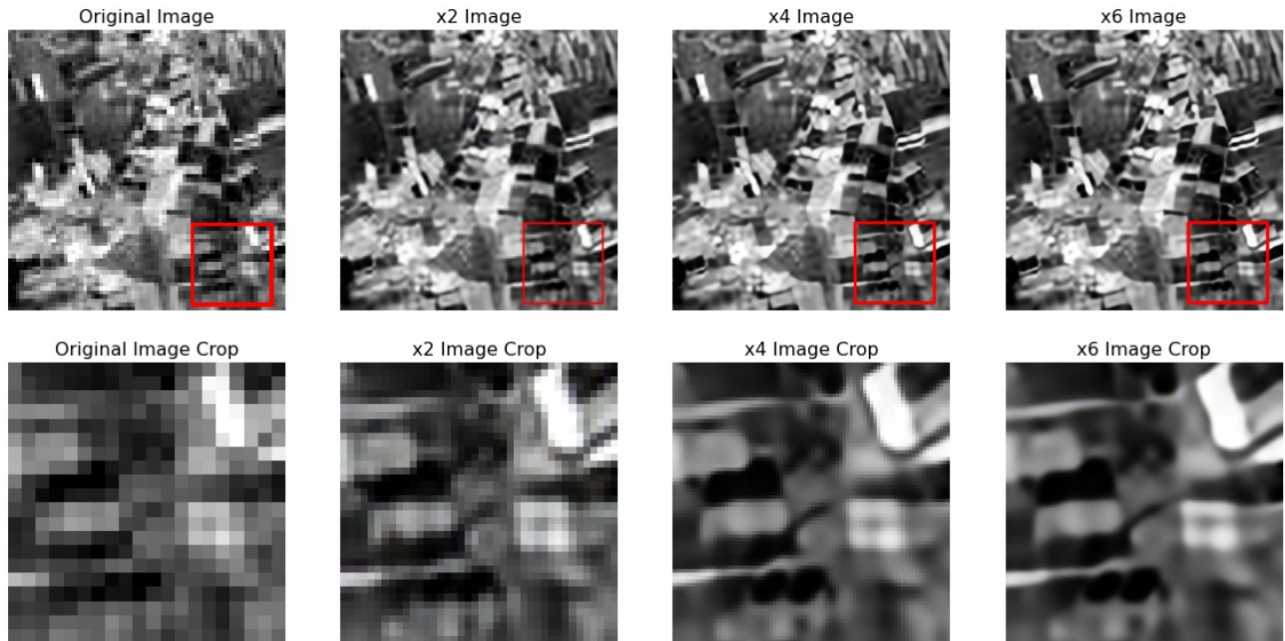
Εικόνα 70: Μεγέθυνση των 60m GSD καναλιών της 2ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.

Εικόνα 3



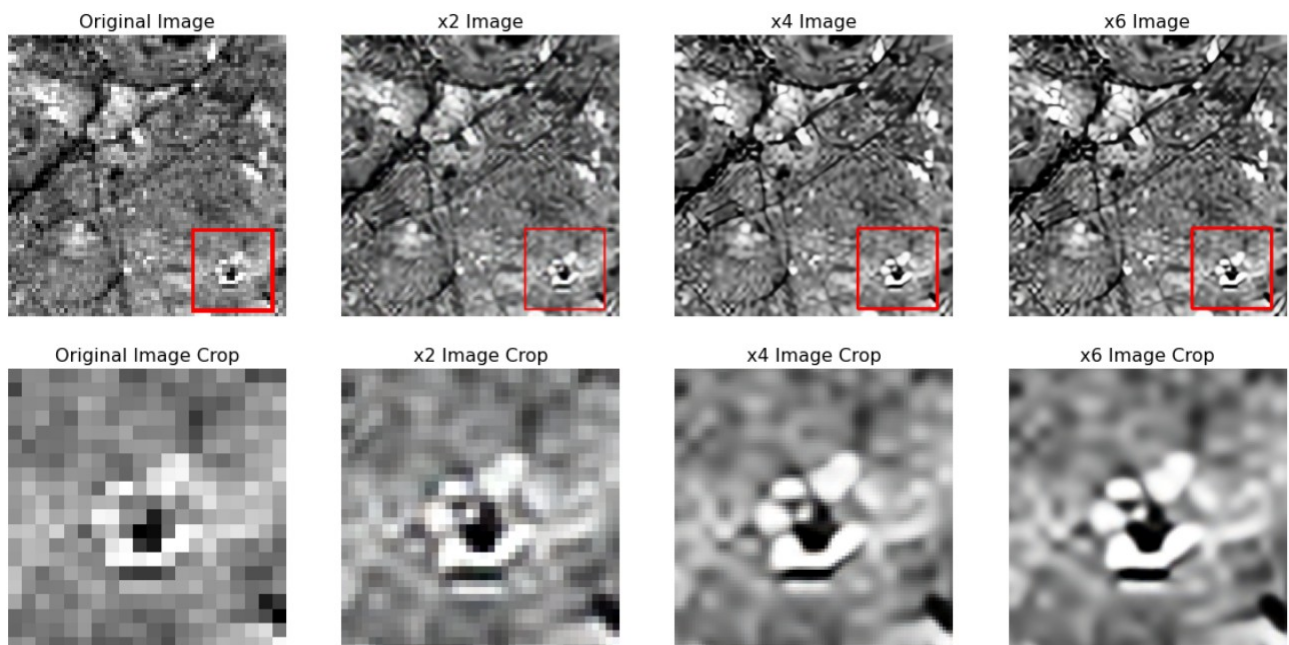
Εικόνα 71: Μεγέθυνση των 60m GSD καναλιών της 3ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.

Εικόνα 4



Εικόνα 72: Μεγέθυνση των 60m GSD καναλιών της 4ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.

Εικόνα 5



Εικόνα 73: Μεγέθυνση των 60m GSD καναλιών της 5ης εικόνας από το μοντέλο DASR κατά παράγοντες 2,4,6.

4.4.4 Σχολιασμός Αποτελεσμάτων

Στην ενότητα αυτή, όπως και στις αντίστοιχες ενότητες που αφορούσαν τα προηγούμενα δύο μοντέλα, παρατίθενται ορισμένα σχόλια και παρατηρήσεις σχετικά με τα αποτελέσματα που λήφθηκαν από το μοντέλο Degradation Aware Super Resolution (DASR). Αρχικά, το

συγκεκριμένο μοντέλο για τις περισσότερες από τις εικόνες του συνόλου ελέγχου με χωρική ανάλυση ίση με 10m GSD, καταφέρνει να δημιουργήσει αποτελέσματα με πιο ευδιάκριτες ακμές και περιγράμματα, καθώς και να περιορίσει σε σημαντικό βαθμό το φαινόμενο του pixelation. Ειδική περίπτωση αποτελεί η εικόνα 3, στην οποία το μοντέλο ανεξάρτητα του συντελεστή μεγέθυνσης αδυνατεί να δημιουργήσει εικόνες αυξημένης χωρικής ανάλυσης στις οποίες τα όρια μεταξύ διαφορετικών περιοχών της εικόνας είναι ορατά και ευδιάκριτα. Επιπρόσθετα, ανεξάρτητα με την αρχική χωρική ανάλυση της υπό εξέταση εικόνας, το μοντέλο επιδεικνύει μία αρκετά καλή συμπεριφορά στην διαχείριση υποπεριοχών των εικόνων, όπου η πυκνότητα των ακμών και κατ' επέκταση των χωρικών δομών είναι αυξημένη, πετυχαίνοντας την δημιουργία νέων εικόνων με αρκετά ξεκάθαρα όρια μεταξύ των χωρικών δομών. Σε αντιδιαστολή με την ικανοποιητική συμπεριφορά του στις περιοχές με πολλαπλές ακμές και σύνορα, στην περίπτωση των πιο ομοιόμορφων περιοχών το μοντέλο έχει την τάση να τις κάνει ακόμα πιο ομοιόμορφες. Η συμπεριφορά αυτή μπορεί να παρομοιαστεί με τοπική εφαρμογή (στις ομοιόμορφες περιοχές της εικόνας χαμηλής ανάλυσης) ενός ισοτροπικού γκαουσσιανού πυρήνα με σχετικά μεγάλη τυπική απόκλιση σ και γίνεται περισσότερο αισθητή καθώς μειώνεται η χωρική ανάλυση της υπό εξέταση εικόνας. Ως χαρακτηριστικό παράδειγμα αναφέρεται η εικόνα 4 με χωρική ανάλυση 60m GSD, στην οποία φαίνεται πως, τόσο για συντελεστή μεγέθυνσης ίσο με τέσσερα όσο για συντελεστή μεγέθυνσης ίσο με έξι, η περιοχές εντός του κόκκινου τετραγώνου έχουν ομαλοποιηθεί με αποτέλεσμα τα όρια των περιοχών αυτών να μην είναι τόσο ξεκάθαρα. Μία ακόμη παρατήρηση σχετικά με τα αποτελέσματα του DASR, είναι ότι σε περιπτώσεις που οι εικόνες χαμηλής ανάλυσης περιέχουν ορισμένες λεπτομέρειες αρκετά μικρής κλίμακας, το μοντέλο έχει την τάση να της αντιμετωπίζει ως προσθετικό θόρυβο και συνεπώς να αφαιρεί τις λεπτομέρειες αυτές από την τελική εικόνα. Η συμπεριφορά αυτή του μοντέλου, δεν αποτελεί υποχρεωτικά κάποιο σφάλμα σχεδίασης, αλλά μπορεί να αποδοθεί στην σύνθετη φύση των δορυφορικών εικόνων. Η σύνθετη αυτή φύση των υπό εξέταση εικόνων, σε συνδυασμό με το γεγονός ότι το σύνολο ελέγχου τροποποιήθηκε έτσι ώστε να περιέχει εικόνες με 8bit ακρίβεια, αντί για 16bit που ήταν αρχικά μπορεί ενδεχομένως να να διαδραμάτισαν καθοριστικό ρόλο στην αδυναμία διατήρησης λεπτομερειών σε μικρές χωρικές κλίμακες.

Στο σημείο αυτό, υπενθυμίζεται πως για την εκτέλεση των πειραμάτων για συντελεστές μεγέθυνσης ίσους με 2 και 4, οι συγγραφείς του [53] παρείχαν τα αντίστοιχα προεκπαιδευμένα μοντέλα στην δημοσίευσή τους. Ωστόσο, στην περίπτωση του συντελεστή μεγέθυνσης με τιμή ίση με 6, όπως προαναφέρθηκε, χρειάστηκαν να γίνουν ορισμένες τροποποιήσεις με αποτέλεσμα την ανάγκη εκ νέου εκπαίδευσης του μοντέλου. Η εκπαίδευση, λοιπόν, του μοντέλου για συντελεστή μεγέθυνσης ίσο με 6 χρειάστηκαν συνολικά περίπου τρεις μέρες. Ωστόσο, η διαδικασία της βελτίωσης της χωρικής ανάλυσης της εκάστοτε εικόνας χρειάστηκε από μερικά δευτερόλεπτα έως και μισό λεπτό για να ολοκληρωθεί. Παρακάτω παρατίθενται οι πίνακες με τους χρόνους εκπαίδευσης των 3 μοντέλων και τους χρόνους σύνθεσης των εικόνων υψηλής χωρικής ανάλυσης

Πίνακας 6: Χρόνοι εκπαίδευσης του μοντέλου DASR, ανά συντελεστή μεγέθυνσης.

Συντελεστής μεγέθυνσης / Ανάλυση εικόνας	10m	20m	60m
x2	Pretrained	Pretrained	Pretrained
x4	Pretrained	Pretrained	Pretrained
x6	3.5 days	3.5 days	3.5 days

Πίνακας 7: Μέσοι χρόνοι σύνθεσης εικόνων από το DASR.

Συντελεστής μεγέθυνσης / Ανάλυση εικόνας	10m	20m	60m
x2	4.98 sec	4.81 sec	4.80 sec
x4	5.23 sec	4.63 sec	4.80 sec
x6	4.87 sec	5.07 sec	4.77 sec

4.5. Ποιοτική αξιολόγηση αποτελεσμάτων μέσω φωτοερμηνείας

Στις προηγούμενες ενότητες παρουσιάστηκαν οι εικόνες βελτιωμένης ανάλυσης που δημιουργήθηκαν από τα τρία μοντέλα τα οποία εξετάζονται στην παρούσα εργασία, καθώς και ορισμένες παρατηρήσεις σχετικές με την ποιότητα των παραγόμενων αποτελεσμάτων και την ικανότητα των μοντέλων να συνθέτουν εικόνες για όλα και μεγαλύτερους συντελεστές μεγέθυνσης. Προκειμένου, λοιπόν, να ολοκληρωθεί η αξιολόγηση των μοντέλων κρίνεται σκόπιμη μία σύγκριση μεταξύ τους, ανά συντελεστή μεγέθυνσης, με σκοπό την ανίχνευση του βέλτιστου από τρία μοντέλα ανά περίπτωση. Στην σύγκριση αυτή θα συμπεριληφθούν παράγοντες όπως η ποιότητα των παραγόμενων εικόνων από το εκάστοτε μοντέλο, αλλά και ο χρόνος που χρειάστηκε για την δημιουργία αυτών.

Αρχικά, εξετάζονται οι επιδόσεις των μοντέλων στην βελτίωση της χωρικής ανάλυσης εικόνων με ανάλυση 10m GSD. Όπως έχει ήδη αναφερθεί στις αντίστοιχες ενότητες, και τα τρία μοντέλα στην περίπτωση των εικόνων με 10m GSD ανάλυση είναι ικανά να δημιουργήσουν ικανοποιητικά αποτελέσματα με ευδιάκριτες ακμές και περιθώρια, συνεπώς για να μπορέσει κανείς να εντοπίσει οποιεσδήποτε υπάρχουσες διαφορές πρέπει να καταφύγει στις αντίστοιχες περικεκομμένες εικόνες. Παρακάτω παρατίθενται ενδεικτικά τα περικεκομμένα τμήματα από την εικόνα 3 για καθένα από τα τρία μοντέλα.



Εικόνα 74: DIP (10m).



Εικόνα 75: ZSSR (10m).

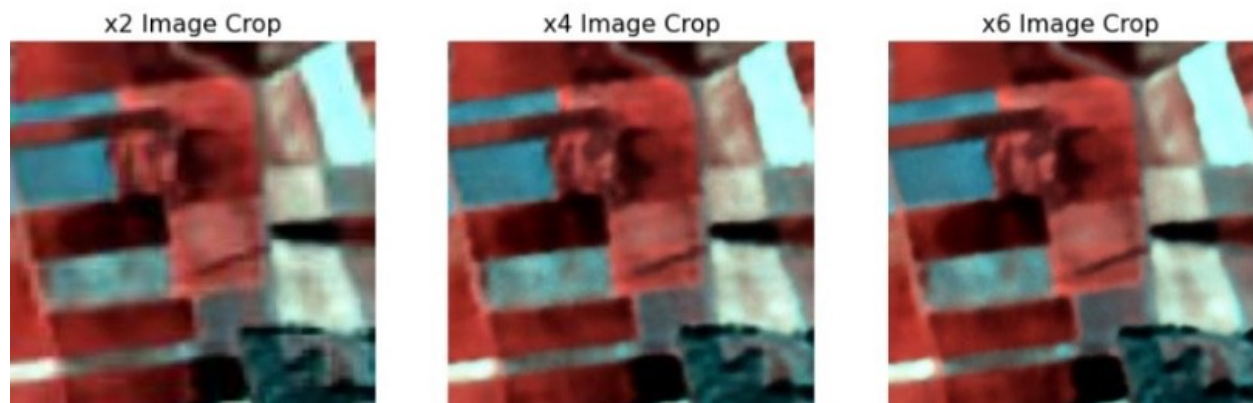


Εικόνα 76: DASR (10m).

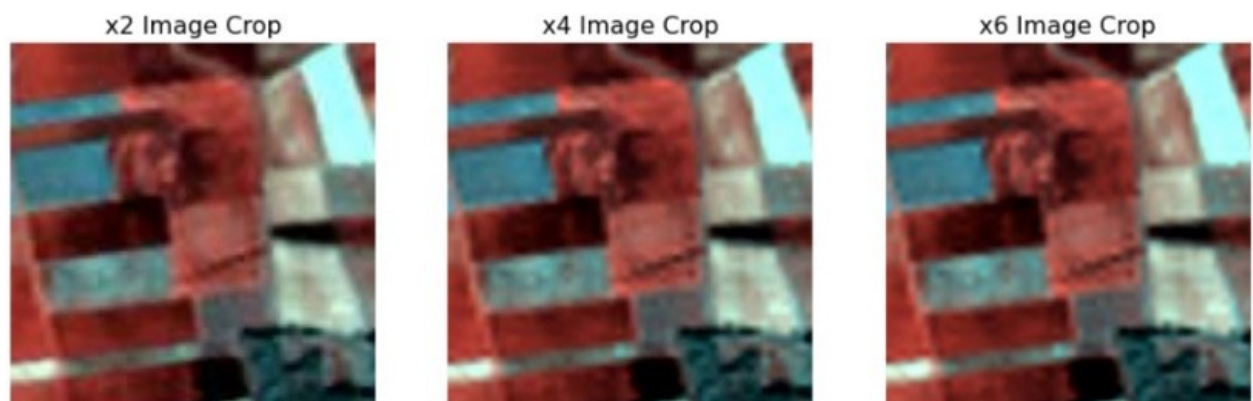
Συγκρίνοντας τις παραπάνω εικόνες, γίνεται άμεσα αντιληπτό πως τα αποτελέσματα του DASR μοντέλου είναι με σχετικά μεγάλη διαφορά τα λιγότερο ικανοποιητικά, καθώς για το συγκεκριμένο μοντέλο οι ακμές των χωρικών δομών είναι αρκετά θολές, ενώ ταυτόχρονα εντοπίζεται και κάποια αδυναμία σύνθεσης των υφών των αντικειμένων και των περιοχών που απεικονίζονται στην εικόνα. Επιπλέον, το DASR αδυνατεί να διατηρήσει τις λεπτομέρειες που υπάρχουν στην εικόνα, όπως αυτές στον κατακόρυφο δρόμο στο αριστερό μέρος αυτής. Συνεχίζοντας την σύγκριση, εξετάζονται αποκλειστικά τα αποτελέσματα των μοντέλων DIP και ZSSR, τα οποία πρέπει να σημειωθεί πως ποιοτικά είναι πολύ κοντά. Συγκεκριμένα, στις περιπτώσεις μεγέθυνσης κατά συντελεστή ίσο με 2, 4 και τα δύο μοντέλα καταφέρνουν να αποδώσουν με μεγάλη ακρίβεια τις χωρικές λεπτομέρειες της εικόνας. Η διαφορά μεταξύ των δύο μοντέλων φαίνεται στην περίπτωση που η μεγέθυνση της εικόνας γίνεται κατά συντελεστή ίσο με 6. Αναλυτικότερα, το DIP μοντέλο φαίνεται να είναι σε θέση να βελτιώσει ακόμη περισσότερο την ποιότητα της εικόνας, πετυχαίνοντας να συνθέσει ακόμη πιο καθαρά και ευδιάκριτα περιγράμματα για τις απεικονιζόμενες χωρικές δομές και να προσθέσει ακόμη περισσότερες πληροφορίες σχετικά με την υφή αυτών. Αντίθετα, τα αποτελέσματα του ZSSR μοντέλου για τον συγκεκριμένο συντελεστή μεγέθυνσης μοιάζουν αρκετά με τα αντίστοιχα αποτελέσματα που προέκυψαν για συντελεστή ίσο με 4, οδηγώντας στο συμπέρασμα πως η ικανότητα γενίκευσής του σε συντελεστές μεγαλύτερους του 4 δεν είναι το ίδιο ικανοποιητική με αυτή του DIP μοντέλου. Επομένως, με βάση τις παραπάνω παρατηρήσεις, μπορεί κανείς να ταξινομήσει τα μοντέλα, ως προς την ποιότητα των αποτελεσμάτων και την ικανότητα γενίκευσης σε μεγαλύτερους συντελεστές μεγέθυνσης, αναγνωρίζοντας το DIP μοντέλο ως το καλύτερο μεταξύ των τριών, ακολουθούμενο από το ZSSR και στο τέλος το DASR. Στο σημείο

αυτό πρέπει να τονιστεί πως η εικόνα 3 επιλέχθηκε, διότι σε αυτή φαίνονται πιο εύκολα η παραπάνω παρατηρήσεις, χωρίς αυτό να σημαίνει πως αυτές δεν ισχύουν και στις υπόλοιπες εικόνες του συνόλου ελέγχου.

Η σύγκριση των μοντέλων συνεχίζεται με την αξιολόγηση των αποτελεσμάτων τους στην περίπτωση εικόνων με χωρική ανάλυση 20m GSD. Για την σύγκριση των μοντέλων θα χρησιμοποιηθούν και πάλι τα περικεκομμένα τμήματα των εικόνων, καθώς ,όπως και στην προηγούμενη περίπτωση, η επισκόπηση ολόκληρων των παραγόμενων εικόνων δεν επαρκεί για την εξαγωγή συμπερασμάτων. Στα πλαίσια της σύγκρισης αυτής θα χρησιμοποιηθεί η εικόνα 4 αντίστοιχης χωρικής ανάλυσης, τα περικεκομμένα τμήματα της οποίας παρατίθενται παρακάτω.



Εικόνα 77: DIP (20m).



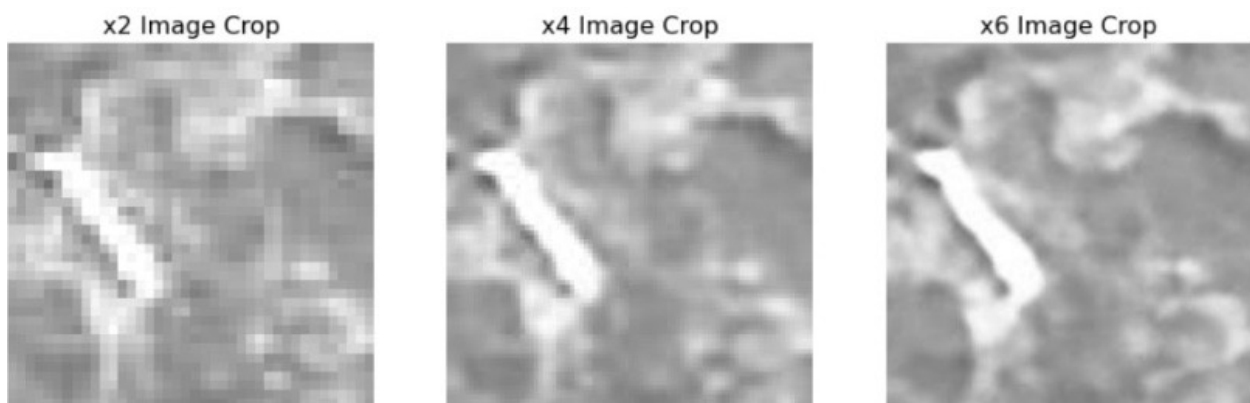
Εικόνα 78: ZSSR (20m).



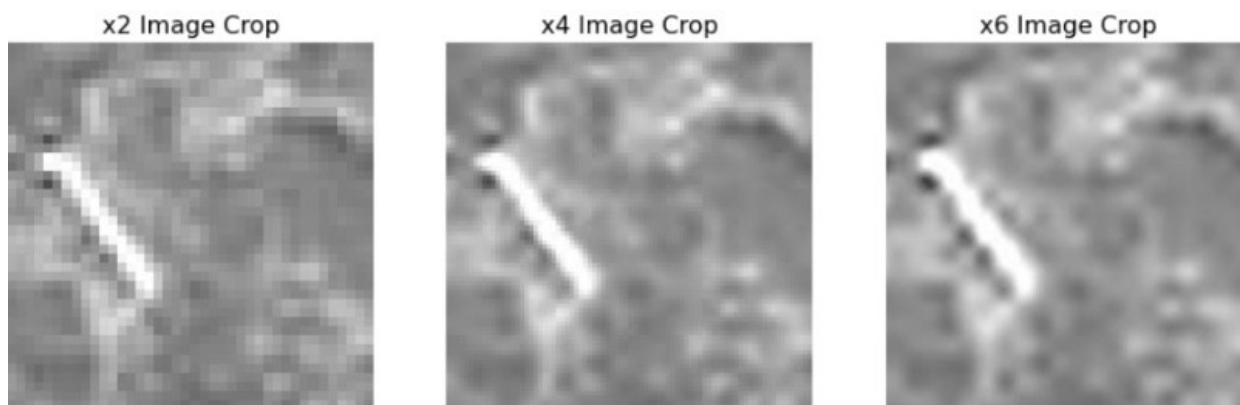
Εικόνα 79: DASR (20m).

Εξετάζοντας τις παραπάνω εικόνες, μπορεί κανείς με ιδιαίτερη ευκολία να καταλήξει στο συμπέρασμα πως το DASR μοντέλο παρουσιάζει, και στην περίπτωση των εικόνων με 20m GSD χωρική ανάλυση, τις χαμηλότερες επιδόσεις. Όπως και στην προηγούμενη περίπτωση, έτσι και τώρα ισχύουν οι παρατηρήσεις σχετικά με τις ακμές, τα περιγράμματα και τις υφές των χωρικών δομών στις εικόνες που συνθέτει το DASR μοντέλο. Επομένως, η σύγκριση θα περιοριστεί ξανά μεταξύ των μοντέλων DIP και ZSSR. Τα αποτελέσματα των δύο μοντέλων για συντελεστή μεγέθυνσης ίσο με 2 μοιάζουν πολύ μεταξύ τους, με το ZSSR να προσφέρει μια ελαφρώς πιο καθαρή συνολική εικόνα. Η διαφορά μεταξύ των παραγόμενων εικόνων των δύο μοντέλων αρχίζει να φαίνεται όταν η μεγέθυνση των εικόνων γίνεται κατά παράγοντα ίσο με 4 και γίνεται αρκετά ξεκάθαρη για συντελεστή μεγέθυνσης ίσο με 6. Αναλυτικότερα, στην πρώτη περίπτωση (συντελεστής μεγέθυνσης ίσος με 4) το DIP φαίνεται να είναι σε θέση να μπορεί να αποδώσει με μεγαλύτερη ακρίβεια και ομοιομορφία την υφή στις εικονιζόμενες χωρικές δομές, ενώ ταυτόχρονα οι ακμές των δομών αυτών αρχίζουν να γίνονται πιο ευδιάκριτες συγκριτικά με το ZSSR. Επίσης, το φαινόμενο του *pixelation* είναι λιγότερο ανιχνεύσιμο στην περίπτωση των εικόνων που δημιουργούνται από το DIP μοντέλο. Οι προαναφερθείσες διαφορές γίνονται ακόμα πιο αισθητές όταν ο συντελεστής μεγέθυνσης γίνεται ίσος με 6, επιβεβαιώνοντας την παρατήρηση ότι το DIP μοντέλο είναι ικανό να συνεχίσει να βελτιώνει την ποιότητα της εικόνας εισόδου του, ενώ το ZSSR για συντελεστές μεγαλύτερους του 4 παράγει περίπου τα ίδια αποτελέσματα. Συνοψίζοντας, λοιπόν, στην περίπτωση που η μεγέθυνση γίνεται κατά παράγοντα ίσο με 2, το ZSSR παράγει ελαφρώς καλύτερα αποτελέσματα από το DIP μοντέλο και, προφανώς, από το DASR μοντέλο. Ωστόσο, για συντελεστές μεγέθυνσης μεγαλύτερους ή ίσους του 4 του DIP μοντέλο επιδεικνύει ξανά την βέλτιστη συμπεριφορά. Συνεπώς, στην περίπτωση των εικόνων με 20m GSD χωρική ανάλυση, δεν καθίσταται εφικτό να οριστεί ένα καθολικά βέλτιστο μοντέλο παρά μόνο κατά περιπτώσεις που σχετίζονται με το συντελεστή μεγέθυνσης.

Στα πλαίσια της σύγκρισης των επιδόσεων των μοντέλων όταν δέχονται ως είσοδο εικόνες χωρικής ανάλυσης 60m GSD θα χρησιμοποιηθεί η εικόνα 1, τα περικεκομμένα τμήματα της οποίας φαίνονται παρακάτω.



Εικόνα 80: DIP (60m).



Εικόνα 81: ZSSR (60m).



Εικόνα 82: DASR (60m).

Προτού γίνει αναφορά στην ποιότητα των παραγόμενων εικόνων, κρίνεται σκόπιμο να σημειωθεί πως η βελτίωση της χωρικής ανάλυσης των συγκεκριμένων αποτελεί ιδιαίτερα δύσκολη διαδικασία και για τα τρία μοντέλα, διότι εξαιτίας του μικρού τους μεγέθους και οι τρεις μέθοδοι έχουν ένα αρκετά περιορισμένο σύνολο πληροφοριών που μπορούν να αξιοποιήσουν για την σύνθεση των ζητούμενων εικόνων. Αναλυτικότερα, στην περίπτωση του DIP μοντέλου, εξαιτίας του μικρού μεγέθους των υπό εξέταση εικόνων, οι εξαρτήσεις σε επίπεδο pixel που προσπαθεί να εξάγει το μοντέλο είναι σημαντικά λιγότερες σε σχέση με τις προηγούμενες δύο αναλύσεις, με αποτέλεσμα η πρότερη γνώση που αποθηκεύεται στα βάρη του δικτύου να μην μοντελοποιεί με ακρίβεια τα στατιστικά χαρακτηριστικά των εικόνων. Στην περίπτωση των DASR και ZSSR το μικρό μέγεθος των εικόνων συνεπάγεται άμεσα μικρό πλήθος τμημάτων από την εκάστοτε εικόνα και κατ' επέκταση μικρό σύνολο εκπαίδευσης των μοντέλων. Με βάση, λοιπόν, την παραπάνω παρατήρηση μπορεί κανείς να προχωρήσει στην αξιολόγηση των παραγόμενων εικόνων από το εκάστοτε μοντέλο, διατηρώντας χαμηλότερες προσδοκίες ως προς την ποιότητα των αποτελεσμάτων.

Αρχικά, αναφέρεται πως το ZSSR στην περίπτωση των εικόνων ανάλυσης 60m GSD φαίνεται να δυσκολεύεται περισσότερο από όλα τα υπόλοιπα μοντέλα να διατηρήσει το ακριβές σχήμα, ή ορθότερα τα ακριβή όρια, των απεικονιζόμενων χωρικών δομών. Ωστόσο, αν και ελαφρώς παραμορφωμένες, οι ακμές των χωρικών δομών των εικόνων παρουσιάζονται με λίγο μεγαλύτερη ακρίβεια από στην περίπτωση των παραγόμενων, από το DASR, εικόνων αντίστοιχης μεγέθυνσης. Επιπρόσθετα, στην περίπτωση του DASR, γίνεται πλέον πολύ εμφανές το φαινόμενο που περιγράφηκε στην ενότητα 4.4.4, πως δηλαδή οι παραγόμενες εικόνες είναι

θολές δίνοντας την εντύπωση ότι έχει εφαρμοστεί σε αυτές γκαουσιανός πυρήνας μεγάλης τυπικής απόκλισης. Συνεπώς, για την επιλογή μεταξύ των δύο αυτών μοντέλων πρέπει να αναλογιστεί κανείς τις προτεραιότητες της εκάστοτε εφαρμογής και να αποφασίσει αν είναι σημαντικότερο μία καθαρότερη εικόνα με ελαφρώς μετατοπισμένα όρια των χωρικών δομών, ή μία πιο θολή εικόνα στην οποία όμως στα σχήματα έχουν διατηρηθεί. Στην συνέχεια, αναφορικά με το μοντέλο DIP σημειώνεται πως μπορεί επιδεικνύει αρκετά ικανοποιητικά αποτελέσματα και για τους τρεις συντελεστές μεγέθυνσης που εξετάστηκαν στην παρούσα εργασία. Αναλυτικότερα, αναφέρεται πως το συγκεκριμένο μοντέλο παρουσιάζει την ικανότητα να διατηρεί τα σχήματα των χωρικών δομών αμετάβλητα σε πολύ μεγάλο βαθμό, ενώ ταυτόχρονα παρέχει εικόνες βελτιωμένης ανάλυσης με αρκετά ξεκάθαρες ακμές και σημαντικά περισσότερη πληροφορία σχετικά με την υφή των εικονιζόμενων περιοχών, συγκριτικά πάντα με τα άλλα δύο μοντέλα. Επιπλέον, αξίζει να σημειωθεί πως το DIP μοντέλο παράγει αρκετά πιο ευκρινή αποτελέσματα για συντελεστή μεγέθυνσης ίσο με 6, με το λιγότερο δυνατό θόλωμα, σε σχέση με τα άλλα δύο μοντέλα. Επομένως, στην περίπτωση των εικόνων με χωρική ανάλυση 60m GSD το DIP μοντέλο μπορεί να θεωρηθεί ως το καλύτερο, μεταξύ των τριών, ενώ μεταξύ των άλλων δύο μοντέλων δεν μπορεί να διακριθεί ένα ολικά καλύτερο μοντέλο για τον λόγο που παρουσιάστηκε παραπάνω.

Ολοκληρώνοντας την ανάλυση των αποτελεσμάτων των τριών μοντέλων, κρίνεται σκόπιμο να ληφθεί υπόψη και ο παράγοντας του απαιτούμενου χρόνου για τον σχηματισμό των εικόνων βελτιωμένης χωρικής ανάλυσης, προκειμένου να μπορέσει κανείς να καταλήξει στο βέλτιστο από τα μοντέλα που μελετήθηκαν στην παρούσα εργασία. Αναλυτικότερα, όπως παρουσιάστηκε προηγουμένως, το μοντέλο DIP χρειάζεται τον περισσότερο χρόνο για την σύνθεση των τελικών αποτελεσμάτων, ακολουθούμενο από το ZSSR και μετά το DASR, δεδομένου ότι υπάρχουν διαθέσιμα προεκπαιδευμένα δίκτυα για τον εκάστοτε συντελεστή μεγέθυνσης. Επομένως, στην περίπτωση εφαρμογών που απαιτούνται τα τελικά αποτελέσματα να είναι διαθέσιμο το γρηγορότερο δυνατόν και ο συντελεστής μεγέθυνσης δεν ξεπερνάει το 4, θα μπορούσα κανείς να θεωρήσει πως η βέλτιστη επιλογή είναι το ZSSR μοντέλο, καθώς για μικρούς συντελεστές μεγέθυνσης παρουσιάζει σχεδόν ίδια συμπεριφορά με το DIP με αποτελέσματα ελαφρώς χειρότερης ποιότητας. Ωστόσο, στην περίπτωση που απαιτούνται συντελεστές μεγέθυνσης μεγαλύτεροι του 4 ή η ποιότητα των αποτελεσμάτων είναι το μοναδικό κριτήριο, τότε η επιλογή του DIP μοντέλου αποτελεί μονόδρομο. Τέλος, σε περίπτωση που υπάρχουν διαθέσιμα προεκπαιδευμένα δίκτυα για το DASR μοντέλο και ζητείται η επεξεργασία εικόνων μικρής χωρικής ανάλυσης (όπως αυτές των 60m), τότε το συγκεκριμένο μοντέλο θα μπορούσε να χρησιμοποιηθεί για την γρήγορη, σχεδόν άμεση, λήψη των εικόνων βελτιωμένης ανάλυσης.

Στο σημείο αυτό κρίνεται σκόπιμο, να συνοψιστούν οι παραπάνω παρατηρήσεις σχετικά με τις επιδόσεις των τριών μοντέλων. Προς διευκόλυνση του αναγνώστη, παρακάτω παρατίθεται ένας πίνακας στον οποίο βαθμολογούνται οι επιδόσεις των τριών μοντέλων, συνολικά και για τους τρεις συντελεστές μεγέθυνσης, ως προς τις κύριες παραμέτρους βάσει των οποίων αξιολογήθηκαν αυτά προηγουμένως. Σημειώνεται πως για την βαθμολογία των επιδόσεων των μοντέλων ακολουθείται μία κλίμακα μεταξύ του 1 και του 5, όπου με 5 βαθμολογείται πάντοτε το μοντέλο με την βέλτιστη συμπεριφορά ως προς την εκάστοτε παράμετρο αξιολόγησης.

Πίνακας 8: Βαθμολόγηση των μοντέλων, ως προς τις κύριες παραμέτρους αξιολόγησης.

Μοντέλο/παραμέτροι αξιολόγησης	Ακμές-Περιγράμματα	Υφή	Λεπτομέρειες μικρότερων κλιμάκων	Διατήρηση σχήματος δομών	Συνολική ευκρίνεια εικόνας	Ικανότητα γενίκευσης συντελεστές μεγέθυνσης ≥ 6
DIP	5	5	5	5	5	5
ZSSR	4	4	5	3	4	2
DASR	3	2	1	4	2	3

4.6. Ποσοτική αξιολόγηση αποτελεσμάτων

Στην ενότητα αυτή πραγματοποιείται η ποσοτική αξιολόγηση των αποτελεσμάτων των μοντέλων για κάθε μία από τις αρχικές αναλύσεις των εικόνων του συνόλου ελέγχου. Στην συγκεκριμένη αξιολόγηση, λαμβάνονται υπόψη οι συντελεστές μεγέθυνσης για τους οποίους τα μοντέλα παρήγαγαν τα βέλτιστα οπτικά αποτελέσματα, όπως αυτά ανιχνεύθηκαν στην προηγούμενη ενότητα. Αναλυτικότερα, για το μοντέλο Deep Image Prior χρησιμοποιούνται οι εικόνες που προέκυψαν μετά από μεγέθυνση κατά παράγοντα ίσο με 6, ενώ για τα υπόλοιπα δύο μοντέλα οι αντίστοιχες μεγεθυμένες εικόνες κατά συντελεστή ίσο με 4.

4.6.1 Αξιολόγηση βάσει μετρικών

Για την ποσοτική αξιολόγηση των αποτελεσμάτων των μοντέλων, χρησιμοποιήθηκαν, σε πρώτη φάση, οι μετρικές *Root Mean Squared Error (RMSE)*, *Peak Signal to Noise Ratio (PSNR)*, *Structural Similar Index Measure (SSIM)*, *Feature-based Similarity Index Measure (FSIM)*, *Signal to Reconstruction Error Ratio (SRE)*, *Spectral Angle Mapper (SAM)* και *Universal Image Quality Index (UIQ)*. Στην συνέχεια, παρατίθεται μία σύντομη παρουσίαση των εννοιών που ποσοτικοποιούν οι παραπάνω μετρικές.

- Το **RMSE** μετρά το μέγεθος της αλλαγής ανά pixel εξαιτίας της επεξεργασίας μίας εικόνας. Οι τιμές του RMSE είναι μη αρνητικές και η τιμή 0 σημαίνει ότι η παραγόμενη εικόνα είναι πανομοιότυπη με την αρχική. Το RMSE μεταξύ της αρχικής εικόνας K και της επεξεργασμένης εικόνας $I(i, j)$ δίνεται από την σχέση :

$$RMSE = \sqrt{\left(\frac{1}{H * W} \sum_{i=0, j=0}^{H-1, W-1} |I(i, j) - K(i, j)|^2 \right)}, \text{ όπου } H, W \text{ το ύψος και το πλάτος της εικόνας}$$

- Το **PSNR** μετρά την αναλογία μεταξύ της μέγιστης δυνατής τιμής ενός σήματος και της τιμής του παραμορφωτικού θορύβου που επηρεάζει την πιστότητα της αναπαράστασης του σήματος αυτού. Το PSNR εκφράζεται συνήθως με όρους λογαριθμικής κλίμακας και μετράται σε ντεσιμπέλ, ενώ υπολογίζεται από την σχέση :

$$PSNR = 10 \log_{10} \left(\frac{R^2}{MSE} \right), \text{ όπου } R \text{ η μέγιστη τιμή του εκάστοτε τύπου δεδομένων και } MSE \text{ η υπόριζη ποσότητα της προηγούμενης μετρικής}$$

- Το **SSIM** ποσοτικοποιεί την υποβάθμιση της ποιότητας της εικόνας που προκαλείται από κάποια επεξεργασία, όπως π.χ. συμπίεση δεδομένων, ή από απώλειες στη μετάδοση δεδομένων. Το SSIM βασίζεται σε ορατές δομές στην εικόνα. Με άλλα λόγια το SSIM μετρά οπτική διαφορά μεταξύ δύο παρόμοιων εικόνων. Οι τιμές του SSIM κυμαίνονται μεταξύ στο διάστημα $[-1,1]$, με την τιμή 1 να υποδηλώνει τέλεια δομική ομοιότητα μεταξύ των δύο εικόνων. Ο υπολογισμός της συγκεκριμένης μετρικής μεταξύ δύο εικόνων/παραθύρων x, y γίνεται μέσω της σχέσης :

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

Στην παραπάνω εξίσωση $\mu_x, \mu_y, \sigma_x, \sigma_y$ είναι η μέση τιμή και η τυπική απόκλιση των παραθύρων x, y αντίστοιχα, ενώ σ_{xy} είναι η συνδιακύμανση των x, y .

Οι σταθερές $c_1 = (k_1 L)^2$ και $c_2 = (k_2 L)^2$ αποτελούν σταθερές σταθεροποίησης της διαίρεσης σε περιπτώσεις πολύ μικρών παρονομαστών, όπου L το εύρος τιμών του εκάστοτε τύπου δεδομένων και $k_1 = 0.01, k_2 = 0.03$.

- Η μετρική **FSIM** έχει αναπτυχθεί με σκοπό τη σύγκριση του δομικού περιεχομένου και των χαρακτηριστικών μεταξύ της αρχικής και της επεξεργασμένης εικόνας. Βασίζεται στην ομοιότητα φάσης και στο μέτρο των παραγώγων αυτών. Η FSIM κυμαίνεται μεταξύ των τιμών 0 και 1, όπου το 1 ισοδυναμεί με τέλεια ομοιότητα χαρακτηριστικών.
- Το **SRE** μετρά το σφάλμα ανακατασκευής συναρτήσει της ισχύος του σήματος. Το SRE αποδεικνύεται ιδιαίτερα χρήσιμη μετρική για την σύγκριση μεταξύ εικόνων στις οποίες δεν επικρατούν ίδιες συνθήκες φωτεινότητας. Η συγκεκριμένη μετρική υπολογίζεται από την σχέση

$$SRE = 10 \log_{10} \frac{\mu_x^2}{|\hat{x} - x|^2/n}, \text{ όπου } \mu_x \text{ η μέση τιμή της αρχικής εικόνας } x \text{ και } \hat{x} \text{ η ανασκευασμένη εικόνα.}$$

- Το **SAM** καθορίζει τη φασματική ομοιότητα μεταξύ δύο εικόνων υπολογίζοντας τη γωνία μεταξύ των φασμάτων. Για τον υπολογισμό της γωνίας αυτής τα φάσματα των εικόνων αντιμετωπίζονται ως διανύσματα σε ένα χώρο με διαστάσεις ίση με τον αριθμό των μπάντων που απεικονίζονται στις εικόνες. Το SAM μετριέται σε μοίρες (ή rad) και μικρότερες γωνίες αντιπροσωπεύουν μεγαλύτερη ομοιότητα μεταξύ των φασμάτων των συγκρινόμενων εικόνων.
- Το **UIQ** μοντελοποιεί την συνολική παραμόρφωση σε μία εικόνα λαμβάνοντας υπόψη την απώλεια συσχέτισης (*Loss of correlation*), την παραμόρφωση φωτεινότητας (*Luminance distortion*) και την παραμόρφωση λόγω αντίθεσης (*Contrast distortion*). Οι τιμές της συγκεκριμένης μετρικής κυμαίνονται στο διάστημα $[-1,1]$, με την τιμή 1 να υποδηλώνει απόλυτη ομοιότητα μεταξύ των συγκρινόμενων εικόνων.

Παρακάτω παρουσιάζονται, υπό την μορφή πίνακα, οι μέσοι όροι των προαναφερθέντων μετρικών για κάθε ένα από τα μοντέλα και για κάθε μία από τις περιπτώσεις χωρικής ανάλυσης των αρχικών εικόνων, όπου τα καλύτερα αποτελέσματα ανά αρχική ανάλυση εικόνας

σημειώνονται με **bold** χαρακτήρες. Στο σημείο αυτό κρίνεται σκόπιμο να σημειωθεί πως για την λήψη των μετρικών, χρησιμοποιήθηκαν οι υποδειγματοληπτημένες εκδοχές των αποτελεσμάτων κάθε μοντέλου προκειμένου να καταστεί δυνατή η σύγκριση αυτών με τις αντίστοιχες αρχικές εικόνες του συνόλου ελέγχου.

Πίνακας 9: Μέσοι όροι μετρικών ανά μοντέλο και ανά χωρική ανάλυση εικόνων.

Μοντέλο/ Μετρική	RMSE	PSNR	SSIM	FSIM	SRE	SAM	UIQ
DIP (10m)	0.0228	33.1094	0.9837	0.9955	27.5207	0.0478	0.9971
ZSSR (10m)	0.0012	60.1528	0.9986	0.9979	19.9796	0.0268	0.9996
DASR (10m)	0.2526	12.2083	0.6908	0.9379	21.7532	0.0887	0.7336
DIP(20m)	0.0201	34.3633	0.9547	0.9431	21.1487	0.1451	0.9400
ZSSR (20m)	0.0014	57.8714	0.9985	0.9970	19.3841	0.0219	0.9997
DASR (20m)	0.0719	22.9488	0.9056	0.9732	20.5006	0.1091	0.9826
DIP(60m)	0.0255	31.9716	0.9889	0.9989	20.3775	0.0511	0.9983
ZSSR (60m)	0.0014	57.6651	0.9985	0.9998	13.1825	0.0243	0.9996
DASR (60m)	0.5202	5.9000	0.0003	0.9686	6.9523	0.0740	5.16×10^{-5}

Συγκρίνοντας τις τιμές των μετρικών στον παραπάνω πίνακα γίνεται άμεσα αντιληπτό πως το μοντέλο DASR παρουσιάζει τις χαμηλότερες επιδόσεις, με μεγάλη διαφορά, μεταξύ των τριών μοντέλων. Η παρατήρηση αυτή βρίσκεται σε απόλυτη συμφωνία με τα συμπεράσματα που προέκυψαν στην ενότητα 4.5, επιβεβαιώνοντας την σχετική, ως προς τα υπόλοιπα μοντέλα, αδυναμία του DASR να επεξεργαστεί δορυφορικές εικόνες και τελικά να συνθέσει νέες εικόνες βελτιωμένης χωρικής ανάλυσης. Ιδιαίτερο ενδιαφέρον παρουσιάζει η περίπτωση επεξεργασίας εικόνων με αρχική χωρική ανάλυση 60m GSD, χρησιμοποιώντας το συγκεκριμένο μοντέλο, όπου παρατηρείται ραγδαία επιδείνωση των τιμών όλων των μετρικών, γεγονός το οποίο επιβεβαιώνει τις παρατηρήσεις που έγιναν κατά την ποιοτική ανάλυση των αποτελεσμάτων, σχετικά με την αδυναμία δημιουργίας ξεκάθαρων ακμών και περιγραμμάτων, την διατήρηση χωρικών λεπτομερειών και γενικά την συνολική ευκρίνεια των παραγόμενων εικόνων.

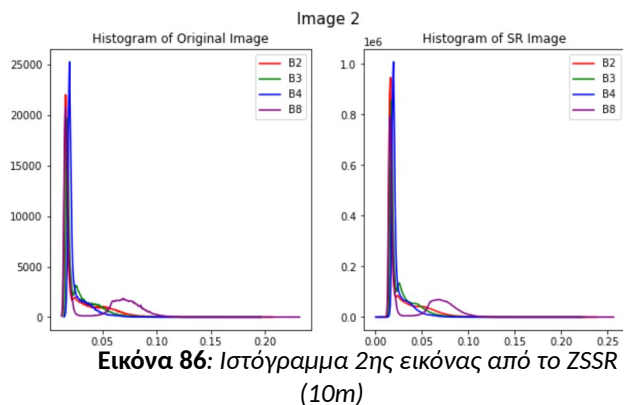
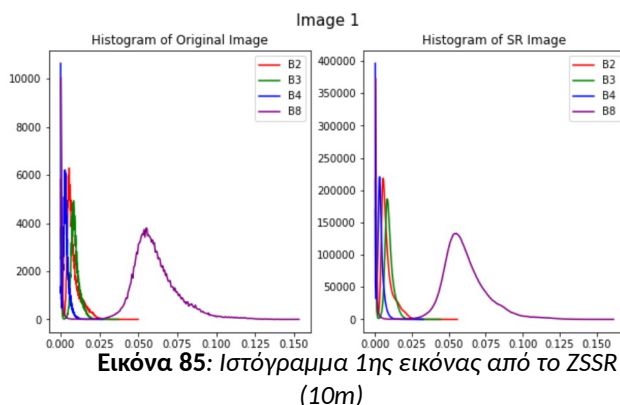
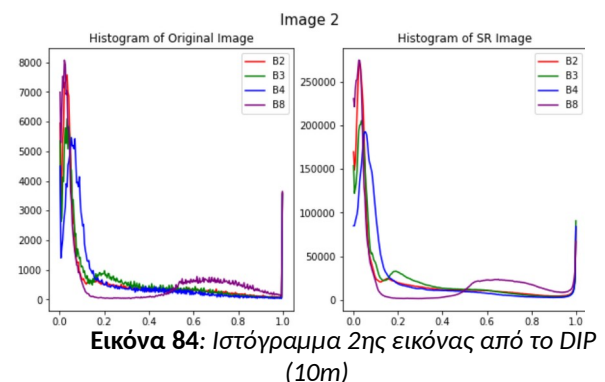
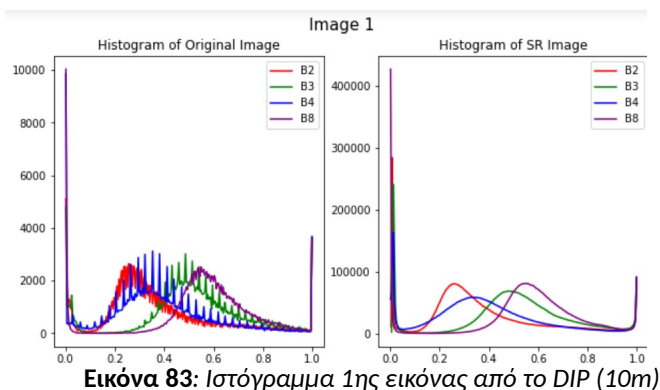
Συνεχίζοντας την ποσοτική αξιολόγηση των μοντέλων, γίνεται αντιληπτό πως το μοντέλο ZSSR ξεπερνά σε επιδόσεις το DIP στην πλειοψηφία των μετρικών, ανεξάρτητα της αρχικής χωρικής ανάλυσης των υπό επεξεργασία εικόνων. Εντυπωσιακή είναι η περίπτωση των μετρικών RMSE και PSNR, όπου το ZSSR πετυχαίνει σχεδόν διπλάσιο σηματοθορυβικό λόγο και μίας τάξης μικρότερο μέσο τετραγωνικό σφάλμα. Η μοναδική μετρική βάσει της οποίας το DIP μοντέλο φαίνεται να επιδεικνύει την βέλτιστη συμπεριφορά είναι αυτή του SRE, γεγονός το οποίο επιβεβαιώνει την ύπαρξη καθαρότερων ακμών και καλύτερης υφής στα παραγόμενα από το μοντέλο αυτό αποτελέσματα, όπως ακριβώς σημειώθηκε και στην προηγούμενη ενότητα.

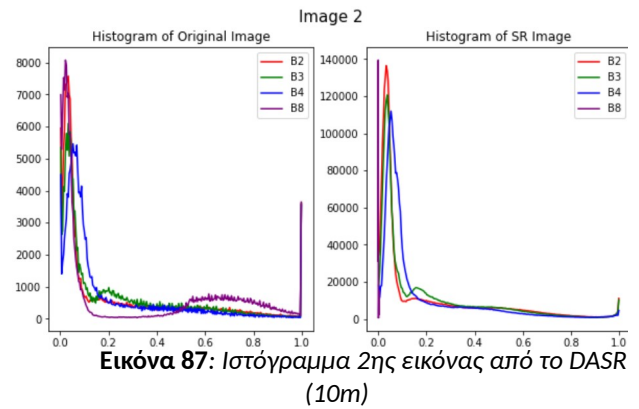
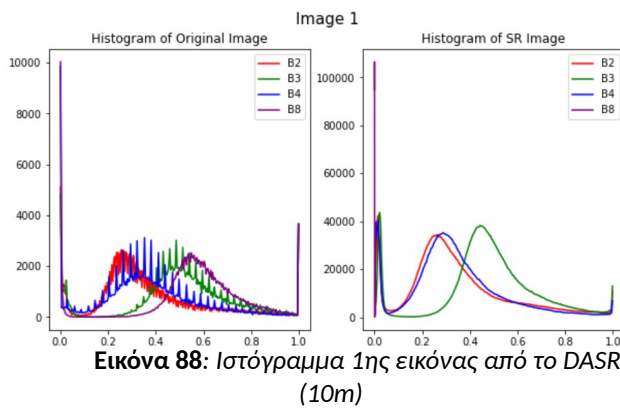
Η ποσοτική υπεροχή του ZSSR μοντέλου έναντι του DIP δεν πρέπει να θεωρηθεί πως αναιρεί τις παρατηρήσεις που έγιναν στην προηγούμενη ενότητα, βάσει των οποίων αναδείχθηκε ως βέλτιστο το μοντέλο DIP. Αντιθέτως, είναι μία ένδειξη πως το ZSSR μοντέλο αποτελεί την βέλτιστη επιλογή σε περιπτώσεις που τα αποτελέσματα του Super Resolution πρόκειται να χρησιμοποιηθούν σε επόμενες εφαρμογές, όπως παραδείγματος χάριν εφαρμογές ταξινόμησης, τμηματοποίησης των περιοχών, ανίχνευσης αντικειμένων κ.α., καθώς, όπως αποδεικνύεται από τις μετρικές, το μοντέλο αυτό είναι ικανό να διατηρεί περισσότερες χωρικές και φασματικές πληροφορίες. Στην περίπτωση, ωστόσο, που ζητούμενο είναι η σύνθεση εικόνων βελτιωμένης χωρικής ανάλυσης για οπτική ανάλυση και επεξεργασία, τότε το DIP μοντέλο συνεχίζει να αποτελεί, όπως φάνηκε στην προηγούμενη ενότητα, την βέλτιστη επιλογή.

4.6.2 Σύγκριση Ιστογραμμάτων

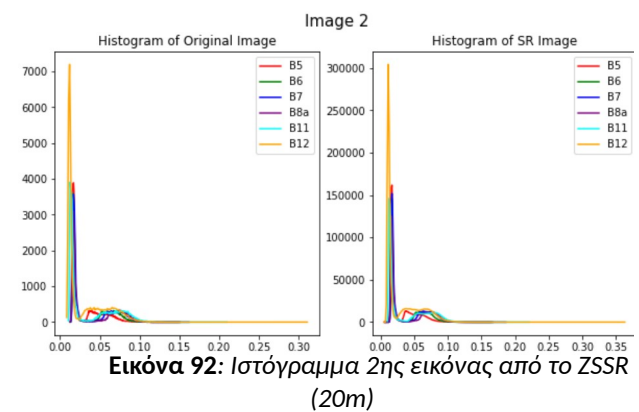
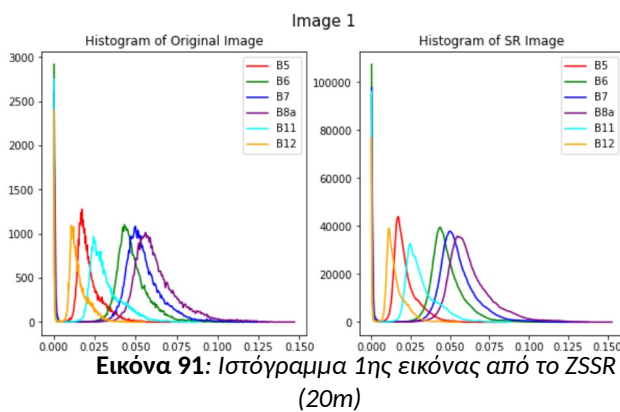
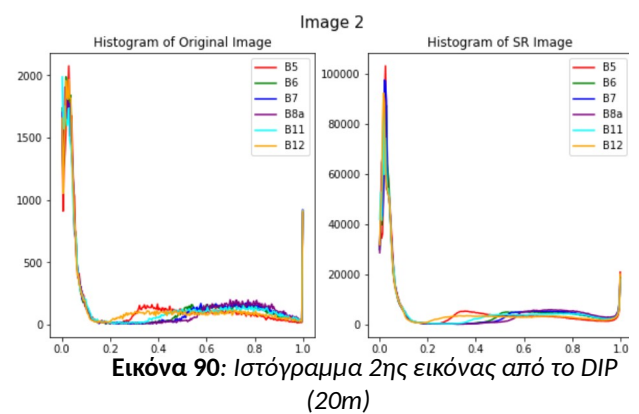
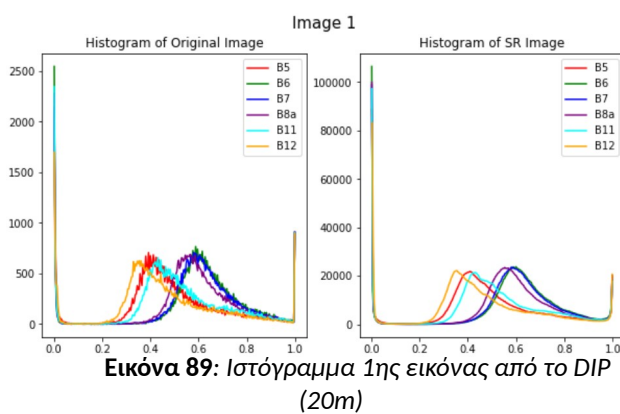
Στην συγκεκριμένη υποενότητα πραγματοποιείται σύγκριση των ιστογραμμάτων μεταξύ των αρχικών εικόνων και των εικόνων που παράγονται από τρία μοντέλα που μελετήθηκαν στην παρούσα εργασία. Συγκεκριμένα, όπως και στην προηγούμενη ενότητα, λαμβάνονται υπόψη τα βέλτιστα ανά μοντέλο αποτελέσματα και συγκρίνονται ξεχωριστά τα ιστογράμματα των δύο πρώτων εικόνων, για κάθε μία από τις τρεις αρχικές χωρικές αναλύσεις των εικόνων του συνόλου ελέγχου.

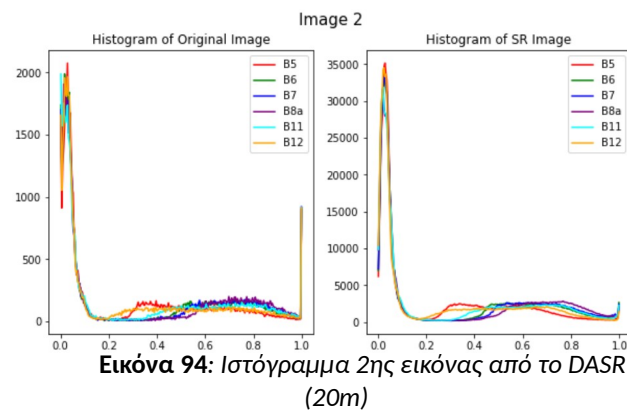
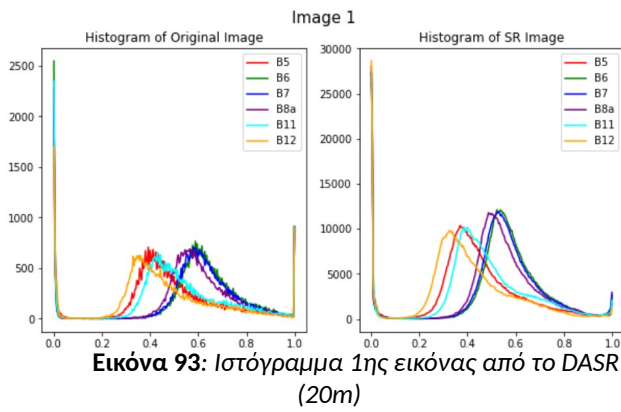
Εικόνες 10m GSD



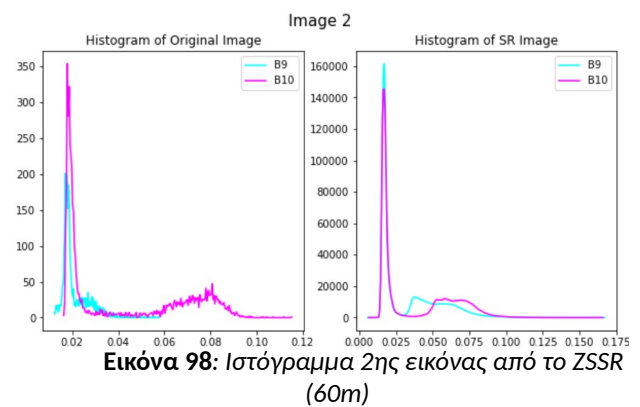
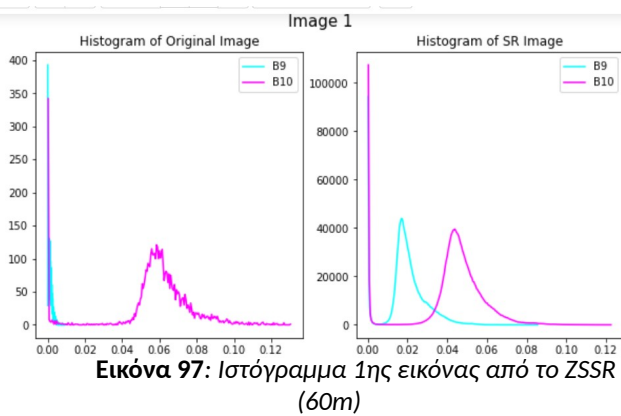
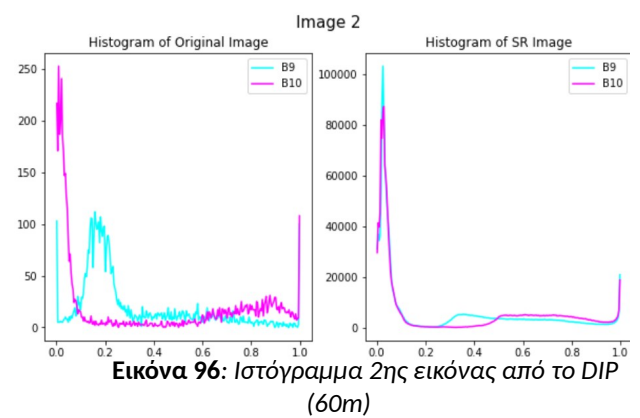
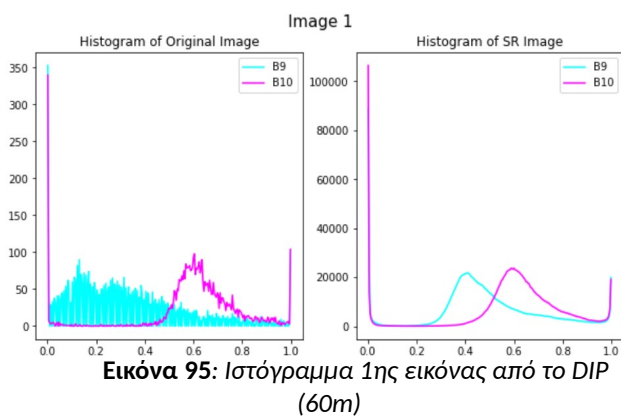


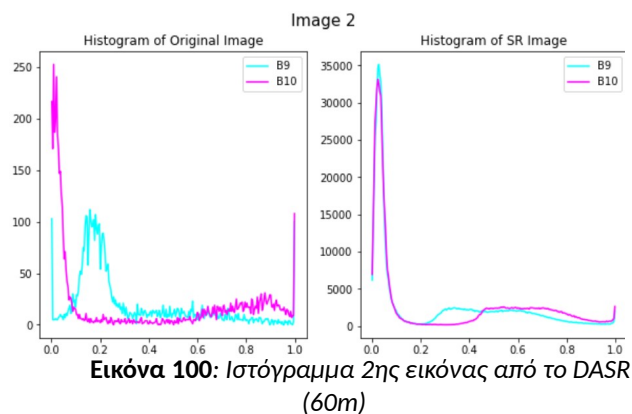
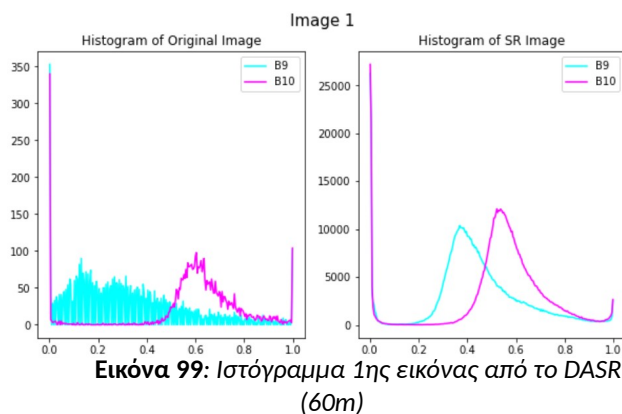
Εικόνες 20m GSD





Εικόνες 60m GSD





Αρχικά, εξετάζοντας τα ιστογράμματα των αρχικών εικόνων και των αποτελεσμάτων των τριών μοντέλων, ανεξάρτητα από την χωρική ανάλυση της εικόνας εισόδου, γίνεται αντιληπτό πως τα ιστογράμματα των δεύτερων αποτελούν, σε κάθε περίπτωση, μία πιο ομαλή εκδοχή των αντιστοιχών ιστογραμμάτων των εικόνων εισόδου. Η συμπεριφορά αυτή είναι επιθυμητή και αναμενόμενη, καθώς με την αύξηση της ανάλυσης των εικόνων, του πλήθους δηλαδή των εικονοστοιχείων που περιέχονται σε αυτές, οι μεταβάσεις μεταξύ διαδοχικών τιμών των pixel γίνονται πιο ομαλές, γεγονός που αντικατοπτρίζεται στην μορφή των ιστογραμμάτων.

Συγκρίνοντας τα ιστογράμματα των παραγόμενων εικόνων με τα αντίστοιχα ιστογράμματα των εικόνων 10m και 20m GSD, μπορεί κανείς να παρατηρήσει πως την μεγαλύτερη απόκλιση μεταξύ αρχικού και παραγόμενου ιστογράμματος παρουσιάζει το μοντέλο DASR, γεγονός το οποίο επιβεβαιώνει για ακόμη μία φορά τις χαμηλές επιδόσεις του αναφορικά με την βελτίωση της χωρικής ανάλυσης των δορυφορικών εικόνων που μελετώνται στην παρούσα εργασία. Επιπλέον, στις περιπτώσεις των εικόνων με χωρική ανάλυση 10m και 20m GSD, αξίζει να σημειωθεί πως τα μοντέλα DIP και ZSSR παράγουν ιστογράμματα τα οποία ταυτίζονται, σχεδόν, απόλυτα με αυτά των αρχικών εικόνων. Η μεγάλη αυτή ομοιότητα μεταξύ αρχικών και τελικών ιστογραμμάτων ήταν αναμενόμενη, με βάση τις μετρικές που μελετήθηκαν στην προηγούμενη ενότητα, υποδεικνύοντας την ισχυρή ικανότητα των μοντέλων DIP και ZSSR να διατηρούν και να μεταφέρουν τόσο το φασματικό όσο και το χωρικό περιεχόμενο των αρχικών εικόνων στις μεγεθυμένες εκδοχές τους. Στην περίπτωση των εικόνων με χωρική ανάλυση 60m GSD και τα τρία μοντέλα φαίνεται να αδυνατούν να διατηρήσουν την ακριβή μορφή των αρχικών ιστογραμμάτων, γεγονός το οποίο μεταφράζεται ως εμφάνιση φασματικών ή και χωρικών αλλοιώσεων στο περιεχόμενο της εκάστοτε εικόνας. Μεταξύ των τριών μοντέλων, τα ιστογράμματα που παράγονται από το DASR μοντέλο είναι αυτά που εμφανίζουν την μεγαλύτερη ασυνέπεια ως προς τα αρχικά, ακολουθούμενα από αυτά του DIP και τέλος του ZSSR.

Στο σημείο αυτό, αξίζει να σημειωθεί πως τα συμπεράσματα στα οποία καταλήγει κανείς μέσω της αξιολόγησης των ιστογραμμάτων των αποτελεσμάτων των τριών μοντέλων έρχονται σε απόλυτη συμφωνία με τα αντίστοιχα συμπεράσματα που εξήχθησαν στην προηγούμενη υποενότητα, πιστοποιώντας με αυτό τον τρόπο την εγκυρότητα της ποσοτικής αξιολόγησης των μοντέλων. Τέλος, κρίνεται σκόπιμο να αναφερθεί πως η ανομοιομορφία μεταξύ των ιστογραμμάτων των αρχικών εικόνων στις περιπτώσεις των μοντέλων DIP, DASR και των αντιστοιχών εικόνων που χρησιμοποιούνται στην περίπτωση του ZSSR μοντέλου, οφείλονται αποκλειστικά και μόνο στο γεγονός ότι στην πρώτη περίπτωση οι εικόνες υπέστησαν histogram stretching, κάτι το οποίο δεν συνέβη στην περίπτωση του ZSSR.

4.7. Συνολική αξιολόγηση

Στην ενότητα αυτή παρουσιάζεται μία συνοπτική επισκόπηση, υπό την μορφή πίνακα, των βέλτιστων μοντέλων ως την οπτική και ποσοτική ποιότητα των αποτελεσμάτων, καθώς και ως προς τον απαιτούμενο (συνολικό) χρόνο για την σύνθεση των εικόνων βελτιωμένης χωρικής ανάλυσης. Στο παρακάτω πίνακα παρατίθενται στην πρώτη διάσταση τα μοντέλα και στην δεύτερη διάσταση οι παράμετροι αξιολόγησης, ενώ με το σύμβολο “✓” σημειώνεται βέλτιστο μοντέλο ως την συγκεκριμένη παράμετρο.

Πίνακας 10: Συνολική αξιολόγηση των μοντέλων.

Μοντέλο / Παράμετροι αξιολόγησης	Οπτική ποιότητα	Ποσοτική ποιότητα	Χρόνος εκτέλεσης
DIP	✓		
ZSSR		✓	✓
DASR			

4.8. Τεχνικά χαρακτηριστικά των υπολογιστικών συστημάτων

Στα πλαίσια της εκτέλεσης του πειραματικού μέρους της παρούσας διπλωματικής εργασίας αξιοποιήθηκαν δύο υπολογιστικά συστήματα, με σκοπό τον παραλληλισμό των διαδικασιών. Αναλυτικότερα, το πρώτο υπολογιστικό σύστημα ήταν εξοπλισμένο με έναν AMD Ryzen 9 3900X επεξεργαστή με 12 φυσικούς και 24 λογικούς πυρήνες χρονισμένους στα 4100Hz, 32GB RAM και κάρτα γραφικών (GPU) NVIDIA GeForce RTX 3070 με 8GB VRAM. Το δεύτερο υπολογιστικό σύστημα, παραχωρήθηκε από το ΕΠΙΣΕΥ με την μορφή Virtual Machine (VM) για απομακρυσμένη πρόσβαση. Το συγκεκριμένο VM διέθετε 16 λογικούς πυρήνες, 16GB RAM και κάρτα γραφικών NVIDIA GeForce RTX 3070 με 8GB VRAM. Στο σημείο αυτό, κρίνεται σκόπιμο να σημειωθεί πως λόγω τεχνικών δυσκολιών και ασυμβατότητας λογισμικού, η εκτέλεση του αλγορίθμου του Zero-Shot Super Resolution μοντέλου δεν πραγματοποιήθηκε σε κανένα από τα προαναφερθέντα υπολογιστικά συστήματα. Αντιθέτως, η εκτέλεση του ZSSR πραγματοποιήθηκε σε ένα hosted περιβάλλον του Google Colab, στο οποίο οι διαθέσιμοι πόροι ήταν 8GB RAM, 30GB σκληρός δίσκος και μία κάρτα γραφικών NVIDIA Tesla K80 με 16GB VRAM.

5. Επίλογος

Στα πλαίσια της παρούσας διπλωματικής εργασίας μελετήθηκαν αλγόριθμοι μη επιβλεπόμενης, βαθιάς μάθησης για την βελτίωση της χωρικής ανάλυσης εικόνων που έχουν ληφθεί από τον αστερισμό δορυφόρων Sentinel-2 του ευρωπαϊκού προγράμματος Copernicus, κατά παράγοντες ίσους με 2, 4 και 6. Αναλυτικότερα, αξιοποιήθηκαν τρία διαφορετικά βαθιά νευρωνικά δίκτυα, καθένα από τα οποία ακολουθεί μία ξεχωριστή μεθοδολογία για την ανάλυση των εικόνων χαμηλής χωρικής ανάλυσης και την σύνθεση των ζητούμενων εικόνων βελτιωμένης χωρικής ανάλυσης.

Η προσέγγιση του Deep Image Prior (DIP) αξιοποιεί ένα βαθύ νευρωνικό δίκτυο με μορφή κλειψύδρας (U-Net) στο οποίο υπάρχουν Skip Connections μεταξύ των επιπέδων υποδειγματοληψίας και υπερδειγματοληψίας, επιβεβαιώνοντας ότι η χρήση της κατάλληλης αρχιτεκτονικής δικτύου είναι αρκετή για την εξαγωγή και μοντελοποίηση, υπό την μορφή των

βαρών του δικτύου, των στατιστικών εξαρτήσεων μεταξύ γειτονικών εικονοστοιχείων μίας εικόνας. Οι εξαρτήσεις αυτές, στην συνέχεια, χρησιμοποιούνται ως ένας είδος πρότερης γνώσης για την επίλυση του προβλήματος της βελτίωσης της χωρικής ανάλυσης της εκάστοτε εικόνας εισόδου.

Το Zero-Shot Super Resolution (ZSSR) μοντέλο, ακολουθεί μια εντελώς διαφορετική προσέγγιση για την σύνθεση των εικόνων βελτιωμένης χωρικής ανάλυσης. Συγκεκριμένα, το μοντέλο αυτό αξιοποιεί την επανεμφάνιση μικρών τμημάτων μίας εικόνας, στο εσωτερικό της, για την σύνθεση ενός συνόλου εκπαίδευσης υπό την μορφή ζευγών εικόνων υψηλής-χαμηλής ευκρίνειας. Τα επιμέρους τμήματα των εικόνων υποβάλλονται σε μία διαδικασία data augmentation προκειμένου να εμπλουτιστεί ακόμη περισσότερο το σύνολο δεδομένων με το οποίο θα εκπαιδευτεί το δίκτυο.

Το Degradation Aware Super Resolution (DASR) μοντέλο, βασίζεται στην ιδέα του contrastive learning για την δημιουργία αναπαραστάσεων παραμόρφωσης, κατά έναν αμιγώς μη επιβλεπόμενο τρόπο, οι οποίες χρησιμοποιούνται στην συνέχεια ως επιπρόσθετη πληροφορία για την επιβλεπόμενη εκπαίδευση ενός αρκετά βαθιού συνελκτικού δικτύου, προκειμένου τελικά να δημιουργηθούν οι εικόνες υψηλής χωρικής ανάλυσης.

Συγκρίνοντας τα αποτελέσματα που προκύπτουν από τις προαναφερθείσες μεθοδολογίες, για συντελεστές μεγέθυνσης 2, 4, 6, μπορεί κανείς εύκολα να καταλήξει στο συμπέρασμα πως ανεξάρτητα του συντελεστή το μοντέλο DASR είναι αυτό που επιδεικνύει τις χαμηλότερες επιδόσεις ως προς την ποιότητα των παραγόμενων αποτελεσμάτων. Η μερική αυτή αποτυχία του μοντέλου, μπορεί να αποδοθεί στο γεγονός ότι η εκπαίδευσή του πραγματοποιήθηκε σε συμβατικές τριχρωματικές εικόνες, των οποίων η χωρική πολυπλοκότητα είναι εμφανώς μικρότερη από αυτή των δορυφορικών εικόνων. Στον αντίποδα, τα μοντέλα DIP και ZSSR παράγουν ιδιαίτερα καλά και παρόμοια αποτελέσματα για συντελεστές μεγέθυνσης 2 και 4, ενώ το μοντέλο DIP είναι αυτό που ξεχωρίζει στην περίπτωση που ο συντελεστής είναι ίσος με 6, υποδεικνύοντας την ικανότητα του συγκεκριμένου μοντέλου να παράγει ποιοτικά ανώτερα αποτελέσματα για μεγαλύτερους συντελεστές μεγέθυνσης.

Η σύγκριση των τριών μοντέλων που εξετάστηκαν στην παρούσα εργασία συνεχίστηκε, πέρα από την ποιοτική φωτοερμηνεία των αποτελεσμάτων, με μία ποσοτική αξιολόγηση των παραγόμενων εικόνων ως προς ορισμένες μετρικές ομοιότητας εικόνων. Από την ποσοτική αξιολόγηση των μοντέλων αναδείχθηκε, ως βέλτιστο, το ZSSR πετυχαίνοντας ταυτόχρονα τους μικρότερους χρόνους σύνθεσης των εικόνων βελτιωμένης χωρικής ανάλυσης. Η εγκυρότητα της ποσοτικής αξιολόγησης των μοντέλων επιβεβαιώθηκε, τέλος, μέσω σύγκρισης των ιστογραμμάτων των παραγόμενων από το εκάστοτε μοντέλο με τις αντίστοιχες αρχικές εικόνες χαμηλής χωρικής ανάλυσης του συνόλου ελέγχου. Η ασυμφωνία μεταξύ της φωτοερμηνείας των αποτελεσμάτων και της ποσοτικής αξιολόγησης αυτών, οδηγεί στο συμπέρασμα πως η επιλογή του βέλτιστου μοντέλου είναι άμεσα συσχετισμένη με το είδος των εφαρμογών που πρόκειται να χρησιμοποιηθούν στην συνέχεια τα αντίστοιχα αποτελέσματα.

Το κύριο κοινό πλεονέκτημα των παραπάνω μεθοδολογιών, συγκριτικά με τις αντίστοιχες μεθόδους επιβλεπόμενης μάθησης, είναι ότι για την σύνθεση των εικόνων αυξημένης χωρικής ανάλυσης χρησιμοποιούνται αποκλειστικά και μόνο δεδομένα τα οποία εξάγονται από την εκάστοτε υπό εξέταση εικόνα. Το γεγονός αυτό, συνεπάγεται την εξαγωγή στατιστικών εξαρτήσεων, μεταξύ των εικονοστοιχείων, που είναι εξειδικευμένες για την εικόνα που αναλύεται σε κάθε περίπτωση, αποφεύγοντας έτσι την πραγματοποίηση υποθέσεων, που

συμβαίνει στην περίπτωση των μοντέλων επιβλεπόμενης μάθησης, σχετικά με τον τρόπο δημιουργίας των εικόνων χαμηλής ανάλυσης από της αντίστοιχες εκδοχές τους υψηλής ανάλυσης. Επιπρόσθετα, εξετάζοντας τα αποτελέσματα των δύο πρώτων μοντέλων μπορεί κανείς να επιβεβαιώσει την ισχύ της μη επιβλεπόμενης μάθησης για την δημιουργία εικόνων υψηλής χωρικής ανάλυσης σε περιπτώσεις που η απόκτηση ενός συνόλου εκπαίδευσης δεν εφικτή. Στο σημείο αυτό, αξίζει να σημειωθεί πως σε περίπτωση διάθεσης ενός κατάλληλα δομημένου συνόλου εκπαίδευσης, οι αλγόριθμοι επιβλεπόμενης μάθησης ενδέχεται να συνέθεταν εικόνες των οποίων η ποιότητα να ξεπερνούσε αυτή των αποτελεσμάτων που λαμβάνονται από τα παραπάνω δύο μοντέλα.

Ολοκληρώνοντας την συγκεκριμένη ενότητα, κρίνεται σκόπιμο να γίνει αναφορά στην συμβολή της παρούσας διπλωματικής εργασίας στην βελτίωση της χωρικής ανάλυσης δορυφορικών εικόνων. Αρχικά, προτείνεται μία γενίκευση των μεθοδολογιών των μοντέλων DIP και ZSSR, τα οποία είναι σχεδιασμένα για την επεξεργασία τριχρωματικών εικόνων με 8bit ακρίβεια, ώστε αυτά να είναι σε θέση να διαχειριστούν εικόνες αυθαίρετου πλήθους χρωματικών καναλιών των οποίων ο τύπος δεδομένων μπορεί πλέον να είναι είτε int8 είτε int16. Επιπρόσθετα, προτείνεται ένας αρκετά απλός τρόπος για την αντιμετώπιση προβλημάτων σχετικών με την διαθέσιμη μνήμη του υπολογιστικού συστήματος, μέσω της διάσπασης της αρχικής εικόνας σε επιμέρους μη επικαλυπτόμενα τμήματα των οποίων η χωρική ανάλυση βελτιώνεται σειριακά από το εκάστοτε μοντέλα. Τα προκύπτοντα επιμέρους τμήματα βελτιωμένης χωρικής ανάλυσης, χρησιμοποιούνται στην συνέχεια για την σύνθεση της τελικής εικόνας. Η μέθοδος αυτή, βέβαια, επηρεάζει σημαντικά τον χρόνο που απαιτείται για την ολοκλήρωση της αλγορίθμου. Τέλος, ειδικά για το DIP μοντέλο, υλοποιήθηκε μία μέθοδος Early Stopping με σκοπό την αποφυγή της υπερεκπαίδευσης του δικτύου πάνω σε μία εικόνα, γεγονός το οποίο οδηγούσε σε μείωση της ποιότητας του τελικού αποτελέσματος.

5.1. Μελλοντικές προοπτικές

Η παρούσα διπλωματική εργασία αξιοποιεί σύγχρονες, με βάση την τωρινή βιβλιογραφική επισκόπηση, μεθόδους βαθιάς, μη επιβλεπόμενης μάθησης για την σύνθεση εικόνων βελτιωμένης χωρικής ανάλυσης, οι οποίες έχουν μελετηθεί σε μικρό βαθμό μέχρι στιγμής. Το πρώιμο στάδιο ανάπτυξης των μη επιβλεπόμενων αλγορίθμων σε εφαρμογές Super Resolution, σε συνδυασμό με τη συνεχή ανάπτυξη της τηλεπισκόπησης, δημιουργεί αρκετές προοπτικές για μελλοντική εξέλιξη της συγκεκριμένης εργασίας και των μεθοδολογιών που προτείνονται σε αυτήν.

Αρχικά, ιδιαίτερο ενδιαφέρον παρουσιάζει η επέκταση των συγκεκριμένων αλγορίθμων σε εικόνες που έχουν ληφθεί από άλλους δορυφόρους, πέρα του Sentinel-2, οι οποίες μπορεί να αποτελούνται από ακόμη περισσότερα κανάλια ή να διαθέτουν καλύτερη αρχική χωρική ανάλυση. Επιπρόσθετα, για την περίπτωση του DIP μοντέλου, το οποίο ήταν το μόνο μεταξύ των τριών που δείγματα γενίκευσης της λειτουργίας του για συντελεστές μεγέθυνσης μεγαλύτερου του 4, θα μπορούσε να δοκιμαστεί η βελτίωση της χωρικής ανάλυσης εικόνων κατά ακόμη μεγαλύτερους παράγοντες, με σκοπό αφενός την λήψη εικόνων ακόμη καλύτερης ανάλυσης και αφετέρου να διερευνηθούν τα όρια μεγέθυνσης του συγκεκριμένου μοντέλου.

Μια διαφορετική κατεύθυνση για την συγκεκριμένη εργασία, θα μπορούσε να είναι η δημιουργία μίας προγραμματιστικής βιβλιοθήκης αποτελούμενη από πολλαπλά διαφορετικά μοντέλα βαθιάς, μη επιβλεπόμενης μάθησης και μεθοδολογίες που θα αποσκοπούν στην επεξεργασία, γενικότερα, δορυφορικών εικόνων, αλλά και στην βελτίωση της χωρικής ανάλυσης

αυτών. Μία τέτοια βιβλιοθήκη θα αποτελούσα μία ολοκληρωμένη λύση, η οποία θα επέτρεπε στους χρήστες να συνθέτουν εικόνες υψηλής χωρικής ανάλυσης με ευκολία και σχετική ταχύτητα.

Επιπρόσθετα, αναφορικά με το DASR μοντέλο, θα μπορούσε να προταθεί μία πλήρως μη επιβλεπόμενη υλοποίησή του. Αναλυτικότερα, διατηρώντας την αρχιτεκτονική του επιβλεπόμενου μοντέλου που συνθέτει τις εικόνες βελτιωμένης ανάλυσης, θα μπορούσε κανείς να τροποποιήσει τον τρόπο υπολογισμού του σφάλματος εξόδου κατά τρόπο παρόμοιο με αυτό που χρησιμοποιείται στην προσέγγιση του Deep Image Prior, έτσι ώστε για την συνολική εκπαίδευση του μοντέλου να απαιτείται μόνο η υπό εξέταση εικόνα χαμηλής χωρικής ανάλυσης. Η προαναφερθείσα τροποποίηση, παρόλο που δεν είναι βέβαιο ότι θα βελτιώσει την ποιότητα των αποτελεσμάτων, θα επέτρεπε σίγουρα την γενίκευση του μοντέλου έτσι ώστε αυτό να είναι σε θέση να διαχειριστεί εικόνες με 16bit ακρίβεια και αυθαίρετο πλήθος καναλιών. Προς την κατεύθυνση γενίκευσης του μοντέλου με σκοπό την επεξεργασία εικόνων με ακρίβεια 16bit, θα μπορούσε κανείς να επιχειρήσει την επανεκπαίδευση του μοντέλου αποκλειστικά με δορυφορικές εικόνες ίδιας ακρίβειας, χωρίς την ανάγκη πραγματοποιήσεις των αλλαγών που αναφέρθηκαν περί αλλαγής του είδους μάθησης του μοντέλου. Η χρήση ενός τέτοιου συνόλου εκπαίδευσης, δεδομένου ότι αυτό μπορεί να βρεθεί, θεωρείται πως θα οδηγούσε σε λήψη ποιοτικότερων αποτελεσμάτων, χωρίς ωστόσο να επιτρέπει την επεξεργασία εικόνων αυθαίρετου πλήθους καναλιών.

Μία, ακόμη, διαφορετική προοπτική της παρούσας εργασίας είναι η ανάπτυξη αλγορίθμων ημι-επιβλεπόμενης μάθησης για την επίλυση του προβλήματος του Super Resolution. Αναλυτικότερα, θα μπορούσαν να αξιοποιηθούν οι αλγόριθμοι μη επιβλεπόμενης μάθησης της παρούσας εργασίας για την δημιουργία συνόλων εκπαίδευσης, μέσω παραδείγματος χάριν μίας διαδικασίας clustering, τα οποία στην συνέχεια θα χρησιμοποιούνταν για την εκπαίδευση ενός επιβλεπόμενου νευρωνικού δικτύου του οποίου στόχος θα είναι η σύνθεση εικόνων υψηλής χωρικής ανάλυσης.

Επιπλέον, με βάση την αρχική ιδέα για επέκταση των συγκεκριμένων μοντέλων σε εικόνες από διαφορετικούς δορυφόρους, προτείνεται μία συμπληρωματική μελλοντική προοπτική. Στα πλαίσια, λοιπόν, της αξιοποίησης δεδομένων από διαφορετικούς δορυφόρους θα μπορούσε να εξεταστεί η δυνατότητα των μοντέλων αυτών να χρησιμοποιούν εικόνες, της ίδιας περιοχής, που έχουν ληφθεί από διαφορετικά δορυφορικά συστήματα για την σύνθεση, τελικά, μίας εικόνας βελτιωμένης χωρικής ανάλυσης.

Τέλος, ιδιαίτερο ενδιαφέρον παρουσιάζει η επέκταση των συγκεκριμένων μοντέλων πέρα από τα όρια της βελτίωσης της χωρικής ανάλυσης εικόνων, προς την κατεύθυνση της βελτίωσης της χρονικής ανάλυσης. Αναλυτικότερα, θα μπορούσε να εξεταστεί η εφαρμογή των μεθοδολογιών αυτών στις περιπτώσεις high temporal δορυφορικών εικόνων, για την μη επιβλεπόμενη σύνθεση νέων εικόνων, που απεικονίζουν την ίδια περιοχή με μία μικρή μετατόπιση, για τον εμπλουτισμό χρονοσειρών.

6. Βιβλιογραφία

- [1] J. B. Campbell and R. H. Wynne., *Introduction to Remote Sensing FIFTH EDITION*, vol. XXXIII, no. 2. 2011.
- [2] S. L. Ozesmi and M. E. Bauer, "Satellite remote sensing of wetlands," 2002.
- [3] Q. Yuan *et al.*, "Deep learning in environmental remote sensing: Achievements and challenges," *Remote Sens. Environ.*, vol. 241, no. February, p. 111716, 2020, doi: 10.1016/j.rse.2020.111716.
- [4] G. Simone, A. Farina, F. C. Morabito, S. B. Serpico, and L. Bruzzone, "Image fusion techniques for remote sensing applications," *Inf. Fusion*, vol. 3, no. 1, pp. 3–15, 2002, doi: 10.1016/S1566-2535(01)00056-2.
- [5] Kumar N, Yamaç SS, and Velmurugan A, "Applications of Remote Sensing and GIS in Natural Resource Management," vol. 20, no. 1, pp. 1–6, 2015.
- [6] ESA, "ESA- ESA facts," 2020. https://www.esa.int/About_Us/Corporate_news/ESA_facts (accessed May 31, 2021).
- [7] "ESA - Member States & Cooperating States." http://www.esa.int/About_Us/Corporate_news/Member_States_Cooperating_States (accessed May 31, 2021).
- [8] "ESA - European Service Module." https://www.esa.int/Science_Exploration/Human_and_Robotic_Exploration/Orion/European_Service_Module (accessed May 31, 2021).
- [9] "ESA - Earth Observation Programmes," 2021, Accessed: Jun. 03, 2021. [Online]. Available: http://www.esa.int/About_Us/Business_with_ESA/Business_Opportunities/Earth_Observation_Programmes.
- [10] "Europe's Copernicus programme," *esa.int*. http://www.esa.int/Applications/Observing_the_Earth/Copernicus/Europe_s_Copernicus_programme (accessed Jun. 03, 2021).
- [11] European Space Agency, "User Guides - Sentinel-2 MSI - Overview - Sentinel Online - Sentinel." <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/overview> (accessed Jun. 03, 2021).

- [12] "ESA - Introducing Sentinel-2." http://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2/Introducing_Sentinel-2 (accessed Jun. 03, 2021).
- [13] "ESA - Facts and figures." http://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2/Facts_and_figures (accessed Jun. 03, 2021).
- [14] ESA, "User Guides - Sentinel-2 MSI - Revisit and Coverage - Sentinel Online - Sentinel," 2021. <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/revisit-coverage> (accessed Jun. 03, 2021).
- [15] ESA, "MSI Instrument - Sentinel-2 MSI Technical Guide - Sentinel Online - Sentinel," 2021. <https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-2-msi/msi-instrument> (accessed Jun. 03, 2021).
- [16] "ESA - Instrument." http://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2/Instrument (accessed Jun. 03, 2021).
- [17] "ESA - Data flow." http://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2/Data_flow (accessed Jun. 03, 2021).
- [18] "Products and Algorithms - Sentinel-2 MSI Technical Guide - Sentinel Online - Sentinel." <https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-2-msi/products-algorithms> (accessed Jun. 03, 2021).
- [19] ESA-Sentinel 2, "User Guides - Sentinel-2 MSI - Sentinel Online - Sentinel Online," 2015. <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/processing-levels> (accessed Jun. 03, 2021).
- [20] J. M. Haut, R. Fernandez-Beltran, M. E. Paoletti, J. Plaza, A. Plaza, and F. Pla, "A new deep generative network for unsupervised remote sensing single-image super-resolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 11, pp. 6792–6810, 2018, doi: 10.1109/TGRS.2018.2843525.
- [21] D. Han, "Comparison of Commonly Used Image Interpolation Methods," 2013, doi: 10.2991/iccsee.2013.391.
- [22] P. Parsania, P. VVirparia, and A. Professor, "A Review: Image Interpolation Techniques for Image Scaling," *Int. J. Innov. Res. Comput. Commun. Eng. (An ISO)*, vol. 3297, 2007, doi: 10.15680/ijircce.2014.0212024.

- [23] A. Prajapati, S. Naik, A. In, and S. Mehta, "Evaluation of Different Image Interpolation Algorithms," *Int. J. Comput. Appl.*, vol. 58, no. 12, pp. 975–8887, 2012, Accessed: Jan. 17, 2022. [Online]. Available: <http://www.dannyruijters>.
- [24] B. Suh, H. Ling, B. B. Bederson, and D. W. Jacobs, "Automatic thumbnail cropping and its effectiveness," in *UIST: Proceedings of the Annual ACM Symposium on User Interface Software and Technology*, 2003, pp. 95–104, doi: 10.1145/964696.964707.
- [25] M. Zhang, L. Zhang, Y. Sun, L. Feng, and W. Ma, "Auto cropping for digital photographs," in *IEEE International Conference on Multimedia and Expo, ICME 2005*, 2005, vol. 2005, pp. 4–7, doi: 10.1109/ICME.2005.1521454.
- [26] G. Vivone *et al.*, "A Critical Comparison Among Pansharpening Algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, 2015, doi: 10.1109/TGRS.2014.2361734.
- [27] X. Li *et al.*, "A Remote-Sensing Image Pan-Sharpener Method Based on Multi-Scale Channel Attention Residual Network," *IEEE Access*, vol. 8, pp. 27163–27177, 2020, doi: 10.1109/ACCESS.2020.2971502.
- [28] P. S. Huang and T. M. Tu, "Reply to Erratum to 'A new look at IHS-like image fusion methods,'" *Information Fusion*, vol. 8, no. 2 SPEC. ISS. p. 218, 2007, doi: 10.1016/j.inffus.2006.10.006.
- [29] K. A. Kalpoma and J. I. Kudoh, "Image fusion processing for IKONOS 1-m color imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3075–3086, 2007, doi: 10.1109/TGRS.2007.897692.
- [30] G. Masi, D. Cozzolino, L. Verdoliva, and G. Scarpa, "Pansharpening by convolutional neural networks," *Remote Sens.*, vol. 8, no. 7, p. 594, Jul. 2016, doi: 10.3390/rs8070594.
- [31] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP Graph. Model. Image Process.*, vol. 53, no. 3, pp. 231–239, 1991, doi: 10.1016/1049-9652(91)90045-L.
- [32] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *IEEE Trans. Image Process.*, vol. 13, no. 10, pp. 1327–1344, 2004, doi: 10.1109/TIP.2004.834669.
- [33] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 349–356, 2009, doi: 10.1109/ICCV.2009.5459271.
- [34] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-Based Super-Resolution," 2002. [Online]. Available: <http://www.altamira.com>.
- [35] K. I. Kim and Y. Kwon, "Example-based Learning for and JPEG Artifact Removal Example-based Learning for Single-Image Super-Resolution and JPEG Artifact

Removal,” *Biol. Cybern.*, no. August, 2008, [Online]. Available: <http://www.kyb.tuebingen.mpg.de/techreports.html>.

[36] S. Baker and T. Kanade, “Hallucinating faces,” in *Proceedings - 4th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2000*, 2000, pp. 83–88, doi: 10.1109/AFGR.2000.840616.

[37] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT Press, 2016.

[38] T. M. Mitchell, *Machine learning*. McGraw-Hill Science/Engineering/Math, 1997.

[39] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*. Elsevier’s Science & Technology Rights Department in Oxford, 2009.

[40] C. M. Bishop, *Pattern Recognition and Machine Learning*. 2006.

[41] F. Rosenblatt, “The perceptron: A probabilistic model for information storage and organization in the brain,” *Psychol. Rev.*, vol. 65, no. 6, pp. 386–408, 1958, doi: 10.1037/h0042519.

[42] R. C. Gonzalez and R. E. Woods, *4TH EDITION Digital image processing*, 4th ed. Pearson, 2008.

[43] Y. LeCun *et al.*, “Backpropagation Applied to Handwritten Zip Code Recognition,” *Neural Comput.*, vol. 1, no. 4, pp. 541–551, Dec. 1989, doi: 10.1162/neco.1989.1.4.541.

[44] R. Patel and S. Patel, “A comprehensive study of applying convolutional neural network for computer vision,” *Int. J. Adv. Sci. Technol.*, vol. 29, no. 6 Special Issue, pp. 2161–2174, 2020.

[45] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” 2016. doi: 10.1109/CVPR.2016.90.

[46] C. Lanaras, J. Bioucas-Dias, S. Galliani, E. Baltsavias, and K. Schindler, “Super-resolution of Sentinel-2 images: Learning a globally applicable deep neural network,” *ISPRS J. Photogramm. Remote Sens.*, vol. 146, no. September, pp. 305–319, 2018, doi: 10.1016/j.isprsjprs.2018.09.018.

[47] D. Ulyanov, A. Vedaldi, and V. Lempitsky, “Deep Image Prior,” Nov. 2017, doi: 10.1007/s11263-020-01303-4.

[48] A. Shocher, N. Cohen, and M. Irani, “Zero-Shot Super-Resolution Using Deep Internal Learning,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2018, pp. 3118–3126, doi: 10.1109/CVPR.2018.00329.

- [49] Y. Xian, B. Schiele, and Z. Akata, "Zero-shot learning - The good, the bad and the ugly," in *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 2017, vol. 2017-Janua, pp. 3077–3086, doi: 10.1109/CVPR.2017.328.
- [50] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced Deep Residual Networks for Single Image Super-Resolution," 2017. doi: 10.1109/CVPRW.2017.151.
- [51] M. Irani and S. Peleg, "Improving Resolution by Image Registration," 1991.
- [52] J. Gu, H. Lu, W. Zuo, and C. Dong, "Blind Super-Resolution With Iterative Kernel Correction."
- [53] L. Wang *et al.*, "Unsupervised Degradation Representation Learning for Blind Super-Resolution," 2021, pp. 10576–10585, doi: 10.1109/cvpr46437.2021.01044.
- [54] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A Simple Framework for Contrastive Learning of Visual Representations," 2020. [Online]. Available: <https://github.com/google-research/simclr>.
- [55] X. Chen, H. Fan, R. Girshick, and K. He, "Improved Baselines with Momentum Contrastive Learning."