



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ ΚΑΙ ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ -
ΜΗΧΑΝΙΚΩΝ ΓΕΩΠΛΗΡΟΦΟΡΙΚΗΣ

ΤΟΜΕΑΣ ΤΟΠΟΓΡΑΦΙΑΣ

Σύγκριση μεθόδων μηχανικής μάθησης για τον εντοπισμό φθορών σε μνημεία

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Χρυσογόνος Νικόλαος

Επιβλέπων : Δουλάμης Αναστάσιος
Αν. Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2022



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ ΚΑΙ ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ -
ΜΗΧΑΝΙΚΩΝ ΓΕΩΠΛΗΡΟΦΟΡΙΚΗΣ

ΤΟΜΕΑΣ ΤΟΠΟΓΡΑΦΙΑΣ

Σύγκριση μεθόδων μηχανικής μάθησης για τον εντοπισμό φθορών σε μνημεία

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Χρυσογόνος Νικόλαος

Επιβλέπων : **Δουλάμης Αναστάσιος**
Αν. Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 21^η Οκτωβρίου 2022.

.....
Γεωργόπουλος Ανδρέας
Καθηγητής Ε.Μ.Π.

.....
Δουλάμης Αναστάσιος
Αν. Καθηγητής Ε.Μ.Π.

.....
Στρατηγέα Αναστασία
Καθηγήτρια Ε.Μ.Π.

Αθήνα, Οκτώβριος 2022

.....
Χρυσογόνος Νικόλαος

Διπλωματούχος Αγρονόμος και Τοπογράφος Μηχανικός Ε.Μ.Π.

Copyright © Χρυσογόνος Νικόλαος, 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 συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Καθώς η κλιματική αλλαγή επιφέρει ακραίες συνθήκες στον πλανήτη, ερευνάται η κατάσταση στρατηγικής για την προστασία των μνημείων, με πρώτο στάδιο τον αυτόματο εντοπισμό παλαιών και νέων φθορών τους. Στο πλαίσιο αυτό, η παρούσα εργασία εξετάζει την εφαρμογή αλγορίθμων μηχανικής μάθησης σε εικόνες μνημείου με φθορές και επιχειρείται η σύγκριση μεθόδων επιβλεπόμενης και μη επιβλεπόμενης μάθησης. Συγκεκριμένα, εξετάζεται η αποτελεσματικότητα ενός μοντέλου συνελκτικού νευρωνικού δικτύου VGG-16 εκπαιδευμένου σε φασματικές εικόνες RGB, ενός μοντέλου συνελκτικού δικτύου VGG-16 εκπαιδευμένου σε φασματικές εικόνες με 3 κανάλια εγγύς υπέρυθρου και ενός αλγορίθμου ομαδοποίησης K-means στον εντοπισμό φθορών πάνω σε εικόνες του φρουρίου Αγίου Νικολάου στη Ρόδο. Η σύγκριση των αποτελεσμάτων γίνεται τόσο με δείκτες ορθότητας, ανάκλησης, ακρίβειας και F1, όσο και οπτικά, συγκρίνοντας τα αποτελέσματα του Class Activation Mapping των συνελκτικών δικτύων και την οπτικοποίηση του αλγορίθμου ομαδοποίησης με τις εικόνες επισημάνσεων φθορών. Τα μοντέλα νευρωνικών δικτύων VGG φαίνεται να έχουν μεγάλο δείκτη ορθότητας στα δεδομένα εκπαίδευσης και επικύρωσης, 85% για το μοντέλο εκπαιδευμένο σε RGB φασματικά κανάλια και 92% για το μοντέλο εκπαιδευμένο σε εγγύς υπέρυθρα φασματικά κανάλια, συγκριτικά με το 61% δείκτη ορθότητας της ομαδοποίησης K-means, οπτικά ωστόσο τα αποτελέσματα του Class Activation Mapping εντοπίζουν μόνο έναν συγκεκριμένο τύπο φθορών, ενώ εμφανίζουν έντονες λανθασμένες ενεργοποιήσεις στα όρια των παραθύρων εκπαίδευσης. Τέλος, προτείνονται εναλλακτικές επιλογές φασματικών καναλιών αλλά και αρχιτεκτονικών νευρωνικών δικτύων για περαιτέρω διερεύνηση, ώστε με πιθανή καλύτερη αξιοποίηση των στοιχείων των δεδομένων χωρίς την απώλεια πληροφορίας να επιτευχθεί η εξαγωγή καλύτερων αποτελεσμάτων στον εντοπισμό φθορών.

Λέξεις κλειδιά: Μηχανική Μάθηση, Συνελκτικά Νευρωνικά Δίκτυα, Ομαδοποίηση, Φθορές, Παρακολούθηση Δομικής Ακεραιότητας, Προστασία Μνημείων

Abstract

As climate change causes extreme conditions on the planet, research focuses on constructing a strategy for the protection of monuments, starting with the automatic detection of past and ongoing deterioration. In this context, this thesis examines the application of Machine Learning Algorithms on images of a monument with present deterioration, while attempting to compare the effectiveness of supervised and unsupervised learning techniques. Specifically, the thesis compares the effectiveness of a VGG-16 convolutional neural network trained on images with only RGB bands, a VGG-16 convolutional neural network trained on hyperspectral images with 3 near infrared spectral bands, and, finally, a K-means clustering algorithm, on locating wears in images of the Saint Nicholas Fortress in Rhodes. The comparison is performed both by metrics, namely accuracy, recall, precision and F1, and visually, by comparing the results of Class Activation Mapping on the convolutional neural networks and the visualization of the clustering algorithm with the annotated images of damages. The VGG Network models seem to achieve high accuracy on the training and validation data, 85% for the RGB trained model and 92% for the near infrared model, compared to the 61% accuracy of the K-means clustering model, although visually the Class Activation Mapping seems to only locate one type of damage, while having multiple errors on the borders of the training segments. Finally, the thesis proposes alternative spectral band and architecture choices for additional research, so that the better possible utilization of the available data, without the loss of information, can lead to better results on the task of locating deterioration.

Keywords: Machine Learning, Convolutional Neural Networks, Clustering, Deterioration, Structural Health Monitoring, Monument Protection

Ευχαριστίες

Θα ήθελα αρχικά να ευχαριστήσω τον κ. Αναστάσιο Δουλάμη, για την εμπιστοσύνη που μου επέδειξε με την ανάθεση της διπλωματικής εργασίας, καθώς και για τις συμβουλές του ανά τα χρόνια μου ως προπτυχιακός φοιτητής της Σ.Α.Τ.Μ. Θα ήθελα επίσης να ευχαριστήσω τον μεταδιδακτορικό ερευνητή Παναγιώτη Αγραφιώτη και τον Υποψήφιο Διδάκτορα Ιωάννη Τζώρτζη για τις παραπομπές τους σε σχετική βιβλιογραφία και την γενικότερη βοήθειά τους τόσο σε προγραμματιστικά θέματα, όσο και στην απόκτηση των απαραίτητων δεδομένων για την εκπόνηση της παρούσας εργασίας.

Θα ήθελα, ακόμα, να ευχαριστήσω το οικογενειακό και φιλικό μου περιβάλλον, για όλη την έμπρακτη υποστήριξή τους όλο αυτό το χρονικό διάστημα, χωρίς την οποία δεν θα ήταν εφικτή η σεμνή ακαδημαϊκή μου πορεία.

Τέλος, οφείλω ένα ευχαριστώ στους συμφοιτητές και τους καθηγητές μου στην σχολή Α.Τ.Μ. του Ε.Μ.Π. για την ανταλλαγή των απόψεων μας, την επιρροή τους στον τρόπο σκέψης μου και την καλλιέργεια ενός ακαδημαϊκού περιβάλλοντος στο οποίο μπόρεσα να εξελιχθώ σαν άτομο.

Χρυσογόνος Νικόλαος

Αθήνα, Οκτώβριος 2022

Πίνακας περιεχομένων

| | |
|--|-----------|
| Περίληψη | iii |
| Abstract | iv |
| Ευχαριστίες | v |
| Εισαγωγή..... | 1 |
| 1. Επίδραση των μνημείων στον περιφερειακό σχεδιασμό | 2 |
| 1.1 Ανάγκη προστασίας των Μνημείων | 2 |
| 1.1.1 Λογική και συναισθηματική αξία..... | 2 |
| 1.1.2 Νομική υποχρέωση | 3 |
| 1.2 Απειλές για τη δομική ακεραιότητα μνημείων | 4 |
| 1.3 Ο εντοπισμός φθορών ως πληροφορία για τον χωρικό σχεδιασμό | 6 |
| 2. Γενικά Στοιχεία Μηχανικής Μάθησης | 8 |
| 2.1 Η ανάπτυξη της μηχανικής μάθησης..... | 8 |
| 2.2 Αρχή λειτουργίας της μηχανικής μάθησης | 10 |
| 2.3 Χρησιμότητα της μηχανικής μάθησης στην καθημερινότητα | 18 |
| 2.4 Επιφυλάξεις για τη χρήση μηχανικής μάθησης | 19 |
| 2.5 Η στάση της ευρωπαϊκής ένωσης | 22 |
| 3. Βασικοί Αλγόριθμοι Μηχανικής Μάθησης | 25 |
| 3.1 Επιβλεπόμενη Μάθηση | 25 |
| 3.1.1 Γενικά..... | 25 |
| 3.1.2 Γραμμική Παλινδρόμηση | 25 |
| 3.1.3 Λογιστική Παλινδρόμηση..... | 27 |
| 3.1.4 Νευρωνικά Δίκτυα..... | 29 |
| 3.1.5 Συνελκτικά νευρωνικά δίκτυα..... | 32 |
| 3.1.6 Μοντέλα εντοπισμού αντικειμένων..... | 36 |
| 3.1.7 Αρχιτεκτονικές νευρωνικών δικτύων | 38 |
| 3.1.8 Μεταφορά Μάθησης | 40 |
| 3.1.9 Κ-Εγγύτερος Γείτονας | 41 |
| 3.2 Μη επιβλεπόμενη μάθηση | 43 |
| 3.2.1 Γενικά..... | 43 |
| 3.2.2 Αλγόριθμος K-Means..... | 43 |
| 3.2.3 DBSCAN και HDBSCAN..... | 45 |
| 3.2.4 Meanshift | 45 |
| 3.2.5 Εκτίμηση βέλτιστου αριθμού ομάδων..... | 47 |
| 3.2.6 Δείκτες εγκυρότητας αλγορίθμων ομαδοποίησης | 49 |

| | |
|---|------------|
| 4. Μηχανική Μάθηση στον Κατασκευαστικό Τομέα | 52 |
| 4.1 Παρακολούθηση δομικής ακεραιότητας | 52 |
| 4.2 Τρόποι προσδιορισμού φθορών | 53 |
| 4.2.1 Ταξινόμηση..... | 53 |
| 4.2.2 Παλινδρόμηση με πλαίσιο οριοθέτησης | 54 |
| 4.2.3 Σηματολογική κατάτμηση..... | 56 |
| 4.3 Πρακτική εφαρμογή της έρευνας μηχανικής μάθησης | 59 |
| 4.4 Πεδία εφαρμογής της έρευνας μηχανικής μάθησης | 60 |
| 4.4.1 Γέφυρες | 60 |
| 4.4.2 Πεζοδρόμια | 61 |
| 4.4.3 Υπόγειες Κατασκευές | 62 |
| 4.4.4 Μνημεία πολιτιστικής κληρονομιάς | 62 |
| 4.4.5 Σύγχρονα κτίρια..... | 63 |
| 4.4.6 Άλλες κατασκευές μεγάλης κλίμακας | 64 |
| 4.5 Δυσκολίες στην εφαρμογή συνελκτικών νευρωνικών δικτύων στην δομική ακεραιότητα..... | 64 |
| 5. Μέθοδος Ανάλυσης | 66 |
| 5.1 Γενικά..... | 66 |
| 5.2 Δεδομένα..... | 66 |
| 5.3 Προεπεξεργασία δεδομένων..... | 70 |
| 6. Εφαρμογή Αλγορίθμων Μηχανικής Μάθησης | 73 |
| 6.1 Επιβλεπόμενη Μάθηση με CNN..... | 73 |
| 6.1.1 Λογισμικό Υλοποίησης | 73 |
| 6.1.2 Προδιαγραφές Ηλεκτρονικού Υπολογιστή..... | 73 |
| 6.1.3 Διαδικασία εκπαίδευσης..... | 74 |
| 6.2 Μη Επιβλεπόμενη Μάθηση με Ομαδοποίηση..... | 82 |
| 6.2.1 Λογισμικό υλοποίησης | 82 |
| 6.2.2 Διαδικασία ομαδοποίησης..... | 82 |
| 6.3 Αποτελέσματα αλγορίθμων | 87 |
| 6.3.1 Επισημασμένες εικόνες..... | 87 |
| 6.3.2 Μη επισημασμένες εικόνες | 101 |
| 6.4 Συγκριτική ανάλυση αποτελεσμάτων | 110 |
| 7. Συμπεράσματα | 115 |
| 7.1 Γενικά..... | 115 |
| 7.2 Δυσκολίες κατά την εκπόνηση εργασίας | 116 |
| 7.3 Κατευθύνσεις μελλοντικής έρευνας | 117 |

| | |
|--------------------------|------------|
| Βιβλιογραφία..... | 119 |
| ΠΑΡΑΡΤΗΜΑ..... | 125 |

Ευρετήριο εικόνων

| | |
|---|-----------|
| Εικόνα 1. Λειτουργία αλγορίθμου boosting | 10 |
| Εικόνα 2. Σχηματική αναπαράσταση των δεικτών precision / recall | 15 |
| Εικόνα 3. Η περιοχή κάτω από την καμπύλη ROC | 17 |
| Εικόνα 4. Η ηθική μηχανή | 21 |
| Εικόνα 5. Τυπική Λογιστική (Σιγμοειδής) Καμπύλη | 28 |
| Εικόνα 6. Multilayer perceptron | 30 |
| Εικόνα 7. Διαδοχικές συνελίξεις φίλτρου πάνω σε εικόνα | 34 |
| Εικόνα 8. Συνέλιξη με βήμα 2 και γέμισμα περιθωρίου 1 | 35 |
| Εικόνα 9. Παλινδρόμηση πλαισίου οριοθέτησης | 38 |
| Εικόνα 10. Επίπεδα συνελκτικού νευρωνικού δικτύου με αρχιτεκτονική VGG-16 | 39 |
| Εικόνα 11. Σχηματική αναπαράσταση συνελκτικού νευρωνικού δικτύου VGG-16..... | 40 |
| Εικόνα 12. Γραφική αναπαράσταση λειτουργίας αλγορίθμου k-means για k=3 | 44 |
| Εικόνα 13. Γραφική αναπαράσταση λειτουργίας αλγορίθμου meanshift | 47 |
| Εικόνα 14. Ταξινόμηση εικόνων σε ρωγμή ή όχι ρωγμή | 53 |
| Εικόνα 15. Ταξινόμηση παραθύρων της εικόνας σε ρωγμή ή όχι ρωγμή | 54 |
| Εικόνα 16. Παράδειγμα class activation mapping για κατηγορία ρωγμής | 55 |
| Εικόνα 17. Σημασιολογική κατάτμηση εικόνας ρωγμής | 56 |
| Εικόνα 18. Δομή πλήρους συνελκτικού νευρωνικού δικτύου | 57 |
| Εικόνα 19. Σημασιολογική κατάτμηση εικόνας μέσω αλγορίθμου ομαδοποίησης | 58 |
| Εικόνα 20. Συνολική απεικόνιση διαφορετικών τρόπων εντοπισμού φθορών | 59 |
| Εικόνα 21. Άποψη του Φρουρίου Αγίου Νικολάου | 67 |
| Εικόνα 22. Εικόνα λιθοδομής του φρουρίου, όπως προκύπτει από τα φασματικά κανάλια κόκκινου, πράσινου και μπλε | 68 |
| Εικόνα 23. Υπέρθωση επισημάνσεων πάνω στην φυσική εικόνα | 69 |
| Εικόνα 24. Κατάτμηση της εικόνας σε 32 παράθυρα για εκπαίδευση, και σύγκριση με τις επισημάνσεις | 71 |
| Εικόνα 25. Στοιχεία πίνακα σύγχυσης μοντέλου RGB στην ταξινόμηση επισημασμένων παραθύρων της εικόνας | 87 |
| Εικόνα 26. Δείκτες απόδοσης ταξινόμησης μοντέλου RGB σε επισημασμένα παράθυρα της εικόνας | 88 |
| Εικόνα 27. Στοιχεία πίνακα σύγχυσης μοντέλου NIR στην ταξινόμηση επισημασμένων παραθύρων της εικόνας | 89 |
| Εικόνα 28. Δείκτες απόδοσης ταξινόμησης μοντέλου NIR σε επισημασμένα παράθυρα της εικόνας | 90 |

| | |
|---|------------|
| Εικόνα 29. Στοιχεία πίνακα σύγχυσης ομαδοποίησης εικονοστοιχείων επισημασμένων εικόνων | 91 |
| Εικόνα 30. Δείκτες απόδοσης ταξινόμησης ομαδοποίησης με K-means σε επισημασμένα εικονοστοιχεία της εικόνας | 92 |
| Εικόνα 31. Δείκτης <i>Calinski – Harabasz</i> της ομαδοποίησης με <i>K-means</i> εικονοστοιχείων των επισημασμένων εικόνων | 93 |
| Εικόνα 32. Δείκτης <i>Davies-Bouldin</i> της ομαδοποίησης με <i>K-means</i> εικονοστοιχείων των επισημασμένων εικόνων | 93 |
| Εικόνα 33. Οπτική παράθεση μεθόδων εντοπισμού φθορών, <i>agiosnik2000</i> | 95 |
| Εικόνα 34. Οπτική παράθεση μεθόδων εντοπισμού φθορών, <i>agiosnik2048</i> | 96 |
| Εικόνα 35. Οπτική παράθεση μεθόδων εντοπισμού φθορών, <i>agiosnik002300</i> | 97 |
| Εικόνα 36. Οπτική παράθεση μεθόδων εντοπισμού φθορών, <i>agiosnik002420</i> | 98 |
| Εικόνα 37. Οπτική παράθεση μεθόδων εντοπισμού φθορών, <i>agiosnik002931</i> | 99 |
| Εικόνα 38. Οπτική παράθεση μεθόδων εντοπισμού φθορών, <i>agiosnik003538</i> | 100 |
| Εικόνα 39. Αποτελέσματα μοντέλου RGB στη ταξινόμηση παραθύρων άγνωστων εικόνων | 101 |
| Εικόνα 40. Αποτελέσματα μοντέλου NIR στη ταξινόμηση παραθύρων άγνωστων εικόνων | 102 |
| Εικόνα 41. Δείκτης <i>Calinski-Harabasz</i> της ομαδοποίησης με <i>K-means</i> εικονοστοιχείων μη επισημασμένων εικόνων | 103 |
| Εικόνα 42. Δείκτης <i>Davies-Bouldin</i> της ομαδοποίησης με <i>K-means</i> εικονοστοιχείων μη επισημασμένων εικόνων | 103 |
| Εικόνα 43. Οπτική παράθεση μεθόδων εντοπισμού φθορών, <i>agiosnik002211</i> | 105 |
| Εικόνα 44. Οπτική παράθεση μεθόδων εντοπισμού φθορών, <i>agiosnik002948</i> | 106 |
| Εικόνα 45. Οπτική παράθεση μεθόδων εντοπισμού φθορών, <i>agiosnik003218</i> | 107 |
| Εικόνα 46. Οπτική παράθεση μεθόδων εντοπισμού φθορών, <i>agiosnik003340</i> | 108 |
| Εικόνα 47. Οπτική παράθεση μεθόδων εντοπισμού φθορών, <i>agiosnik004719</i> | 109 |
| Εικόνα 48. Δείκτης <i>precision</i> για τις ταξινομήσεις μοντέλου RGB, μοντέλου NIR και ομαδοποίησης εικονοστοιχείων με <i>K-means</i> | 110 |
| Εικόνα 49. Δείκτης <i>recall</i> για τις ταξινομήσεις μοντέλου RGB, μοντέλου NIR και ομαδοποίησης εικονοστοιχείων με <i>K-means</i> | 111 |
| Εικόνα 50. Δείκτης <i>accuracy</i> για τις ταξινομήσεις μοντέλου RGB, μοντέλου NIR και ομαδοποίησης εικονοστοιχείων με <i>K-means</i> | 111 |
| Εικόνα 51. Δείκτης <i>F1</i> για τις ταξινομήσεις μοντέλου RGB, μοντέλου NIR και ομαδοποίησης εικονοστοιχείων με <i>K-means</i> | 112 |

Ευρετήριο πινάκων

| | |
|---|------------|
| Πίνακας 1. Ακρωνύμια..... | xii |
| Πίνακας 2. Hardware Υπολογιστή | 74 |
| Πίνακας 3. Αποτελέσματα μοντέλου RGB στην ταξινόμηση γνωστών παραθύρων των εικόνων | 138 |
| Πίνακας 4. Αποτελέσματα μοντέλου NIR στην ταξινόμηση γνωστών παραθύρων των εικόνων | 144 |
| Πίνακας 5. Εξωτερικοί δείκτες ομαδοποίησης εικονοστοιχείων γνωστών εικόνων | 150 |
| Πίνακας 6. Εσωτερικοί δείκτες ομαδοποίησης εικονοστοιχείων γνωστών εικόνων | 150 |
| Πίνακας 7. Αποτελέσματα μοντέλων ταξινόμησης σε άγνωστα δεδομένα | 151 |
| Πίνακας 8. Εσωτερικοί δείκτες ομαδοποίησης εικονοστοιχείων άγνωστων εικόνων..... | 152 |

| Ακρωνύμιο | Πλήρης Ονομασία |
|------------------|-------------------------------------|
| AI | Artificial Intelligence |
| AUC | Area Under ROC Curve |
| CAM | Class Activation Mapping |
| CHI | Calinski – Harabasz Index |
| CNN | Convolutional Neural Network |
| DBI | Davies – Bouldin Index |
| DSLR | Digital Single – Lens Reflex camera |
| FCN | Fully Convolutional Network |
| FFNN | Feed – Forward Neural Network |
| FN | False Negative |
| FP | False Positive |
| FPN | Feature Pyramid Network |
| FPR | False Positive Rate |
| kNN | k-Nearest Neighbor |
| LSTM | Long Short – Term Memory |
| ML | Machine Learning |
| MLP | Multi - Layer Perceptron |
| NIR | Near InfraRed |
| NN | Neural Network |
| R-CNN | Regional CNN |
| ReLU | Rectified Linear Unit |
| RGB | Red Green Blue |
| RNN | Recurring Neural Network |
| ROC | Receiver Operating Characteristic |
| RTK | Real Time Kinematic positioning |
| SHM | Structural Health Monitoring |
| TN | True Negative |
| TP | True Positive |
| TPR | True Positive Rate |
| VGG | Visual Geometry Group |
| YOLO | You Only Look Once |

Πίνακας 1: Ακρωνύμια

Εισαγωγή

Σε μία εποχή που οι επιπτώσεις της κλιματικής αλλαγής είναι ήδη έντονα εμφανείς, πέρα από τον σχεδιασμό νέων υποδομών, μεγάλη σημασία έχει η διατήρηση και η προστασία των ήδη υπαρχόντων κατασκευών, και ιδιαιτέρως των ιστορικών μνημείων. Το πρώτο βήμα στην εγκαθίδρυση της στρατηγικής για προστασία των μνημείων, είναι ο εντοπισμός υπαρχόντων φθορών καθώς και η διαρκής παρακολούθηση των μνημείων για πιθανή εμφάνιση νέων από αυτές.

Εκτενής είναι η έρευνα στο πεδίο της Παρακολούθησης δομικής ακεραιότητας, σχετικά με την εφαρμογή αυτόματων τρόπων εντοπισμού αστοχιών. Η τεχνολογική πρόοδος σε αλγορίθμους όρασης υπολογιστών καθώς και οι αυξημένες δυνατότητες των ηλεκτρονικών εξαρτημάτων, είναι δυνατό να δημιουργήσουν συστήματα αυτοματοποιημένου ελέγχου, χωρίς την απαίτηση διαρκούς ανθρώπινης εργασίας.

Λαμβάνοντας υπόψιν τα παραπάνω, η παρούσα διπλωματική εργασία επιχειρεί να αξιοποιήσει τα γνωστικά πεδία της τηλεπισκόπησης, της φωτογραμμετρίας, του προγραμματισμού και της θεωρίας στατιστικής που διδάσκονται στην Σ.Α.Τ.Μ, ώστε να υπάρξει μία επιφανειακή μεν, διεξοδική δε εισαγωγή στο πεδίο της μηχανικής μάθησης για την προστασία των ιστορικών μνημείων. Ο κύριος στόχος της εργασίας είναι η μελέτη της χρήσης διαφορετικών αλγορίθμων μηχανικής μάθησης για τον εντοπισμό φθορών σε μνημεία, καθώς και η σύγκριση των αποτελεσμάτων όταν αυτοί εφαρμοστούν πάνω στο ίδιο αντικείμενο μελέτης. Η σύγκριση γίνεται τόσο με την παράσταση δεικτών για κάθε αλγόριθμο, όσο και με τεχνικές οπτικής σύγκρισης με την πραγματική κατάσταση.

Στο πρώτο κεφάλαιο, αναλύονται οι υπάρχουσες διατάξεις προστασίας των μνημείων και πώς ο εντοπισμός φθορών σε αυτά μπορεί να συμβάλει σε αποτελεσματικότερο σχεδιασμό στον χώρο. Στο δεύτερο κεφάλαιο της εργασίας, παρατίθεται το συνολικό πλαίσιο που διέπει την λειτουργία της μηχανικής μάθησης στον κόσμο σήμερα. Στο τρίτο κεφάλαιο επεξηγούνται οι βασικοί αλγόριθμοι μηχανικής μάθησης, οι οποίοι αποτελούν τα θεμέλια πάνω στα οποία στηρίζονται τα σύγχρονα λογισμικά και τα πολύπλοκα συστήματα που τα αξιοποιούν. Το τέταρτο κεφάλαιο περιγράφει την επίδραση της μηχανικής μάθησης σε ζητήματα σχετικά με την δομική ακεραιότητα, πεδίο στο οποίο υπάγονται και τα μνημεία και η παρακολούθησή τους.

Το πρακτικό τμήμα της εργασίας ξεκινάει στο πέμπτο κεφάλαιο, με την περαιτέρω ανάλυση του στόχου της εργασίας και του τρόπου επίτευξής του. Ακόμα, επεξηγούνται τα αρχικά δεδομένα που αποκτήθηκαν για την εργασία και πως αυτά επεξεργάστηκαν σε πρώτο βαθμό ώστε να είναι κατάλληλα για αξιοποίηση από τους αλγόριθμους. Στο έκτο κεφάλαιο αναλύεται η εφαρμογή των αλγορίθμων μηχανικής μάθησης στα δεδομένα και οι επιλογές που σχηματίστηκαν κατά την εκκίνηση τους, και παρατίθενται τα αποτελέσματά τους σε διαγράμματα για εποπτεία και σύγκριση των αποτελεσμάτων τους στην αναγνώριση φθορών. Τέλος, στο έβδομο κεφάλαιο αναφέρονται τα συμπεράσματα της ανωτέρω σύγκρισης και προτείνονται ιδέες για περαιτέρω έρευνα παρεμφερή με τον τρόπο που διενεργήθηκε η παρούσα εργασία.

1. Επίδραση των μνημείων στον περιφερειακό σχεδιασμό

1.1 Ανάγκη προστασίας των μνημείων

1.1.1 Λογική και συναισθηματική αξία

Τα ιστορικά μνημεία, ανεξαρτήτου μεγέθους και εμφάνισης, είναι κτίσματα που έχουν ανεγερθεί για να θυμίζουν σημαντικά γεγονότα και φυσιογνωμίες στα άτομα που τα βλέπουν. Καθώς ανά τα χρόνια οι πολιτικά επιφανείς έχουν τόσο την εξουσία, όσο και τους πόρους για την εγκαθίδρυση μνημείων, τα χρησιμοποιούν για να αντιπροσωπεύσουν τις δικές τους κυρίαρχες πεποιθήσεις και την κοσμοθεωρία τους. Ως αποτέλεσμα, τα μνημεία μπορεί να αντιπροσωπεύουν επιλεκτικά τα ιστορικά διηγήματα, εστιάζοντας μόνο σε γεγονότα και προσωπικότητες που είναι βολικά για εκείνους που τα ανοικοδόμησαν. Ωστόσο, δεν μπορούν να ελέγξουν πλήρως τον τρόπο που το κάθε άτομο ξεχωριστά και οι κοινωνικές ομάδες τα ερμηνεύουν. Αφού έχουν ανεγερθεί, τα μνημεία μπορούν να χρησιμοποιηθούν, να αλλάξουν και να ερμηνευτούν με τρόπους διαφορετικούς ή και αντίθετους στους σκοπούς εκείνων που τα εγκαθίδρυσαν.

Η σημασία τους εντοπίζεται σε δύο επίπεδα. Τα ιστορικά μνημεία έχουν αξία τόσο λογική, όσο και συναισθηματική. Η λογική τους αξία έγκειται στην γνώση των ανθρώπων που τα κατασκεύασαν, καθώς και στη γνώση που έχουν οι σημερινοί θεατές τους πάνω σε αυτά που αντιπροσωπεύουν. Μπορούν να θεωρηθούν ως μέσο συλλογικής μνήμης, με το οποίο η γνώση όσων τα έχτισαν μεταβιβάζεται από γενιά σε γενιά, όσο τα μνημεία έχουν υπόσταση στον χώρο. Αν και η γνώση που ενσαρκώνουν τα μνημεία περιέχει προκαταλήψεις σε κάποιο βαθμό, η σκόπιμη καταστροφή μνημείων αποτελεί τρόπο ελέγχου του τι πρέπει να θυμούνται οι πολίτες και τι πρέπει να ξεχάσουν από το παρελθόν τους. Είναι τρόπος των εξουσιαζόντων μία δεδομένη στιγμή να επιβάλλουν ερμηνείες της ιστορίας και να ενθαρρύνουν μελλοντικές κατευθύνσεις της, αλλοιώνοντας αυτή τη συλλογική μνήμη. Ακόμα και τυχαία καταστροφή όμως, μπορεί να μειώσει την μεταβιβαζόμενη γνώση των μνημείων. Άρα η προστασία των μνημείων έχει καθοριστική σημασία στην διατήρηση μίας όσο το δυνατόν ουδέτερης και αντικειμενικής πορείας στην ιστορία, χωρίς περαιτέρω προκαταλήψεις και κενά.

Η συναισθηματική τους αξία έγκειται στο πώς επιδρά η γνώση για το μνημείο στην ψυχολογία κάποιου θεατή, ανάλογα με την θετική άποψη ή την αρνητική άποψη που έχει ως προς αυτό που πρεσβεύει εκείνο. Η συναισθηματική αξιολόγηση του μνημείου προκαλεί σε αυτόν αισθήματα ευφορίας, ή δυσφορίας, ανάλογα με το αν ανακαλεί ευχάριστες ιστορικές συγκυρίες ή ανεπιθύμητες έως τραυματικές εμπειρίες. Ακόμα, το ίδιο μνημείο μπορεί να προκαλεί σε άλλες κοινωνικές ομάδες θετικά συναισθήματα και σε άλλες αρνητικά. Οι συναισθηματικές τάσεις επηρεάζουν και την συμπεριφορά των ατόμων στον χώρο του μνημείου. Ως προς την γενικότερη συναισθηματική αξία, μνημεία μπορούν να χαρακτηριστούν ως «θερμά», αν η παρουσία τους είναι αμφιλεγόμενη και δημιουργούν έντονες συναισθηματικές αντιδράσεις σε κοινωνικές ομάδες που διαφωνούν με την ύπαρξη

τους. Αντιθέτως, «ψυχρά» χαρακτηρίζονται τα μνημεία που έχουν ευρεία αποδοχή από τους θεατές τους και ενσωματώνονται ειρηνικά στις καθημερινές τους συνήθειες ή θεωρούνται κανονικό στοιχείο του περιβάλλοντος τους.¹ Η πλειοψηφία των μνημείων στον Ελλαδικό χώρο μπορεί να χαρακτηριστεί ως «ψυχρά μνημεία», αφού έχουν συνδεθεί πλήρως με την ιστορία της Ελλάδας και αποτελούν πηγή εθνικής υπερηφάνειας, χωρίς ωστόσο να λείπουν και «θερμά» μνημεία, κυρίως όταν αυτά αφορούν διαφορετικές θρησκείες.

Οι πολίτες συμπεριφέρονται γύρω από τα μνημεία σύμφωνα με αυτά που γνωρίζουν για εκείνα, με το αν εκτιμούν θετικά ή αρνητικά την γνώση αυτή καθώς και με το πώς νιώθουν με εκείνα στον χώρο τους. Συνδυάζοντας τις διαφορετικές εκφάνσεις των αξιών των μνημείων, προκύπτει η πραγματική τους αξία στην Ελλάδα. Ως τουριστική χώρα, η Ελλάδα αξιοποιεί έντονα την ύπαρξη των μνημείων και προσελκύει μεγάλο όγκο τουριστών που ενδιαφέρονται για εκείνα. Η γνώση που έχουν για την αρχαία Ελλάδα κυρίως, σε συνδυασμό με τα αισθήματα θαυμασμού και δέους μπροστά σε αυτά μετατρέπονται σε πραγματική χρηματική αξία. Ταυτόχρονα, τα μνημεία κάθε ιστορικής εποχής αποτελούν πηγή υπερηφάνειας σε τοπικό επίπεδο. Τα μνημεία συνεισφέρουν στην δημιουργία ξεχωριστής ταυτότητας για τον τόπο στον οποίο βρίσκονται, και συνδέουν τους πολίτες με αυτόν και την ιστορία του. Με τα θετικά συναισθήματα που προκαλούν, ενισχύουν την κοινωνική συνοχή, προτάσσουν την προστασία του ευρύτερου περιβάλλοντος γύρω από αυτά και βελτιώνουν την ψυχοσύνθεση και την ποιότητα ζωής. Η προστασία τους κρίνεται απαραίτητη, λόγω των πολυάριθμων θετικών επιδράσεων και της συνολικής τους αξίας.

1.1.2 Νομική υποχρέωση

Από τη νομική διάσταση, μεγάλη έμφαση δίνεται στην πολιτική της βιώσιμης ή αειφόρου ανάπτυξης στην Ευρώπη. Αυτή στοχεύει στην βέλτιστη οικονομική ανάπτυξη την παρούσα στιγμή, χωρίς να στερούνται τους πόρους οι μελλοντικές γενιές. Η βιώσιμη ανάπτυξη βασίζεται στην ισόρροπη διαχείριση των τριών πυλώνων της, δηλαδή της οικονομίας, της κοινωνίας και του περιβάλλοντος. Σαν πρωταρχικός στόχος της Ευρωπαϊκής Ένωσης, καθορίζει την στρατηγική της και παρέχει το πλαίσιο πολιτικής για όλα τα συγχρηματοδοτούμενα προγράμματα, στα οποία συμμετέχουν μεταξύ άλλων και προγράμματα προστασίας και ανάδειξης των μνημείων. Επομένως, πρέπει να παραχωρηθεί το δικαίωμα στην επίσκεψη των μνημείων και στην ενασχόληση με αυτά στους επόμενους πολίτες της χώρας, όπως δίνεται στους σημερινούς.

Σε εθνικό επίπεδο, το σύνταγμα ως ανώτερος νόμος του κράτους αναφέρει ρητά στο Άρθρο 24 ότι «Η προστασία του φυσικού και πολιτιστικού περιβάλλοντος αποτελεί υποχρέωση του Κράτους και δικαίωμα του καθενός. Για τη διαφύλαξή του το Κράτος έχει υποχρέωση να παίρνει ιδιαίτερα προληπτικά ή κατασταλτικά μέτρα στο πλαίσιο της αρχής της αειφορίας». Ακόμα, στο ίδιο άρθρο, αναφέρει ότι «Τα μνημεία, οι παραδοσιακές περιοχές και τα παραδοσιακά στοιχεία προστατεύονται από το Κράτος». Σύμφωνα με το άρθρο αυτό οι υπόλοιπες διατάξεις του νόμου οφείλουν να προστατεύουν τα μνημεία, και σε καμία περίπτωση να μην αποτελούν άμεσο ή έμμεσο κίνδυνο για αυτά. Ακόμα, με τον νόμο

3028/2002, τα μνημεία καταγράφονται, τεκμηριώνονται και καταχωρούνται στο Εθνικό Αρχείο Μνημείων, που τηρείται στο Υπουργείο Πολιτισμού.

Ο περιφερειακός σχεδιασμός σε ανώτατο επίπεδο καθορίζεται από το Γενικό πλαίσιο χωροταξικού σχεδιασμού και αειφόρου ανάπτυξης. Εκείνο καθορίζει τις κατευθύνσεις για τα ειδικά και περιφερειακά πλαίσια χωροταξικού σχεδιασμού, καθώς και την ενσωμάτωση τους σε υποκείμενα επίπεδα πιο λεπτομερούς χωρικού σχεδιασμού. Για την διαχείριση πολιτιστικού πλούτου, η γενική κατεύθυνση είναι η ανάδειξη μνημείων πολιτιστικής κληρονομιάς, εξασφαλίζοντας ταυτόχρονα της αξιοποίηση τους ως πόλων έλξης επισκεπτών με τις κατάλληλες υποδομές πρόσβασης, τη λειτουργική σύνδεση τους με τον περιβάλλοντα χώρο και τη σύνδεση των μνημείων μεταξύ τους ανά χωρική ενότητα. Για την προστασία των αρχαιολογικών χώρων, προβλέπει την οριοθέτηση και θεσμοθέτηση ζωνών απόλυτης προστασίας (Α'), όπου περιορίζονται σημαντικά οι επιτρεπόμενες χρήσεις γης και η ανέγερση κτισμάτων, και ζωνών προστασίας της περιβάλλουσας περιοχής (Β'), όπου ισχύουν ειδικοί όροι και περιορισμοί χρήσης και δόμησης, για την προστασία του ευρύτερου τοπίου. Ακόμα διευρύνει την προστασία αρχαιολογικών χώρων με την ένταξη σε αυτούς της διασφάλισης των τοπίων. Τέλος, προβλέπει την υπογειοποίηση δικτύων διανομής ηλεκτρικής ενέργειας σε αρχαιολογικούς χώρους και την αποφυγή της διέλευσης δικτύων μεταφοράς μέσα από αυτούς.

1.2 Απειλές για τη δομική ακεραιότητα μνημείων

Οι ανθρώπινες κατασκευές φθείρονται στο βάθος του χρόνου, από φυσικές διεργασίες της γης. Το σύνολο των ατμοσφαιρικών συμβάντων, που επικρατούν σε έναν τόπο για μακροχρόνια περίοδο, ονομάζεται κλίμα και υποδεικνύει πόσο έντονα επιδρά η φύση στις κατασκευές. Η εναλλαγή μέρας και νύχτας, με τις διαφορές θερμοκρασίας που επιφέρει, οδηγεί σε θερμική διαστολή και συστολή υλικών, ειδικά πετρωμάτων, με αποτέλεσμα ρηγματώσεις και αποχρωματισμό τους. Επιπροσθέτως, η υγρασία της ατμόσφαιρας, μέσω υδρατμών, κατακρημνισμάτων και παγετού, διαλύει σταδιακά τα δομικά στοιχεία των μνημείων. Ο άνεμος μέσω φερτών υλικών όπως σκόνη, ή νερό με άλατα σε περιπτώσεις παραθαλάσσιων μνημείων, αποσαθρώνει την επιφάνειά τους. Τέλος, έμβιοι οργανισμοί, τόσο η χλωρίδα όσο και η πανίδα, μπορούν να προκαλέσουν σταδιακές φθορές σε μνημεία, αν δεν αποτραπούν.²

Οι φθορές αυτές δεν μπορούν να εξουδετερωθούν πλήρως, ακόμα και με προσεκτική διαχείριση. Επομένως, πάντα θα υπάρχει περιορισμένη φθορά, ειδικά σε μνημεία αιώνων. Ωστόσο η κύρια απειλή για αυτά την σημερινή εποχή, είναι η κλιματική αλλαγή, η οποία εντείνει τις διεργασίες που περιγράφονται. Παρουσιάζονται μεγαλύτερες διακυμάνσεις της θερμοκρασίας ανά μονάδα χρόνου, οι μέγιστες και ελάχιστες τιμές της γίνονται εντονότερες και αυξάνεται η σφοδρότητα των καιρικών φαινομένων. Ως αποτέλεσμα, είναι συχνότερη και μεγαλύτερη η φθορά των μνημείων. Η κλιματική αλλαγή, αν και φαινόμενο που εμφανίζεται περιοδικά στην ιστορία σε ηπιότερη μορφή, είναι πλέον αποτέλεσμα ανθρωπογενών παραγόντων.³ Ωστόσο, σαν φαινόμενο που οφείλεται σε δραστηριότητες σε

όλη την υφήλιο και επιδρά και σε πολλές περιοχές ταυτόχρονα, είναι απειλή υπερτοπικού χαρακτήρα, και άρα δεν μπορεί να περιοριστεί μόνο με καταλληλότερο σχεδιασμό σε εθνικό επίπεδο.

Σε πιο τοπική κλίμακα, τα μνημεία πολιτιστικής κληρονομιάς απειλούνται από γεωλογικούς κινδύνους. Η Ελλάδα είναι σεισμογενής περιοχή, και πλήττεται ανά τακτά χρονικά διαστήματα από σεισμούς, μικρής έως και αρκετά μεγάλης έντασης. Ακόμα, συχνά παρουσιάζονται φαινόμενα καταστροφικών κατολισθήσεων και πλημμυρών, που εντείνονται από την προαναφερθείσα κλιματική αλλαγή. Εκτός από τα άμεσα καταστροφικά συμβάντα, οι γεωλογικοί κίνδυνοι μπορεί να είναι αργοκίνητοι και μη καταστροφικοί, όπως τεκτονικά ρήγματα, υποχώρηση του εδάφους και αργές καθιζήσεις του εδάφους.⁴ Για τους φυσικούς κινδύνους, όπως είναι οι σεισμοί, είναι δύσκολο να παρθούν προληπτικά μέτρα, ωστόσο πλημμυρικά φαινόμενα και καθιζήσεις λόγω βιομηχανικών δραστηριοτήτων, όπως είναι οι μεταλλευτικές δραστηριότητες και η λειτουργία λατομείων, μπορούν να αντιμετωπιστούν με στρατηγικές χωρικού σχεδιασμού.

Άλλη πηγή φθορών, ιδιαίτερα αισθητή σε μεγάλα αστικά κέντρα ή σε κοντινή απόσταση από αυτά, είναι οι ατμοσφαιρικοί ρύποι. Πρόκειται για χημικά στοιχεία ή ενώσεις, οι οποίες βρίσκονται σε αέρια μορφή ή σε μορφή στερεών σωματιδίων, και προκαλούν καταστροφές σε διάφορα υλικά είτε με την απευθείας επαφή τους, είτε συχνότερα με την βοήθεια κατακρημνισμάτων, όπου παρουσιάζεται το φαινόμενο της όξινης βροχής. Η όξινη βροχή αλλοιώνει τα μνημεία μέσω χημικών διεργασιών, όπως η διάλυση πετρωμάτων και η διάβρωση μετάλλων. Τα επιβλαβή χημικά στοιχεία που την προκαλούν μπορεί να βρεθούν στην ατμόσφαιρα από φυσικά αίτια σπανιότερα, όπως εκρήξεις ηφαιστειών, αλλά συχνότερα πρόκειται είτε για βιομηχανικά απόβλητα, είτε αποτέλεσμα διαφόρων μέσων μεταφοράς.⁵ Λαμβάνοντας υπόψη την πηγή του κινδύνου αυτού, η χωροθέτηση τόσο της βιομηχανικής δραστηριότητας όσο και των δικτύων μεταφορών οφείλει να περιορίζεται σε λογικά πλαίσια σε περιοχές με παρουσία πολιτιστικών μνημείων και αρχαιολογικών χώρων. Το Γενικό πλαίσιο χωροταξικού σχεδιασμού προβλέπει ορθώς την αποφυγή διέλευσης μεταφορικών δικτύων από τις περιοχές αυτές, όπως αναφέρεται σε προηγούμενο εδάφιο, ωστόσο το Ειδικό πλαίσιο χωροταξικού σχεδιασμού για τη βιομηχανία αποφεύγει πλήρως την αναφορά σε αρχαιολογικούς χώρους, αφήνοντας την εγκατάσταση βιομηχανιών σε αυτούς σε έγκριση του εν ενεργεία υπουργού πολιτισμού (ν. 3028/2020), η οποία όμως ελέγχεται ως προς την αιτιολόγηση της για αποφυγή άμεσης ή έμμεσης βλάβης στο μνημείο.

Το περιβάλλον τοπίο των μνημείων, αν και δεν σχετίζεται άμεσα με τις φθορές σε εκείνο, κρίνεται ότι χρήζει προστασίας. Το ειδικό χωροταξικό πλαίσιο για τις ανανεώσιμες πηγές ενέργειας απαγορεύει την αξιοποίηση των ζωνών προστασίας Α' γύρω από μνημεία με εγκαταστάσεις αιολικής, υδροηλεκτρικής, ηλιακής καθώς και άλλων μορφών ενέργειας. Τέλος, απειλή για τα μνημεία είναι τα ίδια τα άτομα, είτε λόγω κακοπροαίρετων ενεργειών, όπως πολεμικές συρράξεις ή βανδαλισμός, είτε λόγω αμέλειας. Σε ειδικές περιπτώσεις για την καλύτερη μελέτη ή προσβασιμότητα σε μνημεία, μπορούν να χρησιμοποιηθούν τακτικές οι οποίες οδηγούν σε περιορισμένη φθορά τους.

1.3 Ο εντοπισμός φθορών ως πληροφορία για τον χωρικό σχεδιασμό

Σε κάθε διαδικασία επίλυσης ενός προβλήματος, η πληροφορία διακατέχει κρίσιμο ρόλο. Τα δεδομένα, με προσεκτική ανάλυση και επεξεργασία τους, είναι αυτά που οδηγούν στα τελικά πορίσματα και τις επιλογές που επιλύουν το πρόβλημα. Ο σχεδιασμός μπορεί να αντιμετωπιστεί ως μια διαδικασία επίλυσης προβλημάτων, και άρα η πρόσβαση σε πληροφορίες έχει καθοριστική σημασία.

Η διαδικασία του σχεδιασμού περιλαμβάνει την εκτίμηση της υπάρχουσας κατάστασης, τη μελέτη της διαχρονικής εξέλιξης των μεταβλητών που οδήγησαν στην υπάρχουσα κατάσταση, τη θεώρηση των δυνατών εναλλακτικών λύσεων ή σεναρίων και την επιλογή της καταλληλότερης λύσης με βάση υποθέσεις για το μέλλον, την υλοποίηση της λύσης που επιλέχθηκε, τη χάραξη στρατηγικής για το μέλλον και, τέλος, την περιοδική ή συνεχή επανεκτίμηση της κατάστασης μετά την υλοποίηση, που μπορεί να οδηγήσει και σε επανεκτίμηση των προηγούμενων σταδίων.

Η ελλιπής γνώση του περιβάλλοντος που μελετάται, εισάγει αβεβαιότητες τόσο ως προς την αξιολόγηση της υπάρχουσας κατάστασης του όσο και στην αντίδραση του στις παρεμβάσεις σχεδιασμού. Επομένως, οι επαρκείς πληροφορίες λειτουργούν ως μέρος ενός μηχανισμού μείωσης της αβεβαιότητας και βοηθούν στην περιγραφή, την ανάλυση και την λήψη αποφάσεων στα διάφορα στάδια που αναφέρονται.

Ο μεγάλος όγκος δεδομένων που απαιτούνται σε κάθε πρόβλημα σχεδιασμού, δυσχεραίνει την διαχείρισή τους. Ακόμα, σημαντικό πρόβλημα στο πλαίσιο του σχεδιασμού είναι η αδυναμία πρόσβασης των σχεδιαστών στην πληροφορία, ακόμα και αν αυτή έχει ήδη συγκεντρωθεί σε διάφορους φορείς. Επιπλέον, σημαντική είναι η ανάγκη για πληροφορίες συγκρίσιμες χωρικά και χρονικά, τη στιγμή που οι διάφοροι φορείς δεν διαθέτουν μία ολοκληρωμένη βάση πληροφοριών με δεδομένα σε τυποποιημένη μορφή. Για την επίλυση των προβλημάτων αυτών, καταλληλότερη δράση είναι η υλοποίηση ενός ολοκληρωμένου συστήματος πληροφοριών.

Για να είναι χρήσιμη η πληροφορία στον σχεδιασμό, η οργάνωσή της πρέπει να ακολουθεί την οργάνωση της διαδικασίας σχεδιασμού, η οποία με τη σειρά της έχει το χρέος να ακολουθεί την οργάνωση του συστήματος του πραγματικού κόσμου. Σε περιπτώσεις όπου η οργάνωση της πληροφορίας είναι διαχωρισμένη από άλλα στοιχεία της διαδικασίας σχεδιασμού, το σύστημα πληροφοριών που χτίζεται εξυπηρετεί περισσότερο την συλλογή και την οργάνωση της πληροφορίας ως αυτοσκοπό, παρά την χρήση και αξιοποίηση της στον σχεδιασμό. Η ανάκτηση των πληροφοριών σε ένα τέτοιο σύστημα σπάνια συνεισφέρει στη διατύπωση πολιτικής ή στον έλεγχο επιπτώσεων που προκαλούνται από παρεμβάσεις του σχεδιασμού.⁶

Όσον αφορά την δομική ακεραιότητα κατασκευών, έμφαση δίνεται στην κατασκευή βάσεων δεδομένων με εικόνες φθορών σε διάφορα δομικά υλικά. Οι βάσεις αυτές λειτουργούν ως δεδομένα εκπαίδευσης για αλγόριθμους μηχανικής μάθησης, οι οποίοι στη συνέχεια θα

αποτελέσουν τμήματα ενός συστήματος αυτόματου εντοπισμού φθορών σε κάθε υλικό, μεταξύ των οποίων και οι επιφάνειες αρχαίων και ιστορικών μνημείων. Εφαρμόζοντας τις παραπάνω αρχές για τις πληροφορίες σε διαδικασίες σχεδιασμού, είναι εμφανές ότι η δημιουργία βάσεων δεδομένων φθορών, με σκοπό την δημιουργία αυτόματων συστημάτων εντοπισμού φθορών, για περαιτέρω εμπλουτισμό των βάσεων αυτών είναι ένας φαύλος κύκλος, ο οποίος δεν συνεισφέρει ενεργά σε θέματα χωρικού σχεδιασμού.

Για να αποβεί χρήσιμη η πληροφορία της ύπαρξης φθορών σε μνημεία, αντί για την απλή τοπική καταγραφή τους, απαιτείται και η χρονική και ευρύτερη χωρική διάσταση τους, σε συνδυασμό με την χωρική και χρονική καταγραφή άλλων δραστηριοτήτων που εμφανίζονται στο ευρύτερο περιβάλλον των μνημείων. Η διασύνδεση των πληροφοριών αυτών, είναι εκείνη που θα οδηγήσει σε καλύτερη κατανόηση της υπάρχουσας κατάστασης, πού οφείλεται αυτή και με ποια βήματα μπορεί να βελτιωθεί στο μέλλον. Ο αυτόματος εντοπισμός φθορών βοηθά στην συνεχή επανεκτίμηση της κατάστασης μετά την υλοποίηση κάποιας στρατηγικής ή λύσης σε πρόβλημα χωρικού σχεδιασμού, καθώς και στην ακριβέστερη χρονική πληροφορία, σε αντίθεση με τον περιοδικό χειροκίνητο έλεγχο μνημείων για την ύπαρξη φθορών.

Οι καλύτερες πληροφορίες μπορούν να επιδράσουν ευεργετικά σε δύο τομείς της υπάρχουσας στρατηγικής σχεδιασμού. Η οριοθέτηση ζώνων απόλυτης προστασίας Α' και προστασίας περιβάλλουσας περιοχής Β', που καθορίζουν σε μεγάλο βαθμό ποιος χώρος γύρω από μνημεία δεν μπορεί να αξιοποιηθεί, εξαρτάται από το κάθε μνημείο ξεχωριστά, ανάλογα με το τοπίο του. Ωστόσο δεν λαμβάνεται υπόψιν η επίδραση των διαφορετικών χρήσεων σε κάθε μνημείο, αλλά αντιμετωπίζονται όλες συνολικά. Η καλύτερη πληροφορία σχετικά με την επίδραση ή μη κάθε χρήσης σε κάθε μνημείο μπορεί να οδηγήσει σε καλύτερα και ακριβέστερα κριτήρια, καθώς και πιο εξειδικευμένες ζώνες αποκλεισμού κάθε χρήσης, επιτρέποντας αποδοτικότερη αξιοποίηση του χώρου και περιορισμό παραγόντων που απειλούν τα μνημεία.

Ακόμα, οι εγκρίσεις του Υπουργού πολιτισμού για βιομηχανίες σε κοντινές αποστάσεις από μνημεία, μπορούν να τεκμηριωθούν επαρκέστερα με χρονικές και χωρικές πληροφορίες ενός τέτοιου συστήματος παρακολούθησης φθορών από άλλες παρόμοιες περιπτώσεις, ενώ διευκολύνεται και ο έλεγχος επάρκειας των αιτιολογήσεων από το Συμβούλιο της Επικρατείας, σε περίπτωση κίνδυνου για αρχαία μνημεία. Στη συνέχεια, με τη συνεχή παρακολούθηση, βελτιώνεται ο έλεγχος μετά την εγκατάσταση των βιομηχανιών, για πιθανή πρόκληση φθορών από αυτές.

Στα επόμενα κεφάλαια εξετάζεται η εφαρμογή μεθόδων μηχανικής μάθησης, για τον εντοπισμό φθορών σε μνημεία. Ο εντοπισμός αυτός αποτελεί το πρώτο στάδιο στην υλοποίηση ενός αυτόματου συστήματος εντοπισμού, το οποίο μπορεί να λειτουργήσει ως εργαλείο για αποτελεσματικότερο σχεδιασμό, σε θέματα που αφορούν την προστασία των μνημείων στα πλαίσια της βιώσιμης ανάπτυξης.

2. Γενικά Στοιχεία Μηχανικής Μάθησης

2.1 Η ανάπτυξη της μηχανικής μάθησης

Η τεχνητή νοημοσύνη (artificial intelligence, AI) ορίζεται ως η ικανότητα ενός υπολογιστή να εκτελέσει διεργασίες, οι οποίες συνήθως πραγματοποιούνται από τον άνθρωπο λόγω των νοητικών και διακριτικών ικανοτήτων του.⁷ Η μηχανική μάθηση (machine learning), ως υποσύνολο της τεχνητής νοημοσύνης, αναπτύχθηκε μαζί με αυτήν από τα μέσα του 20ού αιώνα έως και σήμερα.

Η μηχανική μάθηση στηρίζεται σε ένα μοντέλο αλληλεπίδρασης εγκεφαλικών κυττάρων, όπως αυτό περιγράφηκε από τον Donald Hebb στο βιβλίο «The organization of behavior. 1949», σύμφωνα με το οποίο, όταν ένα εγκεφαλικό κύτταρο διεγείρει ένα δεύτερο συνεχώς, σχηματίζει ή ενδυναμώνει μία νευρική σύναψη.⁹ Μεταφράζοντας το γεγονός αυτό σε πιθανούς τεχνητούς νευρώνες, αν δύο από αυτούς ενεργοποιούνται ταυτόχρονα ενδυναμώνεται η διασύνδεση μεταξύ τους, έχει δηλαδή μεγαλύτερο βάρος, ενώ αν ενεργοποιούνται ξεχωριστά η διασύνδεση έχει μικρότερο βάρος.⁸ Η προσέγγιση αυτή αποτέλεσε την βάση της μηχανικής μάθησης και χρησιμοποιήθηκε εκτενώς, ιδιαιτέρως στην κατηγορία των νευρωνικών δικτύων, αφού θεωρήθηκε ότι η μίμηση της λειτουργίας του ανθρώπινου εγκεφάλου από υπολογιστές θα οδηγήσει στο επιθυμητό αποτέλεσμα της τεχνητής νοημοσύνης. Ωστόσο, η παρομοίωση νευρωνικών δικτύων με τον ανθρώπινο εγκέφαλο αποδεικνύεται ανακριβής. Στα νευρωνικά δίκτυα η είσοδος της πληροφορίας εκτελείται τμηματικά από επίπεδο σε επίπεδο, ενώ τα λάθη σε αποφάσεις τους διαδίδονται προς τα πίσω από επίπεδο σε επίπεδο με μη βιολογικά εύλογο τρόπο. Η σύγχρονη επεξεργασία πληροφοριών σε κάθε επίπεδο διαφέρει από την λειτουργία του εγκεφάλου, ο οποίος κωδικοποιεί πληροφορίες σε αραιά και ασύγχρονα σήματα.¹⁰

Το 1950, ο Arthur Samuel, απόφοιτος του M.I.T. και εργαζόμενος στην IBM, ανέπτυξε ένα πρόγραμμα που έπαιζε το κλασικό επιτραπέζιο παιχνίδι της ντάμας. Το πρόγραμμα περιείχε μία συνάρτηση βαθμολογίας, η οποία χρησιμοποιούσε τις θέσεις των πιονιών στο τραπέζι για να προσπαθήσει να υπολογίσει της πιθανότητες κάθε πλευράς να κερδίσει. Στη συνέχεια επέλεγε την επόμενη του κίνηση, προσπαθώντας να ελαχιστοποιήσει την πιθανότητα ήττας του. Ο Samuel σχεδίασε και μηχανισμούς που επέτρεπαν στο πρόγραμμα του να βελτιώνεται, καταγράφοντας τις θέσεις που είχε δει σε συνδυασμό με τις τιμές της συνάρτησης βαθμολογίας. Εκείνος έκανε γνωστή την έννοια της μηχανικής μάθησης, σε δημοσίευση του στο «IBM Journal of Research and Development».¹¹

Το 1957, ο Frank Rosenblatt, εργαζόμενος στο εργαστήριο αεροναυπηγίας του Cornell, συνδύασε το μοντέλο αλληλεπίδρασης εγκεφαλικών κυττάρων του Hebb με τις προσπάθειες μηχανικής μάθησης που κατέβαλε ο Samuel και δημιούργησε το perceptron, το οποίο σχεδιάστηκε αρχικά ως μηχανήμα και όχι ως αλγόριθμος. Η πρώτη του υλοποίηση ως λογισμικό έγινε στον υπολογιστή IBM 704, στη συνέχεια υλοποιήθηκε σε έναν ειδικά προσαρμοσμένο υπολογιστή, με την ονομασία «Mark 1 Perceptron». Το μηχανήμα αυτό

σχεδιάστηκε για αναγνώριση εικόνων και διέθετε έναν πίνακα 400 φωτοκυττάρων τυχαία συνδεδεμένων με τεχνητούς νευρώνες.¹² Αν και ανακηρύχθηκε ως ο πρώτος επιτυχής νευρο-υπολογιστής, εν τέλει δεν εκπλήρωσε τις προσδοκίες επειδή δεν μπορούσε να εκπαιδευτεί και να αναγνωρίσει πολλές κατηγορίες μοτίβων. Αυτό οδήγησε την έρευνα πάνω στο πεδία των νευρωνικών δικτύων σε αδράνεια για αρκετά χρόνια.⁸

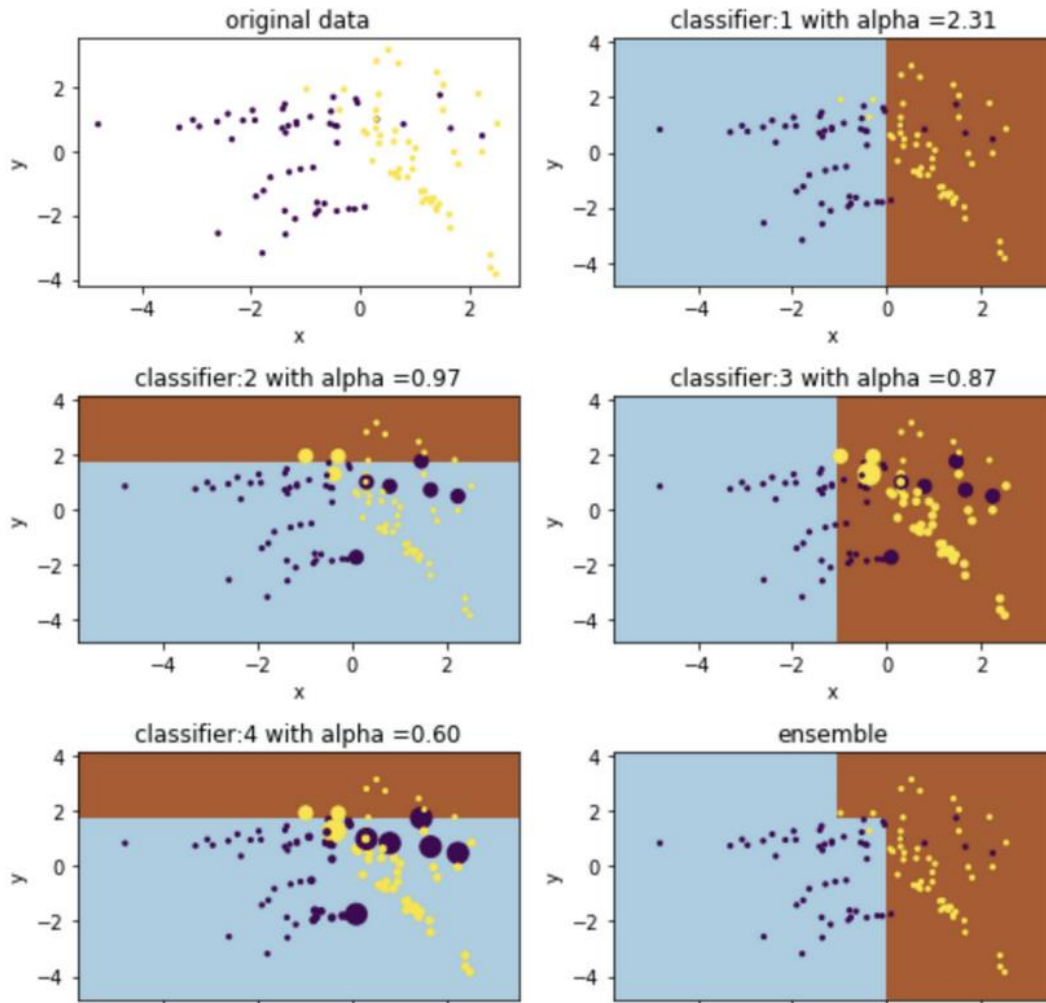
Το 1967 υλοποιήθηκε η ιδέα του αλγορίθμου εγγύτερου γείτονα, ο οποίος ήταν το ξεκίνημα της βασικής αναγνώρισης μοτίβων. Ο αλγόριθμος χρησιμοποιήθηκε για τη χαρτογράφηση διαδρομών και ήταν ένας από τους πρώτους αλγόριθμους που επιστρατεύτηκε για την λύση του προβλήματος της πιο οικονομικής διαδρομής ενός πλανόδιου πωλητή ανάμεσα σε διάφορες πόλεις. Ο Marcello Pelillo έχει αναγνωριστεί ως ο δημιουργός του κανόνα του εγγύτερου γείτονα, ωστόσο ο ίδιος δίνει τα εύσημα στη δημοσίευση «Cover T M & Hart P E. Nearest Neighbor pattern classification. IEEE Trans. Inform. Theory IT – 13:21 – 7, 1967.»

Τη δεκαετία του 1960, η ανακάλυψη και η χρήση πολλών επιπέδων άνοιξε νέο μονοπάτι στην έρευνα των νευρωνικών δικτύων. Διαπιστώθηκε ότι η διάθεση και η χρήση δύο ή περισσότερων στρωμάτων τεχνητών νευρώνων στο perceptron προσέφερε πολύ μεγαλύτερη υπολογιστική ισχύ από ένα perceptron με ένα μόνο στρώμα. Αυτό οδήγησε στην δημιουργία άλλων εκδοχών νευρωνικών δικτύων και σε ποικιλότητα η οποία διευρύνεται ακόμα και σήμερα. Η χρήση πολλών επιπέδων οδήγησε στα ανατροφοδοτούμενα νευρωνικά δίκτυα (feedforward neural networks) και την ανάστροφη αναδιάδοση (backpropagation).

Η ανάστροφη αναδιάδοση, η οποία αναπτύχθηκε την δεκαετία του 1970, επιτρέπει σε ένα νευρωνικό δίκτυο να προσαρμόσει τα κρυφά στρώματα νευρώνων του, μεταδίδοντας τα σφάλματα από το αποτέλεσμα του νευρωνικού δικτύου σε αυτά. Η ανάστροφη αναδιάδοση χρησιμοποιείται σήμερα για την εκπαίδευση βαθέων νευρωνικών δικτύων (deep neural networks).

Στα τέλη της δεκαετίας του 1970 και στις αρχές της δεκαετίας του 1980 η έρευνα μηχανικής μάθησης εγκαταλείφθηκε από ερευνητές επιστήμης των υπολογιστών και τεχνητής νοημοσύνης. Αυτό προκάλεσε σχίσμα μεταξύ της τεχνητής νοημοσύνης και της μηχανικής μάθησης, η οποία χρησιμοποιούνταν μέχρι τότε σαν πρόγραμμα εκπαίδευσης για την τεχνητή νοημοσύνη. Η βιομηχανία μηχανικής μάθησης άλλαξε προτεραιότητα από τις προσεγγίσεις που κληρονόμησε από την έρευνα τεχνητής νοημοσύνης και εστίασε σε μεθόδους και τακτικές από τα πεδία των πιθανοτήτων και της στατιστικής. Συνέχισε την ενασχόληση με τα νευρωνικά δίκτυα και άκμασε την δεκαετία του 1990, με την ανάπτυξη του ίντερνετ και την διαθεσιμότητα ψηφιακών δεδομένων.⁸

Το 1990 εισήχθη η έννοια του boosting, μία γενική μέθοδος βελτιστοποίησης της απόδοσης οποιουδήποτε αλγόριθμου μάθησης, από τον Robert Schapire σε δημοσίευση με τίτλο «The Strength of Weak Learnability¹³». Το boosting δουλεύει τρέχοντας συνεχόμενα έναν δοθέντα αδύναμο αλγόριθμο μάθησης πάνω σε διάφορες κατανομές των δεδομένων εκπαίδευσης (training data) και στη συνέχεια, συνδυάζοντας τους ταξινομητές (classifiers) που προκύπτουν από αυτόν σε έναν μοναδικό σύνθετο ταξινομητή (Εικόνα 1).¹⁴ Αποδείχτηκε αναγκαίο για την εξέλιξη της μηχανικής μάθησης.



Εικόνα 1: Σχηματική αναπαράσταση λειτουργίας αλγορίθμου *boosting**

Στις αρχές του 21ου αιώνα άρχισε η ταχύτατη ανάπτυξη και βελτιστοποίηση αλγορίθμων βαθειάς μάθησης για την αναγνώριση ομιλίας και την αναγνώριση εικόνων και προσώπων, η οποία συνεχίστηκε σε μεγαλύτερο βαθμό την επόμενη δεκαετία με σημαντική πρόοδο. Οι αποτελεσματικότεροι αλγόριθμοι εκείνης της εποχής άνηκαν στις εταιρείες Google και Facebook.

2.2 Αρχή λειτουργίας της μηχανικής μάθησης

Ο όρος μηχανική μάθηση μπορεί να θεωρηθεί παραπλανητικός, καθώς οι μηχανές δεν έχουν την δυνατότητα μάθησης με την έννοια που συναντάται σε έμβιους οργανισμούς. Αυτό που κάνει ένας αλγόριθμος μηχανικής μάθησης, είναι να βρίσκει ένα μαθηματικό μοντέλο, το

*(<https://towardsdatascience.com/boosting-algorithms-explained-d38f56ef3f30>)

οποίο αν εφαρμοστεί σε μία συλλογή δεδομένων εισόδου (input data), επιστρέφει τα επιθυμητά δεδομένα εξόδου (output data). Για να επιτευχθεί η εύρεση αυτού του μαθηματικού μοντέλου, χρειάζεται να προμηθευτεί ο αλγόριθμος μεγάλο αριθμό από τα δεδομένα αυτά, τόσο εισόδου όσο και τα αντίστοιχα δεδομένα εξόδου. Τα δεδομένα αυτά ονομάζονται δεδομένα εκπαίδευσης (training data).

Η μηχανική μάθηση διαχωρίζεται σε επιβλεπόμενη, μη επιβλεπόμενη και ενισχυτική μάθηση. Στην εισαγωγή αυτή επεξηγείται κυρίως η επιβλεπόμενη, η οποία αποτελεί την μεγαλύτερη κατηγορία. Η μη επιβλεπόμενη αναλύεται σε επόμενο κεφάλαιο, λόγω ιδιοτήτων της ως προς την διαδικασία μάθησης. Η ενισχυτική δεν αφορά την παρούσα εργασία, καθώς δεν υπάρχει επιβεβαιωμένος τρόπος ενσωμάτωσης της σε αντικείμενο σχετικό με τη σχολή.

Σε υψηλό επίπεδο, κατανοητό από τον άνθρωπο δηλαδή, αυτό που κάνει ο αλγόριθμος επιβλεπόμενης μηχανικής μάθησης είναι να διαβάζει τα δεδομένα εισόδου, να εφαρμόζει ένα τυχαίο ή υφιστάμενο μαθηματικό μοντέλο σε αυτά και να δημιουργεί εκροές, δηλαδή προβλέψεις, οι οποίες συγκρίνονται με τα δεδομένα εξόδου. Στη συνέχεια τροποποιεί επανειλημμένα το αρχικό εφαρμοσμένο μοντέλο, ανάλογα με την σύγκριση εκροών και δεδομένων εξόδου. Αν αυτά ταυτίζονται, τότε διατηρεί το μοντέλο σταθερό, αν υπάρχουν μεγάλες διαφορές αλλάζει παραμέτρους του μοντέλου τις οποίες θεωρεί υπεύθυνες για τις λανθασμένες εκροές. Ο τελικός στόχος του αλγορίθμου είναι να βρει εκείνο το μαθηματικό μοντέλο που προσαρμόζεται ακριβώς σε όλα τα δεδομένα εισόδου και επιστρέφει εν τέλει τα δεδομένα εξόδου που είχαν δοθεί. Αυτό στην πράξη είναι σχεδόν αδύνατο, κυρίως όταν τα δεδομένα είναι πολλά σε αριθμό και ανομοιογενή. Άρα ο τελικός στόχος είναι να βρεθεί ένα μοντέλο το οποίο προσαρμόζεται όσο το δυνατόν καλύτερα, έχει δηλαδή την μεγαλύτερη ακρίβεια και αστοχεί σε όσο λιγότερα δεδομένα εισόδου γίνεται. Όταν βρεθεί το μοντέλο αυτό, μπορεί να δημιουργήσει σωστές εκροές και για νέα δεδομένα εισόδου διαφορετικά από τα δεδομένα εκπαίδευσης, για τα οποία δεν υπάρχουν αντίστοιχα δεδομένα εξόδου, με την προϋπόθεση να ανήκουν σε ίδιο ή παρόμοιο πληθυσμό με αυτά.¹⁵ Αναμένεται ωστόσο το μοντέλο να έχει χαμηλότερη ακρίβεια στα νέα δεδομένα από αυτήν που έχει στα δεδομένα εκπαίδευσης.

Σε προγραμματιστικό επίπεδο, για να υπάρξει αλγόριθμος μηχανικής μάθησης, χρειάζεται συγκεκριμένα βήματα. Πρώτα, τα δεδομένα εκπαίδευσης πρέπει να υποστούν κατάλληλη επεξεργασία, ώστε να είναι ομοιογενή και να είναι σε μορφή που μπορούν να αξιοποιηθούν από κάποιο μοντέλο. Απαιτείται δηλαδή η δημιουργία ενός συνόλου δεδομένων (dataset). Αυτό περιλαμβάνει την επισήμανση των δεδομένων με τις επιθυμητές τιμές εξόδου (annotation), την μετατροπή ποιοτικών χαρακτηριστικών σε αριθμητική ή διανυσματική μορφή, την εξασφάλιση ότι όλα τα διανύσματα έχουν την ίδια διάσταση και την τροποποίηση του διαστήματος αριθμητικών τιμών ώστε να αντιπροσωπεύεται από τιμές ανάμεσα σε 0 και 1 (normalization) ή να ακολουθεί την τυπική κατανομή (standardization). Στη συνέχεια χρειάζεται να οριστεί ένα βασικό μαθηματικό μοντέλο ή διεργασία ώστε τα δεδομένα εισόδου να χρησιμοποιηθούν από αυτό για να προκύψουν οι εκροές. Το μοντέλο αυτό μπορεί να είναι παραμετρικό, να περιέχει δηλαδή παραμέτρους που μεταβάλλονται, ή μη παραμετρικό. Χρειάζεται επίσης να οριστεί μία συνάρτηση απωλειών, ένας τρόπος δηλαδή να υπάρχει εποπτεία των λάθους εκροών, όπως προκύπτουν σε σύγκριση με τα δεδομένα εξόδου. Εφόσον υπάρχει η συνάρτηση απωλειών, επόμενο βήμα είναι η εύρεση ενός κριτηρίου βελτιστοποίησης το οποίο βασίζεται πάνω στην συνάρτηση απωλειών, όπως για παράδειγμα η ελαχιστοποίηση του μέσου σφάλματος ή η μεγιστοποίηση της πιθανότητας η εκροή να είναι

αληθής. Τέλος, χρειάζεται μία ρουτίνα βελτιστοποίησης, μία διαδικασία η οποία χρησιμοποιεί τα δεδομένα εκπαίδευσης ώστε να βρει την βέλτιστη λύση του κριτηρίου και να μεταβληθούν ανάλογα οι παράμετροι ή όλο το μοντέλο.

Όταν έχουν καθοριστεί οι παραπάνω διεργασίες, απομένει να ρυθμιστούν οι λεγόμενες υπερπαράμετροι (hyperparameters). Οι παράμετροι αυτές εισάγονται από τον χρήστη, δεν υπολογίζονται από τον αλγόριθμο και αφορούν άμεσα την λειτουργία του. Παραδείγματα υπερπαραμέτρων είναι οι εποχές, δηλαδή πόσες φορές θα χρησιμοποιηθούν τα δεδομένα εκπαίδευσης για να μεταβληθούν οι παράμετροι του μοντέλου, ο ρυθμός μάθησης, δηλαδή πόσο μεγάλες θα είναι αυτές οι μεταβολές κάθε φορά και ο αριθμός των ομάδων (clusters) σε αλγόριθμο ομαδοποίησης της μη επιβλεπομένης μηχανικής μάθησης. Ο αλγόριθμος μπορεί να χρησιμοποιήσει πλέον τα δεδομένα εκπαίδευσης, ώστε να σχηματίσει το βέλτιστο μαθηματικό μοντέλο.

Το μαθηματικό μοντέλο δίνει λύση σε ένα πρόβλημα, το οποίο μπορεί να είναι δύο ειδών. Η ταξινόμηση (Classification) είναι ένα πρόβλημα στο οποίο ζητείται να βρεθεί η κατηγορία στην οποία ανήκει ένα στοιχείο εισόδου. Στην μηχανική μάθηση αυτό γίνεται με αλγόριθμο μάθησης ταξινόμησης (Classification learning algorithm), ο οποίος δέχεται παραδείγματα με κατηγορίες και δημιουργεί το μοντέλο. Εκείνο δέχεται ένα παράδειγμα χωρίς κατηγορία και επιστρέφει είτε την κατηγορία στην οποία ανήκει, ή έναν αριθμό από τον οποίο προκύπτει η κατηγορία (π.χ. πιθανότητες κάθε κατηγορίας για το δοθέν παράδειγμα). Αν ο αριθμός κατηγοριών είναι δύο, τότε πρόκειται για δυαδική ταξινόμηση. Αν είναι παραπάνω οι κατηγορίες, η ταξινόμηση είναι πολυταξική (multiclass / multinomial). Σε κάθε περίπτωση, ο αλγόριθμος μάθησης ταξινόμησης σχηματίζει σε κάποιο σημείο σύνορα απόφασης (decision boundary), με βάση τα οποία αποφασίζει σε ποια κατηγορία ανήκει ένα παράδειγμα. Το άλλο είδος προβλήματος είναι η παλινδρόμηση (Regression). Η παλινδρόμηση είναι πρόβλημα εύρεσης μίας πραγματικής τιμής ενός χαρακτηριστικού του παραδείγματος. Ο αλγόριθμος παλινδρόμησης δέχεται ως είσοδο παραδείγματα που έχουν τιμή (target) και δημιουργεί ένα μοντέλο που βρίσκει την τιμή αυτή σε παραδείγματα στα οποία είναι άγνωστη. Η πρόβλεψη της αξίας ενός ακινήτου με βάση τα χαρακτηριστικά του είναι μία απλή εφαρμογή παλινδρόμησης. Ο εντοπισμός ενός αντικειμένου σε μια εικόνα ανάγεται σε αλγόριθμο παλινδρόμησης, με την εύρεση συντεταγμένων των ακμών ενός παραλληλογράμμου γύρω από το αντικείμενο (bounding box regression).

Για τα δεδομένα που έχουν συλλεχθεί για την διαδικασία της εκπαίδευσης και που διαθέτουν τις τιμές εξόδου που αντιστοιχούν, υπάρχει μία σύμβαση με βάση την οποία αυτά χωρίζονται σε τρεις κατηγορίες. Αυτές είναι το σύνολο εκπαίδευσης (training set), το σύνολο επικύρωσης (validation set) και το σύνολο δοκιμής (test set). Το σύνολο των δεδομένων ανακατεύεται και χωρίζεται τυχαία στα τρία υποσύνολα. Το σύνολο εκπαίδευσης είναι τα δεδομένα που θα χρησιμοποιηθούν από τον αλγόριθμο για να σχηματίσει και να προσαρμόσει το μαθηματικό μοντέλο, με βάση τις αποκλίσεις προβλέψεων και αναμενόμενων δεδομένων εξόδου. Αυτό το υποσύνολο είναι το μεγαλύτερο σε όγκο δεδομένων, ώστε το μοντέλο να προσαρμόζεται σε όσο το δυνατό καλύτερο βαθμό και να είναι σε θέση να καλύπτει πολλές περιπτώσεις του υπό μελέτη προβλήματος. Το δεύτερο υποσύνολο, το σύνολο επικύρωσης, αποτελείται από δεδομένα τα οποία προέρχονται από τον πληθυσμό των δεδομένων εκπαίδευσης, ωστόσο δεν συμμετέχουν στην εκπαίδευση του αλγορίθμου. Αυτά χρησιμοποιούνται για μία αρχική εκτίμηση της ακρίβειας του μοντέλου, δηλαδή σε πόσα από αυτά οι προβλέψεις επαληθεύονται από τις δοθείσες τιμές εξόδου. Με βάση την εκτίμηση της ακρίβειας του

μοντέλου, μπορεί είτε να επιλεγεί άλλο μοντέλο το οποίο μπορεί να προσαρμοστεί καλύτερα στις απαιτήσεις του προβλήματος, είτε να τροποποιηθούν οι υπερπαραμέτροι ώστε να βελτιωθεί η απόδοση του επιλεγμένου μοντέλου. Αν η ακρίβεια στο σύνολο επικύρωσης είναι ικανοποιητική και το μοντέλο έχει επιλεχθεί και εκπαιδευτεί, εφαρμόζεται στο τρίτο υποσύνολο, το σύνολο δοκιμής. Το υποσύνολο αυτό αντικατοπτρίζει τα δεδομένα του πραγματικού κόσμου τα οποία δεν έχει δει το μοντέλο ξανά και χρησιμοποιείται για να υπολογιστεί η τελική ακρίβεια του μοντέλου και κατά επέκταση του αλγορίθμου μηχανικής μάθησης.¹⁶

Η τελική ακρίβεια του μοντέλου πάνω στο σύνολο δοκιμής χρησιμοποιείται, μέσω διαφόρων δεικτών και εργαλείων, για την αξιολόγηση της απόδοσης του μοντέλου σε προβλέψεις. Στην περίπτωση της παλινδρόμησης, ως βάση σύγκρισης μπορεί να χρησιμοποιηθεί το μέσο μοντέλο, δηλαδή ένα μοντέλο το οποίο πάντα προβλέπει τον μέσο όρο των τιμών στο σύνολο εκπαίδευσης ως την ζητηθείσα τιμή. Σε κάθε περίπτωση ένα προσαρμοσμένο μοντέλο πρέπει να έχει προβλέψεις που ανταποκρίνονται στην πραγματικότητα καλύτερα από αυτές του μέσου μοντέλου. Επόμενο βήμα είναι η σύγκριση της απόδοσης του μοντέλου στο σύνολο εκπαίδευσης, δηλαδή τις τιμές που έχει ήδη δει και έχει χρησιμοποιήσει για την διαμόρφωσή του, με την απόδοση στο σύνολο δοκιμών, τις τιμές του οποίου δεν έχει δει. Η σύγκριση αυτή μπορεί να γίνει συγκρίνοντας το μέσο τετραγωνικό σφάλμα (1) που παρουσιάζει το μοντέλο στο σύνολο εκπαίδευσης, με αυτό που παρουσιάζει στο σύνολο δοκιμής. Το μέσο τετραγωνικό σφάλμα ορίζεται ως:

$$\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (1)$$

Όπου n το μέγεθος του συνόλου, Y_i η πρόβλεψη της τιμής και \hat{Y}_i η τιμή που έχει δοθεί ως πραγματική τιμή εξόδου. Αν το μέσο τετραγωνικό σφάλμα είναι σημαντικά μεγαλύτερο στο σύνολο δοκιμής από αυτό στο σύνολο εκπαίδευσης, τότε υπάρχει υπερπροσαρμογή (overfitting) του μοντέλου στα δεδομένα εκπαίδευσης και το μοντέλο δεν μπορεί να γενικεύσει τον τρόπο πρόβλεψης του σε δεδομένα που δεν έχει δει. Σε αυτή την περίπτωση πρέπει ή να μεταβληθούν από τον χρήστη του αλγορίθμου οι υπερπαραμέτροι, ή να εφαρμοστεί μία μέθοδος γνωστή ως ομαλοποίηση του μοντέλου (Regularization). Στην περίπτωση της ταξινόμησης, υπάρχει πληθώρα δεικτών που αξιολογούν την ικανότητα του μοντέλου να εντάσσει μία παρατήρηση σε σωστή κατηγορία. Οι δείκτες επεξηγούνται για την περίπτωση της δυαδικής ταξινόμησης.

Ο πρώτος δείκτης, ο οποίος είναι ο πιο εποπτικός και μέσω του οποίου υπολογίζονται άλλοι δείκτες είναι ο πίνακας σύγχυσης. Πρόκειται για έναν πίνακα που δείχνει πόσο επιτυχημένο είναι το μοντέλο ταξινόμησης στην πρόβλεψη παραδειγμάτων που ανήκουν σε διάφορες κατηγορίες. Ο ένας άξονας του πίνακα σύγχυσης περιέχει τις κατηγορίες που το μοντέλο προέβλεψε, ενώ ο άλλος άξονας περιέχει τις πραγματικές κατηγορίες των παραδειγμάτων. Αν υποθέσουμε ένα πρόβλημα κατηγοριοποίησης μίας φωτογραφίας δομικού στοιχείου σε ρωγμή ή όχι ρωγμή, ο πίνακας σύγχυσης θα έχει την παρακάτω μορφή:

| | Ρωγμή (Πρόβλεψη) | Όχι_Ρωγμή (Πρόβλεψη) |
|----------------------------|------------------|----------------------|
| Ρωγμή (Πραγματικότητα) | TP | FN |
| Όχι_Ρωγμή (Πραγματικότητα) | FP | TN |

Κάθε παράδειγμα που εισέρχεται στον αλγόριθμο και χρησιμοποιείται για να προβλεφθεί η κατηγορία στην οποία ανήκει, κατατάσσεται σε ένα από τα τεταρτημόρια του πίνακα σύγχυσης. Φωτογραφίες οι οποίες απεικονίζουν ρωγμές και προβλέπονται από το μοντέλο ως ρωγμές ονομάζονται αληθώς θετικές (True Positive, TP), ενώ αν προβλέπονται ως όχι ρωγμές ονομάζονται ψευδώς αρνητικές (False Negative, FN). Αντιθέτως, φωτογραφίες που απεικονίζουν μη ρηγματωμένα δομικά στοιχεία και προβλέπονται ως ρωγμές ονομάζονται ψευδώς θετικές (False Positive, FP), ενώ αν προβλέπονται ως όχι ρωγμές ονομάζονται αληθώς αρνητικές (True Negative, TN). Ο πίνακας σύγχυσης μπορεί να χρησιμοποιηθεί σε πολυταξική ταξινόμηση, όπου μπορούν να εντοπιστούν μοτίβα κατηγοριών που μπερδεύει συχνότερα το μοντέλο και ταξινομεί παραδείγματα της μίας στις άλλες και αντίστροφα. Από τα στοιχεία του πίνακα, τα αληθώς θετικά και τα ψευδώς αρνητικά (TP, FN) είναι τα στοιχεία που είναι επιθυμητό να εντοπιστούν από τον αλγόριθμο. Ονομάζονται σχετικά στοιχεία (relevant elements) και είναι αυτά που απεικονίζουν ρωγμές. Τα στοιχεία που όντως εντοπίζει και επιστρέφει ως θετικά στοιχεία ο αλγόριθμος είναι τα αληθώς και ψευδώς θετικά (TP, FP). Ονομάζονται ανακτηθέντα στοιχεία (retrieved elements) και είναι αυτά στα οποία το μοντέλο θεωρεί πως υπάρχουν ρωγμές. Τα στοιχεία που έχουν χωριστεί στα τεταρτημόρια του πίνακα σύγχυσης χρησιμοποιούνται για τον υπολογισμό άλλων δεικτών απόδοσης, όπως η ακρίβεια (precision), (2) και η ανάκληση (recall), (3) οι οποίοι είναι και οι πιο συχνά χρησιμοποιούμενοι (Εικόνα 2).

Η ακρίβεια είναι ο λόγος των αληθώς θετικών προβλέψεων προς τον συνολικό αριθμό των θετικών προβλέψεων. Ο δείκτης αυτός δείχνει πόσα από τα ανακτηθέντα στοιχεία του αλγορίθμου είναι όντως σχετικά.

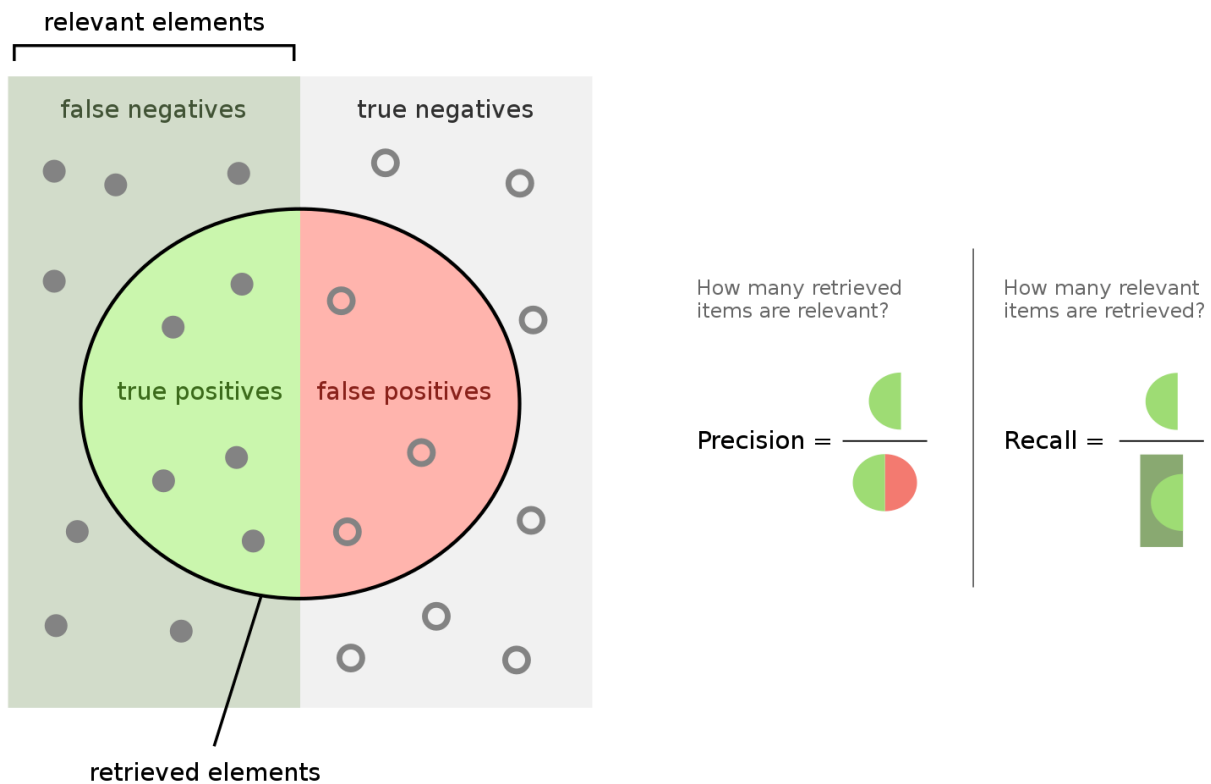
$$precision = \frac{TP}{TP + FP} \quad (2)$$

Η ανάκληση είναι ο λόγος των αληθώς θετικών προβλέψεων προς τον συνολικό αριθμό των σχετικών στοιχείων, αληθώς θετικών και ψευδώς αρνητικών. Ο δείκτης δείχνει πόσα από τα σχετικά στοιχεία μπορεί να αναγνωρίσει και να ανακτήσει ο αλγόριθμος.

$$recall = \frac{TP}{TP + FN} \quad (3)$$

Στην πράξη, ο αλγόριθμος μπορεί να επιτύχει ή υψηλή ακρίβεια, ή υψηλή ανάκληση. Ανάλογα με το πρόβλημα που χρειάζεται να επιλύσει ο αλγόριθμος μηχανικής μάθησης, επιλέγεται ένας από τους δύο δείκτες ως σημαντικότερος και εφαρμόζονται τρόποι αύξησης της τιμής του. Αυτοί είναι η εφαρμογή μεγαλύτερου βάρους σε παραδείγματα κάποιων

κατηγορίας, η τροποποίηση υπερπαραμέτρων ώστε ο επιθυμητός δείκτης να αυξηθεί στο σύνολο επικύρωσης και τέλος, η τροποποίηση του συνόρου απόφασης σε αλγόριθμους που επιστρέφουν πιθανότητες κατηγοριών (για παράδειγμα η επιβολή στον αλγόριθμο ένα παράδειγμα να προβλεφθεί ως θετικό, μόνο αν η πιθανότητα να είναι θετικό που δίνει ο αλγόριθμος στο παράδειγμα υπερβαίνει το 90%). Στην περίπτωση πολυταξικής ταξινόμησης, μελετάται μία κατηγορία ως η θετική και οι υπόλοιπες ως αρνητικές, ώστε να υπολογιστούν οι δείκτες για την μία κατηγορία.



Εικόνα 2: Σχηματική αναπαράσταση των δεικτών *precision* / *recall**

Υπάρχει ωστόσο ο δείκτης F_1 (4), ο οποίος αποτελείται από τον αρμονικό μέσο των δύο δεικτών και μπορεί να χρησιμοποιηθεί αν έχουν σημασία και οι δύο. Η υψηλότερη πιθανή τιμή του είναι το 1, το οποίο δείχνει τέλεια ακρίβεια και ανάκληση, και 0, το οποίο σημαίνει πως είτε η ανάκληση, είτε η ακρίβεια ισούται με 0.

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \quad (4)$$

Η ορθότητα (*accuracy*), (5) είναι άλλος ένας δείκτης που προκύπτει από τον πίνακα σύγχυσης. Δίνεται από τον λόγο του αριθμού των σωστά ταξινομημένων παραδειγμάτων προς τον συνολικό αριθμό των παραδειγμάτων.

* (https://en.wikipedia.org/wiki/Precision_and_recall)

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

Η ορθότητα είναι χρήσιμος δείκτης όταν όλα τα σφάλματα σε προβλέψεις κατηγοριών είναι εξίσου σημαντικά. Στην περίπτωση που διαφορετικές κατηγορίες είναι περισσότερο ή λιγότερο σημαντικές, υπάρχει ο δείκτης της ευαίσθητης σε κόστος ορθότητας (cost-sensitive accuracy). Για τον υπολογισμό της, ορίζεται ένας θετικός αριθμός, το κόστος, για τα σφάλματα, δηλαδή ψευδώς θετικές και αρνητικές προβλέψεις (FP, FN). Στη συνέχεια απαριθμούνται οι προβλέψεις κάθε τεταρτημόριου, πολλαπλασιάζεται ο αριθμός ψευδώς θετικών και αρνητικών προβλέψεων με το κόστος και υπολογίζεται η ορθότητα, σύμφωνα με τον προαναφερθέντα τύπο.

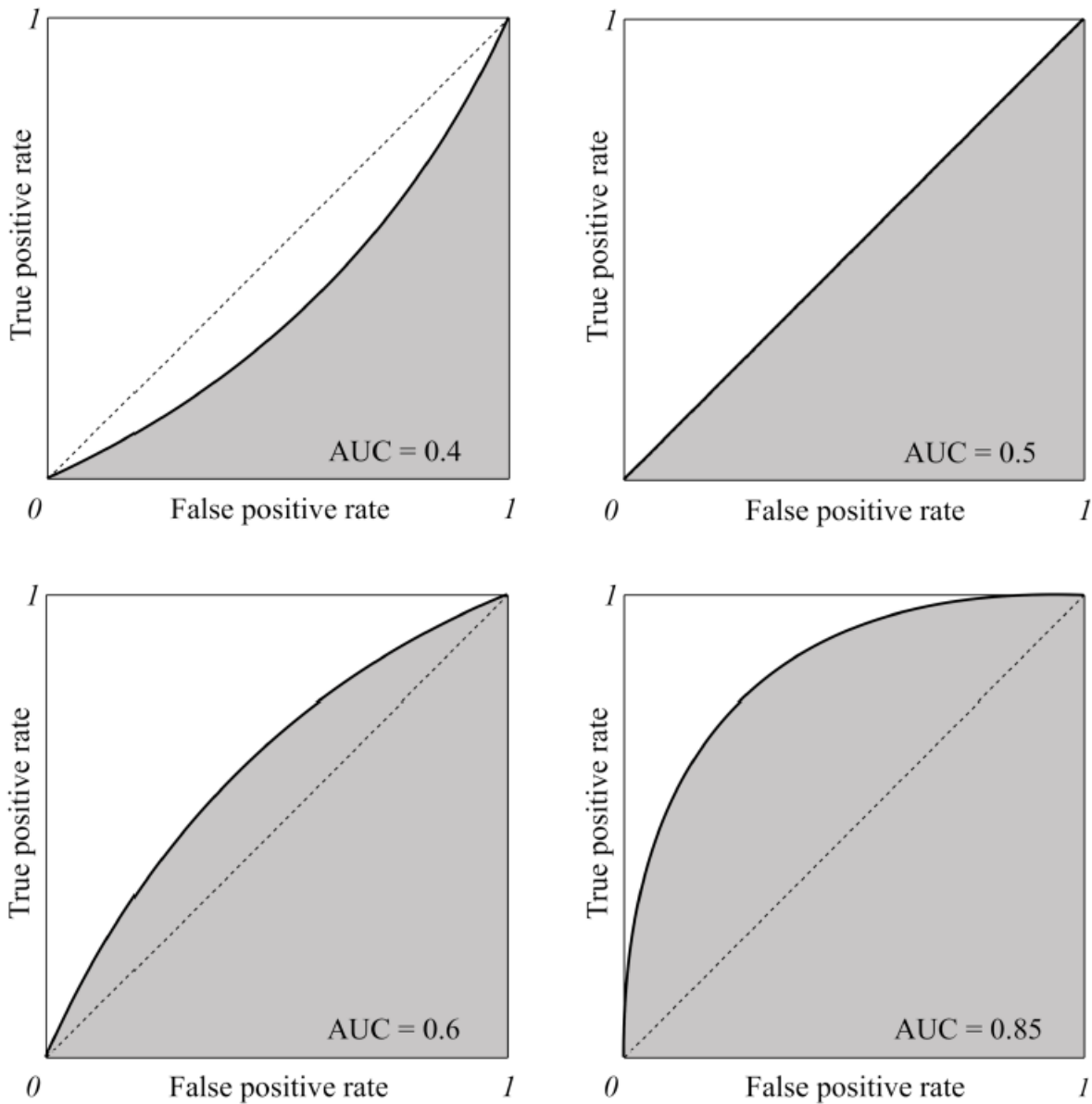
Ένας πιο σύνθετος δείκτης που στηρίζεται στον πίνακα σύγχυσης είναι η επιφάνεια κάτω από καμπύλη ROC (Area Under the roc Curve, AUC). Η καμπύλη ROC (receiver operating characteristic) χρησιμοποιεί έναν συνδυασμό δεικτών, του λόγου αληθώς θετικών (True Positive Rate, TPR) (6) και του λόγου ψευδώς θετικών (False Positive Rate, FPR) (7) οι οποίοι ορίζονται ως εξής:

$$TPR = \frac{TP}{TP + FN} \quad (6)$$

$$FPR = \frac{FP}{FP + TN} \quad (7)$$

Οι καμπύλες ROC μπορούν να χρησιμοποιηθούν μόνο για ταξινομητές οι οποίοι επιστρέφουν πιθανότητες προβλέσεων. Για να γραφεί η καμπύλη, πρέπει πρώτα να γίνει κατάτμηση του εύρους πιθανοτήτων σε τμήματα. Αν οι πιθανότητες που επιστρέφει ο αλγόριθμος ανήκουν σε διάστημα $[0,1]$, μια πιθανή κατάτμηση μπορεί να γίνει σε διαστήματα μεγέθους 0.1. Στη συνέχεια εφαρμόζεται το μοντέλο και γίνεται η ταξινόμηση, χρησιμοποιώντας ως σύνορο απόφασης κάθε μέγιστη τιμή τμήματος της κατάτμησης. Έτσι ένα παράδειγμα προβλέπεται ως θετικό για τιμή της πιθανότητας κατηγορίας > 0.1 , υπολογίζεται ο πίνακας σύγχυσης και άρα οι δείκτες TPR, FPR. Στη συνέχεια επαναλαμβάνεται η ταξινόμηση για σύνορο απόφασης 0.2 και ομοίως μέχρι και την τιμή πιθανότητας κατηγορίας ίση με 1. Εφόσον έχουν υπολογιστεί κατά αυτόν τον τρόπο ζεύγη δεικτών TPR, FPR, δημιουργείται ένα γράφημα με τιμές του TPR από 0 έως 1 στον άξονα των y και τιμές του FPR από 0 έως 1 στον άξονα των x και τοποθετείται πάνω κάθε υπολογισμένο ζεύγος. Είναι σημαντικό να αναφερθεί πως αν το σύνορο απόφασης τεθεί ίσο με 0, τότε όλες οι προβλέψεις είναι θετικές και άρα οι δύο δείκτες ισούνται με 1, ενώ αν το σύνορο απόφασης τεθεί ίσο με πιθανότητα 1 τότε όλες οι προβλέψεις είναι αρνητικές και άρα οι δύο δείκτες ισούνται με 0. Επομένως ζεύγη TPR και FPR θα βρίσκονται στα σημεία $[0,0]$ και $[1,1]$ και κάθε καμπύλη ROC θα περνά από αυτά ανεξαρτήτως μοντέλου. Σαν δείκτης αξιολόγησης, ένα μοντέλο ταξινόμησης έχει καλύτερη απόδοση, όσο μεγαλύτερη είναι η επιφάνεια κάτω από την καμπύλη. Ταξινομητής με επιφάνεια κάτω της καμπύλης μεγαλύτερη από 0.5 είναι καλύτερος από έναν τυχαίο ταξινομητή. Αντίθετα αν η επιφάνεια κάτω από την καμπύλη είναι κάτω από 0.5 τότε το μοντέλο παρουσιάζει κάποιο πρόβλημα. Ένας υποθετικός άψογος ταξινομητής θα έχει επιφάνεια ίση με 1. Αν το μοντέλο έχει ικανοποιητική επιφάνεια κάτω από την καμπύλη

ROC, ένας καλός ταξινομητής είναι αυτός που χρησιμοποιεί ως σύνορο απόφασης μία τιμή της πιθανότητας που δίνει TPR κοντά στο 1, ενώ διατηρεί το FPR όσο είναι δυνατόν κοντά στο 0 (Εικόνα 3).



Εικόνα 3: Η περιοχή κάτω από την καμπύλη ROC*

Όλα τα βήματα ενός αλγορίθμου μηχανικής μάθησης, όπως αυτά περιγράφονται στο εδάφιο αυτό, υλοποιούνται σε κάποια γλώσσα προγραμματισμού και εκτελούνται από ηλεκτρονικούς υπολογιστές. Κατά κανόνα δεν γράφονται από την αρχή¹⁵, αλλά χρησιμοποιούνται υπάρχουσες βιβλιοθήκες, δομές / πλαίσια (frameworks) και πλατφόρμες λογισμικού σε γλώσσες προγραμματισμού C++ και Python. Παραδείγματα αυτών είναι τα Scikit-learn, Pytorch, Tensorflow, Keras, Caffe, Theano.¹⁷

*(Andriy Burkov. (2019). The Hundred-Page Machine Learning Book)

2.3 Χρησιμότητα της μηχανικής μάθησης στην καθημερινότητα

Η μηχανική μάθηση έχει ως σκοπό της δημιουργία ενός μοντέλου πρόβλεψης, το οποίο δίνει λύση στο πρόβλημα εύρεσης μίας τιμής στην περίπτωση της παλινδρόμησης, ή στο πρόβλημα εύρεσης μίας κατηγορίας στην οποία ανήκει μία οντότητα, στην περίπτωση της ταξινόμησης. Τα προβλήματα αυτά συναντώνται στην καθημερινή ζωή και μέσω της μηχανικής μάθησης επιχειρείται είτε να αντικατασταθούν ορισμένες πιθανώς επίπονες ανθρώπινες δραστηριότητες, είτε να συνδράμουν αλγόριθμοι σε ερευνητικές μεθόδους λόγω της ταχύτητας εκτέλεσης πράξεων των υπολογιστών ώστε να λυθούν αναπάντητα ερωτήματα της επιστήμης.

Αλγόριθμοι και μοντέλα μηχανικής μάθησης εφαρμόζονται πλέον σε κάθε πιθανή έκφανση της ανθρώπινης ζωής. Αυτό οφείλεται σε μεγάλο βαθμό στην ύπαρξη φτηνών, υψηλής ποιότητας συλλογών δεδομένων στην ψηφιακή εποχή καθώς και στην ανάπτυξη καρτών γραφικών υψηλής απόδοσης, οι οποίες μπορούν να εκτελούν τεράστιο όγκο πράξεων σε σχετικά μικρό χρονικό διάστημα. Στη συνέχεια αναφέρονται ορισμένα παραδείγματα επίδρασης της μηχανικής μάθησης.

Χαρακτηριστικό παράδειγμα είναι ο αυτόματος πιλότος των ηλεκτρικών αυτοκινήτων που κατασκευάζει η εταιρεία Tesla. Τα νεότερα μοντέλα διαθέτουν 8 εξωτερικές κάμερες και 12 αισθητήρες υπερήχου και χρησιμοποιούν νευρωνικά δίκτυα για τον χωρικό εντοπισμό και αναγνώριση αντικειμένων στην ευρύτερη περιοχή του αυτοκινήτου. Με βάση τα δεδομένα που συλλέγουν οι κάμερες και οι αισθητήρες, το αυτοκίνητο προβαίνει σε συγκεκριμένες πράξεις, όπως αυτόματη οδήγηση με βάση της γραμμές του οδοστρώματος, αυτόματη αλλαγή λωρίδας κυκλοφορίας, αυτόματο παρκάρισμα και αναγνώριση πινακίδων και φωτεινών σηματοδοτών και επιβράδυνση του αυτοκινήτου σε αυτήν την περίπτωση. Υπάρχει ωστόσο η επισήμανση ότι οι δυνατότητες αυτές προορίζονται για χρήση από έναν πλήρως προσεκτικό οδηγό ο οποίος θα είναι σε θέση να πάρει ανά πάσα στιγμή τον έλεγχο του οχήματος. Οι δυνατότητες δηλαδή συμπληρώνουν την δράση του οδηγού, δεν την αντικαθιστούν πλήρως, αν και συνεχίζεται η περαιτέρω ανάπτυξή τους.

Η διαδικτυακή πλατφόρμα Youtube χρησιμοποιεί ανατροφοδοτούμενα νευρωνικά δίκτυα (Recurring Neural Networks, RNN) και πιο συγκεκριμένα δίκτυα μακράς βραχύχρονης μνήμης (Long Short-term Memory, LSTM) για αναγνώριση ομιλίας και ήχου στα βίντεο που φιλοξενεί στην ιστοσελίδα της. Η αλγόριθμοι μηχανικής μάθησης χρησιμοποιούνται για την δημιουργία αυτόματων υποτίτλων στα βίντεο, βελτιώνοντας την προσβασιμότητα για πολλούς θεατές χωρίς να απαιτείται δαπάνη εκατομμυρίων εργατοωρών για χειροκίνητη καταγραφή ομιλίας και απόδοση υποτίτλων.

Εκτενής είναι και η χρήση της μηχανικής μάθησης στο πεδίο της βιοπληροφορικής, με περίπου 8% των δημοσιεύσεων στην βαθειά μάθηση να σχετίζεται με αυτό. Ειδικότερα, στον κλάδο της γονιδιωματικής, μοντέλα βαθειάς μάθησης μπορούν να αναπαραστήσουν ικανοποιητικά τις απολύτως βασικές βιολογικές διεργασίες, όπως το δίπλωμα πρωτεϊνών και άρα να συμβάλλουν στην χαρτογράφηση γονιδιωμάτων.¹⁸

Λογικό επακόλουθο είναι η εφαρμογή τεχνικών μηχανικής μάθησης και στον κατασκευαστικό τεχνικό τομέα, αντικείμενο του οποίου αποτελεί η παρούσα εργασία. Σε επόμενα κεφάλαια γίνεται παράθεση της υπάρχουσας βιβλιογραφίας πάνω στην έρευνα

μηχανικής μάθησης σε θέματα του κατασκευαστικού τομέα και αναλύονται συγκεκριμένα συμπεράσματα και προβληματισμοί που προκύπτουν από αυτήν.

2.4 Επιφυλάξεις για τη χρήση μηχανικής μάθησης

Η χρήση αλγορίθμων για δημιουργία μοντέλων μέσω της μηχανικής μάθησης, αν και επαναστατική και πρωτοπόρα μέθοδος, εγκυμονεί κινδύνους λόγω πιθανών προβλημάτων τόσο στην διαδικασία της εκπαίδευσης όσο και στην εφαρμογή των μοντέλων αυτών σε καταστάσεις του πραγματικού κόσμου. Τα προβλήματα αυτά ανήκουν σε δύο κατηγορίες διαφορετικής φύσεως. Η μία αφορά προβλήματα που παρουσιάζουν οι αλγόριθμοι μηχανικής μάθησης ως επιστημονική μέθοδος, ενώ η άλλη είναι ηθικής φύσεως και αφορά πιθανές αρνητικές επιπτώσεις στο σύνολο της κοινωνίας.

Το κύριο πρόβλημα που αντιμετωπίζουν οι τεχνικές μηχανικής μάθησης ως επιστημονική μέθοδος την παρούσα στιγμή, είναι η αναξιοπιστία των μοντέλων λόγω της σχετικά χαμηλής ακρίβειας σε προβλέψεις. Η επιθυμητή ακρίβεια είναι σε κάθε περίπτωση 100%, κυρίως σε δεδομένα τα οποία ένα μοντέλο δεν έχει ξαναδεί. Ωστόσο αυτό είναι ιδεατή κατάσταση. Σπάνια μπορεί να πραγματοποιηθεί σε πολύ απλά προβλήματα πρόβλεψης, αλλά όσο τα προβλήματα που απαιτείται να λυθούν γίνονται πιο περίπλοκα, η ακρίβεια ακόμα και πολύ ισχυρών μοντέλων πέφτει στο 95% ή σε κάποιες περιπτώσεις και κοντά στο 90%. Επομένως μεγάλο τμήμα της έρευνας πάνω σε αλγορίθμους μηχανικής μάθησης εστιάζει σε τρόπους αύξησης της ακρίβειας, ώστε τα μοντέλα που δημιουργούνται να κάνουν όσο λιγότερα λάθη γίνεται. Όπως αναφέρθηκε είναι αδύνατον να δημιουργηθούν άψογα μοντέλα, αλλά με τον ίδιο τρόπο ο ανθρώπινος εγκέφαλος παρουσιάζει σφάλματα κρίσης και δεν μπορεί να πάρει σωστές αποφάσεις στο 100% των περιπτώσεων κάποιου προβλήματος.

Την μεγαλύτερη επίδραση στην ακρίβεια ενός μοντέλου την έχει ο τρόπος που έχει προσαρμοστεί το μοντέλο στα δεδομένα εκπαίδευσης. Υπάρχουν δύο καταστάσεις οι οποίες κρίνεται σκόπιμο να αποφευχθούν κατά την διαδικασία της εκπαίδευσης, η υποπροσαρμογή (underfitting) ενός μοντέλου και η υπερπροσαρμογή (overfitting) του μοντέλου. Η υποπροσαρμογή χαρακτηρίζεται από χαμηλή ακρίβεια προβλέψεων στα δεδομένα εκπαίδευσης, λόγω αδυναμίας του μοντέλου να προβλέψει σωστά ακόμα και τα δεδομένα στα οποία έχει εκπαιδευτεί ήδη. Αυτό μπορεί να συμβαίνει για δύο λόγους. Η τα δεδομένα εισόδου δεν μπορούν να συσχετιστούν με τα αναμενόμενα δεδομένα εξόδου, ή το μοντέλο που έχει επιλεχθεί για να κάνει προβλέψεις είναι υπερβολικά απλό για να προσαρμοστεί στα δεδομένα και άρα έχει υψηλή μεροληψία, η οποία είναι είδος σφάλματος. Η λύση στην υποπροσαρμογή του μοντέλου είναι ή η κατασκευή καλύτερων δεδομένων εισόδου που μπορούν να χρησιμοποιηθούν για πρόβλεψη, ή η επιλογή πιο περίπλοκου μοντέλου που προσαρμόζεται καλύτερα στα δεδομένα του προβλήματος που επιχειρείται να λυθεί μέσω της μηχανικής μάθησης. Η υπερπροσαρμογή ενός μοντέλου χαρακτηρίζεται από πολύ υψηλή ακρίβεια στο σύνολο εκπαίδευσης, η οποία δεν μπορεί να επιτευχθεί στα σύνολα επικύρωσης και δοκιμής, που παρουσιάζουν σημαντικά χαμηλότερη ακρίβεια. Η υπερπροσαρμογή μπορεί να οφείλεται σε μοντέλο υπερβολικά περίπλοκο για τα δεδομένα που χρησιμοποιούνται στην εκπαίδευση, σε υπερβολικά μεγάλο αριθμό στοιχείων που αξιοποιούνται από κάθε παράδειγμα με μικρό αριθμό παραδειγμάτων ή σε εκπαίδευση πάνω στα ίδια παραδείγματα με ακατάλληλο αριθμό εποχών εκπαίδευσης (συνήθως περισσότερων από όσο χρειάζεται για

να προκύψει ακριβές μοντέλο). Το αποτέλεσμα είναι το μοντέλο να απομνημονεύει στοιχεία των παραδειγμάτων τα οποία μπορεί να μην σχετίζονται με τις προσδοκώμενες προβλέψεις, όπως ο θόρυβος ή τυχαία χαρακτηριστικά των παραδειγμάτων. Όταν το μοντέλο εφαρμοστεί σε δεδομένα που δεν έχει δει, σχηματίζει αποφάσεις με βάση τα στοιχεία που υπάρχουν μόνο στα δεδομένα που έχει δει και άρα εμφανίζει μεγάλες αστοχίες. Η υπερπροσαρμογή μπορεί να αντιμετωπιστεί με εφαρμογή απλούστερου μοντέλου, χρησιμοποίηση λιγότερο πολυδιάστατων παραδειγμάτων, προσθήκη περισσότερων δεδομένων εκπαίδευσης εφόσον αυτό είναι δυνατό, χρήση σωστών υπερπαραμέτρων και κυρίως σωστού αριθμού εποχών και ομαλοποίηση του μοντέλου.

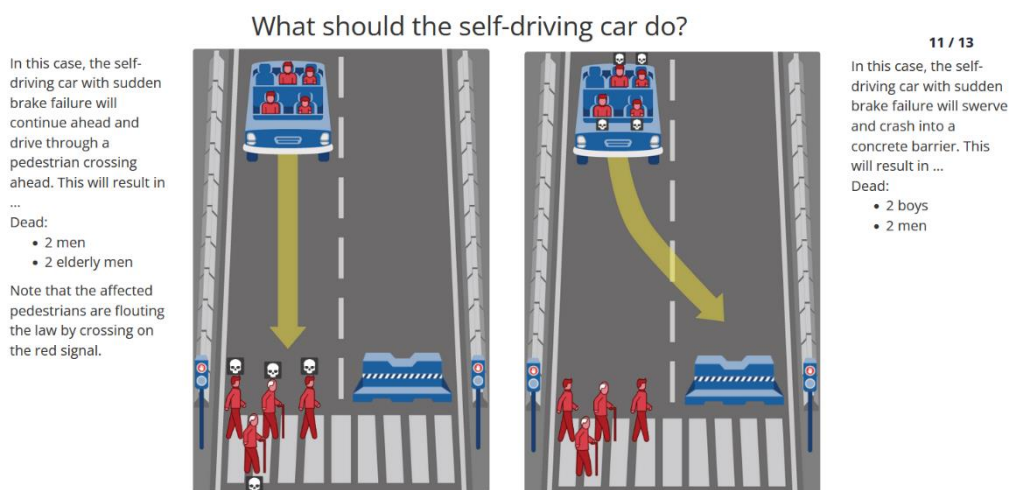
Αν υπάρχει η ένδειξη ότι το μοντέλο δεν υποπροσαρμόζεται ή υπερπροσαρμόζεται στα δεδομένα, η περαιτέρω αύξηση της αποτελεσματικότητας του σε προβλέψεις είναι πιο δύσκολη διαδικασία. Ένα μεγάλο πρόβλημα στην προσπάθεια αύξησης της ακρίβειας, το οποίο είναι ιδιαίτερα έντονο σε περιπτώσεις βαθιάς μάθησης με νευρωνικά δίκτυα, είναι ότι το μοντέλο λειτουργεί ως «μαύρο κουτί». Αυτό σημαίνει ότι στον άνθρωπο μπορούν να είναι κατανοητά τα επίπεδα εισόδου και εξόδου, ωστόσο τα ενδιάμεσα επίπεδα είναι κατανοητά μόνο από τον υπολογιστή και σε πολλά σημεία μπορεί να μην έχουν άμεση φυσική σημασία, σε αντίθεση με μοτίβα που αντιλαμβάνεται για παράδειγμα ο ανθρώπινος εγκέφαλος. Το πρόβλημα αντιμετωπίζεται σε κάποιον βαθμό με την εφαρμογή πιο εκλεπτυσμένων αλγορίθμων που προσπαθούν να αποδώσουν σημασία σε όλα τα επίπεδα, όπως τα συνελκτικά νευρωνικά δίκτυα (Convolutional Neural Networks, CNN) και με το πάγωμα κάποιων επιπέδων ενός νευρωνικού δικτύου και εκπαίδευση μόνο κάποιων με χαμηλούς ρυθμούς μάθησης, ώστε να υπάρξει μια πιο ελεγχόμενη και στοχευμένη παρέμβαση στο μοντέλο που προσαρμόζεται μέσω της εκπαίδευσης. Ωστόσο ο πυλώνας του προβλήματος παραμένει, διότι είναι σχεδόν αδύνατο να είναι εκ των προτέρων γνωστή η επίδραση επιπλέον παραδειγμάτων ή μεταβολής των υπερπαραμέτρων στην ακρίβεια του μοντέλου. Είναι πιθανό οι προσπάθειες βελτίωσης της ακρίβειας να οδηγήσουν αντί αυτού σε μείωση της και ο μόνος τρόπος να εξακριβωθεί η επίδραση είναι ο πειραματισμός.

Από την ηθική διάσταση, η μηχανική μάθηση είναι αμφιλεγόμενη διαδικασία. Στο στάδιο της εκπαίδευσης παρουσιάζει γιγαντιαίες απαιτήσεις ενέργειας, οι οποίες οφείλονται στην μεγάλη υπολογιστική ισχύ που απαιτούν καθώς και στην μεγάλη χρονική περίοδο της εκπαίδευσης, η οποία για πολύπλοκα μοντέλα μπορεί να διαρκέσει από μέρες έως βδομάδες.¹⁰ Η χρήση καρτών γραφικών για τη μηχανική μάθηση, μαζί με άλλες διεργασίες για διαφορετικούς σκοπούς (εξόρυξη κρυπτονομισμάτων, φάρμες απόδοσης εικόνων) καταναλώνει μεγάλα ποσά ηλεκτρικής ενέργειας, η οποία σε πολλές περιπτώσεις παράγεται από φυσικά αποθέματα και όχι από ανανεώσιμες πηγές. Επομένως προτάσσεται το επιχείρημα ότι η μηχανική μάθηση είναι επιζήμια για το περιβάλλον και αντί να βελτιώνει την ανθρώπινη ζωή την υποβαθμίζει. Υπαρκτή είναι η έρευνα για την μείωση του περιβαλλοντικού αποτυπώματός της, ώστε να μεγιστοποιηθούν οι θετικές επιπτώσεις μειώνοντας τις αρνητικές.

Ένα άλλο ζήτημα που προκύπτει είναι η απόκτηση και η διαχείριση προσωπικών δεδομένων. Η διαδικασία της εκπαίδευσης απαιτεί μεγάλο όγκο δεδομένων, τα οποία για να αποκτηθούν χειροκίνητα χρειάζεται μεγάλο χρονικό διάστημα και δαπάνη ανθρώπινου δυναμικού. Με την αύξηση των χρηστών του διαδικτύου και των υπηρεσιών που παρέχονται από αυτό, αυξάνονται και τα δεδομένα που κυκλοφορούν ανά πάσα στιγμή. Χαρακτηριστικό παράδειγμα αποτελούν οι πλατφόρμες κοινωνικής δικτύωσης, όπου οι φωτογραφίες που

αναρτούν οι χρήστες έχουν χρησιμοποιηθεί από πληθώρα μοντέλων για εκπαίδευση και κατηγοριοποίηση. Ωστόσο η πλειοψηφία αυτών των χρηστών δεν γνωρίζει ότι τα δεδομένα τους χρησιμοποιούνται για μηχανική μάθηση, ή αν το γνωρίζουν δεν υπάρχει τρόπος να ξέρουν για ποιόν ακριβώς σκοπό και πόσες φορές χρησιμοποιούνται. Άρα δημιουργείται το ζήτημα του κατά πόσο είναι ηθικό να συμμετέχουν χωρίς την γνώση τους σε αλγορίθμους μηχανικής μάθησης και να παίρνουν μέρος σε διαδικασία η οποία θα αποφέρει πιθανά κέρδη χωρίς να τους δίνεται τμήμα αυτών.

Τέλος, η εφαρμογή μοντέλων μηχανικής μάθησης σε προβλήματα της καθημερινότητας εγείρει επιπρόσθετους ηθικούς δισταγμούς. Κάποιος εύλογα μπορεί να αναρωτηθεί με ποιο δικαίωμα λαμβάνει ένας αλγόριθμος αποφάσεις για καταστάσεις στην ζωή του, καθώς και ποιος θα αναλάβει την ευθύνη για λάθη ή αστοχίες κάποιου μοντέλου μηχανικής μάθησης. Επιπλέον, ένας από τους λόγους που είναι επιθυμητό να έχουν μεγάλη ακρίβεια οι προβλέψεις μοντέλων, είναι ώστε να μπορέσουν στη συνέχεια αλγόριθμοι να πάρουν τις σωστές αποφάσεις. Όμως υπάρχει περίπτωση ένα δοθέν πρόβλημα να μην έχει σωστή και λάθος απόφαση. Σε αυτήν την περίπτωση είναι αδιευκρίνιστο το πώς μπορεί να εκπαιδευτεί ένα μοντέλο μηχανικής μάθησης, πόσο μάλλον όταν ο ίδιος ο άνθρωπος δεν μπορεί να πάρει μία ικανοποιητική απόφαση και βρίσκεται μπροστά σε ένα ηθικό δίλημμα. Για την διερεύνηση του φαινομένου αυτού, δημιουργήθηκε ένα περιβάλλον σε μία διαδικτυακή σελίδα με την συνδρομή των Max Planck Institute, Massachusetts Institute of Technology, University of Exeter, Toulouse School of Economics και University of British Columbia. Η ιστοσελίδα με το όνομα moral machine (ηθική μηχανή) μελετά το υποθετικό πρόβλημα ενός αυτοοδηγούμενου αυτοκινήτου το οποίο έχει υποστεί βλάβη στα φρένα και ως αποτέλεσμα πρέπει να επιλέξει την δράση του σε διάφορα υποθετικά σενάρια. Οι επιλογές του έχουν σε κάθε περίπτωση απώλεια ανθρωπίνων ζωών, οπότε αφήνεται σε όποιον παίρνει μέρος στο πείραμα να επιλέξει αυτό που εκείνος θεωρεί ηθικά πιο σωστό με βάση δικά του κριτήρια. Τα πορίσματα της ιστοσελίδας σχετικά με την άποψη μεγάλου πληθυσμού πάνω σε θέματα, όπως το πώς οφείλουν να παίρνουν αποφάσεις οι μηχανές μπροστά σε ηθικά διλήμματα, και η συζήτηση σχετικά με αυτά έχουν χρησιμοποιηθεί σε δημοσιεύσεις τόσο ψυχολογίας, όσο και τεχνητής νοημοσύνης.



Εικόνα 4: Η ηθική μηχανή*

*(<https://www.moralmachine.net>)

2.5 Η στάση της ευρωπαϊκής ένωσης

Με το κείμενο «Μία ευρωπαϊκή προσέγγιση στη τεχνητή νοημοσύνη» η ευρωπαϊκή ένωση επιχειρεί να καθορίσει ένα πλαίσιο στο οποίο μπορεί να αναπτυχθεί και να αξιοποιηθεί αποτελεσματικά η χρήση αλγορίθμων μηχανικής μάθησης και η εφαρμογή τους μέσω τεχνητής νοημοσύνης. Με δεδομένο πως δεν υπάρχουν νομοθετικές ρυθμίσεις στο ελληνικό δίκαιο που να διέπουν τις μεθόδους μηχανικής μάθησης, καθώς και το γεγονός ότι το κείμενο αυτό περιγράφει πολλές διαδικασίες και πρωτόκολλα σχετικά και με την μηχανική μάθηση, κρίνεται σκόπιμο να αναλυθεί σε αυτή την εργασία η οποία ασχολείται περισσότερο με την μηχανική μάθηση παρά με την τεχνητή νοημοσύνη.

Η προσέγγιση αυτή εστιάζει σε δύο κατηγορίες, τις οποίες ονομάζει υπεροχή στην τεχνητή νοημοσύνη και εμπιστοσύνη στην τεχνητή νοημοσύνη. Σκοπεύει η βελτίωση της τεχνητής νοημοσύνης να βασίζεται στην εφαρμογή κανόνων οι οποίοι προωθούν την έρευνα και την βιομηχανική ικανότητα, ενώ ταυτόχρονα προασπίζουν την λειτουργία της αγοράς, του δημοσίου τομέα και την ανθρώπινη ασφάλεια και δικαιώματα. Το πακέτο στρατηγικών για την τεχνητή νοημοσύνη δημοσιοποιήθηκε τον Απρίλιο του 2021.

Η ευρωπαϊκή προσέγγιση στην υπεροχή στην τεχνητή νοημοσύνη στοχεύει στην ενδυνάμωση των ευρωπαϊκών δυνατοτήτων να είναι ανταγωνιστική παγκόσμια σε θέματα τεχνητής νοημοσύνης. Σε υψηλό επίπεδο, θα το επιτύχει αυτό διευκολύνοντας την ανάδειξη και την ενσωμάτωση της τεχνητής νοημοσύνης στην ευρωπαϊκή ένωση, καθιστώντας την ευρωπαϊκή ένωση περιβάλλον στο οποίο θα ευημερεί η τεχνητή νοημοσύνη από τα εργαστήρια μέχρι την αγορά, διασφαλίζοντας ότι η τεχνητή νοημοσύνη δουλεύει υπέρ των πολιτών και είναι μία δύναμη για το καλό της κοινωνίας και χτίζοντας μια στρατηγική ηγεσία σε τομείς υψηλού αντίκτυπου, δηλαδή περιβάλλον, υγεία, ρομποτική, δημόσιο τομέα, οικιακά, μεταφορές και γεωργία. Η επιτροπή και τα κράτη μέλη συμφώνησαν στην ενίσχυση της υπεροχής ενώνοντας δυνάμεις στις πολιτικές τεχνητής νοημοσύνης και στις επενδύσεις. Κρίσιμα συστατικά της στρατηγικής της ευρωπαϊκής επιτροπής για την τεχνητή νοημοσύνη είναι η μεγιστοποίηση πόρων και ο συντονισμός των επενδύσεων μέσω ευρωπαϊκών προγραμμάτων. Επισημαίνεται ότι η πρόσβαση σε υψηλής ποιότητας δεδομένα είναι καθοριστικός παράγοντας για την δημιουργία ισχυρών συστημάτων τεχνητής νοημοσύνης με υψηλές αποδόσεις και αναφέρονται υπάρχουσες πρωτοβουλίες που παρέχουν τις σωστές δομές για την ανάπτυξη τέτοιων συστημάτων.

Η ευρωπαϊκή προσέγγιση στην εμπιστοσύνη στην τεχνητή νοημοσύνη στοχεύει στην ανάπτυξη έμπιστης τεχνητής νοημοσύνης η οποία θα δημιουργεί ένα ασφαλές και φιλικό στην καινοτομία περιβάλλον για χρήστες, προγραμματιστές και παρόχους. Η επιτροπή προτάσσει τρεις διασυνδεδεμένες νομικές πρωτοβουλίες που θα συνεισφέρουν στην ανάπτυξη της έμπιστης τεχνητής νοημοσύνης. Σύμφωνα με το κείμενο, αυτές είναι ένα ευρωπαϊκό νομικό πλαίσιο για την τεχνητή νοημοσύνη για να διερευνηθούν τα ζητήματα των θεμελιωδών δικαιωμάτων και τα ζητήματα ασφαλείας, ευρωπαϊκοί κανόνες για την αντιμετώπιση θεμάτων υπευθυνότητας σχετικά με τις νέες τεχνολογίες, στις οποίες συμπεριλαμβάνεται η τεχνητή νοημοσύνη και η μηχανική μάθηση, και μία ανασκόπηση της νομοθεσίας περί τομεακής ασφαλείας.¹⁹

Πιο συγκεκριμένα, το νέο ευρωπαϊκό νομικό πλαίσιο για την τεχνητή νοημοσύνη θέτει κανόνες που εξασφαλίζουν ότι οι Ευρωπαίοι μπορούν να εμπιστευτούν τα αποτελέσματα από

την λειτουργία εφαρμογών τεχνητής νοημοσύνης. Ο λόγος για τον οποίο χρειάζονται κανόνες είναι ότι αν και τα περισσότερα συστήματα νοημοσύνης παρουσιάζουν περιορισμένο έως και καθόλου κίνδυνο και άρα χρησιμοποιούνται για την επίλυση κοινωνικών προκλήσεων, κάποια συστήματα ελλοχεύουν κινδύνους οι οποίοι πρέπει να αντιμετωπιστούν για να αποφευχθούν ανεπιθύμητες καταστάσεις. Ως παράδειγμα αναφέρει ότι συχνά δεν είναι δυνατόν να ανακαλυφθεί πως ένα σύστημα τεχνητής νοημοσύνης πήρε μία απόφαση ή ένα μοντέλο μηχανικής μάθησης έκανε μία πρόβλεψη και άρα έφτασε σε κάποιο συγκεκριμένο αποτέλεσμα. Συνεπώς, είναι δύσκολο να διαπιστωθεί πότε κάποιος έχει αδικηθεί σε περιπτώσεις όπως απόφαση πρόσληψης του ή αίτηση για την παροχή κάποιας κοινωφελούς υπηρεσίας. Οι προτεινόμενοι κανόνες θα αντιμετωπίζουν κινδύνους που δημιουργούνται ειδικά από εφαρμογές τεχνητής νοημοσύνης, θα προτείνουν μία λίστα από υψηλής επικινδυνότητας εφαρμογές, θα θέτουν σαφείς απαιτήσεις για συστήματα τεχνητής νοημοσύνης για υψηλής επικινδυνότητας εφαρμογές, θα καθορίζουν συγκεκριμένες υποχρεώσεις για χρήστες και παρόχους αυτών των εφαρμογών, θα προτείνει μία αξιολόγηση της συμμόρφωσης στους κανόνες πριν το σύστημα τεχνητής νοημοσύνης τεθεί σε λειτουργία ή εισέλθει στην αγορά, θα προτείνουν έναν τρόπο επιβολής των κανόνων αφού έχει εισέλθει ένα τέτοιο σύστημα στην αγορά και θα προτείνουν έναν φορέα διοίκησης για το ζήτημα αυτό σε ευρωπαϊκό και εθνικό επίπεδο.

Το νομικό πλαίσιο αυτό προσεγγίζει την εφαρμογή των συστημάτων τεχνητής νοημοσύνης βάσει της επικινδυνότητας τους. Ορίζει λοιπόν τέσσερις κατηγορίες κινδύνου, οι οποίες περιγράφονται από την πιο ριψοκίνδυνη έως την λιγότερο ριψοκίνδυνη, ενώ αναφέρονται παραδείγματα συστημάτων που εντάσσονται σε κάθε κατηγορία. Η υψηλότερη κατηγορία επικινδυνότητας είναι η κατηγορία μη αποδεκτής επικινδυνότητας, η οποία αποτελείται από συστήματα τεχνητής νοημοσύνης που θεωρούνται ως ξεκάθαρη απειλή στην ασφάλεια, την περιουσία και τα δικαιώματα των πολιτών. Αυτά τα συστήματα θεωρείται πως χρησιμοποιούν απαγορευμένες πρακτικές στον τομέα της τεχνητής νοημοσύνης, όπως τεχνικές που απευθύνονται στο υποσυνείδητο του ατόμου ή που αξιολογούν την κοινωνική συμπεριφορά ή προσωπικότητα τους²¹ και απαγορεύεται η ανάπτυξη και εφαρμογή τους εντός της επικράτειας της ευρωπαϊκής ένωσης. Σαν παραδείγματα αυτής της κατηγορίας αναφέρονται συστήματα κοινωνικής βαθμολογίας ή παιχνίδια με χρήση φωνητικής βοήθειας που προτρέπουν σε επικίνδυνη συμπεριφορά. Η επόμενη κατηγορία απαρτίζεται από συστήματα υψηλής επικινδυνότητας, τα οποία ορίζονται ως συστήματα που απαριθμούνται σε συγκεκριμένους τομείς. Αυτοί είναι η βιομετρική ταυτοποίηση και κατηγοριοποίηση φυσικών προσώπων, η διαχείριση και λειτουργία υποδομών ζωτικής σημασίας, η εκπαίδευση και επαγγελματική κατάρτιση, η απασχόληση, η διαχείριση εργαζομένων και πρόσβαση στην αυτοαπασχόληση, η πρόσβαση σε δημόσιες ή ιδιωτικές υπηρεσίες και παροχές, η επιβολή του νόμου, η διαχείριση μετανάστευσης, ασύλου και συνοριακών ελέγχων και, τέλος, η απονομή δικαιοσύνης και οι δημοκρατικές διαδικασίες. Οι εφαρμογές που χρησιμοποιούν συστήματα που εντάσσονται σε αυτή την κατηγορία υποβάλλονται σε αυστηρές υποχρεώσεις πριν κυκλοφορήσουν στην αγορά, οι οποίες είναι επαρκή συστήματα ελέγχου και μετριάσμού της επικινδυνότητας τους, χρήση υψηλής ποιότητας συλλογών δεδομένων για να ελαχιστοποιηθούν οι κίνδυνοι και οι διακρίσεις, καταγραφή της λειτουργίας τους για ιχνηλάτηση των αποτελεσμάτων τους, λεπτομερής τεκμηρίωση λειτουργίας τους και παροχή πληροφοριών για να ελεγχθεί η συμμόρφωση στους κανονισμούς από της αρχές, ξεκάθαρες και επαρκείς πληροφορίες διαθέσιμες στον χρήστη και, τέλος, υψηλή ασφάλεια και ακρίβεια του συστήματος. Η επόμενη κατηγορία περιορισμένης επικινδυνότητας περιέχει συστήματα τα οποία υποχρεούνται να λειτουργούν με δεδομένη διαφάνεια, όπως λογισμικά που

απαντούν με γραπτή συνομιλία σε κάποιον χρήστη, ενώ η τελευταία κατηγορία είναι αυτή της ελάχιστης επικινδυνότητας, όπως φίλτρα ανεπιθύμητης αλληλογραφίας και ηλεκτρονικά παιχνίδια, κατηγορία η οποία όπως υποστηρίζει το κείμενο είναι η πλειοψηφία συστημάτων τεχνητής νοημοσύνης που χρησιμοποιούνται στην ευρωπαϊκή ένωση.

Μετά την κατάθεση της πρότασης αυτής τον Απρίλιο του 2021, οι κανονισμοί λειτουργίας μπορούν να αρχίσουν να πραγματοποιούνται σε μεταβατικό στάδιο στο δεύτερο μισό του 2022. Σε αυτή την περίοδο οι τυπικές διαδικασίες θα τεθούν και θα αναπτυχθούν, και οι κυβερνητικές δομές θα καθίστανται λειτουργικές. Το δεύτερο μισό του 2024 αναμένεται οι τυπικές λειτουργίες να οριστικοποιηθούν, οι κανονισμοί να εφαρμοστούν και οι πρώτες αξιολογήσεις συμμόρφωσης να γίνουν πάνω στα συστήματα.²⁰

Από τα άρθρα της πρότασης γίνεται σαφές ότι όταν αυτή εφαρμοστεί ως κοινοτική οδηγία στο δίκαιο των κρατών μελών, θα αφορά και την λειτουργία αλγορίθμων μηχανικής μάθησης, αφού αναφέρεται σε αυτήν η πρόβλεψη, η ταξινόμηση αλλά και η χρήση συλλογών δεδομένων. Οι αλγόριθμοι μηχανικής μάθησης που εφαρμόζονται στον κατασκευαστικό τεχνικό τομέα, σύμφωνα με το πλαίσιο της πρότασης, θα ανήκουν στην κατηγορία υψηλής επικινδυνότητας ως συστήματα σε τομέα διαχείρισης και λειτουργίας υποδομών ζωτικής σημασίας. Επομένως αναμένεται τα επόμενα χρόνια οι πρακτικές που περιγράφονται και χρησιμοποιούνται στην εργασία αυτή να τεθούν σε περιορισμούς ή απαιτούμενους ελέγχους, λόγω πιθανών συνεπειών στο κοινωνικό σύνολο.

3. Βασικοί Αλγόριθμοι Μηχανικής Μάθησης

3.1 Επιβλεπόμενη Μάθηση

3.1.1 Γενικά

Στην επιβλεπόμενη μάθηση, το σύνολο δεδομένων αποτελείται από επισημασμένα παραδείγματα $\{(x_i, y_i)\}_{i=1}^N$. Κάθε στοιχείο x_i του πληθυσμού N ονομάζεται διάνυσμα χαρακτηριστικών και είναι διάνυσμα του οποίου η κάθε διάσταση περιέχει μία τιμή που περιγράφει με κάποιον τρόπο το παράδειγμα. Η τιμή αυτή ονομάζεται χαρακτηριστικό και συμβολίζεται $x^{(i)}$. Για κάθε παράδειγμα στο σύνολο δεδομένων, το χαρακτηριστικό στη θέση j του διανύσματος χαρακτηριστικών περιέχει το ίδιο είδος πληροφορίας, ακόμα και αν οι τιμές είναι διαφορετικές. Επομένως τα $x_i^{(2)}$ και $x_k^{(2)}$ αναφέρονται στο ίδιο χαρακτηριστικό, με τις ίδιες μονάδες μέτρησης, δύο διαφορετικών παραδειγμάτων x_i και x_k . Η επισήμανση y_i μπορεί να είναι είτε στοιχείο ενός πεπερασμένου συνόλου από κατηγορίες $\{1, 2, \dots\}$, είτε πραγματικός αριθμός, είτε πιο περίπλοκη δομή, όπως διάνυσμα, πίνακας ή γραφική παράσταση. Το σύνολο των κατηγοριών περιέχει κάθε πιθανή κατηγορία στην οποία μπορεί να ανήκει ένα παράδειγμα, σύμφωνα με τον ορισμό του προβλήματος ταξινόμησης. Ο σκοπός της επιβλεπόμενης μάθησης είναι να χρησιμοποιήσει το σύνολο δεδομένων για να δημιουργήσει ένα μοντέλο, το οποίο δέχεται ένα διάνυσμα χαρακτηριστικών x ως είσοδο και βγάζει από αυτό πληροφορίες, οι οποίες επιτρέπουν την εξαγωγή της τιμής επισήμανσης y για το διάνυσμα αυτό.

3.1.2 Γραμμική Παλινδρόμηση

Η γραμμική παλινδρόμηση ως αλγόριθμος μηχανικής μάθησης προσαρμόζει ένα μοντέλο το οποίο είναι γραμμικός συνδυασμός των χαρακτηριστικών του δοθέντος παραδείγματος. Σε κάθε παράδειγμα, x_i είναι το διάνυσμα χαρακτηριστικών και y_i μία τιμή επισήμανσης. Τα χαρακτηριστικά και η επισήμανση ανήκουν στο σύνολο των πραγματικών αριθμών. Το μοντέλο (8) του γραμμικού συνδυασμού των χαρακτηριστικών $x^{(i)}$ ενός διανύσματος χαρακτηριστικών x για κάθε παράδειγμα διατυπώνεται ως

$$f_{w,b}(x) = wx + b \quad (8)$$

όπου w είναι ένα διάνυσμα παραμέτρων με την ίδια διάσταση του διανύσματος χαρακτηριστικών και b ένας πραγματικός αριθμός. Το παραμετρικό μοντέλο θα χρησιμοποιηθεί για την πρόβλεψη μίας άγνωστης τιμής επισήμανσης y για ένα δοθέν διάνυσμα x . Ο αλγόριθμος μηχανικής μάθησης χρησιμοποιείται για την εύρεση των βέλτιστων παραμέτρων w^* , b^* οι οποίες ορίζουν το μοντέλο με την μεγαλύτερη ακρίβεια σε προβλέψεις. Γραφικά το μοντέλο αναπαρίσταται με μία ευθεία καμπύλη η οποία επιλέγεται να βρίσκεται όσο πιο κοντά στα παραδείγματα γίνεται. Στην απλούστερη περίπτωση που

κάθε παράδειγμα έχει μονοδιάστατα διανύσματα χαρακτηριστικών, η ευθεία αποτυπώνεται στις δύο διαστάσεις και σε κάθε x προβλέπεται το y . Ομοίως λειτουργεί η καμπύλη σε περιπτώσεις ανώτερων διαστάσεων διανυσμάτων χαρακτηριστικών.

Για να βρισκεται η ευθεία καμπύλη όσο πιο κοντά στα χαρακτηριστικά γίνεται, η διαδικασία βελτιστοποίησης με την οποία βρίσκονται τα βέλτιστα w^* , b^* είναι η ελαχιστοποίηση του όρου

$$l = \frac{1}{N} \sum_{i=1}^N (f_{w,b}(x_i) - y_i)^2 \quad (9)$$

δηλαδή του μέσου τετραγωνικού σφάλματος, το οποίο υπολογίζεται με αρχικοποιημένες τιμές για w και b . Αυτό είναι και το κριτήριο βελτιστοποίησης με βάση την ελαχιστοποίηση της συνάρτησης απωλειών του αλγορίθμου αυτού. Η λύση της ελαχιστοποίησης ανάγεται στην εύρεση του διανυσματικού διαφορικού τελεστή ∇ (ανάδελτα) της παραπάνω εξίσωσης, ο οποίος στην συνέχεια εξισώνεται με μηδενικό διάνυσμα και επιλύεται το σύστημα για τις τιμές w^* , b^* . Αυτή η μέθοδος λύσης αποτελεί λύση κλειστού προβλήματος, ενώ μπορεί να εφαρμοστεί και η μέθοδος σύγκλισης με ελάττωση της παραγώγου (gradient descent), η οποία αποτελεί λύση ανοιχτού προβλήματος. Στις δύο αυτές μεθόδους, υπολογίζονται οι μερικές παράγωγοι του μέσου τετραγωνικού σφάλματος ως προς τις παραμέτρους w (10) και b (11). Οι μερικές παράγωγοι έχουν την μορφή:

$$\frac{\partial l}{\partial w} = \frac{1}{N} \sum_{i=1}^N 2 x_i (w x_i + b - y_i) \quad (10)$$

$$\frac{\partial l}{\partial b} = \frac{1}{N} \sum_{i=1}^N 2 (w x_i + b - y_i) \quad (11)$$

Στην περίπτωση της κλειστής λύσης, το διάνυσμα $[\partial l / \partial w \quad \partial l / \partial b]$ εξισώνεται με το διάνυσμα $[0 \ 0]$ και λύνεται το σύστημα εξισώσεων $\partial l / \partial w = 0$, $\partial l / \partial b = 0$ ώστε να προκύψουν τα διανύσματα w^* και b^* .

Στην περίπτωση της ανοιχτής λύσης, δηλαδή στην σύγκλιση με ελάττωση της παραγώγου, οι παράμετροι ενημερώνονται με νέες τιμές σε βήματα τα οποία ονομάζονται εποχές. Κάθε εποχή χρησιμοποιεί ολόκληρο ή τμηματικά το σύνολο των δεδομένων εκπαίδευσης, ενώ ο αριθμός εποχών επιλέγεται στην αρχή και αποτελεί υπερπάρμετρο. Στην πρώτη εποχή γίνεται αρχικοποίηση των διανυσμάτων w και b με τιμές ή μηδενικές ή τυχαίες και υπολογίζεται η τιμή των μερικών παραγώγων. Επίσης επιλέγεται ο ρυθμός μάθησης, έστω α , ο οποίος είναι υπερπάρμετρος που ελέγχει πόσο μεγάλες είναι οι μεταβολές στις τιμές των w (12), b (13). Σε κάθε εποχή, οι τιμές μεταβάλλονται ως εξής:

$$w \leftarrow w - \alpha \frac{\partial l}{\partial w} \quad (12)$$

$$b \leftarrow b - \alpha \frac{\partial l}{\partial b} \quad (13)$$

Επειδή οι παράγωγοι είναι ένδειξη αύξησης μίας συνάρτησης, οι παράμετροι μεταβάλλονται κατά την αντίθετη φορά, ώστε το τετραγωνικό σφάλμα να μειώνεται. Στην δεύτερη εποχή υπολογίζονται ξανά οι μερικές παράγωγοι χρησιμοποιώντας τις νέες τιμές των w , b και ενημερώνονται ξανά οι τιμές w , b , συνεχίζοντας την ίδια διαδικασία για τον συνολικό αριθμό εποχών που έχουν οριστεί από τον χρήστη. Ο αλγόριθμος σύγκλισης με ελάττωση της παραγώγου εμφανίζει ευαισθησία στην επιλογή του ρυθμού μάθησης α , δηλαδή εμφανίζει διαφορετικά αποτελέσματα για διαφορετικούς α . Μικρές τιμές του α καθυστερούν την σύγκλιση σε ελάχιστο της συνάρτησης, ενώ υπερβολικά μεγάλες τιμές έχουν τον κίνδυνο να υπερβεί ο αλγόριθμος τα όρια σύγκλισης του κοντινότερου τοπικού ελαχίστου και να συγκλίνει προς άλλο σημείο, μειώνοντας την αποτελεσματικότητα της μεθόδου. Επίσης ο αλγόριθμος είναι πολύ αργός για μεγάλα σύνολα δεδομένων, ωστόσο υπάρχουν προτεινόμενες τροποποιήσεις και βελτιώσεις στην λειτουργία του. Αυτές έχουν ως βάση την υλοποίηση δεσμών του συνόλου εκπαίδευσης, ώστε οι παράμετροι να ενημερώνονται χρησιμοποιώντας ένα τμήμα του συνόλου εκπαίδευσης που περιέχει έναν αριθμό παραδειγμάτων ίσο με την υπερπαράμετρο του μεγέθους δεσμών. Αν το μέγεθος δέσμης ισούται με 1, τότε για κάθε παράδειγμα θα υπολογίζονται μερικές παράγωγοι και θα ενημερώνονται οι παράμετροι, με αποτέλεσμα πολλές μεταβολές των παραμέτρων και μεγάλο απαιτούμενο χρόνο. Από την άλλη, αν το μέγεθος δέσμης είναι υπερβολικά μεγάλο, τότε ο υπολογιστής αναγκάζεται να διατηρεί μεγάλο αριθμό μεταβλητών στην μνήμη λόγω πολύ μεγάλων και πολύπλοκων πράξεων που προκύπτουν από πολλά παραδείγματα. Αυτό σημαίνει και την ύπαρξη μεγάλων απαιτήσεων συστήματος για να λειτουργήσει ο αλγόριθμος.

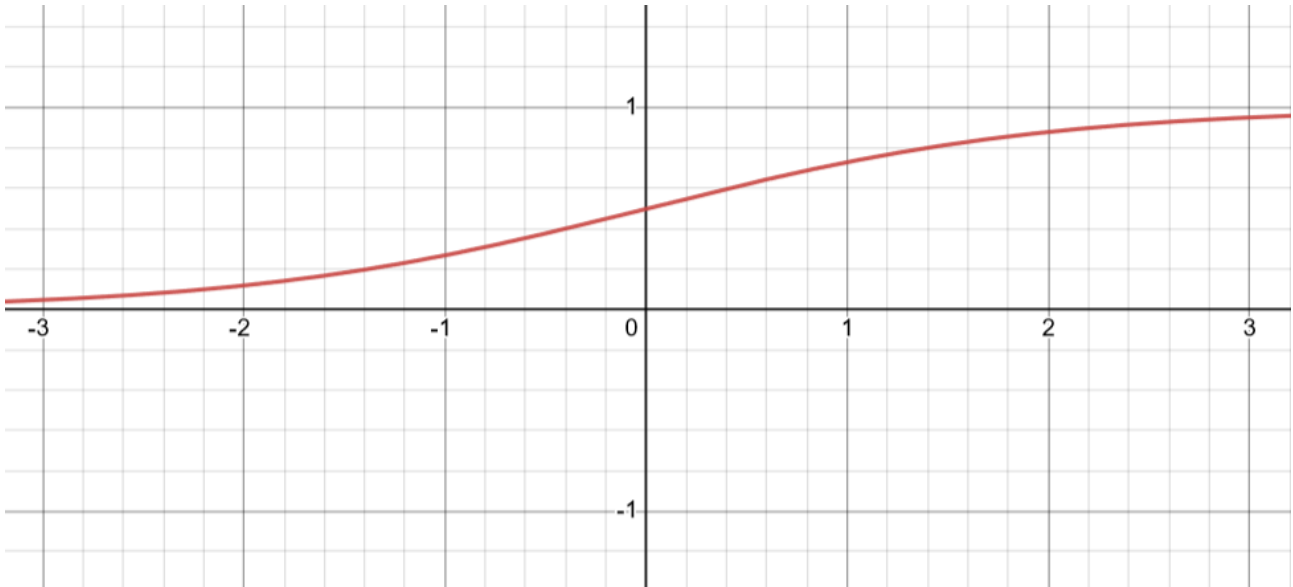
3.1.3 Λογιστική Παλινδρόμηση

Η λογιστική παλινδρόμηση, αντίθετα στο όνομά της, δίνει συνήθως λύση σε πρόβλημα ταξινόμησης και όχι παλινδρόμησης, δηλαδή εύρεσης μίας πραγματικής τιμής. Γίνεται η υπόθεση ότι το πρόβλημα είναι δυαδική ταξινόμηση, άρα η επισήμανση y_i μπορεί να πάρει μόνο δύο πιθανές τιμές. Όμως ο γραμμικός συνδυασμός των χαρακτηριστικών $wx_i + b$ είναι μία συνάρτηση που παίρνει τιμές από το μείον έως το συν άπειρο, όπως είναι γνωστό από την εξίσωση της ευθείας αν δεν είναι σταθερή. Αν οριστούν οι δύο πιθανές τιμές y_i ως 0 ή 1, τότε το μοντέλο χρειάζεται να αποτελείται από συνεχή συνάρτηση φραγμένη σε διάστημα $[0,1]$, ώστε αν το αποτέλεσμα της είναι πλησιέστερα σε μία από τις δύο τιμές να θεωρηθεί ότι ταξινομείται σε αυτήν. Μία συνάρτηση που έχει την ιδιότητα αυτήν, είναι η τυπική λογιστική συνάρτηση (14), γνωστή και ως σιγμοειδής, που απεικονίζεται στην Εικόνα 5 και η οποία ορίζεται ως:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (14)$$

Επομένως το μοντέλο λογιστικής παλινδρόμησης (15) γίνεται η σύνθετη συνάρτηση του γραμμικού μοντέλου όταν σε αυτό εφαρμόζεται η σιγμοειδής καμπύλη:

$$f_{w,b}(x) = \frac{1}{1 + e^{-(wx+b)}} \quad (15)$$



Εικόνα 5: Τυπική Λογιστική (Σιγμοειδής) Καμπύλη

Η ερμηνεία του μοντέλου είναι ότι οι πολύ υψηλές τιμές που προκύπτουν από το γραμμικό μοντέλο πλησιάζουν την τιμή 1, ενώ πολύ αρνητικές τιμές του γραμμικού μοντέλου πλησιάζουν την τιμή 0 στο λογιστικό μοντέλο. Οι ενδιάμεσες τιμές μπορούν να ερμηνευτούν ως η πιθανότητα το παράδειγμα να ανήκει ή στην κατηγορία της τιμής 1 (θετική) ή στην κατηγορία της τιμής 0 (αρνητική), ανάλογα με το αν είναι πάνω ή κάτω από την τιμή 0.5 . Ανάλογα με το πρόβλημα, μπορεί να υλοποιηθεί διαφορετικό σύνορο απόφασης για την προβλεπόμενη κατηγορία.

Για την εύρεση των w^* , b^* στην λογιστική παλινδρόμηση εφαρμόζεται το κριτήριο μεγιστοποίησης των πιθανοτήτων (16) του συνόλου εκπαίδευσης σύμφωνα με το μοντέλο. Σε υψηλό επίπεδο, αυτό συμβαίνει όταν το μοντέλο δίνει μεγάλη πιθανότητα σε ένα παράδειγμα να ανήκει στην κατηγορία που όντως ανήκει σύμφωνα με την επισήμανση του. Έστω ένα παράδειγμα (x_i, y_i) από το σύνολο εκπαίδευσης. Με αρχικές τιμές w, b το μοντέλο επιστρέφει μία τιμή $0 < \pi < 1$. Αν η τιμή της επισήμανσης y_i είναι η θετική κατηγορία 1, τότε είναι επιθυμητό η τιμή π να βρίσκεται όσο πιο κοντά στο 1 γίνεται και ερμηνεύεται ως η πιθανότητα το παράδειγμα να ανήκει όντως στην κατηγορία 1. Αν η τιμή επισήμανσης y_i είναι η αρνητική κατηγορία 0, τότε είναι επιθυμητό η τιμή π να πλησιάζει το 0 και η πιθανότητα το παράδειγμα να ανήκει στην κατηγορία 0 θεωρείται ως $1 - \pi$ από το μοντέλο. Επομένως το κριτήριο βελτιστοποίησης σε αλγεβρική μορφή παρουσιάζεται ως:

$$L_{w,b} = \prod_{i=1}^N \left(f_{w,b}(x_i) \right)^{y_i} \left(1 - f_{w,b}(x_i) \right)^{(1-y_i)} \quad (16)$$

Όπου ο πρώτος όρος του γινομένου υπάρχει μόνο όταν η τιμή επισήμανσης είναι η θετική κατηγορία 1 και ο δεύτερος όρος υπάρχει μόνο όταν η τιμή επισήμανσης είναι η αρνητική κατηγορία 0. Επειδή η πρόβλεψη κάθε παραδείγματος είναι ανεξάρτητη από τα υπόλοιπα παραδείγματα, η συνολική πιθανότητα ισούται με το γινόμενο των πιθανοτήτων κάθε παραδείγματος. Καθώς η μεγιστοποίηση της συνολικής πιθανότητας είναι πρόβλημα ανοιχτού τύπου, συνήθως για την επίλυση του επιστρατεύεται η προαναφερθείσα μέθοδος αριθμητικής βελτιστοποίησης, η σύγκλιση με ελάττωση της παραγώγου.

3.1.4 Νευρωνικά Δίκτυα

Ένα νευρωνικό δίκτυο (Neural Network, NN) είναι μία μαθηματική συνάρτηση από την οποία προκύπτει μία πρόβλεψη y από το διάνυσμα χαρακτηριστικών x . Η συνάρτηση αυτή είναι σύνθετη και μπορεί να αναπαρασταθεί από επίπεδα. Ένα υποτιθέμενο νευρωνικό δίκτυο τριών επιπέδων (17), το οποίο επιστρέφει μία αριθμητική τιμή ως πρόβλεψη, έχει την μορφή:

$$y = f_{NN}(x) = f_3 \left(f_2(f_1(x)) \right) \quad (17)$$

όπου οι f_1 και f_2 (18) είναι διανυσματικές συναρτήσεις της μορφής:

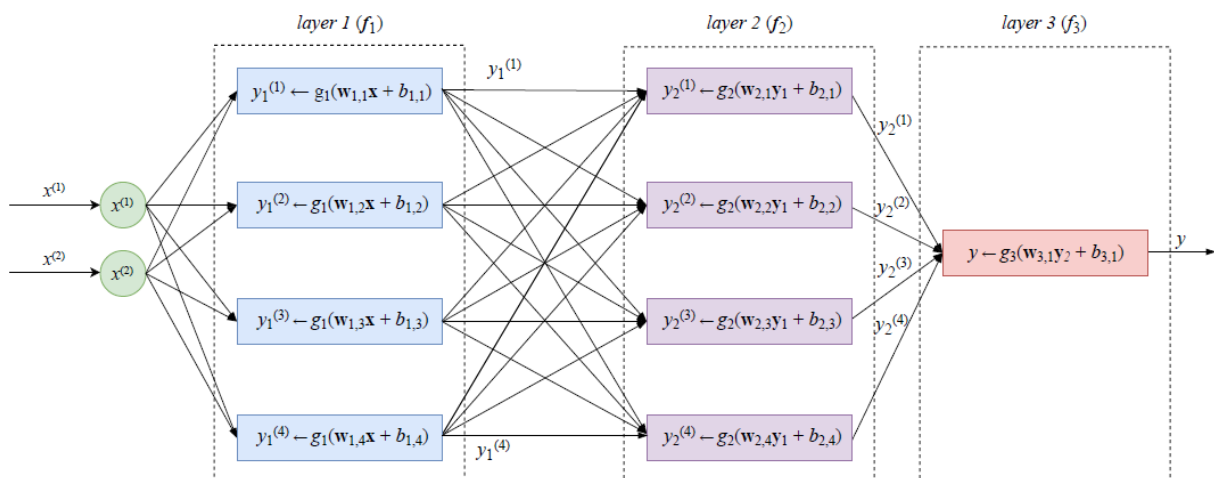
$$f_l(t) = g_l(W_l t + b_l) \quad (18)$$

Το t είναι διάνυσμα με αριθμητικές τιμές. Το l ονομάζεται δείκτης επιπέδου και μπορεί να πάρει τιμή από 1 έως οποιονδήποτε πιθανό αριθμό επιπέδων. Η συνάρτηση g_l που εφαρμόζεται στη γραμμική συνάρτηση ονομάζεται συνάρτηση ενεργοποίησης (activation function) και είναι μία καθορισμένη μη γραμμική συνάρτηση η οποία επιλέγεται πριν ξεκινήσει η διαδικασία μηχανικής μάθησης. Η παράμετρος W_l είναι ένας πίνακας και η παράμετρος b_l ένα διάνυσμα, που για κάθε επίπεδο μαθαίνονται από τον αλγόριθμο εφαρμόζοντας την μέθοδο σύγκλισης με ελάττωση της παραγώγου, με βελτιστοποίηση μίας συγκεκριμένης συνάρτησης απωλειών. Το αποτέλεσμα της συνάρτησης $f_l(t)$ είναι ένα διάνυσμα που περιέχει αριθμητικές τιμές, το οποίο με την σειρά του αποτελεί μεταβλητή για συνάρτηση επόμενου επιπέδου.

Η τελική συνάρτηση f_3 μπορεί να είναι ή αριθμητική ή διανυσματική, ανάλογα με το πρόβλημα της μάθησης.

Μια απλή διαμόρφωση νευρωνικών δικτύων είναι τα εμπροσθοτροφοδοτούμενα νευρωνικά δίκτυα (feed-forward neural networks, FFNN). Έστω ότι το υποτιθέμενο νευρωνικό δίκτυο χρησιμοποιεί μία συγκεκριμένη ακολουθία επιπέδων, μία αρχιτεκτονική δηλαδή, η οποία ονομάζεται Multilayer Perceptron (MLP) και απεικονίζεται σχηματικά στην Εικόνα 6. Το δίκτυο δέχεται ως είσοδο ένα διάνυσμα δύο διαστάσεων, χρησιμοποιεί τρία επίπεδα για να κάνει εκτίμηση και επιστρέφει έναν αριθμό.

Γραφικά το νευρωνικό δίκτυο αναπαρίσταται ως ένας συνδυασμός μονάδων ή κόμβων με δείκτη u , οι οποίοι είναι οργανωμένοι σε επίπεδα με δείκτη l . Οι τιμές εισόδου σε κάποιον κόμβο ενώνονται και σχηματίζουν ένα διάνυσμα εισόδου. Στη συνέχεια εφαρμόζεται ένας γραμμικός μετασχηματισμός στο διάνυσμα αυτό και εφαρμόζεται μία συνάρτηση ενεργοποίησης, από την οποία προκύπτει ένας πραγματικός αριθμός, η τιμή εξόδου για τον κόμβο αυτόν. Η τιμές που προκύπτουν από τις συναρτήσεις των κόμβων ενός επιπέδου σχηματίζουν ένα νέο διάνυσμα, το οποίο χρησιμοποιείται ως στοιχείο εισόδου για τις συναρτήσεις κόμβων του επόμενου επιπέδου. Η συγκεκριμένη αρχιτεκτονική ονομάζεται πλήρως συνδεδεμένη, αν οι κάθε κόμβος του επιπέδου δέχεται ως στοιχεία εισόδου τα στοιχεία εξόδου όλων των κόμβων του προηγούμενου επιπέδου. Τέλος, το τελευταίο επίπεδο, το επίπεδο εκροών, χρησιμοποιεί έναν μόνο κόμβο στην περίπτωση απλής παλινδρόμησης ή ταξινόμησης. Αν η συνάρτηση ενεργοποίησης του κόμβου αυτού είναι γραμμική, το μοντέλο εκτελεί παλινδρόμηση. Αν η συνάρτηση ενεργοποίησης είναι λογιστική, το δίκτυο είναι μοντέλο που εκτελεί δυαδική ταξινόμηση.



Εικόνα 6: Multilayer perceptron*

Οποιαδήποτε συνάρτηση μπορεί να χρησιμοποιηθεί ως συνάρτηση ενεργοποίησης, αρκεί να είναι παραγωγίσιμη ώστε να μπορούν να βρεθούν οι βέλτιστες παράμετροι $W_{l,u}^*$, $b_{l,u}^*$ με ελάττωση της παραγώγου. Ο σκοπός της εισαγωγής συναρτήσεων ενεργοποίησης ανάμεσα στα επίπεδα του νευρωνικού δικτύου είναι η δημιουργία μοντέλου που μπορεί να κάνει προβλέψεις σε δεδομένα που δεν λειτουργούν γραμμικά. Οι συχνότερες συναρτήσεις ενεργοποίησης είναι η ήδη επεξηγημένη λογιστική καμπύλη (σιγμοειδής), η υπερβολική εφαπτομένη \tanh και η συνάρτηση διορθωμένης γραμμικής μονάδας ReLU.

*(Andriy Burkov (2019) The Hundred-Page Machine Learning Book)

Η υπερβολική εφαπτομένη (19) είναι συνάρτηση παρόμοια γραφικά στην λογιστική καμπύλη, ωστόσο λαμβάνει τιμές από -1 έως 1 ασυμπτωτικά. Ορίζεται ως:

$$\tanh(t) = \frac{e^t - e^{-t}}{e^t + e^{-t}} \quad (19)$$

Η συνάρτηση διορθωμένης γραμμικής μονάδας (rectified linear unit, ReLU) (20) επιστρέφει τιμές 0 αν η είσοδος είναι αρνητικός αριθμός, αλλιώς επιστρέφει τον ίδιο αριθμό. Ορίζεται ως:

$$\text{relu}(t) = \begin{cases} 0, & t < 0 \\ t, & t \geq 0 \end{cases} \quad (20)$$

Στην περίπτωση που το πρόβλημα είναι πολυταξική ταξινόμηση, μία τυπική συνάρτηση ενεργοποίησης του τελευταίου επιπέδου είναι η συνάρτηση softmax (21). Η softmax έχει την ιδιότητα να αποδίδει σε κάθε κατηγορία τιμή μεγαλύτερη από 0, ενώ το άθροισμα των τιμών όλων των κατηγοριών ισούται με 1. Άρα είναι μία άμεση εκτίμηση των πιθανοτήτων να ανήκει το παράδειγμα σε καθεμία κατηγορία. Η softmax ορίζεται ως:

$$\sigma(t) = [\sigma^{(1)}, \sigma^{(2)}, \dots, \sigma^{(D)}], \quad (21)$$

$$\sigma^{(j)} = \frac{e^{t^{(j)}}}{\sum_{k=1}^D e^{t^{(k)}}}$$

Για την ενημέρωση των παραμέτρων χρησιμοποιείται ο αλγόριθμος ανάστροφης αναδιάδοσης, ο οποίος χρησιμοποιεί τον κανόνα της αλυσίδας για τον υπολογισμό των παραγώγων της συνάρτησης απώλειας ως προς τις παραμέτρους σε κάθε κόμβο, ώστε να εφαρμοστεί στη συνέχεια η σύγκλιση με ελάττωση της παραγώγου.

Όταν ένα νευρωνικό δίκτυο έχει παραπάνω από 2 επίπεδα, εξαιρουμένων των επιπέδων εισόδου και εξόδου, η διαδικασία εκπαίδευσης ονομάζεται βαθιά μάθηση. Τα επίπεδα μεταξύ των δύο αυτών επιπέδων ονομάζονται και κρυφά επίπεδα. Η βαθιά μάθηση αντιμετωπίζει το πρόβλημα του εκρηκτικού ρυθμού αύξησης ή μείωσης των παραγώγων. Επειδή οι παράγωγοι υπολογίζονται με τον κανόνα της αλυσίδας, πολλαπλασιάζονται συνεχόμενα μερικές παράγωγοι για να βρεθεί η μερική παράγωγος των αρχικών επιπέδων. Ως αποτέλεσμα, μπορούν να επιτευχθούν παράγωγοι κοντά στο 0 ή υπερβολικά μεγάλες σε τιμή και άρα η ανάλογη ενημέρωση των παραμέτρων να σταματήσει ή να βγει εκτός ορίων. Τα προβλήματα αυτά έχουν σε μεγάλο βαθμό επιλυθεί με σύγχρονες βελτιώσεις όπως η ομαλοποίηση, η εφαρμογή της ReLU αλλά και άλλων τεχνικών. Επομένως, ο όρος βαθιά

μάθηση αναφέρεται πλέον σε εκπαίδευση νευρωνικών δικτύων με τις συγκεκριμένες αλγοριθμικές και μαθηματικές μεθόδους, ανεξαρτήτως του αριθμού κρυφών επιπέδων.

3.1.5 Συνελικτικά νευρωνικά δίκτυα

Οι παράμετροι ενός νευρωνικού δικτύου αυξάνονται με μεγάλο βαθμό όσο αυξάνονται τα επίπεδα του δικτύου και τα διανύσματα χαρακτηριστικών που εισέρχονται. Η βελτιστοποίηση μεγάλων μοντέλων είναι υπολογιστικά έντονη διαδικασία. Όταν σαν παράδειγμα εκπαίδευσης χρησιμοποιούνται εικόνες, κάθε εικονοστοιχείο (pixel) της εικόνας είναι χαρακτηριστικό. Επομένως τα διανύσματα εισόδου έχουν μεγάλες διαστάσεις και το πρόβλημα βελτιστοποίησης κινδυνεύει να καταστεί μη επιλύσιμο. Για να αποφευχθεί η κατάσταση αυτή, αναπτύχθηκε μία ειδική κατηγορία εμπροσθοτροφοδοτούμενων νευρωνικών δικτύων, τα συνελικτικά νευρωνικά δίκτυα (Convolutional Neural Networks, CNN). Λειτουργούν μειώνοντας σημαντικά τον αριθμό παραμέτρων σε βαθιά νευρωνικά δίκτυα, χωρίς να χαμηλώνουν πολύ την ποιότητα του μοντέλου. Τα συνελικτικά δίκτυα έχουν ευρεία εφαρμογή σε επεξεργασία εικόνων.

Στις εικόνες, εικονοστοιχεία που βρίσκονται κοντά το ένα σε άλλο περιέχουν τις ίδιες τιμές, όταν απεικονίζουν την ίδια πληροφορία. Εξαιρέση είναι η παρουσία ακμών στην εικόνα, όπου δύο διαφορετικά αντικείμενα βρίσκονται σε γειτονικά εικονοστοιχεία που έχουν διαφορετικές τιμές σε αυτή την περίπτωση. Το ζητούμενο στα νευρωνικά δίκτυα που δουλεύουν πάνω σε εικόνες είναι η εκπαίδευση τους ώστε να αναγνωρίζουν ομοιογενείς περιοχές αλλά και ακμές και να τις χρησιμοποιήσουν για πρόβλεψη του απεικονιζόμενου αντικειμένου.

Με την υπόθεση ότι τα σημαντικότερα στοιχεία για την αναγνώριση της εικόνας είναι τοπικά, γίνεται κατάτμηση της εικόνας σε μικρά τετράγωνα τμήματα χρησιμοποιώντας μία προσέγγιση κινούμενου παραθύρου. Στη συνέχεια γίνεται εκπαίδευση πολλών μικρών μοντέλων παλινδρόμησης που το καθένα δέχεται ως εισροή ένα τετραγωνικό τμήμα, με σκοπό να μάθει να αναγνωρίζει ένα χαρακτηριστικό μοτίβο μέσα στο τμήμα αυτό. Το μοτίβο μπορεί να είναι συγκεκριμένη ομοιογενής περιοχή ή συγκεκριμένη ακμή, το σύνολο των οποίων δημιουργεί ένα ή πολλά αναγνωρίσιμα αντικείμενα στην ενιαία εικόνα.

Αναλυτικότερα, κάθε μικρό μοντέλο παλινδρόμησης έχει ένα μόνο επίπεδο εκτός από το επίπεδο εισόδου. Για να αναγνωρίσει κάποιο μοτίβο, το μοντέλο μαθαίνει τις παραμέτρους μέσα σε έναν πίνακα F (γνωστός και ως φίλτρο), με μέγεθος $p \cdot p$, όπου p μια υπερπαραμέτρος διαστάσεων του τετραγωνικού τμήματος σε εικονοστοιχεία (pixel). Έστω ότι το συνελικτικό νευρωνικό δίκτυο εφαρμόζεται σε μία καθαρά ασπρόμαυρη εικόνα, στην οποία τα εικονοστοιχεία είναι μόνο λευκά ή μαύρα και αντιπροσωπεύονται αντίστοιχα από τις τιμές 1 και 0. Επίσης έστω ότι κάθε τετραγωνικό τμήμα έχει μέγεθος 3×3 εικονοστοιχεία ($p = 3$). Ένα τυχαίο τμήμα της εικόνας μπορεί να έχει την υποθετική μορφή:

$$P = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

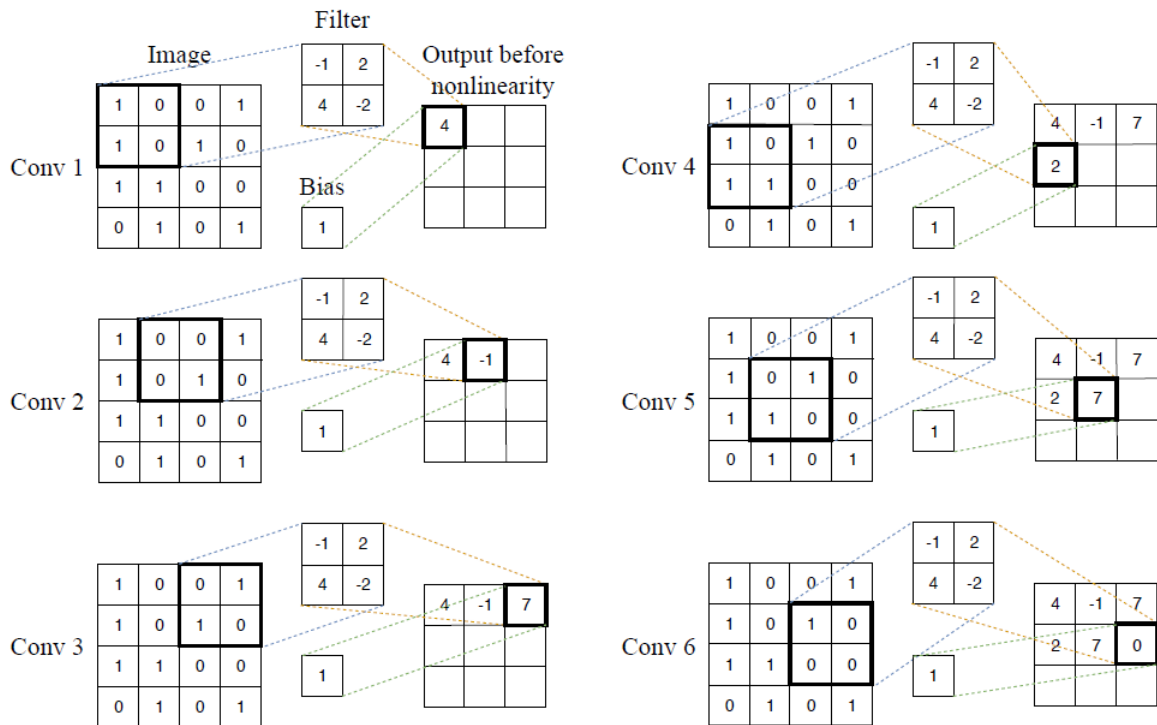
Το τμήμα απεικονίζει ένα μοτίβο με τη μορφή T . Αν το μοτίβο αυτό είναι ενδεικτικό χαρακτηριστικό του αντικειμένου που πρέπει να αναγνωρίσει ο αλγόριθμος, το μικρό μοντέλο παλινδρόμησης που αναγνωρίζει παρόμοια μοτίβα και μόνο θα πρέπει να μάθει το φίλτρο F του οποίου οι παράμετροι αντιστοιχούν στα στοιχεία του παραθύρου P . Αυτό συμβαίνει όταν οι παράμετροι του F έχουν θετικούς αριθμούς σε θέσεις που αντιστοιχούν στα στοιχεία 1 στο P , και τιμές κοντά στο μηδέν σε θέσεις που αντιστοιχούν στα στοιχεία 0 του P . Για τον υπολογισμό της ομοιότητας των δύο πινάκων, χρησιμοποιείται η συνέλιξη των P και F , από την οποία προκύπτει μία τιμή που είναι υψηλότερη όσο πιο όμοιοι είναι. Για την οπτικοποίηση της συνέλιξης, έστω πως το φίλτρο F έχει αρχικοποιηθεί με τις παρακάτω τιμές:

$$F = \begin{bmatrix} 2 & 2 & 3 \\ 0 & 4 & 1 \\ 0 & 3 & 0 \end{bmatrix}$$

Η συνέλιξη πινάκων (22) ορίζεται μόνο για πίνακες με τον ίδιο αριθμό γραμμών και στηλών και στην περίπτωση των πινάκων P και F υπολογίζεται ως εξής:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \text{ συνέλιξη } \begin{bmatrix} 2 & 2 & 3 \\ 0 & 4 & 1 \\ 0 & 3 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 \cdot 2 & 1 \cdot 2 & 1 \cdot 3 \\ 0 \cdot 0 & 1 \cdot 4 & 0 \cdot 1 \\ 0 \cdot 0 & 1 \cdot 3 & 0 \cdot 0 \end{bmatrix} \rightarrow \begin{matrix} \text{άθροισμα} \\ \text{στοιχείων} \end{matrix} \rightarrow 14 \quad (22)$$

Σε κάθε φίλτρο F υπάρχει και μία σχετιζόμενη παράμετρος μεροληψίας b , η οποία προστίθεται στο αποτέλεσμα της συνέλιξης πριν εφαρμοστεί σε αυτό η μη γραμμικότητα μέσω της συνάρτησης ενεργοποίησης. Ένα επίπεδο συνελκτικού νευρωνικού δικτύου αποτελείται από πολλαπλά φίλτρα συνέλιξης, το καθένα από τα οποία έχει την δική του παράμετρο b και λειτουργεί με τρόπο αντίστοιχο των κόμβων σε ένα απλό εμπροσθοτροφοδοτούμενο δίκτυο. Επομένως ο πίνακας F και η παράμετρος b βελτιστοποιούνται χρησιμοποιώντας σύγκλιση με ελάττωση της παραγώγου μέσω του αλγορίθμου ανάστροφης αναδιάδοσης. Συνήθως εφαρμόζεται η ReLU ως συνάρτηση ενεργοποίησης στο άθροισμα της συνέλιξης και της παραμέτρου μεροληψίας b σε όλα τα κρυφά επίπεδα, ενώ η συνάρτηση ενεργοποίησης του επιπέδου εκροών εξαρτάται από την εργασία που καλείται να εκτελέσει το μοντέλο.



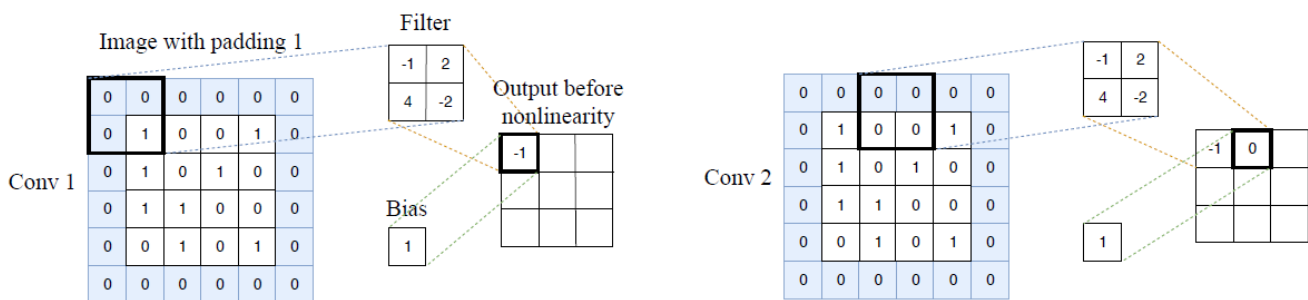
Εικόνα 7: Διαδοχικές συνελίξεις φίλτρου πάνω σε εικόνα*

Κάθε φίλτρο του πρώτου επιπέδου περνάει από όλη την εικόνα που εισάγεται στο δίκτυο, από αριστερά προς τα δεξιά και από πάνω προς τα κάτω, εκτελώντας μία συνέλιξη σε κάθε βήμα. Η εκροή ενός φίλτρου που έχει εφαρμοστεί σε όλη την εικόνα τμηματικά είναι ένας χάρτης χαρακτηριστικών (feature map), που είναι ένας πίνακας με στοιχεία τα αποτελέσματα κάθε συνέλιξης και δείχνει κατά πόσο και σε ποια σημεία αναγνωρίζεται το μοτίβο του φίλτρου στην εικόνα αυτήν. Επειδή κάθε επίπεδο l έχει αριθμό n φίλτρων, η εκροή του επιπέδου αποτελείται από n πίνακες - χάρτες χαρακτηριστικών.

Στην περίπτωση που το συνελικτικό νευρωνικό δίκτυο έχει ένα επίπεδο συνέλιξης ακριβώς μετά από άλλο επίπεδο συνέλιξης, το δεύτερο επίπεδο $l+1$ δέχεται την εκροή του προηγούμενου επιπέδου l , δηλαδή τους χάρτες χαρακτηριστικών, σαν μία συλλογή από n εικόνες. Η συλλογή αυτή ονομάζεται τόμος, ενώ το μέγεθος της συλλογής σε εικόνες ονομάζεται βάθος του τόμου. Κάθε φίλτρο του επιπέδου $l+1$ περνά από όλον τον τόμο, με την συνέλιξη πάνω σε ένα παράθυρο του τόμου να ισούται με το άθροισμα των συνελίξεων πάνω στα αντίστοιχα παράθυρα στις εικόνες του τόμου. Επειδή συχνά οι εικόνες δεν είναι μονόχρωμες, όπως είναι για παράδειγμα η ασπρόμαυρη εικόνα που χρησιμοποιήθηκε για επεξήγηση, το πρώτο επίπεδο του δικτύου δέχεται και αυτό έναν τόμο ως είσοδο. Ο τόμος αποτελείται από τα τρία χρωματικά κανάλια κόκκινο, πράσινο, μπλε (RGB) στην περίπτωση έγχρωμης εικόνας, ενώ σε άλλες περιπτώσεις μπορεί να δέχεται τόμους μεγαλύτερου βάθους όπως στην περίπτωση πολυφασματικών εικόνων. Κάθε κανάλι παριστάνεται ως μία μονοχρωματική εικόνα.

* (Andriy Burkov (2019) *The Hundred-Page Machine*)

Δύο σημαντικές ιδιότητες της συνέλιξης σε δίκτυα είναι το βήμα και το γέμισμα του περιθωρίου, τα οποία προσδιορίζονται αριθμητικά ως υπερπαραμέτροι. Το βήμα ορίζει την απόσταση σε εικονοστοιχεία που θα έχουν τα παράθυρα της εικόνας που συνελίσσονται με το φίλτρο για να φτιαχτεί ο χάρτης χαρακτηριστικών. Λογικό επακόλουθο της αύξησης του βήματος, όταν το μέγεθος του φίλτρου παραμένει ίδιο, είναι ο εξαγόμενος πίνακας του χάρτη χαρακτηριστικών να είναι μικρότερος όσο μεγαλύτερο είναι το βήμα. Το γέμισμα του περιθωρίου είναι η διαδικασία με την οποία περιβάλλεται η εικόνα ή οι εικόνες ενός τόμου με επιπρόσθετα εικονοστοιχεία, ώστε το αποτέλεσμα της συνέλιξης να είναι ένας μεγαλύτερος πίνακας. Τα εικονοστοιχεία που προστίθενται περιέχουν συνήθως μηδενικές τιμές. Το γέμισμα του περιθωρίου επιτρέπει επίσης την καλύτερη αναγνώριση μοτίβων όταν αυτά βρίσκονται ακριβώς στα όρια μίας εικόνας, ειδικά όταν τα φίλτρα είναι σχετικά μεγάλα σε διαστάσεις.



Εικόνα 8: Συνέλιξη με βήμα 2 και γέμισμα περιθωρίου 1

Τέλος, μία τεχνική που συχνά χρησιμοποιείται μέσα σε συνελκτικά νευρωνικά δίκτυα είναι η παρεμβολή επιπέδων συγκέντρωσης. Τα επίπεδα συγκέντρωσης (pooling) λειτουργούν με τρόπο παρόμοιο με την συνέλιξη, ορίζοντας ένα παράθυρο το οποίο σταδιακά περνάει από τμήματα της εικόνας ή του χάρτη χαρακτηριστικών. Η διαφορά με την συνέλιξη είναι ότι αντί να υπάρχει ένας αντίστοιχος πίνακας φίλτρου, το παράθυρο της συγκέντρωσης εφαρμόζει έναν σταθερό τελεστή στα εικονοστοιχεία που περιέχει, συνήθως αριθμητικό μέγιστο (max pooling) ή αριθμητικό μέσο (average pooling). Στην πρώτη περίπτωση δημιουργείται ένα εικονοστοιχείο στον πίνακα εξόδου που έχει τιμή ίση με την μεγαλύτερη τιμή που εμφανίζεται στα εικονοστοιχεία της εικόνας εισόδου που επεξεργάζεται το παράθυρο. Στην δεύτερη δημιουργείται εικονοστοιχείο με τιμή ίση με τον μέσο όρο των τιμών των εικονοστοιχείων της εικόνας. Τα επίπεδα συγκέντρωσης έχουν τις υπερπαραμέτρους μεγέθους του παραθύρου και βήματος, και κατά κανόνα βρίσκονται μετά από επίπεδα συνέλιξης στα δίκτυα, όπου δέχονται ως είσοδο τις εκροές των συνελίξεων. Όταν δέχονται ως είσοδο τόμους, κάθε πίνακας του τόμου επεξεργάζεται ξεχωριστά από το επίπεδο, και η εκροή του επιπέδου συγκέντρωσης είναι ένας άλλος τόμος με το ίδιο βάθος του αρχικού τόμου.

Τυπικές τιμές των υπερπαραμέτρων του επιπέδου συγκέντρωσης είναι 2 ή 3 για το μέγεθος του κινούμενου παραθύρου και 2 για το βήμα του. Η συγκέντρωση με βάση τον αριθμητικό μέγιστο είναι συχνότερη λόγω καλύτερων επιδόσεων στα συνελκτικά νευρωνικά δίκτυα,

* (Andriy Burkov (2019) *The Hundred-Page Machine Learning Book*)

ωστόσο και η συγκέντρωση με βάση τον αριθμητικό μέσο μπορεί να χρησιμοποιηθεί για ορισμένες λειτουργίες οι οποίες περιγράφονται σε επόμενο κεφάλαιο. Η χρήση επιπέδων συγκέντρωσης έχει ως κύριο σκοπό την μείωση των διαστάσεων των πινάκων που διαχειρίζεται το δίκτυο και άρα των παραμέτρων που πρέπει να μαθευτούν, βελτιώνοντας την ταχύτητα εκπαίδευσης. Επίσης συμβάλλει στην ακρίβεια του μοντέλου σε προβλέψεις.

3.1.6 Μοντέλα εντοπισμού αντικειμένων

Στην περίπτωση που απαιτείται για κάποια εφαρμογή ο ακριβής εντοπισμός ενός στοιχείου πάνω σε μία εικόνα, μπορούν να αξιοποιηθούν τα χαρακτηριστικά μοτίβα που έχει προσδιορίσει ένα συνελκτικό δίκτυο σε αυτήν. Με βάση την ενεργοποίηση μοτίβων, όπως αυτή περιγράφεται σε χάρτες χαρακτηριστικών του συνελκτικού νευρωνικού δικτύου, μπορεί σε μία υποπεριοχή της εικόνας να οριστεί ένα πλαίσιο που οριοθετεί ένα επιθυμητό στοιχείο, το οποίο και ταξινομείται από τον αλγόριθμο σε κάποια κατηγορία. Επομένως αντί να ταξινομηθεί ολόκληρη η εικόνα, ταξινομείται μόνο τμήμα της, ενώ άλλο τμήμα μπορεί να ταξινομηθεί σε άλλη κατηγορία. Η συνολική διεργασία εντοπισμού και ταξινόμησης αποτελεί μοντέλο εντοπισμού αντικειμένων.

Η βασική διαδικασία με την οποία εντοπίζεται το πλαίσιο ενός στοιχείου εντός της εικόνας, ονομάζεται παλινδρόμηση πλαισίου οριοθέτησης (bounding box regression, Εικόνα 9) και υφιστάται σε νευρωνικά δίκτυα ως τελικό επίπεδο ταξινόμητη. Οι παράμετροι που βελτιστοποιούνται με εκπαίδευση καθορίζουν την θέση του πλαισίου, και ανάλογα τον αλγόριθμο είναι ή οι συντεταγμένες x , y της άνω αριστερής κορυφής και της κάτω δεξιάς κορυφής, ή οι συντεταγμένες x , y του κέντρου του πλαισίου καθώς και το ύψος του και το πλάτος του, σε εικονοστοιχεία.

Έστω η περίπτωση που το πλαίσιο ορίζεται από τις συντεταγμένες του κέντρου του, το ύψος και το πλάτος. Για τις απαιτήσεις της εκπαίδευσης, τα δεδομένα εισόδου είναι ένα σύνολο N παραδειγμάτων $\{(P_i, G_i)\}_{i=1}^N$, όπου $P = (P_x, P_y, P_w, P_h)$ το διάνυσμα που ορίζει το προτεινόμενο από τον αλγόριθμο πλαίσιο και $G = (G_x, G_y, G_w, G_h)$ το διάνυσμα που ορίζει το πλαίσιο βασικής πραγματικότητας, δηλαδή ένα πλαίσιο που έχει επισημανθεί χειροκίνητα από τον χρήστη του αλγορίθμου εκπαίδευσης και απεικονίζει όντως το επιθυμητό στοιχείο στο κάθε παράδειγμα. Ο στόχος του αλγορίθμου είναι να μάθει έναν μετασχηματισμό ο οποίος συνδέει το προτεινόμενο πλαίσιο P σε ένα πλαίσιο βασικής πραγματικότητας G .

Ο μετασχηματισμός παραμετροποιείται ως προς τέσσερις συναρτήσεις $d_x(P)$, $d_y(P)$, $d_w(P)$, $d_h(P)$. Οι πρώτες δύο ορίζουν έναν ανεξάρτητο από κλίμακα μετασχηματισμό του κέντρου του προτεινόμενου πλαισίου P , ενώ οι τελευταίες δύο ορίζουν εκθετικούς μετασχηματισμούς του ύψους και του πλάτους. Όταν μαθευτούν από την εκπαίδευση αυτές οι συναρτήσεις, ένα πλαίσιο πρότασης P που εισάγεται μπορεί να μετασχηματιστεί σε πρόβλεψη πλαισίου βασικής αλήθειας \tilde{G} , το οποίο είναι το ζητούμενο του προβλήματος εντοπισμού, μέσω του εξής μετασχηματισμού (23):

$$\begin{aligned}
\hat{G}_x &= P_w d_x(P) + P_x \\
\hat{G}_y &= P_h d_y(P) + P_y \\
\hat{G}_w &= P_w e^{d_w(P)} \\
\hat{G}_h &= P_h e^{d_h(P)}
\end{aligned} \tag{23}$$

Κάθε συνάρτηση $d^*(P)$ (όπου * ένα από τα x, y, w, h) είναι γραμμική συνάρτηση επιπέδων συγκέντρωσης των χαρτών χαρακτηριστικών της πρότασης P , τα οποία συμβολίζονται με $\Phi_5(P)$ και συσχετίζονται προφανώς με τα δεδομένα της εικόνας. Επομένως λόγω γραμμικότητας ισχύει $d^*(P) = w^{*T} \Phi_5(P)$, όπου w^* διάνυσμα παραμέτρων που μπορούν να μαθευτούν. Οι παράμετροι βρίσκονται με βελτιστοποίηση του κριτηρίου κανονικοποιημένων ελαχίστων τετραγώνων (24):

$$w_* = \underset{\hat{w}_*}{\operatorname{argmin}} \sum_{i=1}^N \left(t_*^i - \hat{w}_*^T \Phi_5(P^i) \right)^2 + \lambda \|\hat{w}_*\|^2 \tag{24}$$

όπου $\lambda = 1000$ με βάση πειραματικό σύνολο επικύρωσης και t^* (25) οι τιμές που προκύπτουν από την παλινδρόμηση, με

$$\begin{aligned}
t_x &= (G_x - P_x)/P_w \\
t_y &= (G_y - P_y)/P_h \\
t_w &= \log(G_w/P_w) \\
t_h &= \log(G_h/P_h)
\end{aligned} \tag{25}$$

Ένα θέμα που προκύπτει από τον παραπάνω αλγόριθμο είναι ότι αν το P απέχει μεγάλη απόσταση από τα πλαίσια βασικής πραγματικότητας G , ο μετασχηματισμός του σε πλαίσιο βασικής πραγματικότητας δεν βγάζει νόημα και το πρόβλημα μάθησης δεν επιδέχεται λύση. Επομένως, αφού το P μπορεί να μάθει μόνο από κοντινό του G , απαιτείται η επιλογή ζεύγους (P, G) με υψηλό λόγο τομής προς ένωσης, που δείχνει την επικάλυψη, ενώ τα P που δεν μπορούν να ανατεθούν σε G απορρίπτονται.²⁷

Η εξαγωγή των προτεινόμενων πλαισίων P μπορεί να γίνει με διάφορες μεθόδους, χρησιμοποιώντας κάποιον αλγόριθμο πρότασης περιοχών πριν την συνέλιξη. Ο αργότερος είναι η επιλογή προτεινόμενων πλαισίων με την μέθοδο του κινούμενου παραθύρου, όπου σε κάθε βήμα ελέγχεται αν το πλαίσιο μπορεί να μετασχηματιστεί σε πλαίσιο αντικειμένου και

αν το αντικείμενο μπορεί να ταξινομηθεί μέσω συνέλιξης. Στο μοντέλο Regional Convolutional Neural Network (R-CNN) χρησιμοποιείται η επιλεκτική αναζήτηση 2000 πλαισίων, όπου οι προτεινόμενες περιοχές πλαισίων βρίσκονται βάση της έντασης των εικονοστοιχείων τους, ενώ στο μοντέλο You only look once (YOLO) εφαρμόζεται ένα πλέγμα σε ολόκληρη την εικόνα και επεξεργάζονται πλαίσια σε κάθε παράθυρο του πλέγματος.



Εικόνα 9: Παλινδρόμηση πλαισίου οριοθέτησης*

3.1.7 Αρχιτεκτονικές νευρωνικών δικτύων

Η ύπαρξη πολλών επιπέδων και πολλών κόμβων μέσα στα επίπεδα αυτά προσφέρει μεγάλη ευελιξία στην μορφή που μπορεί να έχει ένα νευρωνικό δίκτυο. Ανάλογα με τις διαστάσεις του διανύσματος χαρακτηριστικών αλλά και την επίλυση που δέχεται το πρόβλημα για το οποίο υλοποιείται ένα μοντέλο, μπορεί να χρησιμοποιηθεί διαφορετικό πλήθος επιπέδων, με διαφορετικό πλήθος και λειτουργίες κόμβων και διαφορετικές συναρτήσεις ενεργοποίησης στους κόμβους. Ακόμα, όλοι οι κόμβοι ενός επιπέδου δεν είναι απαραίτητο να ενώνονται με τους κόμβους του επόμενου, δηλαδή να είναι πλήρως συνδεδεμένα τα επίπεδα. Στην περίπτωση συνελκτικών νευρωνικών δικτύων, πληθαίνουν ακόμα περισσότερο οι πιθανές διαφορετικές διατάξεις λόγω παρουσίας επιπέδων συνέλιξης και συγκέντρωσης, τα οποία μπορεί να συνδυάζονται με επίπεδα απλών εμπροσθοτροφοδοτούμενων δικτύων. Η ακριβής διαμόρφωση ενός νευρωνικού δικτύου, που ορίζει τον αριθμό και την ακολουθία επιπέδων, κόμβων, συναρτήσεων ενεργοποίησης αλλά και υπερπαραμέτρων των επιπέδων (βήμα,

* (<https://towardsdatascience.com/yolo2-walkthrough-with-examples-e40452ca265f>)

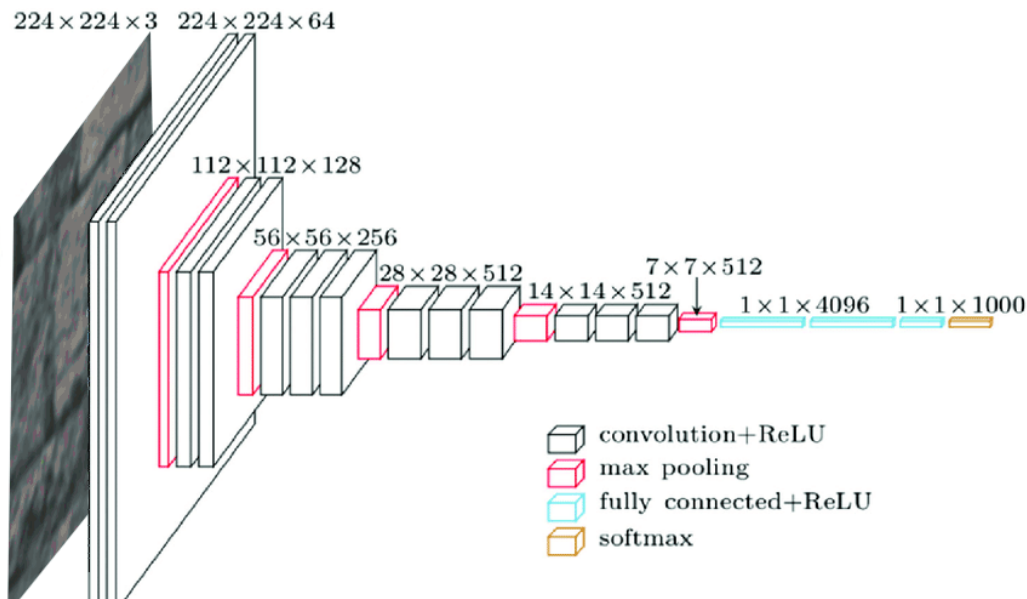
περιθώριο, μέγεθος παραθύρων) ονομάζεται αρχιτεκτονική του δικτύου. Αναμένεται σε γενικές γραμμές δύο μοντέλα που εφαρμόζουν την ίδια αρχιτεκτονική νευρωνικών δικτύων να έχουν παρόμοια αποτελέσματα όταν εκπαιδεύονται στα ίδια δεδομένα.

Οι σύγχρονες αρχιτεκτονικές νευρωνικών δικτύων είναι αποτελέσματα της συνεργασίας επιστημόνων από πολλά εργαστήρια ή εταιρείες, και συνήθως είναι υπερβολικά περίπλοκα για να σχεδιαστούν για ατομική χρήση και απαιτούν πολλή υπολογιστική ισχύ και πολλά παραδείγματα για να εκπαιδευτούν από την αρχή. Μεγάλο τμήμα της έρευνας πάνω σε νευρωνικά δίκτυα είναι η εφεύρεση νέων αρχιτεκτονικών για υπάρχοντα προβλήματα, με νέες αρχιτεκτονικές να προτείνονται και να δημοσιεύονται συνεχώς¹⁵. Η ανάπτυξη νέων συνελκτικών νευρωνικών δικτύων ακολουθεί μία τάση χρήσης όλο και περισσότερων επιπέδων, δηλαδή βαθιών αρχιτεκτονικών¹⁷.

Για το πρόβλημα αναγνώρισης και κατηγοριοποίησης εικόνων, μία αρχιτεκτονική που χρησιμοποιείται ευρέως είναι η προτεινόμενη από τους Simonyan και Zisserman²² αρχιτεκτονική του Visual Geometry Group, VGGNet, που αναπτύχθηκε στο πανεπιστήμιο της Οξφόρδης. Χρησιμοποιούνται δύο μοντέλα ανοιχτού κώδικα που σχεδιάστηκαν, ονομαζόμενα VGG-16 και VGG-19. Το μοντέλο VGG-16 αποτελείται από 16 επίπεδα, από τα οποία 13 είναι συνελκτικά με μέγεθος φίλτρου 3x3 και 3 είναι πλήρως συνδεδεμένα. Τα 13 συνελκτικά επίπεδα χωρίζονται σε 5 υποομάδες και στο τέλος της καθεμίας από αυτές τοποθετείται ένα επίπεδο συγκέντρωσης αριθμητικού μεγίστου. Η είσοδος στο μοντέλο είναι μία εικόνα με 224 επί 224 εικονοστοιχεία σε καθένα από τα τρία κανάλια RGB. Στο τέλος των 13 επιπέδων συνέλιξης, εξέρχεται ένας τόμος 512 χαρτών χαρακτηριστικών 7 επί 7 εικονοστοιχείων, από τον οποίο σχηματίζεται ένα ενιαίο διάνυσμα χαρακτηριστικών μέσω της διαδικασίας επιπλάτυνσης (flattening). Το διάνυσμα στη συνέχεια εισέρχεται στα 3 πλήρως συνδεδεμένα επίπεδα, από τα οποία προκύπτει ένα τελικό διάνυσμα που μπορεί να χρησιμοποιηθεί από ένα επίπεδο softmax για ταξινόμηση (Εικόνες 10 και 11).



Εικόνα 10: Επίπεδα συνελκτικού νευρωνικού δικτύου με αρχιτεκτονική VGG-16



Εικόνα 11: Σχηματική αναπαράσταση συνελκτικού νευρωνικού δικτύου VGG-16*

Το μοντέλο VGG19 αποτελείται από 19 επίπεδα, από τα οποία τα 16 είναι συνελκτικά με μέγεθος φίλτρου 3×3 εικονοστοιχεία, βήμα και γέμισμα περιθωρίου μεγέθους 1 εικονοστοιχείο, και τα 3 είναι πλήρως συνδεδεμένα. Ακόμα, μετά από καθεμία από τις 5 υποομάδες συνελκτικών επιπέδων, η πράξη συγκέντρωσης αριθμητικού μεγίστου εκτελείται με παράθυρο 2×2 και βήμα 2 εικονοστοιχεία. Στο τέλος χρησιμοποιούνται ομοίως τα πλήρως συνδεδεμένα επίπεδα με το επίπεδο softmax. Η Αρχιτεκτονική VGGNet ενίσχυσε την πεποίθηση ότι τα συνελκτικά νευρωνικά δίκτυα πρέπει να έχουν βαθιά διαμόρφωση επιπέδων, ώστε τα οπτικά δεδομένα να ερμηνεύονται ιεραρχικά, από απλά μοτίβα τα οποία συνδυάζονται σε πιο σύνθετα μοτίβα, τα οποία με την σειρά τους συνδυάζονται σε ακόμα πιο σύνθετα.

Άλλα ανεπτυγμένα μοντέλα συνελκτικών νευρωνικών δικτύων για διάφορες εφαρμογές είναι τα Resnet-50 με 50 συνελκτικά επίπεδα, Inception V3 με 48 συνελκτικά επίπεδα²³, Googlenet με 22 επίπεδα και τα παλαιότερα LeNet και AlexNet. Το τελευταίο χρησιμοποιεί διαφορετικού μεγέθους φίλτρα ανά επίπεδο, με το πρώτο επίπεδο να έχει φίλτρο συνέλιξης 11×11 .¹⁷

3.1.8 Μεταφορά Μάθησης

Η μεταφορά μάθησης (transfer learning) είναι ένα πλεονέκτημα των νευρωνικών δικτύων έναντι των ρηχών μοντέλων. Σε αυτή, επιλέγεται ένα υπάρχον μοντέλο το οποίο έχει εκπαιδευτεί πάνω σε κάποιο σύνολο δεδομένων και προσαρμόζεται για να κάνει προβλέψεις σε παραδείγματα ενός άλλου, διαφορετικού συνόλου δεδομένων. Το δεύτερο σύνολο

* (https://www.researchgate.net/figure/Typical-architecture-of-the-VGG-model-Simonyan-and-Zisserman-2014_fig1_355901164)

δεδομένων εκπροσωπεί κάποιο άλλο φαινόμενο ή πρόβλημα και προέρχεται από διαφορετική στατιστική κατανομή.¹⁵

Ο λόγος για τον οποίο μπορεί να χρειαστεί μεταφορά μάθησης, είναι η εξάρτηση της βαθειάς μάθησης σε δεδομένα, αφού χρειάζεται γιγαντιαίος όγκος πληροφοριών ώστε τα δίκτυα να καταλάβουν μοτίβα. Στη βαθειά μάθηση, τα σύνολα εκπαίδευσης και δοκιμής υπόκεινται στην υπόθεση ότι έχουν την ίδια κατανομή και την ίδια μορφή χαρακτηριστικών. Ωστόσο στην πράξη, μπορεί να υπάρχουν επαρκή δεδομένα σε μία περιοχή, ενώ το μοντέλο να εφαρμόζει επίλυση ενός προβλήματος σε άλλη περιοχή. Επίσης υπαρκτό είναι το σενάριο η κατανομή των παραδειγμάτων του προβλήματος να αλλάζει με τον χρόνο, κατάσταση που απαιτεί επανεκπαίδευση του μοντέλου με εκ νέου συλλογή δεδομένων για τη δημιουργία συνόλου εκπαίδευσης. Επειδή η κατασκευή μεγάλου όγκου δεδομένων με σωστές επισημάνσεις είναι χρονοβόρα και σε πολλές περιπτώσεις δαπανηρή, η αξιοποίηση ήδη υπάρχουσας γνώσης σε μοντέλα σε διαφορετικές περιοχές προσφέρει σημαντικά οφέλη.

Η εφαρμογή της μεταφοράς μάθησης γίνεται συχνά μέσω της διαδικασίας fine-tuning. Επιλέγεται ένα μοντέλο το οποίο έχει εκπαιδευτεί πάνω σε κάποιο πρόβλημα και μεταβάλλεται η αρχιτεκτονική του για να επιλύσει ένα παρόμοιο πρόβλημα. Η μεταβολή αυτή γίνεται με αφαίρεση ενός ή περισσοτέρων επιπέδων του παλαιού μοντέλου και αντικατάσταση τους με νέα επίπεδα, προοριζόμενα για το νέο πρόβλημα. Ενδεχομένως να χρειαστεί και αλλαγή του αριθμού κόμβων στα επίπεδα εισόδου και εξόδου, ώστε να μπορούν να χρησιμοποιηθούν τα νέα δεδομένα. Όταν η αρχιτεκτονική έχει μεταβληθεί σύμφωνα με τις ανάγκες του προβλήματος, πραγματοποιείται το πάγωμα (freezing) των επιπέδων του παλαιού εκπαιδευμένου μοντέλου που χρησιμοποιούνται στο νέο μοντέλο. Το πάγωμα διατηρεί αμετάβλητες τις παραμέτρους (βάρη, μεροληψίες) των κόμβων σε αυτά τα επίπεδα, με αποτέλεσμα στην εκπαίδευση του νέου μοντέλου με τα νέα δεδομένα να αλλάζουν μόνο οι παράμετροι των νέων επιπέδων όταν εφαρμόζεται η ανάστροφη αναδιάδοση.

Ο βαθμός παρέμβασης στο αρχικό μοντέλο εξαρτάται από το μέγεθος του νέου διαθέσιμου συνόλου δεδομένων, καθώς και τον βαθμό ομοιότητας των συνόλων δεδομένων για παλιό και νέο πρόβλημα. Όταν τα σύνολα είναι όμοια, η αντικατάσταση μόνο του τελευταίου επιπέδου είναι επαρκής και το νέο μοντέλο λειτουργεί μόνο σε επίπεδο ταξινομητή, αξιοποιώντας εξαγωγή χαρακτηριστικών, δηλ. μοτίβων, από το προεκπαιδευμένο δίκτυο και χρησιμοποιώντας τα για ταξινόμηση νέων παραδειγμάτων. Σε διαφορετική περίπτωση μπορεί να απαιτείται η αφαίρεση παραπάνω επιπέδων και η αντικατάστασή τους με παραπάνω από ένα νέα επίπεδα. Τότε η μεταφορά μάθησης αφορά προγενέστερα επίπεδα και η μεταβολή των νέων πιθανότατα απαιτεί μεγαλύτερο όγκο νέων δεδομένων εκπαίδευσης.²⁴

3.1.9 K-Εγγύτερος Γείτονας

Ένας αλγόριθμος επιβλεπόμενης μηχανικής μάθησης ο οποίος διαφέρει σε μεθοδολογία, είναι ο αλγόριθμος K-εγγύτερου γείτονα (k-Nearest Neighbor, kNN). Πρόκειται για έναν μη παραμετρικό αλγόριθμο, ο οποίος αντικατοπτρίζει σε μεγάλο βαθμό την λειτουργία μοντέλων μη επιβλεπόμενης μάθησης. Σε αντίθεση με τους άλλους αλγόριθμους μηχανικής μάθησης που έχουν παρατεθεί έως αυτό το σημείο, οι οποίοι επιτρέπουν την εγκατάλειψη των δεδομένων μετά την δημιουργία του μοντέλου, ο αλγόριθμος kNN διατηρεί όλα τα

παραδείγματα εκπαίδευσης στη μνήμη. Όταν ένα νέο, απαραίτητο παράδειγμα x εισέρχεται σε αυτόν, ο αλγόριθμος βρίσκει τα κοντινότερα k παραδείγματα εκπαίδευσης στο x , όπου k ακέραιος αριθμός. Στη συνέχεια επιστρέφει την επισημασμένη τιμή της πλειοψηφίας στην περίπτωση της ταξινόμησης, ή την μέση τιμή επισήμανσης στην περίπτωση της παλινδρόμησης.

Το πόσο κοντά βρίσκονται δύο παραδείγματα, η εγγύτητα δηλαδή, δίνεται από μία συνάρτηση απόστασης. Η ευκλείδεια απόσταση (26) χρησιμοποιείται συχνά στην πράξη και για δύο παραδείγματα x_i, x_k με ίδιο αριθμό διαστάσεων N ορίζεται ως:

$$d(x_i, x_k) = \sqrt{(x_i^{(1)} - x_k^{(1)})^2 + (x_i^{(2)} - x_k^{(2)})^2 + \dots + (x_i^{(N)} - x_k^{(N)})^2} = \sqrt{\sum_{j=1}^N (x_i^{(j)} - x_k^{(j)})^2} \quad (26)$$

Άλλη δημοφιλής επιλογή για την συνάρτηση απόστασης είναι η αρνητική συνημιτονική ομοιότητα (27). Η συνημιτονική ομοιότητα είναι δείκτης της ομοιότητας των κατευθύνσεων δύο διανυσμάτων x_i, x_k και ορίζεται ως:

$$s(x_i, x_k) = \cos(\theta) = \frac{\sum_{j=1}^N x_i^{(j)} x_k^{(j)}}{\sqrt{\sum_{j=1}^N (x_i^{(j)})^2} \sqrt{\sum_{j=1}^N (x_k^{(j)})^2}} \quad (27)$$

όπου θ η γωνία που σχηματίζουν τα δύο διανύσματα στον ευκλείδειο χώρο εφόσον αυτά ορίζονται σε αυτόν. Αν η γωνία μεταξύ των διανυσμάτων είναι 0 μοίρες, τότε τα διανύσματα δείχνουν προς την ίδια κατεύθυνση και η συνημιτονική ομοιότητα ισούται με 1. Αν τα διανύσματα είναι κάθετα μεταξύ τους, η συνημιτονική ομοιότητα ισούται με 0. Αν τα διανύσματα δείχνουν σε αντίθετες κατευθύνσεις, η συνημιτονική ομοιότητα ισούται με -1. Άρα για να χρησιμοποιηθεί ως συνάρτηση απόστασης, πολλαπλασιάζεται με το -1, ώστε τα διανύσματα με μέγιστη ομοιότητα να εμφανίζουν ελάχιστη απόσταση.

Τέλος, δημοφιλείς συναρτήσεις απόστασης είναι μεταξύ άλλων η απόσταση Chebychev, η απόσταση Mahalanobis και η απόσταση Hamming. Η επιλογή της συνάρτησης απόστασης, καθώς και ο καθορισμός της τιμής του k είναι επιλογές του χρήστη του αλγορίθμου πριν την εκτέλεση του και άρα αποτελούν υπερπαραμέτρους. Η συνάρτηση απόστασης μπορεί επίσης να μαθευτεί από τα δεδομένα, αν παραμετροποιηθεί και στη συνέχεια οι παράμετροι της βελτιστοποιηθούν με εκπαίδευση σε αυτά.

3.2 Μη επιβλεπόμενη μάθηση

3.2.1 Γενικά

Στη μη επιβλεπόμενη μάθηση, το σύνολο δεδομένων αποτελείται από μία συλλογή μη επισημασμένων παραδειγμάτων $\{x_i\}_{i=1}^N$. Το x σε κάθε παράδειγμα είναι το διάνυσμα χαρακτηριστικών του και ο σκοπός του αλγορίθμου μη επιβλεπόμενης μάθησης είναι να χρησιμοποιήσει το διάνυσμα αυτό και να το μεταμορφώσει είτε σε ένα άλλο διάνυσμα είτε σε μία τιμή, που μπορεί έπειτα να αξιοποιηθεί για την επίλυση κάποιου πρακτικού προβλήματος. Τα προβλήματα αυτά μπορεί να είναι ομαδοποίηση, κατά την οποία τα παραδείγματα κατατάσσονται σε διαφορετικές υποκατηγορίες και το μοντέλο επιστρέφει την τιμή ταυτότητας της υποκατηγορίας για κάθε διάνυσμα χαρακτηριστικών ενός συνόλου δεδομένων, μείωση διαστάσεων, όπου η εκροή του μοντέλου είναι ένα διάνυσμα με λιγότερα χαρακτηριστικά από αυτά του διανύσματος χαρακτηριστικών x , ή ανίχνευση ακραίων τιμών, όπου η εκροή είναι μία πραγματική τιμή που φανερώνει κατά πόσο το διάνυσμα χαρακτηριστικών x διαφέρει από ένα τυπικό παράδειγμα του συνόλου δεδομένων. Η παρούσα εργασία εστιάζει στην ομαδοποίηση και για αυτόν τον λόγο αναλύονται μόνο αλγόριθμοι και τεχνικές που χρησιμοποιούνται στην συγκεκριμένη υποκατηγορία μηχανικής μάθησης.

Η ομαδοποίηση (clustering) είναι το πρόβλημα της απόδοσης τιμών σε παραδείγματα χρησιμοποιώντας μη επισημασμένο σύνολο δεδομένων. Επειδή το σύνολο δεδομένων δεν περιέχει καμία απολύτως επισήμανση, η αξιολόγηση του κατά πόσο ένα μοντέλο είναι βέλτιστο είναι πολύ πιο περίπλοκη από ότι στην επιβλεπόμενη μάθηση. Υπάρχει πλήθος αλγορίθμων ομαδοποίησης από τους οποίους προκύπτουν διαφορετικής ποιότητας μοντέλα, ανάλογα με το σύνολο δεδομένων που χρησιμοποιείται. Συνήθως, η απόδοση ενός αλγορίθμου εξαρτάται από τις άγνωστες ιδιότητες της στατιστικής κατανομής του συνόλου δεδομένων.

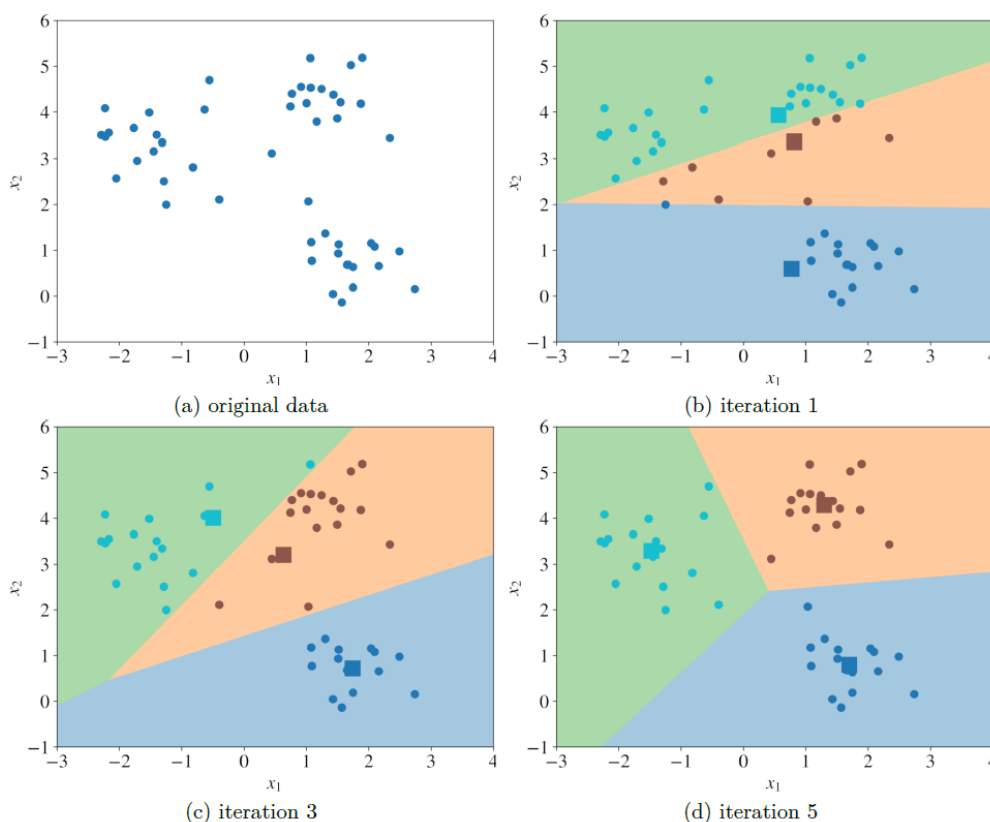
3.2.2 Αλγόριθμος K-Means

Ο αλγόριθμος ομαδοποίησης k-means λειτουργεί με επιλογή ενός ακέραιου αριθμού K . Στη συνέχεια ορίζονται K τυχαία διανύσματα χαρακτηριστικών, τα επονομαζόμενα κεντροειδή, στον διανυσματικό χώρο των χαρακτηριστικών. Ακολουθεί η μέτρηση της απόστασης κάθε παραδείγματος x_i από κάθε κεντροειδές χρησιμοποιώντας κάποια από τις συναρτήσεις απόστασης που έχουν απαριθμηθεί στο προηγούμενο εδάφιο του αλγορίθμου kNN (26) (27). Σε κάθε παράδειγμα ανατίθεται το κοντινότερο σε αυτό κεντροειδές, με τρόπο ανάλογο της επισήμανσης του παραδείγματος με μια τιμή ταυτότητας κεντροειδούς. Για κάθε κεντροειδές υπολογίζεται το μέσο διάνυσμα χαρακτηριστικών των παραδειγμάτων που έχουν επισημανθεί με αυτό. Τα μέσα διανύσματα που υπολογίζονται με αυτόν τον τρόπο γίνονται στη συνέχεια οι νέες τοποθεσίες των κεντροειδών.

Στην επόμενη επανάληψη υπολογίζεται η απόσταση κάθε παραδείγματος από κάθε κεντροειδές, τροποποιείται η επισήμανση του κάθε παραδείγματος με την ταυτότητα του

κοντινότερου πλέον κεντροειδούς και υπολογίζεται το νέο κεντροειδές. Η διαδικασία επαναλαμβάνεται μέχρι οι επισημάνσεις των παραδειγμάτων να μην αλλάζουν μετά τον επανυπολογισμό των κεντροειδών. Το τελικό μοντέλο είναι η λίστα των επισημάνσεων των παραδειγμάτων με ταυτότητες κεντροειδών (Εικόνα 12). Σημαντική παρατήρηση είναι ότι η αρχική τυχαία θέση των κεντροειδών επηρεάζει την τελική θέση των ομάδων και άρα δύο εκτελέσεις του αλγορίθμου k -means πάνω στα ίδια δεδομένα μπορεί να καταλήξουν σε δύο διαφορετικά μοντέλα. Κάποιες παραλλαγές του αλγορίθμου υπολογίζουν τις αρχικές θέσεις των κεντροειδών με βάση κάποιες ιδιότητες του συνόλου δεδομένων, ώστε να προκύπτουν μοναδικά μοντέλα για κάθε σύνολο δεδομένων.

Η τιμή του K , που αντιπροσωπεύει τον αριθμό των ομάδων, είναι μία υπερπαραμέτρος που καθορίζεται από τον χρήστη. Η επιλογή αυτή μπορεί να γίνει αυθαίρετα, με βάση δείκτες ή με οπτική εξέταση της κατανομής των παραδειγμάτων σε ομάδες, στην περίπτωση που ο διανυσματικός χώρος τους έχει μέχρι τρεις διαστάσεις. Σε επόμενο εδάφιο παρουσιάζεται μία προσέγγιση για επιλογή μίας καλής τιμής K .



Εικόνα 12: Γραφική αναπαράσταση λειτουργίας αλγόριθμου k -means για $K=3$ *

* (Andriy Burkov (2019) *The Hundred-Page Machine Learning Book*)

3.2.3 DBSCAN και HDBSCAN

Σε αντίθεση με τον k -means και παρόμοιους αλγόριθμους, οι οποίοι βασίζονται σε κεντροειδή, ο DBSCAN είναι αλγόριθμος ομαδοποίησης με βάση την πυκνότητα. Αντί για τον καθορισμό του αριθμού ομάδων εξαρχής, για την λειτουργία του DBSCAN καθορίζονται δύο παράμετροι, η απόσταση ϵ και το ελάχιστο πλήθος ομάδας n . Αρχικά επιλέγεται τυχαία ένα παράδειγμα x από το σύνολο δεδομένων και ανατίθεται στην ομάδα με αριθμό 1. Έπειτα υπολογίζεται πόσα από τα παραδείγματα απέχουν από το x απόσταση λιγότερη ή ίση με ϵ . Αν αυτή η ποσότητα παραδειγμάτων είναι μεγαλύτερη ή ίση με n , τότε όλοι αυτοί οι ϵ -γείτονες τοποθετούνται στην ίδια ομάδα 1. Στη συνέχεια εξετάζονται όλα τα μέλη της ομάδας και βρίσκονται οι ϵ -γείτονες τους. Αν κάποιο μέλος έχει n ή περισσότερους ϵ -γείτονες, η ομάδα 1 επεκτείνεται εντάσσοντας όσους από αυτούς τους ϵ -γείτονες δεν ήταν ήδη μέλη της. Η επέκταση αυτή συνεχίζεται μέχρι να μην υπάρχουν άλλα παραδείγματα που να μπορούν να ενταχθούν. Όταν σταματήσει η διαδικασία αυτή για την πρώτη ομάδα, επιλέγεται τυχαία από το σύνολο δεδομένων ένα νέο παράδειγμα που δεν ανήκει ακόμα σε ομάδα και ανατίθεται στην ομάδα 2. Ο αλγόριθμος συνεχίζει ομοίως να επεκτείνει τις επόμενες ομάδες, μέχρι όλα τα παραδείγματα να ανήκουν σε κάποια ομάδα ή να επισημανθούν ως εξαιρέσεις. Εξαιρέση νοείται ένα παράδειγμα του οποίου η ϵ -γειτονιά περιέχει λιγότερα από n παραδείγματα και δεν μπορεί να προσαρτηθεί σε ομάδα. Αν ένα παράδειγμα μπορεί να προσαρτηθεί σε ομάδα παρά την απουσία n αριθμού ϵ -γειτόνων, ονομάζεται οριακό σημείο της ομάδας στην οποία ανήκει.

Το πλεονέκτημα του αλγορίθμου DBSCAN έγκειται στην ικανότητα του να χτίζει ομάδες με αυθαίρετο σχήμα, όταν ο k -means και άλλοι βασιζόμενοι σε κεντροειδή αλγόριθμοι χτίζουν ομάδες με το σχήμα υπερσφαίρας. Αντιθέτως, το μειονέκτημα του DBSCAN εντοπίζεται στην παρουσία δύο παραμέτρων, των οποίων η επιλογή τιμής μπορεί να αποτελέσει πρόβλημα. Ακόμα, επειδή η παράμετρος της απόστασης ϵ παραμένει σταθερή κατά την εκτέλεση του αλγορίθμου, δεν μπορεί να ανταπεξέλθει σωστά σε ομάδες με διαφορετικές πυκνότητες παραδειγμάτων.

Ο αλγόριθμος HDBSCAN αποτελεί μία σημαντική βελτίωση επί του DBSCAN, αφού υλοποιεί πληθώρα περίπλοκων τεχνικών σε συνδυασμό με αυτόν. Διατηρεί τα πλεονεκτήματά του, ενώ ταυτόχρονα απαλλάσσεται από την επιλογή της παραμέτρου ϵ και άρα η μόνη σημαντική παράμετρός του είναι το ελάχιστο πλήθος παραδειγμάτων ανά ομάδα, n . Έτσι μπορεί να δημιουργήσει ομάδες με ποικίλες πυκνότητες. Επίσης έχει πολλή γρήγορη εφαρμογή ακόμα και σε πολύ μεγάλους όγκους δεδομένων. Αν και ο k -means εξακολουθεί να είναι ταχύτερος, τα πλεονεκτήματά του HDBSCAN πιθανότατα τον καθιστούν καλύτερη επιλογή για την επίλυση πρακτικών προβλημάτων.¹⁵

3.2.4 Meanshift

Ο αλγόριθμος ομαδοποίησης meanshift λειτουργεί με εύρεση τοπικών κεντροειδών με βάση την πυκνότητα των παραδειγμάτων, χωρίς την ανάγκη τυχαίας επιλογής οποιουδήποτε

παραδείγματος. Κάθε παράδειγμα «ακολουθεί» τον τοπικό μέσο μέχρι τον τελικό προορισμό του, ο οποίος είναι το κεντροειδές που ορίζει την ομάδα του.

Ο τρόπος με τον οποίο υπολογίζεται ο τοπικός μέσος των παραδειγμάτων, αντί για τον ολικό, είναι η εύρεση ενός σταθμισμένου μέσου όρου (28), ο οποίος προκύπτει με βάση το εξεταζόμενο κάθε φορά από τον αλγόριθμο σημείο. Ένας σταθμισμένος μέσος όρος διανυσμάτων ορίζεται ως:

$$x_{\mu} = \frac{\sum_{i=1}^N w_i x_i}{\sum_{i=1}^N w_i} \quad (28)$$

όπου w_i το βάρος που έχει ανατεθεί σε κάθε παράδειγμα x_i του συνόλου δεδομένων. Το βάρος σχετίζεται με την ευκλείδεια απόσταση μεταξύ των παραδειγμάτων και υπολογίζεται ακριβώς με την χρήση κάποιας συνάρτησης βάρους. Η πιο διαδεδομένη συνάρτηση βάρους που υλοποιείται στον meanshift είναι η απλή συνάρτηση (29):

$$w(d) = \begin{cases} 1, & d \leq R \\ 0, & d > R \end{cases} \quad (29)$$

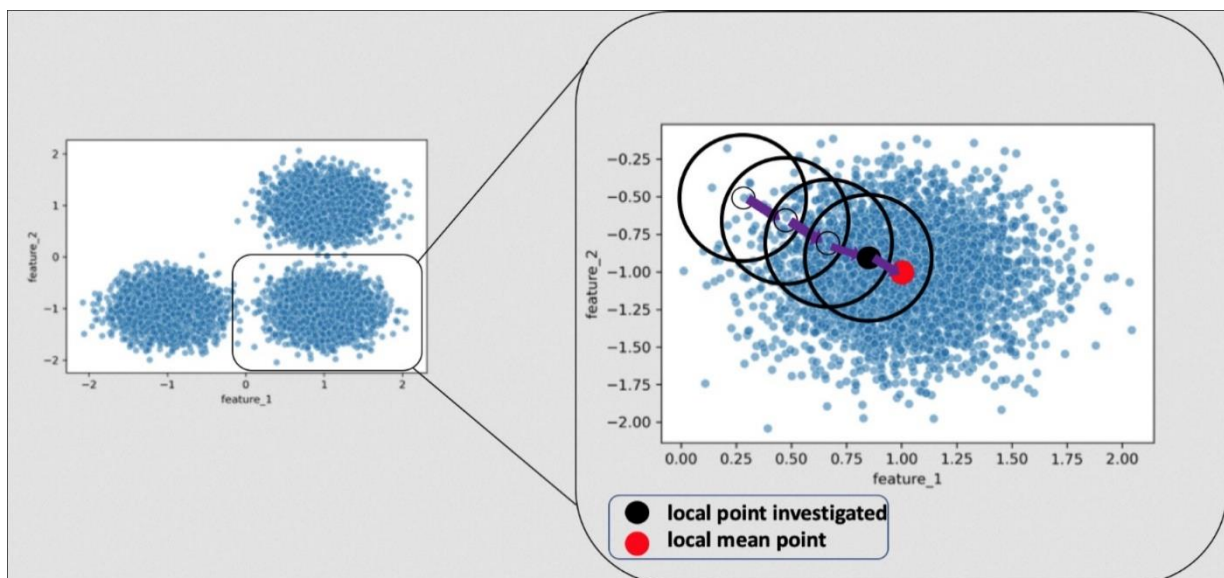
όπου d η απόσταση μεταξύ του υπό μελέτη παραδείγματος και του κάθε παραδείγματος x_i , και R η ακτίνα μιας υποθετικής υπερσφαίρας με κέντρο το υπό εξέταση σημείο. Το αποτέλεσμα του σταθμισμένου αυτού μέσου είναι να υπολογίζεται το μέσο τοπικό διάνυσμα μόνο από τα παραδείγματα εντός της ακτίνας, αγνοώντας όσα απέχουν απόσταση μεγαλύτερη από R και εκφράζοντας το σημείο με την μεγαλύτερη πυκνότητα παραδειγμάτων στη συγκεκριμένη περιοχή μόνο. Αξίζει να σημειωθεί ότι η συνάρτηση βάρους μπορεί να αντικατασταθεί με μια γκαουσιανή συνάρτηση βάρους, η οποία εξακολουθεί να υπολογίζει τα βάρη συναρτήσει απόστασης των παραδειγμάτων, αλλά χρησιμοποιώντας όλα αντί για πλήρη εξαίρεση κάποιων. Η παράμετρος R είναι η μόνη παράμετρος του αλγορίθμου meanshift και ονομάζεται εύρος ζώνης.

Στην συνέχεια, από τον τοπικό μέσο που έχει βρεθεί επαναλαμβάνεται η διαδικασία χρησιμοποιώντας το εύρος ζώνης R , και βρίσκεται ένας νέος τοπικός μέσος με βάση την πυκνότητα των σημείων στην νέα περιοχή. Μετά από αρκετές επαναλήψεις με μετακίνηση των τοπικών μέσων, ο αριθμός παραδειγμάτων μέσα στην περιοχή που ορίζει το R δεν αυξάνεται και έχει βρεθεί ο τοπικός μέσος της περιοχής με την μέγιστη πυκνότητα παραδειγμάτων. Αυτός είναι το κεντροειδές, με το οποίο χαρακτηρίζεται το αρχικό υπό εξέταση παράδειγμα και εντάσσεται σε ομάδα (Εικόνα 13).

Θεωρητικά επαναλαμβάνεται όλη η διαδικασία για κάθε παράδειγμα, ώστε να βρεθεί το κεντροειδές του και να επισημανθεί με την ομάδα του. Στην πράξη, όταν ένας τοπικός μέσος που υπολογίζεται από ένα νέο παράδειγμα βρίσκεται σε πολύ κοντινή θέση με τον τοπικό

μέσο ενός προηγούμενου ήδη εξετασμένου παραδείγματος, θεωρείται ότι θα ακολουθήσει την ίδια πορεία προς κεντροειδές και εντάσσεται σε ομάδα χωρίς να χρειαστεί περαιτέρω εύρεση και μετακίνηση τοπικών μέσων.

Το πλεονέκτημα του meanshift είναι η εξαγωγή ομάδων με την εκμετάλλευση της χωρικής κατανομής των παραδειγμάτων στον διανυσματικό χώρο, με αξιοποίηση τόσο της πυκνότητας παραδειγμάτων όσο και των θέσεων στις οποίες αυτή παρουσιάζεται. Επειδή έχει μόνο μια παράμετρο χρειάζεται προσοχή στην επιλογή της τιμής της. Υπερβολικά μεγάλο εύρος ζώνης μετακινεί κάθε τοπικό μέσο κοντά στον ολικό μέσο όλων των παραδειγμάτων, με αποτέλεσμα την κατάργηση της τοπικής δομής του συνόλου δεδομένων και τη δημιουργία μίας μόνο ενιαίας ομάδας. Από την άλλη η χρήση πολύ μικρού εύρους ζώνης δημιουργεί πολλά κεντροειδή και πολλές μικρές ομάδες χωρίς ιδιαίτερο νόημα, αφού οι τοπικοί μέσοι θα υπολογίζονται από εξαιρετικά τοπικές πυκνότητες αγνοώντας την μεγαλύτερη εικόνα των δεδομένων.²⁵



Εικόνα 13: Γραφική αναπαράσταση λειτουργίας αλγόριθμου meanshift *

3.2.5 Εκτίμηση βέλτιστου αριθμού ομάδων

Στους αλγόριθμους ομαδοποίησης που επιλέγεται από τον χρήστη ο αριθμός ομάδων, η επιλογή της κατάλληλης τιμής αποτελεί σημαντικό πρόβλημα. Όταν ο χώρος των διανυσμάτων χαρακτηριστικών των παραδειγμάτων είναι έως τρισδιάστατος, μπορούν να παρατηρηθούν εποπτικά συγκεντρώσεις σημείων οι οποίες αποτελούν υποψήφιες ομάδες. Όταν όμως τα διανύσματα χαρακτηριστικών έχουν παραπάνω διαστάσεις δεν μπορούν να απεικονιστούν και να εξαχθούν συμπεράσματα οπτικά. Η μόνη οπτική αναπαράστασή τους είναι η μερική αποτύπωση τους ανά δυάδες χαρακτηριστικών σε δισδιάστατα διαγράμματα, προσέγγιση η οποία είναι πολλές φορές υποκειμενική και ασαφής.

*(<https://towardsdatascience.com/understanding-mean-shift-clustering-and-implementation-with-python-6d5809a2ac40>)

Ένας τρόπος προσδιορισμού του αριθμού των ομάδων βασίζεται στην έννοια της ισχύος πρόβλεψης. Η διαδικασία αυτή περιλαμβάνει τον διαχωρισμό του συνόλου δεδομένων σε σύνολο εκπαίδευσης Σ_ϵ , με αριθμό παραδειγμάτων N_ϵ , και σύνολο δοκιμής Σ_δ με αριθμό παραδειγμάτων N_δ , που θεωρητικά θα έχουν την ίδια στατιστική κατανομή. Στη συνέχεια τρέχει ο αλγόριθμος ομαδοποίησης για διαφορετικές τιμές του αριθμού ομάδων K στα δύο υποσύνολα. Έστω ότι τα αποτελέσματα κάθε ομαδοποίησης συμβολίζονται με $C(\Sigma_\epsilon, k) = A$ και $C(\Sigma_\delta, k) = B$ αντίστοιχα.

Ο σκοπός είναι η χωρική σύγκριση των ομάδων που προκύπτουν από τις δύο ομαδοποιήσεις σε σχέση με τα δεδομένα δοκιμής. Με βάση την ομαδοποίηση A , ο διανυσματικός χώρος καταμετράται σε k υποπεριοχές. Ορίζεται επίσης ο πίνακας συνύπαρξης $D[A, \Sigma_\delta]$ με μέγεθος $N_\delta \cdot N_\delta$ ως εξής: $D[A, \Sigma_\delta]^{(i,i')} = 1$ αν και μόνο αν δύο παραδείγματα x_i και $x_{i'}$ του συνόλου δοκιμής ανήκουν στην ίδια υποπεριοχή, όπως αυτή ορίζεται από τις ομάδες του A . Σε διαφορετική περίπτωση $D[A, \Sigma_\delta]^{(i,i')} = 0$. Η διαγώνιος του πίνακα δηλώνει την σχέση του παραδείγματος με τον εαυτό του και άρα μπορεί να αγνοηθεί, αφού δεν αποτελεί ζευγάρι παραδειγμάτων.

Αν ο αριθμός ομάδων K είναι λογικός για την κατανομή των παραδειγμάτων, δύο παραδείγματα δοκιμής που ανήκουν στην ίδια ομάδα με βάση την ομαδοποίηση B ανήκουν στην ίδια ομάδα, άρα και υποπεριοχή με βάση την ομαδοποίηση A . Αν ο αριθμός ομάδων δεν είναι κατάλληλος, είναι δηλαδή πολύ υψηλός ή χαμηλός, τότε οι δύο ομαδοποιήσεις θα επιδρούν διαφορετικά στα δεδομένα εκπαίδευσης και δοκιμής. Αναλυτικά, η ισχύς πρόβλεψης (30) δίνεται από τον τύπο:

$$IP(k) = \min_{j=1,\dots,k} \frac{1}{|A_j|(|A_j| - 1)} \sum_{i,i' \in A_j} D[A, \Sigma_\delta]^{(i,i')} \quad (30)$$

όπου A_j η ομάδα j της ομαδοποίησης B και $|A_j|$ ο αριθμός των παραδειγμάτων που έχουν ενταχθεί στην ομάδα j . Η ερμηνεία του τύπου είναι ότι για τα παραδείγματα που ανήκουν πλέον στην ίδια ομάδα με βάση το B , αθροίζονται τα στοιχεία του πίνακα συνύπαρξης που δηλώνουν ζευγάρια παραδειγμάτων που ανήκουν στην ίδια ομάδα του A και διαιρούνται με τον υποθετικό όμοιο πίνακα που περιέχει 1 σε όλα τα στοιχεία. Το αποτέλεσμα ισούται με το ποσοστό παραδειγμάτων που προβλέφθηκαν σε ίδια ομάδα τόσο του A , όσο και του B . Από όλες τις ομάδες υπολογίζεται εκείνη με το χειρότερο ποσοστό και θεωρείται αντιπροσωπευτική για το συγκεκριμένο αριθμό ομάδων K . Μία λογική τιμή K θεωρείται εκείνη που έχει $IP(K)$ μεγαλύτερη από 0,8. Για $K = 1$, ο τύπος θα ισούται με 1, οπότε προφανώς δεν συνυπολογίζεται η περίπτωση της μίας ενιαίας ομάδας.

Για αλγορίθμους με μεταβλητά αποτελέσματα, όπως ο k -means, θεωρείται σκόπιμο να διερευνηθεί η ισχύς πρόβλεψης για περισσότερες από μία εκτελέσεις του αλγορίθμου και να υπολογιστεί η μέση ισχύς πρόβλεψης. Άλλες μέθοδοι που μπορούν να χρησιμοποιηθούν για εκτίμηση του αριθμού ομάδων είναι η μέθοδος στατιστικού κενού, η μέθοδος του αγκώνα και η μέθοδος μέσης σιλουέτας.¹⁵

3.2.6 Δείκτες εγκυρότητας αλγορίθμων ομαδοποίησης

Όταν έχει επιλεγεί ο αριθμός K του πλήθους των ομάδων και έχει πραγματοποιηθεί ομαδοποίηση πάνω σε κάποιο σύνολο δεδομένων, το αποτέλεσμα της μπορεί να αξιολογηθεί σχετικά με την απόδοση της ως προς κάποια κριτήρια, όπως το πόσο πυκνές είναι η ομάδες ή πόσο καλά διαχωρισμένες είναι η μία από την άλλη. Τα κριτήρια αυτά εκφράζονται μέσω των δεικτών που περιγράφονται σε αυτό το εδάφιο.

Οι δείκτες εγκυρότητας χωρίζονται σε τρεις κατηγορίες. Αυτές είναι η εσωτερική εγκυρότητα ομάδας, όπου το αποτέλεσμα αξιολογείται μόνο σε σχέση με το ίδιο το σύνολο δεδομένων που ομαδοποιείται, η εξωτερική εγκυρότητα, όπου το αποτέλεσμα συγκρίνεται με βάση κάποιο εξωτερικό γνωστό χαρακτηριστικό όπως μια επισήμανση, και τέλος η συγκριτική εγκυρότητα, όπου το αποτέλεσμα συγκρίνεται με εκτελέσεις του ίδιου αλγορίθμου για διαφορετικές τιμές κάποιας παραμέτρου, ώστε να επιλεγεί η καταλληλότερη τιμή. Παράδειγμα της τελευταίας κατηγορίας είναι η ισχύς πρόβλεψης που αναφέρθηκε στο προηγούμενο εδάφιο. Για την τελική αξιολόγηση της ομαδοποίησης όταν έχει επιλεγεί η παράμετρος K του πλήθους ομάδων, χρησιμοποιούνται δείκτες ή τεχνικές εσωτερικής και εξωτερικής εγκυρότητας.

Για τον υπολογισμό των δεικτών, χρησιμοποιείται η προαναφερθείσα συνάρτηση βάρους (29) για να υλοποιηθεί η παρουσία ή η απουσία ενός παραδείγματος μέσα σε συγκεκριμένη ομάδα, καθώς στους συνήθεις αλγόριθμους ομαδοποίησης κάθε παράδειγμα ανήκει σε μία μόνο ομάδα. Στην προκειμένη περίπτωση, η συνάρτηση βάρους (31) εκφράζεται ως:

$$w_{ki} = \begin{cases} 1, & x_i \in C_k \\ 0, & x_i \notin C_k \end{cases} \quad (31)$$

όπου w_{ki} το βάρος κάθε παραδείγματος x_i και C_k η υπό εξέταση ομάδα, με $k = 1, 2, \dots, K$.

Ο δείκτης Calinski - Harabasz (CHI) (32) αποτελεί έναν εσωτερικό δείκτη εγκυρότητας της ομάδας και περιγράφεται από τον τύπο:

$$CHI(K) = \frac{T_B / (K - 1)}{T_W / (K - 1)} \quad (32)$$

$$\text{όπου } T_B = \sum_{k=1}^K |C_k| \cdot \|\bar{C}_k - \bar{x}\|^2, \quad T_W = \sum_{k=1}^K \sum_{i=1}^N w_{ki} \cdot \|x_i - \bar{C}_k\|^2$$

Στον τύπο το $|C_k|$ είναι ο αριθμός παραδειγμάτων που ανήκουν στην ομάδα C_k , \bar{C}_k είναι το κεντροειδές διάνυσμα της ομάδας C_k , \bar{x} είναι το μέσο διάνυσμα από όλα τα παραδείγματα του συνόλου δεδομένων. Ο δείκτης Calinski - Harabasz είναι η αναλογία του διαχωρισμού

των ομάδων μεταξύ τους προς την συνοχή κάθε ομάδας. Η μέγιστη τιμή του CHI δείχνει βέλτιστο διαχωρισμό των ομάδων, χωρίς να υπάρχει μη αποδεκτή τιμή του δείκτη. Στην περίπτωση που διατηρείται σταθερός ο αριθμός ομάδων σε εκτελέσεις πολλών διαφορετικών αλγορίθμων ομαδοποίησης, καταλληλότερος αλγόριθμος είναι αυτός που επιστρέφει την μεγαλύτερη τιμή του δείκτη CHI.

Ο δείκτης Davies – Bouldin (DBI) (33) είναι επίσης εσωτερικός δείκτης εγκυρότητας με τύπο:

$$DBI(K) = \frac{1}{K} \sum_{k=1}^K R_k \quad (33)$$

$$\text{όπου } R_k = \max \left(\frac{S_k + S_j}{d_{kj}} \right), j = 1, 2, \dots, K, j \neq k$$

$$d_{kj} = \|\bar{x}_k - \bar{x}_j\|, S_k = \frac{1}{\sum_{i=1}^N w_{ki}} \cdot \sum_{i=1}^N w_{ki} \|x_i - \bar{x}_k\|$$

Από τον παραπάνω τύπο ο δείκτης ορίζεται ως συνάρτηση της αναλογίας της εσωτερικής διασποράς των ομάδων προς την απόσταση μεταξύ τους. Άρα διατηρώντας σταθερό αριθμό ομάδων σε εκτελέσεις διαφορετικών αλγορίθμων ομαδοποίησης, καταλληλότερος αποδεικνύεται ο αλγόριθμος με την χαμηλότερη τιμή του δείκτη DBI.

Άλλος δείκτης εσωτερικής εγκυρότητας είναι η τιμή σιλουέτας (34). Ο τύπος της για ένα υπό εξέταση παράδειγμα είναι:

$$S(x_i) = \frac{b(x_i) - a(x_i)}{\max(b(x_i), a(x_i))} \quad (34)$$

όπου $a(x_i)$ είναι η μέση απόσταση του παραδείγματος x_i από τα υπόλοιπα παραδείγματα που ανήκουν στην ομάδα του και $b(x_i)$ η μέση απόσταση του παραδείγματος x_i από τα παραδείγματα της κοντινότερης ομάδας στην οποία δεν ανήκει. Η τιμή σιλουέτας δείχνει πόσο όμοιο είναι ένα παράδειγμα με την ομάδα του, δηλαδή την συνοχή της ομάδας, σε σχέση με άλλες ομάδες, δηλαδή τον διαχωρισμό του από αυτές. Η μέση τιμή σιλουέτας που προκύπτει από όλα τα παραδείγματα μίας ομάδας είναι ένδειξη της πυκνότητας της συγκεκριμένης ομάδας, ενώ η μέση τιμή σιλουέτας που προκύπτει από όλα τα παραδείγματα του συνόλου δεδομένων είναι ένδειξη της αποτελεσματικότητας του αλγορίθμου στην ομαδοποίηση. Η τιμή σιλουέτας μπορεί να κυμαίνεται από -1 έως 1. Αν ένα παράδειγμα έχει τιμή σιλουέτας κοντά στο 1, τότε έχει ενταχθεί σε σωστή ομάδα καθώς $a(x_i) < b(x_i)$. Αντιθέτως αν η τιμή σιλουέτας του πλησιάζει το -1 τότε έχει ομαδοποιηθεί λάθος και βρίσκεται πιο κοντά στη γειτονική ομάδα, με μέσες αποστάσεις $a(x_i) > b(x_i)$. Κατά επέκταση, αν η μέση τιμή σιλουέτας από όλα τα παραδείγματα του συνόλου πλησιάζει την τιμή 1, δηλώνει καταλληλότερο αλγόριθμο για το συγκεκριμένο πρόβλημα ομαδοποίησης.²⁶

Οι παραπάνω εσωτερικοί δείκτες δίνουν μια εκτίμηση για την ποιότητα της ομαδοποίησης λαμβάνοντας υπόψη την κατανομή των παραδειγμάτων του συνόλου δεδομένων και την διάταξη των ομάδων που δημιουργήθηκαν. Ωστόσο σε προβλήματα με φυσική σημασία, στα οποία υπάρχουν αναμενόμενα αποτελέσματα και μπορούν να συγκριθούν οι δημιουργηθείσες ομάδες με αυτά, είναι πιθανό να μην ανταποκρίνεται ικανοποιητικά μια ομαδοποίηση ακόμα και αν παρουσιάζει καλούς δείκτες. Σε αυτές τις περιπτώσεις εφαρμόζονται κριτήρια εξωτερικής εγκυρότητας.

Για την εξαγωγή αυτών των εξωτερικών δεικτών, γίνεται σύγκριση της ομαδοποίησης με κάποια άλλη ομαδοποίηση. Ιδιαίτερο ενδιαφέρον παρουσιάζει η περίπτωση της επαλήθευσης της βασικής πραγματικότητας (ground truth verification). Πρώτα καθορίζεται από τον χρήστη ένας επιθυμητός τρόπος διαχωρισμού του συνόλου δεδομένων σε ομάδες. Με βάση αυτό το ανθρωπογενές κριτήριο διαχωρισμού, επισημαίνονται τα παραδείγματα με μία τιμή που καθορίζει την ομάδα τους. Η επισήμανση γίνεται δηλαδή με βάση μία χειροκίνητη ομαδοποίηση, τα αποτελέσματα της οποίας θεωρείται πως ανταποκρίνονται στην πραγματικότητα του προβλήματος. Χωρίς να εμπλακούν οι τιμές επισήμανσης των παραδειγμάτων, διενεργείται η ομαδοποίηση στο ίδιο σύνολο δεδομένων από κάποιον αλγόριθμο ομαδοποίησης και προκύπτουν οι ομάδες στις οποίες ανήκουν αυτά.

Όταν τα παραδείγματα του συνόλου δεδομένων έχουν ομαδοποιηθεί τόσο χειροκίνητα όσο και από έναν ή περισσότερους αλγόριθμους ομαδοποίησης, συγκρίνονται τα αποτελέσματα των ομαδοποιήσεων. Επομένως προκύπτουν εξωτερικοί δείκτες, οι οποίοι αξιοποιούν με κάποια μορφή την διαφορά στην συνοχή και των διαχωρισμό των ομάδων μεταξύ των διαφορετικών ομαδοποιήσεων. Ωστόσο για επαλήθευση της βασικής πραγματικότητας, όπως αυτή ορίζεται από την χειροκίνητη ομαδοποίηση, μπορεί να υποτεθεί ότι οι ομάδες που έχει υποδείξει ο αλγόριθμος έχουν τις τιμές ταυτότητας της χειροκίνητης επισήμανσης του χρήστη. Σε αυτή την περίπτωση τα παραδείγματα μπορούν να ορίσουν έναν πίνακα σύγχυσης, κατά αντιστοιχία με την επιβλεπόμενη μάθηση.

Παραδείγματα τα οποία ανήκουν στην ίδια ομάδα στις δύο ομαδοποιήσεις ορίζονται ως αληθώς μέλη της ομάδας εκείνης, παραδείγματα που δεν ανήκουν στην ομάδα στη βασική πραγματικότητα αλλά τα κατατάσσει ο αλγόριθμος σε εκείνη είναι ψευδώς μέλη της ομάδας εκείνης. Ομοίως για παραδείγματα που δεν κατατάσσονται στην συγκεκριμένη ομάδα από τον αλγόριθμο ομαδοποίησης, αληθώς όχι μέλη της ομάδας αν δεν ήταν μέλη ούτε στην βασική πραγματικότητα, ψευδώς όχι μέλη αν ήταν στην ομάδα στην βασική πραγματικότητα. Με την δημιουργία του πίνακα σύγχυσης, μπορούν ως εξωτερικοί δείκτες να χρησιμοποιηθούν οι βασιζόμενοι σε αυτήν δείκτες μηχανικής μάθησης που έχουν αναλυθεί στην εργασία, δηλαδή precision (2), recall (3), accuracy (5) και F1 (4).²⁶

Συνδυάζοντας τους εσωτερικούς με τους εξωτερικούς δείκτες, ένας καλός αλγόριθμος ομαδοποίησης πάνω σε κάποιο πρόβλημα είναι αυτός που έχει καθαρή δομή και σαφή διαχωρισμό ομάδων, ενώ ταυτόχρονα δημιουργεί ομάδες που αντιστοιχούν στις κατηγορίες που εντοπίζει ο ανθρώπινος εγκέφαλος αντικρίζοντας τα δεδομένα. Ανάλογα με την φύση του προβλήματος αλλά και τις απαιτήσεις σε ακρίβεια, κάποιοι δείκτες από όλους τους χρησιμοποιούμενους έχουν μεγαλύτερο βάρος στην αξιολόγηση των αλγορίθμων ομαδοποίησης.

4. Μηχανική Μάθηση στον Κατασκευαστικό Τομέα

4.1 Παρακολούθηση δομικής ακεραιότητας

Η παρακολούθηση δομικής ακεραιότητας (Structural Health Monitoring, SHM) είναι η παρακολούθηση και ανάλυση ενός δομικού συστήματος στην πάροδο του χρόνου μέσω μίας συλλογής πρωτοπόρων και δυναμικών διαγνωστικών εργαλείων. Οι περισσότερες τεχνικές που εντάσσονται σε αυτήν αφορούν την μέτρηση δυνάμεων, που απαιτεί την εγκατάσταση αισθητήρων επαφής όπως είναι τα επιταχυνσιόμετρα, οι μετρητές τάσεων, οι αισθητήρες οπτικών ινών και οι αισθητήρες υπερηχητικών κυμάτων, με σημαντικό κόστος εγκατάστασης.

Με την πρόσφατη ανάπτυξη αισθητήρων νέας τεχνολογίας, όπως οι ψηφιακές κάμερες και κάμερες υψηλών καρτέ ανά δευτερόλεπτο, τα μη επανδρωμένα οχήματα εδάφους, οι κινητοί αισθητήρες, έχει υπάρξει ραγδαία στροφή προς ανέπαφες τεχνικές ανίχνευσης για την παρακολούθηση δομικής ακεραιότητας. Αυτές εφαρμόζονται ευκολότερα, απαιτούν λιγότερο φόρτο εργασίας και έχουν μικρότερο κόστος, καθώς επιτρέπουν την απόκτηση δεδομένων για κτίρια μέσω χρονικών και χωρικών δεδομένων υψηλής ανάλυσης. Σε αντίθεση με τους κλασσικούς αισθητήρες επαφής, οι ανέπαφοι αισθητήρες επιστρέφουν εικόνες και βίντεο που προαπαιτούν σημαντική πρόοδο στη ρομποτική, την επεξεργασία εικόνων, την «όραση» υπολογιστών και την βαθιά μηχανική μάθηση, τομείς στους οποίους οι μηχανικοί του κατασκευαστικού τομέα δεν έχουν εκτεταμένη εμπειρία και αντιμετωπίζουν ακόμα αρκετές δυσκολίες. Πρόσφατα, οι ερευνητές παρακολούθησης δομικής ακεραιότητας έχουν εξερευνήσει τεχνικές τεχνητής νοημοσύνης για την επίλυση των δυσκολιών αυτών και για να επιτύχουν πρωτοπόρες αυτόνομες και έξυπνες στρατηγικές παρακολούθησης χρησιμοποιώντας τα ανέπαφα μηχανήματα.

Τα πορίσματα της συνολικής έρευνας επιταχύνουν τις διεργασίες παρακολούθησης και συντήρησης για τους ιδιοκτήτες των κατασκευών, αλλά και επιτρέπουν τον έγκαιρο εντοπισμό φθορών αποτρέποντας τις καταστροφικές στατικές αστοχίες που μπορεί να συμβούν μελλοντικά. Η προώθηση της έρευνας στο πεδίο αυτό επιτρέπει την αποτελεσματικότερη συντήρηση, με λιγότερα ανθρώπινα λάθη, χαμηλότερο κόστος, πιθανή υψηλότερη ακρίβεια και παρέχει ένα σύστημα από την μία άκρη στην άλλη, δηλαδή από το κτίριο απευθείας στους ιδιοκτήτες του. Μέχρι την στιγμή της συγγραφής της εργασίας αυτής, τα αποτελέσματα της έρευνας είναι πλήθος δημοσιεύσεων σε κορυφαία περιοδικά δομικής μηχανικής. Αφορούν κυρίως την ανάπτυξη αλγορίθμων επεξεργασίας εικόνων για αξιολόγηση της κατάστασης διαφορετικών τύπων κατασκευών, καθώς και διαφορετικών τύπων φθορών σε αυτές.¹⁷

4.2 Τρόποι προσδιορισμού φθορών

4.2.1 Ταξινόμηση

Με την υπόθεση ότι ένα σύστημα αυτόματου εντοπισμού φθορών βασίζεται σε συνελκτικά νευρωνικά δίκτυα, θα λειτουργεί λαμβάνοντας ως είσοδο μία εικόνα. Ο εντοπισμός της ύπαρξης μίας φθοράς μπορεί να γίνει με διάφορους τρόπους, με τον πιο απλό να είναι η ταξινόμηση (classification) ολόκληρης της εικόνας σε κατηγορία με φθορές ή κατηγορία χωρίς φθορές (Εικόνα 14).

Αφού συλλεχθούν εικόνες από τις φθορές που θα ήταν επιθυμητό να αναγνωρίζει ο αλγόριθμος, καθώς και εικόνες από τμήματα χωρίς φθορές, επισημαίνονται κατάλληλα και χρησιμοποιούνται για εκπαίδευση του αλγορίθμου μηχανικής μάθησης, όπως έχει περιγραφεί στα προηγούμενα κεφάλαια. Στην διαδικασία της δοκιμής του δικτύου με βάση ένα σύνολο δοκιμής, όταν δημιουργηθεί ο πίνακας σύγχυσης, είναι σημαντικό να ληφθεί υπόψη η φύση του προβλήματος εντοπισμού φθορών. Καθώς τα κτίρια είναι υποδομές ζωτικής σημασίας και οι αστοχίες σε αυτά εγκυμονούν τεράστιους κινδύνους, είναι σημαντικό να μην παραβλέπει το δίκτυο εικόνες που όντως παρουσιάζουν πρόβλημα. Επομένως, κρίνεται σκόπιμο να αποφεύγεται όσο το δυνατόν γίνεται η ταξινόμηση εικόνων ως ψευδώς αρνητικές. Από τους δείκτες απόδοσης της ταξινόμησης, σημαντικότερος για τον λόγο αυτόν είναι η ανάκληση (recall) στις περισσότερες περιπτώσεις.



No Crack



Crack

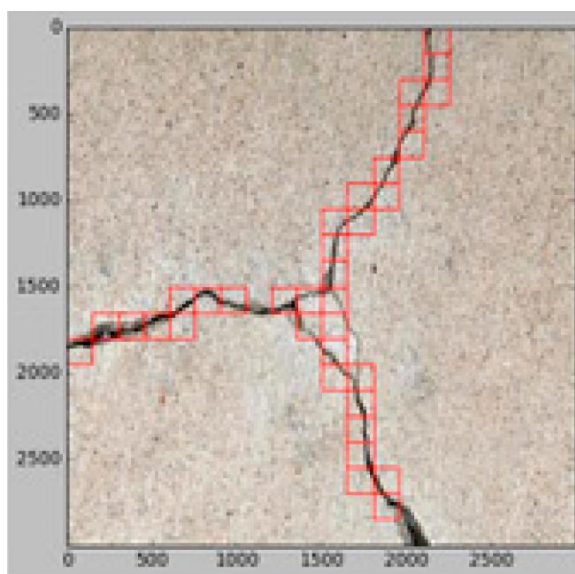
*Εικόνα 14: Ταξινόμηση εικόνων σε ρωγμή ή όχι ρωγμή**

Επειδή τα κτίσματα αντιμετωπίζουν συνήθως πολλαπλούς τύπους φθορών, η ταξινόμηση μπορεί να γίνει και για περισσότερες κατηγορίες φθορών¹⁷. Αν σε μια εικόνα καταγράφεται μόνο ένας τύπος φθοράς από τους πολλούς πιθανούς, γίνεται πολυταξική ταξινόμηση, και χρησιμοποιείται ένα επίπεδο softmax στο τέλος του δικτύου ώστε να επιστραφεί η πιθανότερη κατηγορία φθοράς. Αν υπάρχει παρουσία πολλών τύπων φθορών ταυτόχρονα,

*2018 – Özgenel, Ç.F., Gönenç Sorguç, A. “Performance Comparison of Pretrained Convolutional Neural Networks on Crack Detection in Buildings”, ISARC 2018, Berlin.)

τότε η ταξινόμηση ολόκληρης της εικόνας είναι πιο δύσκολη διαδικασία. Χρησιμοποιείται ή η συνάρτηση κόστους της δυαδικής εντροπίας, με κάποιο όριο στο οποίο ενεργοποιείται η κάθε κατηγορία, ή ξανά πολυταξική ταξινόμηση όταν υπάρχει συσχέτιση των κατηγοριών φθορών, με την δημιουργία νέας κατηγορίας που τις περιέχει ταυτόχρονα.¹⁵

Αν κάποια εικόνα θεωρηθεί ότι περιλαμβάνει φθορά, επόμενο βήμα είναι ο εντοπισμός της μέσα σε αυτήν. Μία τεχνική για την επίτευξη του εντοπισμού είναι η ταξινόμηση παραθύρων της εικόνας (image patch classification). Χρησιμοποιώντας ένα κινούμενο παράθυρο, με συγκεκριμένο μέγεθος και βήμα, εκτελούνται διαδοχικές ταξινομήσεις σε τμήματα της εικόνας, χρησιμοποιώντας κάποιο συνελκτικό νευρωνικό δίκτυο. Ανάλογα με το αποτέλεσμα της ταξινόμησης, κάθε τμήμα που ορίζεται από το παράθυρο επισημαίνεται θετικό ή αρνητικό σε παρουσία φθοράς. Όταν το κινούμενο παράθυρο έχει περάσει από όλη την εικόνα, τα μικρά υποτμήματα που έχουν ταξινομηθεί σε κατηγορία φθοράς ορίζουν την τοποθεσία της φθοράς με το σύνολο των πλαισίων τους (Εικόνα 15). Πιθανά μειονεκτήματα αυτής της μεθόδου εντοπισμού είναι η παρουσία ψευδώς θετικών και ψευδώς αρνητικών ταξινομημένων τμημάτων της εικόνας, καθώς και ο σχετικά μεγάλος απαιτούμενος χρόνος για εντοπισμό που καθιστά τον αλγόριθμο ακατάλληλο για εφαρμογές εντοπισμού φθορών σε πραγματικό χρόνο.²³



Εικόνα 15: Ταξινόμηση παραθύρων της εικόνας σε ρωγμή ή όχι ρωγμή *

4.2.2 Παλινδρόμηση με πλαίσιο οριοθέτησης

Ο συνήθης τρόπος εντοπισμού με χρήση αλγορίθμου μηχανικής μάθησης είναι η παλινδρόμηση με πλαίσιο οριοθέτησης. Ελέγχονται υποπεριοχές της εικόνας, και αν εντοπιστεί φθορά σε κάποια από αυτές, περικλείεται ολόκληρη με ένα πλαίσιο και ταξινομείται σε κάποια πιθανή κατηγορία φθοράς. Το τελικό αποτέλεσμα είναι η παρουσία

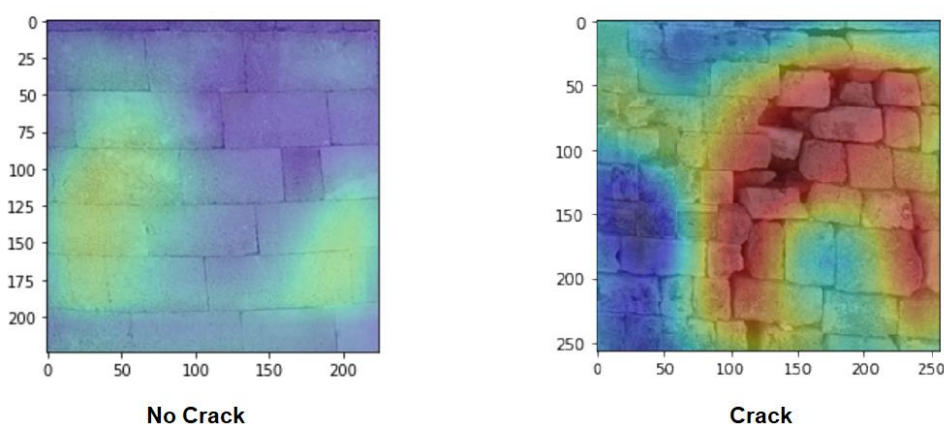
*(Ali, L.; Alnajjar, F.; Jassmi, H.A.; Gocho, M.; Khan, W.; Serhani, M.A. Performance Evaluation of Deep CNN-Based Crack Detection and Localization Techniques for Concrete Structures. Sensors 2021, 21, 1688.)

πλαισίων πάνω στην εικόνα, με εμφάνιση της κατηγορίας φθοράς που απεικονίζεται μέσα σε κάθε πλαίσιο, για άμεσο εντοπισμό και εποπτικότητα.

Ωστόσο στην περίπτωση φθορών, ειδικά όταν αυτές έχουν ακανόνιστο σχήμα, όπως για παράδειγμα εκτενείς ρωγμές, ένα πλαίσιο οριοθέτησης υπάρχει περίπτωση να μην ορίζει ικανοποιητικά την τοποθεσία που παρουσιάζεται η φθορά, ή να περιέχει υπερβολικά μεγάλο τμήμα χωρίς φθορές στην προσπάθεια να επιτύχει την πλήρη κάλυψη της φθοράς μέσα στο πλαίσιο.²⁸ Δεδομένων των μειονεκτημάτων στον εντοπισμό των φθορών σαν αντικείμενα με την χρήση πλαισίων οριοθέτησης, μπορούν να χρησιμοποιηθούν παραλλαγές του αλγορίθμου αυτού ή παραπλήσιες τεχνικές για τον πιο ακριβή εντοπισμό των φθορών σε εικόνες.

Η μέθοδος χαρτογράφησης της ενεργοποίησης κατηγοριών (Class Activation Mapping, CAM) αξιοποιεί τους χάρτες χαρακτηριστικών που προκύπτουν από το τελευταίο επίπεδο συνέλιξης ενός συνελκτικού νευρωνικού δικτύου. Για την ταξινόμηση, χρησιμοποιεί παγκόσμια συγκέντρωση αριθμητικού μέσου αντί για πλήρως συνδεδεμένα επίπεδα, δηλαδή αντικαθιστά κάθε χάρτη χαρακτηριστικών του τελευταίου συνελκτικού επιπέδου με μία μόνο τιμή, ίση με τον μέσο όρο των τιμών του. Στην συνέχεια το διάνυσμα με τον μέσο όρο κάθε χάρτη εισέρχεται σε επίπεδο συνάρτησης softmax για να προκύψει η κατηγορία. Για τον εντοπισμό, αφού εφαρμοστεί ένα βάρος σε κάθε χάρτη χαρακτηριστικών του τελευταίου συνελκτικού επιπέδου ανάλογα με την προβλεπόμενη κατηγορία ταξινόμησης, αθροίζονται οι χάρτες και το αποτέλεσμα είναι ένας συνολικός θερμικός χάρτης. Αυξάνοντας την ανάλυση του ώστε οι διαστάσεις του να είναι ίσες με αυτές της αρχικής εικόνας, ο χάρτης δίνει μία θερμική απεικόνιση των περιοχών της εικόνας που ευθύνονται για το αποτέλεσμα της ταξινόμησης (Εικόνα 16).

Τα μεγάλα προτερήματα της μεθόδου είναι η πιο ακριβής εφαρμογή σε περίπλοκα σχήματα αντικειμένων, καθώς και οι μειωμένες απαιτήσεις σε επισημάνσεις, αφού αποτελεί αδύναμα επιβλεπόμενο εντοπισμό αντικειμένων. Αντί να χρειάζεται η χωρική επισήμανση, όπως σε αλγορίθμους πλαισίου οριοθέτησης, ο εντοπισμός γίνεται με αυτόματα εξαγόμενα χαρακτηριστικά συνελκτικού δικτύου.^{24, 29} Πιθανό αρνητικό της μεθόδου είναι η λιγότερη χωρική ακρίβεια, επειδή ο θερμικός χάρτης υπολογίζεται για μικρότερη διάσταση εικόνας από την αρχική, λόγω περάσματος από επίπεδα συγκέντρωσης.

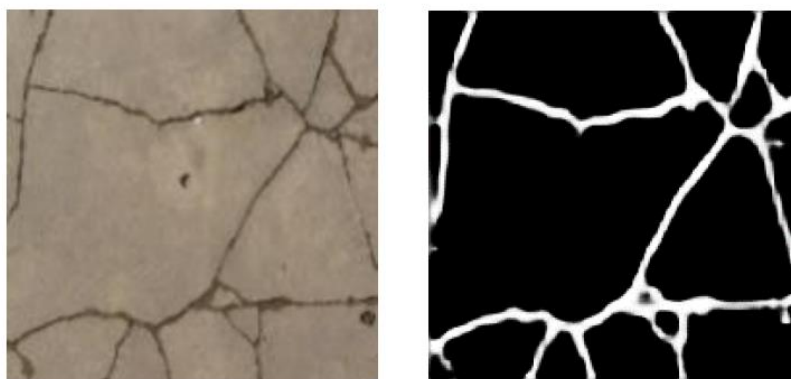


Εικόνα 16: Παράδειγμα class activation mapping για κατηγορία ρωγμής*

*P. Agrafiotis, "HYPERION, D6.3, Section: Deep Learning Paradigms, Crack detection")

4.2.3 Σημασιολογική κατάτμηση

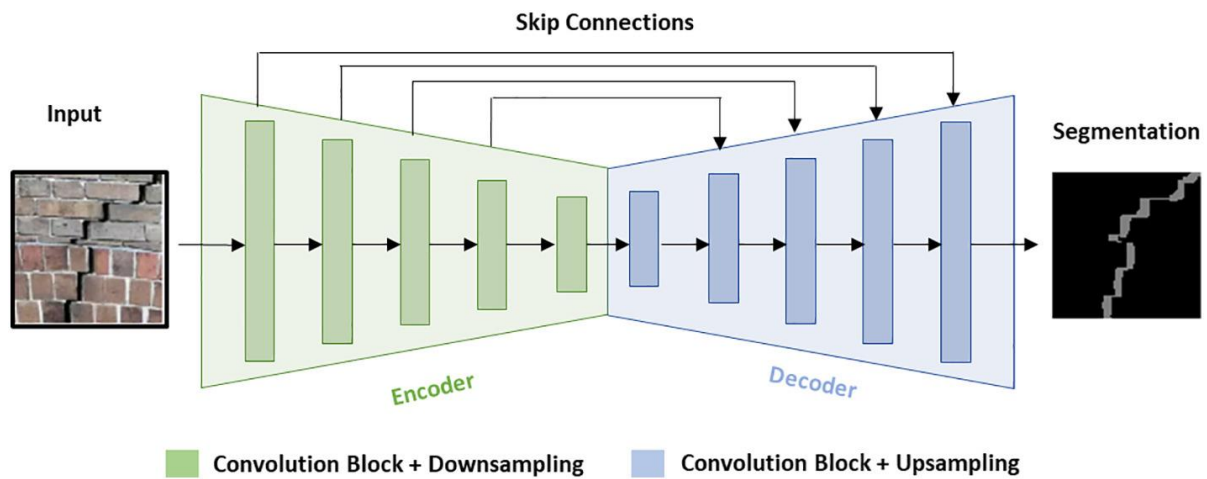
Σε αντίθεση με τις παραπάνω μεθόδους, οι οποίες χρησιμοποιούν μία γενικευμένη προσέγγιση ολόκληρης της εικόνας για εντοπισμό της ευρύτερης περιοχής μίας φθοράς, η σημασιολογική κατάτμηση (semantic segmentation) στοχεύει στον εντοπισμό κάποιας φθοράς σε επίπεδο εικονοστοιχείου. Καθώς κάθε εικονοστοιχείο ταξινομείται σε μία κατηγορία, προκύπτει πληροφορία για την ακριβή θέση, το μήκος ή και το πλάτος της φθοράς (Εικόνα 17). Ωστόσο η σημασιολογική κατάτμηση αντιμετωπίζει μία αντίφαση ανάμεσα στον χώρο και τη σημασία, αφού η γενική πληροφορία της εικόνας είναι αυτή που δείχνει το τι απεικονίζεται, ενώ η ειδική τοπική πληροφορία δείχνει το πού ακριβώς.



Εικόνα 17: Σημασιολογική κατάτμηση εικόνας ρωγμής *

Πρόσφατα έχουν χρησιμοποιηθεί πλήρως συνελκτικά δίκτυα (Fully Convolutional Networks, FCN) για σημασιολογική κατάτμηση. Τα δίκτυα αυτά λειτουργούν σαν διευρυμένα συνελκτικά δίκτυα, των οποίων η τελική πρόβλεψη είναι μία σημασιολογικά κατατμημένη εικόνα αντί για την ταυτοποίηση της κατηγορίας της. Η αρχιτεκτονική τους διαθέτει σχήμα πυραμίδας, όπου το πρώτο σκέλος έχει το σύνηθες συρρικνούμενο σχήμα των συνελκτικών νευρωνικών δικτύων για ταξινόμηση εικόνων, ενώ το δεύτερο σκέλος προστίθεται στο πρώτο και περιέχει μια ακολουθία από συνελκτικά δίκτυα και τελεστές αύξησης των διαστάσεων (upsampling operators). Το πρώτο σκέλος, ο κωδικοποιητής, βρίσκει την σημασία, ενώ το δεύτερο σκέλος, ο αποκωδικοποιητής, επιτρέπει τον ακριβή εντοπισμό. Για να αποφευχθεί η απώλεια πληροφορίας από τα βαθιά στρώματα, υλοποιούνται συνδέσεις παράκαμψης, οι οποίες επιτρέπουν στον αποκωδικοποιητή την πρόσβαση σε βαθύτερα χαρακτηριστικά που έχει αποκτήσει το σκέλος του κωδικοποιητή (Εικόνα 18).

* (Z. Qu, J. Mei, L. Liu and D. Zhou, "Crack Detection of Concrete Pavement With Cross-Entropy Loss Function and Improved VGG16 Network Model," in *IEEE Access*, vol. 8, pp. 54564-54573, 2020, doi:10.1109/ACCESS.2020.2981561.)



Εικόνα 18: Δομή πλήρους συνελκτικού νευρωνικού δικτύου*

Οι έρευνες με δίκτυα FCN έχουν χρησιμοποιήσει πρότερα εκπαιδευμένες υπάρχουσες αρχιτεκτονικές σαν βάση για το σκέλος κωδικοποιητή για σημασιολογική κατάτμηση εικόνων με ρωγμές σε σκυρόδεμα. Επίσης έχουν εφαρμοστεί FCN για ταυτοποίηση απεικονιζόμενων υλικών αλλά και διαφόρων τύπων φθορών σε αυτά. Το U-net είναι ένα βαθύ FCN που αναπτύχθηκε για βιοϊατρική κατάτμηση εικόνων. Λόγω της καλύτερης απόδοσης του σε σχέση με άλλα μοντέλα θεωρήθηκε ως το βασικό σημείο σύγκρισης για πολλά επιστημονικά πεδία, και η αποτελεσματικότητά του στην ανίχνευση λεπτών ακμών οδήγησε στην εκτενή εφαρμογή του στην επιθεώρηση κτιρίων.

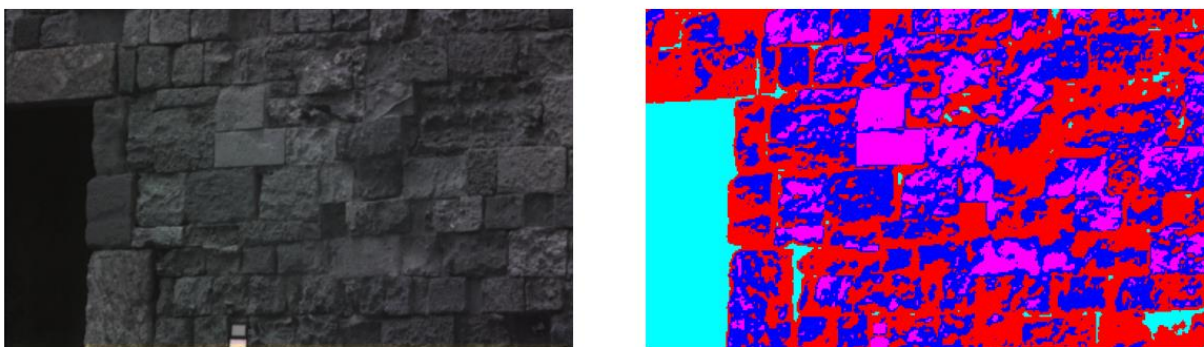
Τα δίκτυα πυραμίδων χαρακτηριστικών (Feature Pyramid Networks, FPN) είναι μία τυπική αρχιτεκτονική μοντέλων που χρησιμοποιείται στην ανίχνευση αντικειμένων. Η αρχιτεκτονική αυτή εξάγει χαρακτηριστικά σε διαφορετικές κλίμακες και στη συνέχεια τα συγχωνεύει, το οποίο φέρεται να δίνει προβλέψεις μεγαλύτερης ακρίβειας. Τα FPN μπορούν να χρησιμοποιήσουν διάφορες αρχιτεκτονικές συνελκτικών νευρωνικών δικτύων σαν κωδικοποιητή. Τα χαρακτηριστικά υψηλού επιπέδου, που έχουν έντονη σημασία αλλά χαμηλή ανάλυση, αυξάνουν τις διαστάσεις και συνδυάζονται με χαρακτηριστικά υψηλότερης ανάλυσης για να δημιουργήσουν απεικονίσεις χαρακτηριστικών που έχουν υψηλή ανάλυση και δυνατή σημασία. Χάρτες χαρακτηριστικών διαφόρων κλιμάκων έχουν αποκτηθεί από διαφορετικά εκπαιδευμένα συνελκτικά νευρωνικά δίκτυα και έχουν συνενωθεί με την χρήση FPN, για ανίχνευση φθορών σε σκυρόδεμα γεφύρων και σημασιολογική κατάτμηση ρωγμών σε εικόνες σκυροδέματος.

Τελευταία, αξιοσημείωτα αποτελέσματα στην σημασιολογική κατάτμηση έχουν επιτευχθεί με την εφαρμογή δικτύων προτάσεων περιοχών, τα οποία ακολουθούνται από αλγόριθμους για εντοπισμό ρωγμών σε επίπεδο εικονοστοιχείου. Μία τέτοια υβριδική μέθοδος, που συνδυάζει τα πλαίσια οριοθέτησης με την σημασιολογική κατάτμηση, είναι η κατάτμηση με τροποποιημένο αλγόριθμο tubularity flow field περιοχών ρωγμών που εντόπισε ένα δίκτυο

* (Dimitris Dais, İhsan Engin Bal, Eleni Smyrou, Vasilis Sarhosis, Automatic crack classification and segmentation on masonry surfaces using convolutional neural networks and transfer learning, Automation in Construction, (<https://www.sciencedirect.com/science/article/pii/S0926580521000571>))

Faster R-CNN (Kang et al, 2020)³². Η μέθοδος, αν και απαιτεί μόνο επισημάνση παραδειγμάτων εκπαίδευσης με πλαίσιο οριοθέτησης και μπορεί να εντοπίσει ρωγμές ακόμα και σε περίπλοκα υπόβαθρα, είναι χρήσιμη μόνο σε επιφάνειες από σκυρόδεμα.³⁰

Πέρα από τις μεθόδους επιβλεπόμενης μηχανικής μάθησης που περιγράφονται προηγουμένως, για την σημασιολογική κατάτμηση μπορούν να χρησιμοποιηθούν αλγόριθμοι ομαδοποίησης (clustering), οι οποίοι ανήκουν στην μη επιβλεπόμενη μηχανική μάθηση. Πρώτα διασπάται η εικόνα σε πίνακα με εικονοστοιχεία από τα οποία αποτελείται. Κάθε εικονοστοιχείο της θεωρείται ως παράδειγμα, με διάνυσμα χαρακτηριστικών τις τιμές που έχει είτε στα κανάλια του ορατού φάσματος (RGB), είτε σε υπερφασματικά κανάλια. Τα παραδείγματα αυτά εισέρχονται σε κάποιον αλγόριθμο ομαδοποίησης, ο οποίος δημιουργεί ομάδες βάσει κριτηρίων και αναθέτει μία από τις δημιουργηθείσες ομάδες σε κάθε παράδειγμα, δηλαδή εικονοστοιχείο. Το τελικό αποτέλεσμα είναι ένας πίνακας με τιμές ταυτότητας ομάδων, ο οποίος μπορεί να μεταφραστεί σε μία νέα τμηματοποιημένη εικόνα, αν αποδοθεί σε κάθε τιμή ταυτότητας διαφορετική απόχρωση (Εικόνα 19).

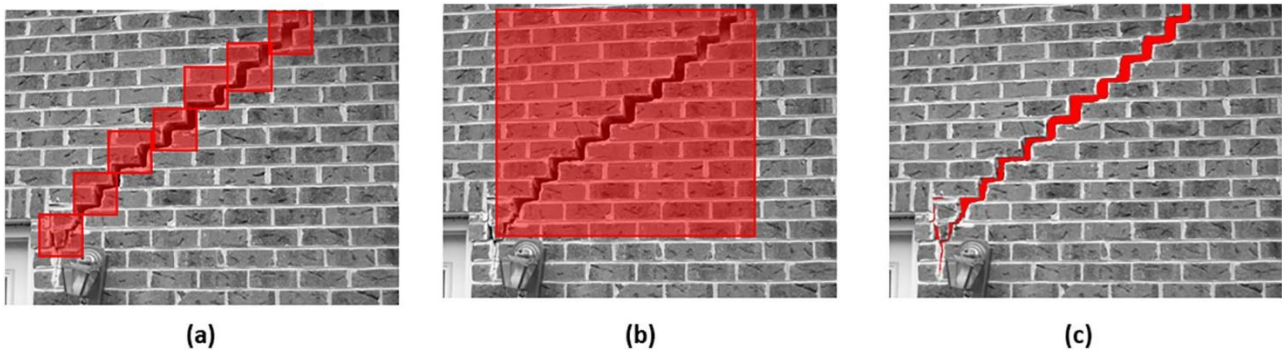


Εικόνα 19: Σημασιολογική κατάτμηση εικόνας μέσω αλγορίθμου ομαδοποίησης*

Επομένως, με την διαδικασία της ομαδοποίησης, κάθε εικόνα που εισέρχεται κατατμείται σημασιολογικά σε κατηγορίες ανάλογα με τα φασματικά της δεδομένα. Το επιθυμητό αποτέλεσμα είναι η κατανομή των εικονοστοιχείων που απεικονίζουν κάποιον τύπο φθοράς σε μία ομάδα. Η μέθοδος είναι ανεξάρτητη από χωρικά μοτίβα και εστιάζει μόνο σε διαφορές στην ακτινοβολία που εκπέμπεται από τα απεικονιζόμενα αντικείμενα. Άρα ενδείκνυται σε περιπτώσεις όπου οι φθορές δεν είναι άμεσα ορατές ώστε να εξαχθούν μοτίβα μέσω συνελκτικών δικτύων. Οι μέθοδοι αξιολόγησης είναι τόσο οι δείκτες διαχωρισμού και συνοχής των ομάδων, όσο και η σύγκριση με εικόνες στις οποίες έχουν επισημανθεί χειροκίνητα οι φθορές για να φανεί αν και κατά πόσο κάποια ομάδα του αλγορίθμου ανταποκρίνεται στην εύρεσή τους.²⁶

* (C. Zafeiropoulos, I. N. Tzortzis, I. Rallis, E. Protopapadakis, N. Doulamis, A. Doulamis, "Evaluating the Usefulness of Unsupervised monitoring in Cultural Heritage Monuments", arXiv:2107.00964 [cs.CV])

Στην εικόνα 20 παρουσιάζονται συνοπτικά οι διαφορετικοί τρόποι εντοπισμού φθορών που επεξηγήθηκαν.



Εικόνα 20: Συνολική απεικόνιση διαφορετικών τρόπων εντοπισμού φθορών:*

α) ταξινόμηση παραθύρων της εικόνας, β) παλινδρόμηση πλαισίου οριοθέτησης, γ) σημασιολογική κατάτμηση

4.3 Πρακτική εφαρμογή της έρευνας μηχανικής μάθησης

Συνδυάζοντας τα παραπάνω εδάφια, καθώς και όλες τις γενικές τεχνικές της μηχανικής μάθησης που έχουν αναλυθεί σε προηγούμενα κεφάλαια, προκύπτει ο τρόπος που ενσωματώνεται πρακτικά η έρευνα μηχανικής μάθησης στο πεδίο της παρακολούθησης δομικής ακεραιότητας. Η ενσωμάτωση αυτή μπορεί να διακριθεί σε δύο φάσεις, οι οποίες προσδιορίζονται ως η συλλογή δεδομένων και η αξιολόγηση της υπάρχουσας κατάστασης.

Η φάση της συλλογής δεδομένων σχετίζεται άμεσα με τις συνθήκες ενός προβλήματος και είναι προγενέστερη της αξιολόγησης, ενώ είναι καθοριστικής σημασίας για την κατανόηση του τύπου δεδομένων και σε ποιο αλγόριθμο ταιριάζουν. Πρώτα μελετάται ο τύπος της κατασκευής για την οποία θα συλλεχθούν τα δεδομένα, καθώς και οι ιδιαιτερότητες που αυτά παρουσιάζουν. Το συνολικό μέγεθος, τα υλικά κατασκευής καθώς και η τοποθεσία μίας κατασκευής πάνω στην γήινη επιφάνεια (υπόγεια, υπέργεια) επιδρούν άμεσα στον τρόπο απόκτησης δεδομένων και άρα στην αξιοποίηση μεθόδων μηχανικής μάθησης. Έπειτα καθορίζονται οι αισθητήρες που θα χρησιμοποιηθούν για την εξαγωγή δεδομένων από μετρήσεις. Ανάλογα με τις απαιτήσεις ενός προβλήματος επιλέγονται αισθητήρες επαφής ή ανέπαφοι, ενώ μπορούν να εξειδικευτούν ακόμα περισσότερο οι μέθοδοι ανάλογα με τα χαρακτηριστικά του κτιρίου, επιλέγοντας την λήψη εικόνων από ανθρώπινο δυναμικό, μη επανδρωμένα σκάφη ή και εναέριες εικόνες μέσω άλλων μηχανημάτων.

Όταν έχουν αποφασιστεί οι μέθοδοι λήψης των δεδομένων, ακολουθεί η συλλογή τους. Εφαρμόζεται στη συνέχεια το στάδιο της επεξεργασίας των δεδομένων και της κατάλληλης επισήμανσης τους, ανάλογα με το πρόβλημα ταξινόμησης ή πρόβλεψης τιμής που απαιτείται από μία συγκεκριμένη εφαρμογή. Τέλος, ένας στόχος της έρευνας μηχανικής μάθησης σε θέματα κατασκευαστικού τομέα είναι η δημιουργία μίας ενιαίας βάσης δεδομένων για πληροφορίες δομικής ακεραιότητας. Σε μεταγενέστερο χρόνο, τα δεδομένα που έχουν

* (Dimitris Dais, İhsan Engin Bal, Eleni Smyrou, Vasilis Sarhosis, *Automatic crack classification and segmentation on masonry surfaces using convolutional neural networks and transfer learning*, *Automation in Construction*, (<https://www.sciencedirect.com/science/article/pii/S0926580521000571>))

συλλεχθεί μπορούν να εμπλουτίσουν την βάση, για μελλοντική χρήση μέσω μεταφοράς μάθησης ή για την δοκιμή νέων αρχιτεκτονικών πάνω σε αυτά τα δεδομένα.

Η φάση της αξιολόγησης της υπάρχουσας κατάστασης αφορά την παρακολούθηση της δομικής ακεραιότητας με την χρήση σύγχρονων τεχνολογιών μηχανικής μάθησης. Αρχικά ακολουθούνται τα βήματα για την δημιουργία ικανού μοντέλου, δηλαδή η επιλογή συγκεκριμένης αρχιτεκτονικής του, η εκπαίδευση και η επικύρωση του, η λεπτομερής ρύθμιση των υπερπαραμέτρων, η δοκιμή του σε δεδομένα του πραγματικού κόσμου και η τελική εφαρμογή του σε αυτά. Όταν έχει εκπαιδευτεί στον βέλτιστο βαθμό το μοντέλο και έχει επιτύχει ικανοποιητική ακρίβεια, μπορεί να χρησιμοποιηθεί σε πρακτικά προβλήματα και να συνεισφέρει στην έρευνα.

Το τελικό στάδιο της αξιολόγησης υπάρχουσας κατάστασης, το οποίο αποτελεί και τον στόχο της έρευνας, είναι η αυτόνομη παρακολούθηση των κτιρίων, χωρίς ανθρώπινη παρέμβαση ή και επισήμανση δεδομένων που έχουν συλλεχθεί. Μέσω των μοντέλων μηχανικής μάθησης ζητείται η ανίχνευση ελαττωμάτων και ανωμαλιών, η προειδοποίηση σε αρχικό στάδιο εμφάνισης τους και η διερεύνηση της κατάστασης μετά από φυσικές καταστροφές. Το επιθυμητό αποτέλεσμα της έρευνας μηχανικής μάθησης είναι η υλοποίηση ενός συστήματος παρακολούθησης δομικής ακεραιότητας με χαμηλό κόστος, με μεγάλη ακρίβεια και από άκρη σε άκρη, μειώνοντας τα ενδιάμεσα επίπεδα δράσης και λήψης αποφάσεων.¹⁷

4.4 Πεδία εφαρμογής της έρευνας μηχανικής μάθησης

4.4.1 Γέφυρες

Οι υποδομές γεφύρωσης είναι κρίσιμες για τον τομέα των μεταφορών και άρα απαιτούν συνεχή παρακολούθηση. Τα σημαντικά τμήματα μίας γέφυρας που έχουν την τάση να εμφανίζουν ζημιές χρησιμοποιούνται για την απόκτηση δεδομένων, τα οποία μπορεί να έχουν την μορφή χρονοσειράς επιταχύνσεων (Guo et al., 2014)³³, εικόνων (Zhao et al., 2018³⁴, Liang et al., 2018³⁵) ή συνεχών μεταδόσεων βίντεο. Οι μέθοδοι βαθειάς μάθησης, όπως συνελκτικά νευρωνικά δίκτυα (CNN), πλήρως συνελκτικά δίκτυα (FCN) χρησιμοποιούνται για την αναγνώριση, ταξινόμηση αλλά και ποσοτικοποίηση των φθορών.

Τα πλεονεκτήματα της χρήσης μεθόδων συνελκτικών νευρωνικών δικτύων είναι η μεγάλη ποικιλία τύπων δεδομένων που παρέχονται για εκείνα, καθώς οι χρονοσειρές και οι οπτικές εικόνες και βίντεο καθιστούν αποτελεσματικά τα νευρωνικά δίκτυα 1 διάστασης (1D CNN) και δύο διαστάσεων (2D CNN) αντίστοιχα. Επίσης μέσω της χρήσης συνελκτικών δικτύων είναι δυνατός ο εντοπισμός τόσο γενικών φθορών (Fallahian et al., 2018)³⁶, όσο και τοπικών δομικών φθορών σε γέφυρες ή τμήματα αυτών (Narazaki et al., 2017³⁷, Kim et al., 2018³⁸)

Ένα μειονέκτημα των μεθόδων είναι η συχνή απαίτηση στην εφαρμογή ενός αλγορίθμου αραιής κωδικοποίησης πριν την εξαγωγή χαρακτηριστικών, σε συνδυασμό με την εφαρμογή συνελκτικών δικτύων, για να αντιμετωπιστεί η δυσκολία επισήμανσης των δεδομένων. Η αραιή κωδικοποίηση είναι ένα μη επιβλεπόμενο επίπεδο για την μάθηση χαρακτηριστικών από μη επισημασμένα παραδείγματα. Άλλο μειονέκτημα είναι η δυσκολία λήψης οπτικών δεδομένων από μεμονωμένα εξαρτήματα γεφυρών, και άρα τα συνελκτικά νευρωνικά δίκτυα

χρησιμοποιούνται στην εκπαίδευση μοντέλων ταξινόμησης με βάση την κατάτμηση της απεικόνισης και την αναγνώριση εξαρτημάτων γεφυρών από μία εικόνα μεγάλης κλίμακας.

4.4.2 Πεζοδρόμια

Τα πεζοδρόμια είναι επιρρεπή σε φθορές λόγω της υψηλής κινητικότητας πάνω σε αυτά και των ακραίων καιρικών συνθηκών. Το σύνολο δεδομένων για αυτές τις εφαρμογές αποτελείται συνήθως από εικόνες που λαμβάνονται με κάμερες σε ταμπλό αυτοκινήτων (Maeda et al., 2018)³⁹ ή μη επανδρωμένα αεροσκάφη (Gopalakrishnan et al., 2017⁴⁰, Kim and Cho, 2018⁴¹), αλλά σε έρευνες έχουν χρησιμοποιηθεί και φωτογραφίες από νέες τεχνολογίας κινητά τηλέφωνα (Fan et al., 2018)⁴². Η κυριότερη φθορά που εμφανίζεται σε πεζοδρόμια είναι η παρουσία ρωγμών.

Στη δημοσίευση του Zhang (Zhang et al., 2017)⁴³ προτάθηκε ένα συνελικτικό δίκτυο επιπέδου εικονοστοιχείου για την ανίχνευση ρωγμών σε τρισδιάστατες επιφάνειες πεζοδρομίου. Το συνελικτικό δίκτυο ονομάστηκε “CrackNet”, αποτελούταν από δύο πλήρως συνδεδεμένα επίπεδα, ένα επίπεδο συνέλιξης, ένα επίπεδο συνέλιξης 1*1 και ένα επίπεδο εξόδου. Το δίκτυο ήταν πιο αποτελεσματικό από παραδοσιακά συνελικτικά νευρωνικά δίκτυα λόγω της απουσίας επιπέδων συγκέντρωσης που μειώνουν τις διαστάσεις προηγούμενων επιπέδων. Άλλες αξιοσημείωτες έρευνες με υψηλή ακρίβεια είναι οι δημοσιεύσεις των Cha (Cha et al., 2017)⁴⁴ και Fan (Fan et al., 2019)⁴⁵.

Τα πλεονεκτήματα χρήσης μεθόδων συνελικτικών νευρωνικών δικτύων για τον εντοπισμό ρωγμών και την παρακολούθηση των πεζοδρομίων είναι η δυνατότητα λήψης συνόλου δεδομένων με εικόνες κάτω από ποικίλες περιβαλλοντικές συνθήκες, η πληθώρα των επιλύσιμων προβλημάτων, όπως η ανίχνευση της ύπαρξής τους, η εύρεση του μεγέθους τους και ο εντοπισμός του χώρου που βρίσκονται, και ο αποτελεσματικός υπολογισμός του μήκους των ρωγμών με την ικανότητα συνδυασμού συνελικτικών δικτύων και ομαδοποίησης k-means (Tong et al., 2017)⁴⁶.

Τα κύρια μειονεκτήματα είναι η παρουσία θορύβου και η πολυπλοκότητα των ρωγμών σε εικόνες, που οδηγούν στην συμπλήρωση των δικτύων με αυξημένη προεπεξεργασία όπως η εφαρμογή διμερούς φίλτρου για εξομάλυνση θορύβου και η υλοποίηση προσαρμοζόμενων ορίων απόφασης, το γεγονός ότι τα σύνολα δεδομένων περιέχουν συχνά ανισοβαρείς παρατηρήσεις για εκπαίδευση λόγω της σπανιότητας αλλά και των πολύ διαφορετικών σχημάτων ρωγμών και, τέλος, η πιθανή παρουσία διαφορετικών τύπων ρωγμών που μοιάζουν, που για σωστή ταξινόμηση χρειάζονται εφαρμογή μεταφοράς μάθησης και την υλοποίηση πιθανοτήτων πρόβλεψης της κατηγορίας. Για παράδειγμα οι πραγματικές ρωγμές, τα όρια μεταξύ πλακών στο πεζοδρόμιο, οι γρατζουνιές από ανθρώπινες δραστηριότητες και οι μπαλωμένες ρωγμές έχουν κοινά επιμήκη χαρακτηριστικά, ωστόσο υπάρχει πιθανότητα να μην είναι επιθυμητή η καταγραφή όλων ως ρωγμή. Σε αυτή την περίπτωση είναι προτιμότερη η πολυταξική ταξινόμηση, παρά η δυαδική ταξινόμηση ρωγμής ή όχι ρωγμής.³¹

4.4.3 Υπόγειες Κατασκευές

Οι υπόγειες κατασκευές, όπως οι σωλήνες υπονόμου και οι σήραγγες, είναι μη προσβάσιμες για επιθεώρηση. Επομένως η παρακολούθηση τους γίνεται μέσω βίντεο, σε συνδυασμό με τεχνικές βαθειάς μάθησης, όπως συνελικτικά δίκτυα, kNN αλλά και φασματική ομαδοποίηση (spectral clustering). Τα σύνολα δεδομένων για εκπαίδευση μπορούν να ληφθούν με κάμερες DSLR (Stentoumis et al., 2016)⁴⁷, κάμερες κλειστού κυκλώματος (Cheng and Wang, 2018)⁴⁸ ή και αυτόνομα ρομποτικά συστήματα, όπως ρομποτικό όχημα με ρομποτικό βραχίονα που λαμβάνει φωτογραφίες κατά την κίνηση του (Doulamis et al., 2018)⁴⁹.

Η εφαρμογή μεθόδων μηχανικής μάθησης που χρησιμοποιούν ως δεδομένα εισόδου εικόνες, πλεονεκτεί έναντι των παραδοσιακών μεθόδων παρακολούθησης δομικής ακεραιότητας που βασίζονται σε δονήσεις. Αυτό οφείλεται στην δυσκολία επιθεώρησης με αυτές των υπόγειων κατασκευών, όπως σωλήνες υπονόμου, ύδρευσης και συστήματα υδροηλεκτρικών φραγμάτων λόγω του μεγάλου βάθους και των παχιών διαστάσεων τους. Για τις υπερβολικά εκτενείς υδρολογικές κατασκευές, μπορούν να χρησιμοποιηθούν μη επανδρωμένα αεροσκάφη με κινηματικό εντοπισμό πραγματικού χρόνου (Real Time Kinematic positioning, RTK) για την συλλογή δεδομένων και την ανίχνευση φθορών. Επίσης όταν τα δεδομένα είναι εξαρτημένα μεταξύ τους, όπως δεδομένα από ραντάρ (Kang et al., 2020)⁵⁰, τα συνελικτικά νευρωνικά δίκτυα έχουν καλύτερη απόδοση, με την προϋπόθεση ότι έχει αφαιρεθεί ο θόρυβος από τα σήματα.

Στα μειονεκτήματα συγκαταλέγονται οι διαφορετικές προσεγγίσεις που απαιτεί η συγκέντρωση των δεδομένων από τις κατασκευές, καθώς διαφέρουν τόσο σε λειτουργία όσο σε υποδομές. Σε σήραγγες μεταφορών μπορούν να χρησιμοποιηθούν οι φωτογραφικές κάμερες DSLR και τα ρομποτικά συστήματα, ωστόσο σε υπονόμους καλύτερη μέθοδος είναι οι κάμερες κλειστού συστήματος. Ακόμα, εικόνες σε υπόγειες κατασκευές έχουν συχνά κακές συνθήκες φωτισμού, γεγονός το οποίο εισάγει θόρυβο σε αυτές και δυσχεραίνει την συγχώνευση συνελικτικών δικτύων με μη επιβλεπόμενη ομαδοποίηση που απαιτείται για καλύτερο προσδιορισμό της έκτασης των ρωγμών. Τέλος, το σύνολο δεδομένων σε αυτές τις περιπτώσεις τείνει να περιέχει περιορισμένο αριθμό παραδειγμάτων, ωστόσο οι τεχνικές μπορούν να ενισχυθούν με την χρήση μεταφοράς μάθησης (Βλ. 3.1.8).¹⁷

4.4.4 Μνημεία πολιτιστικής κληρονομιάς

Ως μνημεία ορίζονται τα κτίσματα που δημιουργούνται από ιστορικά άτομα ή γεγονότα και συμβολίζουν μία ιστορική περίοδο της τοποθεσίας τους μέσω της καλλιτεχνικής, ιστορικής, πολιτικής ή αρχιτεκτονικής σημασίας τους. Τα μνημεία, όπως και άλλα ευρήματα πολιτιστικής κληρονομιάς, υποφέρουν από διαρκή αλλοίωση λόγω φυσικών καταστροφών, της κλιματικής αλλαγής αλλά και ανθρώπινη παρέμβαση ή αμέλεια. Οι μέθοδοι ανίχνευσης και αξιολόγησης των φθορών σε πέτρινα μνημεία διακρίνονται σε καταστροφικές και μη καταστροφικές μεθόδους. Στις καταστροφικές μεθόδους, ένα τμήμα του μνημείου αφαιρείται για μελέτη, ενώ στις μη καταστροφικές χρησιμοποιούνται διάφοροι αισθητήρες και τεχνολογίες όπως φωτογραμμετρία και υπερφασματικοί αισθητήρες, για αποτύπωση των χαρακτηριστικών της υπό μελέτη επιφάνειας. Αυτά μπορούν να χρησιμοποιηθούν στην συνέχεια σαν είσοδος τόσο για συνελικτικά νευρωνικά δίκτυα, όσο και για αλγορίθμους μη

επιβλεπόμενης μάθησης όπως φασματική ομαδοποίηση εικονοστοιχείων. Ο σκοπός τους είναι η ανίχνευση των ρωγμών, των φθορών στην αρχιτεκτονική επιφάνεια και της αποσύνθεσης του υλικού από τα οποία αποτελούνται τα κτίσματα.

Τα πλεονεκτήματα της χρήσης μεθόδων μηχανικής μάθησης είναι ότι δεν απαιτούν την καταστροφή ή αφαίρεση τμήματος από ένα μνημείο, όπως οι καταστροφικές μέθοδοι αξιολόγησης της κατάστασης του, και άρα μπορεί να διατηρηθεί στην ίδια κατάσταση μετά την ολοκλήρωση της δομικής αξιολόγησης. Ακόμα η λήψη πληροφοριών από απόσταση μειώνει τον κίνδυνο στην δομική διάρθρωση του μνημείου από ανθρώπινο λάθος. Σημαντικό μειονέκτημα των μεθόδων επιβλεπόμενης, αλλά και μη επιβλεπόμενης μηχανικής μάθησης που εφαρμόζονται σε μνημεία είναι η περίπλοκη διαδικασία επισήμανσης των παραδειγμάτων. Καθώς συχνά δεν είναι άμεσα ορατές οι φθορές που πρέπει να εντοπιστούν από τον αλγόριθμο, χρειάζεται εξειδικευμένη επισήμανση τους, η οποία είναι χρονοβόρα και απαιτεί προσεκτική εργασία.^{26,28}

4.4.5 Σύγχρονα κτίρια

Τα ψηλά κτίρια εμφανίζουν δυσκολίες στην χειροκίνητη επιθεώρηση και επομένως χρειάζονται έναν προσβάσιμο τρόπο για αυτόνομη παρακολούθηση. Πέρα από την συνήθη προσπάθεια ανίχνευσης ρωγμών, έρευνες έχουν γίνει τόσο για την επιθεώρηση κτιρίων μετά από καταστροφικά γεγονότα (Yeum et al., 2018)⁵¹, όπως σεισμοί ή τυφώνες, όσο και για φθορές και αστοχίες που προκαλούνται από τη μακροχρόνια επίδραση υγρασίας, όπως μούχλα, φθορές τοιχωμάτων και λεκέδες. Οι κύριες μέθοδοι απόκτησης δεδομένων είναι μέσω καμερών και μη επανδρωμένων αεροσκαφών (Chaiyasarn et al., 2018)⁵², ενώ έχει χρησιμοποιηθεί και ένα μη επανδρωμένο όχημα ανάβασης σε τοίχο, το οποίο κατέγραψε βίντεο από το οποίο προέκυψαν εικόνες με ρωγμές για εκπαίδευση νευρωνικού δικτύου (Jiang and Zhang, 2019)⁵³.

Η χρήση μεθόδων μηχανικής μάθησης έχει το πλεονέκτημα ότι μπορεί να εφαρμοστεί εσωτερικά των κτιρίων και, κυριότερα, εξωτερικά των κτιρίων με σχετικά φθηνές ψηφιακές κάμερες, κάμερες από κινητά τηλέφωνα ή από μη επανδρωμένα αεροσκάφη. Ακόμα, πέρα από την απλή ανίχνευση ρωγμών, με την χρήση αλγορίθμων εντοπισμού μπορεί να διαπιστωθεί κατά πόσο οι φθορές απειλούν δομικά στοιχεία που δέχονται στατικά φορτία ή που δεν ευθύνονται άμεσα για την δομική υγεία. Επομένως μπορεί να γίνει άμεση αξιολόγηση της επικινδυνότητας μίας φθοράς, αν επιδρά σε φέροντα οργανισμό της δομής. Τα αρνητικά των μεθόδων είναι ότι δεν υπάρχει άμεσος τρόπος μελέτης της τρισδιάστατης φύσης των κτιρίων, και άρα απαιτούνται επιπρόσθετες τεχνικές για την ανάλυση κατασκευών στην τρίτη διάσταση, ενώ γενικότερο πρόβλημα είναι η ανεπάρκεια των δεδομένων εκπαίδευσης σε δομικά ζητήματα, που σημαίνει ότι τα δίκτυα είναι υποχρεωμένα να στηρίζονται σε ήδη εκπαιδευμένα βασικά μοντέλα, όπως το VGG16.

4.4.6 Άλλες κατασκευές μεγάλης κλίμακας

Οι κατασκευές εκτεταμένης κλίμακας παρακολουθούνται δύσκολα από ανθρώπινο εργατικό δυναμικό και για τον λόγο αυτόν τεχνικές με χρήση εικόνων είναι σημαντικό εργαλείο στην παρακολούθηση της δομικής τους υγείας. Παραδείγματα έρευνας πάνω σε αυτές είναι η εφαρμογή συνελκτικών νευρωνικών δικτύων σε φωτογραμμετρικά στερεοζεύγη για την ανίχνευση φθορών σε σιδηροτροχιές υπό διάφορες συνθήκες φωτισμού (Soukup and Huber-Mork, 2014)⁵⁴, η ανίχνευση μικροσκοπικών ρωγμών σε μεταλλικές επιφάνειες για επιθεώρηση πυρηνικών αντιδραστήρων μέσω συνελκτικού δικτύου σε συνδυασμό με πιθανολογικό ταξινομητή Bayes (Chen and Jahanshahi, 2018)⁵⁵ και, τέλος, η χρήση τριών διαφορετικών συνελκτικών νευρωνικών δικτύων για ανίχνευση ρωγμών σε λεπίδες ανεμογεννητριών (Xu et al, 2019)⁵⁶ καθώς και η εφαρμογή συνελκτικών δικτύων με πρόταση περιοχών για τον έλεγχο χαλάρωσης βιδών κάτω από διαφορετικές συνθήκες λειτουργίας (Zhang et al, 2020)⁵⁷. Ιδιαίτερο ενδιαφέρον παρουσιάζει μία έρευνα για επιθεώρηση κρίσιμων εγκαταστάσεων ηλεκτροδότησης. Μέσω μεταφοράς μάθησης και συνελκτικών δικτύων, εικόνες από εγκατεστημένες σε οχήματα κάμερες ταξινομήθηκαν σε πέντε κατηγορίες, δηλαδή αυτοκινητόδρομους, πεύκα, χωράφια, φορτηγά και κατασκευές ηλεκτροδότησης. Η τεχνική προσέφερε αυτόματο εντοπισμό βλάβστησης, η οποία εθεωρείτο μεγάλος κίνδυνος για το δίκτυο υποδομών ηλεκτροδότησης.

Οι αλγόριθμοι μηχανικής μάθησης σε αυτήν την περίπτωση έχουν το θετικό της αρκετά καλής απόδοσης, σε πληθώρα κατασκευών και περιβαλλοντικών συνθηκών. Ωστόσο ο θόρυβος μπορεί να αλλοιώσει τα δεδομένα μεγάλων σε έκταση κατασκευών, ενώ λόγω της ιδιοτροπίας πολλών κατασκευών χρειάζεται μεγάλος όγκος δεδομένων εκπαίδευσης ώστε να συγκλίνουν οι αλγόριθμοι και να αποφευχθεί η υπερπροσαρμογή ενός μοντέλου.¹⁷

4.5 Δυσκολίες στην εφαρμογή συνελκτικών νευρωνικών δικτύων στην δομική ακεραιότητα

Με τις αυξημένες υπολογιστικές ικανότητες της εποχής της πληροφορίας, οι τεχνικές συνελκτικών νευρωνικών δικτύων έχουν εξελιχθεί σημαντικά στο πεδίο της από απόσταση και αυτόνομης παρακολούθησης δομικής ακεραιότητας σε κρίσιμες αστικές υποδομές. Παρά τις ευκαιρίες που παρουσιάζει η έρευνα για την δημιουργία αποτελεσματικών και φθηνών συστημάτων παρακολούθησης, η υπάρχουσα βιβλιογραφία συνελκτικών νευρωνικών δικτύων παρουσιάζει ορισμένες δυσκολίες που πρέπει να αντιμετωπιστούν τα επόμενα χρόνια πριν μπορέσει να εφαρμοστεί γενικά σε παρακολούθηση και συντήρηση ποικίλων δομών.

Μερικές από τις δυσκολίες είναι αυτές που παρουσιάζονται σε όλες τις εφαρμογές νευρωνικών δικτύων, ανεξαρτήτως πεδίου, και άρα επιδρούν και σε εφαρμογές δομικής ακεραιότητας. Αυτές είναι το μεγάλο κόστος της εκπαίδευσης ενός δικτύου σε μία βάση δεδομένων δομικής ακεραιότητας, λόγω υψηλών απαιτήσεων σε χώρους αποθήκευσης και σε ενέργεια για την χρήση καρτών γραφικών, και η λειτουργία των δικτύων σαν μαύρα κουτιά, καθώς τα ενδιάμεσα επίπεδα δεν έχουν άμεσα αναγνωρίσιμα χαρακτηριστικά, ενώ η αξιολόγηση με βάση μόνο την τελική ορθότητα μπορεί να έχει καταστροφικά αποτελέσματα.

Ειδικότερα στο υπό μελέτη πεδίο, ένα θέμα στις υποδομές μεγάλης κλίμακας είναι η ανισορροπία των δεδομένων. Σε αντίθεση με συστήματα σε άλλα πεδία μηχανικής, η αστικές

υποδομές έχουν κολοσσιαίο μέγεθος και σχεδιάζονται για δεκαετίες διάρκειας ζωής. Επομένως οι παροδικές στιγμιαίες μετρήσεις της στατικής κατάστασης μπορούν να παρουσιάσουν μεγάλο εύρος διαφορετικών σταδίων φθορών. Αν συνυπολογιστούν και τα πολλά διαφορετικά ακραία φαινόμενα, δυσχεραίνεται ακόμα περισσότερο η δημιουργία ενιαίας βάσης δεδομένων εκπαίδευσης. Μάλιστα σε περιπτώσεις μνημείων, που έχουν υποστεί φθορές σε πολύ μεγάλο βάθος χρόνου, το πρόβλημα μπορεί να είναι ακόμη πιο έντονο.

Η μεγάλη ποικιλία δεδομένων στην παρακολούθηση δομικής υγείας οφείλεται και στον τύπο των κατασκευών και τον τύπο των αισθητήρων για την απόκτηση δεδομένων. Κατά συνέπεια, οι υπάρχουσες προσεγγίσεις των ερευνών εστιάζουν στην επιλογή της καταλληλότερης αρχιτεκτονικής για τα συγκεκριμένα διαθέσιμα δεδομένα. Για παράδειγμα, κατασκευές από μέταλλο ή σκυρόδεμα υπό τις ίδιες συνθήκες λειτουργίας δεν ανταποκρίνονται απαραίτητα στην ίδια αρχιτεκτονική συνελκτικών δικτύων. Η προσαρμογή και η κάλυψη αρχιτεκτονικών σε πολλαπλές υποδομές αποτελεί ακόμα πρόβλημα, αν και μπορεί ενδεχομένως να επιλυθεί σε κάποιον βαθμό με έρευνες για πιο αποδοτική μεταφορά μάθησης μεταξύ δικτύων.

Τέλος, συχνά εντοπίζονται από μοντέλα πρόβλεψης ψευδώς θετικά στοιχεία τα οποία οφείλονται στο διαφορετικό υπόβαθρο των εικόνων λόγω περιβαλλοντικών συνθηκών. Αυτές μπορεί να είναι σκιές, φωτισμός, καιρικά φαινόμενα, διαφορετικά χρώματα και ανεπιθύμητα για εντοπισμό αντικείμενα, όπως πεζοί και οχήματα. Αυτός ο θόρυβος κατά την εκπαίδευση μπορεί να οδηγήσει σε ανεπαρκή προσδιορισμό φθορών σε περιπτώσεις γεφυρών, πεζοδρομίων, οδοστρώματος και σωλήνων. Συγκεκριμένα στην μελέτη ρωγμών, δεν έχει μελετηθεί εκτενώς η επίδραση φωτισμού, καιρικών φαινομένων και η απόσταση της κάμερας από κατασκευές. Πέρα από την αρχιτεκτονική των δικτύων, η προετοιμασία και ο όγκος των δεδομένων εκπαίδευσης είναι επίσης αντικείμενο έρευνας.¹⁷

5. Μέθοδος Ανάλυσης

5.1 Γενικά

Για τον απώτερο σκοπό της παρακολούθησης και χαρτογράφησης παθογενειών των υλικών σε μνημεία, αναπτύσσονται αυτόματα συστήματα που χρησιμοποιούν εικόνες και μηχανική μάθηση και τα οποία είναι επιθυμητό να αντικαταστήσουν σταδιακά τη χειροκίνητη εργασία. Στο πλαίσιο αυτό, ο σκοπός της παρούσας εργασίας είναι η σύγκριση της επιβλεπόμενης μάθησης με την μη επιβλεπόμενη, και άρα ο έλεγχος της υπόθεσης ότι μπορεί να εφαρμοστεί η μέθοδος των συνελκτικών νευρωνικών δικτύων πάνω στην ανίχνευση δυσεύρετων φθορών σε μνημεία. Επομένως, κρίνεται σκόπιμο να ελεγχθούν τα αποτελέσματα διαφόρων αλγορίθμων πάνω στο ίδιο σύνολο δεδομένων, στον μέγιστο βαθμό που αυτό είναι εφικτό.

Αναλυτικότερα, θα ελεγχθεί το αποτέλεσμα αλγορίθμου ομαδοποίησης K-means στα εικονοστοιχεία εικόνας, για την μη επιβλεπόμενη μάθηση, με το αποτέλεσμα συνελκτικών νευρωνικών δικτύων VGG-16 στην ταξινόμηση παραθύρων της εικόνας για την επιβλεπόμενη. Το κοινό σύνολο δεδομένων αποτελείται από πολυφασματικές εικόνες του φρουρίου του Αγίου Νικολάου στη Ρόδο. Σύμφωνα με τους συγγραφείς της δημοσίευσης «Evaluating the Usefulness of Unsupervised monitoring in Cultural Heritage Monuments²⁶», θεωρητικά καταλληλότερος αλγόριθμος ομαδοποίησης για την περίπτωση αυτή είναι ο Meanshift, ενώ πρακτικά πιο αποτελεσματικός είναι ο DBSCAN. Για την εφαρμογή συνελκτικού δικτύου, η πιο διαδεδομένη αρχιτεκτονική σε παρόμοιες εφαρμογές είναι η VGG-16, με αξιοποίηση μεταφοράς μάθησης λόγω μικρού όγκου δεδομένων. Οι εικόνες πρέπει να υποστούν κατάλληλη προεπεξεργασία για να εισέλθουν στον αλγόριθμο επιβλεπόμενης μηχανικής μάθησης, η οποία επεξηγείται παρακάτω.

Τέλος, για να είναι δυνατή η σύγκριση των δύο διαφορετικών μεθόδων, στα αποτελέσματα των μοντέλων συνελκτικών δικτύων πρέπει να υπάρχει και εντοπισμός των φθορών, πέρα από την απλή κατηγοριοποίηση ενός παραθύρου της εικόνας σε φθορά. Επειδή η σημασιολογική κατάτμηση με συνελκτικό δίκτυο κρίνεται σχεδόν αδύνατη στο παρόν σύνολο δεδομένων, λόγω της μορφής των φθορών και των λίγων σχετικά παραδειγμάτων, επιχειρείται ο εντοπισμός και η επακόλουθη σύγκριση με την εφαρμογή της μεθόδου Class Activation Mapping (βλ. 4.2.2).

5.2 Δεδομένα

Οι αλγόριθμοι μηχανικής μάθησης, θα εφαρμοστούν πάνω σε πολυφασματικές εικόνες, οι οποίες απεικονίζουν τμήματα των τοιχωμάτων του φρουρίου του Αγίου Νικολάου (Εικόνα 21). Το φρούριο αποτελεί γνωστό αξιοθέατο της πόλης της Ρόδου, ενώ η ανέγερσή του χρονολογείται στον 15^ο αιώνα μ.Χ. Βρίσκεται στην είσοδο του παλαιού λιμανιού της πόλης, και η τοποθεσία του αυτή το αφήνει εκτεθειμένο στην θάλασσα, τους ανέμους και άρα το θαλασσινό νερό, με αποτέλεσμα την διάβρωση σε βάθος χρόνου πολλών από τις πέτρες που σχηματίζουν τα δομικά στοιχεία.



Εικόνα 21: Άποψη του Φρουρίου Αγίου Νικολάου, Ρόδος, Ελλάδα*

Οι εικόνες περιέχουν 41 φασματικά κανάλια, τα οποία αντιπροσωπεύουν διάφορα ένα μικρό τμήμα του φάσματος ηλεκτρομαγνητικής ακτινοβολίας. Για την λήψη τους, χρησιμοποιήθηκε ένα σύστημα δύο καμερών, μίας για το ορατό φάσμα και μίας για το εγγύς υπέρυθρο. Οι κάμερες αυτές λαμβάνουν την τιμή ενός τμήματος φάσματος μόνο για κάθε εικονοστοιχείο, αντί για όλα τα τμήματα για όλα τα εικονοστοιχεία. Λαμβάνουν όμως τις τιμές όλων των τμημάτων φάσματος για παράθυρα εικονοστοιχείων, 4×4 για το ορατό φάσμα και 5×5 για το εγγύς υπέρυθρο. Επομένως, οι λήψεις μετατρέπονται σε μία εικόνα χαμηλής ανάλυσης ($1/4$ της συνολικής ανάλυσης στην περίπτωση του ορατού φάσματος και $1/5$ στην περίπτωση του εγγύς υπέρυθρου). Η πολυφασματική εικόνα χαμηλής ανάλυσης συνδυάζεται με την πανχρωματική εικόνα υψηλής ανάλυσης μέσω της διαδικασίας *pansharpening*, και προκύπτει μία τελική υπερφασματική εικόνα.

Οι υπερφασματικές εικόνες αυτές αποθηκεύονται σε μορφή *.tif*, που αντιπροσωπεύει ψηφιδωτές εικόνες σε πληθώρα χρωματικών χώρων. Μέσω της προγραμματιστικής γλώσσας *python*, οι εικόνες μπορούν να μετατραπούν σε τρισδιάστατους πίνακες, όπου η τρίτη διάσταση αντιστοιχεί στα φασματικά κανάλια. Χρησιμοποιώντας κάποιο πρόγραμμα μπορούν να αποδοθούν 3 από τα κανάλια με τα χρωματικά κανάλια RGB. Θέτοντας τα ορατά φασματικά κανάλια του κόκκινου, του πράσινου και του μπλε της εικόνας στα αντίστοιχα χρωματικά κανάλια, απεικονίζεται μία προσέγγιση της φυσικής εμφάνισης του αντικειμένου (Εικόνα 22).²⁶

* (Arno Peeraer (2022) Google Maps)

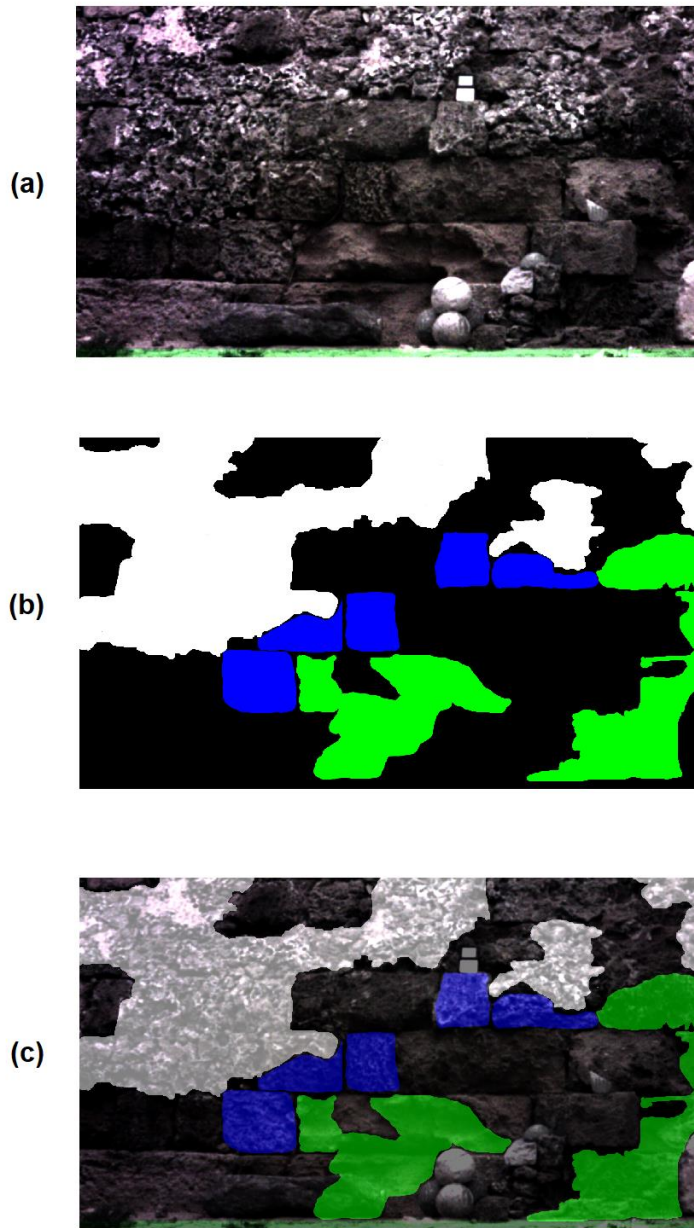


Εικόνα 22: Εικόνα λιθοδομής του φρουρίου Αγίου Νικολάου, όπως προκύπτει από τα φασματικά κανάλια κόκκινου, πράσινου και μπλε

Μαζί με τις εικόνες, δίνονται σε ορισμένες περιπτώσεις και οι αντίστοιχες εικόνες επισημάνσεων, όπου εμφανίζεται η βασική πραγματικότητα της ύπαρξης ή μη φθορών. Διαφορετικοί τύποι φθορών απεικονίζονται με διαφορετικά χρώματα, ενώ η περιοχή που απεικονίζεται με μαύρο θεωρείται ότι δεν έχουν φθαρεί ουσιαστικά. Οι εικόνες επισημάνσεων είναι απλές εικόνες με μία μόνο τιμή χρώματος ανά εικονοστοιχείο, άρα αποθηκεύονται σε format .png, ενώ για προγραμματιστικούς σκοπούς έχουν δοθεί και σε πίνακες numpy, μορφή πινάκων διαχειρίσιμων από τη γλώσσα προγραμματισμού python. Με υπέρθεση της εικόνας επισημάνσεων πάνω στην εικόνα από τα τρία φασματικά κανάλια, γίνεται κατανοητή η τοποθεσία των φθορών σε αυτή (Εικόνα 23).

Συνολικά οι πολυφασματικές εικόνες είναι 11. Για τις 6 από αυτές, δίνονται και οι αντίστοιχες εικόνες επισημάνσεων. Επομένως για την περίπτωση της επιβλεπόμενης μηχανικής μάθησης μέσω συνελκτικών δικτύων, κάποια τμήματα από τις 6 πολυφασματικές εικόνες μπορούν να χρησιμοποιηθούν σαν σύνολο δεδομένων εκπαίδευσης, τα υπόλοιπα από αυτά σαν σύνολο δεδομένων επικύρωσης ενώ οι 5 πολυφασματικές εικόνες για τις οποίες δεν υπάρχουν οι εικόνες επισημάνσεων μπορούν να χρησιμοποιηθούν μόνο ως σύνολο δοκιμής. Για την περίπτωση της μη επιβλεπόμενης μηχανικής μάθησης, οι 6 πολυφασματικές εικόνες για τις οποίες διατίθενται οι αντίστοιχες επισημασμένες μπορούν να ελεγχθούν ως προς τα αποτελέσματα της ομαδοποίησης και να προκύψουν οι δείκτες εξωτερικής εγκυρότητας, ενώ για τις υπόλοιπες 5 μπορούν να βρεθούν μόνο δείκτες εσωτερικής εγκυρότητας. Όσον αφορά την σύγκριση των δύο μεθόδων εκπαίδευσης, για τις 6 επισημασμένες εικόνες μπορούν να συγκριθούν άμεσα οι κοινοί δείκτες που βασίζονται στον πίνακα σύγχυσης, αν και στα νευρωνικά δίκτυα αφορούν ταξινόμηση παραθύρων της εικόνας ενώ στην ομαδοποίηση αφορούν ταξινόμηση εικονοστοιχείων. Οι υπόλοιπες 5 μπορούν να συγκριθούν μόνο οπτικά,

με βάση την ομαδοποίηση και το αποτέλεσμα του Class Activation Mapping, χωρίς να είναι δυνατός ο έλεγχος με βάση την πραγματική κατάσταση.



Εικόνα 23: Υπέρθυση επισημάνσεων πάνω στην φυσική εικόνα:

α) *rgb* εικόνα λιθοδομής, β) επισημάνσεις φθορών, γ) υπέρθεση των επισημάνσεων φθορών στην εικόνα *rgb*

Οι εικόνες και οι επισημάνσεις έχουν καταγραφεί κατά την διάρκεια εργασιών του HYPERION project. Το HYPERION είναι τμήμα του ευρωπαϊκού προγράμματος χρηματοδότησης έρευνας και καινοτομίας, και προσφέρει μία συνολική στρατηγική για την αποτροπή ζημιών σε μνημεία πολιτιστικής κληρονομιάς, την αναγνώριση και αποτροπή επιπρόσθετων απειλών και την ανοικοδόμηση και επαναφορά κανονικών συνθηκών σε ιστορικές τοποθεσίες. Συμπεριλαμβάνει την διαχείριση κινδύνων, την προστασία και την παραμονή σε ετοιμότητα σαν συμπληρωματικές στρατηγικές, καθώς και μακροπρόθεσμες

στρατηγικές προσεγγίσεις για την προσαρμογή στην κλιματική αλλαγή και τον ορισμό πολιτικών για οικονομική ανθεκτικότητα.

5.3 Προεπεξεργασία δεδομένων

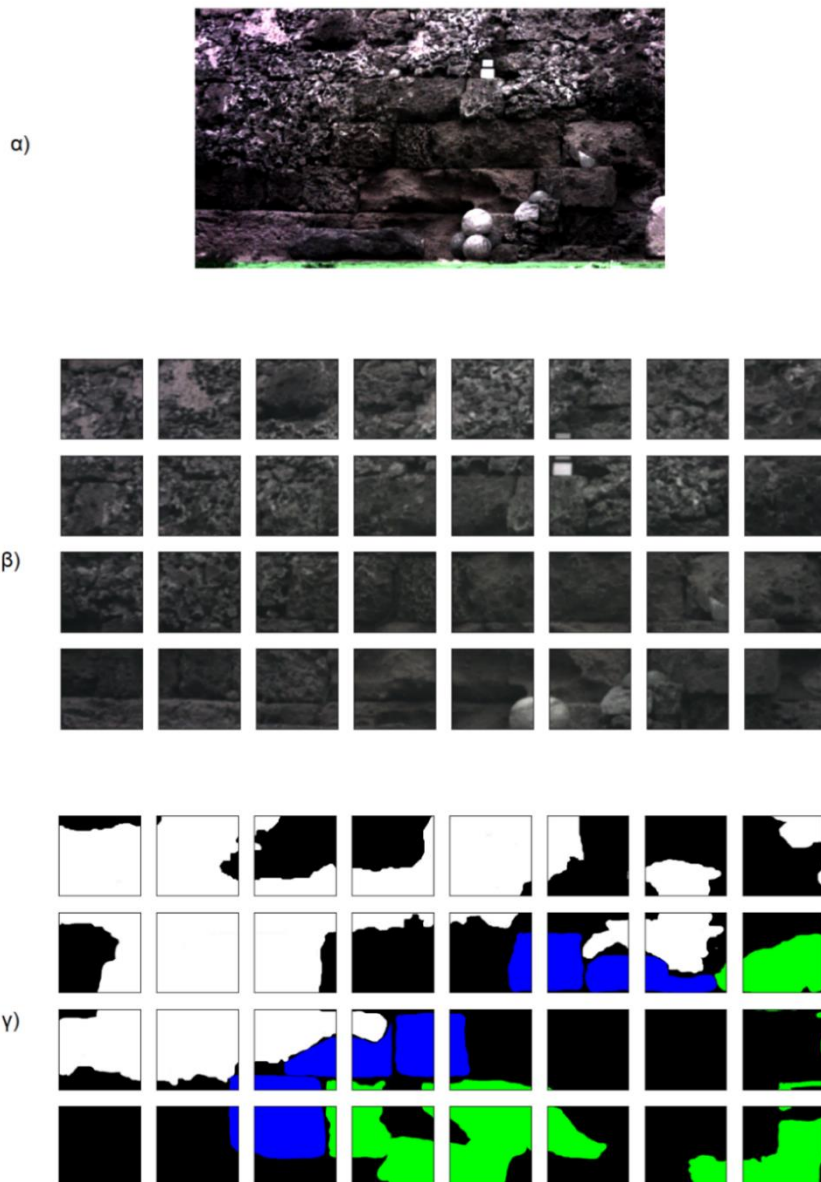
Η αρχιτεκτονική συνελκτικού νευρωνικού δικτύου VGG-16 δέχεται ως είσοδο εικόνες με 224×224 εικονοστοιχεία, στα 3 χρωματικά κανάλια RGB. Επομένως, τα δεδομένα πρέπει να μετασχηματιστούν, ώστε να έχουν τις διαστάσεις (224,224,3). Οι εικόνες του συνόλου δεδομένων που έχει περιγραφεί στο προηγούμενο εδάφιο, έχουν διαστάσεις 1032 εικονοστοιχεία σε ύψος, 1851 εικονοστοιχεία σε πλάτος, ενώ διαθέτουν 41 φασματικά κανάλια. Ο μετασχηματισμός τους για τις ανάγκες του νευρωνικού δικτύου περιλαμβάνει την μείωση των καναλιών της εικόνας, και την κατάτμηση της σε υποπεριοχές των 224×224 εικονοστοιχείων.

Για την αφαίρεση των καναλιών εκτός από 3, υλοποιήθηκε απλός κώδικας `rython`, ο οποίος δημιούργησε νέες εικόνες `.tif` αντιγράφοντας μόνο 3 από τα 41 κανάλια κάθε εικόνας. Επιλέχθηκε να δημιουργηθούν 2 υποσύνολα δεδομένων, ένα με τα κανάλια RGB των υπερφασματικών εικόνων και ένα χρησιμοποιώντας 3 κανάλια του εγγύς υπέρυθρου, τα 17, 27, και 37, ώστε να συγκριθούν τα αποτελέσματα της αρχιτεκτονικής συνελκτικού νευρωνικού δικτύου σε εικόνες που περιέχουν μόνο ορατό και μόνο υπέρυθρο φάσμα ακτινοβολίας (Παράρτημα, κώδικας, Αφαίρεση πρόσθετων καναλιών).

Όσον αφορά την κατάτμηση, οι εικόνες δεν μπορούν να διαιρεθούν τέλεια σε παράθυρα 224×224 , καθώς τόσο το ύψος των 1032 εικονοστοιχείων όσο και το πλάτος των 1851 εικονοστοιχείων δεν είναι πολλαπλάσια του 224. Για την δημιουργία των υποπαραθύρων με σκοπό την βέλτιστη παραγωγή δεδομένων εκπαίδευσης, παρουσιάζονται δύο επιλογές. Η πρώτη είναι κατάτμηση σε παράθυρα που περιλαμβάνουν ξεχωριστές υποπεριοχές της εικόνας, όπου κάθε παράθυρο ξεκινάει αμέσως μετά το όριο του προηγούμενου, για να καλυφθεί όλη η εικόνα. Λόγω των διαστάσεων της εικόνας όμως, με την επιλογή αυτή υπολείπονται ορισμένα εικονοστοιχεία, 59 κατά πλάτος και 136 κατά ύψος, τα οποία δεν θα συμμετέχουν στην εκπαίδευση με αυτόν τον τρόπο. Η δεύτερη επιλογή είναι η δημιουργία παραθύρων τα οποία θα συμπεριλαμβάνουν όλα τα εικονοστοιχεία της εικόνας, παρουσιάζοντας όμως επικάλυψη, δηλαδή εικονοστοιχεία ενός παραθύρου 224×224 θα εμφανίζονται και σε άλλα. Η επικάλυψη αυτή είναι 7 περίπου κατά πλάτος και 34 κατά ύψος. Ο κίνδυνος σε αυτήν την περίπτωση είναι ότι τα επικαλυπτόμενα εικονοστοιχεία, ορισμένα από τα οποία συμμετέχουν σε 4 διαφορετικά παράθυρα, έχουν πολύ μεγαλύτερο βάρος στον υπολογισμό των παραμέτρων του νευρωνικού δικτύου από τα υπόλοιπα. Ακόμα, η σύνθεση μίας συνολικής απεικόνισης από πολλούς τμηματικούς χάρτες CAM είναι πιο εύκολη όταν τα παράθυρα που χρησιμοποιούνται για εκπαίδευση είναι διακριτά, παρά όταν υπάρχουν επικαλυπτόμενα εικονοστοιχεία τα οποία παρουσιάζουν διαφορετική ένταση ενεργοποίησης σε κάθε παράθυρο που ανήκουν. Άρα υλοποιείται η πρώτη επιλογή, παρόλο που σημαίνει απώλεια εικονοστοιχείων και παραμέληση τους στα συμπεράσματα σχετικά με την εφαρμογή CNN στις εικόνες.

Αυτό πραγματοποιήθηκε μέσω απλού αλγόριθμου `rython`, ο οποίος χώρισε τις εικόνες σε 32 παράθυρα 224×224 . Αυτό έγινε για τις εικόνες με τα τρία κανάλια RGB, για τις εικόνες με τα τρία κανάλια του εγγύς υπέρυθρου (NIR) και για τις εικόνες επισημάνσεων (Εικόνα 24). Κάθε

παράθυρο ονομάστηκε με το όνομα της εικόνας από την οποία προήρθε, μαζί με έναν ξεχωριστό αριθμό από το 1 έως το 32, με τους αριθμούς να αυξάνονται πρώτα προς τα δεξιά και μετά προς τα κάτω (Παράρτημα, κώδικας, Κατάτμηση επισημάνσεων, Κατάτμηση tiff). Στη συνέχεια πραγματοποιήθηκε σύγκριση των παραθύρων που διαθέτουν 3 φασματικά κανάλια με τα αντίστοιχα παράθυρα επισημάνσεων, ώστε να κατηγοριοποιηθούν σε παράθυρα τα οποία απεικονίζουν φθορές και παράθυρα τα οποία δεν απεικονίζουν φθορές. Τα παράθυρα με ελάχιστη κάλυψη φθορών μέσα (μικρότερη από 15% της εικόνας), επιλέχθηκε να καταταχθούν σε αυτά χωρίς φθορές, εντάσσοντας έτσι ένα υποκειμενικό όριο απόφασης ως προς τι πρέπει να θεωρήσει ο αλγόριθμος μηχανικής μάθησης εικόνα με φθορά και τι όχι. Τέλος, τα παράθυρα που δημιουργήθηκαν από τις εικόνες χωρίς επισήμανση, λαμβάνονται ως το σύνολο δοκιμής.



Εικόνα 24: Κατάτμηση της εικόνας σε 32 παράθυρα για εκπαίδευση, και σύγκριση με τις επισημάνσεις:

α) Αρχική εικόνα με 3 κανάλια β) παράθυρα εκπαίδευσης γ) αντίστοιχα παράθυρα επισημάνσεων

Τα τελικά υποσύνολα των αρχικών δεδομένων διαμορφώθηκαν ως εξής. Από την διαδικασία της σύγκρισης με τα παράθυρα με τις επισημάνσεις, για την κατηγορία των εικόνων με RGB φασματικά κανάλια, το κοινό σύνολο εκπαίδευσης και επικύρωσης χωρίστηκε σε 46 εικόνες φθορών και 146 εικόνες χωρίς παρουσία φθορών. Το σύνολο δοκιμής αποτελείται από τις 5 αρχικές εικόνες × 32 παράθυρα, δηλαδή 160 εικόνες. Ομοίως, την ίδια κατανομή ακολουθούν και οι εικόνες των εγγύς υπέρυθρων καναλιών, οι οποίες θα χρησιμοποιηθούν για εκπαίδευση άλλου μοντέλου με την ίδια αρχιτεκτονική. Με αυτόν τον διαχωρισμό των παραθύρων, μπορεί να υλοποιηθεί αλγόριθμος εκπαίδευσης,

Ωστόσο, το σύνολο δεδομένων εκπαίδευσης εμφανίζει ανισορροπία, καθώς το 76% των παραθύρων εμφανίζουν φθορές και μόνο το 24% είναι αρνητικά ως προς την παρουσία φθορών. Το πρόβλημα σε αυτήν την περίπτωση είναι ότι το μοντέλο θα είναι ελαφρώς προκατειλημμένο σε εικόνες φθορών κατά την εκπαίδευση, και άρα είναι πιθανό να αποφασίζει πιο εύκολα πως μία εικόνα που δεν έχει ξαναδεί περιέχει φθορές. Οι συνήθεις τρόποι άρσης της ανισορροπίας αυτής, είναι ή η αφαίρεση από το σύνολο δεδομένων εικόνων της επικρατούσας κατηγορίας, δηλαδή των εικόνων με φθορές, ή η ενίσχυση του συνόλου δεδομένων με συνθετικά παραδείγματα. Αυτά προέρχονται από τις εικόνες χωρίς φθορές όταν εφαρμοστούν πάνω τους διάφοροι μετασχηματισμοί, όπως περιστροφή ή αντικατοπτρισμός. Η παράλειψη εικόνων θετικών ως προς την παρουσία φθορών δεν ενδείκνυται, λόγω του ήδη μικρού αριθμού παραδειγμάτων εκπαίδευσης. Επομένως, η καταλληλότερη λύση για εξισορρόπηση του συνόλου δεδομένων, είναι η ενίσχυση με συνθετικές εικόνες χωρίς φθορές. Η δημιουργία των συνθετικών εικόνων υλοποιήθηκε με απλούς αλγόριθμους `rython` για αριστερόστροφη περιστροφή 90 μοιρών και οριζόντια αναστροφή των `.tif` αρχείων ξεχωριστά (Παράρτημα, κώδικας, Ενίσχυση συνόλου δεδομένων χωρίς φθορές). Το τελικό σύνολο δεδομένων εκπαίδευσης και επικύρωσης, αποτελείται από 146 εικόνες με παρουσία φθορών (51%) και 138 εικόνες με απουσία φθορών (49%). Από αυτές, 35 από τις θετικά επισημασμένες και 35 από τις αρνητικά επισημασμένες, που αντιστοιχούν στο 25% των παραδειγμάτων εκπαίδευσης, λαμβάνονται ως σύνολο επικύρωσης (Παράρτημα, κώδικας, Διαχωρισμός `training set / validation set`).

Τέλος, για την εκμετάλλευση του συνόλου δεδομένων από προεκπαιδευμένα μοντέλα VGG-16 μέσω μεταφοράς μάθησης, οι εικόνες που περιγράφονται παραπάνω αποθηκεύονται σε μορφή `.png` μέσω αλγορίθμου `rython`. Οι τιμές των εικονοστοιχείων αντί για δεκαδική μορφή, από 0 έως 1, μετατρέπονται σε ακέραιους αριθμούς από 0 έως 255 (Παράρτημα, κώδικας, Αποθήκευση ως `.png` των ανωτέρω αρχείων).

6. Εφαρμογή Αλγορίθμων Μηχανικής Μάθησης

6.1 Επιβλεπόμενη Μάθηση με CNN

6.1.1 Λογισμικό Υλοποίησης

Όπως και οι διεργασίες στο στάδιο της προεπεξεργασίας των δεδομένων, έτσι και η εκπαίδευση του συνελκτικού νευρωνικού δικτύου πραγματοποιείται σε γλώσσα προγραμματισμού `python`. Η `python` είναι μία γλώσσα υψηλού επιπέδου, γενικού σκοπού, η οποία ως λογισμικό ανοιχτού κώδικα είναι δωρεάν διαθέσιμη για χρήση και διανομή, ακόμη και για εμπορικούς σκοπούς. Λόγω του μεγάλου αριθμού χρηστών της διαθέτει πληθώρα εξωτερικών βιβλιοθηκών για διάφορες χρήσεις, μεταξύ των οποίων και η ακαδημαϊκή / ερευνητική.

Λόγω της μεγάλης δημοφιλίας της ως γλώσσα, η `python` χρησιμοποιείται ευρέως και στο πεδίο της μηχανικής μάθησης, με πλήθος πλαϊσίων και βιβλιοθηκών που διατίθενται για τον σκοπό αυτόν. Η παρούσα εργασία και ο απαιτούμενος κώδικας για αυτήν στηρίζονται σε μεγάλο βαθμό στην εξωτερική βιβλιοθήκη `pytorch 1.10.0`, η οποία είναι επίσης λογισμικό ανοιχτού κώδικα. Η `pytorch` είναι μία βελτιστοποιημένη βιβλιοθήκη βαθιάς μάθησης, που λειτουργεί με τανυστές, δηλαδή γεωμετρικά αντικείμενα με μορφή διανυσμάτων. Κάθε σύνολο δεδομένων μετατρέπεται σε τανυστές, ώστε να χρησιμοποιηθεί για εκπαίδευση μοντέλου μηχανικής μάθησης. Οι πράξεις εύρεσης των βέλτιστων παραμέτρων εκτελούνται σε κάρτα γραφικών στην προκειμένη περίπτωση, η οποία προτιμάται για την μεγαλύτερη ταχύτητά της σε υπολογισμούς. Η `pytorch` απαιτεί συγκεκριμένο πρόγραμμα – οδηγό για τη λειτουργία του, το οποίο παρέχεται από την `nvidia` και ονομάζεται `cuda`. Η έκδοση `cuda` που χρησιμοποιήθηκε στην εργασία είναι η 11.5 .

Ακόμα, η συγγραφή του κώδικα έγινε στο ολοκληρωμένο περιβάλλον ανάπτυξης `PyCharm` της εταιρείας `JetBrains S.R.O.` Προσφέρει μεγάλη συλλογή από καθοριστικά εργαλεία για προγραμματιστές `python`, βαθιά ενσωματωμένα για να δημιουργήσουν ένα βολικό περιβάλλον για ανάπτυξη προγραμμάτων επιστήμης των δεδομένων. Η `pycharm` διαθέτει τρεις διαφορετικές εκδόσεις. Οι εκδόσεις `Community` και `Edu` αποτελούν λογισμικό ανοιχτού κώδικα με περιορισμούς και διατίθενται δωρεάν. Η έκδοση `Professional` διαθέτει πολλές περαιτέρω διευκολύνσεις και εργαλεία, και διατίθεται επί πληρωμή. Ο κώδικας της παρούσας εργασίας γράφηκε στο σύνολο του στην έκδοση `PyCharm Community Edition 2021.1.3`.

Όλα τα αναφερόμενα λογισμικά εφαρμόστηκαν σε λειτουργικό σύστημα `Windows 10`, έκδοση 64 bit.

6.1.2 Προδιαγραφές Ηλεκτρονικού Υπολογιστή

Πέρα από το χρησιμοποιηθέν λογισμικό, σημασία για την εκπαίδευση νευρωνικών δικτύων έχουν και τα εξαρτήματα από τα οποία αποτελείται ο ηλεκτρονικός υπολογιστής, στον οποίο πραγματοποιείται η διαδικασία. Η εκπαίδευση μπορεί να πραγματοποιηθεί τόσο σε

επεξεργαστές όσο και σε κάρτες γραφικών, με τις δεύτερες να είναι σημαντικά ταχύτερες. Στην περίπτωση που επιλέγεται κάρτα γραφικών για εκπαίδευση, το μέγεθος της μνήμης της καθορίζει το μέγεθος δεσμών (batch size) που μπορεί να χρησιμοποιηθεί, καθώς και πόσο μεγάλο συνολικά δίκτυο μπορεί να εκπαιδευτεί. Η μνήμη RAM δεν επιδρά άμεσα στην εκπαίδευση δικτύων, ωστόσο μεγάλη μνήμη επιτρέπει την ομαλή εκτέλεση κώδικα και την ταυτόχρονη χρήση άλλων προγραμμάτων. Τέλος, σημασία έχει και ο χρόνος ανάκτησης των δεδομένων για εκπαίδευση και η αποθήκευση/φόρτωση του ίδιου του μοντέλου, τα οποία βασίζονται στη λειτουργία του σκληρού δίσκου. Λαμβάνοντας υπόψιν τα παραπάνω, ο φορητός υπολογιστής στον οποίο διενεργήθηκε η μηχανική μάθηση έχει τα εξής χαρακτηριστικά:

| | |
|-------------------------|--|
| Επεξεργαστής | AMD Ryzen 5 4600H, 3.00 GHz |
| Αυτόνομη Κάρτα γραφικών | Nvidia GeForce GTX 1650 Ti laptop, 4.00 GB |
| Μνήμη RAM | 8.00 GB |
| Σκληρός δίσκος | SSD 512 GB, PCIe 3.0, NVMe |

Πίνακας 2: Hardware Υπολογιστή

6.1.3 Διαδικασία εκπαίδευσης

Η αρχιτεκτονική που έχει επιλεγεί για εκπαίδευση είναι η VGG-16 (βλ. 3.1.7), ενώ οι παράμετροι προέρχονται από την προεκπαίδευση του μοντέλου στη βάση δεδομένων imagenet για 1000 κατηγορίες. Το προεκπαιδευμένο μοντέλο διατίθεται από τη βιβλιοθήκη pytorch. Εφαρμόζεται μεταφορά μάθησης αξιοποιώντας τις παραμέτρους αυτές, και μεταβάλλοντας μερικές μόνο με εκπαίδευση πάνω στο σύνολο δεδομένων τις εργασίες για πρόβλεψη σε 2 πιθανές κατηγορίες μόνο (φθορά ή όχι φθορά). Ο κώδικας στηρίζεται σε μεγάλο βαθμό στο εκπαιδευτικό άρθρο για την χρήση μεταφοράς μάθησης στην όραση υπολογιστών με pytorch, του Sasank Chilamkurthy.⁵⁸ Πρώτα περιγράφεται κάθε βήμα του αλγορίθμου αναλυτικά, μετά παρατίθενται οι σχετικές γραμμές κώδικα γλώσσας python και επεξηγούνται αντίστοιχα στην ελληνική γλώσσα.

Αρχικά, χωρίζονται τα δεδομένα εκπαίδευσης κατάλληλα σε φακέλους. Πρώτα δημιουργείται ένας κύριος φάκελος με δύο υποφακέλους, train και val, για τα σύνολα εκπαίδευσης και επικύρωσης αντίστοιχα. Κάθε υποφάκελος από αυτούς περιέχει έναν φάκελο για κάθε κατηγορία ταξινόμησης, που στην περίπτωση της εργασίας είναι negative (για απουσία φθορών) και positive (για παρουσία φθορών). Επομένως ο αλγόριθμος αναγνωρίζει ποιες εικόνες ανήκουν στο σύνολο εκπαίδευσης, ποιες ανήκουν στο σύνολο επικύρωσης καθώς και τις επισημάνσεις τους με κατηγορίες, από το όνομα των φακέλων. Στη συνέχεια χτίζεται το σύνολο δεδομένων στην κατάλληλη για τον αλγόριθμο μορφή, και ορίζεται η ρουτίνα με την οποία αυτά εισάγονται στο μοντέλο μηχανικής μάθησης.


```
data_dir = 'C:/Diplomatiki/Dataset/AgiosNikolaos/trainfolder/RGB'
```

```
image_datasets = {x: datasets.ImageFolder(os.path.join(data_dir, x),  
data_transforms[x])  
for x in ['train', 'val']}
```

```
dataloaders = {x: torch.utils.data.DataLoader(image_datasets[x],  
batch_size=8, shuffle=True, num_workers=1)  
for x in ['train', 'val']}
```

Φάκελος δεδομένων = διεύθυνση κύριου φακέλου στον υπολογιστή

Σύνολα δεδομένων εικόνων = διεύθυνση υποφακέλων στον υπολογιστή (Φάκελος δεδομένων + train, val), προαιρετικοί μετασχηματισμοί, για train, val

Ρουτίνα εισαγωγής = (Σύνολα δεδομένων εικόνων, μέγεθος δέσμης = 8, ανακάτεμα εικόνων = Αληθές, παράλληλες διεργασίες εισαγωγής εικόνων = 1, για train, val)

Τα δεδομένα (και το μοντέλο παρακάτω) μεταφέρονται στην κάρτα γραφικών, από τον επεξεργαστή όπου βρίσκονται ως προεπιλογή. Στην συνέχεια ορίζεται η ρουτίνα εκπαίδευσης του μοντέλου, με βάση το ίδιο το μοντέλο, το κριτήριο βελτιστοποίησης, την ρουτίνα βελτιστοποίησης, και τις εποχές που θα οριστούν. Μία ιδιαιτερότητα στον ορισμό της διαδικασίας εκπαίδευσης, είναι ότι ο αλγόριθμος αποθηκεύει το μοντέλο με την καλύτερη επίδοση στο σύνολο επικύρωσης, ώστε ακόμα και αν γίνει χειρότερο σε επόμενες εποχές εκπαίδευσης λόγω overfitting, να διατηρηθεί η καλύτερη πιθανή εκδοχή του. Επομένως το θετικό της υλοποίησης αυτής είναι ότι ο αριθμός εποχών δεν επιδρά άμεσα στο πόσο καλό είναι το μοντέλο που προκύπτει, αφού δεν αποθηκεύονται απαραίτητα οι παράμετροι της τελευταίας εποχής. Το αρνητικό της όμως, είναι ότι μπορεί τυχαία η ακρίβεια του μοντέλου στα δεδομένα επικύρωσης στην αρχή να είναι πολύ υψηλή, ενώ στα δεδομένα εκπαίδευσης να είναι χαμηλή, και άρα να αποθηκευτεί ένα μη προσαρμοσμένο μοντέλο απλά επειδή έτυχε να έχει καλά αποτελέσματα επικύρωσης. Για να αντισταθμιστεί αυτός ο κίνδυνος, επιλέχθηκε στην εργασία αυτή να αποθηκεύεται η καλύτερη εκδοχή του μοντέλου μετά το πέρας 6 εποχών. Κατά την εκπαίδευση, εκτυπώνονται ο αριθμός της εποχής, τα αποτελέσματα της συνάρτησης απωλειών στα σύνολα εκπαίδευσης και επικύρωσης καθώς και η ακρίβεια σε καθένα από τα σύνολα. Μετά το τέλος της εκπαίδευσης, εκτυπώνεται η καλύτερη ακρίβεια επικύρωσης και σε ποια εποχή επιτεύχθηκε αυτή.

```
"use GPU"  
device = torch.device("cuda:0" if torch.cuda.is_available() else "cpu")  
print(device)
```

```
def train_model(model, criterion, optimizer, scheduler,  
num_epochs=25):
```

```
{  
for epoch in range(num_epochs):  
...
```

```
for phase in ['train', 'val']:  
for inputs, labels in dataloaders[phase]:  
inputs = inputs.to(device)  
labels = labels.to(device)
```

```
loss = criterion(outputs, labels)
```

```
if phase == 'train':  
loss.backward()
```

```
optimizer.step()
```

«Χρήση κάρτας γραφικών»
Μηχάνημα εκπαίδευσης = «κάρτα γραφικών» ΑΝ υπάρχει οδηγός κάρτας γραφικών, ΑΛΛΙΩΣ «επεξεργαστής»

Ορισμός εκπαίδευσης μοντέλου(μοντέλο, κριτήριο βελτιστοποίησης, ρουτίνα βελτιστοποίησης, χρονοπρογραμματιστής, μέγιστος αριθμός εποχών =25 από προεπιλογή)

```
{  
Για αριθμό εποχής του μέγιστου αριθμού εποχών:  
...
```

Για φάση [εκπαίδευσης, επικύρωσης]:
Για εικόνες, επισημάνσεις της Ρουτίνας εισαγωγής:
εικόνες προς Μηχάνημα εκπαίδευσης
επισημάνσεις προς Μηχάνημα εκπαίδευσης

απώλειες = κριτήριο βελτιστοποίησης(προβλέψεις, επισημάνσεις)

Αν φάση = εκπαίδευση
ανάστροφη αναδιάδοση σφαλμάτων στις παραμέτρους του μοντέλου
Βελτιστοποίηση παραμέτρων με τη ρουτίνα βελτιστοποίησης

```

if phase == 'val' and epoch_acc > best_acc and epoch > 5:

    best_acc = epoch_acc

    best_model_wts = copy.deepcopy(model.state_dict())

    best_epoch = epoch

    print(f'Epoch {epoch}/{num_epochs - 1}')
    print('-' * 10)
    print(f'{phase} Loss: {epoch_loss:.4f} Acc: {epoch_acc:.4f}')

print(f'Training complete in {...}s')
print(f'Best val Acc: {best_acc:.4f}')
print(f'Epoch: {best_epoch}')

model.load_state_dict(best_model_wts)
return model
}

```

ΑΝ φάση = επικύρωση ΚΑΙ ορθότητα προβλέψεων εποχής > καλύτερη ορθότητα προβλέψεων ΚΑΙ αριθμός εποχής > 5

καλύτερη ορθότητα = ορθότητα εποχής

καλύτερες παράμετροι μοντέλου =
αντιγραφή(παράμετροι μοντέλου τώρα)

αριθμός καλύτερης εποχής = αριθμός εποχής

Εκτύπωσε «Αριθμό εποχής / Συνολικό αριθμό εποχών», Εκτύπωσε «γραμμές -----»
Εκτύπωσε Απώλειες εποχής, ορθότητα εποχής

«Μετά το τέλος όλων των εποχών»

Εκτύπωσε διάρκεια εκπαίδευσης

Εκτύπωσε καλύτερη ορθότητα

Εκτύπωσε αριθμό καλύτερης εποχής

Επαναφορά καλύτερων παραμέτρων του μοντέλου

Παράδοση εκπαιδευμένου μοντέλου

}

ΤΕΛΟΣ ΟΡΙΣΜΟΥ ΕΚΠΑΙΔΕΥΣΗΣ

ΜΟΝΤΕΛΟΥ

Αφού έχουν οριστεί τα παραπάνω, επιλέγονται τα χαρακτηριστικά στοιχεία της εκπαίδευσης που εισάγονται στον ορισμό της. Για το μοντέλο των RGB φασματικών κανάλιων, τα παραδείγματα εκπαίδευσης φορτώθηκαν σε δέσμες των 8 εικόνων, δηλαδή υπολογίστηκε η συνάρτηση απωλειών συγκεντρωτικά από 8 εικόνες πριν μεταβληθούν οι παράμετροι. Η αρχιτεκτονική μεταβάλλεται στο τελευταίο πλήρως συνδεδεμένο επίπεδο, όπου αντί για 1000 κατηγορίες εξόδου τοποθετούνται μόνο 2. Πραγματοποιήθηκε πάγωμα των παραμέτρων των πρώτων 2 συνελικτικών μπλοκ του προεκπαιδευμένου μοντέλου VGG16, και αφέθηκαν μεταβλητά τα υπόλοιπα 3 συνελικτικά μπλοκ, δηλαδή συνολικά 9 συνελικτικά επίπεδα μαζί με ενδιάμεσα επίπεδα, καθώς και οι παράμετροι στα επίπεδα του ταξινομητή.

```

model_conv = torchvision.models.vgg16(pretrained=True)
for param in model_conv.parameters():
    param.requires_grad = False

for i in range(10,31):
    for param in model_conv.features[i].parameters():
        param.requires_grad = True

for param in model_conv.avgpool.parameters():

    param.requires_grad = True

model_conv.classifier[6] = nn.Linear(4096,num_classes)

model_conv = model_conv.to(device)

print(model_conv)

```

«Πάγωμα προεκπαιδευμένου μοντέλου»

Μοντέλο = VGG-16(προεκπαίδευση = αληθές)

Για κάθε παράμετρο στις παραμέτρους μοντέλου:

Υπολογισμός παραγώγου παράμετρου = ψευδής

«Ξεπάγωμα τελευταίων 3 συνελικτικών μπλοκ»

Για i από 10 μέχρι και 30:

Για παράμετρο στις παραμέτρους επιπέδου i

Υπολογισμός παραγώγου παραμέτρου = αληθής

Για κάθε παράμετρο στις παραμέτρους τελικού επιπέδου συγκέντρωσης

Υπολογισμός παραγώγου παραμέτρου = αληθής

«Μεταβολή τελευταίου πλήρως συνδεδεμένου επιπέδου»

επίπεδο ταξινομητή 6 = Συνδεδεμένο(4096 είσοδοι, αριθμός κατηγοριών έξοδοι (2, φθορά ή όχι φθορά))

Μοντέλο = Μοντέλο προς Μηχάνημα εκπαίδευσης

Εκτύπωσε Μοντέλο

Η τελική μορφή του μοντέλου τυπώνεται και είναι η εξής:

VGG(

(features): Sequential(
 (0): Conv2d(3, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
 (1): ReLU(inplace=True)
 (2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
 (3): ReLU(inplace=True)
 (4): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
 (5): Conv2d(64, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
 (6): ReLU(inplace=True)
 (7): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
 (8): ReLU(inplace=True)
 (9): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
 (10): Conv2d(128, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
 (11): ReLU(inplace=True)
 (12): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
 (13): ReLU(inplace=True)
 (14): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
 (15): ReLU(inplace=True)
 (16): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
 (17): Conv2d(256, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
 (18): ReLU(inplace=True)
 (19): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
 (20): ReLU(inplace=True)
 (21): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
 (22): ReLU(inplace=True)
 (23): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
 (24): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
 (25): ReLU(inplace=True)
 (26): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
 (27): ReLU(inplace=True)
 (28): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
 (29): ReLU(inplace=True)
 (30): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False))
(avgpool): AdaptiveAvgPool2d(output_size=(7, 7))
(classifier): Sequential(
 (0): Linear(in_features=25088, out_features=4096, bias=True)
 (1): ReLU(inplace=True)
 (2): Dropout(p=0.5, inplace=False)
 (3): Linear(in_features=4096, out_features=4096, bias=True)
 (4): ReLU(inplace=True)
 (5): Dropout(p=0.5, inplace=False)
 (6): Linear(in_features=4096, out_features=2, bias=True)

Όπου Conv2d είναι ένα συνελκτικό επίπεδο και kernel size το μέγεθος φίλτρου, stride το βήμα, padding το γέμισμα του περιθωρίου σε εικονοστοιχεία.

Σαν κριτήριο βελτιστοποίησης, χρησιμοποιήθηκε η ελαχιστοποίηση της συνάρτησης απώλειας εγκάρσιας εντροπίας (cross entropy loss) (35) η οποία ορίζεται ως:

$$\ell(x, y) = L = \{l_1, \dots, l_N\}^T, \quad l_n = \log \frac{e^{x_n, y_n}}{\sum_{c=1}^C e^{x_n, c}} \quad (35)$$

όπου N ο αριθμός παραδειγμάτων σε μία δέσμη (batch), n ένα από τα παραδείγματα, c κάθε πιθανή κατηγορία, C συνολικός αριθμός κατηγοριών, x_n η τιμή χαρακτηριστικών του και y_n

η τιμή επισήμανσης.⁵⁹ Επεξηγώντας την συνάρτηση απωλειών, η συνολική απώλεια είναι η μέση τιμή των απωλειών των παραδειγμάτων, με την απώλεια για ένα παράδειγμα να προκύπτει ως ο λογάριθμος του κλάσματος της εκθετικής συνάρτησης του αριθμού που δίνει ο αλγόριθμος πρόβλεψης στην σωστή κατηγορία προς το άθροισμα των εκθετικών συναρτήσεων των αριθμών που δίνει ο αλγόριθμος σε κάθε κατηγορία.

```
critterion = nn.CrossEntropyLoss()
```

Κριτήριο βελτιστοποίησης = ελαχιστοποίηση συνάρτησης απωλειών εγκάρσιας εντροπίας

Σαν ρουτίνα βελτιστοποίησης επιλέχθηκε η στοχαστική σύγκλιση με ελάττωση της παραγώγου (βλ. 3.1.2), η οποία αφού υπολογίσει την τιμή συνάρτησης απωλειών για μία δέσμη παραδειγμάτων, διορθώνει τις παραμέτρους σύμφωνα με τις μερικές παραγώγους τους ως προς αυτήν με σκοπό την μείωση των απωλειών. Στην στοχαστική σύγκλιση με ελάττωση παραγώγου υπεισέρχεται σαν προαιρετική παράμετρος η ορμή (momentum), η οποία σε κάθε διόρθωση παραμέτρων συμπεριλαμβάνει και ένα ποσοστό της μεταβολής κατά την προηγούμενη διόρθωση, ώστε να έχουν μία γενική κατεύθυνση οι διορθώσεις και να μην εξουδετερώνονται.

```
optimizer_conv = optim.SGD(model_conv.parameters(), lr=0.001, momentum=0.9)
```

Ρουτίνα βελτιστοποίησης = Στοχαστική σύγκλιση με ελάττωση παραγώγου (παράμετροι μοντέλου, ρυθμός μάθησης = 0.001, συντελεστής ορμής = 0.9)

Προσθέτως, υλοποιείται ένας χρονοπρογραμματιστής μάθησης (scheduler). Η λειτουργία του είναι να μειώνει τον ρυθμό μάθησης όσο αυξάνονται οι εποχές εκπαίδευσης. Η ιδέα πίσω από την τοποθέτηση ενός τέτοιου συστήματος είναι ότι αν ο ρυθμός μάθησης είναι μόνιμα ψηλός τότε κινδυνεύει να βρει την βέλτιστη λύση, αλλά στην επόμενη δέσμη να μεταβληθεί σε τέτοιο βαθμό που να την χάσει πάλι, ενώ αν είναι μόνιμα χαμηλός κινδυνεύει να μην βρει ικανοποιητική λύση ποτέ, λόγω χαμηλής κινητικότητάς του. Εφαρμόζοντας τον χρονοπρογραμματιστή, υπάρχει η ελπίδα να βρει ικανοποιητική λύση όσο είναι μεγάλος ο ρυθμός μάθησης, και στη συνέχεια όσο προχωράνε οι εποχές να συνεχίζει η εκπαίδευση με χαμηλό ρυθμό μάθησης για πλήρη βελτιστοποίηση της λύσης.⁶⁰ Στην εκπαίδευση αυτή, το βήμα ορίστηκε στις 7 εποχές και ο νέος ρυθμός μάθησης ορίστηκε στο 10% του προηγούμενου. Επομένως κάθε 7 εποχές, ο ρυθμός μάθησης πολλαπλασιάζεται με 0,1 .

```
exp_lr_scheduler = lr_scheduler.StepLR(optimizer_conv, step_size=7, gamma=0.1)
```

Χρονοπρογραμματιστής = χρονοπρογραμματιστής(ρουτίνα βελτιστοποίησης, βήμα σε εποχές = 7, συντελεστής = 0,1)

Η εκπαίδευση πραγματοποιήθηκε με υπερπαραμέτρους αριθμού εποχών μάθησης ίσο με 15 και ρυθμό μάθησης ίσο με 0,001 . Το καλύτερο αποτέλεσμα του μοντέλου για τα RGB φασματικά κανάλια προέκυψε στην 10^η εποχή εκπαίδευσης, και είχε ακρίβεια 93% στο σύνολο εκπαίδευσης και 87% στο σύνολο επικύρωσης. Η αρχιτεκτονική του μοντέλου καθώς και τα βάρη του αποθηκεύτηκαν για χρήση σε μελλοντικές προβλέψεις.

```
model_conv = train_model(model_conv, criterion, optimizer_conv,
                          exp_lr_scheduler, num_epochs=15)

torch.save(model_conv, 'RGBdeteriorationVGG16.pth')
```

«Κλήση της συνάρτησης ορισμού του μοντέλου, με εισόδους τις παραπάνω επιλογές»

Μοντέλο = εκπαίδευση μοντέλου (Μοντέλο, κριτήριο βελτιστοποίησης, ρουτίνα βελτιστοποίησης, χρονοπρογραμματιστής, αριθμός εποχών = 15)

Αποθήκευση μοντέλου(Μοντέλο, αρχείο στον υπολογιστή)

Τα αποτελέσματα που τυπώνει ο αλγόριθμος στην οθόνη έχουν την παρακάτω μορφή:

Epoch 0/14

train Loss: 0.7036 Acc: 0.5748
val Loss: 0.6467 Acc: 0.6429

...

Epoch 14/14

train Loss: 0.0739 Acc: 0.9813
val Loss: 0.2844 Acc: 0.8714

Training complete in 5m 26s

Best val Acc: 0.871428

Epoch: 9

Όπου για κάθε εποχή τυπώνει τιμή συνάρτησης απώλειας και ορθότητα προβλέψεων, τόσο για την φάση εκπαίδευσης με το σύνολο εκπαίδευσης όσο και την φάση επικύρωσης με το σύνολο επικύρωσης.

Όσον αφορά το μοντέλο για τα εγγύς υπέρυθρα φασματικά κανάλια, το οποίο εκπαιδεύεται με εικόνες χαμηλότερης ανάλυσης, επιλέχθηκαν 4 εικόνες ανά δέσμη εκπαίδευσης, καθώς με 8 εικόνες ανά δέσμη το μοντέλο παρουσίαζε πολύ χαμηλή ορθότητα στο σύνολο εκπαίδευσης. Στην ίδια αρχιτεκτονική μοντέλου, παγώθηκαν όλες οι παράμετροι εκτός από το τελευταίο συνελικτικό μπλοκ και έπειτα, δηλαδή τα 3 συνελικτικά επίπεδα και τα πλήρως συνδεδεμένα επίπεδα ταξινόμησης. Διατηρώντας ίδιο αριθμό εποχών και ρυθμό μάθησης, το καλύτερο αποτέλεσμα του μοντέλου ήταν στην 8^η εποχή, με ακρίβεια εκπαίδευσης 98,6% και ακρίβεια επικύρωσης μόλις 80%. Τα μεγέθη αυτά υποδεικνύουν υπερπροσαρμογή του μοντέλου, καθώς εμφανίζουν πολύ υψηλή ακρίβεια στα δεδομένα που έχει δει το μοντέλο και σαφώς χαμηλότερη σε αυτά που δεν έχει δει. Η πιθανή αιτία για αυτό είναι ότι το μοντέλο είναι υπερβολικά περίπλοκο για να αναγνωρίσει στοιχεία σε εικόνες χαμηλής ανάλυσης, και εντοπίζει στοιχεία χωρίς άμεση φυσική σημασία για να αποφασίσει την κατηγορία των εικόνων. Παρά την χαμηλή ποιότητα του μοντέλου, αποθηκεύτηκε για να εφαρμοστεί στο σύνολο δοκιμής, και να συγκριθούν τα αποτελέσματα του με τις άλλες μεθόδους. Ο κώδικας εκπαίδευσης αυτού του μοντέλου, ο οποίος ακολουθεί την δομή εκπαίδευσης του μοντέλου RGB, παρατίθεται πλήρως στο τέλος της εργασίας (Παράρτημα, κώδικας, Μεταφορά μάθησης και εκπαίδευση τροποποιημένου VGG16).

Όταν έχουν αποθηκευτεί τα μοντέλα, μπορούν να χρησιμοποιηθούν για προβλέψεις, καθώς και να εφαρμοστεί το Class Activation Mapping για να εντοπιστούν τα στοιχεία που οδήγησαν στην κάθε πρόβλεψη. Η υλοποίηση του CAM γίνεται μέσω της βιβλιοθήκης torchcam, και ο κώδικας βασίζεται σε μεγάλο βαθμό στον κώδικα του δημιουργού της βιβλιοθήκης, Francois-Guillaume Fernandez.

Η αρχική μέθοδος CAM (βλ. 4.2.2) δεν μπορεί να εφαρμοστεί σε δίκτυα με πλήρως συνδεδεμένα επίπεδα, όπως αυτά που εκπαιδεύτηκαν στην εργασία, επειδή απαιτεί ένα παγκόσμιο επίπεδο συγκέντρωσης αριθμητικού μέσου στο τέλος. Ωστόσο, υπάρχει πληθώρα τροποποιημένων αλγορίθμων CAM, από δημοσιεύσεις που έχουν στηριχθεί στην αρχική μέθοδο. Αυτές περιλαμβάνουν τα score-CAM, SS-CAM, IS-CAM που δουλεύουν επίσης με άμεση χρήση των βαρών ενεργοποίησης στο τελευταίο επίπεδο συγκέντρωσης, και τα Grad-CAM, Grad-CAM++, Smooth-Grad-CAM++, XGrad-CAM και Layer-CAM, τα οποία αξιοποιούν τις μερικές παραγώγους των παραμέτρων σε συνελκτικά επίπεδα, και άρα μπορούν να παραλάβουν βάρη ενεργοποίησης ακόμα και μέσα από πλήρως συνδεδεμένα επίπεδα.⁶¹

Ο κώδικας της εργασίας χρησιμοποιεί την μέθοδο CAM ως εξής. Αρχικά φορτώνεται ένα από τα εκπαιδευμένα μοντέλα της προηγούμενης παραγράφου, το οποίο θα χρησιμοποιηθεί μόνο για προβλέψεις και άρα δεν θα μεταβληθεί. Στη συνέχεια ορίζεται ποια από τις παραπάνω μεθόδους CAM θα εφαρμοστεί, και σε ποιο συνελκτικό επίπεδο. Κάθε παράθυρο 224x224 εισέρχεται στο μοντέλο και προκύπτει το διάνυσμα των 2 κατηγοριών, μη φθορά ή φθορά. Ανάλογα με τις τιμές του διανύσματος προβλέπεται σε ποια κατηγορία ανήκει το παράθυρο, και το διάνυσμα εισέρχεται στην επιλεγμένη μέθοδο CAM μαζί με τον αριθμό κατηγορίας που είναι επιθυμητό να εντοπιστεί σε κάθε εικόνα, δηλαδή η φθορά. Τέλος, τυπώνεται η προβλεπόμενη κατηγορία του παραθύρου, η οποία χρησιμοποιείται για σύγκριση με την πραγματική κατηγορία του παραθύρου για να προκύψει ο πίνακας σύγχυσης και οι δείκτες ταξινόμησης precision (2), recall (3), accuracy (5), F1 (4). Ομοίως τυπώνεται και αποθηκεύεται το παράθυρο με υπέρθεση του θερμικού χάρτη από το CAM. Τα αποθηκευμένα παράθυρα της εικόνας μετά ενώνονται σε μία ενιαία εικόνα που απεικονίζει τον εντοπισμό των φθορών από το μοντέλο, μέσω CAM.

```
model = torch.load('RGBdeteriorationVGG16.pth')
model.to(torch.device('cuda'))
model = model.eval()
print(model)
cam_extractor = GradCAMpp(model, target_layer= model.avgpool)
...
for imgIndx in range(0, num_imgs):
    image = read_image(imgPath, ImageReadMode.RGB)
    input_tensor = normalize(resize(image, (224, 224))...)
    input_tensor = input_tensor.to(torch.device('cuda'))
    out = model(input_tensor.unsqueeze(0))
```

Μοντέλο = Φόρτωση από μνήμη(αρχείο στον υπολογιστή)
Μοντέλο προς Μηχάνημα εκπαίδευσης (Κάρτα Γραφικών)
Μοντέλο = Φάση επικύρωσης (Αμετάβλητες παράμετροι)
Εκτύπωσε Μοντέλο
Μέθοδος CAM = Grad-CAM++ (Μοντέλο, επίπεδο εφαρμογής = τελευταίο συγκεντρωτικό επίπεδο)
...
Για κάθε αριθμό παραθύρου στις εικόνες
Εικόνα = διάβασε εικόνα(διεύθυνση εικόνας στη μνήμη, τρόπος διαβάσματος εικόνας RGB)
τανυστής = κανονικοποίηση(αλλαγή διαστάσεων εικόνας σε 224x224 εικονοστοιχεία, ...)
τανυστής = τανυστής προς Μηχάνημα εκπαίδευσης (Κάρτα Γραφικών)
διάνυσμα προβλέψεων κατηγοριών = Εφαρμογή μοντέλου(τανυστής)

Με δοκιμές και σύγκριση με τις επισημάνσεις, καταλληλότερη μέθοδος CAM αποδείχτηκε η Grad-CAM++ η οποία και επιλέχθηκε να εφαρμοστεί στα παράθυρα των εικόνων. Ο λόγος που δεν εφαρμόζεται η μέθοδος CAM στις ενιαίες εικόνες, είναι ότι θα πρέπει η ενιαία εικόνα να μετασχηματιστεί σε παράθυρο 224x224, και πάνω σε αυτό να γίνει ο εντοπισμός, το οποίο δεδομένου του μεγέθους των εικόνων σημαίνει απώλεια πληροφορίας, και μη ακριβή εντοπισμό. Αντίθετα, με παράθυρα μεγέθους 224x224 εικονοστοιχεία, θεωρητικά ένα σωστό μοντέλο μπορεί να υποδείξει τις φθορές με τη μεγαλύτερη πιθανή ακρίβεια. Το επίπεδο του οποίου οι χάρτες χαρακτηριστικών θα συνθέσουν τον θερμικό χάρτη, επιλέχθηκε να είναι το επίπεδο συγκέντρωσης αριθμητικού μέσου που βρίσκεται ακριβώς πριν τα πλήρως συνδεδεμένα επίπεδα των μοντέλων. Ο λόγος επιλογής του τελευταίου επιπέδου, είναι ότι οι μέθοδοι CAM που βασίζονται στις παραγώγους βγάζουν χειρότερα αποτελέσματα, όσο πιο βαθιά βρίσκεται το στοχευόμενο επίπεδο. Χρησιμοποιώντας τις εκροές του μοντέλου και θέτοντας την επιθυμητή κατηγορία εντοπισμού ίση με 1, θα εμφανιστούν σε όλα τα παράθυρα τα μοτίβα που αναγνώρισε ο αλγόριθμος ως φθορά, ανεξάρτητα από την τελική πρόβλεψη. Τέλος, ο θερμικός χάρτης υπερτίθεται πάνω στην εικόνα του παραθύρου με διαφάνεια 80% (Παράρτημα, κώδικας, Εφαρμογή CAM σε προβλέψεις των μοντέλων).

```

for imgIndx in range(0, num_imgs):
    ...

    activation_map = cam_extractor(scores=out, class_idx=1)

    result = overlay_mask(to_pil_image(image),
                        to_pil_image(activation_map[0].squeeze(0),
                                    mode='F'), alpha=0.8)

plt.imshow(result)
...
plt.savefig(...)

```

Για κάθε αριθμό παραθύρου στις εικόνες:

...

Θερμικός χάρτης = Μέθοδος CAM(τιμές προβλέψεων ανά κατηγορία = διάνυσμα προβλέψεων κατηγοριών, αριθμός κατηγορίας = 1 (φθορά))

«Υπέρθωση θερμικού χάρτη στο αρχικό παράθυρο»

Αποτέλεσμα = Υπέρθωση(μορφή εικόνας(Εικόνα), μορφή εικόνας(Θερμικός χάρτης, εικονοστοιχεία θερμικού χάρτη = δεκαδική μορφή), διαφάνεια=0,8)

Εκτύπωση (Αποτελέσματος Υπέρθωσης)

...

Αποθήκευση (Αποτελέσματος Υπέρθωσης)

Στη συνέχεια ενώνονται όλα τα παράθυρα που έχουν απεικόνιση CAM, ώστε να συντεθεί μία συνολική εικόνα CAM με πιο λεπτομερή εντοπισμό φθορών, από όση θα διέθετε αν εφαρμοζόταν το CAM στην αρχική ενιαία εικόνα. Αυτό πραγματοποιείται για καθεμία από τις 5 εικόνες χωρίς επισημάνσεις και 6 εικόνες με επισημάνσεις (Παράρτημα, κώδικας, Συνένωση τμημάτων CAM).

```

for RGBimage in Images:
    CAMsegments = []
    columns = 8
    rows = 4
    for i in range(1, columns * rows + 1):
        img = Image.open(FILEPATH+RGBimage+"_" + str(i) + ".png")
        CAMsegments.append(img)

```

«Ένταξη 32 παραθύρων με θερμικούς χάρτες σε πίνακα»

Για κάθε ενιαία εικόνα σε φάκελο εικόνων

Παράθυρα εικόνας = κενός πίνακας []

Στήλες = 8

Γραμμές = 4

Για i από 1 έως γραμμές επί στήλες (=32):

Παράθυρο = άνοιγμα Αποτελέσματος υπέρθεσης Προσθήκη παραθύρου στον πίνακα παραθύρων εικόνας

```

for RGBimage in Images:
    ...
    width,height = CAMsegments[0].size

    NewImage = Image.new('RGB',(width*8,height*4))

    i=0
    for row in range(0,height*4,height):
        for column in range(0,width*8,width):
            NewImage.paste(CAMsegments[i],[column, row])

            i+=1

    NewImage.save(FILEPATH+'ImageCAM'+RGBimage+'.png')

```

«Επικόλληση των 32 τμημάτων με σωστή σειρά σε νέα κενή εικόνα»

Για κάθε ενιαία εικόνα σε φάκελο εικόνων

...
 Πλάτος, ύψος = διαστάσεις παραθύρου

Νέα Εικόνα = κενή εικόνα(RGB, διαστάσεις
 (πλάτος παραθύρου*8, ύψος παραθύρου*4)

i=0
 για γραμμή από 0 έως (ύψος*4) με βήμα ύψος
 για στήλη από 0 έως (πλάτος*8) με βήμα πλάτος
 Επικόλληση στη Νέα εικόνα (παραθύρου[i] ,
 συντεταγμένες (X = στήλη, Y = γραμμή))
 i = i +1

Αποθήκευση Νέας εικόνας(Διεύθυνση μνήμης H/Y)

6.2 Μη Επιβλεπόμενη Μάθηση με Ομαδοποίηση

6.2.1 Λογισμικό υλοποίησης

Καθώς η βιβλιοθήκη pytorch διαθέτει εργαλεία κυρίως για επιβλεπόμενη μάθηση, για την μη επιβλεπόμενη μηχανική μάθηση με αλγόριθμους ομαδοποίησης επιλέχθηκε η βιβλιοθήκη scikit-learn 1.1.2. Η scikit-learn είναι εξωτερική βιβλιοθήκη της rythop και ενσωματώνει μεγάλη ποικιλία από σύγχρονους αλγόριθμους για μεσαίας κλίμακας προβλήματα επιβλεπόμενης και μη επιβλεπόμενης μηχανικής μάθησης. Το πλαίσιο αυτό στοχεύει να κάνει προσιτή τη μηχανική μάθηση σε μη ειδικούς, χρησιμοποιώντας μία υψηλού επιπέδου προγραμματιστική γλώσσα για γενικούς σκοπούς. Δίνει έμφαση στην ευκολία χρήσης, την απόδοση, την εποπτικότητα και την συνέπεια των διεπαφών με τον κώδικα. Έχει ελάχιστες εξαρτήσεις από άλλες βιβλιοθήκες και εκδίδεται κάτω από απλή άδεια ελεύθερου λογισμικού.⁶²

6.2.2 Διαδικασία ομαδοποίησης

Οι πλήρεις εικόνες είναι αρχεία .tif με διαστάσεις 1032 εικονοστοιχεία ύψος, 1851 εικονοστοιχεία πλάτος και 41 φασματικά κανάλια. Το επιθυμητό αποτέλεσμα μέσω της ομαδοποίησης, είναι νέες εικόνες οι οποίες έχουν τις ίδιες διαστάσεις ύψους και πλάτους, αλλά μόνο 3 κανάλια RGB, τα οποία εμφανίζουν χρώματα ανάλογα με ποια κατηγορία ανήκει το κάθε εικονοστοιχείο. Η επεξεργασία της αρχικής εικόνας ώστε να μπορεί να εκτελεστεί η ομαδοποίηση, στηρίζεται στον κώδικα του παραδείγματος κατάτμησης εικόνων RGB με rythop και εφαρμογή K-means της Bala Priya C.⁶³ Μετά την εφαρμογή της ομαδοποίησης με αλγόριθμο K-means με τη βιβλιοθήκη scikit-learn, η ανοικοδόμηση των ταξινομημένων εικονοστοιχείων βασίστηκε στον κώδικα του Huynh Ngoc Anh για κατάτμηση εικόνων.⁶⁴

Η ομαδοποίηση συμβαίνει για κάθε μεμονωμένη εικόνα ξεχωριστά. Αρχικά, κάθε εικονοστοιχείο που περιέχει 41 κανάλια, χρησιμοποιείται ως ένα μόνο παράδειγμα προς ταξινόμηση. Επομένως, η εικόνα 1032x1851x41 μετασχηματίζεται σε μορφή 1.910.323 x 41,

και το σύνολο δεδομένων για την πράξη ομαδοποίησης είναι 1.910.323 παραδείγματα με διάνυσμα χαρακτηριστικών τις τιμές του παραδείγματος στα 41 φασματικά κανάλια. Το διάνυσμα που περιέχει τα διανύσματα χαρακτηριστικών των παραδειγμάτων εισέρχεται στον αλγόριθμο ομαδοποίησης K-means, όπως αυτός περιγράφεται σε προηγούμενο εδάφιο (βλ. 3.2.2) και υλοποιείται στην βιβλιοθήκη scikit-learn. Ο αριθμός που επιλέχθηκε για τα κεντροειδή ομαδοποίησης είναι 4, με την ελπίδα ότι αυτά θα διαχωρίσουν τις κατηγορίες φθορών, τοιχωμάτων χωρίς φθορές, περιοχές με σκιές και έδαφος/ουρανό που απεικονίζονται σε κάποιες από αυτές. Αφού εφαρμοστεί ο αλγόριθμος ομαδοποίησης, υπολογίζονται επί τόπου και οι δείκτες εσωτερικής εγκυρότητας της ομαδοποίησης Calinski-Harabasz (32) και Davies-Bouldin (33).

```

for imgIndx in range(0, num_imgs):
    ...
    img = tiff.imread(imgPath)

    (h, w, c) = img.shape

    img2D = img.reshape(h * w, c)

    kmeans_model = KMeans(n_clusters=4, random_state=1)

    cluster_labels = kmeans_model.fit_predict(img2D)

    print("Calinski-Harabasz:" +
          str(metrics.calinski_harabasz_score(img2D, cluster_labels)))

    print("Davies-Bouldin:" +
          str(metrics.davies_bouldin_score(img2D, cluster_labels)))

```

Για κάθε αριθμό εικόνας στον φάκελο εικόνων
 ...
 Εικόνα = διάβασε εικόνα (Διεύθυνση εικόνας στη μνήμη H/Y)

(ύψος, πλάτος, κανάλια) = διαστάσεις εικόνας
 «ύψος = 1032, πλάτος = 1851, κανάλια = 41»

Εικόνα2D = μετασχηματισμός Εικόνας
 (ύψος * πλάτος, κανάλια)
 «ύψος = 1.910.323, κανάλια = 41»

Μοντέλο Kmeans = Kmeans(αριθμός κεντροειδών=4, τυχαία κατάσταση = 1)

«Η συνάρτηση Kmeans ορίζεται εντός της βιβλιοθήκης scikit-learn, η τυχαία κατάσταση ορίζει ένα πρότυπο τυχαίας τοποθέτησης κεντροειδών, ώστε να μπορεί να ξαναεπιτευχθεί αυτή»

λίστα κατηγοριών = πρόβλεψη Μοντέλου Kmeans (Εικόνα2D)

Εκτύπωση (Υπολογισμός CHI (Εικόνα2D, λίστα κατηγοριών)

Εκτύπωση (Υπολογισμός DBI (Εικόνα2D, λίστα κατηγοριών)

Με κάθε παράδειγμα να έχει πλέον μία ομάδα, χτίζεται η λίστα από ομαδοποιημένα παραδείγματα και τυπώνεται πόσα από αυτά ανήκουν σε κάθε ομάδα. Στη συνέχεια, μετασχηματίζεται η λίστα στις αρχικές διαστάσεις ύψους και πλάτους, ώστε πλέον κάθε εικονοστοιχείο της αρχικής εικόνας να αντιστοιχείται σε ένα στοιχείο με αριθμό ομάδας στον νέο αυτόν δισδιάστατο πίνακα. Ο πίνακας με τον αριθμό ομάδας κάθε εικονοστοιχείου αποθηκεύεται σε μορφή πίνακα numpy, .npy, για σύγκριση με τους πίνακες numpy των επισημάνσεων. Ακόμα, για οπτική σύγκριση, αποθηκεύεται ως εικόνα .png με διαφορετικό χρώμα για κάθε ομάδα, με την ομάδα με αριθμό 0 να απεικονίζεται μαύρη, την ομάδα με αριθμό 1 να απεικονίζεται μωβ, την ομάδα με αριθμό 2 να απεικονίζεται με πορτοκαλί και την ομάδα με αριθμό 3 να απεικονίζεται με λευκό (Παράρτημα, κώδικας, Εφαρμογή αλγόριθμου ομαδοποίησης και υπολογισμός εσωτερικών δεικτών).

```

for imgIndx in range(0, num_imgs):
    ...

    labels_count = Counter(cluster_labels)

    print(labels_count)

    label_array = np.uint8(cluster_labels.reshape(img.shape[:2]))

    print(label_array.shape)

    ...

    np.save(TIFF_PATH + ".npz/" + img_name, label_array)

    ...

plt.imsave(fname=TIFF_PATH + "clustered/" + img_name +
            '.png', arr=label_array, cmap='CMRmap', format='png')

```

Για κάθε αριθμό εικόνας στον φάκελο εικόνων:

Αριθμός κατηγοριών = μετρητής(Λίστα κατηγοριών)
 Εκτύπωση (Αριθμό κατηγοριών)

Πίνακας κατηγοριών = ακέραιοι(μετασχηματισμός Λίστας κατηγοριών (αρχικό ύψος πλάτος εικόνας)

Εκτύπωση(διαστάσεις Πίνακα κατηγοριών)
 «1032,1851»

«αποθήκευση πίνακα ως .npz»
 Αποθήκευση(Διεύθυνση μνήμης H/Y, Πίνακας κατηγοριών)

«αποθήκευση πίνακα ως εικόνα»
 Αποθήκευση εικόνας(Διεύθυνση μνήμης H/Y, Πίνακας = Πίνακας κατηγοριών, χρωματική κλίμακα = CMRκλίμακα, τύπος αρχείου = .png)

Ακολουθεί το στάδιο σύγκρισης των αποτελεσμάτων ομαδοποίησης με τις εικόνες επισημάνσεων, που θεωρείται ότι δηλώνουν την βασική πραγματικότητα της ύπαρξης φθορών. Στις εικόνες επισημάνσεων, με μαύρο χρώμα εμφανίζονται τα μη φθαρμένα δομικά στοιχεία, και με λευκό, πράσινο, κόκκινο και μπλε τα φθαρμένα δομικά στοιχεία με διαφορετικούς τύπους φθοράς ανά χρώμα. Για κάθε εικονοστοιχείο, γίνεται η σύγκριση μεταξύ της παρουσίας φθοράς στην ομαδοποίηση και στην βασική πραγματικότητα, ώστε να προκύψει ο πίνακας σύγχυσης.

Αυτό επιτυγχάνεται με την μετατροπή τόσο του πίνακα επισημάνσεων για κάθε εικόνα, όσο και του πίνακα των αριθμών ομάδων κάθε εικόνας, σε δυαδικούς πίνακες θετικών ή αρνητικών σε παρουσία φθοράς στοιχείων. Στους πίνακες επισημάνσεων, τα εικονοστοιχεία με μαύρο χρώμα ορίζονται ως False, δηλαδή μη έχοντα φθορές, και όλα τα υπόλοιπα ορίζονται ως True, δηλαδή έχοντα φθορές. Στην περίπτωση της ομαδοποίησης εικόνων, ο διαχωρισμός δεν είναι τόσο απλός και πραγματοποιείται με διαφορετικό τρόπο για κάθε εικόνα.

```

Annotation_array = np.load('C:/Diplomatiki/...npz', ...)

```

```

Annotation_array=Annotation_array.astype(bool)

```

```

print(Annotation_array.shape)

```

Πίνακας επισημάνσεων = Φόρτωση(Θέση μνήμης H/Y, άλλες προεπιλεγμένες παράμετροι)

Πίνακας επισημάνσεων = Πίνακας επισημάνσεων σε μορφή (Ψευδής/Αληθής)

«Τιμές του πίνακα = 0 γίνονται ψευδείς, τιμές του πίνακα > 0 γίνονται αληθείς»

Εκτύπωση(διαστάσεις Πίνακα επισημάνσεων)

Επειδή η ομαδοποίηση πραγματοποιήθηκε για κάθε εικόνα ξεχωριστά, και επειδή ο αλγόριθμος K-means επιλέγει τυχαία τις θέσεις των αρχικών κεντροειδών στον διανυσματικό χώρο 41 διαστάσεων, τα εικονοστοιχεία που κατατάσσονται στην ίδια ομάδα μπορεί να απεικονίζουν εντελώς διαφορετικές φασματικές οντότητες σε διαφορετικές φωτογραφίες. Απλούστερα, εικονοστοιχεία που ανήκουν στην ομάδα 0 και απεικονίζονται με μαύρο, μπορεί υποθετικά στην μία εικόνα να απεικονίζουν φθορές και στην άλλη να απεικονίζουν υγιή δομή. Για να οριστεί ποιες από τις ομάδες της ομαδοποίησης αντιστοιχούν σε φθορές και ποιες όχι, αρχικά συγκρίνονται οπτικά με τις εικόνες επισημάνσεων, και αν δεν είναι άμεσα εμφανές υπολογίζονται οι εξωτερικοί δείκτες υποθέτοντας κάθε φορά διαφορετικές ομάδες ως φθορές και όχι φθορές. Με βάση τα καλύτερα αποτελέσματα στους εξωτερικούς δείκτες, αποφασίζεται ποιες από τις ομάδες δείχνουν φθορές σε κάθε εικόνα. Έτσι, για κάθε εικόνα ομαδοποίησης ξεχωριστά μετατρέπονται τα εικονοστοιχεία των ομάδων σε False αν δεν θεωρούνται φθορές, ή True αν θεωρούνται φθορές.

```

Clustered_array = np.load('C:/Diplomatiki/...npz', ...)

bool_clustered_array = np.empty_like(Clustered_array, dtype=bool)

for i in range(Clustered_array.shape[0]):
    for j in range(Clustered_array.shape[1]):
        if Clustered_array[i,j] == 1 or Clustered_array[i,j] == 3:
            bool_clustered_array[i,j]=True
        else:
            bool_clustered_array[i,j]=False

```

«Περίπτωση εικόνας agiosnik003538, όπου η φθορές απεικονίζονται με μωβ και λευκό»

Ομαδοποιημένος πίνακας = Φόρτωση(Θέση μνήμης H/Y, άλλες προεπιλεγμένες παράμετροι)

Ομαδοποιημένος πίνακας Αληθές/Ψευδές = κενός πίνακας με διαστάσεις (Ομαδοποιημένος πίνακας, τύπος = Ψευδές/Αληθές)

Για i από 0 μέχρι ύψος Ομαδοποιημένου πίνακα
 Για ξ από 0 μέχρι πλάτος Ομαδοποιημένου πίνακα
 Αν στοιχείο ομαδοποιημένου πίνακα = 1 ή 3
 στοιχείο ομαδοποιημένου πίνακα A/Ψ = Αληθές
 Αλλιώς
 στοιχείο ομαδοποιημένου πίνακα A/Ψ = Ψευδές

«Αληθή τα εικονοστοιχεία με φθορές, μωβ(1) ή λευκά (3)»

Με τους δύο δυαδικούς πίνακες ίσων διαστάσεων, κάθε εικονοστοιχείο του πρώτου συγκρίνεται με το αντίστοιχό του στον δεύτερο πίνακα. Αν και τα δύο είναι True, τότε και τα δύο ανήκουν σε κατηγορία φθοράς και το αποτέλεσμα του συγκριτικού ελέγχου ισούται με αληθώς θετικό (TP), αν και τα δύο ανήκουν σε κατηγορία μη φθοράς το αποτέλεσμα του ελέγχου θεωρείται αληθώς αρνητικό (TN) και αν διαφωνούν τότε θεωρείται ψευδώς θετικό η αρνητικό το αποτέλεσμα στη συγκεκριμένη θέση του εικονοστοιχείου. Αφού υπολογιστεί το αποτέλεσμα του ελέγχου, τυπώνονται συνολικά οι τιμές του πίνακα σύγχυσης και υπολογίζονται και τυπώνονται οι τιμές των εξωτερικών δεικτών απόδοσης της ταξινόμησης εικονοστοιχείων, precision (2), recall (3), accuracy (5), F1 (4). (Παράρτημα, κώδικας, Σύγκριση εικόνων που προκύπτουν από ομαδοποίηση με τις εικόνες επισημάνσεων).

```

TP=0
TN=0
FP=0
FN=0

for i in range(bool_clustered_array.shape[0]):
    for j in range(bool_clustered_array.shape[1]):
        if (bool_clustered_array[i,j] == True) and (Annotation_array[i,j]
        == True):
            TP+=1
        elif (bool_clustered_array[i, j] == False) and
        (Annotation_array[i, j] == False):
            TN+=1
        elif (bool_clustered_array[i, j] == True) and
        (Annotation_array[i, j] == False):
            FP+=1
        elif (bool_clustered_array[i, j] == False) and
        (Annotation_array[i, j] == True):
            FN+=1

precision = TP/(TP+FP)
recall = TP/(TP+FN)
accuracy = (TP+TN)/(TP+TN+FP+FN)

F1 = 2*precision*recall/(precision+recall)

print(f'TP: {TP}, TN: {TN}, FP: {FP}, FN: {FN}')

print(f'precision: {round(precision,2)}, recall: {round(recall,2)},
accuracy: {round(accuracy,2)}, F1: {round(F1,2)}')

```

Αληθώς θετικά = 0
 Αληθώς αρνητικά = 0
 Ψευδώς θετικά = 0
 Ψευδώς αρνητικά = 0

Για i από 0 μέχρι ύψος Ομαδοποιημένου πίνακα A/Ψ
 Για ξ από 0 μέχρι πλάτος Ομαδοποιημένου πίνακα A/Ψ
 Αν στοιχείο ομαδοποιημένου πίνακα A/Ψ =
 Αληθές και στοιχείο Πίνακα επισημάνσεων =
 Αληθές:
 $\text{Αληθώς θετικά} = \text{Αληθώς θετικά} + 1$
 Αλλιώς αν στοιχείο ομαδοποιημένου πίνακα A/Ψ
 = Ψευδές και στοιχείο Πίνακα επισημάνσεων =
 Ψευδές:
 $\text{Αληθώς αρνητικά} = \text{Αληθώς αρνητικά} + 1$
 Αλλιώς αν στοιχείο ομαδοποιημένου πίνακα A/Ψ
 = Αληθές και στοιχείο Πίνακα επισημάνσεων =
 Ψευδές:
 $\text{Ψευδώς θετικά} = \text{Ψευδώς θετικά} + 1$
 Αλλιώς αν στοιχείο ομαδοποιημένου πίνακα A/Ψ
 = Ψευδές και στοιχείο Πίνακα επισημάνσεων =
 Αληθές:
 $\text{Ψευδώς αρνητικά} = \text{Ψευδώς αρνητικά} + 1$

$\text{Ακρίβεια} = \text{Αληθώς θετικά} / (\text{Αληθώς θετικά} + \text{Ψευδώς θετικά})$
 $\text{Ανάκληση} = \text{Αληθώς θετικά} / (\text{Αληθώς θετικά} + \text{Ψευδώς Αρνητικά})$
 $\text{Ορθότητα} = (\text{Αληθώς θετικά} + \text{Αληθώς αρνητικά}) / (\text{Αληθώς θετικά} + \text{Αληθώς αρνητικά} + \text{Ψευδώς θετικά} + \text{Ψευδώς αρνητικά})$
 $\text{F1} = 2 * \text{Ακρίβεια} * \text{Ανάκληση} / (\text{Ακρίβεια} + \text{Ανάκληση})$

Εκτύπωσε (Αριθμούς Αληθώς και Ψευδώς θετικών και αρνητικών)

Εκτύπωσε (Στρογγυλοποίηση στα δύο δεκαδικά (Ακρίβεια, Ανάκληση, Ορθότητα, F1))

Τα αποτελέσματα που τυπώνει ο αλγόριθμος στην οθόνη για το παράδειγμα της εικόνας agiosnik003538 έχουν την παρακάτω μορφή:

TP: 402521, TN: 737988, FP: 479565, FN: 290158

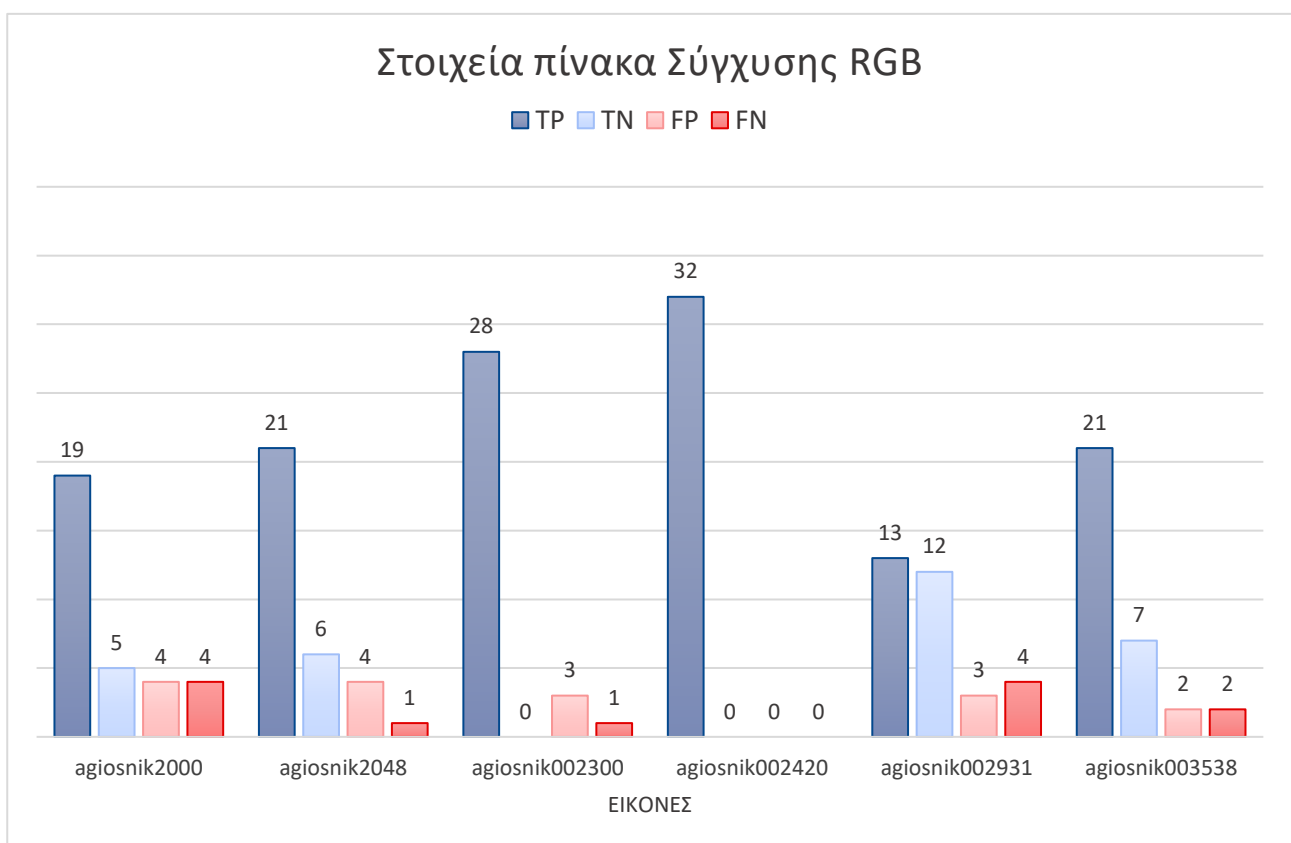
precision: 0.46, recall: 0.58, accuracy: 0.6, F1: 0.51

6.3 Αποτελέσματα αλγορίθμων

6.3.1 Επισημασμένες εικόνες

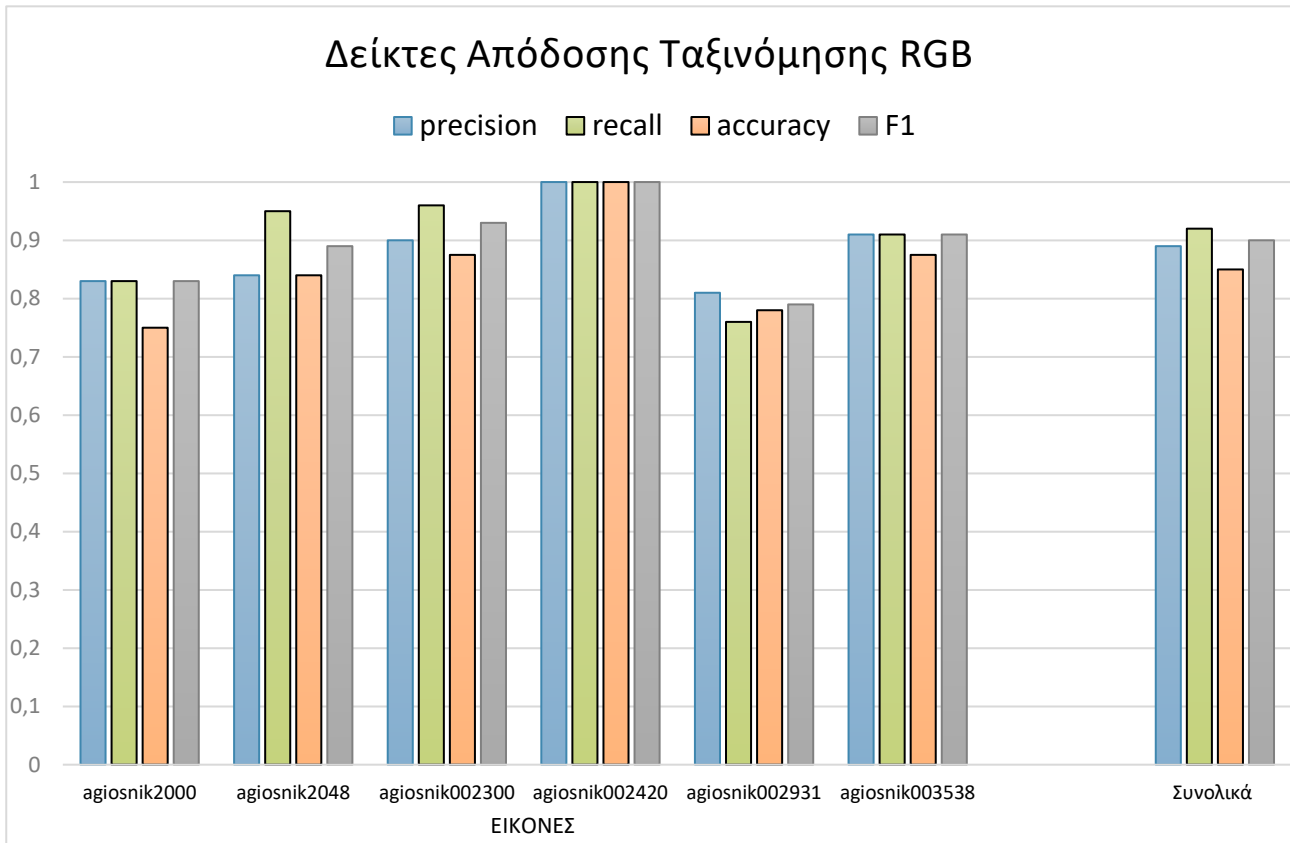
Από την εφαρμογή των μοντέλων προβλέψεων στα παράθυρα των εικόνων του συνόλου δεδομένων, προκύπτουν τα αποτελέσματα όπως αναφέρονται στα παρακάτω διαγράμματα. Η παρουσία φθορών κατονομάζεται ως θετική περίπτωση, ενώ η μη εμφάνιση φθορών κατονομάζεται αρνητική περίπτωση. Από την σύγκριση των προβλέψεων των παραθύρων εικόνων και των παραθύρων επισημάνσεων, προκύπτει ο πίνακας σύγχυσης (Εικόνα 25) και άρα οι δείκτες απόδοσης της ταξινόμησης (Εικόνα 26) για κάθε εικόνα ξεχωριστά. Πρέπει να σημειωθεί ότι τα παράθυρα είναι αυτά στα οποία έχουν εκπαιδευτεί τα μοντέλα, και άρα θα εμφανίζουν σημαντικά μεγαλύτερη ακρίβεια από όση θα εμφάνιζαν σε δεδομένα που δεν έχουν ξαναδεί.

Μοντέλο VGG-16 εκπαιδευμένο σε φασματικά κανάλια RGB:



Εικόνα 25: Στοιχεία πίνακα σύγχυσης μοντέλου RGB στην ταξινόμηση επισημασμένων παραθύρων της εικόνας

Συνολικά για όλες τις εικόνες RGB, τα αληθώς θετικά παράθυρα (True positive) είναι 134, τα αληθώς αρνητικά (True negative) είναι 30, τα ψευδώς θετικά (False positive) είναι 16 και τα ψευδώς αρνητικά (False negative) είναι 12.

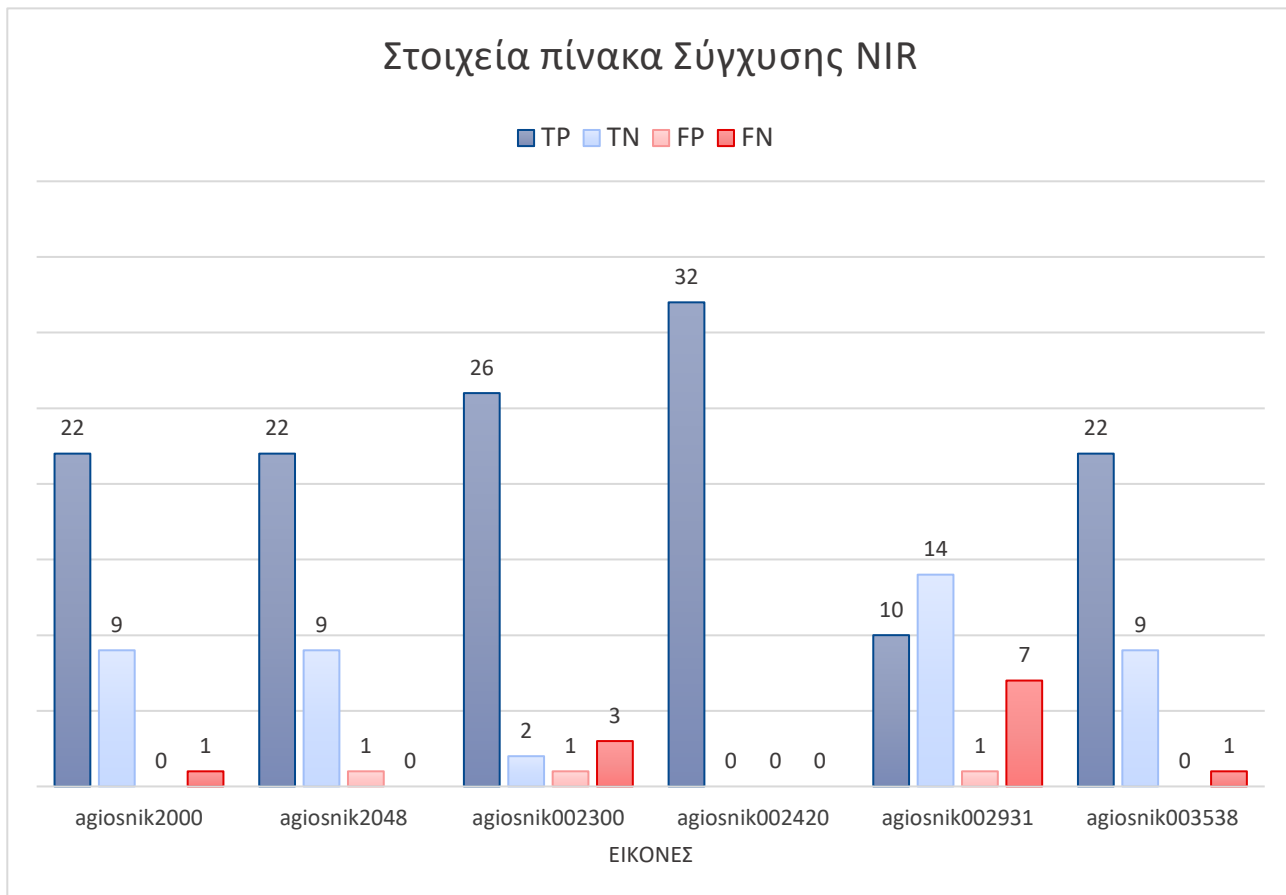


Εικόνα 26: Δείκτες απόδοσης ταξινόμησης μοντέλου RGB σε επισημασμένα παράθυρα της εικόνας

Τα αναλυτικά αποτελέσματα των συγκρίσεων για κάθε παράθυρο κάθε εικόνας RGB παρατίθενται στο τέλος της εργασίας (Παράρτημα, Πίνακες Αποτελεσμάτων Ταξινόμησης, Αποτελέσματα μοντέλου RGB στην ταξινόμηση επισημασμένων παραθύρων των εικόνων σε μορφή πλήρους πίνακα).

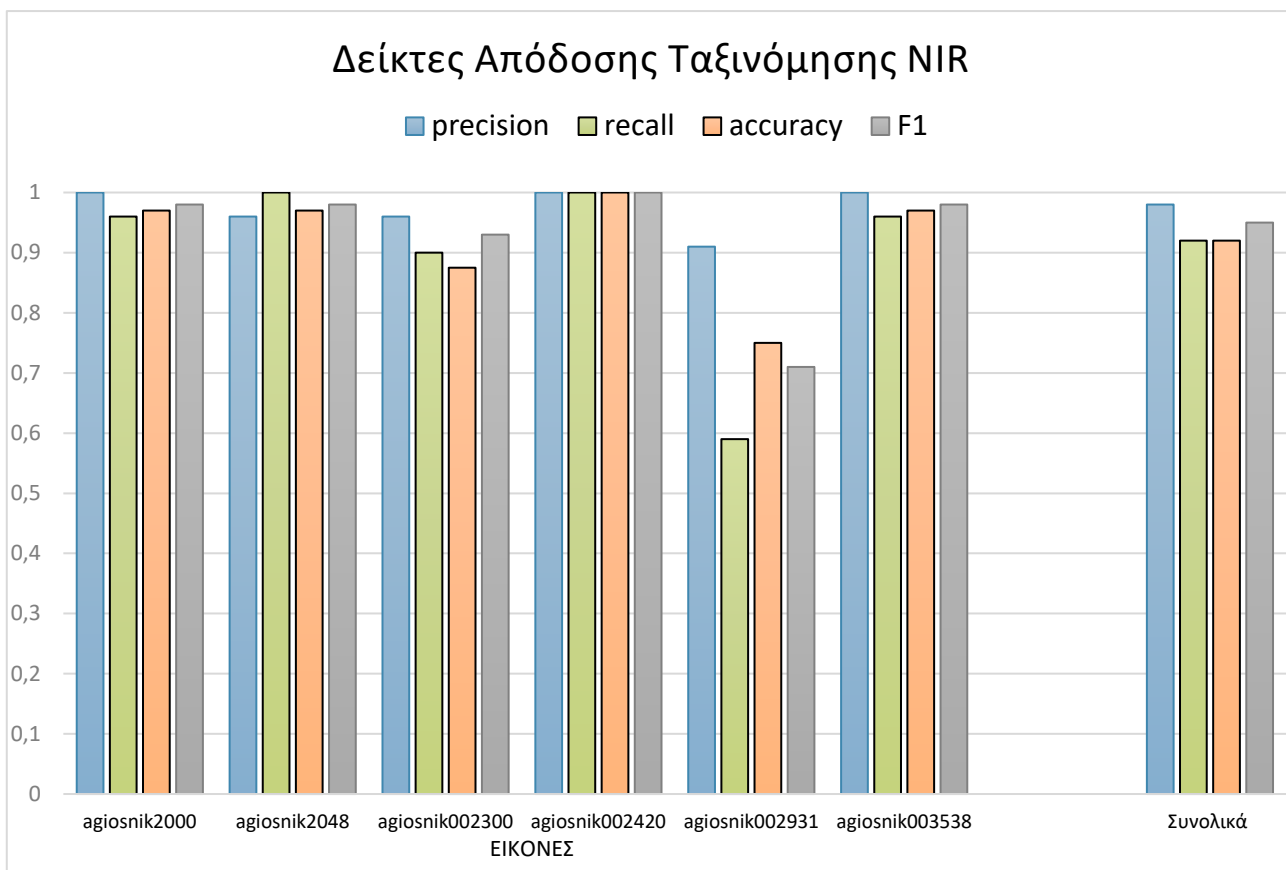
Όσον αφορά το μοντέλο εκπαιδευμένο σε εγγύς υπέρυθρα φασματικά κανάλια, όπως έχει προαναφερθεί έχει παρουσιάσει έντονο φαινόμενο υπερπροσαρμογής. Επομένως αν και οι προβλέψεις του είναι ακριβείς, όπως φαίνεται στα παρακάτω διαγράμματα (Εικόνες 27, 28), αυτό οφείλεται περισσότερο στην απομνημόνευση συγκεκριμένων χαρακτηριστικών σε κάθε εικόνα παρά στην δυνατότητά του να εντοπίζει γενικευμένα φθορές.

Μοντέλο VGG-16 εκπαιδευμένο σε φασματικά κανάλια NIR:



Εικόνα 27: Στοιχεία πίνακα σύγκρισης μοντέλου NIR στην ταξινόμηση επισημασμένων παραθύρων της εικόνας

Συνολικά για όλες τις εικόνες NIR, τα αληθώς θετικά παράθυρα (True positive) είναι 134, τα αληθώς αρνητικά (True negative) είναι 43, τα ψευδώς θετικά (False positive) είναι 3 και τα ψευδώς αρνητικά (False negative) είναι 12.

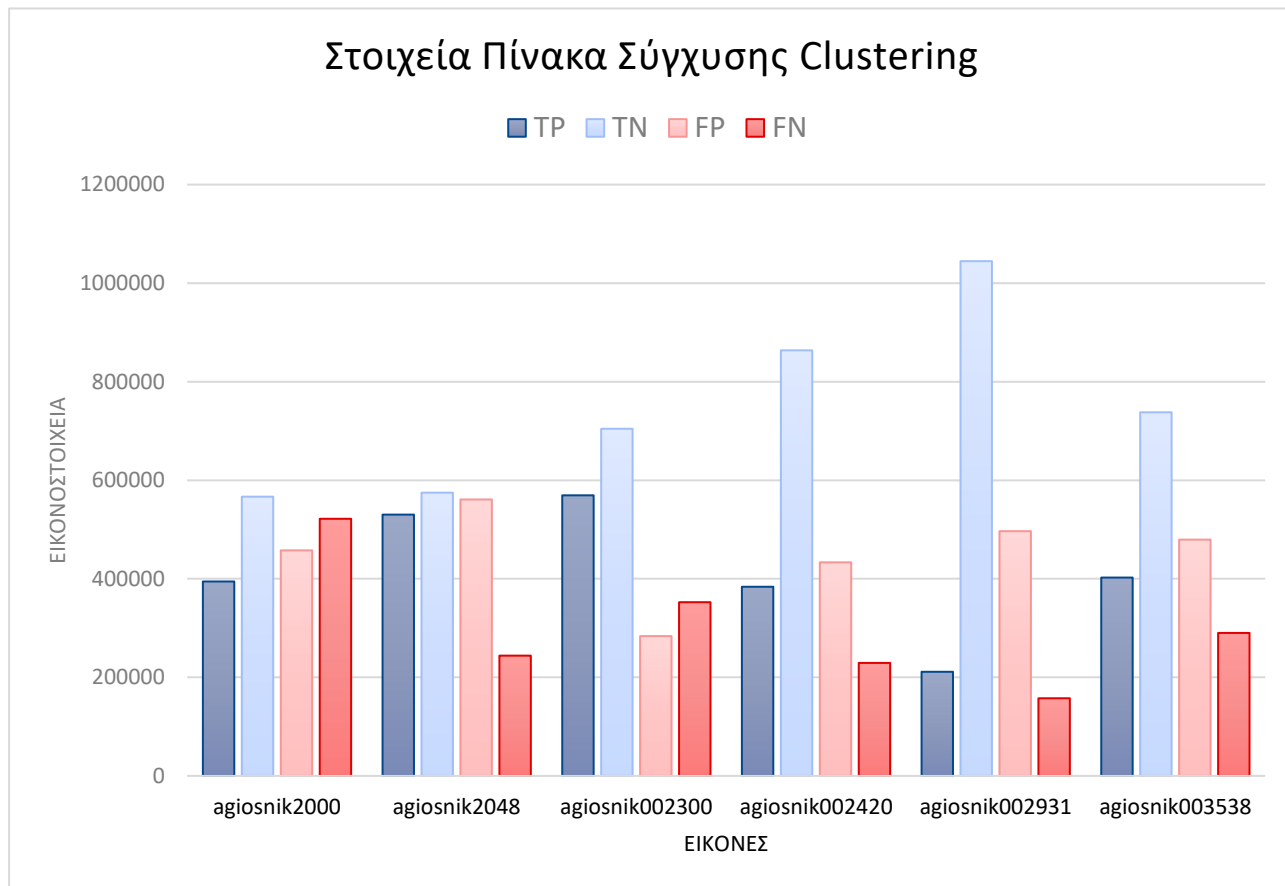


Εικόνα 28: Δείκτες απόδοσης ταξινόμησης μοντέλου NIR σε επισημασμένα παράθυρα της εικόνας

Τα αποτελέσματα των συγκρίσεων για κάθε παράθυρο κάθε εικόνας NIR παρατίθενται αναλυτικά στο τέλος της εργασίας (Παράρτημα, Πίνακες Αποτελεσμάτων Ταξινόμησης, Αποτελέσματα μοντέλου NIR στην ταξινόμηση επισημασμένων παραθύρων των εικόνων σε μορφή πλήρους πίνακα).

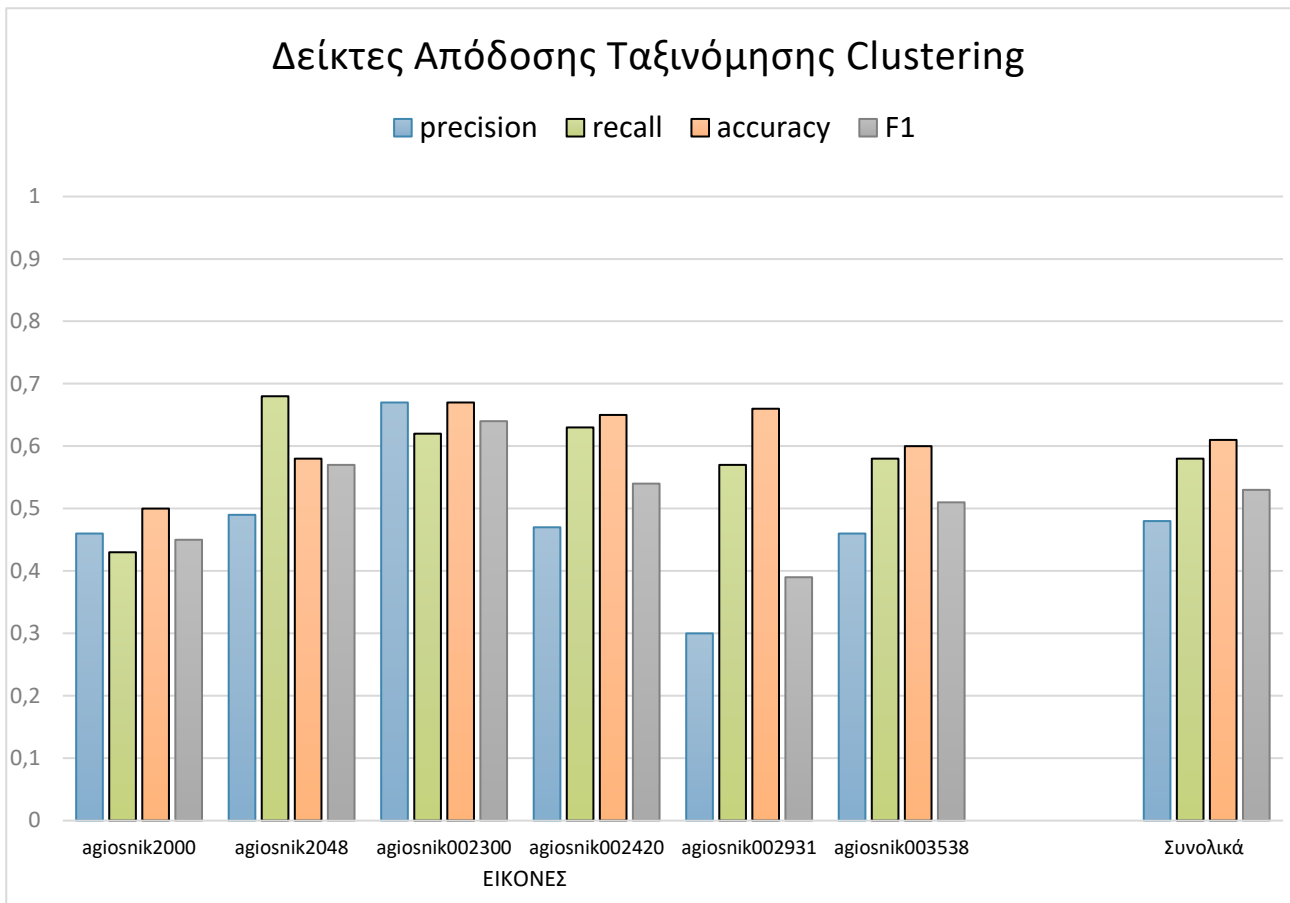
Όσον αφορά την ομαδοποίηση, ο υπολογισμός των εξωτερικών δεικτών ως προς τις επισημάνσεις γίνεται σε επίπεδο εικονοστοιχείου. Επομένως οι αθροιστικοί αριθμοί των στοιχείων πίνακα σύγκρισης (Εικόνα 29) είναι πολλές τάξεις κλίμακας μεγαλύτεροι, με κάθε εικόνα να έχει 1.910.232 στοιχεία, εξαιρουμένης της εικόνας agiosnik2000 που είναι ελάχιστα μεγαλύτερη. Οι δείκτες απόδοσης της ταξινόμησης εικονοστοιχείων (Εικόνα 30) ακολουθούν την ίδια κλίμακα, και άρα μπορούν να συγκριθούν άμεσα με τους δείκτες των συνελκτικών μοντέλων (Εικόνες 26, 28).

Ομαδοποίηση εικονοστοιχείων με K-means:



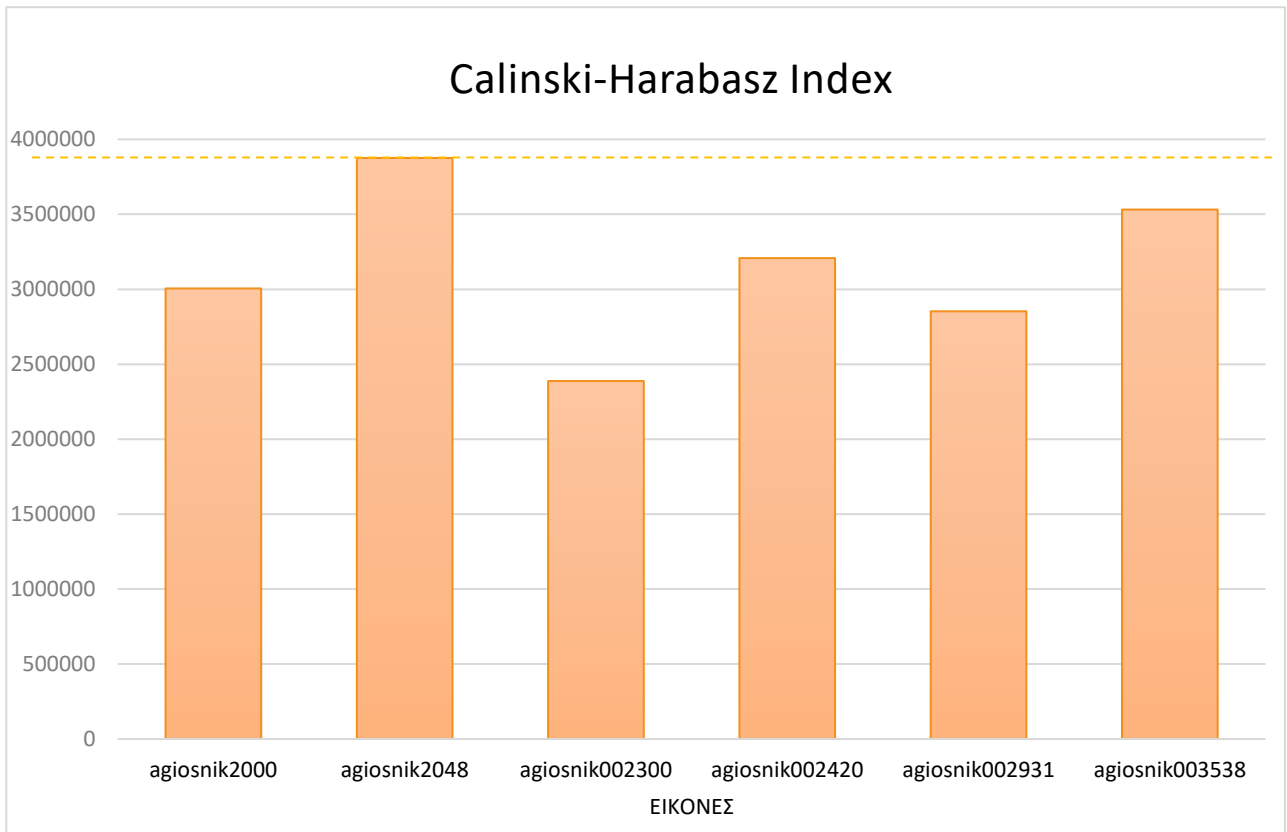
Εικόνα 29: Στοιχεία πίνακα σύγκυσης ομαδοποίησης εικονοστοιχείων επισημασμένων εικόνων

Συνολικά για όλες τις εικόνες tiff με τα 41 φασματικά κανάλια, τα αληθώς θετικά καταταγμένα εικονοστοιχεία (True positive) είναι 2.491.963, τα αληθώς αρνητικά (True negative) είναι 4.492.482, τα ψευδώς θετικά (False positive) είναι 2.712.106 και τα ψευδώς αρνητικά (False negative) είναι 1.795.405 .

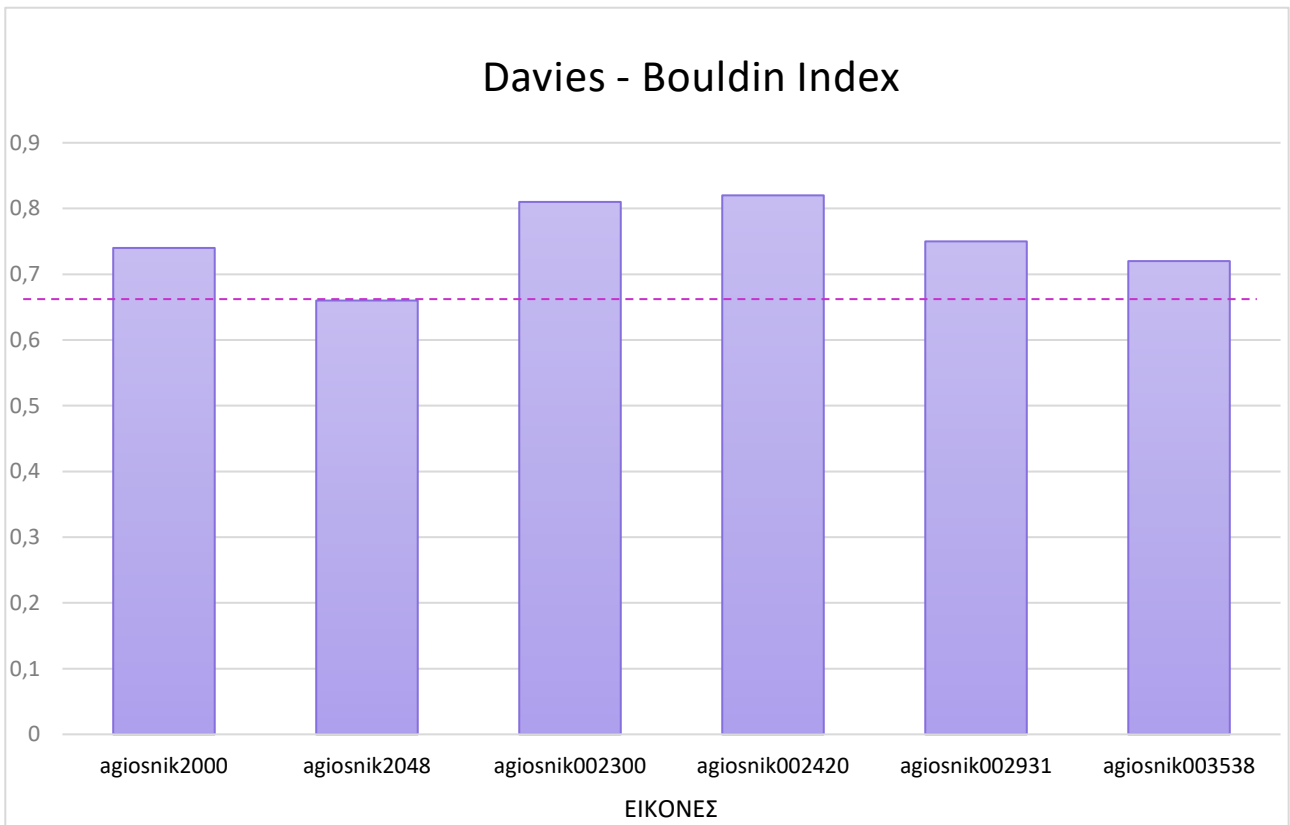


Εικόνα 30: Δείκτες απόδοσης ταξινόμησης ομαδοποίησης με K-means σε επισημασμένα εικονοστοιχεία της εικόνας

Πέρα από τους εξωτερικούς δείκτες ως προς τις επισημάνσεις σε επίπεδο εικονοστοιχείου, υπολογίζονται και οι εσωτερικοί δείκτες εγκυρότητας της ομαδοποίησης, δηλαδή ο δείκτης Calinski-Harabasz (CHI) (32) και Davies-Bouldin (DBI) (33). Υπενθυμίζεται ότι καταλληλότερος αλγόριθμος είναι αυτός που επιστρέφει την μεγαλύτερη τιμή του δείκτη CHI, ενώ καταλληλότερος αλγόριθμος είναι αυτός που επιστρέφει χαμηλότερη τιμή του δείκτη DBI. Με την επιφύλαξη ότι ο αλγόριθμος K-means δεν ενδείκνυται για αυτά τα δεδομένα, οι δείκτες για κάθε επισημασμένη εικόνα απεικονίζονται στα παρακάτω διαγράμματα, με την καλύτερη παρούσα τιμή του δείκτη να καθορίζει ένα διακεκομμένο ευθύγραμμο τμήμα (Εικόνες 31 και 32).



Εικόνα 31: Δείκτης Calinski – Harabasz της ομαδοποίησης με K-means εικονοστοιχείων των επισημασμένων εικόνων



Εικόνα 32: Δείκτης Davies – Bouldin της ομαδοποίησης με K-means εικονοστοιχείων των επισημασμένων εικόνων

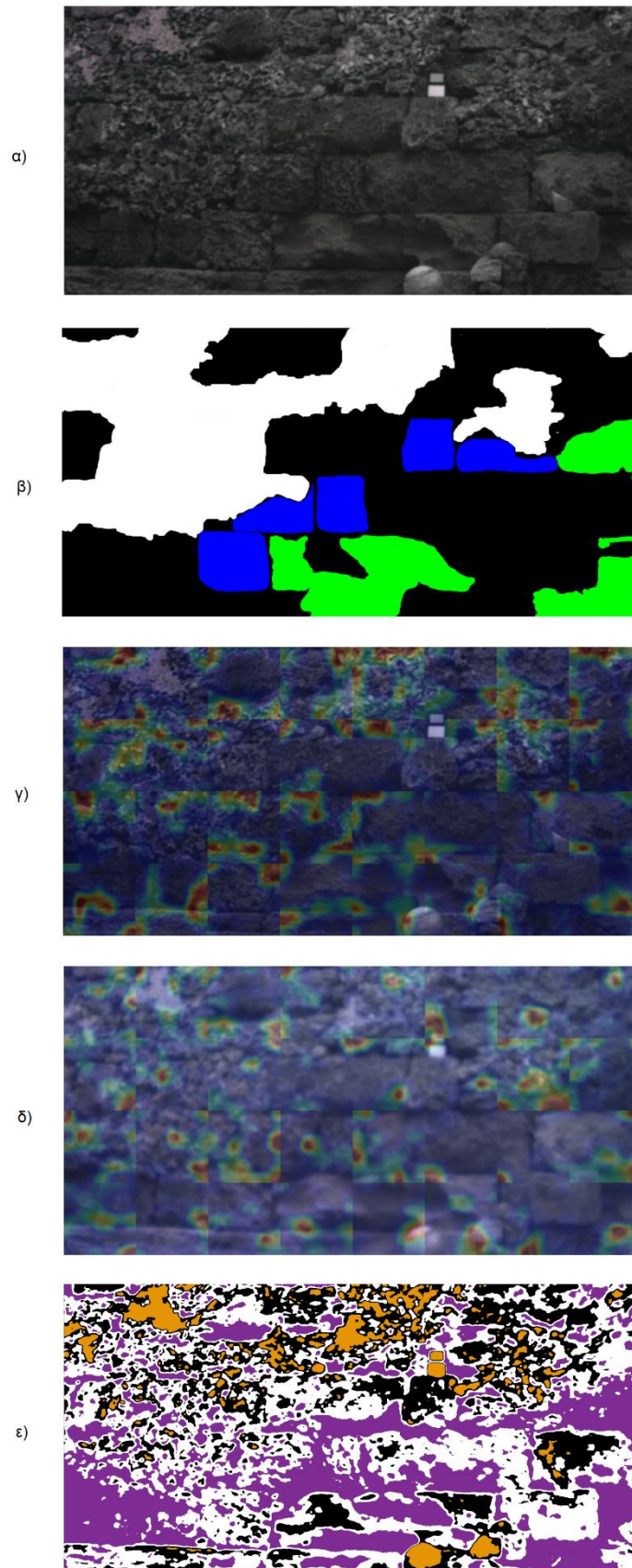
Οι βέλτιστοι δείκτες εσωτερικής εγκυρότητας για τις επισημασμένες εικόνες αναδεικνύονται με την ομαδοποίηση των εικονοστοιχείων της εικόνας agiosnik2048. Για τον δείκτη Calinski-Harabasz είναι η τιμή 3.875.641 , για τον δείκτη Davies-Bouldin είναι η τιμή 0,66.

Τα αναλυτικά αποτελέσματα των συγκρίσεων εικονοστοιχείων κάθε ομαδοποιημένης εικόνας με την αντίστοιχη εικόνα επισήμανσης παρατίθενται στο τέλος της εργασίας, μαζί με τους πίνακες υπολογισμένων εσωτερικών δεικτών (Παράρτημα, Πίνακες Αποτελεσμάτων Ταξινόμησης, Εξωτερικοί και εσωτερικοί δείκτες ομαδοποίησης εικονοστοιχείων επισημασμένων εικόνων σε μορφή πλήρους πίνακα).

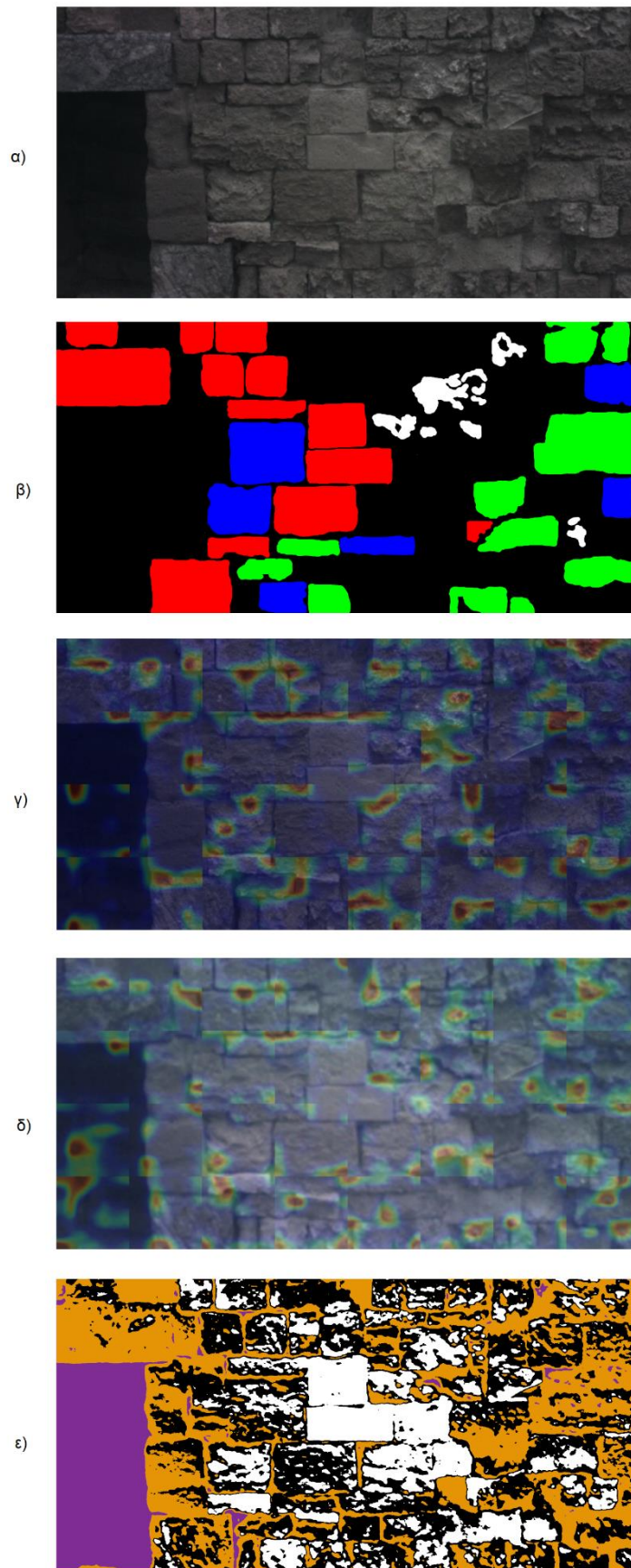
Για την οπτική σύγκριση των αλγορίθμων μηχανικής μάθησης πάνω στα δεδομένα, παρατίθενται οι αρχικές εικόνες αφού έχουν μετατραπεί σε .png τριών χρωματικών καναλιών RGB, οι εικόνες με τις επισημάνσεις όπως έχουν δοθεί μαζί με τα δεδομένα, οι ενωμένες εικόνες παραθύρων με Class Activation Mapping τόσο για το μοντέλο εκπαιδευμένο σε φασματικά κανάλια RGB όσο και για το μοντέλο εκπαιδευμένο σε φασματικά κανάλια NIR, και οι εικόνες με την απεικόνιση των ομάδων που προέκυψαν από την ομαδοποίηση.

Στις εικόνες επισημάνσεων με μαύρο χρώμα εμφανίζονται οι περιοχές χωρίς φθορές, ενώ με άλλα χρώματα εμφανίζονται φθορές διαφορετικού είδους ανά χρώμα. Στις εικόνες που φαίνεται το CAM των παραθύρων, με μπλε χρώμα εμφανίζονται τα στοιχεία που δεν συνεισφέρουν στην πρόβλεψη κατηγορίας φθοράς, ενώ με κίτρινο – κόκκινο εμφανίζονται τα στοιχεία που εντοπίζει ο αλγόριθμος ως σημαντικά στην ταξινόμηση του παραθύρου σε κατηγορία φθοράς. Στις εικόνες με τα αποτελέσματα του αλγορίθμου ομαδοποίησης, κάθε χρώμα δηλώνει μία ομάδα από τις 4 στις οποίες έχουν κατηγοριοποιηθεί τα εικονοστοιχεία. Αν και τα χρώματα των ομάδων είναι σταθερά για τον αριθμό κατηγορίας (μαύρο για την ομάδα 0, μωβ για την ομάδα 1, πορτοκαλί για την ομάδα 2 και λευκό για την ομάδα 3), πρέπει να υπογραμμιστεί ότι σε κάθε εικόνα οι ίδιες ομάδες απεικονίζουν διαφορετικά στοιχεία, καθώς η ομαδοποίηση εκτελείται σε κάθε εικόνα ξεχωριστά. Επομένως, στην περιγραφή κάθε εικόνας, αναφέρεται με ποια χρώματα εμφανίζονται οι κατηγορίες που είναι πιο πιθανό να δείχνουν φθορές, και με βάση τις οποίες υπολογίστηκαν οι εξωτερικοί δείκτες εγκυρότητας.

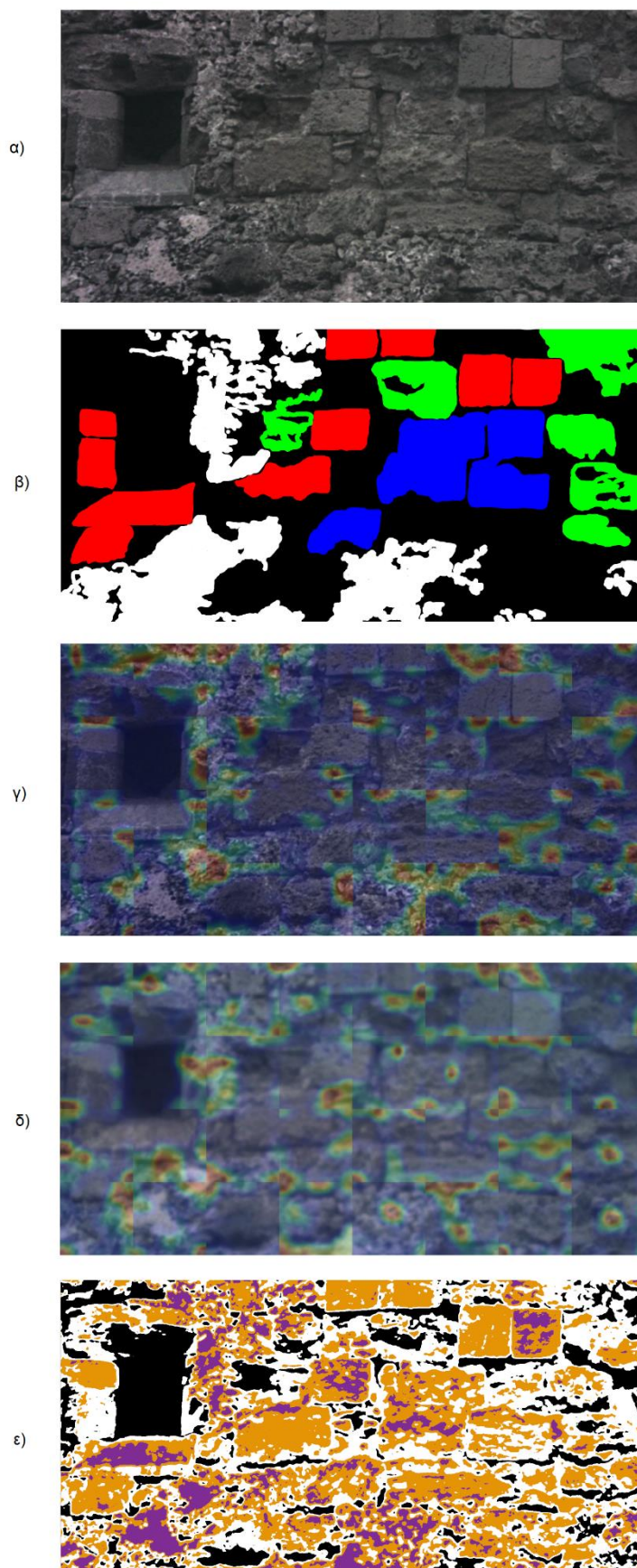
Η οπτική σύγκριση των μεθόδων μηχανικής μάθησης για εντοπισμό φθορών παρουσιάζεται για κάθε εικόνα ξεχωριστά, παρακάτω, στις εικόνες 33 έως και 38.



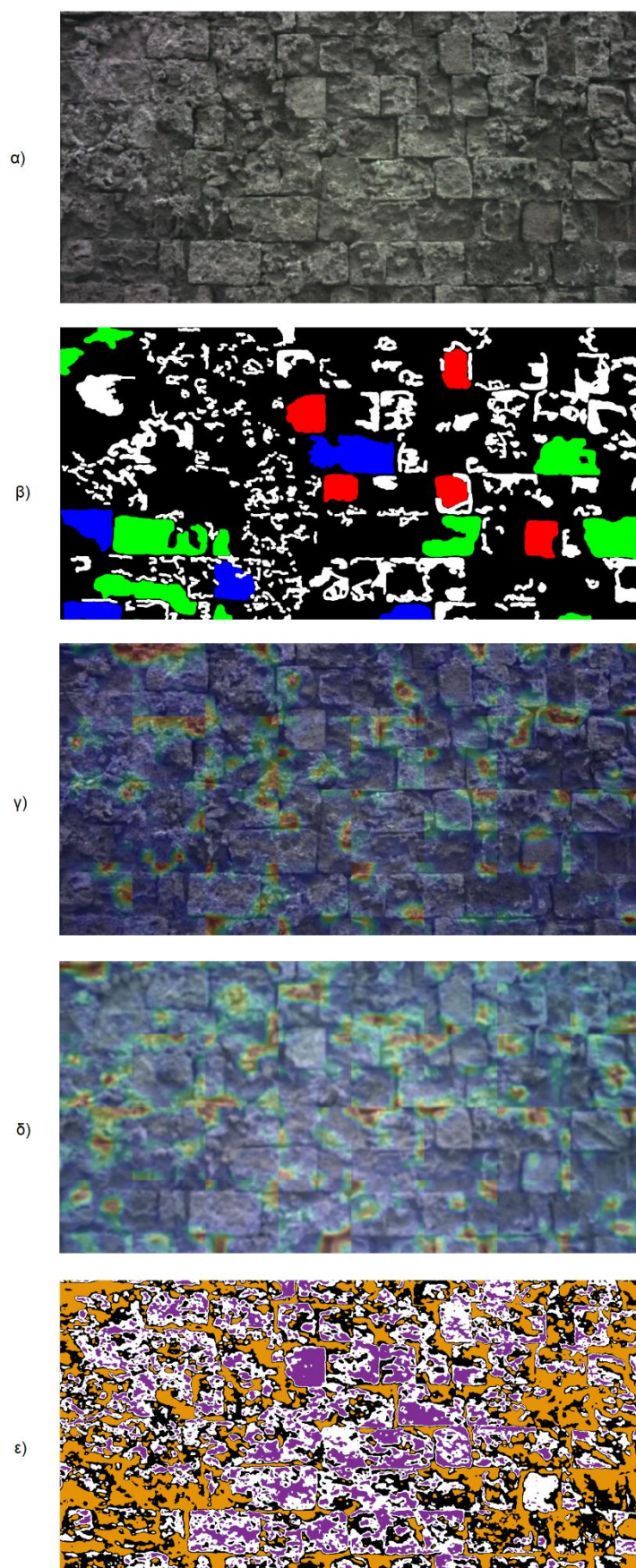
Εικόνα 33: Οπτική παράθεση μεθόδων εντοπισμού φθορών, *agiosnik2000*
α) εικόνα RGB, β) Επισημάνσεις φθορών, γ) CAM μοντέλου RGB, δ) CAM μοντέλου NIR,
ε) Αποτέλεσμα ομαδοποίησης. Οι φθορές εμφανίζονται με πορτοκαλί και λευκό χρώμα.



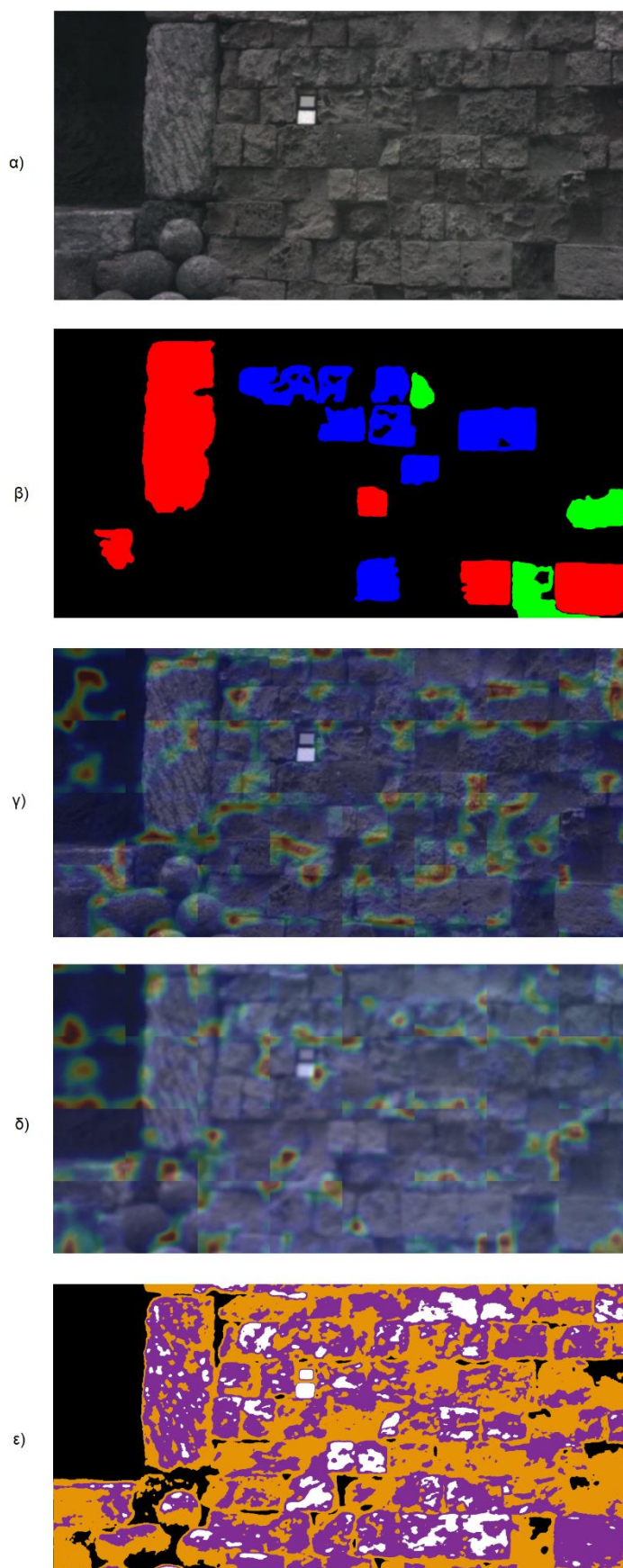
Εικόνα 34: Οπτική παράθεση μεθόδων εντοπισμού φθορών, agiosnik2048
α) εικόνα RGB, β) Επισημάνσεις φθορών, γ) CAM μοντέλου RGB, δ) CAM μοντέλου NIR,
ε) Αποτέλεσμα ομαδοποίησης. Οι φθορές εμφανίζονται με μαύρο και λευκό χρώμα.



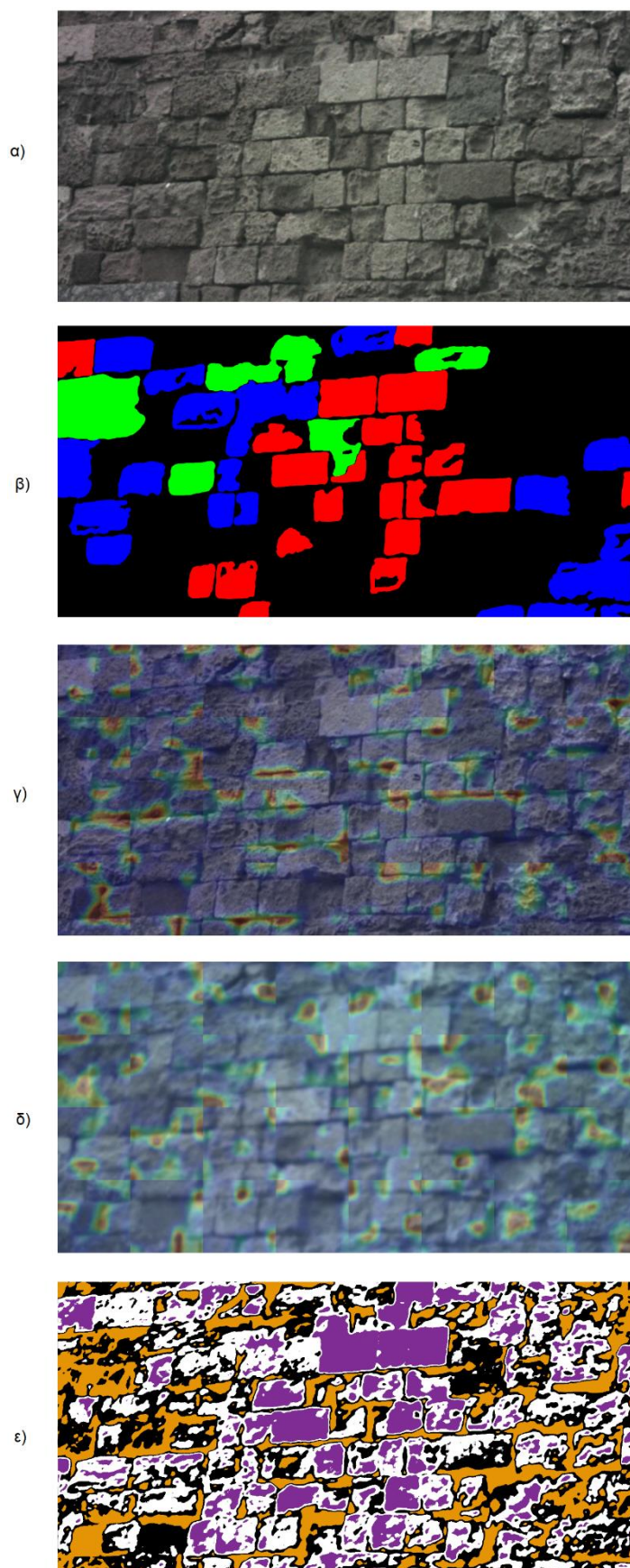
Εικόνα 35: Οπτική παράθεση μεθόδων εντοπισμού φθορών, agiosnik002300
α) εικόνα RGB, β) Επισημάνσεις φθορών, γ) CAM μοντέλου RGB, δ) CAM μοντέλου NIR,
ε) Αποτέλεσμα ομαδοποίησης. Οι φθορές εμφανίζονται με πορτοκαλί και μωβ χρώμα.



Εικόνα 36: Οπτική παράθεση μεθόδων εντοπισμού φθορών, agiosnik002420
α) εικόνα RGB, β) Επισημάνσεις φθορών, γ) CAM μοντέλου RGB, δ) CAM μοντέλου NIR,
ε) Αποτέλεσμα ομαδοποίησης. Οι φθορές εμφανίζονται με λευκό και μωβ χρώμα.



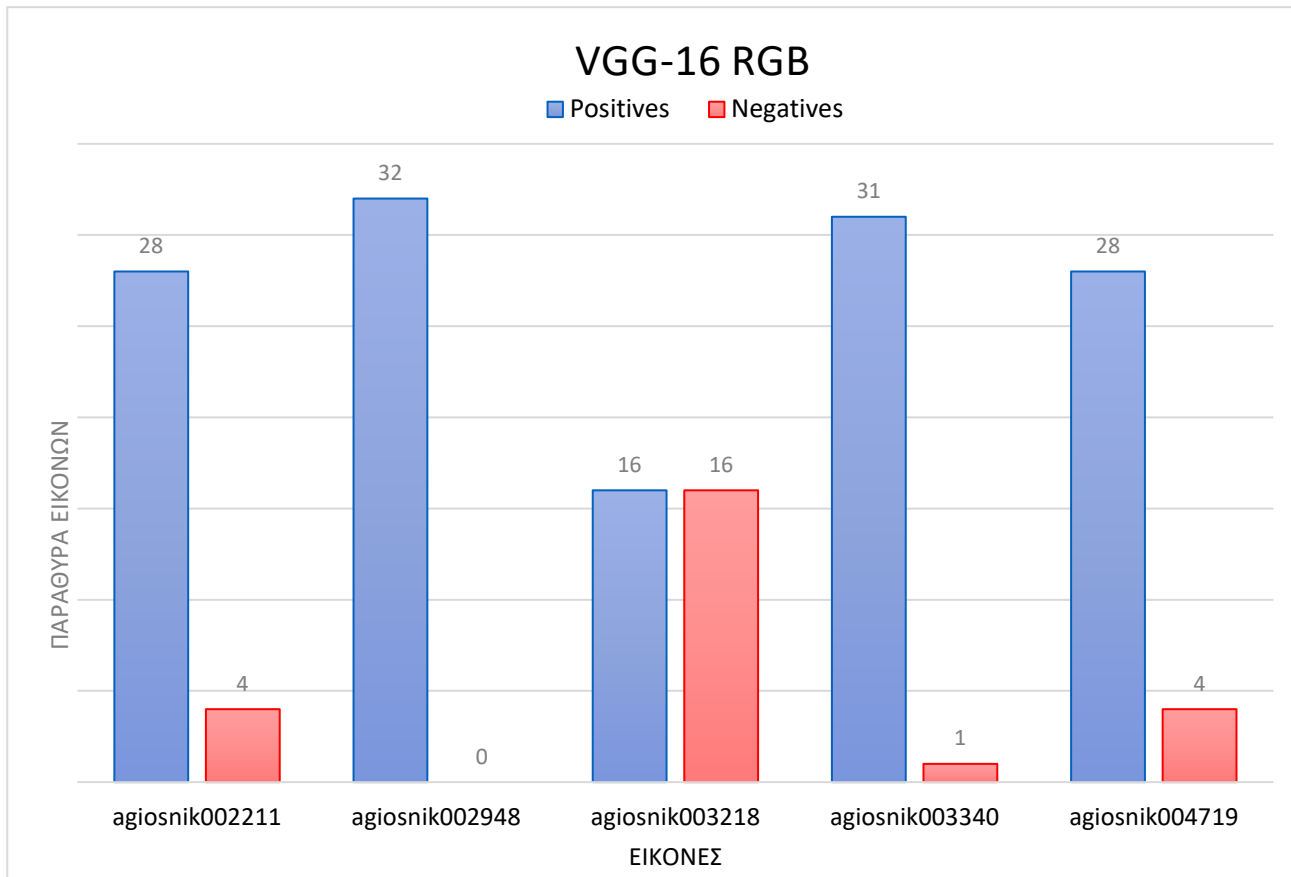
Εικόνα 37: Οπτική παράθεση μεθόδων εντοπισμού φθορών, agiosnik002931
α) εικόνα RGB, β) Επισημάνσεις φθορών, γ) CAM μοντέλου RGB, δ) CAM μοντέλου NIR,
ε) Αποτέλεσμα ομαδοποίησης. Οι φθορές εμφανίζονται με μωβ χρώμα.



Εικόνα 38: Οπτική παράθεση μεθόδων εντοπισμού φθορών, agiosnik003538
α) εικόνα RGB, β) Επισημάνσεις φθορών, γ) CAM μοντέλου RGB, δ) CAM μοντέλου NIR,
ε) Αποτέλεσμα ομαδοποίησης. Οι φθορές εμφανίζονται με μωβ και λευκό χρώμα.

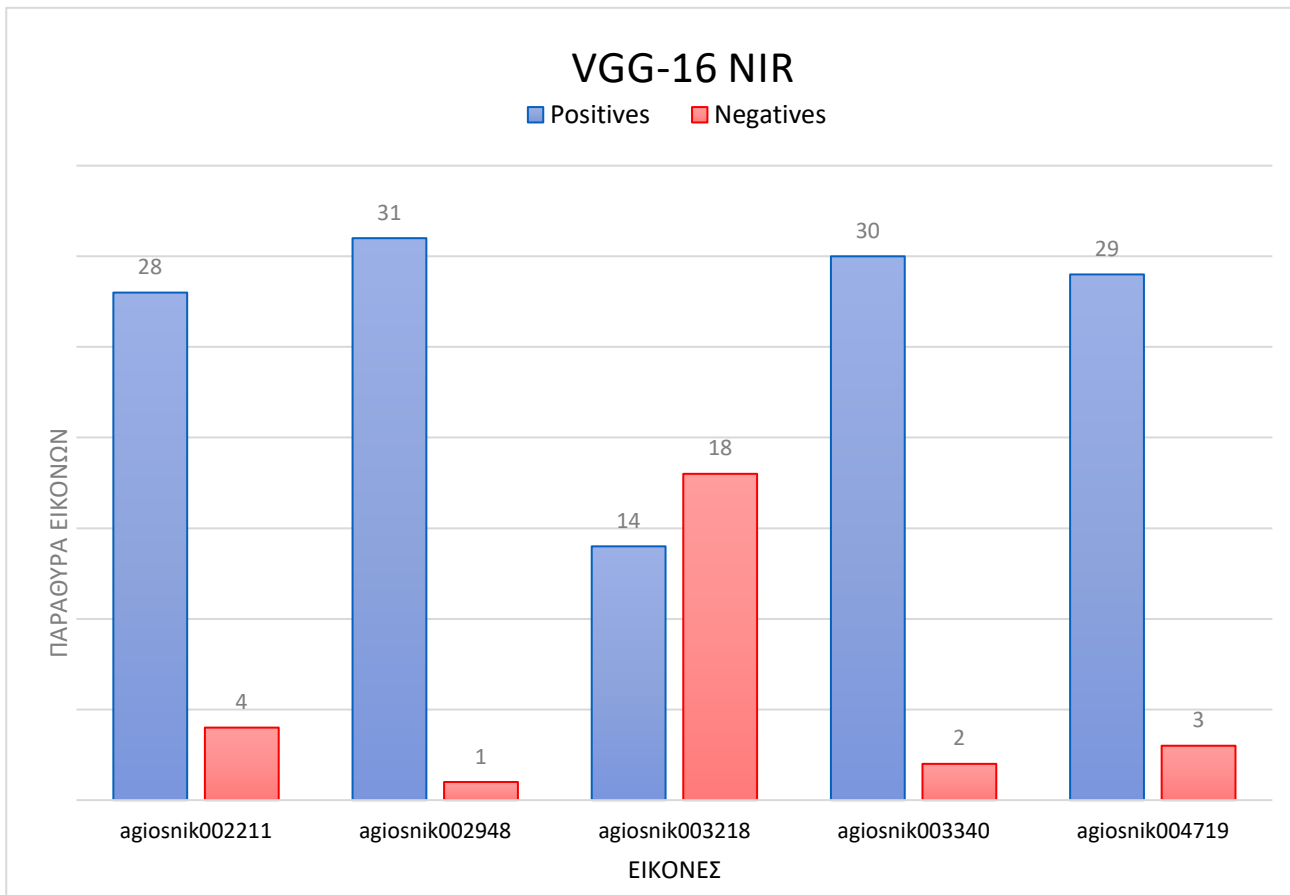
6.3.2 Μη επισημασμένες εικόνες

Επειδή οι παρακάτω εικόνες δεν διαθέτουν επισημάνσεις, παρατίθενται μόνο τα αποτελέσματα της ταξινόμησης των παραθύρων τους, τόσο από το μοντέλο που εφαρμόζεται στα φασματικά κανάλια RGB (Εικόνα 39) όσο και από το μοντέλο που εφαρμόζεται στα φασματικά κανάλια NIR (Εικόνα 40). Επίσης υπολογίζονται τα ποσοστά θετικών και αρνητικών σε φθορές προβλέψεων, για σύγκριση των τάσεων των μοντέλων στην πρόβλεψη των δεδομένων που δεν έχουν ξαναδεί.



Εικόνα 39: Αποτελέσματα μοντέλου RGB στη ταξινόμηση παραθύρων άγνωστων εικόνων

Συνολικά, κατατάσσονται 135 παράθυρα ως εικόνες που περιέχουν φθορές, ενώ τα υπόλοιπα 25 κατατάσσονται ως μη έχοντα φθορές. Το ποσοστό των παραθύρων με φθορές διαμορφώνεται στο 84,4% του συνόλου των παραθύρων μη επισημασμένων εικόνων, αντίστοιχα το ποσοστό των παραθύρων χωρίς φθορές διαμορφώνεται στο 15,6% του συνόλου, σύμφωνα με το fine-tuned με RGB εικόνες μοντέλο VGG-16.

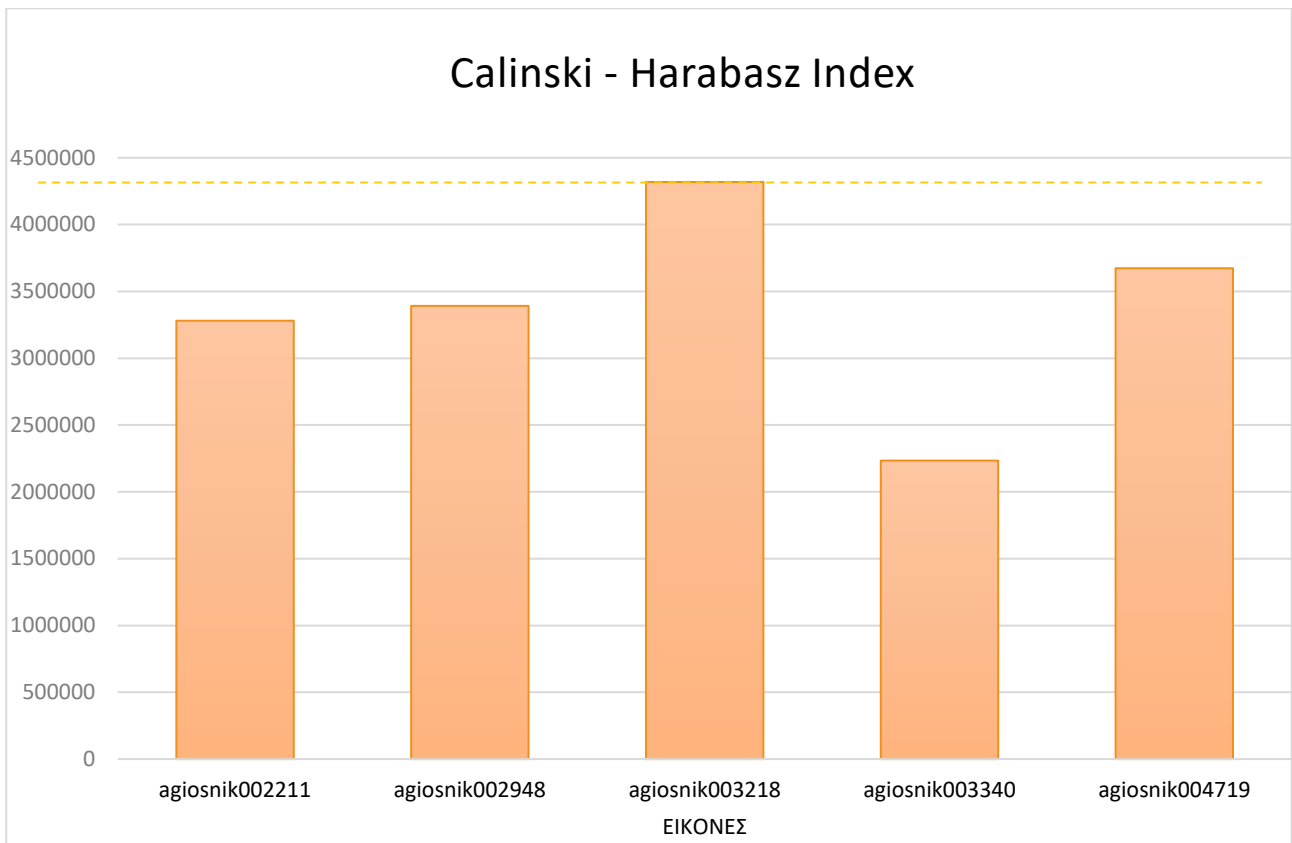


Εικόνα 40: Αποτελέσματα μοντέλου NIR στη ταξινόμηση παραθύρων άγνωστων εικόνων

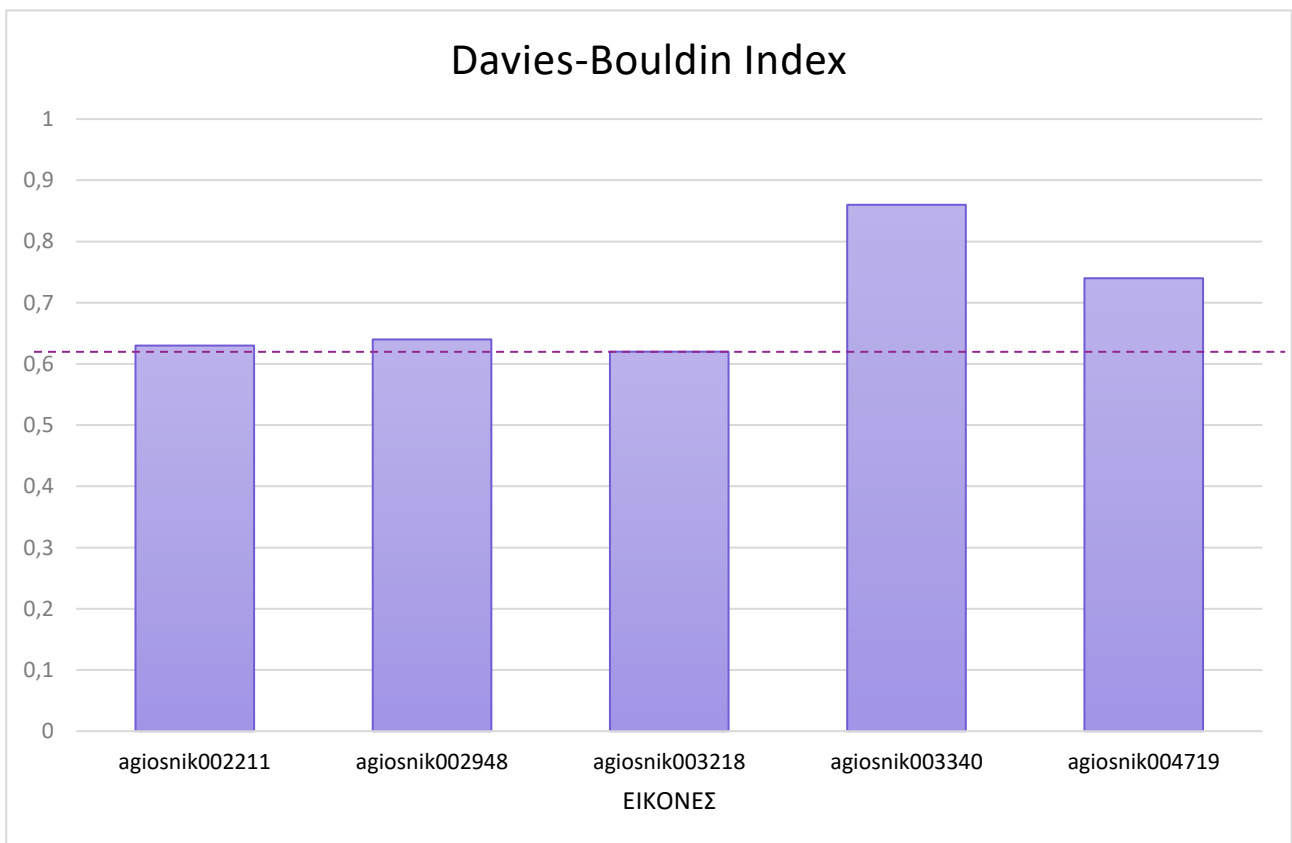
Σύμφωνα με το fine-tuned μοντέλο VGG-16 που έχει εκπαιδευτεί σε εικόνες εγγύς υπέρυθρου, 132 παράθυρα κατατάσσονται ως φθορές, ενώ τα υπόλοιπα 28 κατατάσσονται ως μη έχοντα φθορές. Το ποσοστό των παραθύρων με φθορές διαμορφώνεται στο 82,5% του συνόλου σε αυτή την ταξινόμηση, ενώ αντίστοιχα το ποσοστό των παραθύρων χωρίς φθορές διαμορφώνεται στο 17,5% του συνόλου μη επισημασμένων παραθύρων.

Ο αναλυτικός πίνακας προβλέψεων των δύο μοντέλων σε εικόνες χωρίς επισημάνσεις παρατίθεται στο τέλος της εργασίας (Παράρτημα, Πίνακες Αποτελεσμάτων Ταξινόμησης, Αποτελέσματα μοντέλων RGB και NIR στην ταξινόμηση παραθύρων μη επισημασμένων εικόνων σε μορφή πίνακα).

Αντίστοιχα, στις εικόνες αυτές δεν μπορεί να γίνει σύγκριση με την βασική πραγματικότητα των ομάδων που προκύπτουν από αλγόριθμους ομαδοποίησης. Επομένως μπορούν να υπολογιστούν μόνο εσωτερικοί δείκτες της ποιότητας ομαδοποίησης. Στα διαγράμματα που ακολουθούν που ακολουθεί αναγράφονται οι δείκτες Calinski-Harabasz (CHI) στην εικόνα 41 και Davies – Bouldin (DBI) στην εικόνα 42, για καθεμία από τις εικόνες χωρίς επισημάνσεις. Το διακεκομμένο ευθύγραμμο τμήμα ορίζει την καλύτερη παρούσα τιμή κάθε δείκτη.



Εικόνα 41: Δείκτης Calinski – Harabasz της ομαδοποίησης με K-means εικονοστοιχείων μη επισημασμένων εικόνων

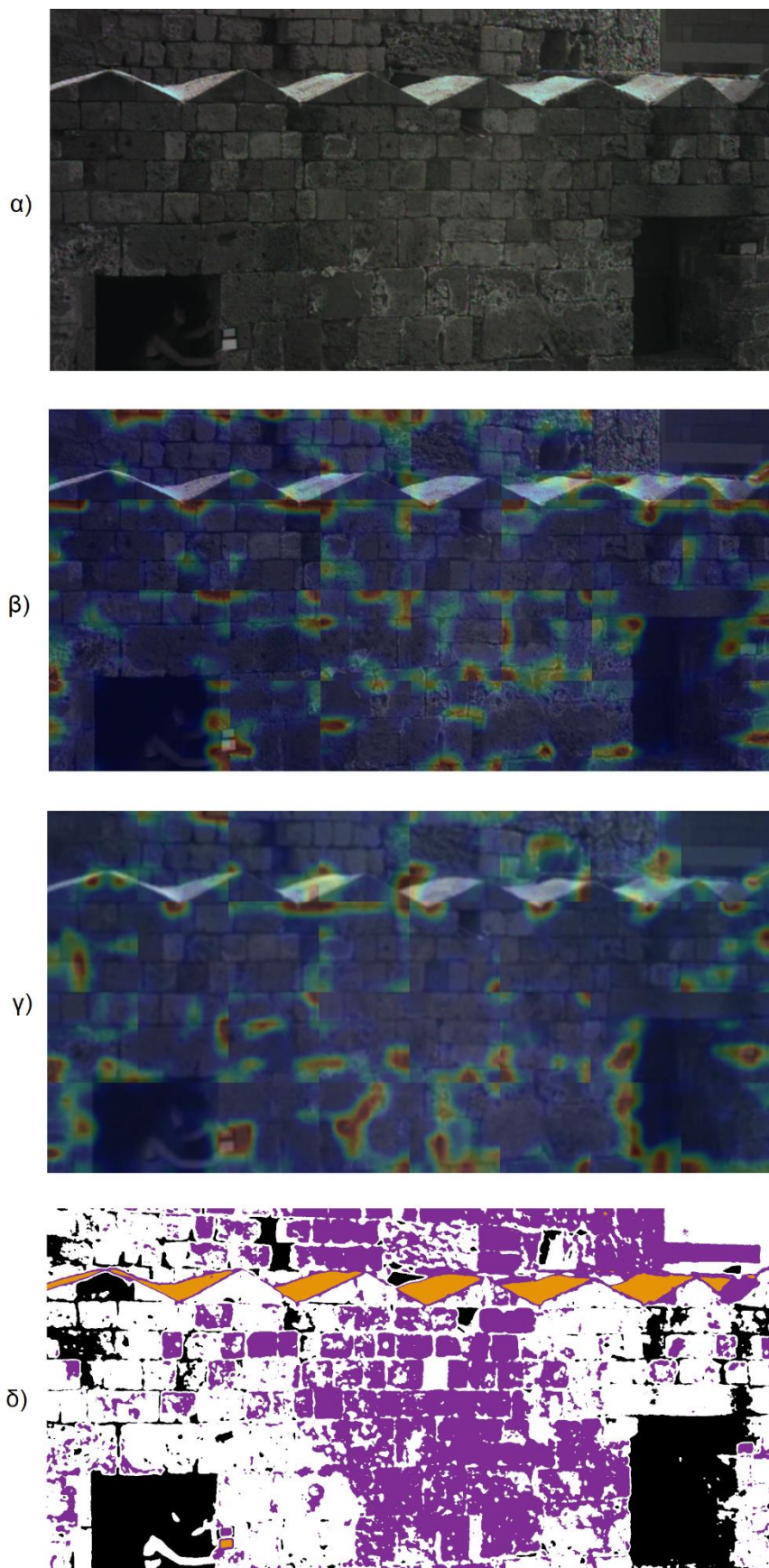


Εικόνα 42: Δείκτης Davies-Bouldin της ομαδοποίησης με K-means εικονοστοιχείων μη επισημασμένων εικόνων

Οι βέλτιστες τιμές των δεικτών εσωτερικής εγκυρότητας για τις εικόνες χωρίς επισημάνσεις αναδεικνύονται κατά την ομαδοποίηση των εικονοστοιχείων της εικόνας agiosnik003218. Για τον δείκτη Calinski-Harabasz είναι 4.319.148, για τον δείκτη Davies-Bouldin είναι 0,62. Οι ακριβείς τιμές των εσωτερικών δεικτών παρατίθενται στο τέλος της εργασίας (Παράρτημα, Πίνακες Αποτελεσμάτων Ταξινόμησης, Εσωτερικοί δείκτες ομαδοποίησης εικονοστοιχείων μη επισημασμένων εικόνων σε μορφή πίνακα)

Όσον αφορά την οπτικοποίηση των παραπάνω αλγόριθμων, δεν μπορεί να διαπιστωθεί επακριβώς η ορθότητα της. Ωστόσο παρατίθενται οι αρχικές εικόνες σε μορφή RGB, τα αποτελέσματα των CAM στα δύο μοντέλα συνελκτικών δικτύων, καθώς και το αποτέλεσμα της ομαδοποίησης, χωρίς ωστόσο να ονομάζεται κάποια από τις ομάδες ως φθορά.

Τα αποτελέσματα παρουσιάζονται για κάθε εικόνα ξεχωριστά παρακάτω, στις εικόνες 43 έως και 47.

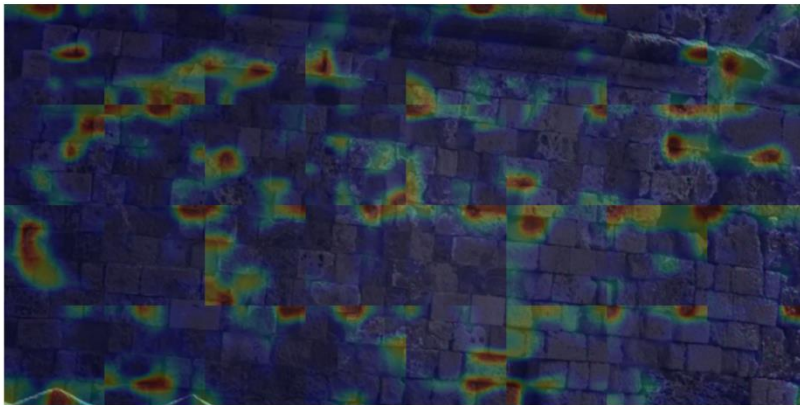


Εικόνα 43: Οπτική παράθεση μεθόδων εντοπισμού φθορών, agiosnik002211
α) εικόνα RGB, β) CAM μοντέλου RGB, γ) CAM μοντέλου NIR, δ) Αποτέλεσμα ομαδοποίησης.

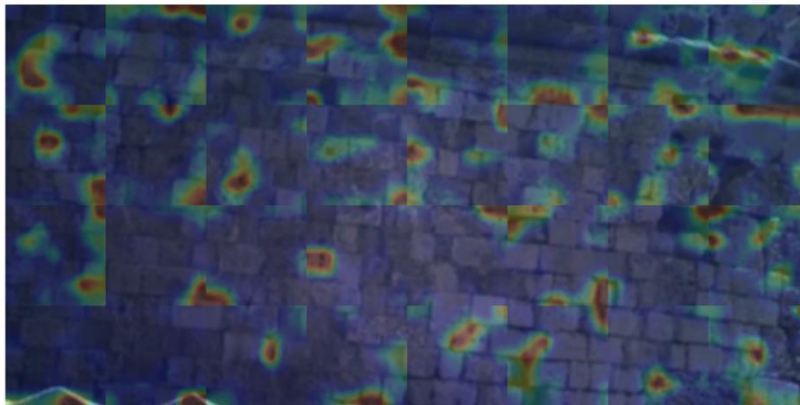
α)



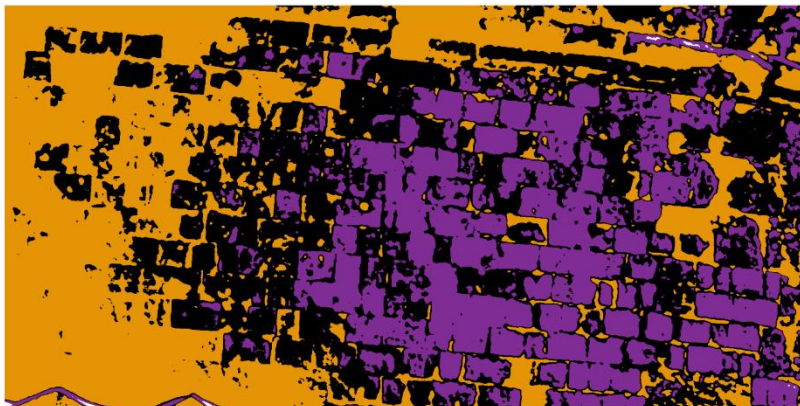
β)



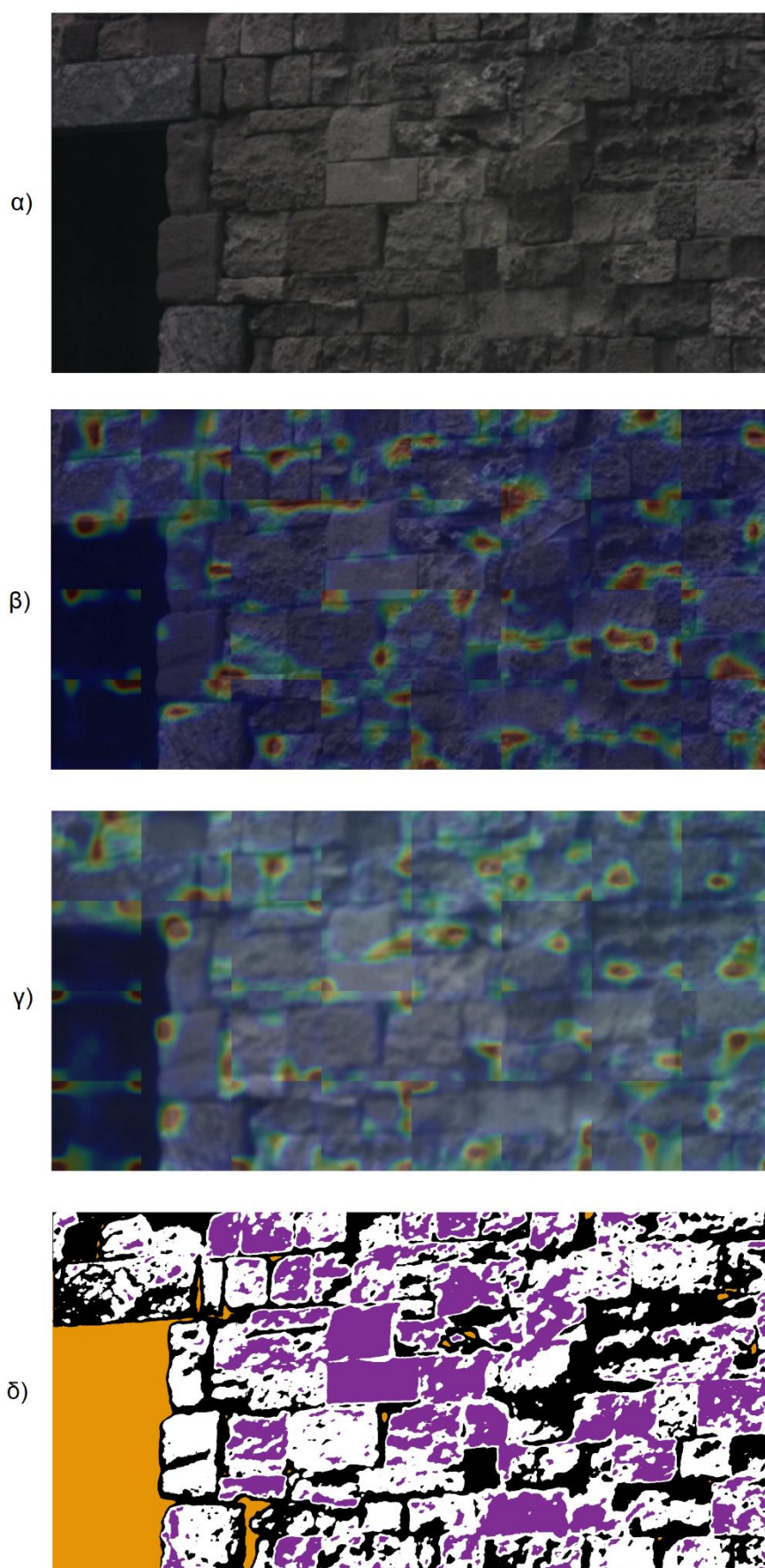
γ)



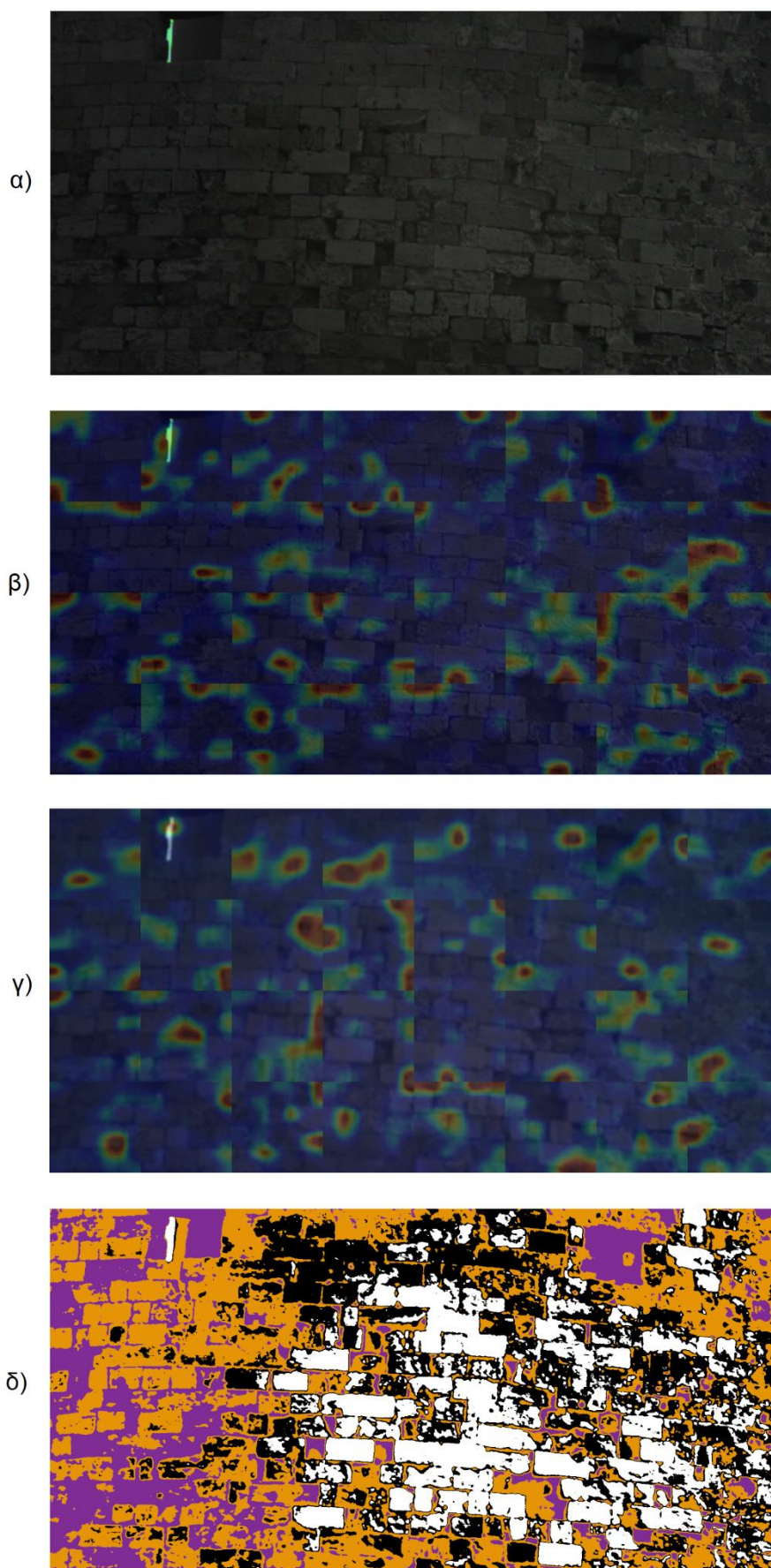
δ)



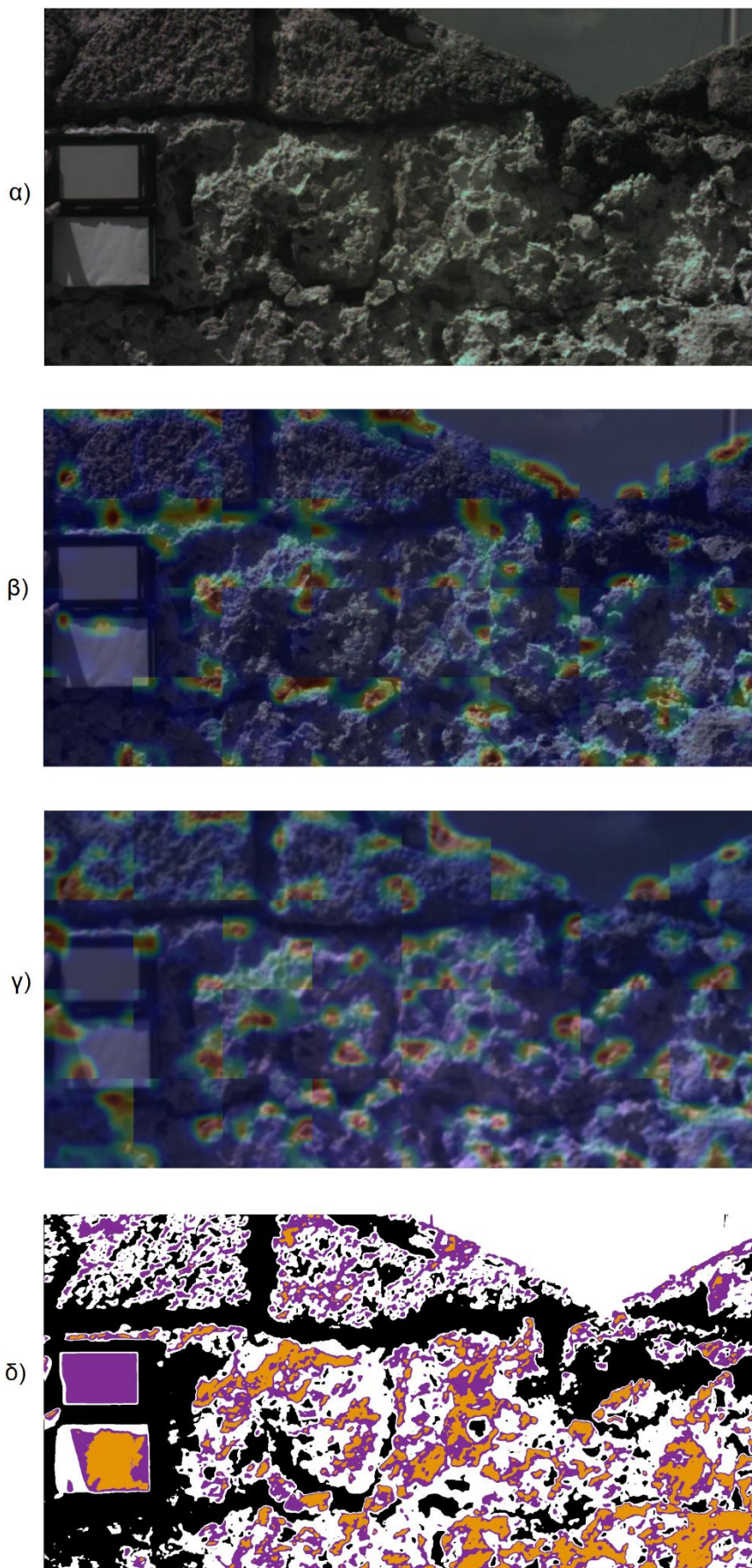
Εικόνα 44: Οπτική παράθεση μεθόδων εντοπισμού φθορών, agiosnik002948
α) εικόνα RGB, β) CAM μοντέλου RGB, γ) CAM μοντέλου NIR, δ) Αποτέλεσμα ομαδοποίησης



Εικόνα 45: Οπτική παράθεση μεθόδων εντοπισμού φθορών, agiosnik003218
α) εικόνα RGB, β) CAM μοντέλου RGB, γ) CAM μοντέλου NIR, δ) Αποτέλεσμα ομαδοποίησης



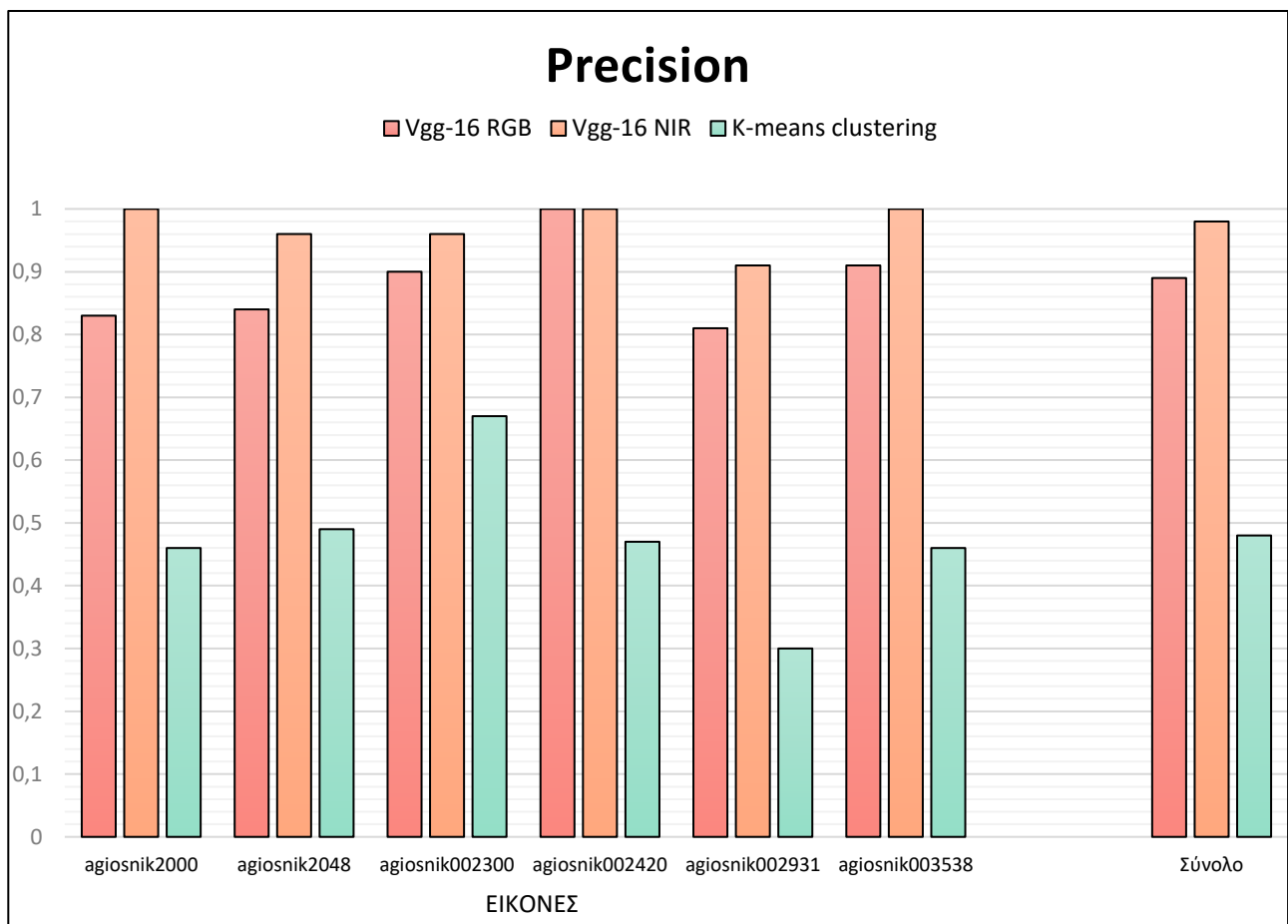
Εικόνα 46: Οπτική παράθεση μεθόδων εντοπισμού φθορών, agiosnik003340
α) εικόνα RGB, β) CAM μοντέλου RGB, γ) CAM μοντέλου NIR, δ) Αποτέλεσμα ομαδοποίησης



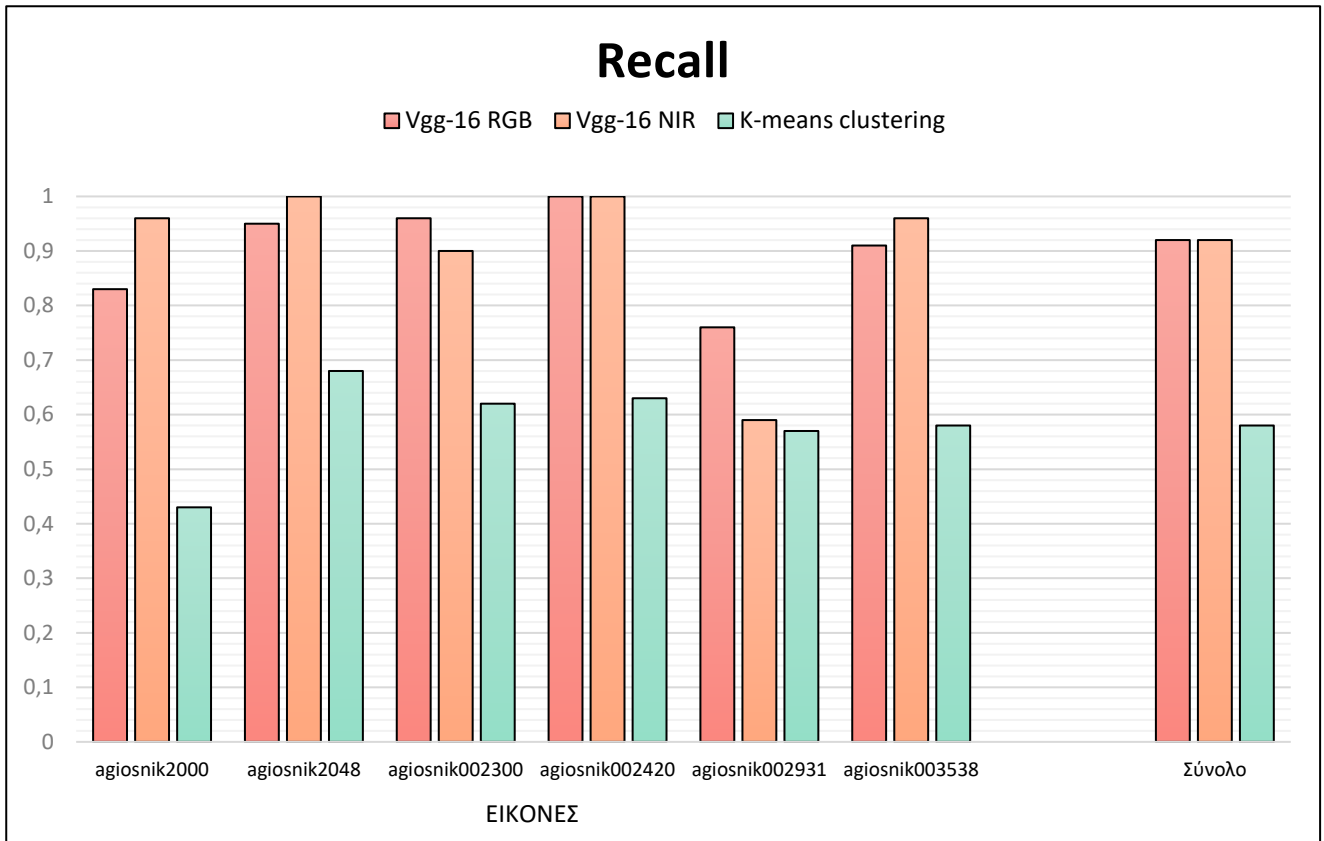
Εικόνα 47: Οπτική παράθεση μεθόδων εντοπισμού φθορών, agiosnik004719
α) εικόνα RGB, β) CAM μοντέλου RGB, γ) CAM μοντέλου NIR, δ) Αποτέλεσμα ομαδοποίησης

6.4 Συγκριτική Ανάλυση αποτελεσμάτων

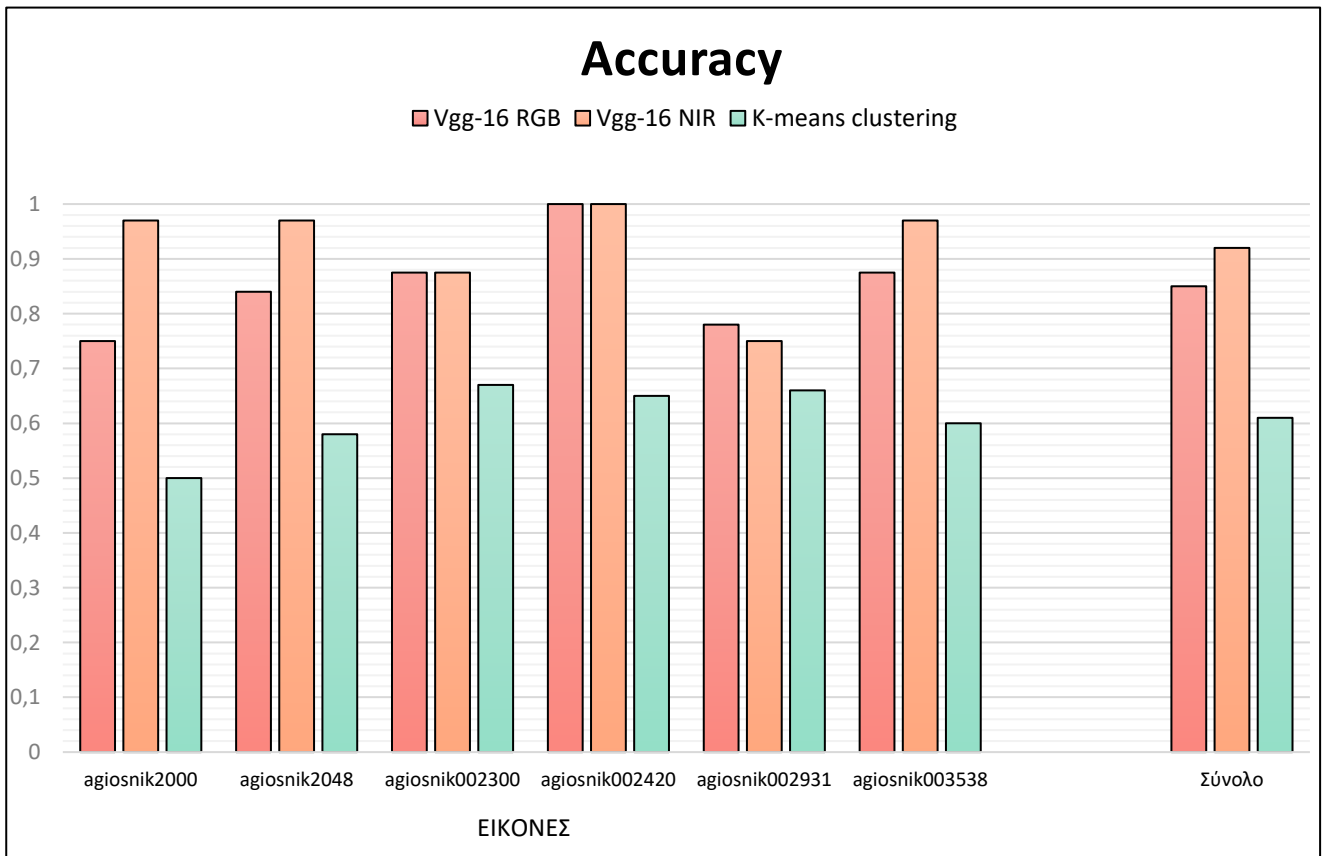
Καθώς έχουν υπολογιστεί οι δείκτες απόδοσης της ταξινόμησης για τις εικόνες που διαθέτουν επισημάνσεις, παρατίθενται συγκεντρωτικά στα παρακάτω διαγράμματα, ώστε να αναλυθούν στη συνέχεια οι τιμές τους. Κάθε διάγραμμα των εικόνων 36, 37, 38 και 39 απεικονίζει έναν από τους δείκτες precision, recall, accuracy και F1 score των 3 μεθόδων ταξινόμησης των δεδομένων, δηλ. των δύο συνελκτικών δικτύων και της ομαδοποίησης, για κάθε εικόνα ξεχωριστά και στη συνέχεια συνολικά για όλες τις εικόνες. Οι τιμές των μοντέλων συνελκτικών νευρωνικών δικτύων αναμένονται να είναι υψηλότερες αφού εφαρμόζονται σε δεδομένα εκπαίδευσης με πιθανή υπερπροσαρμογή, ενώ οι τιμές της ομαδοποίησης αναμένεται να είναι χαμηλότερες αφού εφαρμόζεται ο αλγόριθμος K-means, ο οποίος είναι ταχύτερος αλλά λιγότερο ακριβής.



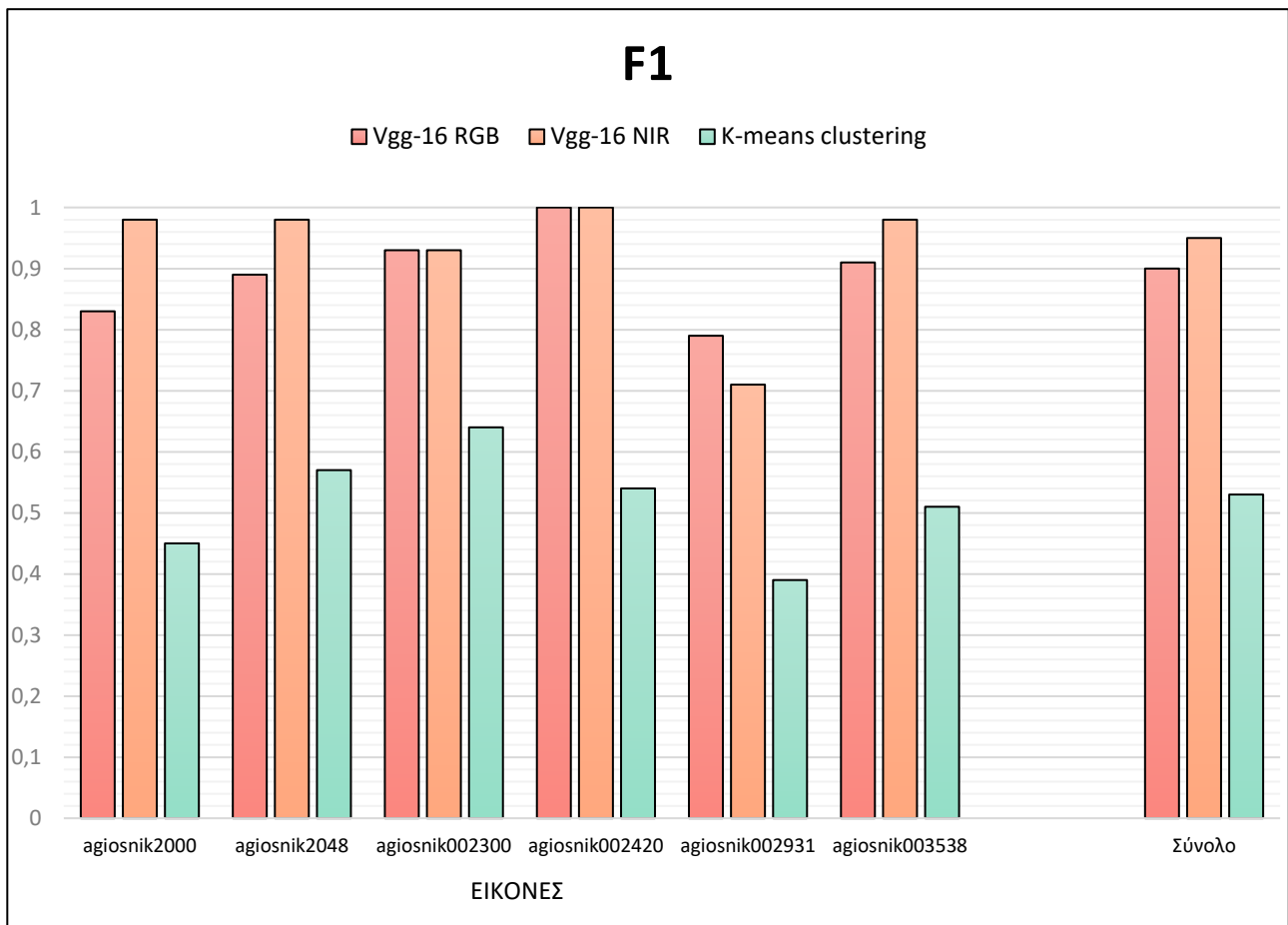
Εικόνα 48: Δείκτης precision για τις ταξινομήσεις μοντέλου RGB, μοντέλου NIR και ομαδοποίησης εικονοστοιχείων με K-means



Εικόνα 49: Δείκτης recall για τις ταξινομήσεις μοντέλου RGB, μοντέλου NIR και ομαδοποίησης εικονοστοιχείων με K-means



Εικόνα 50: Δείκτης accuracy για τις ταξινομήσεις μοντέλου RGB, μοντέλου NIR και ομαδοποίησης εικονοστοιχείων με K-means



Εικόνα 51: Δείκτης F1 για τις ταξινομήσεις μοντέλου RGB, μοντέλου NIR και ομαδοποίησης εικονοστοιχείων με K-means

Με μια πρώτη ματιά φαίνεται ότι οι δείκτες της ταξινόμησης με τα συνελκτικά δίκτυα είναι καλύτεροι από αυτούς της ομαδοποίησης. Ωστόσο, πέρα από τις συνθήκες που έχουν αναφερθεί, το φαινόμενο αυτό είναι σε μεγάλο βαθμό αναμενόμενο από τη φύση του προβλήματος. Όσο μικρότερη είναι η μονάδα που ταξινομείται, τόσο χειροτερεύει η ακρίβεια της ταξινόμησης, καθώς γίνεται πιο δύσκολο το εγχείρημα. Η ταξινόμηση μιας ολόκληρης εικόνας σε φθορά ή όχι φθορά είναι λογικό να γίνει με μεγαλύτερη ακρίβεια από ότι η ταξινόμηση μικρότερων παραθύρων σε φθορά ή όχι φθορά, η οποία με την σειρά της θα είναι πιο ακριβής από τη σημασιολογική κατάτμηση σε επίπεδο εικονοστοιχείου.

Τα δύο μοντέλα συνελκτικών δικτύων φαίνεται να ακολουθούν παρόμοιες διακυμάνσεις δεικτών ανά εικόνα, με το δίκτυο εκπαιδευμένο σε εγγύς υπέρυθρα να εμφανίζει συνήθως καλύτερους δείκτες λόγω υπερπροσαρμογής. Η καλύτερη εικόνα των δύο δικτύων είναι η εικόνα agiosnik002420, η οποία σύμφωνα με τις επισημάνσεις παρουσιάζει φθορές σε όλη την επιφάνεια της, το οποίο εκτιμούν και τα μοντέλα, που προβλέπουν ως φθορά όλα τα παράθυρα της εικόνας. Αντίθετα, η χειρότερη εικόνα τους από άποψη δεικτών είναι η εικόνα agiosnik002931, καθώς εκεί εμφανίζουν τους χειρότερους δείκτες τους, με μόνη εξαίρεση την ορθότητα (accuracy) του μοντέλου εκπαιδευμένου σε εικόνες RGB, καθώς εμφανίζει ακόμα χαμηλότερη στην εικόνα agiosnik2000. Είναι ενδεικτικό ότι μόνο στην εικόνα agiosnik002931 η ορθότητα της ομαδοποίησης με K-means είναι σχεδόν ίση με τις ορθότητες των συνελκτικών δικτύων.

Η ομαδοποίηση των εικονοστοιχείων με K-means εμφανίζει τους καλύτερους της δείκτες στην εικόνα agiosnik002300, όπου επιδεικνύει σχετικά χαμηλότερη σύγχυση φθορών με μη φθορές, ενώ χειρότερους δείκτες εμφανίζει στην εικόνα agiosnik2000, με μόνη εξαίρεση την ύπαρξη χαμηλής ακρίβειας (precision) στην εικόνα agiosnik002931 λόγω μεγάλης παρουσίας ψευδώς θετικών (False Positives). Παρόλα αυτά, η ορθότητα (accuracy) της ομαδοποίησης συνολικά στην εικόνα agiosnik002931 είναι καλή.

Μεγάλο ενδιαφέρον έχει η σύγκριση των εσωτερικών δεικτών εγκυρότητας της ομαδοποίησης με τους εξωτερικούς δείκτες απόδοσης. Οι εσωτερικοί δείκτες δείχνουν πόσο ξεκάθαρες και διαχωρισμένες είναι οι ομάδες, ενώ οι εξωτερικοί δείκτες δείχνουν πόσο ανταποκρίνονται οι ομάδες στην πραγματικότητα. Τους καλύτερους δείκτες εσωτερικής εγκυρότητας τους παρουσιάζει η ομαδοποίηση της εικόνας agiosnik2048, ενώ συγκεντρωτικά τους χειρότερους η ομαδοποίηση της εικόνας agiosnik002300. Ωστόσο η τιμές των εξωτερικών δεικτών είναι αντιστρόφως ανάλογες με τις τιμές των παραπάνω εσωτερικών δεικτών. Η εικόνα agiosnik002300 έχει την μεγαλύτερη ορθότητα ταξινόμησης εικονοστοιχείων, ενώ η εικόνα agiosnik2048 έχει σημαντικά χαμηλότερη ορθότητα. Επομένως δεν υπάρχει άμεση σύνδεση των καλών εσωτερικών δεικτών με την καλή ομαδοποίηση με βάση τις πραγματικές κατηγορίες.

Όπως έχει αναφερθεί, στο πρόβλημα παρακολούθησης δομικής ακεραιότητας σημαντικότερο σφάλμα είναι η ύπαρξη ψευδώς αρνητικών προβλέψεων (υποτίμηση των κινδύνων). Επομένως σημαντικότερος δείκτης είναι η ανάκληση (recall) από την ακρίβεια (precision). Στα διαγράμματα των δεικτών, το συνελικτικό μοντέλο NIR είναι το μόνο που παρουσιάζει υψηλότερη ακρίβεια, ενώ το μοντέλο RGB και η ομαδοποίηση K-means εμφανίζουν υψηλότερη ανάκληση. Αξίζει να σημειωθεί ότι όσον αφορά την ομαδοποίηση, μπορούν να μειωθούν σημαντικά τα ψευδώς αρνητικά εις βάρος των ψευδώς θετικών, θέτοντας απλά περισσότερες από τις ομάδες ως κατηγορίες φθορών. Έτσι, βελτιώνεται ο δείκτης ανάκλησης, χειροτερεύοντας ωστόσο τον δείκτη ακρίβειας και πιθανότατα την συνολική ορθότητα των προβλέψεων. Ομοίως, κατά την εκπαίδευση των μοντέλων, παράθυρα που εμφάνιζαν υπερβολικά μικρή κάλυψη φθορών θεωρήθηκαν ως παράθυρα χωρίς φθορές, τα οποία στην θεωρία εισάγουν ψευδώς αρνητικά στοιχεία στην εκπαίδευση. Αυτό χειροτερεύει τις επιδόσεις των δικτύων στην ανάκληση, και υποθέτοντας τα παράθυρα αυτά ως φθορές στην εκπαίδευση μπορεί να αυξηθεί ο συγκεκριμένος δείκτης, χειροτερεύοντας ωστόσο την ακρίβεια. Πρακτικά όμως τα στοιχεία αυτά είναι αδιάφορα ως προς την παρουσία φθορών και άρα είναι πιθανότερο να μπερδέψουν τα μοντέλα και την εκπαίδευσή τους, αν θεωρηθούν ως έχοντα φθορές.

Στο ζήτημα του εντοπισμού, τα συνελικτικά μοντέλα αγνοούν επαρκώς κατηγορίες που προφανέστατα δεν είναι φθαρμένα δομικά στοιχεία, όπως ο ουρανός ή οι σκιές μέσα σε πύλες, ωστόσο ξεκάθαρες φθορές που μπορούν να φανούν με γυμνό μάτι δεν τις διαχωρίζουν ικανοποιητικά από τα υπόλοιπα υγιή δομικά στοιχεία. Το μοντέλο RGB φαίνεται να εντοπίζει μέσω CAM τις φθορές που έχουν επισημανθεί με λευκό χρώμα σε γενικές γραμμές, ενώ το NIR τις εντοπίζει σχετικά χειρότερα. Ακόμα, το CAM έχει την τάση να δείχνει μόνο τις ακμές των φθορών και όχι την συνολική τους επιφάνεια, άλλος ένας λόγος που οι λευκές φθορές εντοπίζονται συχνότερα λόγω της διακύμανσης τους στην υφή. Τα αποτελέσματα χωρικού εντοπισμού μέσω CAM στις υπόλοιπες κατηγορίες φθορών όμως είναι απογοητευτικά, ενώ εμφανίζονται πολλές λανθασμένες ενεργοποιήσεις στα όρια των παραθύρων εκπαίδευσης που συνενώνονται για την δημιουργία της συνολικής απεικόνισης. Η κατηγοριοποίηση των

εικονοστοιχείων σε φθορές μέσω ομαδοποίησης K-means φαίνεται να έχει καλύτερα αποτελέσματα σε ότι αφορά τον ακριβή διαχωρισμό φθαρμένων τοιχωμάτων από τα υγιή στις περισσότερες κατηγορίες φθορών, ειδικά όταν αυτές καλύπτουν μεγάλες επιφάνειες. Ωστόσο, η κατηγορία φθορών που έχει επισημανθεί με πράσινο χρώμα, δεν κατηγοριοποιείται ως φθορά στις περισσότερες περιπτώσεις και δεν διαχωρίζεται από τα υγιή δομικά στοιχεία ούτε στην υλοποίηση μη επιβλεπόμενης μάθησης, τουλάχιστον με ομαδοποίηση σε 4 φασματικές κατηγορίες.

Για τις μη επισημασμένες εικόνες, φαίνεται από την ταξινόμηση των δύο συνελικτικών δικτύων ότι και τα δύο μοντέλα είναι προκατειλημμένα στο να βρίσκουν φθορές, το οποίο είναι άλλο ένα σημάδι ότι δεν μπορούν να τα διαχωρίσουν επαρκώς από τα απλά δομικά στοιχεία. Από τα δύο, το μοντέλο των υπέρυθρων φασματικών κανάλιων εμφανίζει περισσότερα αρνητικά αποτελέσματα, δηλαδή κατατάσσει περισσότερα τμήματα εικόνων ως μη έχοντα φθορές. Ωστόσο σε γενικές γραμμές τα αποτελέσματα των δύο μοντέλων στην ταξινόμηση φαίνεται να συμφωνούν, καθώς μόνο 17 από τα 160 παράθυρα εικόνων έχουν ταξινομηθεί σε διαφορετικές κατηγορίες.

Όσον αφορά τον εντοπισμό φθορών σε αυτές, αν και δεν μπορούν να ερμηνευτούν με βεβαιότητα τα αποτελέσματα, οπτικά φαίνεται να διαχωρίζονται τα δομικά στοιχεία σε κάποιες διαφορετικές κατηγορίες. Στην εικόνα agiosnik002211, τόσο τα συνελικτικά δίκτυα όσο και η ομαδοποίηση διαχωρίζουν τις πέτρινες επιφάνειες από τις σκιασμένες πύλες. Τα συνελικτικά δίκτυα εντοπίζουν φθορές στα δομικά στοιχεία, ενώ η ομαδοποίηση διαχωρίζει τα δομικά στοιχεία σε δύο κατηγορίες, που θα μπορούσαν να αντιστοιχούν σε φθαρμένες και μη επιφάνειες. Στις υπόλοιπες μη επισημασμένες εικόνες τα αποτελέσματα των συνελικτικών δικτύων στον εντοπισμό δεν φαίνεται να βγάζουν ιδιαίτερο νόημα, πέρα από πολύ φανερά μοτίβα στα πετρώματα. Ακόμα η ομαδοποίηση φαίνεται να επηρεάζεται περισσότερο από την εναλλαγή σκιασμένων και φωτισμένων από τον ήλιο επιφανειών, παρά από την σύνθεση των πετρωμάτων. Επομένως και τα αποτελέσματα της ομαδοποίησης είναι αμφισβητήσιμα, ειδικά όταν υπάρχουν εικονοστοιχεία που απεικονίζουν τον ουρανό.

7. Συμπεράσματα

7.1 Γενικά

Τα κύρια ευρήματα τις εργασίας είναι οι τιμές των σημαντικών δεικτών ανάκλησης (recall) και ορθότητας (accuracy) που προκύπτουν από την ταξινόμηση στοιχείων πολυφασματικών εικόνων σε φθορές ή μη φθορές. Για το μοντέλο συνελκτικών δικτύων εκπαιδευμένο σε φασματικά κανάλια RGB αυτές είναι συνολικά για όλες τις εικόνες 92% recall, 85% accuracy, για το μοντέλο εκπαιδευμένο σε υπέρυθρα φασματικά κανάλια αυτές είναι 92% recall και 92% accuracy, ενώ για την μέθοδο σημασιολογικής κατάτμησης με ομαδοποίηση K-means αυτές είναι 58% recall και 61% accuracy.

Από τα αποτελέσματα τόσο μεμονωμένα των αλγορίθμων μηχανικής μάθησης, όσο και της σύγκρισης των μεθόδων, προκύπτει ότι οι διαφορετικοί τύποι φθορών που έχουν επισημανθεί δεν μπορούν να ενταχθούν στην ίδια κατηγορία προβλέσεων για ταξινόμηση, καθώς παρουσιάζουν μεγάλες διαφορές μεταξύ τους. Μεγαλύτερη επιτυχία ενδέχεται να έχει η πολυταξική ταξινόμηση σε 4 τύπους φθορών όπως επισημαίνονται με διαφορετικά χρώματα, σε 5^η κατηγορία όχι φθορά, και σε πιθανή 6^η κατηγορία που είναι ούτε άφθαρτη λιθοδομή, ούτε φθαρμένη (παράδειγμα ο ουρανός, το έδαφος, τα πλήρως σκιασμένα εσωτερικά δωμάτια).

Φαίνεται ακόμα ότι η υπόθεση ότι μπορούν να αξιοποιηθούν εγγύς υπέρυθρα φασματικά κανάλια μέσω κλασσικών αρχιτεκτονικών συνελκτικών δικτύων δεν ευσταθεί. Παρά τους υψηλούς δείκτες λόγω υπερπροσαρμογής, η προσπάθεια να χτιστεί ένα μοντέλο το οποίο θα βρίσκει μοτίβα στο υπέρυθρο φάσμα δεν απέδωσε ικανοποιητικά στον εντοπισμό, καθώς οποιοδήποτε θετικό στοιχείο μπορεί να προκύψει από την αξιοποίηση υπέρυθρων φασματικών καναλιών, αντισταθμίζεται από την απώλεια ξεκάθαρης πληροφορίας λόγω της χαμηλότερης ανάλυσης των εικόνων σε αυτό το φάσμα.

Ένα συμπέρασμα που προκύπτει για την ομαδοποίηση των εικονοστοιχείων της εικόνας, είναι ότι η καλοί δείκτες εσωτερικής εγκυρότητας, δεν μεταφράζονται σε καλά αποτελέσματα εξωτερικών δεικτών και εντοπισμού των φθορών, τουλάχιστον όταν η ομαδοποίηση εφαρμόζεται στο σύνολο των 41 φασματικών καναλιών. Αντιπροσωπευτικά, η εικόνα agiosnik2048 που παρουσιάζει τους βέλτιστους δείκτες CHI (3.875.641) και DBI (0,66) για τα δεδομένα της εργασίας, επιτυγχάνει μόνο 58% ορθότητα προβλέσεων. Επίσης, η υψηλή ακρίβεια κατά την εκπαίδευση συνελκτικών νευρωνικών δικτύων δεν σημαίνει και καλό εντοπισμό των φθορών, ακόμα και αν οι δείκτες ταξινόμησης είναι υψηλοί. Ενδέχεται αυτοί να είναι αποτέλεσμα υπερπροσαρμογής των μοντέλων, η οποία τα οδηγεί να εστιάζουν και να εντοπίζουν τα λάθος μοτίβα στις εικόνες που αποτελούν παραδείγματα εκπαίδευσης.

Τέλος, η κατηγορία φθορών που έχει επισημανθεί με πράσινο στις εικόνες επισήμανσης είναι σχεδόν αδύνατο να προσδιορισθεί τόσο με τα συνελκτικά δίκτυα της εργασίας, όσο και με την φασματική ομαδοποίηση. Αν συνυπολογιστεί και το γεγονός ότι οι φθορές αυτού του τύπου δεν διακρίνονται άμεσα με το ανθρώπινο μάτι στις εικόνες RGB μερικές φορές, είναι πιθανό να χρειάζεται εντελώς διαφορετική προσέγγιση ανίχνευσης από απόσταση για αυτές.

Αλλιώς ο μόνος τρόπος εντοπισμού τους θα εξακολουθεί να είναι η εργασία πεδίου και οι καταστροφικές μέθοδοι αξιολόγησης της κατάστασης των πετρωμάτων μνημείων.

7.2 Δυσκολίες κατά την εκπόνηση εργασίας

Ιδανικά, η σύγκριση των μεθόδων επιβλεπόμενης και μη επιβλεπόμενης μηχανικής μάθησης, θα γινόταν σε όσο το δυνατό πιο όμοια δεδομένα εισόδου. Αυτό προϋποθέτει ότι τόσο το συνελκτικό νευρωνικό δίκτυο, όσο και η ομαδοποίηση, θα εφαρμόζονταν στα ίδια ακριβώς αρχεία ή, τουλάχιστον, σε αρχεία με το ίδιο format, δηλαδή αρχεία .tif. Ωστόσο, η πλειοψηφία των προεκπαιδευμένων διαθέσιμων συνελκτικών μοντέλων δέχεται ως είσοδο αρχεία με format .png ή .jpeg. Επομένως για την αξιοποίησή τους, τα δεδομένα μεταμορφώθηκαν από το αρχικό τους format στο άλλο, με πιθανή αλλοίωση τιμών λόγω συμπίεσης τους. Η αδυναμία στην χρήση του αρχικού τύπου αρχείου, έγκειται στην ανάγκη να χτιστεί ένα μοντέλο το οποίο διαχειρίζεται δεδομένα διαφορετικού τύπου από τα προεκπαιδευμένα μοντέλα που χρησιμοποιήθηκαν στην εργασία. Αυτό απαιτεί αφενός αυξημένες γνώσεις επί της επιστήμης υπολογιστών και της επιστήμης δεδομένων, αφετέρου μεγάλο όγκο δεδομένων εκπαίδευσης, ώστε να υπολογιστούν βάρη σε πολύ περισσότερα επίπεδα, από αυτά που απαιτεί η μεταφορά μάθησης που πραγματοποιήθηκε. Ακόμα, η εκπαίδευση διαφορετικού μοντέλου απαιτεί τεράστιο υπολογιστικό χρόνο και ισχύ, πόροι οι οποίοι την περίοδο παρουσίας της εργασίας είναι περιορισμένοι.

Παρομοίως, ο ηλεκτρονικός υπολογιστής στον οποίον διενεργήθηκε η επεξεργασία των πολυφασματικών εικόνων και η μηχανική μάθηση, αν και επαρκής σε γενικές γραμμές, σε συγκεκριμένες συγκυρίες εισήγαγε περιορισμούς λόγω των εξαρτημάτων του. Αυτό παρουσιάστηκε σε ορισμένους αλγόριθμους ομαδοποίησης, όπως ο DBSCAN, και η εκπαίδευση πλήρους μοντέλου VGG χωρίς πάγωμα παραμέτρων, καθώς οι διεργασίες αυτές είναι υπολογιστικά πολύ έντονες λόγω του πλήθους των παραμέτρων που πρέπει να υπολογιστούν. Ενδεικτικά για την ομαδοποίηση, κάθε εικόνα περιέχει 1.910.232 εικονοστοιχεία, με 41 κανάλια το καθένα, άρα σχεδόν 2 εκατομμύρια διανύσματα με 41 χαρακτηριστικά το καθένα. Πέρα από ορισμένους αλγόριθμους ομαδοποίησης, το γεγονός αυτό δυσχεραίνει και τον υπολογισμό ορισμένων δεικτών ομαδοποίησης, όπως για παράδειγμα τη μέση τιμή σιλουέτας.

Όσον αφορά την εφαρμογή συνελκτικών δικτύων στα δεδομένα της εργασίας, οι φθορές είναι υπερβολικά ανόμοιες και με ποικιλία σχημάτων και υφών επιφανειών. Πολλές φορές δεν διακρίνονται οι φθορές από τις «υγιείς» πέτρες με το ανθρώπινο μάτι, το οποίο καθιστά την εφαρμογή συνελκτικών δικτύων δύσκολη. Τέλος, το σύνολο δεδομένων είναι υπερβολικά μικρό για το πόσο περίπλοκο ήταν το εγχείρημα που επιχειρήθηκε στην εργασία, ενώ οι φωτογραφίες επισημάνσεων δεν είναι σαφείς ως προς τι επισημαίνουν και με πιο χρώμα κάθε φορά.

7.3 Κατευθύνσεις μελλοντικής έρευνας

Όσον αφορά την επιβλεπόμενη μάθηση, ένα μοντέλο προβλέψεων με νευρωνικά δίκτυα ποτέ δεν είναι άψογο, και άρα πάντα επιδέχεται βελτίωση με κάποιον τρόπο. Επομένως εξ ορισμού, η φυσική πορεία των ερευνών είναι η βελτίωση των μοντέλων, μέσω δοκιμής υπερπαραμέτρων, περισσότερων παραδειγμάτων εκπαίδευσης και καλύτερη προσαρμογή στα επιθυμητά δεδομένα εισόδου.

Στο αντικείμενο της παρούσας εργασίας, πέρα από συνεχείς δοκιμές εκπαίδευσης για καλύτερα όμοια μοντέλα, η έρευνα μπορεί να επικεντρωθεί σε άλλα στοιχεία. Αρχικά, αντί για τα φασματικά κανάλια RGB και τα επιλεγμένα κανάλια NIR 17,27,37, μπορούν να δοκιμαστούν στην ίδια αρχιτεκτονική VGG-16 διαφορετικοί συνδυασμοί φασματικών καναλιών, ακόμα και μεικτά, δηλαδή 2 από το ορατό και 1 από το εγγύς υπέρυθρο ή αντίστροφα. Κάποιος συγκεκριμένος συνδυασμός μπορεί να αποδώσει καλύτερα στην υπόθεση της εργασίας, και να εντοπίζει ορθότερα τις φθορές στις εικόνες. Χρησιμοποιώντας την εξίσωση πλήθους συνδυασμών προκύπτει ο αριθμός των διαφορετικών πιθανών συνδυασμών φασματικών καναλιών που μπορούν να αξιοποιηθούν:

$$\binom{41}{3} = \frac{41!}{3!(41-3)!} = 10.660 \text{ Συνδυασμοί}$$

Όπου $n = 41$ ο συνολικός αριθμός καναλιών και $k = 3$ αυτά που επιλέγονται για το μοντέλο. Ακόμα, μπορεί να διερευνηθεί η εφαρμογή των μοντέλων αυτών για πολυταξική ταξινόμηση και εντοπισμό των διαφορετικών ειδών φθορών, όπως αυτά προσδιορίζονται από τα διαφορετικά χρώματα των εικόνων επισημάνσεων. Αν αντί για το τελικό διάνυσμα δύο στοιχείων, δηλαδή φθορά ή όχι φθορά, χρησιμοποιηθεί ένα διάνυσμα επισημάνσεων με περισσότερα στοιχεία, το μοντέλο μπορεί να εντοπίσει ποιες φθορές και σε ποιο βαθμό εντοπίζονται σε κάθε εικόνα. Για τον σκοπό αυτόν όμως, απαιτείται σημαντικά μεγαλύτερος όγκος δεδομένων εκπαίδευσης.

Σε επόμενο επίπεδο, μπορεί διαφορετική αρχιτεκτονική να μπορεί να προσαρμοστεί καλύτερα στις εικόνες και να τις συνδέει με τη σωστή ταξινόμηση και εντοπισμό των φθορών. Πολύ δημοφιλής επιλογή σε διαφορετικά επιστημονικά πεδία αποτελεί η αρχιτεκτονική Resnet-50, η οποία ωστόσο δέχεται τις ίδιες διαστάσεις δεδομένων εισόδου με το VGG, δηλαδή $224 \times 224 \times 3$. Ωστόσο νέες πειραματικές αρχιτεκτονικές μελετώνται διαρκώς, και μερικές φιλοδοξούν να έχουν σεβαστά αποτελέσματα ακόμα και σε υπερφασματικές εικόνες. Ενδεικτικά παρατίθεται το συνελκτικό νευρωνικό δίκτυο της δημοσίευσης «Convolutional neural networks for hyperspectral image classification»⁶⁵, το οποίο έχει σχεδιαστεί για εφαρμογές σε πολυφασματικά δεδομένα κάλυψης γης. Το συνελκτικό δίκτυο αυτό δέχεται ως είσοδο παράθυρα εικόνων με $5 \times 5 \times N$ εικονοστοιχεία, όπου N ο αριθμός καναλιών, 41 για τα αρχικά δεδομένα αυτής της εργασίας. Αν και δεν έχει σχεδιαστεί για τα δεδομένα των φθορών σε μνημεία, οι συγγραφείς της δημοσίευσης δηλώνουν αισιόδοξοι ως προς την πιθανή εφαρμογή συνελκτικών δικτύων σε πλήρη υπερφασματικά δεδομένα, τα οποία παραμελήθηκαν με την χρήση των προτύπων συνελκτικών δικτύων στην προκειμένη περίπτωση.

Για την μη επιβλεπόμενη μάθηση, η μελλοντική έρευνα μπορεί να επικεντρωθεί στον καταλληλότερο αριθμό των φασματικών καναλιών που συμμετέχουν στην ομαδοποίηση, καθώς και ποια είναι αυτά. Ακόμα, μπορεί να μελετηθεί ο καλύτερος αλγόριθμος ομαδοποίησης για κάθε συνδυασμό φασματικών καναλιών. Το αποσκοπούμενο αποτέλεσμα της έρευνας, είναι ο αλγόριθμος ομαδοποίησης ο οποίος λαμβάνοντας υπόψιν τα κατάλληλα φασματικά κανάλια, μπορεί να προσδιορίσει ακριβώς τις φασματικές υπογραφές των σημείων που έχουν φθαρεί στα μνημεία και άρα να τα επισημάνει στις σημασιολογικά καταταμημένες εικόνες που προκύπτουν από την ομαδοποίηση.

Τέλος, ο απώτερος σκοπός όλων των ερευνών είναι η δημιουργία ενός αυτοματοποιημένου συστήματος, που μπορεί να αναλάβει τόσο τον εντοπισμό φθορών όσο και την ειδοποίηση φορέων, χωρίς την απαίτηση ανθρώπινης παρέμβασης ανάμεσα στα δύο σημεία. Όλες οι συγκριτικές μελέτες οδηγούν στην αύξηση της ακρίβειας και τη μείωση των σφαλμάτων και επιπλοκών, τα οποία είναι το πρώτο βήμα στην προσπάθεια για την εγκαθίδρυση ενός γενικότερου συστήματος διαχείρισης μνημείων πολιτιστικής κληρονομιάς.

Βιβλιογραφία

1. Bellentani, F., Panico, M. The meanings of monuments and memorials: toward a semiotic approach. *Punctum*, 2(1): 28-46, 2016
DOI: 10.18680/hss.2016.0004
2. T. Stabolov, J.R.J van Asperen de Boer. (1976). "The Deterioration and Conservation of Porous Building Materials in Monuments. A Review of Literature". International Centre for the Study of the Preservation of Cultural Property, Rome.
3. Lynas, Mark & Houlton, Benjamin & Perry, Simon. (2021). Greater than 99% consensus on human caused climate change in the peer-reviewed scientific literature. *Environmental Research Letters*. 16. 114005. 10.1088/1748-9326/ac2966.
4. Themistocleous, Kyriacos & Danezis, Chris. (2018). Monitoring cultural heritage sites affected by geohazards. 9. 10.1117/12.2325455
5. Marcos L.S. Oliveira, Alcindo Neckel, Diana Pinto, Laércio Stolfo Maculan, Matheus Roberto Dalmagro Zanchett, Luis F.O. Silva, "Air pollutants and their degradation of a historic building in the largest metropolitan area in Latin America", *Chemosphere*, Volume 277, 2021, 130286,
<https://doi.org/10.1016/j.chemosphere.2021.130286>.
6. Γιαουτζή, Μ., Στρατηγέα, Α. (2011). Χωροταξικός σχεδιασμός - Θεωρία και πράξη. Εκδόσεις Κριτική ΑΕ
7. Copeland, B. artificial intelligence. *Encyclopedia Britannica*.
<https://www.britannica.com/technology/artificial-intelligence>
[Accessed 19 1 2022]
8. Foote, K. D. A Brief History of Machine Learning. DATAVERSITY.
<https://www.dataversity.net/a-brief-history-of-machine-learning/>
[Accessed 19 1 2022]
9. Hebb, D. (1949). *The organization of behavior*. Psychology Press.
10. Αλεξάκης, Γ., Κορακοβούνης, Δ. (2021). Εφαρμογή τεχνητών νευρωνικών δικτύων αιχμών σε δεδομένα δυναμικών αισθητήρων όρασης.
11. Samuel, A. L. (1959). Some Studies in Machine Learning using the Game of Checkers. *IBM Journal of Research and Development*.
12. Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
13. Schapire, R.E. The strength of weak learnability. *Mach Learn* 5, 197–227 (1990).
<https://doi.org/10.1007/BF00116037>

14. Freund, Y., Schapire, R. E., Experiments with a new boosting algorithm. In *icml* (vol. 96, pp. 148 – 156).
15. Burkov, A. (2019). *The Hundred-Page Machine Learning Book*
16. James, G. (2013). *An Introduction to Statistical Learning: with Applications in R*. Springer. p. 176. ISBN 978-1461471370. <https://doi.org/10.1007/978-1-4614-7138-7>
17. Sony, Sandeep; Dunphy, Kyle; Sadhu, Ayan; and Capretz, Miriam A M, "A Systematic Review of Convolutional Neural Network-Based Structural Condition Assessment Techniques" (2021). *Electrical and Computer Engineering Publications*. 184. <https://ir.lib.uwo.ca/electricalpub/184>
18. Prof. Manolis Kellis. *Deep Learning in the Life Sciences / Computational Systems Biology* 6.874/6.802/20.390/20.490/HST.506 Spring 2021
19. A European approach to artificial intelligence [Online]. Available at: <https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence> [Accessed 8 2 2022]
20. Regulatory framework proposal on artificial intelligence. Available at: <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai> [Accessed 8 2 2022]
21. proposal for a regulation of the European parliament and of the council laying down harmonized rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts, COM/2021/206 final. <https://eur-lex.europa.eu/legal-content/EL/TXT/PDF/?uri=CELEX:52021PC0206&from=EN>
22. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* 2014, arXiv:1409.1556
23. Ali, L.; Alnajjar, F.; Jassmi, H.A.; Gocho, M.; Khan, W.; Serhani, M.A. Performance Evaluation of Deep CNN-Based Crack Detection and Localization Techniques for Concrete Structures. *Sensors* 2021, 21, 1688. <https://doi.org/10.3390/s21051688>
24. Perez, H., Tah, J. H. M., & Mosavi, A. (2019). Deep Learning for Detecting Building Defects Using Convolutional Neural Networks. *Sensors*, 19(16), 3556. MDPI AG. Retrieved from <http://dx.doi.org/10.3390/s19163556>
25. Jianan Lin, Understanding meanshift clustering and implementation with python. *Towards Data Science* [Online]. Available at: <https://towardsdatascience.com/understanding-mean-shift-clustering-and-implementation-with-python-6d5809a2ac40> [Accessed 28 3 2022]

26. C. Zafeiropoulos, I. N. Tzortzis, I. Rallis , E. Protopapadakis , N. Doulamis , A. Doulamis, "Evaluating the Usefulness of Unsupervised monitoring in Cultural Heritage Monuments", arXiv:2107.00964 [cs.CV], <https://doi.org/10.48550/arXiv.2107.00964>
27. Girshick, R. Donahue, J. Darrell, T. Malik, J. (2013). "Rich feature hierarchies for accurate object detection and semantic segmentation", [arXiv:1311.2524](https://doi.org/10.48550/arXiv.1311.2524) [cs.CV], <https://doi.org/10.48550/arXiv.1311.2524>
28. Agrafiotis, P. "HYPERION, D6.3, Section: Deep Learning Paradigms, Crack detection"
29. Zhou, B.; Khosla, A.; Lapedriza, A.; Oliva, A.; Torralba, A. Learning deep features for discriminative localization. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition; 2016; pp. 2921–2929
<https://doi.org/10.48550/arXiv.1512.04150>
30. Dimitris Dais, İhsan Engin Bal, Eleni Smyrou, Vasilis Sarhosis, Automatic crack classification and segmentation on masonry surfaces using convolutional neural networks and transfer learning, Automation in Construction, (<https://www.sciencedirect.com/science/article/pii/S0926580521000571>)
31. Z. Qu, J. Mei, L. Liu and D. Zhou, "Crack Detection of Concrete Pavement With Cross-Entropy Loss Function and Improved VGG16 Network Model," in IEEE Access, vol. 8, pp. 54564-54573, 2020, doi: 10.1109/ACCESS.2020.2981561.
32. D. Kang, S.S. Benipal, D.L. Gopal, Y. Cha, Hybrid pixel-level concrete crack segmentation and quantification across complex backgrounds using deep learning, Automation in Construction 118 (2020), p. 103291, <https://doi.org/10.1016/j.autcon.2020.103291>
33. Guo, J., Xie, X., Bie, R. and Sun, L. (2014), "Structural health monitoring by using a sparse coding-based deep learning algorithm with wireless sensor networks", Pers Ubiquit Comput, 18, 1977 – 1987.
34. Zhao, X., Li, S., Su, H., Zhou, L. & Loh, K. J. (2018). "Image-based comprehensive maintenance and inspection method for bridges using deep learning", SMASIS2018- 8268. Proceedings of the ASME 2018 Conference on Smart Materials, Adaptive Structures and Intelligent Systems SMASIS2018 September 10-12, 2018, San Antonio, TX, USA, 1–7
35. Liang, X. (2018), 'Image-based post-disaster inspection of reinforced concrete bridge systems using deep learning with Bayesian optimization', Computer-Aided Civil and Infrastructure Engineering, 34, 415 – 430.
36. Fallahian, M., Khoshnoudian, F., & Meruane, V. (2018). Ensemble classification method for structural damage assessment under varying temperatures. Structural Health Monitoring, 17(4),747-762. <https://doi.org/10.1177/1475921717717311>

37. Narazaki, Y., Hoskere, V., Hoang, T. A. and Spencer Jr., B. F. (2017), "Automated vision-based bridge component extraction using multiscale convolutional neural networks", 3rd Huixian International Forum on Earthquake Engineering for Young Researchers, August 2017.
38. Kim, I. H., Jeon, H., Baek, S. C., Hong, W. H., & Jung, H. J. (2018). Application of crack identification techniques for an aging concrete bridge inspection using an unmanned aerial vehicle. *Sensors (Switzerland)*, 18(6), 1–14. <https://doi.org/10.3390/s18061881>
39. Maeda, H., Sekimoto, Y., Seto, T., Kashiya, T., & Omata, H. (2018b). Road Damage Detection and Classification Using Deep Neural Networks with Smartphone Images. *Computer-Aided Civil and Infrastructure Engineering*, 33(12), 1127–1141. <https://doi.org/10.1111/mice.12387>
40. Gopalakrishnan, K., Khaitan, S. K., Choudhary, A. and Agrawal, A. (2017), "Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection", *Construction and Building Materials*, 157, 322 – 330.
41. Kim, B. and Cho, S. (2018), "Automated vision-based detection of cracks on concrete surface using a deep-learning technique", *Sensors*, 18(10), 1 -18.
42. Fan, Z., Wu, Y., Lu, J. and Li, W. (2018), "Automatic pavement crack detection based on structural prediction with the convolutional neural network", arXiv preprint arXiv:1802.02208
43. Zhang, A., Wang, K. C. P., Li, B., Yang, E., Dai, X., Peng, Y., Fei, Y., Liu, Y., Li, J. Q. and Chen, C. (2017), "Automated pixel-level pavement crack detection on 3D asphalt surfaces using a deep-learning network", *Computer-Aided Civil and Infrastructure Engineering*, 32, 805 – 819.
44. Cha, Y. J., Choi, W. and Buyukozturk, O. (2017), "Deep learning-based crack damage detection using convolutional neural networks", *Computer-Aided Civil and Infrastructure Engineering*, 32, 361 – 378.
45. Fan, R., Bocus, M. J., Zhu, Y., Jiao, J., Wang, L., Ma, F., Cheng, S., and Liu, M. (2019), "Road crack detection using deep convolutional neural network and adaptive thresholding", arXiv preprint arXiv: 1904.08582
46. Tong, Z., Gao, J., Han, Z. and Wang, Z. (2017), "Recognition of asphalt pavement crack length using deep convolutional neural networks", *Road Materials and Pavement Design*, 19(6), 1334 –1349.
47. Stentoumis, C., Protopapadakis, E., Doulamis, A., & Doulamis, N. (2016). A holistic approach for inspection of civil infrastructures based on computer vision techniques. In *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences –ISPRS Archives (Vol. 41, pp. 131–138)*. <https://doi.org/10.5194/isprsarchives-XLI-B5-131-2016>.

48. Cheng, J. C. P., & Wang, M. (2018). Automated detection of sewer pipe defects in closed-circuit television images using deep learning techniques. *Automation in Construction*, 95, 155–171. <https://doi.org/10.1016/j.autcon.2018.08.006>
49. Doulamis, A., Doulamis, N., Protopapadakis, E., & Voulodimos, A. (2018). Combined convolutional neural networks and fuzzy spectral clustering for real-time crack detection in tunnels. *Proceedings - International Conference on Image Processing, ICIP*, 4153–4157. <https://doi.org/10.1109/ICIP.2018.8451758>
50. Kang, M-S., Kim, N., Lee, J. J, and An, Y-K. (2020). Deep learning-based automated underground cavity detection using three-dimensional ground penetration radar. *Structural Health Monitoring*, 2020, 19(1), 173-185. <https://doi.org/10.1177/1475921719838081>
51. Yeum, C. M., Dyke, S. J., & Ramirez, J. (2018). Visual data classification in post-event building reconnaissance. *Engineering Structures*, 155, 16–24. <https://doi.org/10.1016/j.engstruct.2017.10.057>
52. Chaiyasarn, K., Sharma, M., Ali, L., Khan, W., & Poovarodom, N. (2018). Crack detection in historical structures based on Convolutional Neural Network. *International Journal of GEOMATE*, 15(51), 240–251. <https://doi.org/10.21660/2018.51.35376>.
53. Jiang, S., and Zhang, J. (2019), “Real-time crack assessment using deep neural networks with the wall-climbing unmanned aerial system”, *Computer-Aided Civil and Infrastructure Engineering*. <https://doi.org/10.1111/mice.12519>.
54. Soukup, D. and Huber-Mork, R. (2014), “Convolutional neural networks for steel surface defect detection from photometric stereo images”, *International Symposium on Visual Computing (ISVC 2014): Advances in Visual Computing*, 668 – 677.
55. Chen, F. C. and Jahanshahi, M. R. (2018), “NB-CNN: Deep learning-based crack detection using convolutional neural network and naïve Bayes data fusion”, *IEEE Transactions on Industrial Electronics*, 65 (5), 4392 – 4400
56. Xu, D., Wen, C., & Liu, J. (2019). Wind turbine blade surface inspection based on deep learning and UAV-taken images. *Journal of Renewable and Sustainable Energy*, 11, 053305.
57. Zhang, Y., Sun, X., Loh, K. J., Su, W., Xue, Z. and Zhao, X. (2020). “Autonomous bolt loosening detection using deep learning”, *Structural Health Monitoring*, 19(1), 105-122.
58. Sasank Chilamkurthy, *Transfer Learning for Computer Vision Tutorial*. Available at: https://pytorch.org/tutorials/beginner/transfer_learning_tutorial.html [Accessed 10 9 2022]

59. Cross Entropy Loss – Pytorch 1.12 documentation. Available at:
<https://pytorch.org/docs/stable/generated/torch.nn.CrossEntropyLoss.html#torch.nn.CrossEntropyLoss>
[Accessed 11 9 2022]
60. Shreenidhi Sudhakar, Learning rate scheduler. Towards Data Science. Available at:
<https://towardsdatascience.com/learning-rate-scheduler-d8a55747dd90>
[Accessed 11 9 2022]
61. François-Guillaume Fernandez, TorchCAM: class activation explorer. Available at:
<https://github.com/frgfm/torch-cam>
[Accessed 11 9 2022]
62. Scikit-learn: Machine Learning in Python, Pedregosa *et al.*, JMLR 12, pp. 2825-2830, 2011.
<https://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html>
63. Bala Priya C. Learn K-Means Clustering by Quantizing Color Images in Python. Available at:
<https://hackernoon.com/learn-k-means-clustering-by-quantizing-color-images-in-python>
[Accessed 20 9 2022]
64. Huynh Ngoc Anh. Image segmentation. Available at:
<https://experiencor.github.io/segmentation.html>
[Accessed 20 9 2022]
65. Shiqi Yu, Sen Jia, Chunyan Xu, Convolutional neural networks for hyperspectral image classification, Neurocomputing, Volume 219, 2017, Pages 88-98, ISSN 0925-2312,
<https://doi.org/10.1016/j.neucom.2016.09.010>

ΠΑΡΑΡΤΗΜΑ

Εξωτερικές βιβλιοθήκες python

Matplotlib: Εκτύπωση εικόνων και διαγραμμάτων στην Οθόνη

Numpy: Δημιουργία και διαχείριση πινάκων τιμών

Tiff file: Ανάγνωση αρχείων .tif σαν πίνακες numpy, συγγραφή πινάκων numpy ως .tif

Pillow / PIL: Επεξεργασία εικόνων .png

Pytorch: Πακέτο μηχανικής μάθησης, όπως περιγράφεται στο Κεφ. 5

Torchcam: Εφαρμογή Class Activation Mapping σε προβλέψεις μοντέλου

Scikit-learn: Πακέτο μηχανικής μάθησης, όπως περιγράφεται στο Κεφ. 5

Κώδικας

Αφαίρεση πρόσθετων καναλιών

```
import os
# import cv2
import numpy as np
import tifffile as tiff

TIFF_PATH = "C:/Diplomatiki/Dataset/AgiosNikolaos/tiff/"

_, _, imgFiles = next(os.walk(TIFF_PATH))
num_imgs = imgFiles.__len__()

for imgIndx in range(0, num_imgs):
    filename = imgFiles[imgIndx]
    img_name = filename[:-4]
    print("Processing image: ", img_name)

    imgPath = TIFF_PATH + filename

    img = tiff.imread(imgPath)

    """RGB tifs"""
    print(img.shape)
    RGB_img = img[..., [14, 5, 2]]
    tiff.imwrite('C:/Diplomatiki/Dataset/AgiosNikolaos/RGBtif/'+
img_name+'_RGB'+'.tif', RGB_img, photometric='rgb')
    print(RGB_img.shape)
```

Κατάτμηση επισημάνσεων

```
from PIL import Image
import sys
import os
from pathlib import Path
import matplotlib.pyplot as plt

img_path = "C:/Diplomatiki/Dataset/AgiosNikolaos/annotations/"
p = Path(img_path)

#segment PNGs
for imagefile in p.glob('*.png'):
    image = Image.open(imagefile)
    i=1
    for y in range(0,image.height,224):
        for x in range(0,image.width,224):
            if(y+224<=image.height)&(x+224<=image.width):
                segment = image.crop((x,y,(x+224),(y+224)))
                segment.save(img_path+imagefile.stem+"_"+str(i)+".png")
                i+=1

#plot segmented image
fig = plt.figure(figsize=(8, 8))
columns = 8
rows = 4
```

```

fig.suptitle("agiosnik003538", fontsize=14)
for i in range(1, columns*rows +1):

    img =
Image.open("C:/Diplomatiki/Dataset/AgiosNikolaos/annotations/agiosnik003538_RGB_
"+str(i)+".png")

    fig.add_subplot(rows, columns, i)
    ax = plt.gca()

    # hide x-axis
    ax.get_xaxis().set_visible(False)

    # hide y-axis
    ax.get_yaxis().set_visible(False)

    plt.imshow(img)
plt.show()

```

Κατάτμηση tiff

```

from matplotlib import pyplot as plt
import os
#import cv2
import numpy as np
import tiff as tiff

TIFF_PATH = "C:/Diplomatiki/Dataset/AgiosNikolaos/3NIRtif/not used/"

_, _, imgFiles = next(os.walk(TIFF_PATH))
num_imgs = imgFiles.__len__()

for imgIndx in range(0, num_imgs):
    filename = imgFiles[imgIndx]
    img_name = filename[:-4]
    print("Processing image: ", img_name)

    imgPath = TIFF_PATH + filename

    image = tiff.imread(imgPath)

    #segmenting tiffs

    i = 1
    for y in range(0, image.shape[0], 224):
        for x in range(0, image.shape[1], 224):
            if (y + 224 <= image.shape[0]) & (x + 224 <= image.shape[1]):
                segment = image[y:y+224,x:x+224,:]

tiff.imwrite("C:/Diplomatiki/Dataset/AgiosNikolaos/3NIRtif/testingsegments/" +
img_name + '_' + str(i) + '.tif', segment,
            photometric='rgb')

        i += 1

    print(segment.shape)

#plot segmented image

```

```

fig = plt.figure(figsize=(8, 8))
columns = 8
rows = 4
fig.suptitle(img_name, fontsize=14)
for i in range(1, columns*rows + 1):

    img = tiff.imread(TIFF_PATH+"segmentedNIRs/"+img_name+"_"+str(i)+".tif")

    fig.add_subplot(rows, columns, i)
    ax = plt.gca()

    # hide x-axis
    ax.get_xaxis().set_visible(False)

    # hide y-axis
    ax.get_yaxis().set_visible(False)

    plt.imshow(img)
plt.show()

```

Ενίσχυση συνόλου δεδομένων χωρίς φθορές

```

import os
import numpy as np
import tiff as tiff

TIFF_PATH =
"C:/Diplomatiki/Dataset/AgiosNikolaos/3NIRtif/trainingsegments/negative/"

_, _, imgFiles = next(os.walk(TIFF_PATH))
num_imgs = len(imgFiles)

for imgIdx in range(0, num_imgs):
    filename = imgFiles[imgIdx]
    img_name = filename[:-4]
    print("Processing image: ", img_name)

    imgPath = TIFF_PATH + filename

    img = tiff.imread(imgPath)

    #flip images
    flipped_img = np.fliplr(img)

    tiff.imwrite('C:/Diplomatiki/Dataset/AgiosNikolaos/3NIRtif/trainingsegments/augmentednegative/'+ img_name+'_flipped'+'.tif', flipped_img, photometric='rgb')

    #rotate images
    rotated_img = np.rot90(img, axes=(0,1))

    tiff.imwrite('C:/Diplomatiki/Dataset/AgiosNikolaos/3NIRtif/trainingsegments/augmentednegative/'+ img_name+'_rotated'+'.tif', rotated_img, photometric='rgb')

```

Διαχωρισμός training set / validation set

```
import os
from pathlib import Path
import shutil
import random

ImgPathSource =
'C:/Diplomatiki/Dataset/AgiosNikolaos/trainfolder/RGB/train/negative'
ImgPathDestination =
'C:/Diplomatiki/Dataset/AgiosNikolaos/trainfolder/RGB/val/negative'
ImgList = os.listdir(ImgPathSource)

for i in range(35):
    x = random.randint(0, len(ImgList)-1)
    print(f"Moving {ImgList[x]} to val folder")
    shutil.move(os.path.join(ImgPathSource, ImgList[x]), ImgPathDestination)
    del ImgList[x]

ImgPathSource =
'C:/Diplomatiki/Dataset/AgiosNikolaos/trainfolder/RGB/train/positive'
ImgPathDestination =
'C:/Diplomatiki/Dataset/AgiosNikolaos/trainfolder/RGB/val/positive'

ImgList = os.listdir(ImgPathSource)

for i in range(35):
    x = random.randint(0, len(ImgList)-1)
    print(f"Moving {ImgList[x]} to val folder")
    shutil.move(os.path.join(ImgPathSource, ImgList[x]), ImgPathDestination)
    del ImgList[x]
```

Αποθήκευση ως .png των ανωτέρω αρχείων

```
from matplotlib import pyplot as plt
import os
# import cv2
import numpy as np
import tiff file as tiff

TIFF_PATH = "C:/Diplomatiki/Dataset/AgiosNikolaos/RGBtif/not used/"

_, _, imgFiles = next(os.walk(TIFF_PATH))
num_imgs = len(imgFiles)

for imgIndx in range(0, num_imgs):
    filename = imgFiles[imgIndx]
    img_name = filename[:-4]
    print("Processing image: ", img_name)

    imgPath = TIFF_PATH + filename

    img = tiff.imread(imgPath)
    formatted = (img * 255 / np.max(img)).astype('uint8')

    #save as png

    plt.imsave("C:/Diplomatiki/Dataset/AgiosNikolaos/RGBtif/" + "pngs/" +
img_name+".png", formatted, format='png')
```

Μεταφορά μάθησης και εκπαίδευση τροποποιημένου VGG16

```
from __future__ import print_function, division

import torch
import torch.nn as nn
import torch.optim as optim
from torch.optim import lr_scheduler
import torch.backends.cudnn as cudnn
import numpy as np
import torchvision
from torchvision import datasets, models, transforms
import matplotlib.pyplot as plt
import time
import os
import copy
from torch.utils import data

if __name__ == '__main__':
    cudnn.benchmark = True
    num_classes = 2

    data_transforms = {
        'train': transforms.Compose([
            transforms.ToTensor(),
            transforms.Normalize([0.485, 0.456, 0.406], [0.229, 0.224, 0.225])
        ]),
        'val': transforms.Compose([
            transforms.ToTensor(),
            transforms.Normalize([0.485, 0.456, 0.406], [0.229, 0.224, 0.225])
        ]),
    }

    data_dir = 'C:/Diplomatiki/Dataset/AgiosNikolaos/trainfolder/NIR'
    image_datasets = {x: datasets.ImageFolder(os.path.join(data_dir, x),
                                                    data_transforms[x])
                      for x in ['train', 'val']}
    dataloaders = {x: torch.utils.data.DataLoader(image_datasets[x],
                                                    batch_size=4, shuffle=True, num_workers=1)
                   for x in ['train', 'val']}
    dataset_sizes = {x: len(image_datasets[x]) for x in ['train', 'val']}
    class_names = image_datasets['train'].classes
    print(image_datasets['train'].class_to_idx)

    """use GPU"""
    device = torch.device("cuda:0" if torch.cuda.is_available() else "cpu")
    print(device)

    #define training

    def train_model(model, criterion, optimizer, scheduler, num_epochs=25):
        since = time.time()

        best_model_wts = copy.deepcopy(model.state_dict())
        best_acc = 0.0

        for epoch in range(num_epochs):
            print(f'Epoch {epoch}/{num_epochs - 1}')
            print('-' * 10)
```



```

# Each epoch has a training and validation phase
for phase in ['train', 'val']:
    if phase == 'train':
        model.train() # Set model to training mode
    else:
        model.eval() # Set model to evaluate mode

running_loss = 0.0
running_corrects = 0

# Iterate over data.
for inputs, labels in dataloaders[phase]:
    inputs = inputs.to(device)
    labels = labels.to(device)

    # zero the parameter gradients
    optimizer.zero_grad()

    # forward
    # track history if only in train
    with torch.set_grad_enabled(phase == 'train'):
        outputs = model(inputs)
        _, preds = torch.max(outputs, 1)
        loss = criterion(outputs, labels)

    # backward + optimize only if in training phase
    if phase == 'train':
        loss.backward()
        optimizer.step()

    # statistics
    running_loss += loss.item() * inputs.size(0)
    running_corrects += torch.sum(preds == labels.data)
    if phase == 'train':
        scheduler.step()

epoch_loss = running_loss / dataset_sizes[phase]
epoch_acc = running_corrects.double() / dataset_sizes[phase]

print(f'{phase} Loss: {epoch_loss:.4f} Acc: {epoch_acc:.4f}')

# deep copy the model
if phase == 'val' and epoch_acc > best_acc and epoch > 5:
    best_acc = epoch_acc
    best_model_wts = copy.deepcopy(model.state_dict())
    best_epoch = epoch

print()

time_elapsed = time.time() - since
print(f'Training complete in {time_elapsed // 60:.0f}m {time_elapsed %
60:.0f}s')
print(f'Best val Acc: {best_acc:4f}')
print(f"Epoch: {best_epoch}")

# load best model weights
model.load_state_dict(best_model_wts)
return model

#freeze model parameters
model_conv = torchvision.models.vgg16(pretrained=True)
for param in model_conv.parameters():

```

```

    param.requires_grad = False

#unfreeze last 3 convolutional blocks
for i in range (24,31):
    for param in model_conv.features[i].parameters():
        param.requires_grad = True

for param in model_conv.avgpool.parameters():
    param.requires_grad = True

# Parameters of newly constructed modules have requires_grad=True by default
model_conv.classifier[6] = nn.Linear(4096,num_classes)

model_conv = model_conv.to(device)

criterion = nn.CrossEntropyLoss()

# Observe that only parameters of final layer are being optimized as
# opposed to before.
optimizer_conv = optim.SGD(model_conv.parameters(), lr=0.001, momentum=0.9)

# Decay LR by a factor of 0.1 every 7 epochs
exp_lr_scheduler = lr_scheduler.StepLR(optimizer_conv, step_size=7,
gamma=0.1)

model_conv = train_model(model_conv, criterion, optimizer_conv,
                        exp_lr_scheduler, num_epochs=15)

#save model
torch.save(model_conv, 'NIRdeteriorationVGG16v2.pth')

```

Εφαρμογή CAM σε προβλέψεις των μοντέλων

```

import torch
import torch.nn as nn
from torchvision.io.image import read_image, ImageReadMode
from torchvision.transforms.functional import normalize, resize, to_pil_image
from torchvision.models import vgg16
from torchcam.methods import GradCAMpp
import matplotlib.pyplot as plt
from torchcam.utils import overlay_mask
import numpy as np
from PIL import Image
import os

model = torch.load('RGBdeteriorationVGG16.pth')
model.to(torch.device('cuda'))
model = model.eval()
print(model)
cam_extractor = GradCAMpp(model, target_layer= model.avgpool)

deterioration_classes = ["negative", "positive"]

FILE_PATH = "C:/Diplomatiki/Dataset/AgiosNikolaos/RGBtif/pngs/segments/"

, , imgFiles = next(os.walk(FILE_PATH))

```

```

num_imgs = imgFiles.__len__()

for imgIndx in range(0, num_imgs):
    filename = imgFiles[imgIndx]
    img_name = filename[:-4]
    print("Processing image: ", img_name)

    imgPath = FILE_PATH + filename

    # Get your input
    image = read_image(imgPath, ImageReadMode.RGB)
    input_tensor = normalize(resize(image, (224, 224)) / 255., [0.485, 0.456,
0.406], [0.229, 0.224, 0.225])
    input_tensor = input_tensor.to(torch.device('cuda'))
    # Preprocess your data and feed it to the model
    out = model(input_tensor.unsqueeze(0))
    _, preds = torch.max(out, dim=1)

    # Retrieve the CAM by passing the class index and the model output
    activation_map = cam_extractor( scores=out, class_idx=1)

    # Resize the CAM and overlay it
    result = overlay_mask(to_pil_image(image),
to_pil_image(activation_map[0].squeeze(0), mode='F'), alpha=0.8)

    # Save and Display it

    plt.imshow(result);
    plt.axis('off');
    plt.savefig("C:/Diplomatiki/Dataset/AgiosNikolaos/RGBcam/training/"+
filename, format='png', bbox_inches="tight", pad_inches=0)
    plt.suptitle(filename)
    plt.title(f"predicted class:
{deterioration_classes[out.squeeze(0).argmax().item()]}")
    plt.show()

```

Συνένωση τμημάτων CAM

```

from PIL import Image

FILEPATH = 'C:/Diplomatiki/Dataset/AgiosNikolaos/RGBcam2/train/'

Images = ['agiosnik2000RGB', 'agiosnik2048RGB',
          'agiosnik002300RGB', 'agiosnik002420RGB',
          'agiosnik002931RGB', 'agiosnik003538RGB'
          ]

for RGBimage in Images:
    CAMsegments = []
    columns = 8
    rows = 4
    for i in range(1, columns * rows + 1):
        img = Image.open(FILEPATH+RGBimage+"_" + str(i) + ".png")
        CAMsegments.append(img)

```

```

width,height = CAMsegments[0].size

NewImage = Image.new('RGB', (width*8,height*4))
i=0
for row in range(0,height*4,height):
    for column in range(0,width*8,width):
        NewImage.paste(CAMsegments[i],(column, row))
        i+=1

NewImage.save(FILEPATH+'ImageCAM/'+RGBimage+'.png')

```

Εφαρμογή αλγόριθμου ομαδοποίησης και υπολογισμός εσωτερικών δεικτών

```

import tifffile as tiff
import numpy as np
import matplotlib
import matplotlib.pyplot as plt
import os
from sklearn import metrics
import hdbscan
from sklearn.cluster import KMeans
from collections import Counter

TIFF_PATH = "C:/Diplomatiki/Dataset/AgiosNikolaos/tiff/not_used/"

_, _, imgFiles = next(os.walk(TIFF_PATH))
num_imgs = imgFiles.__len__()

for imgIndx in range(0, num_imgs):
    filename = imgFiles[imgIndx]
    img_name = filename[:-4]
    print("Processing image: ", img_name)

    imgPath = TIFF_PATH + filename

    img = tiff.imread(imgPath)
    #print(img.shape)

    (h, w, c) = img.shape
    img2D = img.reshape(h * w, c)
    #print(img2D.shape)

    kmeans_model = KMeans(n_clusters=4, random_state=1)
    cluster_labels = kmeans_model.fit_predict(img2D)

    print("Calinski-Harabasz:" + str(metrics.calinski_harabasz_score(img2D,
cluster_labels)))
    print("Davies-Bouldin:" + str(metrics.davies_bouldin_score(img2D,
cluster_labels)))
    #print("Silhouette score:" + str(metrics.silhouette_score(img2D,
cluster_labels, metric='euclidean')))

    labels_count = Counter(cluster_labels)
    print(labels_count)

```

```

label_array = np.uint8(cluster_labels.reshape(img.shape[:2]))
print(label_array.shape)
np.save(TIFF_PATH + "npy/" + img_name, label_array)

fig = plt.figure(figsize=(10, 10))

plt.imshow(label_array[224*4,224*8,:], cmap='CMRmap')
plt.axis('off')
plt.imsave(fname=TIFF_PATH + "clustered/" + img_name
+'.png',arr=label_array[224*4,224*8,:], cmap='CMRmap', format='png')
plt.suptitle(img_name)
plt.show()

```

Εκ νέου δημιουργία numpy αρχείων επισημάνσεων επειδή αυτά που δόθηκαν είχαν λάθος διαστάσεις

```

from PIL import Image
import numpy as np
from pathlib import Path
import matplotlib.pyplot as plt

#np.set_printoptions(threshold=np.inf)

img_path = "C:/Diplomatiki/Dataset/AgiosNikolaos/annotations/"
p = Path(img_path)

for imagefile in p.glob('*.png'):
    image = Image.open(imagefile)
    numpydata = np.asarray(image)
    print(numpydata.shape)

    for i in range(numpydata.shape[0]):
        for j in range(numpydata.shape[1]):
            for k in range(numpydata.shape[2]):
                if numpydata[i,j,k]>0:
                    numpydata[i,j,k]=255

    New_annotation_npy =
np.empty((numpydata.shape[0],numpydata.shape[1]), dtype=int)

    for i in range(numpydata.shape[0]):
        for j in range(numpydata.shape[1]):
            if np.array_equal(numpydata[i,j],np.array([0, 0, 0, 255])) or
np.array_equal(numpydata[i,j] , np.array([0, 0, 0, 0])):
                New_annotation_npy[i,j] = 0
            elif np.array_equal(numpydata[i,j],np.array([255, 255, 255, 255])):
                New_annotation_npy[i,j] = 1
            elif numpydata[i, j, 0] > 0:
                New_annotation_npy[i, j] = 2
            elif numpydata[i,j,1] > 0:
                New_annotation_npy[i,j] = 3
            elif numpydata[i, j, 2] > 0:

```

```
New_annotation_npy[i, j] = 4
```

```
np.save(img_path+imagefile.stem+"new.npy",New_annotation_npy)  
print(New_annotation_npy.shape)
```

Σύγκριση εικόνων που προκύπτουν από ομαδοποίηση με τις εικόνες επισημάνσεων

```
import numpy as np  
import pprint  
import sys  
  
Clustered_array =  
np.load('C:/Diplomatiki/Dataset/AgiosNikolaos/tiff/npy/agiosnik002300.npy',  
mmap_mode=None, allow_pickle=False, fix_imports=True, encoding='ASCII')  
  
Annotation_array =  
np.load('C:/Diplomatiki/Dataset/AgiosNikolaos/annotations/agiosnik002300_RGBnew.  
npy', mmap_mode=None, allow_pickle=False, fix_imports=True, encoding='ASCII')  
  
#Annotation array: True for positive. False for negative  
Annotation_array=Annotation_array.astype(bool)  
print(Annotation_array.shape)  
  
#Clustered array: True for positive, False for negative  
bool_clustered_array = np.empty_like(Clustered_array, dtype=bool)  
  
for i in range(Clustered_array.shape[0]):  
    for j in range(Clustered_array.shape[1]):  
        if Clustered_array[i,j] == 2 or Clustered_array[i,j] == 1:  
            bool_clustered_array[i,j]=True  
        else:  
            bool_clustered_array[i,j]=False  
  
#Comparison between annotation and clustering, positives/negatives  
TP=0  
TN=0  
FP=0  
FN=0  
  
for i in range(bool_clustered_array.shape[0]):  
    for j in range(bool_clustered_array.shape[1]):  
        if (bool_clustered_array[i,j] == True) and (Annotation_array[i,j] ==  
True):  
            TP+=1  
        elif (bool_clustered_array[i, j] == False) and (Annotation_array[i, j]  
== False):  
            TN+=1  
        elif (bool_clustered_array[i, j] == True) and (Annotation_array[i, j] ==  
False):  
            FP+=1  
        elif (bool_clustered_array[i, j] == False) and (Annotation_array[i, j]
```

```
== True):  
    FN+=1  
  
precision = TP/(TP+FP)  
recall = TP/(TP+FN)  
accuracy = (TP+TN)/(TP+TN+FP+FN)  
F1 = 2*precision*recall/(precision+recall)  
  
print(f'TP: {TP}, TN: {TN}, FP: {FP}, FN: {FN}')  
print(f'precision: {round(precision,2)}, recall: {round(recall,2)}, accuracy:  
{round(accuracy,2)}, F1: {round(F1,2)}')
```

Πίνακες Αποτελεσμάτων Ταξινόμησης

Αποτελέσματα μοντέλου RGB στην ταξινόμηση επισημασμένων παραθύρων των εικόνων σε μορφή πλήρους πίνακα

| Εικόνα | Παράθυρο | Πραγματική κατάσταση | Πρόβλεψη μοντέλου | Συγκριτικός έλεγχος | |
|-----------------|------------------|----------------------|-------------------|---------------------|----|
| agiosnik2000RGB | _1 | Positive | Positive | TP | |
| | _2 | Positive | Positive | TP | |
| | _3 | Positive | Positive | TP | |
| | _4 | Positive | Positive | TP | |
| | _5 | Positive | Positive | TP | |
| | _6 | Positive | Positive | TP | |
| | _7 | Positive | Positive | TP | |
| | _8 | Positive | Positive | TP | |
| | _9 | Positive | Positive | TP | |
| | _10 | Positive | Positive | TP | |
| | _11 | Positive | Positive | TP | |
| | _12 | | Negative | Positive | FP |
| | _13 | | Negative | Negative | TN |
| | _14 | | Positive | Positive | TP |
| | _15 | | Positive | Positive | TP |
| | _16 | | Positive | Positive | TP |
| | _17 | | Positive | Positive | TP |
| | _18 | | Positive | Positive | TP |
| | _19 | | Positive | Positive | TP |
| | _20 | | Positive | Positive | TP |
| | _21 | | Negative | Negative | TN |
| | _22 | | Negative | Negative | TN |
| | _23 | | Negative | Positive | FP |
| | _24 | | Negative | Negative | TN |
| | _25 | | Negative | Positive | FP |
| | _26 | | Negative | Positive | FP |
| | _27 | | Positive | Positive | TP |
| | _28 | | Positive | Negative | FN |
| | _29 | | Positive | Negative | FN |
| | _30 | | Positive | Negative | FN |
| | _31 | | Negative | Negative | TN |
| | _32 | | Positive | Negative | FN |
| | | | | | |
| | TP | TN | FP | FN | |
| | 19 | 5 | 4 | 4 | |
| | | | | | |
| | precision | recall | accuracy | F1 | |
| | 0,83 | 0,83 | 0,75 | 0,83 | |

| Εικόνα | Παράθυρο | Πραγματική κατάσταση | Πρόβλεψη μοντέλου | Συγκριτικός έλεγχος |
|-----------------|------------------|----------------------|-------------------|---------------------|
| agiosnik2048RGB | _1 | Positive | Positive | TP |
| | _2 | Positive | Positive | TP |
| | _3 | Positive | Positive | TP |
| | _4 | Negative | Positive | FP |
| | _5 | Negative | Positive | FP |
| | _6 | Positive | Positive | TP |
| | _7 | Positive | Positive | TP |
| | _8 | Positive | Positive | TP |
| | _9 | Negative | Negative | TN |
| | _10 | Negative | Negative | TN |
| | _11 | Positive | Positive | TP |
| | _12 | Positive | Positive | TP |
| | _13 | Positive | Positive | TP |
| | _14 | Positive | Positive | TP |
| | _15 | Positive | Positive | TP |
| | _16 | Positive | Positive | TP |
| | _17 | Negative | Negative | TN |
| | _18 | Negative | Negative | TN |
| | _19 | Positive | Positive | TP |
| | _20 | Positive | Positive | TP |
| | _21 | Negative | Positive | FP |
| | _22 | Negative | Negative | TN |
| | _23 | Positive | Positive | TP |
| | _24 | Positive | Positive | TP |
| | _25 | Negative | Negative | TN |
| | _26 | Positive | Positive | TP |
| | _27 | Positive | Positive | TP |
| | _28 | Positive | Positive | TP |
| | _29 | Negative | Positive | FP |
| | _30 | Positive | Negative | FN |
| | _31 | Positive | Positive | TP |
| | _32 | Positive | Positive | TP |
| | | | | |
| | TP | TN | FP | FN |
| | 21 | 6 | 4 | 1 |
| | | | | |
| | precision | recall | accuracy | F1 |
| | 0,84 | 0,95 | 0,84 | 0,89 |

| Εικόνα | Παράθυρο | Πραγματική κατάσταση | Πρόβλεψη μοντέλου | Συγκριτικός έλεγχος |
|-------------------|------------------|----------------------|-------------------|---------------------|
| agiosnik002300RGB | _1 | Negative | Positive | FP |
| | _2 | Positive | Positive | TP |
| | _3 | Positive | Positive | TP |
| | _4 | Positive | Positive | TP |
| | _5 | Positive | Positive | TP |
| | _6 | Positive | Positive | TP |
| | _7 | Positive | Positive | TP |
| | _8 | Positive | Positive | TP |
| | _9 | Positive | Negative | FN |
| | _10 | Negative | Positive | FP |
| | _11 | Positive | Positive | TP |
| | _12 | Positive | Positive | TP |
| | _13 | Positive | Positive | TP |
| | _14 | Positive | Positive | TP |
| | _15 | Positive | Positive | TP |
| | _16 | Positive | Positive | TP |
| | _17 | Positive | Positive | TP |
| | _18 | Positive | Positive | TP |
| | _19 | Positive | Positive | TP |
| | _20 | Positive | Positive | TP |
| | _21 | Positive | Positive | TP |
| | _22 | Positive | Positive | TP |
| | _23 | Positive | Positive | TP |
| | _24 | Positive | Positive | TP |
| | _25 | Positive | Positive | TP |
| | _26 | Positive | Positive | TP |
| | _27 | Positive | Positive | TP |
| | _28 | Positive | Positive | TP |
| | _29 | Positive | Positive | TP |
| | _30 | Positive | Positive | TP |
| | _31 | Negative | Positive | FP |
| | _32 | Positive | Positive | TP |
| | | | | |
| | TP | TN | FP | FN |
| | 28 | 0 | 3 | 1 |
| | | | | |
| | precision | recall | accuracy | F1 |
| | 0,90 | 0,96 | 0,875 | 0,93 |

| Εικόνα | Παράθυρο | Πραγματική κατάσταση | Πρόβλεψη μοντέλου | Συγκριτικός έλεγχος |
|-------------------|-----------|----------------------|-------------------|---------------------|
| agiosnik002420RGB | _1 | Positive | Positive | TP |
| | _2 | Positive | Positive | TP |
| | _3 | Positive | Positive | TP |
| | _4 | Positive | Positive | TP |
| | _5 | Positive | Positive | TP |
| | _6 | Positive | Positive | TP |
| | _7 | Positive | Positive | TP |
| | _8 | Positive | Positive | TP |
| | _9 | Positive | Positive | TP |
| | _10 | Positive | Positive | TP |
| | _11 | Positive | Positive | TP |
| | _12 | Positive | Positive | TP |
| | _13 | Positive | Positive | TP |
| | _14 | Positive | Positive | TP |
| | _15 | Positive | Positive | TP |
| | _16 | Positive | Positive | TP |
| | _17 | Positive | Positive | TP |
| | _18 | Positive | Positive | TP |
| | _19 | Positive | Positive | TP |
| | _20 | Positive | Positive | TP |
| | _21 | Positive | Positive | TP |
| | _22 | Positive | Positive | TP |
| | _23 | Positive | Positive | TP |
| | _24 | Positive | Positive | TP |
| | _25 | Positive | Positive | TP |
| | _26 | Positive | Positive | TP |
| | _27 | Positive | Positive | TP |
| | _28 | Positive | Positive | TP |
| | _29 | Positive | Positive | TP |
| | _30 | Positive | Positive | TP |
| | _31 | Positive | Positive | TP |
| | _32 | Positive | Positive | TP |
| | | | | |
| | TP | TN | FP | FN |
| | 32 | 0 | 0 | 0 |
| | | | | |
| | precision | recall | accuracy | F1 |
| | 1 | 1 | 1 | 1 |

| Εικόνα | Παράθυρο | Πραγματική κατάσταση | Πρόβλεψη μοντέλου | Συγκριτικός έλεγχος |
|-------------------|------------------|----------------------|-------------------|---------------------|
| agiosnik002931RGB | _1 | Negative | Negative | TN |
| | _2 | Positive | Positive | TP |
| | _3 | Positive | Positive | TP |
| | _4 | Positive | Positive | TP |
| | _5 | Positive | Positive | TP |
| | _6 | Negative | Negative | TN |
| | _7 | Negative | Negative | TN |
| | _8 | Negative | Positive | FP |
| | _9 | Negative | Negative | TN |
| | _10 | Positive | Positive | TP |
| | _11 | Positive | Positive | TP |
| | _12 | Positive | Positive | TP |
| | _13 | Positive | Negative | FN |
| | _14 | Positive | Positive | TP |
| | _15 | Positive | Positive | TP |
| | _16 | Negative | Negative | TN |
| | _17 | Negative | Negative | TN |
| | _18 | Positive | Positive | TP |
| | _19 | Negative | Positive | FP |
| | _20 | Negative | Negative | TN |
| | _21 | Positive | Positive | TP |
| | _22 | Negative | Positive | FP |
| | _23 | Negative | Negative | TN |
| | _24 | Positive | Negative | FN |
| | _25 | Negative | Negative | TN |
| | _26 | Negative | Negative | TN |
| | _27 | Negative | Negative | TN |
| | _28 | Negative | Negative | TN |
| | _29 | Positive | Positive | TP |
| | _30 | Positive | Positive | TP |
| | _31 | Positive | Negative | FN |
| | _32 | Positive | Negative | FN |
| | | | | |
| | TP | TN | FP | FN |
| | 13 | 12 | 3 | 4 |
| | | | | |
| | precision | recall | accuracy | F1 |
| | 0,81 | 0,76 | 0,78 | 0,79 |

| Εικόνα | Παράθυρο | Πραγματική κατάσταση | Πρόβλεψη μοντέλου | Συγκριτικός έλεγχος |
|-------------------|------------------|----------------------|-------------------|---------------------|
| agiosnik003538RGB | _1 | Positive | Positive | TP |
| | _2 | Positive | Positive | TP |
| | _3 | Positive | Positive | TP |
| | _4 | Positive | Positive | TP |
| | _5 | Positive | Negative | FN |
| | _6 | Positive | Positive | TP |
| | _7 | Negative | Positive | FP |
| | _8 | Negative | Negative | TN |
| | _9 | Positive | Positive | TP |
| | _10 | Positive | Positive | TP |
| | _11 | Positive | Positive | TP |
| | _12 | Positive | Positive | TP |
| | _13 | Positive | Positive | TP |
| | _14 | Positive | Positive | TP |
| | _15 | Negative | Negative | TN |
| | _16 | Positive | Positive | TP |
| | _17 | Positive | Positive | TP |
| | _18 | Positive | Negative | FN |
| | _19 | Positive | Positive | TP |
| | _20 | Positive | Positive | TP |
| | _21 | Positive | Positive | TP |
| | _22 | Positive | Positive | TP |
| | _23 | Positive | Positive | TP |
| | _24 | Negative | Negative | TN |
| | _25 | Negative | Negative | TN |
| | _26 | Negative | Negative | TN |
| | _27 | Positive | Positive | TP |
| | _28 | Negative | Negative | TN |
| | _29 | Positive | Positive | TP |
| | _30 | Negative | Negative | TN |
| | _31 | Negative | Positive | FP |
| | _32 | Positive | Positive | TP |
| | | | | |
| | TP | TN | FP | FN |
| | 21 | 7 | 2 | 2 |
| | | | | |
| | precision | recall | accuracy | F1 |
| | 0,91 | 0,91 | 0,875 | 0,91 |

| | | | | |
|------------------|------------------|---------------|-----------------|-----------|
| Συνολικά: | TP | TN | FP | FN |
| | 134 | 30 | 16 | 12 |
| | | | | |
| | precision | recall | accuracy | F1 |
| | 0,89 | 0,92 | 0,85 | 0,90 |

Πίνακας 3: Αποτελέσματα μοντέλου RGB στην ταξινόμηση επισημασμένων παραθύρων των εικόνων

Αποτελέσματα μοντέλου NIR στην ταξινόμηση επισημασμένων παραθύρων των εικόνων σε μορφή πλήρους πίνακα

| Εικόνα | Παράθυρο | Πραγματική κατάσταση | Πρόβλεψη μοντέλου | Συγκριτικός έλεγχος |
|------------------|----------|----------------------|-------------------|---------------------|
| agiosnik2000_NIR | _1 | Positive | Positive | TP |
| | _2 | Positive | Positive | TP |
| | _3 | Positive | Positive | TP |
| | _4 | Positive | Positive | TP |
| | _5 | Positive | Positive | TP |
| | _6 | Positive | Positive | TP |
| | _7 | Positive | Positive | TP |
| | _8 | Positive | Positive | TP |
| | _9 | Positive | Positive | TP |
| | _10 | Positive | Positive | TP |
| | _11 | Positive | Positive | TP |
| | _12 | Negative | Negative | TN |
| | _13 | Negative | Negative | TN |
| | _14 | Positive | Positive | TP |
| | _15 | Positive | Positive | TP |
| | _16 | Positive | Positive | TP |
| | _17 | Positive | Positive | TP |
| | _18 | Positive | Positive | TP |
| | _19 | Positive | Positive | TP |
| | _20 | Positive | Negative | FN |
| | _21 | Negative | Negative | TN |
| | _22 | Negative | Negative | TN |
| | _23 | Negative | Negative | TN |
| | _24 | Negative | Negative | TN |
| | _25 | Negative | Negative | TN |
| | _26 | Negative | Negative | TN |
| | _27 | Positive | Positive | TP |
| | _28 | Positive | Positive | TP |
| | _29 | Positive | Positive | TP |
| | _30 | Positive | Positive | TP |
| | _31 | Negative | Negative | TN |
| | _32 | Positive | Positive | TP |
| | | | | |
| | | TP | TN | FP |
| | | 22 | 9 | 0 |
| | | | | FN |
| | | | | 1 |
| | | | | |
| | | precision | recall | accuracy |
| | | 1 | 0,96 | 0,97 |
| | | | | F1 |
| | | | | 0,98 |

| Εικόνα | Παράθυρο | Πραγματική κατάσταση | Πρόβλεψη μοντέλου | Συγκριτικός έλεγχος |
|------------------|------------------|----------------------|-------------------|---------------------|
| agiosnik2048_NIR | _1 | Positive | Positive | TP |
| | _2 | Positive | Positive | TP |
| | _3 | Positive | Positive | TP |
| | _4 | Negative | Negative | TN |
| | _5 | Negative | Positive | FP |
| | _6 | Positive | Positive | TP |
| | _7 | Positive | Positive | TP |
| | _8 | Positive | Positive | TP |
| | _9 | Negative | Negative | TN |
| | _10 | Negative | Negative | TN |
| | _11 | Positive | Positive | TP |
| | _12 | Positive | Positive | TP |
| | _13 | Positive | Positive | TP |
| | _14 | Positive | Positive | TP |
| | _15 | Positive | Positive | TP |
| | _16 | Positive | Positive | TP |
| | _17 | Negative | Negative | TN |
| | _18 | Negative | Negative | TN |
| | _19 | Positive | Positive | TP |
| | _20 | Positive | Positive | TP |
| | _21 | Negative | Negative | TN |
| | _22 | Negative | Negative | TN |
| | _23 | Positive | Positive | TP |
| | _24 | Positive | Positive | TP |
| | _25 | Negative | Negative | TN |
| | _26 | Positive | Positive | TP |
| | _27 | Positive | Positive | TP |
| | _28 | Positive | Positive | TP |
| | _29 | Negative | Negative | TN |
| | _30 | Positive | Positive | TP |
| | _31 | Positive | Positive | TP |
| | _32 | Positive | Positive | TP |
| | | | | |
| | TP | TN | FP | FN |
| | 22 | 9 | 1 | 0 |
| | | | | |
| | precision | recall | accuracy | F1 |
| | 0,96 | 1 | 0,97 | 0,98 |

| Εικόνα | Παράθυρο | Πραγματική κατάσταση | Πρόβλεψη μοντέλου | Συγκριτικός έλεγχος |
|--------------------|------------------|----------------------|-------------------|---------------------|
| agiosnik002300_NIR | _1 | Negative | Negative | TN |
| | _2 | Positive | Positive | TP |
| | _3 | Positive | Positive | TP |
| | _4 | Positive | Positive | TP |
| | _5 | Positive | Positive | TP |
| | _6 | Positive | Positive | TP |
| | _7 | Positive | Positive | TP |
| | _8 | Positive | Positive | TP |
| | _9 | Positive | Negative | FN |
| | _10 | Negative | Positive | FP |
| | _11 | Positive | Positive | TP |
| | _12 | Positive | Positive | TP |
| | _13 | Positive | Positive | TP |
| | _14 | Positive | Positive | TP |
| | _15 | Positive | Positive | TP |
| | _16 | Positive | Negative | FN |
| | _17 | Positive | Positive | TP |
| | _18 | Positive | Positive | TP |
| | _19 | Positive | Positive | TP |
| | _20 | Positive | Positive | TP |
| | _21 | Positive | Positive | TP |
| | _22 | Positive | Positive | TP |
| | _23 | Positive | Negative | FN |
| | _24 | Positive | Positive | TP |
| | _25 | Positive | Positive | TP |
| | _26 | Positive | Positive | TP |
| | _27 | Positive | Positive | TP |
| | _28 | Positive | Positive | TP |
| | _29 | Positive | Positive | TP |
| | _30 | Positive | Positive | TP |
| | _31 | Negative | Negative | TN |
| | _32 | Positive | Positive | TP |
| | | | | |
| | TP | TN | FP | FN |
| | 26 | 2 | 1 | 3 |
| | | | | |
| | precision | recall | accuracy | F1 |
| | 0,96 | 0,90 | 0,875 | 0,93 |

| Εικόνα | Παράθυρο | Πραγματική κατάσταση | Πρόβλεψη μοντέλου | Συγκριτικός έλεγχος |
|--------------------|------------------|----------------------|-------------------|---------------------|
| agiosnik002420_NIR | _1 | Positive | Positive | TP |
| | _2 | Positive | Positive | TP |
| | _3 | Positive | Positive | TP |
| | _4 | Positive | Positive | TP |
| | _5 | Positive | Positive | TP |
| | _6 | Positive | Positive | TP |
| | _7 | Positive | Positive | TP |
| | _8 | Positive | Positive | TP |
| | _9 | Positive | Positive | TP |
| | _10 | Positive | Positive | TP |
| | _11 | Positive | Positive | TP |
| | _12 | Positive | Positive | TP |
| | _13 | Positive | Positive | TP |
| | _14 | Positive | Positive | TP |
| | _15 | Positive | Positive | TP |
| | _16 | Positive | Positive | TP |
| | _17 | Positive | Positive | TP |
| | _18 | Positive | Positive | TP |
| | _19 | Positive | Positive | TP |
| | _20 | Positive | Positive | TP |
| | _21 | Positive | Positive | TP |
| | _22 | Positive | Positive | TP |
| | _23 | Positive | Positive | TP |
| | _24 | Positive | Positive | TP |
| | _25 | Positive | Positive | TP |
| | _26 | Positive | Positive | TP |
| | _27 | Positive | Positive | TP |
| | _28 | Positive | Positive | TP |
| | _29 | Positive | Positive | TP |
| | _30 | Positive | Positive | TP |
| | _31 | Positive | Positive | TP |
| | _32 | Positive | Positive | TP |
| | | | | |
| | TP | TN | FP | FN |
| | 32 | 0 | 0 | 0 |
| | | | | |
| | precision | recall | accuracy | F1 |
| | 1 | 1 | 1 | 1 |

| Εικόνα | Παράθυρο | Πραγματική κατάσταση | Πρόβλεψη μοντέλου | Συγκριτικός έλεγχος |
|--------------------|------------------|----------------------|-------------------|---------------------|
| agiosnik002931_NIR | _1 | Negative | Negative | TN |
| | _2 | Positive | Positive | TP |
| | _3 | Positive | Negative | FN |
| | _4 | Positive | Positive | TP |
| | _5 | Positive | Negative | FN |
| | _6 | Negative | Negative | TN |
| | _7 | Negative | Negative | TN |
| | _8 | Negative | Negative | TN |
| | _9 | Negative | Negative | TN |
| | _10 | Positive | Negative | FN |
| | _11 | Positive | Positive | TP |
| | _12 | Positive | Positive | TP |
| | _13 | Positive | Negative | FN |
| | _14 | Positive | Positive | TP |
| | _15 | Positive | Positive | TP |
| | _16 | Negative | Positive | FP |
| | _17 | Negative | Negative | TN |
| | _18 | Positive | Positive | TP |
| | _19 | Negative | Negative | TN |
| | _20 | Negative | Negative | TN |
| | _21 | Positive | Negative | FN |
| | _22 | Negative | Negative | TN |
| | _23 | Negative | Negative | TN |
| | _24 | Positive | Negative | FN |
| | _25 | Negative | Negative | TN |
| | _26 | Negative | Negative | TN |
| | _27 | Negative | Negative | TN |
| | _28 | Negative | Negative | TN |
| | _29 | Positive | Positive | TP |
| | _30 | Positive | Positive | TP |
| | _31 | Positive | Negative | FN |
| | _32 | Positive | Positive | TP |
| | | | | |
| | TP | TN | FP | FN |
| | 10 | 14 | 1 | 7 |
| | | | | |
| | precision | recall | accuracy | F1 |
| | 0,91 | 0,59 | 0,75 | 0,71 |

| Εικόνα | Παράθυρο | Πραγματική κατάσταση | Πρόβλεψη μοντέλου | Συγκριτικός έλεγχος |
|--------------------|----------|----------------------|-------------------|---------------------|
| agiosnik003538_NIR | _1 | Positive | Positive | TP |
| | _2 | Positive | Positive | TP |
| | _3 | Positive | Positive | TP |
| | _4 | Positive | Positive | TP |
| | _5 | Positive | Positive | TP |
| | _6 | Positive | Positive | TP |
| | _7 | Negative | Negative | TN |
| | _8 | Negative | Negative | TN |
| | _9 | Positive | Positive | TP |
| | _10 | Positive | Positive | TP |
| | _11 | Positive | Positive | TP |
| | _12 | Positive | Positive | TP |
| | _13 | Positive | Positive | TP |
| | _14 | Positive | Positive | TP |
| | _15 | Negative | Negative | TN |
| | _16 | Positive | Negative | FN |
| | _17 | Positive | Positive | TP |
| | _18 | Positive | Positive | TP |
| | _19 | Positive | Positive | TP |
| | _20 | Positive | Positive | TP |
| | _21 | Positive | Positive | TP |
| | _22 | Positive | Positive | TP |
| | _23 | Positive | Positive | TP |
| | _24 | Negative | Negative | TN |
| | _25 | Negative | Negative | TN |
| | _26 | Negative | Negative | TN |
| | _27 | Positive | Positive | TP |
| | _28 | Negative | Negative | TN |
| | _29 | Positive | Positive | TP |
| | _30 | Negative | Negative | TN |
| | _31 | Negative | Negative | TN |
| | _32 | Positive | Positive | TP |
| | | | | |
| | | TP | TN | FP |
| | | 22 | 9 | 0 |
| | | | | FN |
| | | | | 1 |
| | | | | |
| | | precision | recall | accuracy |
| | | 1 | 0,96 | 0,97 |
| | | | | F1 |
| | | | | 0,98 |

| | | | | |
|-----------------|------------------|---------------|-----------------|-----------|
| Συνολικά | TP | TN | FP | FN |
| | 134 | 43 | 3 | 12 |
| | | | | |
| | precision | recall | accuracy | F1 |
| | 0,98 | 0,92 | 0,92 | 0,95 |

Πίνακας 4: Αποτελέσματα μοντέλου NIR στην ταξινόμηση επισημασμένων παραθύρων των εικόνων

Εξωτερικοί και εσωτερικοί δείκτες ομαδοποίησης εικονοστοιχείων επισημασμένων εικόνων σε μορφή πλήρους πίνακα

| Εικόνα | TP | TN | FP | FN | Εξωτερικοί δείκτες | | | |
|----------------|---------|---------|---------|---------|--------------------|--------|----------|------|
| | | | | | precision | recall | accuracy | F1 |
| agiosnik2000 | 394559 | 566698 | 457713 | 521826 | 0,46 | 0,43 | 0,5 | 0,45 |
| agiosnik2048 | 530289 | 574843 | 561050 | 244050 | 0,49 | 0,68 | 0,58 | 0,57 |
| agiosnik002300 | 569449 | 704560 | 283692 | 352531 | 0,67 | 0,62 | 0,67 | 0,64 |
| agiosnik002420 | 383856 | 863680 | 433401 | 229295 | 0,47 | 0,63 | 0,65 | 0,54 |
| agiosnik002931 | 211289 | 1044713 | 496685 | 157545 | 0,3 | 0,57 | 0,66 | 0,39 |
| agiosnik003538 | 402521 | 737988 | 479565 | 290158 | 0,46 | 0,58 | 0,6 | 0,51 |
| Σύνολο | 2491963 | 4492482 | 2712106 | 1795405 | 0,48 | 0,58 | 0,61 | 0,53 |

Πίνακας 5: Εξωτερικοί δείκτες ομαδοποίησης εικονοστοιχείων επισημασμένων εικόνων

| Εικόνα | Εσωτερικοί δείκτες | |
|----------------|--------------------|------|
| | CHI | DBI |
| agiosnik2000 | 3005286 | 0,74 |
| agiosnik2048 | 3875641 | 0,66 |
| agiosnik002300 | 2387634 | 0,81 |
| agiosnik002420 | 3207839 | 0,82 |
| agiosnik002931 | 2852866 | 0,75 |
| agiosnik003538 | 3531619 | 0,72 |

Πίνακας 6: Εσωτερικοί δείκτες ομαδοποίησης εικονοστοιχείων επισημασμένων εικόνων

Αποτελέσματα μοντέλων RGB και NIR στην ταξινόμηση παραθύρων μη επισημασμένων εικόνων σε μορφή πίνακα

| εικόνες τμήματα | agiosnik002211 | | agiosnik002948 | | agiosnik003218 | | agiosnik003340 | | agiosnik004719 | |
|--------------------|----------------|----------|----------------|----------|----------------|----------|----------------|----------|----------------|----------|
| | RGB | NIR | RGB | NIR | RGB | NIR | RGB | NIR | RGB | NIR |
| _1 | Positive | Positive | Positive | Positive | Positive | Negative | Positive | Positive | Positive | Positive |
| _2 | Positive | Positive | Positive | Positive | Negative | Negative | Positive | Negative | Positive | Positive |
| _3 | Positive | Positive | Positive | Positive | Positive | Positive | Positive | Positive | Positive | Positive |
| _4 | Positive | Positive | Positive | Positive | Negative | Negative | Positive | Negative | Positive | Positive |
| _5 | Positive | Positive | Positive | Positive | Positive | Positive | Positive | Positive | Positive | Positive |
| _6 | Positive | Positive | Positive | Positive | Positive | Negative | Positive | Positive | Positive | Positive |
| _7 | Negative | Negative | Positive | Positive | Positive | Positive | Positive | Positive | Negative | Negative |
| _8 | Negative | Positive | Positive | Positive | Positive | Positive | Negative | Positive | Negative | Negative |
| _9 | Positive | Positive | Positive | Positive | Negative | Negative | Positive | Positive | Negative | Positive |
| _10 | Positive | Positive | Positive | Positive | Negative | Negative | Positive | Positive | Positive | Positive |
| _11 | Positive | Positive | Positive | Positive | Positive | Positive | Positive | Positive | Positive | Positive |
| _12 | Positive | Positive | Positive | Positive | Positive | Positive | Positive | Positive | Positive | Positive |
| _13 | Positive | Positive | Positive | Positive | Negative | Positive | Positive | Positive | Positive | Positive |
| _14 | Positive | Positive | Positive | Positive | Negative | Negative | Positive | Positive | Positive | Positive |
| _15 | Positive | Positive | Positive | Positive | Negative | Positive | Positive | Positive | Positive | Positive |
| _16 | Positive | Positive | Positive | Positive | Positive | Negative | Positive | Positive | Positive | Positive |
| _17 | Positive | Positive | Positive | Positive | Negative | Negative | Positive | Positive | Negative | Negative |
| _18 | Positive | Positive | Positive | Positive | Positive | Positive | Positive | Positive | Positive | Positive |
| _19 | Positive | Positive | Positive | Positive | Positive | Positive | Positive | Positive | Positive | Positive |
| _20 | Positive | Positive | Positive | Positive | Positive | Negative | Positive | Positive | Positive | Positive |
| _21 | Positive | Positive | Positive | Positive | Negative | Negative | Positive | Positive | Positive | Positive |
| _22 | Positive | Positive | Positive | Positive | Negative | Negative | Positive | Positive | Positive | Positive |
| _23 | Positive | Positive | Positive | Positive | Negative | Positive | Positive | Positive | Positive | Positive |
| _24 | Positive | Positive | Positive | Positive | Positive | Negative | Positive | Positive | Positive | Positive |
| _25 | Negative | Negative | Positive | Positive | Negative | Negative | Positive | Positive | Positive | Positive |
| _26 | Positive | Negative | Positive | Positive | Positive | Positive | Positive | Positive | Positive | Positive |
| _27 | Positive | Positive | Positive | Negative | Positive | Positive | Positive | Positive | Positive | Positive |
| _28 | Positive | Positive | Positive | Positive | Negative | Positive | Positive | Positive | Positive | Positive |
| _29 | Positive | Positive | Positive | Positive | Negative | Negative | Positive | Positive | Positive | Positive |
| _30 | Positive | Positive | Positive | Positive | Negative | Negative | Positive | Positive | Positive | Positive |
| _31 | Negative | Negative | Positive | Positive | Negative | Negative | Positive | Positive | Positive | Positive |
| _32 | Positive | Positive | Positive | Positive | Positive | Negative | Positive | Positive | Positive | Positive |

| | agiosnik002211 | | agiosnik002948 | | agiosnik003218 | | agiosnik003340 | | agiosnik004719 | |
|--------------------|----------------|--------|----------------|--------|----------------|--------|----------------|--------|----------------|--------|
| | RGB | NIR | RGB | NIR | RGB | NIR | RGB | NIR | RGB | NIR |
| Positives | 28 | 28 | 32 | 31 | 16 | 14 | 31 | 30 | 28 | 29 |
| Negatives | 4 | 4 | 0 | 1 | 16 | 18 | 1 | 2 | 4 | 3 |
| % Positives | 87,50% | 87,50% | 100,00% | 96,88% | 50,00% | 43,75% | 96,88% | 93,75% | 87,50% | 90,63% |
| % Negatives | 12,50% | 12,50% | 0,00% | 3,13% | 50,00% | 56,25% | 3,13% | 6,25% | 12,50% | 9,38% |

| | Σύνολο | |
|--------------------|--------|--------|
| | RGB | NIR |
| Positives | 135 | 132 |
| Negatives | 25 | 28 |
| % Positives | 84,38% | 82,50% |
| % Negatives | 15,63% | 17,50% |

Πίνακας 7: Αποτελέσματα μοντέλων ταξινόμησης σε άγνωστα δεδομένα

Εσωτερικοί δείκτες ομαδοποίησης εικονοστοιχείων μη επισημασμένων εικόνων σε μορφή πίνακα

| Εικόνα | Εσωτερικοί δείκτες | |
|-----------------------|--------------------|------|
| | CHI | DBI |
| agiosnik002211 | 3280680 | 0,63 |
| agiosnik002948 | 3391835 | 0,64 |
| agiosnik003218 | 4319148 | 0,62 |
| agiosnik003340 | 2233848 | 0,86 |
| agiosnik004719 | 3673392 | 0,74 |

Πίνακας 8: Εσωτερικοί δείκτες ομαδοποίησης εικονοστοιχείων άγνωστων εικόνων