



**ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ
ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ**

**ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
«ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΚΑΙ ΕΦΑΡΜΟΓΕΣ»**

**«MASTER OF SCIENCE IN ARTIFICIAL INTELLIGENCE
AND APPLICATIONS»**

**«Ανίχνευση και ταξινόμηση ρωγμών σε δομικά στοιχεία
με χρήση Συνελικτικών Νευρωνικών Δικτύων (CNNs)»**

Ζάχου Ευαγγελία του Θεοδώρου

Σεπτέμβριος 2026



ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ «ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΚΑΙ ΕΦΑΡΜΟΓΕΣ»

«Master of Science in Artificial Intelligence and Applications»

«Ανίχνευση και ταξινόμηση ρωγμών σε δομικά στοιχεία με χρήση Συνελικτικών Νευρωνικών Δικτύων (CNNs)»

Διπλωματική Εργασία που υποβλήθηκε στο Τμήμα Πληροφορικής και
Τηλεπικοινωνιών του Πανεπιστημίου Θεσσαλίας ως μέρος των απαιτήσεων για την
απόκτηση Διπλώματος Μεταπτυχιακών Σπουδών Τεχνητή Νοημοσύνη και Εφαρμογές
από τον/την

Ζάχου Ευαγγελία του Θεοδώρου

Δήλωση Αυθεντικότητας, ζητήματα Copyright

«Ο/Η μεταπτυχιακός/ή φοιτητής/τρια που εκπόνησε την παρούσα διπλωματική εργασία φέρει ολόκληρη την ευθύνη προσδιορισμού της δίκαιης χρήσης του υλικού, η οποία ορίζεται στη βάση των εξής παραγόντων: του σκοπού και χαρακτήρα της χρήσης (μη-εμπορικός, μη-κερδοσκοπικός, αλλά εκπαιδευτικός-ερευνητικός), της φύσης του υλικού που χρησιμοποιεί (τμήμα του κειμένου, πίνακες, σχήματα, εικόνες κ.λπ.), του ποσοστού και της σημαντικότητας του τμήματος που χρησιμοποιεί σε σχέση με το όλο κείμενο υπό copyright, και των πιθανών συνεπειών της χρήσης αυτής στην αγορά ή την γενικότερη αξία του υπό copyright κειμένου».

Ο/Η Μεταπτυχιακός/ή Φοιτητής/τρια

Σεπτέμβριος 2026

Τριμελής Εξεταστική Επιτροπή

Η παρούσα Μ.Δ.Ε. εγκρίθηκε ομόφωνα από την τριμελή εξεταστική επιτροπή η οποία ορίστηκε από την Σ.Ε. / Συνέλευση του Τμήματος Πληροφορικής και Τηλεπικοινωνιών του Πανεπιστημίου Θεσσαλίας, σύμφωνα με το νόμο και τον εγκεκριμένο Οδηγό Σπουδών του Π.Μ.Σ. «Τεχνητή Νοημοσύνη και Εφαρμογές». Τα μέλη της Επιτροπής ήταν:

1 ^ο Μέλος (Επιβλέπων):	Κολομβάτσος Κωνσταντίνος
2 ^ο Μέλος:	
3 ^ο Μέλος:	

Η έγκριση της Μ.Δ.Ε. από το Τμήμα Πληροφορικής και Τηλεπικοινωνιών του Πανεπιστημίου Θεσσαλίας, δεν υποδηλώνει αποδοχή των απόψεων του συγγραφέα.

Περίληψη

Η διατήρηση της δομικής ακεραιότητας των κτιριακών υποδομών αποτελεί κρίσιμη παράμετρο για την ασφάλεια των κατασκευών, με την παρουσία ρωγμών σε δομικά στοιχεία να αποτελεί συχνά την πρώτη ένδειξη δομικής παθογένειας ή και αστοχίας των υλικών. Η παραδοσιακή οπτική επιθεώρηση από εξειδικευμένο προσωπικό είναι πολύ χρονοβόρα, δαπανηρή και επιρρεπής σε υποκειμενικότητα. Σκοπός της παρούσας διπλωματικής εργασίας είναι η αυτοματοποίηση του εντοπισμού και του χαρακτηρισμού δομικών βλαβών μέσω της ανάπτυξης και εκπαίδευσης μοντέλων Βαθιάς Μάθησης, αξιοποιώντας Συνελικτικά Νευρωνικά Δίκτυα (CNNs) πάνω σε αρχιτεκτονική YOLO11. Ειδικότερα, αναπτύχθηκε ένα MultiTaskYOLO, ένα πολυκαθηκοντικό μοντέλο ταξινόμησης βασισμένο σε backbone YOLO11m με μηχανισμό προσοχής C2PSA, το οποίο εκπαιδεύτηκε στο δημόσιο σύνολο δεδομένων φ-Net (PEER Hub ImageNet) για την ταυτόχρονη πρόβλεψη τεσσάρων χαρακτηριστικών: της κατάστασης βλάβης, της ύπαρξης αποφλοιώσης, του τύπου δομικού στοιχείου και του τύπου βλάβης. Παράλληλα, αναπτύχθηκε συμπληρωματικό pipeline εντοπισμού και μέτρησης ρωγμών μέσω fine-tuning ενός μοντέλου YOLO11x-seg σε εξειδικευμένο σύνολο δεδομένων κατάτμησης. Η μεθοδολογία περιέλαβε προεπεξεργασία εικόνων, σχεδίαση αρχιτεκτονικής, εκπαίδευση δύο φάσεων (πάγωμα και πλήρες ξεπάγωμα του backbone) και συστηματικές μελέτες αφαίρεσης (ablation studies) για την αξιολόγηση σχεδιαστικών επιλογών, όπως το μέγεθος του backbone, ο μηχανισμός δειγματοληψίας ανά task και η στάθμιση της συνάρτησης απώλειας. Η αξιολόγηση πραγματοποιήθηκε με μετρικές accuracy, F1-score και AUC, καθώς και μέσω πινάκων σύγχυσης και τεχνικών οπτικοποίησης. Το τελικό μοντέλο πέτυχε μέση ακρίβεια περίπου 76,85% στα τέσσερα καθήκοντα, με τη διάκριση του τύπου βλάβης να παραμένει το πιο απαιτητικό task, εύρημα συνεπές με αντίστοιχες μελέτες της διεθνούς βιβλιογραφίας.

Abstract

Maintaining the structural integrity of building infrastructure is a critical parameter for construction safety, with the presence of cracks in structural elements often being the first indication of structural pathology or material fatigue. Traditional visual inspection by specialized personnel remains time-consuming, costly, and prone to subjectivity. The aim of this thesis is to automate the detection and characterization of structural damage through the development and training of Deep Learning models, leveraging Convolutional Neural Networks (CNNs) built on the YOLO11 architecture.

Specifically, MultiTaskYOLO was developed, a multi-task classification model based on a YOLO11m backbone with a C2PSA attention mechanism, trained on the public ϕ -Net (PEER Hub ImageNet) dataset to simultaneously predict four characteristics: damage state, presence of spalling, structural component type, and damage type. In parallel, a complementary crack localization and measurement pipeline was developed through fine-tuning a YOLO11x-seg model on a specialized segmentation dataset.

The methodology included image preprocessing, architecture design, two-phase training (frozen and fully unfrozen backbone), and systematic ablation studies evaluating design choices such as backbone size, per-task sampling strategy, and loss weighting. Evaluation was performed using accuracy, F1-score, and AUC, as well as confusion matrices and visualization techniques. The final model achieved a mean accuracy of approximately 76.85% across the four tasks, with damage-type discrimination remaining the most demanding task, a finding consistent with corresponding studies in the international literature.

Πίνακας Περιεχομένων

Κεφάλαιο 1. Εισαγωγή	11
1.1 Γενικά	11
1.2 Κίνητρο της Εργασίας	11
1.3 Στόχοι της Εργασίας	12
1.4 Συνεισφορά της Εργασίας	12
1.5 Δομή της Εργασίας	13
Κεφάλαιο 2. Θεωρητικό Υπόβαθρο	14
2.1. Τεχνητή Νοημοσύνη, Μηχανική Μάθηση και Βαθιά Μάθηση	14
2.2. Όραση Υπολογιστών και Ψηφιακή Αναπαράσταση Εικόνας	14
2.3. Νευρωνικά Δίκτυα και Συνελκτικά Νευρωνικά Δίκτυα	14
2.4. Μεταφορά Μάθησης (Transfer Learning)	15
2.5. Πολυκαθηκοντική Μάθηση (Multi-Task Learning)	16
2.6. Ανίχνευση Αντικειμένων και η Οικογένεια Μοντέλων YOLO	16
2.7. Semantic segmentation και Μοντέλα Θεμελίωσης (Foundation Models)	17
2.8. Ανίχνευση Δομικής Βλάβης με Βαθιά Μάθηση: Ανασκόπηση Βιβλιογραφίας .	17
Κεφάλαιο 3. Δεδομένα και Μεθοδολογία	20
3.1 Ανίχνευση Δομικής Βλάβης με Βαθιά Μάθηση: Ανασκόπηση Βιβλιογραφίας .	20
3.2 Σύνολο Δεδομένων: φ-Net	20
3.3 Μηχανική Ερμηνεία των Χαρακτηριστικών (Tasks) του φ-Net	21
3.4 Προεπεξεργασία Δεδομένων	24
3.5 Διαχωρισμός Δεδομένων	24
3.6 Προτεινόμενη Αρχιτεκτονική: MultiTaskYOLO	24
3.7 Αρχιτεκτονική Segmentation Track: YOLO11x-seg	27
3.8 Διαδικασία Εκπαίδευσης	29
3.9 Μετρικές Αξιολόγησης	31
3.10 Ενοποιημένο Pipeline Ταξινόμησης-Κατάτμησης και Επίπεδο Σύντηξης Σοβαρότητας (Severity Fusion)	31
Κεφάλαιο 4. Πειραματικά Αποτελέσματα	34
4.1 Αποτελέσματα Ταξινόμησης (Classification Track)	34
4.2 Επίδραση Κανόνα HEA-style Gating	40
4.3 Αποτελέσματα Segmentation Track	41
4.4 Αξιολόγηση του Segmentation ως Ταξινομητή (Image-Level)	42
4.5 Μελέτες Αφαίρεσης (Ablation Studies)	43

4.6	Ενοποιημένο Pipeline: Ποιοτική Αξιολόγηση με Severity Fusion	44
Κεφάλαιο 5. Συζήτηση		54
5.1	Σύγκριση με τη Σχετική Βιβλιογραφία	54
5.2	Η Δυσκολία του Task damage_type	54
5.3	Το Εύρημα του HEA-style Gating	55
5.4	Περιορισμοί του Segmentation Track	55
5.5	Περιορισμοί της Μελέτης	56
Κεφάλαιο 6. Συμπεράσματα		57
6.1	Σύνοψη της Εργασίας	57
6.2	Κύρια Ευρήματα	57
6.3	Συνεισφορά της Εργασίας	58
6.4	Μελλοντικές Κατευθύνσεις	58

Κατάλογος Εικόνων

Εικόνα 1. Χαρακτηριστικά πρότυπα ρωγμών σε δοκό: καμπτική (κάθετες ρωγμές στο μέσο ανοίγματος), διατμητική (λοξές ρωγμές κοντά στις στηρίξεις) και συνδυασμένη αστοχία.....	24
Εικόνα 2. Χαρακτηριστικά πρότυπα ρωγμών σε υποστύλωμα: οριζόντιες ρωγμές βάσης (καμπτική), διαγώνιο «X» (διατμητική) και συνδυασμένη αστοχία.	25
Εικόνα 3. Χαρακτηριστικά πρότυπα ρωγμών σε τοίχωμα: οριζόντιες ρωγμές βάσης / rocking (καμπτική), πλήρες διαγώνιο «X» (διατμητική) και συνδυασμένη αστοχία.....	26
Εικόνα 4. Backbone ταξινόμησης YOLO11m-cls (MultiTaskYOLO): layers 0–10 (BACKBONE_CUTOFF=10) και κοινή έξοδος 512 καναλιών @ 7×7, πάνω από την οποία τοποθετούνται οι τέσσερις κεφαλές ταξινόμησης.....	28
Εικόνα 5. Πλήρης αρχιτεκτονική YOLO11x-seg (segmentation track): backbone (layers 0–10), neck τύπου FPN με Concat/Upsample (layers 11–22) και segment head (layer 23), χωρίς καμία δομική τροποποίηση.....	31
Εικόνα 6. Πίνακας σύγκρισης - damage_state.....	37
Εικόνα 7. Καμπύλη ROC - damage_state (AUC=0,933).....	37
Εικόνα 8. Πίνακας σύγκρισης - spalling.....	38
Εικόνα 9. Καμπύλη ROC - spalling (AUC=0,881).....	38
Εικόνα 10. Πίνακας σύγκρισης - component.....	39
Εικόνα 11. Καμπύλες ROC (one-vs-rest) - component, ανά κλάση και macro-average (AUC=0,858).	40
Εικόνα 12. Πίνακας σύγκρισης - damage_type.....	41
Εικόνα 13. Καμπύλες ROC (one-vs-rest) - damage_type, ανά κλάση και macro-average (AUC=0,797).	41
Εικόνα 14. Πίνακας σύγκρισης - damage_type μετά το HEA-style gating.....	43
Εικόνα 15. Πίνακας σύγκρισης - αξιολόγηση του segmentation track ως ταξινομητή σε επίπεδο εικόνας (σύνολο ελέγχου, 3 κλάσεις).....	45
Εικόνα 16. Spalling override: Beam, Combined, κάλυψη 26,7%.....	47
Εικόνα 17. Spalling override: Wall, Shear, κάλυψη 19,9%.....	48
Εικόνα 18. Spalling override: Wall, Combined, κάλυψη 12,0%.....	48
Εικόνα 19. Spalling override: Beam, Shear, κάλυψη 3,9%.....	49
Εικόνα 20. Extensive→Red: Wall, Shear, κάλυψη 94,3%.....	49

Εικόνα 21. Extensive→Red: Wall, Shear, κάλυψη 84,7% (damage_state confidence 0,68).....	50
Εικόνα 22. Βασικός πίνακας: Column, Combined, κάλυψη 29,4%.....	50
Εικόνα 23. Βασικός πίνακας: Wall, Shear, κάλυψη 2,2%.....	51
Εικόνα 24. Βασικός πίνακας: Column, Shear, κάλυψη 2,0%.....	51
Εικόνα 25. Οριακή περίπτωση: μηδενικές ανιχνεύσεις segmentation παρά damage_type=Combined.....	52
Εικόνα 26. Forced Green: Beam, Undamaged.....	53
Εικόνα 27. Forced Green: Column, Undamaged (spalling πρόβλεψη αγνοείται).....	53
Εικόνα 28. Forced Green: Wall, Undamaged.....	54
Εικόνα 29. Forced Green: Wall, Undamaged.....	54

Κατάλογος Πινάκων

Πίνακας 1. Πίνακας σύντηξης σοβαρότητας (severity fusion table): Component × Damage Type × Coverage Zone.....	32
Πίνακας 2. Συγκεντρωτικά αποτελέσματα ταξινόμησης στο σύνολο ελέγχου, ανά task και macro average.....	33
Πίνακας 3. Σύγκριση μετρικών damage_type πριν και μετά το HEA-style gating.....	36
Πίνακας 4. Μετρικές segmentation track (YOLO11x-seg) — συνολικές μετρικές στο σύνολο ελέγχου (test set) και ανάλυση ανά κλάση στο σύνολο επικύρωσης, καλύτερο checkpoint.....	38
Πίνακας 5. Συνοπτικά αποτελέσματα μελετών αφαίρεσης (qualitative summary)....	40
Πίνακας 6. Two-Model Pipeline + Severity Fusion - συγκεντρωτικός πίνακας δεκατεσσάρων ενδεικτικών παραδειγμάτων.....	41

Κεφάλαιο 1. Εισαγωγή

1.1 Γενικά

Η οπτική επιθεώρηση των κτιρίων αποτελεί μέχρι σήμερα μια κύρια μέθοδο αξιολόγησης της κατάστασης φέρουσας κατασκευής μετά από κάποιο σεισμικό γεγονός, κάποια φυσική καταστροφή γενικότερα ή στο πλαίσιο τακτικής συντήρησης, καθώς επιτρέπει τον άμεσο εντοπισμό εμφανών βλαβών χωρίς την ανάγκη εξειδικευμένου ειδικού εξοπλισμού [11]. Η διαδικασία αυτή παραμένει ωστόσο πολύ χρονοβόρα, καθώς απαιτεί έμπειρο ανθρώπινο δυναμικό και είναι επιρρεπής σε υποκειμενικότητα, ειδικά όταν οι ενδείξεις βλάβης είναι ασθενείς, μερικώς καλυμμένες ή δυσδιάκριτες από την εικόνα [16]. Σε συνθήκες έπειτα από σεισμό, όπου πλήθος κατασκευών χρειάζεται άμεση αξιολόγηση σε σύντομο χρονικό διάστημα, ο απαιτούμενος όγκος επιθεωρήσεων καθιστά τις παραδοσιακές μεθόδους ανεπαρκείς για την έγκαιρη λήψη αποφάσεων [8].

Τα τελευταία χρόνια, οι τεχνικές βαθιάς μάθησης και ειδικότερα τα συνελικτικά νευρωνικά δίκτυα (Convolutional Neural Networks, CNNs) έχουν αναδειχθεί σε ένα αποτελεσματικό εργαλείο για την αυτοματοποιημένη ανάλυση εικόνων δομικής βλάβης [1]. Η δημιουργία πολλών συνόλων δεδομένων εικόνων, με χαρακτηριστικότερο παράδειγμα το φ-Net (PEER Hub ImageNet), έδωσε τη δυνατότητα να εκπαιδεύονται μοντέλα ικανά να αναγνωρίζουν ταυτόχρονα διαφορετικές διαστάσεις της βλάβης, όπως την ύπαρξη ή μη βλάβης, την αποφλοιώση σκυροδέματος, το είδος του δομικού στοιχείου και τον τύπο της βλάβης - απευθείας από φωτογραφικό υλικό, χωρίς να απαιτείται χειροκίνητη εξαγωγή χαρακτηριστικών [2].

1.2 Κίνητρο της Εργασίας

Παρά τη σημαντική πρόοδο στο πεδίο αυτό, οι περισσότερες υπάρχουσες προσεγγίσεις επικεντρώνονται σε μεμονωμένα καθήκοντα αναγνώρισης, όπως για παράδειγμα αποκλειστικά στην ανίχνευση ρωγμών ή αποκλειστικά στην ταξινόμηση αποφλοιώσης, χωρίς να λαμβάνουν υπόψη τη συνολική εικόνα της κατάστασης μιας κατασκευής [8], [16]. Όπως επισημαίνεται και σε πρόσφατη έρευνα που αξιοποιεί το ίδιο σύνολο δεδομένων φ-Net [13], λίγες μελέτες εξετάζουν ταυτόχρονα πολλαπλές διαστάσεις της βλάβης, την κατάσταση βλάβης, τον τύπο δομικού στοιχείου, την ύπαρξη αποφλοιώσης, τον τύπο βλάβης μέσα από ένα ενιαίο μοντέλο. Οι περισσότερες εργασίες προσεγγίζουν

κάθε task ξεχωριστά, πολλαπλασιάζοντας το υπολογιστικό κόστος εκπαίδευσης και συντήρησης και δυσχεραίνοντας την πρακτική εφαρμογή σε πραγματικές συνθήκες πεδίου, όπου απαιτείται ταχεία και μια πιο συνολική αξιολόγηση.

Ένα δεύτερο πρόβλημα που εντοπίζεται στη βιβλιογραφία αφορά τη φύση των διαθέσιμων συνόλων δεδομένων. Σύνολα όπως το φ-Net παρέχουν επισημειώσεις σε επίπεδο εικόνας (image-level labels) και όχι σε επίπεδο pixel ή bounding box [2], γεγονός που περιορίζει τη χρήση τους σε σενάρια ανίχνευσης αντικειμένων (object detection) ή σημασιολογικής κατάτμησης (segmentation), όπως αυτά που εφαρμόζονται σε άλλα σύνολα δεδομένων μετασεισμικής καταγραφής βλαβών [8]. Παράλληλα, η εξέλιξη των μοντέλων ανίχνευσης αντικειμένων της οικογένειας YOLO, και ειδικότερα η κυκλοφορία της έκδοσης YOLO11 με ενσωματωμένο μηχανισμό προσοχής C2PSA στο neck της αρχιτεκτονικής, έχει κάνει εφικτή την ταυτόχρονη εξαγωγή καλύτερων οπτικών αναπαραστάσεων με σχετικά μικρό υπολογιστικό κόστος [7]. Αυτό είναι ένα στοιχείο που την καθιστά υποψήφια όχι μόνο για κλασική ανίχνευση αντικειμένων αλλά και ως backbone για σενάρια πολυκαθηκοντικής ταξινόμησης (multi-task classification).

1.3 Στόχοι της Εργασίας

Στόχος της παρούσας εργασίας είναι η ανάπτυξη ενός συστήματος αυτόματης ανίχνευσης και ταξινόμησης ρωγμών σε δομικά στοιχεία με χρήση συνελκτικών νευρωνικών δικτύων, το οποίο να αξιοποιεί αποκλειστικά το σύνολο δεδομένων φ-Net και να επεκτείνει τη λογική της multi-task ταξινόμησης που έχει προταθεί σε προηγούμενες εργασίες σε μία ενιαία αρχιτεκτονική βασισμένη σε YOLO11 backbone [2], [13]. Ειδικότερα, επιδιώκεται: (α) η ταυτόχρονη πρόβλεψη τεσσάρων ανεξάρτητων αλλά συσχετισμένων ιδιοτήτων της εικόνας (κατάσταση βλάβης, αποφλοίωση, τύπος δομικού στοιχείου, τύπος βλάβης) μέσω κοινού backbone και ξεχωριστών κεφαλών ταξινόμησης, (β) η ανάπτυξη συμπληρωματικού pipeline εντοπισμού και μέτρησης ρωγμών (segmentation track) μέσω fine-tuning με το YOLO11x-seg, ώστε το τελικό σύστημα να καλύπτει τόσο την ταξινόμηση όσο και τη χωρική οριοθέτηση της βλάβης.

1.4 Συνεισφορά της Εργασίας

Η κύρια συνεισφορά της εργασίας αφορά την ενοποίηση τεσσάρων καθηκόντων ταξινόμησης σε ένα ενιαίο multi-task μοντέλο με κοινό YOLO11m backbone και C2PSA neck, αντί της μεμονωμένης εκπαίδευσης ξεχωριστών μοντέλων ανά task, μια πρακτική

που ακολουθείται στο μεγαλύτερο μέρος της σχετικής βιβλιογραφίας [8], [9], [16]. Τέλος, η εργασία συνεισφέρει με τη συστηματική σύγκριση εναλλακτικών σχεδιαστικών επιλογών (μέγεθος backbone, μηχανισμός δειγματοληψίας ανά task, στάθμιση συνάρτησης απώλειας, Spatial Attention Pooling έναντι global average pooling) μέσω δομημένων ablation studies.

1.5 Δομή της Εργασίας

Η εργασία οργανώνεται ως εξής: το Κεφάλαιο 2 παρουσιάζει το θεωρητικό υπόβαθρο της τεχνητής νοημοσύνης, της βαθιάς μάθησης και της υπολογιστικής όρασης, με έμφαση στην εξέλιξη των συνελκτικών δικτύων και της οικογένειας μοντέλων YOLO, καθώς και στην ανασκόπηση της σχετικής βιβλιογραφίας για την ανίχνευση δομικής βλάβης. Το Κεφάλαιο 3 περιγράφει το σύνολο δεδομένων φ-Net, τη διαδικασία προεπεξεργασίας, τον διαχωρισμό σε υποσύνολα train/validation/test και την προτεινόμενη αρχιτεκτονική MultiTaskYOLO. Το Κεφάλαιο 4, παρουσιάζει τα πειραματικά αποτελέσματα και τις μελέτες αφαίρεσης (ablation studies). Το Κεφάλαιο 5, αναλύει και συζητά τα ευρήματα σε σχέση με τη βιβλιογραφία. Τέλος, το Κεφάλαιο 6 συνοψίζει τα συμπεράσματα και προτείνει κατευθύνσεις μελλοντικής έρευνας.

Κεφάλαιο 2. Θεωρητικό Υπόβαθρο

2.1.Τεχνητή Νοημοσύνη, Μηχανική Μάθηση και Βαθιά Μάθηση

Η τεχνητή νοημοσύνη αναφέρεται στο ευρύτερο πεδίο ανάπτυξης υπολογιστικών συστημάτων ικανών να εκτελούν καθήκοντα που παραδοσιακά απαιτούν την ανθρώπινη νοημοσύνη. Η μηχανική μάθηση αποτελεί υποσύνολό της, στο οποίο τα συστήματα βελτιώνουν την απόδοσή τους μέσα από την έκθεση σε δεδομένα αντί για ρητά προγραμματισμένους κανόνες. Η βαθιά μάθηση, με τη σειρά της, αξιοποιεί τα πολυεπίπεδα νευρωνικά δίκτυα για την αυτόματη εξαγωγή ιεραρχικών αναπαραστάσεων από ακατέργαστα δεδομένα, χωρίς χειροκίνητη μηχανική εξαγωγή χαρακτηριστικών. Η καθιέρωση της βαθιάς μάθησης στην υπολογιστική όραση συνδέεται στενά με την επιτυχία αρχιτεκτονικών όπως το AlexNet, το οποίο απέδειξε ότι τα βαθιά συνελκτικά δίκτυα μπορούν να ξεπεράσουν σημαντικά τις παραδοσιακές μεθόδους σε προβλήματα ταξινόμησης εικόνων μεγάλης κλίμακας [3].

2.2.Όραση Υπολογιστών και Ψηφιακή Αναπαράσταση Εικόνας

Η υπολογιστική όραση ασχολείται με την εξαγωγή νοήματος από οπτικά δεδομένα. Μια ψηφιακή εικόνα αναπαρίσταται ως πλέγμα εικονοστοιχείων (pixels), καθένα από τα οποία φέρει τιμές έντασης σε ένα ή περισσότερα κανάλια χρώματος (π.χ. RGB). Πριν από την εδραίωση της βαθιάς μάθησης, οι κλασικές τεχνικές όρασης υπολογιστών βασίζονταν σε χειροκίνητους αλγορίθμους ανίχνευσης ακμών και εξαγωγής χαρακτηριστικών, οι οποίοι όμως παρουσίαζαν περιορισμένη ανθεκτικότητα σε μεταβολές φωτισμού, υψής επιφάνειας και θορύβου φόντου, ιδίως σε εικόνες πραγματικών συνθηκών πεδίου [16]. Η μετάβαση σε μεθόδους βαθιάς μάθησης επέτρεψε την αυτόματη εκμάθηση χαρακτηριστικών απευθείας από τα δεδομένα εκπαίδευσης, βελτιώνοντας σημαντικά τη γενίκευση σε ετερογενή, μη ελεγχόμενα περιβάλλοντα λήψης.

2.3.Νευρωνικά Δίκτυα και Συνελκτικά Νευρωνικά Δίκτυα

Ένα τεχνητό νευρωνικό δίκτυο αποτελείται από επίπεδα διασυνδεδεμένων κόμβων (νευρώνων), καθένας από τους οποίους υπολογίζει έναν σταθμισμένο συνδυασμό των εισόδων του και εφαρμόζει μια μη γραμμική συνάρτηση ενεργοποίησης (π.χ. ReLU,

Sigmoid). Η εκπαίδευση πραγματοποιείται μέσω οπισθοδιάδοσης του σφάλματος (backpropagation) και αλγορίθμων βελτιστοποίησης τύπου gradient descent, οι οποίοι ενημερώνουν τα βάρη του δικτύου ώστε να ελαχιστοποιείται μια συνάρτηση απώλειας. Τα συνελκτικά νευρωνικά δίκτυα διαφοροποιούνται από τα πλήρως συνδεδεμένα δίκτυα μέσω της χρήσης συνελκτικών επιπέδων, τα οποία εφαρμόζουν κοινά, εκπαιδευσιμα φίλτρα σε τοπικές περιοχές της εικόνας, αξιοποιώντας την τοπική χωρική συσχέτιση των εικονοστοιχείων και μειώνοντας δραστικά τον αριθμό των εκπαιδευσιμων παραμέτρων σε σχέση με πλήρως συνδεδεμένα δίκτυα ίδιου μεγέθους εισόδου.

Η εξέλιξη των αρχιτεκτονικών CNN για ταξινόμηση εικόνων περιλαμβάνει “σταθμούς” όπως το AlexNet, το οποίο εισήγαγε τη χρήση βαθιών συνελκτικών δικτύων σε προβλήματα μεγάλης κλίμακας, το VGG, το οποίο κατέδειξε ότι η αύξηση του βάθους του δικτύου με μικρά φίλτρα 3×3 βελτιώνει σημαντικά την ακρίβεια αναγνώρισης, και το ResNet, το οποίο εισήγαγε τις υπολειμματικές συνδέσεις (residual connections) για την αντιμετώπιση του προβλήματος της εξαφάνισης της κλίσης (vanishing gradient) σε πολύ βαθιά δίκτυα, επιτρέποντας την εκπαίδευση δικτύων δεκάδων ή και εκατοντάδων επιπέδων [3], [4], [5]. Οι αρχιτεκτονικές αυτές αποτέλεσαν τη βάση πάνω στην οποία αναπτύχθηκαν μετέπειτα εξειδικευμένα μοντέλα ανίχνευσης αντικειμένων και κατάτμησης εικόνας.

2.4.Μεταφορά Μάθησης (Transfer Learning)

Η μεταφορά μάθησης αξιοποιεί βάρη προεκπαιδευμένα σε ένα μεγάλο, γενικό σύνολο δεδομένων (π.χ. ImageNet) ως σημείο εκκίνησης για την εκπαίδευση σε ένα μικρότερο, εξειδικευμένο σύνολο δεδομένων, μειώνοντας τον χρόνο εκπαίδευσης και τον κίνδυνο υπερπροσαρμογής όταν τα διαθέσιμα δεδομένα είναι περιορισμένα. Στο πεδίο της αναγνώρισης δομικής βλάβης, η μεταφορά μάθησης από γενικά σύνολα δεδομένων εικόνας σε εξειδικευμένα σύνολα δομικών βλαβών αποδείχθηκε ιδιαίτερα αποτελεσματική, ήδη από τις πρώτες συστηματικές μελέτες του πεδίου, οι οποίες κατέδειξαν ότι προεκπαιδευμένα δίκτυα μπορούν να προσαρμοστούν επιτυχώς στην αναγνώριση τύπου βλάβης, κατάστασης δομικού στοιχείου και επιπέδου σοβαρότητας βλάβης με σχετικά περιορισμένο όγκο δεδομένων εξειδικευμένης επισήμανσης [1].

2.5. Πολυκαθηκοντική Μάθηση (Multi-Task Learning)

Το Multi-Task Learning αναφέρεται στην ταυτόχρονη εκπαίδευση ενός μοντέλου για την εκτέλεση περισσότερων του ενός σχετιζόμενων καθηκόντων, μοιράζοντας ένα κοινό τμήμα εξαγωγής χαρακτηριστικών (backbone) και διαφοροποιώντας μόνο τις τελικές κεφαλές πρόβλεψης ανά καθήκον. Η προσέγγιση αυτή έχει εφαρμοστεί στο πεδίο της δομικής βλάβης μέσω αρχιτεκτονικών semantic segmentation ικανών να αναγνωρίζουν ταυτόχρονα πολλαπλούς τύπους υλικού και βλάβης σε εικόνες υποδομών, καθώς και μέσω πλαισίων που συνδυάζουν πολλαπλά διαδοχικά μοντέλα ταξινόμησης (υλικό, κατάσταση βλάβης, τύπος δομικού στοιχείου, τύπος βλάβης) πάνω στο ίδιο σύνολο δεδομένων φ-Net [17], [13]. Σε σχέση με την εκπαίδευση ξεχωριστών μοντέλων ανά καθήκον, η multi-task μάθηση μειώνει το συνολικό υπολογιστικό κόστος και μπορεί να βελτιώσει τη γενίκευση όταν τα καθήκοντα μοιράζονται υποκείμενα οπτικά χαρακτηριστικά, ενώ ταυτόχρονα εισάγει την πρόκληση της ισορροπημένης στάθμισης της συνάρτησης απώλειας ανάμεσα σε καθήκοντα με διαφορετικό βαθμό δυσκολίας.

2.6. Ανίχνευση Αντικειμένων και η Οικογένεια Μοντέλων YOLO

Σε αντίθεση με την ταξινόμηση εικόνας, η ανίχνευση αντικειμένων στοχεύει στον ταυτόχρονο εντοπισμό της θέσης (bounding box) και της κατηγορίας ενός ή περισσότερων αντικειμένων εντός μιας εικόνας. Η οικογένεια μοντέλων YOLO (You Only Look Once) αποτελεί ένα από τα πλέον διαδεδομένα single-stage πλαίσια ανίχνευσης, προσφέροντας ταχύτητα αποτελέσματος συγκρίσιμη ή καλύτερη από πιο αργές two-stage μεθόδους όπως τα Faster R-CNN, καθιστώντας τα κατάλληλα για εφαρμογές πεδίου με περιορισμένους υπολογιστικούς πόρους [8]. Η συστηματική συγκριτική μελέτη εκδόσεων YOLOv3 έως YOLOv10 σε σύνολο δεδομένων δομικής βλάβης με σύνθετο φόντο απέδειξε ότι δεν υπάρχει μία καθολικά βέλτιστη έκδοση για όλα τα σενάρια, καθώς η απόδοση εξαρτάται από τον αριθμό των περιστατικών βλάβης ανά εικόνα και τη σύνθεση του συνόλου δεδομένων εκπαίδευσης [16].

Η έκδοση YOLO11 διατηρεί τη βασική δομή backbone–neck–head των προηγούμενων εκδόσεων, εισάγοντας όμως το block C3k2 στο backbone και έναν μηχανισμό αυτοπροσοχής μέσω του module C2PSA (Cross-Stage Partial with Self-Attention) στο neck [7]. Ο μηχανισμός αυτός επιτρέπει στο δίκτυο να σταθμίζει δυναμικά τη σημασία διαφορετικών χωρικών περιοχών του feature map, ενισχύοντας την εστίαση σε πληροφοριακά σημαντικές περιοχές και μειώνοντας τον θόρυβο του φόντου [7]. Το head

της αρχιτεκτονικής έχει σχεδιαστεί ως αρθρωτό πλαίσιο που υποστηρίζει, πέρα από την κλασική ανίχνευση αντικειμένων, και εργασίες σημασιολογικής/instance κατάτμησης, ταξινόμησης εικόνας και pose estimation, μέσω εναλλακτικών κεφαλών πρόβλεψης πάνω στο ίδιο backbone [7]. Η αρθρωτή αυτή σχεδίαση αποτελεί βασικό λόγο για τον οποίο το backbone του YOLO11 μπορεί να αξιοποιηθεί ως κοινή βάση εξαγωγής χαρακτηριστικών για multi-task σενάρια ταξινόμησης, πέρα από την αρχική του χρήση για ανίχνευση αντικειμένων.

2.7.Semantic segmentation και Μοντέλα Θεμελίωσης (Foundation Models)

Η Semantic segmentation αποδίδει σε κάθε εικονοστοιχείο μιας εικόνας μια ετικέτα κατηγορίας, επιτρέποντας ακριβέστερο εντοπισμό των ορίων μιας βλάβης σε σχέση με τη πιο χονδρική οριοθέτηση που παρέχει ένα bounding box, κάτι ιδιαίτερα χρήσιμο για τη μέτρηση μήκους, πλάτους και εμβαδού ρωγμών. Πρόσφατες εργασίες έχουν διερευνήσει την αξιοποίηση μοντέλων ανίχνευσης της οικογένειας YOLO11 σε συνδυασμό με το Segment Anything Model (SAM), ένα μοντέλο θεμελίωσης προεκπαιδευμένο σε τεράστιο όγκο μασκών κατάτμησης, ικανό να παράγει μάσκες υψηλής ακρίβειας από απλά prompts (σημεία ή bounding boxes) χωρίς εξειδικευμένη εκπαίδευση ανά εφαρμογή [14], [15]. Η προσέγγιση αυτή έχει ιδιαίτερο ενδιαφέρον για σύνολα δεδομένων όπως το φ-Net, στα οποία δεν υπάρχουν annotations σε επίπεδο pixel ή bounding box παρά μόνο ετικέτες σε επίπεδο εικόνας[14]. Αυτό επιτρέπει τη μερική υπέρβαση αυτού του περιορισμού μέσω semi-supervised σχημάτων που απαιτούν μόνο ετικέτες ανίχνευσης [14].

2.8. Ανίχνευση Δομικής Βλάβης με Βαθιά Μάθηση: Ανασκόπηση Βιβλιογραφίας

Η βιβλιογραφία γύρω από την αυτοματοποιημένη ανίχνευση δομικής βλάβης έχει αναπτυχθεί κατά κύριο λόγο γύρω από δύο άξονες: την κατασκευή κατάλληλων συνόλων δεδομένων και τη σχεδίαση αρχιτεκτονικών ικανών να αξιοποιήσουν αποτελεσματικά τα δεδομένα αυτά. Ως προς τον πρώτο άξονα, το φ-Net παραμένει ένα από τα ευρύτερα διαθέσιμα σύνολα δεδομένων, περιλαμβάνοντας πάνω από 30.000 εικόνες δομικής βλάβης από πραγματικά σεισμικά γεγονότα, με annotation ως προς επτά διαφορετικά χαρακτηριστικά (υλικό κατασκευής, ύπαρξη βλάβης, αποφλοίωση, επίπεδο κατάρρευσης, τύπος δομικού στοιχείου, επίπεδο βλάβης, τύπος βλάβης) [13], [2]. Το γεγονός ότι τα annotation αυτά δίνονται σε επίπεδο εικόνας και όχι σε επίπεδο bounding

box έχει οδηγήσει άλλες ερευνητικές ομάδες στη δημιουργία εναλλακτικών, μικρότερων συνόλων δεδομένων με πλήρη annotation bounding box, ειδικά προσαρμοσμένων σε σενάρια ανίχνευσης αντικειμένων μετά από σεισμό [8].

Ως προς τον δεύτερο άξονα, μια σημαντική κατεύθυνση αφορά τη χρήση synthetic ή semi-synthetic δεδομένων για την αντιμετώπιση της έλλειψης μεγάλου όγκου επισημειωμένων εικόνων μετασεισμικής καταγραφής. Χαρακτηριστικό παράδειγμα αποτελεί η μέθοδος tunable semi-synthetic image generation, στην οποία τεχνητές βλάβες τεσσάρων κατηγοριών (ρωγμή, αποφλοίωση, διάβρωση/εκτεθειμένος οπλισμός, διαρροή υγρασίας) εφαρμόζονται πάνω σε πραγματικά τρισδιάστατα μοντέλα κτιρίων και γεφυρών μέσω παραμετρικών meta-annotations, με στόχο την επαύξηση ενός μοντέλου ανίχνευσης YOLO εκπαιδευμένου στο σύνολο δεδομένων IDEA [8]. Η μελέτη αυτή κατέδειξε συστηματική βελτίωση της ακρίβειας ανίχνευσης όταν τα πραγματικά δεδομένα συνδυάζονται με τα ελεγχόμενα παραγόμενα δείγματα, ιδίως στην κλάση της διάβρωσης/εκτεθειμένου οπλισμού, όπου τα πραγματικά δεδομένα ήταν πιο περιορισμένα σε πλήθος.

Μια δεύτερη κατεύθυνση επικεντρώνεται στη σύγκριση αρχιτεκτονικών ανίχνευσης και κατάτμησης ειδικά για ρωγμές σκυροδέματος. Μελέτες που εκπαιδεύουν ξεχωριστά μοντέλα YOLO11-seg για ρωγμές και αποφλοίωση, συνδυάζοντάς τα εκ των υστέρων σε ενιαία πολυκλασική μάσκα (unified multi-class mask), έχουν δείξει ότι η εξειδίκευση ανά τύπο βλάβης αποδίδει καλύτερα από ένα ενιαίο πολυκλασικό μοντέλο για annotations, ιδίως όταν οι δύο τύποι βλάβης εμφανίζονται οπτικά παρόμοιοι [9]. Παράλληλα, εργασίες που τροποποιούν την εσωτερική αρχιτεκτονική του YOLO11, για παράδειγμα μέσω δυναμικής συνέλιξης τύπου KernelWarehouse, μηχανισμού τριπλής προσοχής (Triple Attention) στο feature pyramid network και εξειδικευμένης συνάρτησης απώλειας FP-IoU, έχουν καταδείξει μετρήσιμη βελτίωση στην ανίχνευση ρωγμών μικρής κλίμακας μέσω συστηματικών ablation studies, στα οποία η συνεισφορά κάθε τροποποίησης αξιολογείται ξεχωριστά και σε συνδυασμό με τις υπόλοιπες [10].

Ιδιαίτερο ενδιαφέρον για την παρούσα εργασία παρουσιάζουν μελέτες που αξιοποιούν άμεσα το σύνολο δεδομένων φ-Net για την εκπαίδευση υβριδικών πλαισίων ταξινόμησης. Μία τέτοια προσέγγιση συνδυάζει διαδοχικά μοντέλα YOLOv11 (πρώτα για ταξινόμηση εσωτερικού/εξωτερικού χώρου, στη συνέχεια για αναγνώριση δομικού στοιχείου και τέλος για ανίχνευση τύπου βλάβης) με ένα σύστημα κανόνων που αποδίδει τελικό επίπεδο βλάβης, επιτυγχάνοντας ακρίβεια έως 73,7% (εντός ενός επιπέδου) σε ανεξάρτητα σύνολα ελέγχου από διαφορετικούς σεισμούς [11], [12]. Μία άλλη

προσέγγιση συνδυάζει το φ-Net με ένα προεκπαιδευμένο μεγάλο γλωσσικό μοντέλο όρασης (vision-language model) για την ταυτόχρονη εκτέλεση επτά καθηκόντων αξιολόγησης βλάβης, αναδεικνύοντας ότι η ακρίβεια διαφέρει σημαντικά ανά task -από 96,4% στην αναγνώριση υλικού έως μόλις 34,6% στην αναγνώριση τύπου βλάβης-γεγονός που αποδίδεται στη μεγαλύτερη οπτική ασάφεια και υποκειμενικότητα που χαρακτηρίζει τη διάκριση μεταξύ καμπτικής, διατμητικής και σύνθετης βλάβης, ακόμη και μεταξύ ανθρώπινων ειδικών [13].

Από τη συνολική επισκόπηση της βιβλιογραφίας προκύπτουν δύο κύρια κενά τα οποία η παρούσα εργασία επιχειρεί να καλύψει εν μέρει: πρώτον, ο περιορισμένος αριθμός εργασιών που εκπαιδεύουν ένα ενιαίο μοντέλο πολυκατηγορικής ταξινόμησης πάνω σε backbone της τελευταίας γενιάς YOLO (YOLO11) με ενσωματωμένο μηχανισμό προσοχής C2PSA, αντί για διαδοχικά ή ανεξάρτητα μοντέλα ανά task [11], [13]. Δεύτερον, η περιορισμένη διερεύνηση της σύνδεσης ανάμεσα σε μοντέλα ταξινόμησης εικόνας - επιπέδου, όπως αυτά που εκπαιδεύονται στο φ-Net, και μοντέλα χωρικής κατάτμησης/μέτρησης βλάβης, όπως αυτά που εκπαιδεύονται σε εξειδικευμένα σύνολα δεδομένων segmentation [9], [14], [15].

Κεφάλαιο 3. Δεδομένα και Μεθοδολογία

3.1 Ανίχνευση Δομικής Βλάβης με Βαθιά Μάθηση: Ανασκόπηση Βιβλιογραφίας

Η προτεινόμενη μεθοδολογία οργανώνεται σε δύο παράλληλα σκέλη. Το πρώτο σκέλος αφορά την ανάπτυξη ενός πολυκαθηκοντικού μοντέλου ταξινόμησης (MultiTaskYOLO), το οποίο δέχεται ως είσοδο μια εικόνα και παράγει ταυτόχρονα προβλέψεις για τέσσερα ανεξάρτητα χαρακτηριστικά: την κατάσταση βλάβης (`damage_state`), την ύπαρξη αποφλοίωσης (`spalling`), τον τύπο δομικού στοιχείου (`component`) και τον τύπο βλάβης (`damage_type`). Το δεύτερο σκέλος, αφορά την ανάπτυξη ενός συμπληρωματικού pipeline εντοπισμού και μέτρησης ρωγμών μέσω fine-tuning ενός μοντέλου YOLO11x-seg, το οποίο παρέχει χωρική πληροφορία (εμβαδόν επικάλυψης σε ποσοστό της εικόνας) που δεν είναι διαθέσιμη από το μοντέλο ταξινόμησης. Η επιλογή δύο παράλληλων σκελών αντί ενός ενιαίου μοντέλου δικαιολογείται από τη διαφορετική φύση των διαθέσιμων δεδομένων για κάθε σκέλος: το φ-Net παρέχει μόνο ετικέτες σε επίπεδο εικόνας, ενώ η χωρική κατάτμηση απαιτεί `annotations pixel-επιπέδου` που δεν είναι διαθέσιμα στο φ-Net και αντλούνται από ξεχωριστό, εξειδικευμένο σύνολο δεδομένων ρωγμών[2].

3.2 Σύνολο Δεδομένων: φ-Net

Το φ-Net (PEER Hub ImageNet) αποτελεί ένα ευρέως αναγνωρισμένο σύνολο δεδομένων εικόνων δομικής βλάβης, το οποίο δημιουργήθηκε από τους Gao και Mosalam με σκοπό την υποστήριξη έρευνας μηχανικής/βαθιάς μάθησης στο πεδίο της μετασεισμικής αξιολόγησης κατασκευών [2]. Περιλαμβάνει πάνω από 30.000 εικόνες, με `annotations` ως προς επτά διακριτά χαρακτηριστικά (υλικό κατασκευής, ύπαρξη βλάβης, αποφλοίωση, επίπεδο κατάρρευσης, τύπο δομικού στοιχείου, επίπεδο βλάβης και τύπο βλάβης [13]) με τις ετικέτες να δίνονται σε επίπεδο εικόνας και όχι μέσω `bounding box` ή `pixel-mask annotations`. Στην παρούσα εργασία αξιοποιούνται τέσσερα από τα επτά διαθέσιμα `tasks` (κατάσταση βλάβης, αποφλοίωση, τύπος δομικού στοιχείου, τύπος βλάβης).

3.3 Μηχανική Ερμηνεία των Χαρακτηριστικών (Tasks) του φ-Net

Πέρα από την τεχνική διάσταση της πολυκαθηκοντικής ταξινόμησης, τα τέσσερα tasks του φ-Net που αξιοποιούνται στην παρούσα εργασία φέρουν σαφή μηχανική σημασία, η οποία καθορίζει και τον τρόπο ερμηνείας των αποτελεσμάτων του μοντέλου από την πλευρά του πολιτικού μηχανικού.

Το task `damage_state` αντιστοιχεί στη βασική διάκριση ύπαρξης ή απουσίας ορατής βλάβης σε μια εικόνα και λειτουργεί ως πρώτο επίπεδο διαλογής (triage) σε ένα σενάριο μετασεισμικής επιθεώρησης: μια κατασκευή ή δομικό στοιχείο που ταξινομείται ως μη ρηγματωμένο μπορεί να εξαιρεθεί από περαιτέρω λεπτομερή έλεγχο, επιταχύνοντας τη διαδικασία προτεραιοποίησης [2].

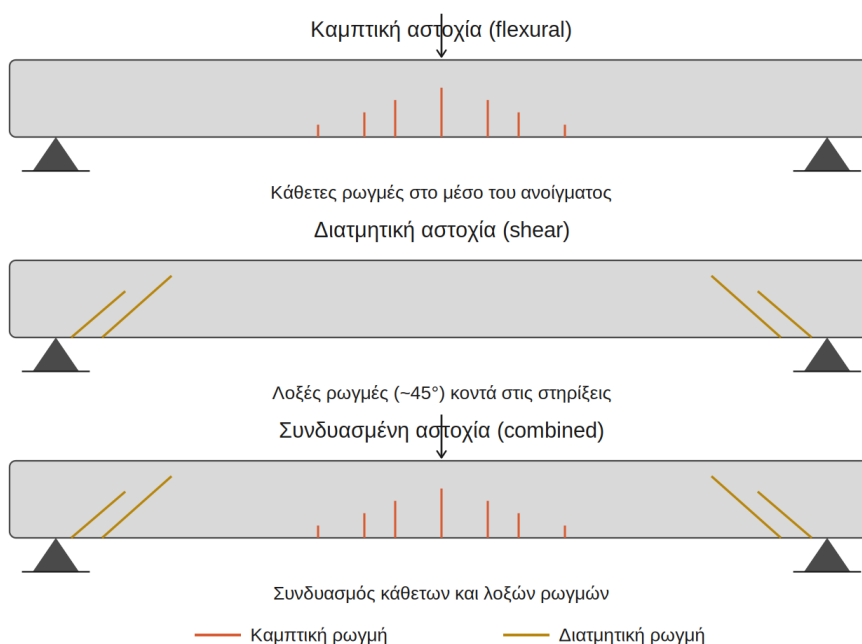
Το task `spalling` αφορά την αποφλοίωση του σκυροδέματος, δηλαδή την τοπική απώλεια της επικάλυψης σκυροδέματος γύρω από τον οπλισμό. Η αποφλοίωση αποτελεί σοβαρότερη μορφή βλάβης σε σχέση με μια απλή ρωγμή, καθώς συχνά συνοδεύεται από έκθεση του χαλύβδινου οπλισμού σε περιβαλλοντικές συνθήκες (υγρασία, οξείδωση) και υποδηλώνει προχωρημένο στάδιο υποβάθμισης του δομικού στοιχείου [2].

Το task `component` προσδιορίζει τον τύπο του δομικού στοιχείου στο οποίο εντοπίζεται η βλάβη (δοκός, υποστύλωμα, τοίχείο ή άλλο στοιχείο) σύμφωνα με την κατηγοριοποίηση που ακολουθεί και το φ-Net [13]. Η πληροφορία αυτή είναι κρίσιμη για την εκτίμηση της δομικής σημασίας μιας βλάβης, καθώς η ίδια οπτική ρωγμή μπορεί να έχει τελείως διαφορετική βαρύτητα ανάλογα με το στοιχείο στο οποίο εμφανίζεται: μια βλάβη σε υποστύλωμα θεωρείται γενικά πιο κρίσιμη από αντίστοιχη βλάβη σε τοιχοποιία πλήρωσης, καθώς τα υποστυλώματα φέρουν άμεσα τα κατακόρυφα φορτία της κατασκευής και η αστοχία τους μπορεί να οδηγήσει σε τοπική ή ολική κατάρρευση. Τέλος, το task `damage_type` διακρίνει τον μηχανισμό αστοχίας της ρωγμής σε καμπτικό (flexural), διατμητικό (shear) ή συνδυασμένο (combined) τύπο. Οι καμπτικές ρωγμές εμφανίζονται συνήθως κάθετα στον άξονα του δομικού στοιχείου, στις περιοχές μέγιστης καμπτικής ροπής, και σχετίζονται με πιο όλκιμη (ductile) συμπεριφορά αστοχίας. Αντίθετα, οι διατμητικές ρωγμές εμφανίζονται συνήθως υπό γωνία περίπου 45 μοιρών ως προς τον άξονα του στοιχείου και σχετίζονται με ψαθυρή (brittle), αιφνίδια αστοχία, η οποία θεωρείται μηχανικά πιο επικίνδυνη καθώς παρέχει μικρότερη προειδοποίηση πριν την τελική αστοχία [18]. Η διάκριση αυτή έχει αποδειχθεί ιδιαίτερα απαιτητική τόσο για ανθρώπινους επιθεωρητές όσο και για αυτοματοποιημένα μοντέλα, καθώς οι δύο κατηγορίες ρωγμών μπορεί να εμφανίζουν παρόμοια οπτικά χαρακτηριστικά, ιδίως σε φωτογραφίες χαμηλής ανάλυσης ή υπό μη ιδανικές συνθήκες φωτισμού· το εύρημα αυτό

επιβεβαιώνεται και από τη σημαντικά χαμηλότερη ακρίβεια που επιτυγχάνουν μοντέλα βαθιάς μάθησης στο συγκεκριμένο task σε σύγκριση με τα υπόλοιπα [13].

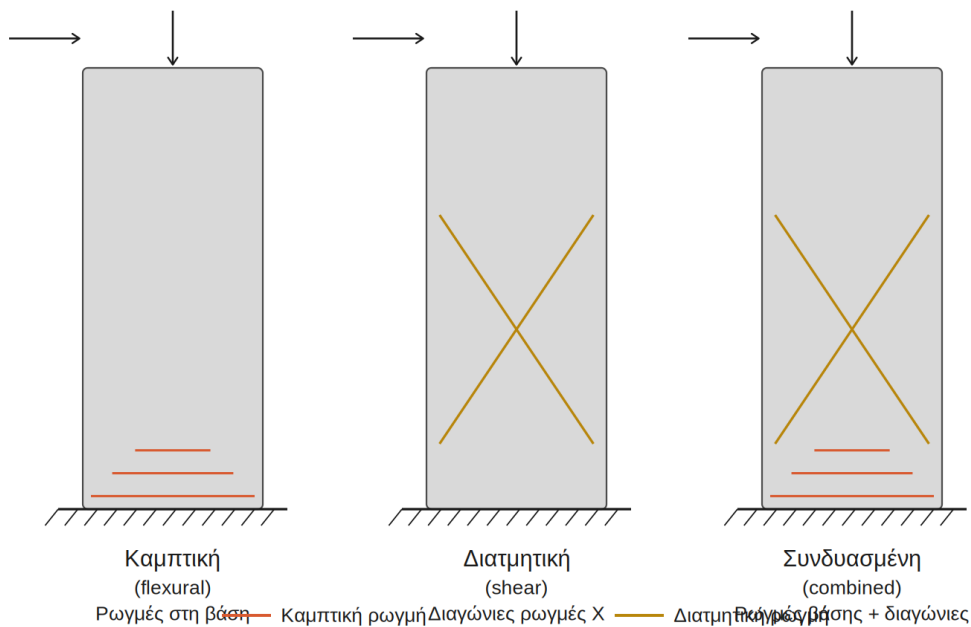
Οι παραπάνω μορφές αστοχίας ορίζονται θεωρητικά μέσω της μηχανικής των τάσεων και παραμορφώσεων σε κάμψη και διάτμηση [19], και αναγνωρίζονται στην πράξη μέσω χαρακτηριστικών οπτικών προτύπων ρωγμών, τα οποία διαφέρουν ανάλογα με τον τύπο του δομικού στοιχείου [20].

Σε δοκούς, η καμπτική αστοχία εκδηλώνεται με κάθετες ρωγμές συγκεντρωμένες στο μέσο του ανοίγματος, εκεί όπου αναπτύσσεται η μέγιστη καμπτική ροπή. Η διατμητική αστοχία, αντίθετα, εμφανίζεται ως λοξές ρωγμές υπό γωνία περίπου 45 μοιρών κοντά στις στηρίξεις, όπου η διατμητική δύναμη είναι μεγαλύτερη.



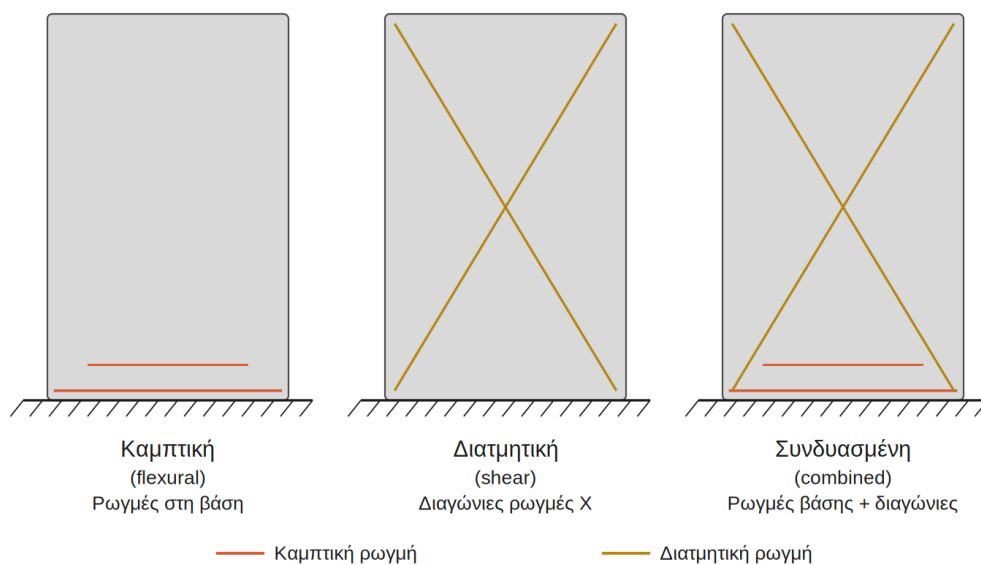
Εικόνα 1. Χαρακτηριστικά πρότυπα ρωγμών σε δοκό: καμπτική (κάθετες ρωγμές στο μέσο ανοίγματος), διατμητική (λοξές ρωγμές κοντά στις στηρίξεις) και συνδυασμένη αστοχία.

Σε υποστυλώματα, η καμπτική αστοχία εκδηλώνεται με οριζόντιες ρωγμές συγκεντρωμένες στη βάση, όπου εντοπίζεται η μέγιστη ροπή λόγω πλευρικής (σεισμικής) φόρτισης. Η διατμητική αστοχία εμφανίζεται ως διαγώνιες ρωγμές σε σχήμα «X», χαρακτηριστικές ιδίως σε κοντά υποστυλώματα [20].



Εικόνα 2. Χαρακτηριστικά πρότυπα ρωγμών σε υποστύλωμα: οριζόντιες ρωγμές βάσης (καμπτική), διαγώνιο «X» (διατμητική) και συνδυασμένη αστοχία.

Σε τοιχώματα, η καμπτική αστοχία (rocking) εκδηλώνεται επίσης με οριζόντιες ρωγμές στη βάση, ενώ η διατμητική αστοχία εμφανίζεται ως εκτεταμένο διαγώνιο «X» σε ολόκληρη την επιφάνεια του τοιχώματος, χαρακτηριστική μορφή ρηγμάτωσης τοιχείων υπό σεισμική διέγερση [20]. Και στις τρεις περιπτώσεις δομικών στοιχείων, η συνδυασμένη αστοχία περιλαμβάνει ταυτόχρονη εμφάνιση και των δύο προτύπων ρωγμών, γεγονός που καθιστά τη διάκρισή της ιδιαίτερα απαιτητική τόσο για ανθρώπινους επιθεωρητές όσο και για αυτοματοποιημένα μοντέλα βαθιάς μάθησης [13].



Εικόνα 3. Χαρακτηριστικά πρότυπα ρωγμών σε τοίχωμα: οριζόντιες ρωγμές βάσης / rocking (καμπτική), πλήρες διαγώνιο «X» (διατμητική) και συνδυασμένη αστοχία.

3.4 Προεπεξεργασία Δεδομένων

Οι εικόνες του φ-Net παρέχονται σε προεπεξεργασμένη μορφή float32 BGR, με αφαίρεση μέσης τιμής τύπου Caffe/VGG ανά κανάλι χρώματος (B=103.939, G=116.779, R=123.68), μια σύμβαση προεπεξεργασίας που ακολουθεί τη λογική που εισήγαγε αρχικά η αρχιτεκτονική VGG για την κανονικοποίηση εισόδου [4]. Στο πλαίσιο της παρούσας εργασίας, η μορφή αυτή αναλύθηκε και αντιστράφηκε (reverse-engineered), ώστε οι εικόνες να εξαχθούν εκ νέου σε τυπική μορφή JPEG, συμβατή με τα εργαλεία φόρτωσης δεδομένων της βιβλιοθήκης Ultralytics που χρησιμοποιούνται για την εκπαίδευση του μοντέλου YOLO11 [6]. Η αναδόμηση της αρχικής δομής φακέλων ακολουθεί το πρότυπο ImageFolder (extracted_images/task{N}_{name}/{train,val,test}/{ClassName}/*.jpg), με αυτόματη ανίχνευση κλάσεων ανά task.

3.5 Διαχωρισμός Δεδομένων

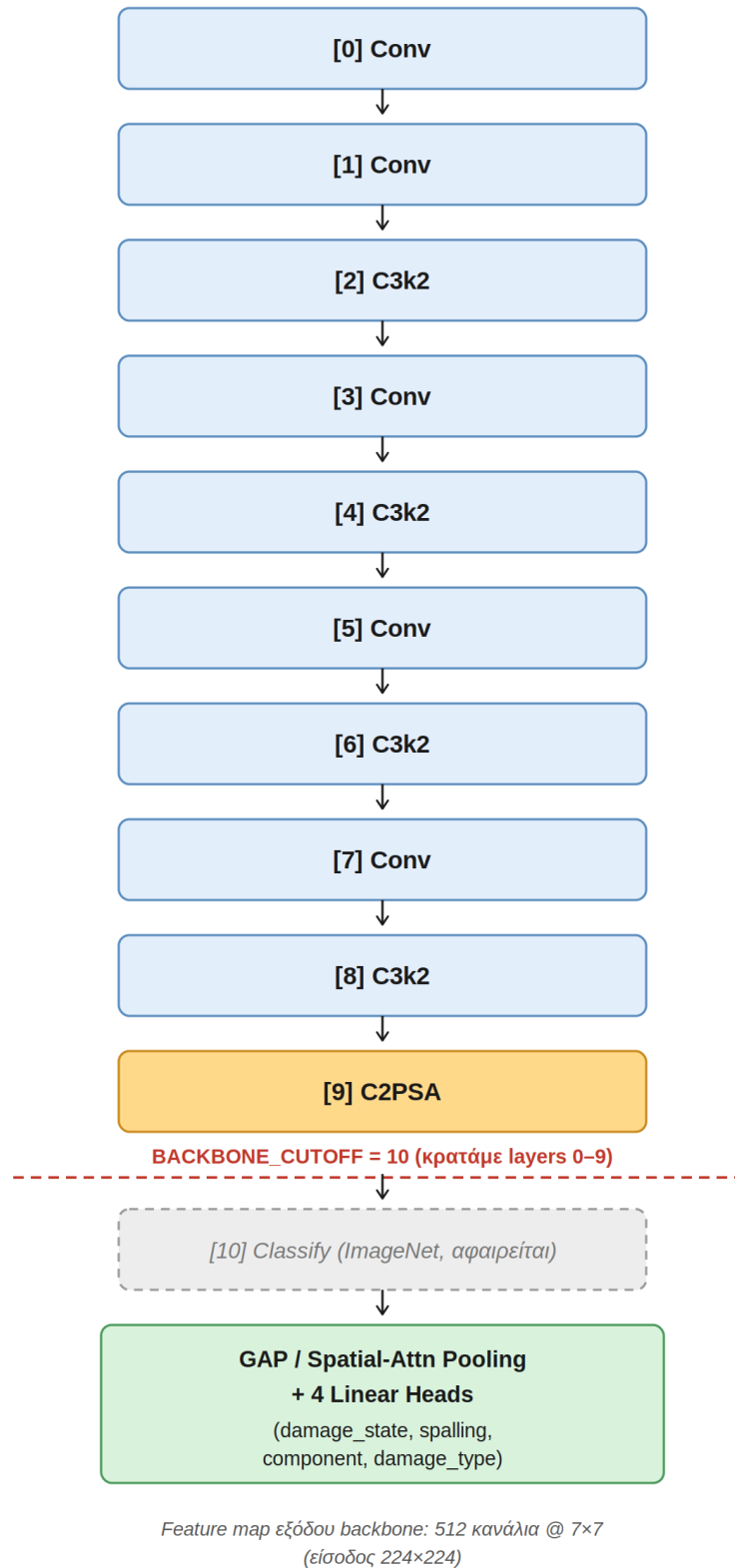
Το φ-Net παρέχει μόνο υποσύνολα εκπαίδευσης (train) και ελέγχου (test), χωρίς ξεχωριστό υποσύνολο επικύρωσης (validation). Για την κάλυψη αυτού του κενού εφαρμόστηκε στρωματοποιημένος (stratified) διαχωρισμός 15% του υποσυνόλου εκπαίδευσης προς επικύρωση, με σταθερό seed τυχαιότητας (seed=42) για λόγους αναπαραγωγιμότητας, μια πρακτική ανάλογη με τον χειροκίνητο διαχωρισμό 80:20 που έχει εφαρμοστεί για την εξασφάλιση ισορροπημένης αντιπροσώπευσης κλάσεων βλάβης και τύπων κατασκευής σε άλλα σύνολα δεδομένων μετασεισμικής καταγραφής [8].

3.6 Προτεινόμενη Αρχιτεκτονική: MultiTaskYOLO

Η προτεινόμενη αρχιτεκτονική αξιοποιεί ως κοινό backbone εξαγωγής χαρακτηριστικών το backbone του προεκπαιδευμένου μοντέλου YOLO11m, συμπεριλαμβανομένου του C2PSA module στο τέλος του backbone, το οποίο ενισχύει τις εξαγόμενες αναπαραστάσεις μέσω μηχανισμού αυτοπροσοχής πάνω στον χάρτη χαρακτηριστικών χαμηλότερης χωρικής ανάλυσης [7]. Πάνω από το κοινό backbone τοποθετούνται τέσσερις ανεξάρτητες κεφαλές ταξινόμησης, μία ανά task, οι οποίες εκπαιδεύονται ταυτόχρονα μέσω σταθμισμένου αθροίσματος επιμέρους συναρτήσεων απώλειας. Για εικόνες στις οποίες ένα συγκεκριμένο task δεν έχει έγκυρη ετικέτα, λόγω της μερικής επικάλυψης των tasks στο φ-Net, εφαρμόζεται στρατηγική masked loss, όπου η ετικέτα ορίζεται ως -1 και το αντίστοιχο task εξαιρείται από τον υπολογισμό της απώλειας για τη

συγκεκριμένη εικόνα, επιτρέποντας την αξιοποίηση όλων των διαθέσιμων εικόνων ανεξαρτήτως πλήρους επισήμανσης σε όλα τα tasks. Η επιλογή του μεγέθους backbone (YOLO11m έναντι YOLO11s/x) αξιολογήθηκε και επιλέχθηκε έπειτα από δοκιμές, ακολουθώντας τη λογική συγκριτικής αξιολόγησης μεγεθών μοντέλου που έχει εφαρμοστεί και σε άλλες μελέτες ανίχνευσης δομικής βλάβης [16].

YOLO11m-cls — Backbone Ταξινόμησης
(MultiTaskModel)



Εικόνα 4. Backbone ταξινόμησης YOLO11m-cls (MultiTaskYOLO): layers 0–10 (BACKBONE_CUTOFF=10) και κοινή έξοδος 512 καναλιών @ 7×7, πάνω από την οποία τοποθετούνται οι τέσσερις κεφαλές ταξινόμησης.

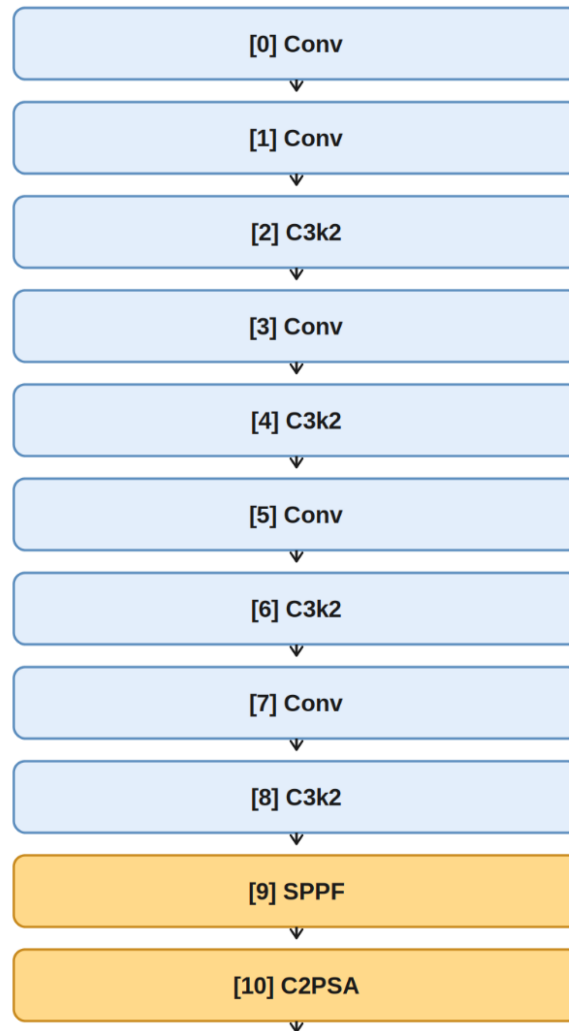
3.7 Αρχιτεκτονική Segmentation Track: YOLO11x-seg

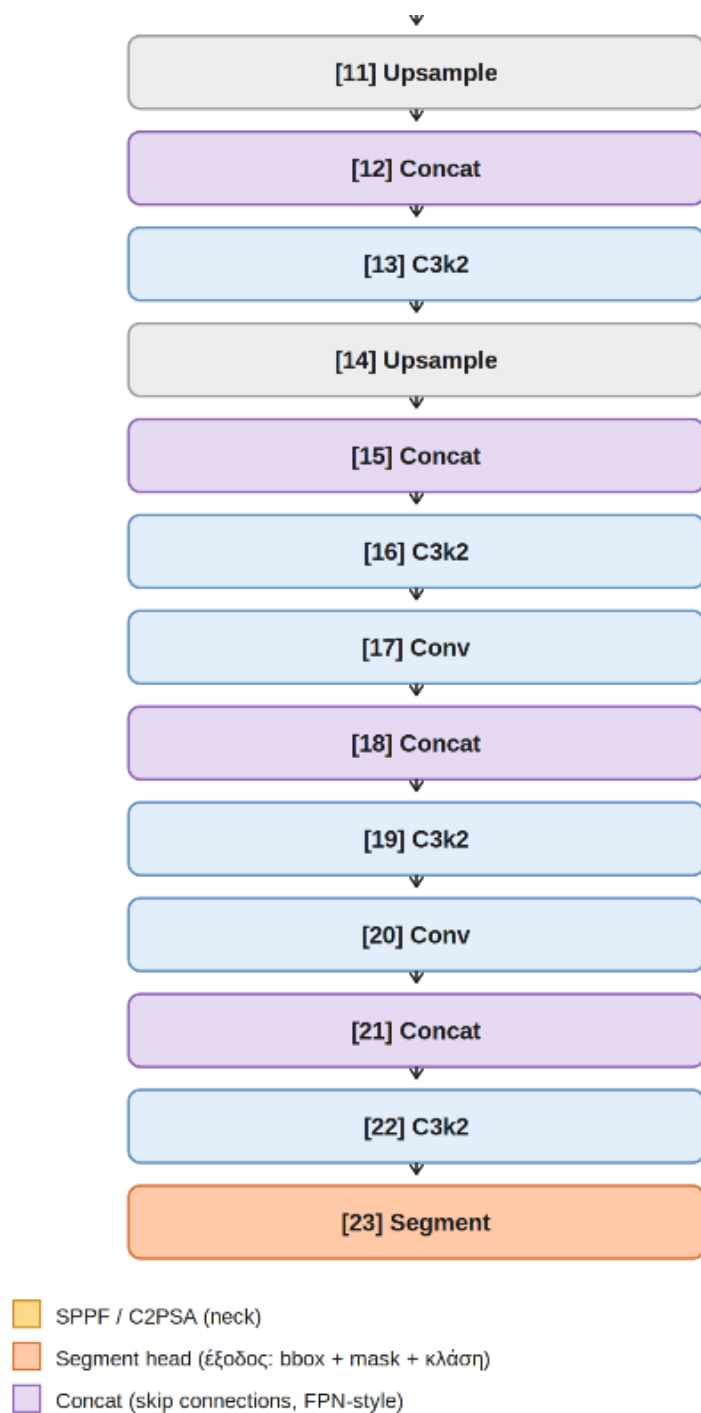
Το σύνολο δεδομένων κατάτμησης προέρχεται εξ ολοκλήρου από το task damage_type του φ-Net και κατασκευάστηκε και διαχειρίστηκε μέσω της πλατφόρμας Roboflow, η οποία χρησιμοποιήθηκε τόσο για την οργάνωση των εικόνων ανά κλάση (Shear, Flexural, Combined) όσο και για την επισημείωση (annotation) των μασκών ρωγμών. Η επισημείωση πραγματοποιήθηκε με τη λειτουργία Auto-Label and Review του Roboflow: αρχικά παρήχθη ένα πρώτο πέρασμα μασκών πολυγώνου (polygon masks) πάνω σε ολόκληρο το σύνολο εικόνων μέσω του μοντέλου SAM 3, και στη συνέχεια οι προτεινόμενες μάσκες ελέγχθηκαν και διορθώθηκαν χειροκίνητα πριν την τελική εξαγωγή του συνόλου δεδομένων. Μετά την εκκαθάριση, το τελικό σύνολο δεδομένων κατάτμησης αποτελείται από 2.374 εικόνες συνολικά, κατανεμημένες σε 1.821 εικόνες εκπαίδευσης (train), 326 εικόνες επικύρωσης (validation) και 227 εικόνες ελέγχου (test).

Το συμπληρωματικό σκέλος εντοπισμού και μέτρησης ρωγμών αξιοποιεί αυτούσια την πλήρη αρχιτεκτονική YOLO11x-seg, χωρίς δομική τροποποίηση, μέσω fine-tuning πάνω σε προεκπαιδευμένα βάρη με τη διαδικασία `model.train()` της βιβλιοθήκης Ultralytics [6]. Το backbone (layers 0–8) ακολουθεί την ίδια λογική διαδοχικών blocks Conv/C3k2 με το backbone ταξινόμησης, ενισχυμένο στο τέλος με SPPF και C2PSA (layers 9–10) για την ενίσχυση των εξαγόμενων χαρακτηριστικών μέσω self-attention [7]. Το neck ακολουθεί δομή τύπου FPN (Feature Pyramid Network), με διαδοχικά upsampling και concatenation στρώματα (layers 11–22) που συνδυάζουν χαρακτηριστικά διαφορετικής χωρικής ανάλυσης μέσω skip connections, επιτρέποντας στο δίκτυο να αξιοποιεί ταυτόχρονα λεπτομέρειες υψηλής ανάλυσης και σημασιολογική πληροφορία υψηλού επιπέδου. Το segmentation head (layer 23) παράγει ταυτόχρονα bounding box, μάσκα κατάτμησης και κλάση για κάθε ανιχνευόμενη ρωγμή.

YOLO11x-seg — Πλήρης Αρχιτεκτονική

(Segmentation Track, χωρίς τροποποίηση)





Ολόκληρο το δίκτυο εκπαιδεύεται (fine-tuning) μέσω

Εικόνα 5. Πλήρης αρχιτεκτονική YOLO11x-seg (segmentation track): backbone (layers 0–10), neck τύπου FPN με Concat/Upsample (layers 11–22) και segment head (layer 23), χωρίς καμία δομική τροποποίηση.

3.8 Διαδικασία Εκπαίδευσης

Η εκπαίδευση του μοντέλου MultiTaskYOLO ακολούθησε διαδικασία δύο φάσεων. Στην πρώτη φάση (Phase 1), το backbone διατηρήθηκε παγωμένο (frozen), με τα βάρη του να

προέρχονται από την προεκπαίδευση της αρχιτεκτονικής YOLO11m-cls, ενώ εκπαιδεύτηκαν αποκλειστικά οι τέσσερις κεφαλές ταξινόμησης, για 5 epochs, με ρυθμό μάθησης 1×10^{-3} και βελτιστοποιητή AdamW. Η στρατηγική αυτή αποσκοπεί στην ταχεία προσαρμογή τυχαία αρχικοποιημένων κεφαλών πάνω σε ήδη ποιοτικά χαρακτηριστικά, αποφεύγοντας τη μετάδοση θορυβωδών κλίσεων προς το backbone στα πρώτα στάδια της εκπαίδευσης.

Στη δεύτερη φάση (Phase 2) το σύνολο του δικτύου ξεπαγώνεται και εκπαιδεύεται με discriminative learning rates: το backbone λαμβάνει μικρότερο ρυθμό μάθησης (1×10^{-5}) σε σχέση με τις κεφαλές ταξινόμησης (1×10^{-4}), ώστε οι ήδη προσαρμοσμένες αναπαραστάσεις του backbone να υποστούν μόνο μικρές διορθώσεις, ενώ οι κεφαλές συνεχίζουν να προσαρμόζονται πιο δυναμικά. Ο ρυθμός μάθησης μειώνεται σταδιακά μέσω cosine annealing scheduler για έως 60 epochs, ενώ εφαρμόζεται αποκοπή κλίσης (gradient clipping, norm=5.0) για την αποφυγή έκρηξης κλίσεων μετά το πλήρες ξεπάγωμα του δικτύου. Η εκπαίδευση διακόπτεται νωρίτερα (early stopping) εάν δεν παρατηρηθεί βελτίωση της μέσης ακρίβειας επικύρωσης για 15 συνεχόμενα epochs, με διατήρηση πάντα του καλύτερου checkpoint.

Κατά την εκπαίδευση εφαρμόστηκε ελαφριά επαύξηση δεδομένων, περιλαμβάνοντας οριζόντια αναστροφή (πιθανότητα 50%) και τυχαίες διαταραχές φωτεινότητας/αντίθεσης, ενώ αποφεύχθηκε σκόπιμα η τυχαία περιστροφή, καθώς στρεβλώνει τη γεωμετρική πληροφορία προσανατολισμού που διακρίνει τις καμπτικές από τις διατμητικές ρωγμές [18], μια επιλογή ιδιαίτερα κρίσιμη για το task damage_type. Οι εικόνες προσαρμόστηκαν σε ανάλυση 224×224 pixel, συμβατή με το προεκπαιδευμένο backbone, και κλιμακώθηκαν στο εύρος [0,1] χωρίς αφαίρεση μέσης τιμής τύπου ImageNet, ώστε να παραμείνει συνεπής η είσοδος με αυτή στην οποία εκπαιδεύτηκε αρχικά το backbone μέσω της βιβλιοθήκης Ultralytics [6].

Η χρήση εναλλακτικών συναρτήσεων απώλειας, όπως το Focal Loss [21] για τα δύο tasks με τη χαμηλότερη απόδοση (component και damage_type), διερευνήθηκε μέσω ablation study (ενότητα 4.5), όμως δεν υιοθετήθηκε στο τελικό μοντέλο καθώς υποβάθμιζε την απόδοση σε σχέση με το baseline weighted cross-entropy, εύρημα συνεπές με την παρατήρηση ότι τεχνικές επαναστάθμισης εστίασης τείνουν να βλάπτουν tasks χαμηλού

όγκου δεδομένων όταν ο περιοριστικός παράγοντας είναι ο ίδιος ο όγκος δεδομένων και όχι η κατανομή δυσκολίας δειγμάτων.

Κατά την εκπαίδευση χρησιμοποιήθηκε επίσης ο μηχανισμός TaskBalancedSampler, ο οποίος εξισορροπεί τη συχνότητα εμφάνισης δειγμάτων από κάθε task εντός κάθε batch, αντισταθμίζοντας τη φυσική ανισορροπία στο πλήθος διαθέσιμων ετικετών ανά task στο σύνολο δεδομένων φ-Net.

3.9 Μετρικές Αξιολόγησης

Η αξιολόγηση του μοντέλου πραγματοποιήθηκε με πέντε συμπληρωματικές μετρικές ανά task, macro-σταθμισμένες όπου εφαρμόζεται. Η Accuracy εκφράζει το ποσοστό σωστών προβλέψεων. Precision και Recall δείχνουν αντίστοιχα πόσο αξιόπιστες είναι οι θετικές προβλέψεις του μοντέλου και πόσο καλά εντοπίζει τα πραγματικά θετικά δείγματα ανά κλάση, ενώ το F1-score, υπολογισμένο σε macro και weighted μορφή, αποδίδει τον αρμονικό τους μέσο με ίση ή σταθμισμένη βαρύτητα ανά κλάση αντίστοιχα, δίνοντας πιο αξιόπιστη εικόνα από την Accuracy όταν οι κλάσεις είναι ανισόρροπες. Τέλος, υπολογίστηκε το εμβαδόν κάτω από την καμπύλη ROC (AUC), με προσέγγιση one-vs-rest και macro-σταθμισμένη μέση τιμή για τα multi-class tasks.

Για τη συνολική αξιολόγηση του μοντέλου υπολογίστηκε ο μη σταθμισμένος μέσος όρος (macro average) κάθε μετρικής πάνω στα τέσσερα tasks, ώστε κάθε task να συνεισφέρει εξίσου στην τελική βαθμολογία, ανεξάρτητα από το πλήθος διαθέσιμων δειγμάτων του στο σύνολο ελέγχου, αποτρέποντας έτσι τα tasks με περισσότερα δεδομένα (π.χ. damage_state) να κυριαρχούν στατιστικά έναντι tasks με λιγότερα δεδομένα (π.χ. component).

3.10 Ενοποιημένο Pipeline Ταξινόμησης-Κατάτμησης και Επίπεδο Σύντηξης Σοβαρότητας (Severity Fusion)

Το μοντέλο ταξινόμησης παρέχει ετικέτες σε επίπεδο εικόνας για τα τέσσερα tasks, χωρίς όμως καμία πληροφορία για τη χωρική έκταση της βλάβης, ενώ το segmentation track παρέχει ακριβή χωρική οριοθέτηση της ρωγμής, αλλά μόνο για damage_type εντός του συνόλου {Shear, Flexural, Combined}, χωρίς να καλύπτει τα υπόλοιπα τρία χαρακτηριστικά (κατάσταση βλάβης, αποφλοιώση, τύπος δομικού στοιχείου). Κανένα από τα δύο μοντέλα δεν παράγει από μόνο του μια πρακτικά αξιοποιήσιμη ετικέτα

σοβαρότητας του τύπου που χρησιμοποιείται σε πρωτόκολλα μετασεισμικής επιθεώρησης. Για τον λόγο αυτό αναπτύχθηκε ένα τρίτο, μη εκπαιδεύσιμο συστατικό του συστήματος, το επίπεδο σύντηξης σοβαρότητας (severity fusion layer), το οποίο συνδυάζει τις εξόδους και των δύο μοντέλων σε μία ενιαία ετικέτα τριών επιπέδων.

Η ενοποίηση των δύο σκελών υλοποιείται μέσω ενός pipeline δύο μοντέλων υπό συνθήκη (two-model conditional pipeline). Το μοντέλο ταξινόμησης εκτελείται πάντοτε, για κάθε εικόνα εισόδου. Το segmentation track ενεργοποιείται μόνο όταν η πρόβλεψη damage_type του μοντέλου ταξινόμησης διαφέρει από Undamaged, τόσο για λόγους υπολογιστικής αποδοτικότητας, (αποφεύγεται η εκτέλεση του πολύ βαρύτερου YOLO11x-seg σε εικόνες χωρίς ορατή βλάβη) όσο και για λόγους λογικής συνέπειας, καθώς μια μη ρηγματωμένη εικόνα δεν έχει νόημα να αξιολογηθεί ως προς ποσοστό κάλυψης ρωγμής. Όταν το segmentation track ενεργοποιείται, υπολογίζεται το ποσοστό κάλυψης της εικόνας από τις προβλεπόμενες μάσκες ρωγμών (coverage_pct), το οποίο κατηγοριοποιείται σε τέσσερις ζώνες: Minimal (<5%), Light (5–15%), Moderate (15–30%) και Extensive (>30%).

Ο πυρήνας του επιπέδου σύντηξης είναι ένας κανονιστικός (rule-based) πίνακας αναζήτησης διαστάσεων 12×4, ο οποίος αντιστοιχίζει κάθε συνδυασμό τύπου δομικού στοιχείου (Column, Beam, Wall, Other) και τύπου βλάβης (Shear, Combined, Flexural) ανά ζώνη κάλυψης σε μία από τρεις ετικέτες σοβαρότητας: Green (Κατοικήσιμο), Yellow (Χρειάζεται Παρεμβάσεις) και Red (Κατεδαφιστέο). Ο πίνακας παρουσιάζεται αναλυτικά στον Πίνακα 1.

Component	Damage Type	Minimal (<5%)	Light (5–15%)	Moderate (15–30%)	Extensive (>30%)
Column	Shear	Yellow	Yellow	Yellow	Red
Column	Combined	Yellow	Yellow	Red	Red
Column	Flexural	Green	Green	Yellow	Red
Beam	Shear	Yellow	Yellow	Yellow	Red
Beam	Combined	Yellow	Yellow	Yellow	Red
Beam	Flexural	Green	Green	Yellow	Red
Wall	Shear	Yellow	Yellow	Yellow	Red
Wall	Combined	Yellow	Yellow	Yellow	Red
Wall	Flexural	Green	Green	Yellow	Red
Other	Shear	Green	Green	Yellow	Red
Other	Combined	Green	Green	Yellow	Red
Other	Flexural	Green	Green	Green	Red

Πίνακας 1. Πίνακας σύντηξης σοβαρότητας (severity fusion table): Component × Damage Type × Coverage Zone.

Πέραν του βασικού πίνακα, εφαρμόζονται τρεις κανόνες παράκαμψης (hard overrides), οι οποίοι προηγούνται της αναζήτησης στον πίνακα:

(α) αν damage_state ή damage_type ισούται με Undamaged, η ετικέτα ορίζεται αναγκαστικά ως Green, ανεξάρτητα από τυχόν θορυβώδεις προβλέψεις άλλων tasks·

(β) αν spalling=Yes, η ετικέτα που προκύπτει από τον βασικό πίνακα αναβαθμίζεται κατά ένα επίπεδο σοβαρότητας (Green→Yellow ή Yellow→Red), καθώς η αποφλοίωση σκυροδέματος αποτελεί ένδειξη προχωρημένης υποβάθμισης του δομικού στοιχείου, όπως αναλύθηκε στην ενότητα 3.3·

(γ) κάθε συνδυασμός με ζώνη κάλυψης Extensive (>30%) αντιστοιχίζεται καθολικά σε Red, ανεξάρτητα από τύπο δομικού στοιχείου ή τύπο βλάβης.

Η σχεδίαση αυτή ακολουθεί σε γενικές γραμμές τη λογική τρι-επίπεδης ταξινόμησης (habitable / needs-intervention / demolition) που υιοθετούν επίσημα πρωτόκολλα ταχείας μετασεισμικής επιθεώρησης, όπως το ελληνικό πλαίσιο του Γ.Δ.Α.Ε.Φ.Κ. (2017) και το αντίστοιχο διεθνές πρότυπο ATC-20, χωρίς ωστόσο να αποτελεί πιστή εφαρμογή κάποιου συγκεκριμένου επίσημου πρωτοκόλλου, καθώς οι συγκεκριμένες τιμές ορίων και η στάθμιση ανά δομικό στοιχείο ορίστηκαν ευρετικά, με βάση τη μηχανική ερμηνεία των tasks (ενότητα 3.3), και δεν έχουν επικυρωθεί έναντι πραγματικών ετικετών σοβαρότητας από ειδικό.

Κεφάλαιο 4. Πειραματικά Αποτελέσματα

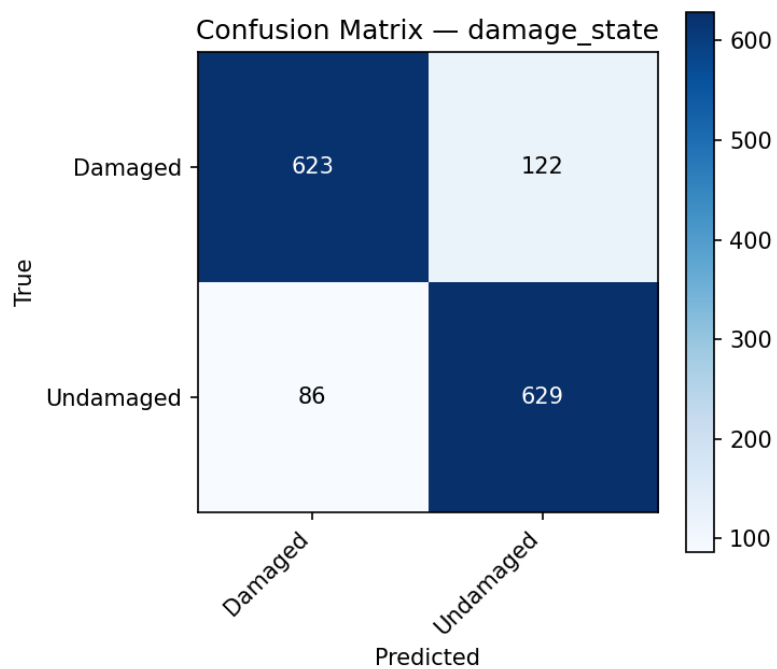
4.1 Αποτελέσματα Ταξινόμησης (Classification Track)

Η τελική αξιολόγηση του μοντέλου MultiTaskYOLO πραγματοποιήθηκε στο σύνολο ελέγχου (test set) του φ-Net (3.368 εικόνες συνολικά, με διαφορετικό αριθμό έγκυρων ετικετών ανά task λόγω της μερικής επικάλυψης tasks), χρησιμοποιώντας το καλύτερο checkpoint της Φάσης 2 (epoch 19, val_mean_acc=0,8024). Ο Πίνακας 2 συνοψίζει τις μετρικές ανά task και τον macro-σταθμισμένο μέσο όρο τους.

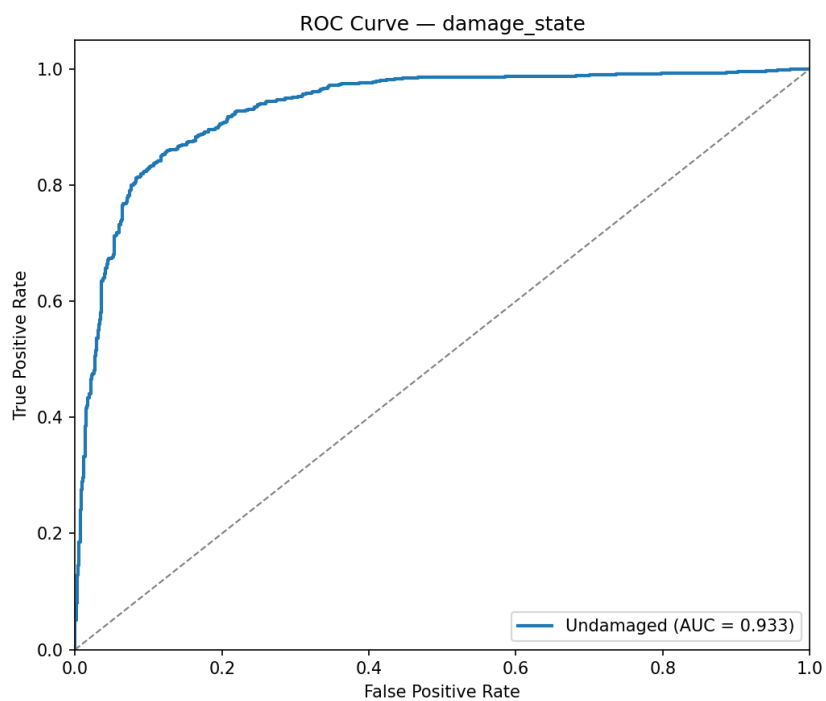
Task	Accuracy	Precision	Recall	F1 macro	F1 weighted	AUC
damage_state	0,858	0,86	0,86	0,858	0,858	0,933
spalling	0,817	0,80	0,81	0,805	0,817	0,881
component	0,743	0,62	0,65	0,631	0,751	0,856
damage_type	0,657	0,58	0,58	0,577	0,659	0,795
MEAN (macro avg)	0,768	0,715	0,725	0,717	0,771	0,866

Πίνακας 2. Συγκεντρωτικά αποτελέσματα ταξινόμησης στο σύνολο ελέγχου, ανά task και macro average.

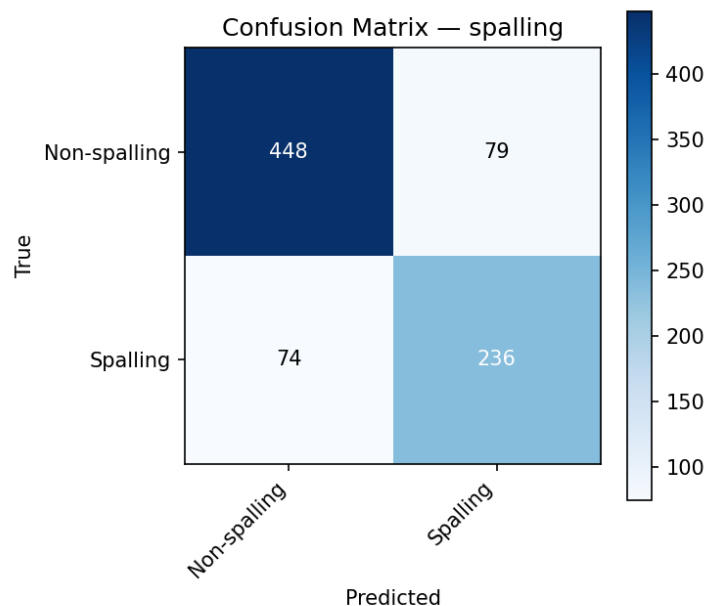
Το damage_state πέτυχε την υψηλότερη επίδοση (Accuracy=85,75%, AUC=0,933), αναμενόμενο για ένα δυαδικό task με σαφή οπτικά χαρακτηριστικά· η αντίστοιχη καμπύλη ROC (Εικόνα 7) πλησιάζει έντονα την πάνω αριστερή γωνία, επιβεβαιώνοντας οπτικά την υψηλή διαχωριστική ικανότητα του μοντέλου. Ο Πίνακας σύγκρισης (Εικόνα 6) δείχνει ισορροπημένα λάθη και στις δύο κατευθύνσεις (122 false negatives, 86 false positives σε 1.460 δείγματα). Το spalling ακολούθησε με Accuracy=81,72% και F1 macro=0,805 (Εικόνα 8), με το recall της κλάσης Spalling (0,76) χαμηλότερο από αυτό της Non-spalling (0,85), η καμπύλη ROC (Εικόνα 9) παραμένει ωστόσο σαφώς πάνω από τη διαγώνιο τυχαίας πρόβλεψης σε όλο το εύρος.



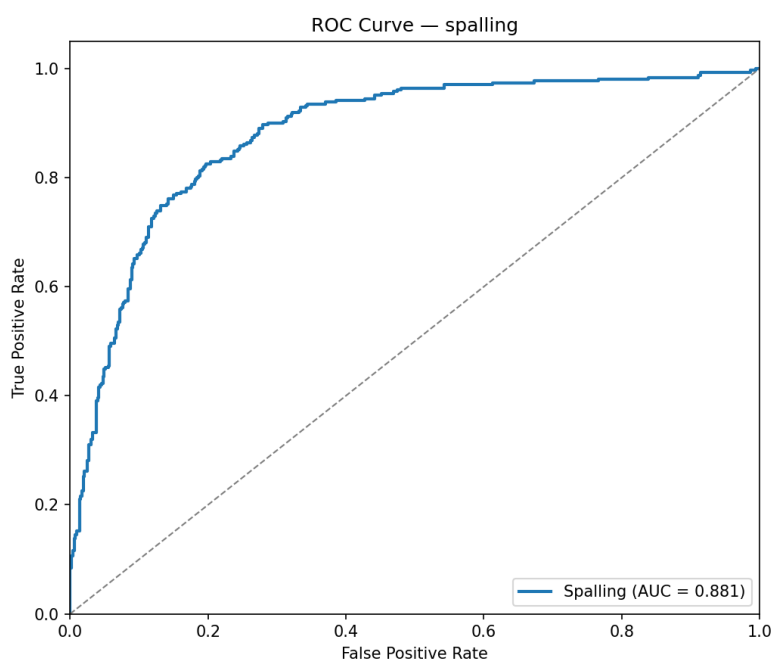
Εικόνα 6. Πίνακας σύγκρισης - damage_state.



Εικόνα 7. Καμπύλη ROC - damage_state (AUC=0,933).



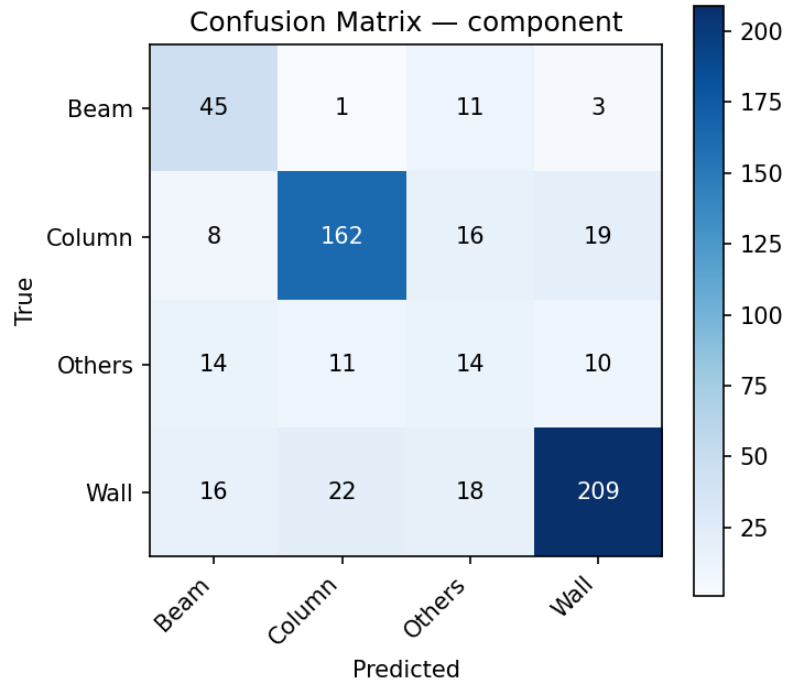
Εικόνα 8. Πίνακας σύγχυσης - spalling



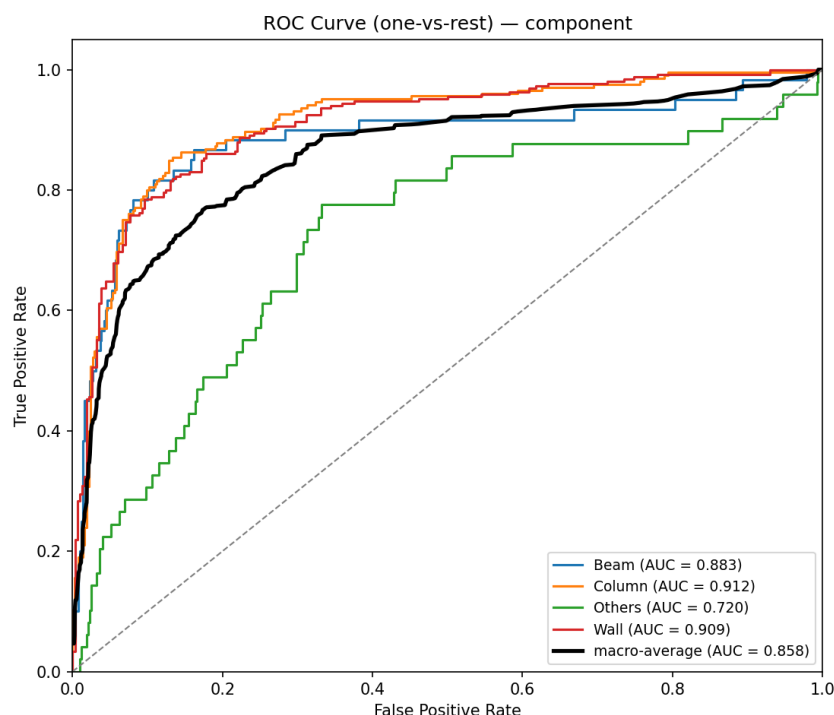
Εικόνα 9. Καμπύλη ROC - spalling (AUC=0,881).

Το component πέτυχε Accuracy =74,27% αλλά μόλις F1 macro=0,631 (Εικόνα 10), ένα σημαντικό χάσμα που οφείλεται σχεδόν αποκλειστικά στην κλάση Others: με precision=0,24 και recall=0,29 σε 49 δείγματα ελέγχου, η κλάση αυτή λειτουργεί ως ετερογενής «κάδος» δομικών στοιχείων χωρίς σαφές κοινό οπτικό πρότυπο, σε αντίθεση

με τις Beam, Column και Wall, οι οποίες πέτυχαν F1-score 0,63 - 0,83. Το ίδιο πρότυπο επιβεβαιώνεται καθαρά στις καμπύλες ROC ανά κλάση (Εικόνα 11): η κλάση Others έχει σαφώς την πιο χαμηλή καμπύλη (AUC=0,720), ενώ οι Column, Wall και Beam κινούνται σε παρόμοιο, πολύ υψηλότερο επίπεδο (AUC=0,88 - 0,91).

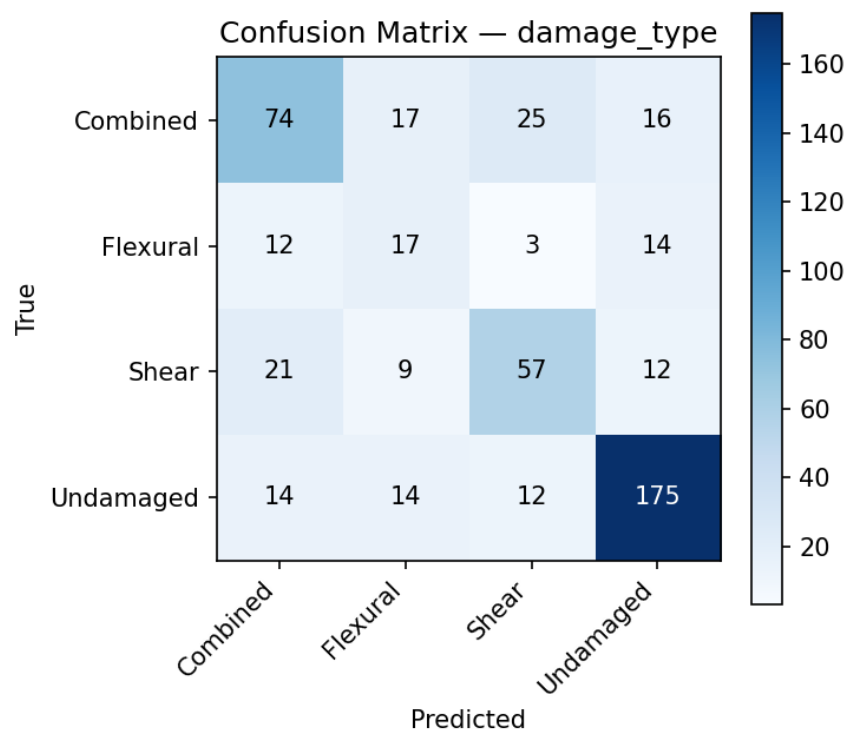


Εικόνα 10. Πίνακας σύγχυσης - component.

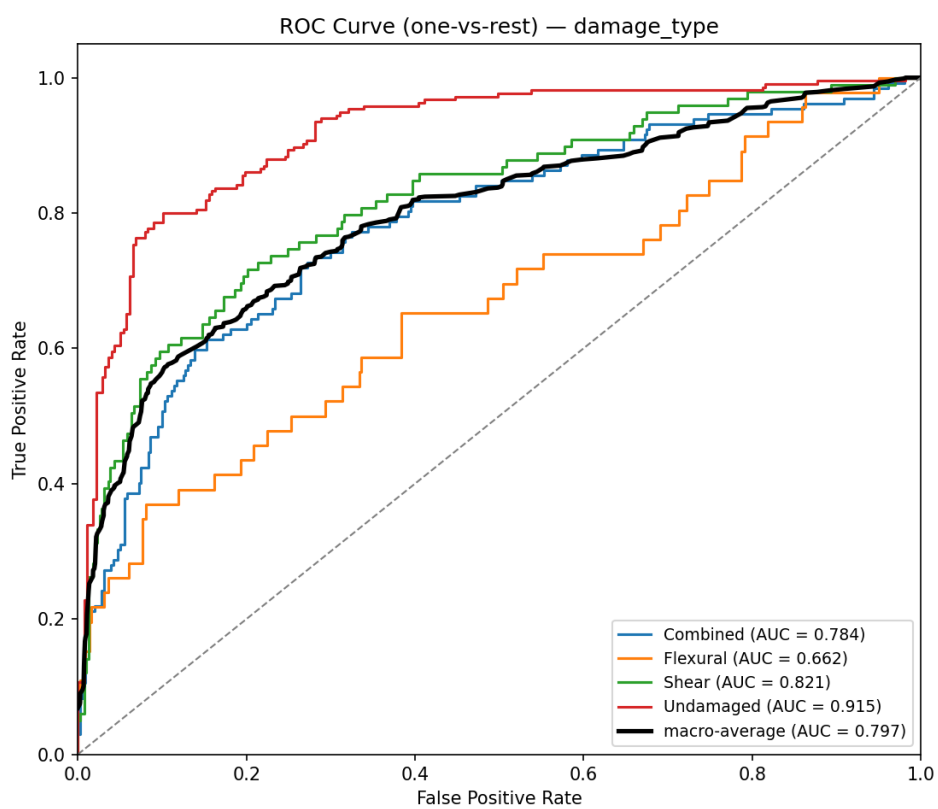


Εικόνα 11. Καμπύλες ROC (one-vs-rest) - component, ανά κλάση και macro-average (AUC=0,858).

Το damage_type παρέμεινε το πιο απαιτητικό task (Accuracy=65,65%, F1 macro=0,577, AUC=0,795, Εικόνα 12). Η ανάλυση ανά κλάση αναδεικνύει την κλάση Flexural ως το κύριο σημείο αδυναμίας (precision=0,30, recall=0,34, μόλις 17/46 σωστές προβλέψεις στο σύνολο ελέγχου), με τα λάθη να κατανέμονται σχεδόν εξίσου προς Combined (12) και Undamaged (14). Οι κλάσεις Shear και Combined πέτυχαν καλύτερη, αν και όχι υψηλή, απόδοση (F1≈0,58), με σημαντική αμοιβαία σύγχυση μεταξύ τους (21 Shear→Combined, 17 Combined→Shear), αναμενόμενη λόγω της οπτικής επικάλυψης μεταξύ διατμητικών και σύνθετων ρωγμών. Οι καμπύλες ROC ανά κλάση (Εικόνα 13) επιβεβαιώνουν την ίδια ιεράρχηση: η Flexural έχει σαφώς τη χαμηλότερη καμπύλη (AUC=0,662), η Undamaged την υψηλότερη (AUC=0,915), ενώ οι Shear και Combined κινούνται σε ενδιάμεσο επίπεδο (AUC=0,82 και 0,78 αντίστοιχα).



Εικόνα 12. Πίνακας σύγχυσης - damage_type.



Εικόνα 13. Καμπύλες ROC (one-vs-rest) - damage_type, ανά κλάση και macro-average (AUC=0,797).

4.2 Επίδραση Κανόνα HEA-style Gating

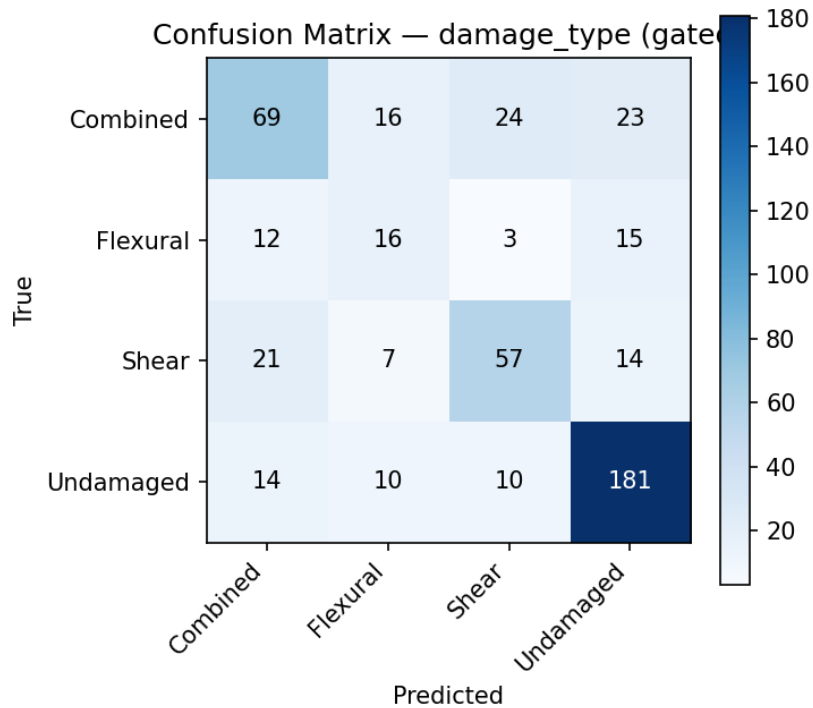
Ο κανόνας HEA-style gating εφαρμόστηκε μετά την εκπαίδευση, ως post-hoc διόρθωση των προβλέψεων damage_type με βάση τη συνέπειά τους με τις προβλέψεις damage_state, με στόχο τη μείωση λογικά ασυνεπών συνδυασμών (π.χ. damage_state=Undamaged με damage_type≠Undamaged). Η επίδρασή του στο σύνολο ελέγχου (492 εικόνες) συνοψίζεται στον Πίνακα 3.

HEA-style gating — damage_type (πριν vs μετά)

	Accuracy	F1 macro	Αλλαγμένες προβλέψεις
Πριν το gating	0.657	0.577	-
Μετά το gating	0.657	0.573	16 / 492

Πίνακας 3. Σύγκριση μετρικών damage_type πριν και μετά το HEA-style gating.

Το gating τροποποίησε 16 από τις 492 προβλέψεις (3,3%), χωρίς να μεταβάλει το Accuracy (65,65% και στις δύο περιπτώσεις) και με ελαφρά υποβάθμιση του F1 macro (από 0,577 σε 0,573). Ο πίνακας σύγκρισης μετά το gating (Εικόνα 14) δείχνει ότι ο κανόνας κέρδισε λίγες επιπλέον σωστές προβλέψεις Undamaged (175→181) αλλά «θυσίασε» ορισμένες σωστές προβλέψεις Combined (74→69) και Flexural (17→16), με καθαρό αποτέλεσμα ελαφρώς αρνητικό έως ουδέτερο.



Εικόνα 14. Πίνακας σύγχυσης - damage_type μετά το HEA-style gating.

4.3 Αποτελέσματα Segmentation Track

Το μοντέλο YOLO11x-seg εκπαιδεύτηκε για 100 epochs (2,62 ώρες συνολικά) στο εξειδικευμένο σύνολο δεδομένων κατάτμησης τριών κλάσεων. Ο Πίνακας 4 παρουσιάζει τις μετρικές του καλύτερου checkpoint στο πραγματικό σύνολο ελέγχου (test set, 227 εικόνες, 2.947 instances), τόσο συνολικά όσο και ανά κλάση, σε επανάληψη της αρχικής αξιολόγησης ώστε να είναι απευθείας συγκρίσιμες με το classification track (ενότητα 4.1).

	Precision	Recall	mAP50	mAP50-95
Box — Overall (σύνολο ελέγχου)	0,360	0,348	0,338	0,242
Mask — Overall (σύνολο ελέγχου)	0,274	0,286	0,250	0,124
Mask — Shear (σύνολο ελέγχου)	0,312	0,138	0,194	0,0951
Mask — Flexural (σύνολο ελέγχου)	0,218	0,461	0,334	0,1701
Mask — Combined (σύνολο ελέγχου)	0,291	0,259	0,221	0,1078

Πίνακας 4. Μετρικές segmentation track (YOLO11x-seg) στο πραγματικό σύνολο ελέγχου (test set), συνολικά και ανά κλάση, καλύτερο checkpoint.

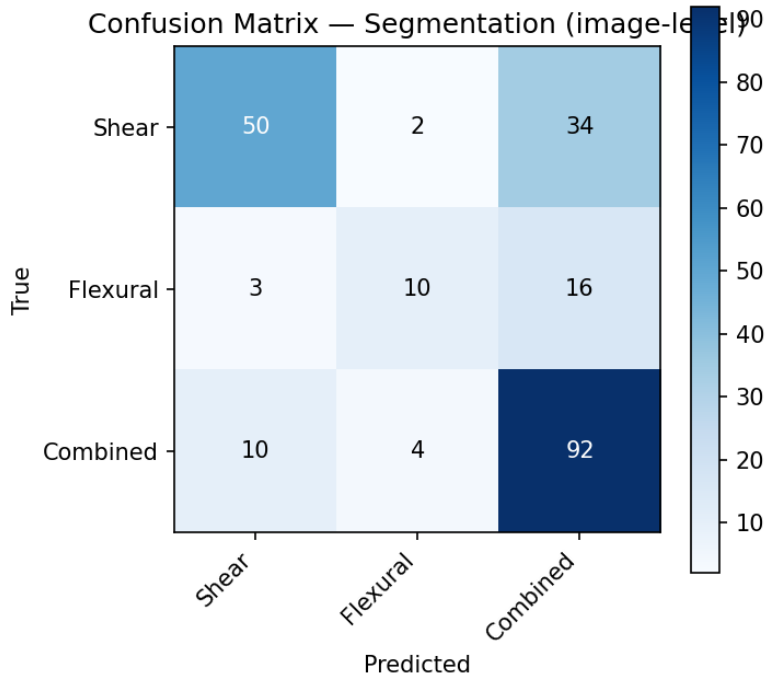
Στο πραγματικό σύνολο ελέγχου, η ανάκληση μάσκας κυμαίνεται από 13,8% (Shear) έως 46,1% (Flexural), με το Combined στο ενδιάμεσο (25,9%) - εύρημα συνεπές με το ήδη τεκμηριωμένο συμπέρασμα ότι ο περιορισμένος όγκος δεδομένων κατάτμησης αποτελεί σημαντικό περιοριστικό παράγοντα. Αξίζει να σημειωθεί ότι η κλάση Flexural πέτυχε εδώ το υψηλότερο Mask mAP50 (0,334), παρά τον μικρό αριθμό δειγμάτων της στο test set (89 instances σε 30 εικόνες, έναντι 1.178 instances σε 91 εικόνες για το Shear και 1.680 σε 106 για το Combined) - αποτέλεσμα αντίθετο από ό,τι παρατηρήθηκε στο σύνολο επικύρωσης, όπου το Flexural ήταν η ασθενέστερη κλάση, και που πρέπει να ερμηνευτεί με προσοχή λόγω του μικρού μεγέθους δείγματος. Η κλάση Shear, αντίθετα, παρουσιάζει εδώ το χαμηλότερο recall (13,8%) παρά τον μεγάλο αριθμό δειγμάτων της, υποδεικνύοντας ότι η δυσκολία εντοπισμού της δεν οφείλεται σε έλλειψη δεδομένων αλλά πιθανώς σε οπτική ασάφεια του προτύπου ρωγμής.

Επιπλέον, δοκιμάστηκε Test-Time Augmentation (TTA - πολυκλιμακωτή αξιολόγηση με οριζόντια αναστροφή, ενσωματωμένη λειτουργία του Ultralytics) καθώς και σάρωση confidence threshold (0,05–0,50), και τα δύο στο πραγματικό σύνολο ελέγχου, με στόχο τη βελτίωση του recall χωρίς επανεκπαίδευση. Το TTA δεν υποστηρίζεται από το συγκεκριμένο μοντέλο/έκδοση της βιβλιοθήκης (μήνυμα «Model does not support augment=True», με αυτόματη επιστροφή σε single-scale inference), οπότε δεν κατέστη δυνατή η αξιολόγησή του. Η σάρωση threshold επιβεβαίωσε ότι η προεπιλεγμένη τιμή (conf=0,25) αποδίδει ήδη το καλύτερο recall μάσκας (0,286) σε όλο το εύρος που εξετάστηκε, χαμηλότερα thresholds (0,05 - 0,20) οδήγησαν αντιθέτως σε ελαφρώς χειρότερο recall (0,247), ενώ υψηλότερα (0,40 - 0,50) βελτίωσαν το precision με σημαντική απώλεια recall. Το εύρημα αυτό, αν και δεν προσέφερε βελτίωση, επιβεβαιώνει ότι η προεπιλεγμένη διαμόρφωση του μοντέλου βρίσκεται ήδη κοντά στο βέλτιστο σημείο precision/recall trade-off για το συγκεκριμένο checkpoint.

4.4 Αξιολόγηση του Segmentation ως Ταξινομητή (Image-Level)

Για να καταστεί δυνατή η άμεση σύγκριση του segmentation track με το classification track ως προς το task damage_type, αναπτύχθηκε ξεχωριστό σενάριο αξιολόγησης στο οποίο η πλειοψηφική κλάση των instances που ανιχνεύονται ανά εικόνα λαμβάνεται ως πρόβλεψη image-level damage_type, με τις ίδιες μετρικές (Accuracy, F1, AUC) που χρησιμοποιήθηκαν στην ενότητα 4.1. Η αξιολόγηση πραγματοποιήθηκε στο πραγματικό σύνολο ελέγχου (test set) της κατάτμησης (227 εικόνες, εκ των οποίων 221

αξιολογήθηκαν ενώ 6 εικόνες εξαιρέθηκαν λόγω μηδενικών ανιχνεύσεων). Σημειώνεται ότι το σενάριο αυτό εξ ορισμού δεν περιλαμβάνει την κλάση Undamaged, καθώς το segmentation track δεν εκτελείται σε μη ρηγματωμένες εικόνες στο πλαίσιο του ενοποιημένου pipeline (ενότητα 3.10), οπότε η σύγκριση με το 4-κλασικό αποτέλεσμα του classification track (ενότητα 4.1) είναι ενδεικτική και όχι απευθείας συγκρίσιμη.



Εικόνα 15. Πίνακας σύγκρισης - αξιολόγηση του segmentation track ως ταξινομητή σε επίπεδο εικόνας (σύνολο ελέγχου, 3 κλάσεις).

Το segmentation-ως-classifier πέτυχε Accuracy=68,78%, F1 macro=0,619 και AUC=0,755. Η κλάση Combined παρουσίασε το πιο ασύμμετρο προφίλ (precision=0,65, recall=0,87): το μοντέλο τείνει να ταξινομεί ως Combined εικόνες με πολλαπλές ή ασαφείς ανιχνεύσεις, «τραβώντας» έτσι σημαντικό μέρος των πραγματικών Shear (34/86) και Flexural (16/29) περιπτώσεων. Η Flexural παρέμεινε η πιο αδύναμη κλάση και σε αυτό το σενάριο (recall=0,34), συνεπές εύρημα με το classification track. Η σύγκριση αυτή συζητείται περαιτέρω στην ενότητα 5.4.

4.5 Μελέτες Αφαίρεσης (Ablation Studies)

Πραγματοποιήθηκαν έξι συστηματικές μελέτες αφαίρεσης, καθεμία μεταβάλλοντας μία μόνο σχεδιαστική παράμετρο σε σχέση με το baseline (weighted cross-entropy, χωρίς

πρόσθετες τεχνικές επαναστάθμισης), ώστε να καταστεί δυνατή η καθαρή απόδοση κάθε μεταβολής. Ο Πίνακας 5 συνοψίζει ποιοτικά την κατεύθυνση κάθε αποτελέσματος. Η πλήρης ποσοτική τεκμηρίωση (ακριβείς μετρικές ανά ablation) αποτελεί εκκρεμότητα που αναφέρεται ρητά στις μελλοντικές κατευθύνσεις (ενότητα 6.4).

Ablation	Task-στόχος	Αποτέλεσμα έναντι baseline
Focal Loss	component, damage_type	Υποβάθμιση
Εκτεταμένο Spatial Attention Pooling	όλα τα tasks	Υποβάθμιση
Label Smoothing	όλα τα tasks	Υποβάθμιση
Στοχευμένη υπερδειγματοληψία κλάσεων (class-targeted oversampling)	component, damage_type	Υποβάθμιση
Συνδυασμός παραπάνω μεταβολών	όλα τα tasks	Υποβάθμιση
Segmentation imgsz=960 + retina_masks	segmentation track	Υποβάθμιση

Πίνακας 5. Συνοπτικά αποτελέσματα μελετών αφαίρεσης (qualitative summary).

Και οι έξι μελέτες υποβάθμισαν την απόδοση σε σχέση με το baseline, ένα συστηματικό εύρημα το οποίο ενισχύει την ερμηνεία ότι ο περιοριστικός παράγοντας της απόδοσης, ιδίως στα tasks component και damage_type, είναι ο όγκος των διαθέσιμων δεδομένων και όχι η αρχιτεκτονική ή η συνάρτηση απώλειας: τεχνικές που ανακατανέμουν την έμφαση του μοντέλου σε δύσκολα δείγματα (Focal Loss, oversampling) τείνουν να βλάπτουν tasks χαμηλού όγκου δεδομένων, καθώς ενισχύουν τον θόρυβο μαζί με το σήμα.

4.6 Ενοποιημένο Pipeline: Ποιοτική Αξιολόγηση με Severity Fusion

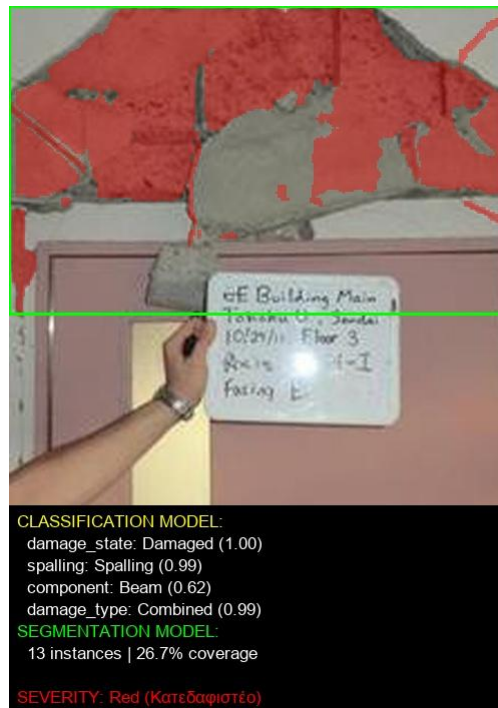
Για την ποιοτική επαλήθευση της ορθής λειτουργίας του ενοποιημένου pipeline (ενότητα 3.10) εκτελέστηκε το πλήρες pipeline (ταξινόμηση → υπό συνθήκη segmentation → severity fusion) σε δεκατέσσερις ενδεικτικές εικόνες του συνόλου ελέγχου, επιλεγμένες ώστε να καλύπτουν και τις τρεις ετικέτες σοβαρότητας καθώς και όλους τους κανόνες παράκαμψης. Ο Πίνακας 6 συνοψίζει τα αποτελέσματα.

Εικόν α	Damage State	Spallin g	Componen t	Damage Type	Κάλυψ η	Severit y	Ενεργοποιημένο ς κανόνας
00079	Damaged	Spalling	Beam	Combined	26,7%	Red	Spalling override (Yellow→Red)
00420	Damaged	Spalling	Wall	Shear	19,9%	Red	Spalling override (Yellow→Red)
00096	Damaged	Spalling	Wall	Combined	12,0%	Red	Spalling override (Yellow→Red)

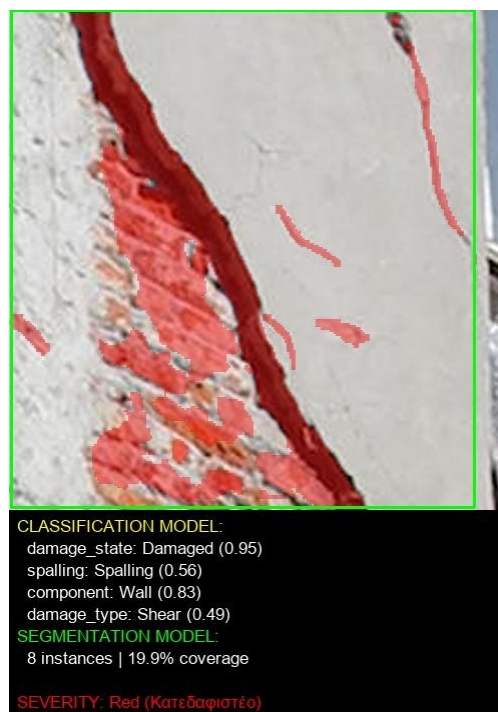
00104	Damaged	Spalling	Beam	Shear	3,9%	Red	Spalling override (Yellow→Red)
00458	Damaged	Spalling	Wall	Shear	94,3%	Red	Extensive→Red (καθολικός)
00017	Damaged	Spalling	Wall	Shear	84,7%	Red	Extensive→Red (καθολικός)
00439	Damaged	Spalling	Column	Combined	29,4%	Red	Βασικός πίνακας (χωρίς override)
00011	Damaged	Non-spalling	Wall	Shear	2,2%	Yellow	Βασικός πίνακας (χωρίς override)
00033	Damaged	Non-spalling	Column	Shear	2,0%	Yellow	Βασικός πίνακας (χωρίς override)
00026	Damaged	Non-spalling	Column	Combined	0,0%*	Yellow	Μηδενικές ανιχνεύσεις → default Minimal
00327	Undamaged	—	Beam	Undamaged	—	Green	Forced Green (Undamaged)
00392	Undamaged	Spalling	Column	Undamaged	—	Green	Forced Green (Undamaged)
00050	Undamaged	Non-spalling	Wall	Undamaged	—	Green	Forced Green (Undamaged)
00263	Undamaged	Non-spalling	Wall	Undamaged	—	Green	Forced Green (Undamaged)

Πίνακας 6. Two-Model Pipeline + Severity Fusion. Συγκεντρωτικός πίνακας δεκατεσσάρων ενδεικτικών παραδειγμάτων.

Τέσσερα από τα δεκατέσσερα παραδείγματα (00079, 00420, 00096, 00104) ενεργοποιούν τον κανόνα παράκαμψης του spalling, αναβαθμίζοντας την ετικέτα από Yellow σε Red· πρόκειται για περιπτώσεις όπου η αποφλοίωση σκυροδέματος (ενδεικτικό εύρημα προχωρημένης υποβάθμισης, ενότητα 3.3) μεταβάλλει ουσιαστικά την τελική απόφαση σοβαρότητας σε σχέση με αυτή που θα προέκυπτε μόνο από τον βασικό πίνακα.



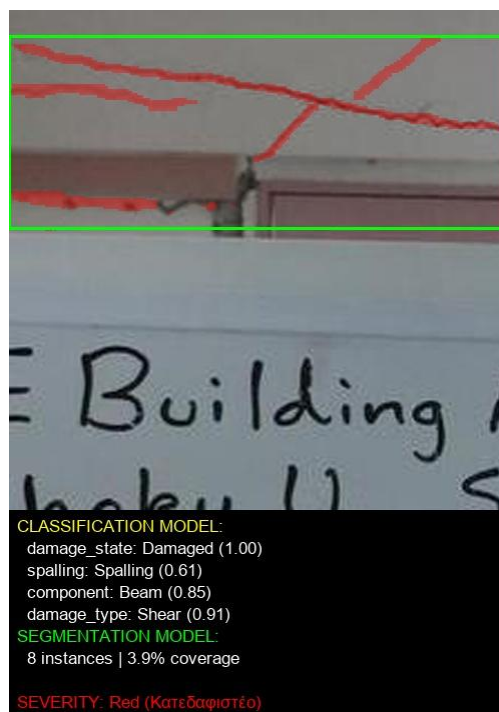
Εικόνα 16. Spalling override: Beam, Combined, κάλυψη 26,7%.



Εικόνα 17. Spalling override: Wall, Shear, κάλυψη 19,9%.



Εικόνα 18. Spalling override: Wall, Combined, κάλυψη 12,0%.



Εικόνα 19. Spalling override: Beam, Shear, κάλυψη 3,9%.

Δύο παραδείγματα (00458, 00017) ενεργοποιούν τον καθολικό κανόνα Extensive→Red λόγω κάλυψης βλάβης άνω του 30%, ανεξάρτητα από τύπο δομικού στοιχείου ή τύπο βλάβης.



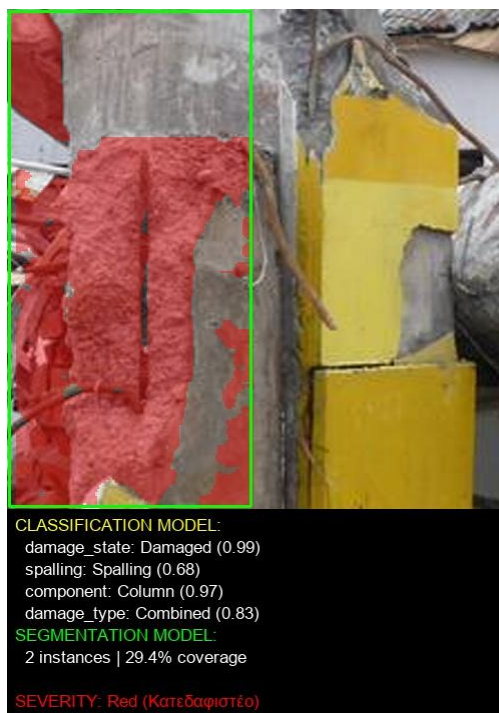
Εικόνα 20. Extensive→Red: Wall, Shear, κάλυψη 94,3%.



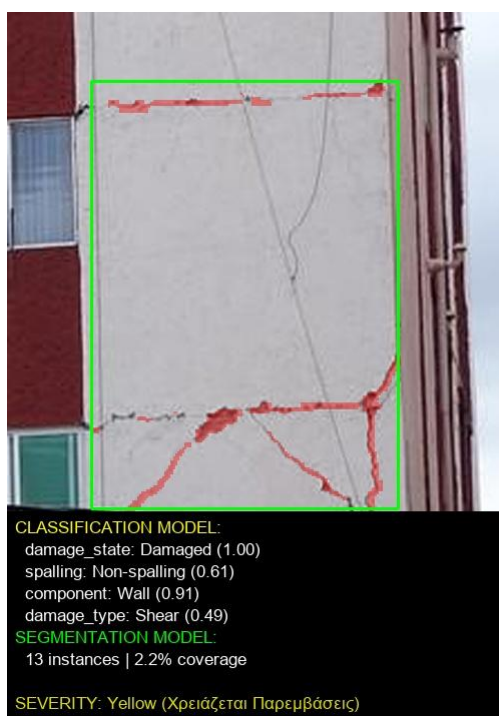
Εικόνα 21. Extensive→Red: Wall, Shear, κάλυψη 84,7% (damage_state confidence 0,68).

Τρία παραδείγματα (00439, 00011, 00033) αντιστοιχίζονται απευθείας από τον βασικό πίνακα, χωρίς να απαιτείται κανένας κανόνας παράκαμψης - το 00439 σε Red (Column,

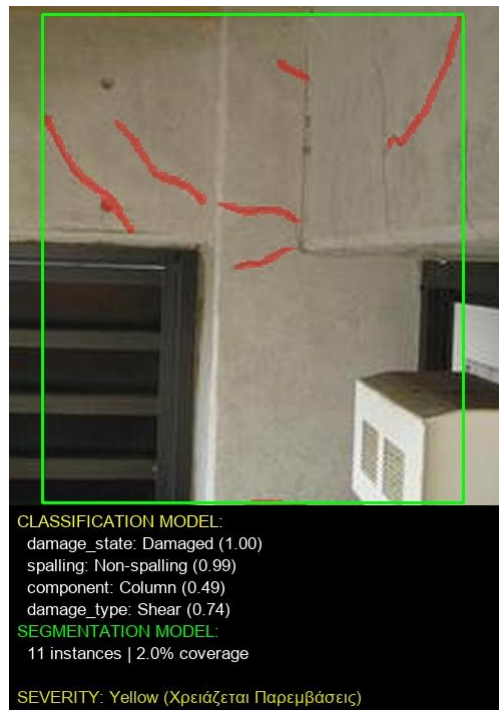
Combined, ζώνη Moderate) και τα 00011, 00033 σε Yellow (χαμηλή κάλυψη, χωρίς spalling).



Εικόνα 22. Βασικός πίνακας: Column, Combined, κάλυψη 29,4%.



Εικόνα 23. Βασικός πίνακας: Wall, Shear, κάλυψη 2,2%.



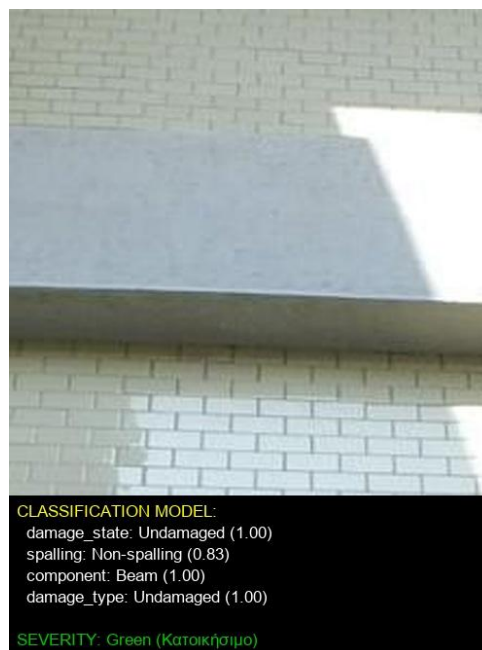
Εικόνα 24. Βασικός πίνακας: Column, Shear, κάλυψη 2,0%.

Το παράδειγμα 00026 αναδεικνύει μια οριακή περίπτωση: παρότι το classification track προέβλεψε `damage_type=Combined` (άρα το segmentation track ενεργοποιήθηκε κανονικά), το YOLO11x-seg δεν εντόπισε καμία instance στη συγκεκριμένη εικόνα. Η απουσία ανιχνεύσεων μεταφράζεται σε κάλυψη 0%, η οποία εμπίπτει από default στη ζώνη Minimal, οδηγώντας σε ετικέτα Yellow αντί για μια πιθανώς υψηλότερη σοβαρότητα αν η ρωγμή είχε εντοπιστεί σωστά. Το ζήτημα αυτό συζητείται περαιτέρω ως περιορισμός του pipeline στην ενότητα 5.5.



Εικόνα 25. Οριακή περίπτωση: μηδενικές ανιχνεύσεις segmentation παρά damage_type=Combined.

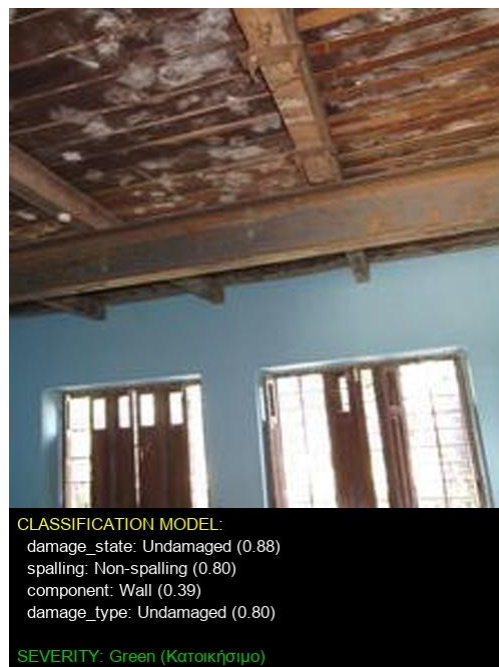
Τέλος, τέσσερα παραδείγματα (00327, 00392, 00050, 00263) επιβεβαιώνουν τον κανόνα forced Green για εικόνες με damage_state/damage_type=Undamaged, ανεξάρτητα από τυχόν θορυβώδεις προβλέψεις άλλων tasks - χαρακτηριστικά, στο 00392 το spalling προβλέπεται (εσφαλμένα, με χαμηλή βεβαιότητα 0,50) ως Spalling, χωρίς ωστόσο αυτό να επηρεάζει την τελική ετικέτα, ακριβώς όπως προβλέπει ο κανόνας.



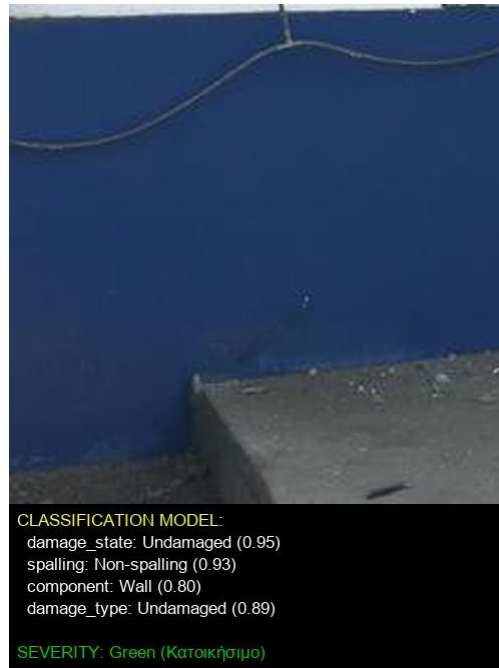
Εικόνα 26. Forced Green: Beam, Undamaged.



Εικόνα 27. Forced Green: Column, Undamaged (spalling πρόβλεψη αγνοείται).



Εικόνα 28. Forced Green: Wall, Undamaged.



Εικόνα 29. Forced Green: Wall, Undamaged.

Η παρούσα αξιολόγηση παραμένει ποιοτική, σε διευρυμένο πλέον δείγμα δεκατεσσάρων εικόνων, και δεν αντικαθιστά μια συστηματική ποσοτική κατανομή Green/Yellow/Red σε ολόκληρο το σύνολο ελέγχου, ούτε επικύρωση έναντι πραγματικών ετικετών σοβαρότητας από ειδικό μηχανικό — αμφότερα παραμένουν ανοιχτά και αναφέρονται ως περιορισμοί (ενότητα 5.5) και ως μελλοντική κατεύθυνση (ενότητα 6.4).

Κεφάλαιο 5. Συζήτηση

5.1 Σύγκριση με τη Σχετική Βιβλιογραφία

Στο task `damage_type`, το μοντέλο πέτυχε $\text{Accuracy}=65,65\%$, σημαντικά υψηλότερο από το $34,6\%$ που αναφέρουν οι Jiang et al. [13] για προσέγγιση βασισμένη σε μεγάλο γλωσσικό μοντέλο όρασης πάνω στο ίδιο σύνολο δεδομένων φ-Net. Η διαφορά αυτή ενισχύει την υπόθεση ότι μια εξειδικευμένη, συμπαγής αρχιτεκτονική CNN, εκπαιδευμένη ειδικά για το συγκεκριμένο, στενά ορισμένο πρόβλημα ταξινόμησης, μπορεί να υπερτερεί γενικού σκοπού μοντέλων μεγάλης κλίμακας που δεν έχουν εξειδικευτεί σε αυτό. Στο task `component`, το $74,27\%$ Accuracy βρίσκεται σε γενικές γραμμές στην ίδια τάξη μεγέθους με το $81,6\%$ που αναφέρουν οι Türier et al. [11], [12], αν και η δική μας μετρική $F1 \text{ macro}$ ($0,631$) αναδεικνύει καθαρότερα την επίδραση της ανισόρροπης κλάσης Others, ζήτημα που δεν φαίνεται να σχολιάζεται ρητά στην αντίστοιχη εργασία.

Ως προς το `segmentation track`, η άμεση αριθμητική σύγκριση με εργασίες όπως οι Mallepalli et al. [9] είναι δύσκολη λόγω διαφορετικών συνόλων δεδομένων και ορισμού κλάσεων, ωστόσο το ποιοτικό εύρημα της χαμηλής ανάκλησης σε συνθήκες περιορισμένου όγκου δεδομένων κατάτμησης συμφωνεί με παρατηρήσεις άλλων εργασιών που τονίζουν τη σημασία του μεγέθους του συνόλου δεδομένων έναντι αρχιτεκτονικών τροποποιήσεων για την ανίχνευση ρωγμών μικρής κλίμακας [8], [10].

5.2 Η Δυσκολία του Task `damage_type`

Η συστηματικά χαμηλότερη απόδοση του `damage_type` σε σχέση με τα υπόλοιπα τρία tasks, τόσο στο `classification track` ($\text{Accuracy}=65,65\%$) όσο και στο `segmentation`, ως classifier σενάριο ($\text{Accuracy}=68,78\%$), επιβεβαιώνει τη μηχανική παρατήρηση της ενότητας 3.3 ότι η διάκριση καμπτικής, διατμητικής και σύνθετης βλάβης βασίζεται σε λεπτές γεωμετρικές διαφορές προσανατολισμού ρωγμής, οι οποίες μπορεί να είναι δυσδιάκριτες σε φωτογραφίες χαμηλής ανάλυσης, υπό μη ιδανικό φωτισμό, ή όταν συνυπάρχουν πολλαπλά πρότυπα ρωγμών στην ίδια εικόνα. Η κλάση `Flexural` αναδείχθηκε ως το κύριο σημείο αδυναμίας και στα δύο σενάρια αξιολόγησης, πιθανότατα λόγω του μικρότερου αριθμού διαθέσιμων δειγμάτων εκπαίδευσης (405 εικόνες, έναντι 1.015 για `Combined` και 1.359 για `Undamaged`) σε συνδυασμό με την

οπτική επικάλυψή της τόσο με το Combined όσο και με το Undamaged σε ήπιες περιπτώσεις βλάβης.

5.3 Το Εύρημα του HEA-style Gating

Το γεγονός ότι ο κανόνας HEA-style gating δεν βελτίωσε -και οριακά υποβάθμισε- το F1 macro του damage_type αποτελεί εύρημα με ευρύτερο ενδιαφέρον για τη σχεδίαση πλαισίων πολυκατηγορικής σύντηξης. Η λογική υπόθεση πίσω από τον κανόνα, ότι η πρόβλεψη damage_state μπορεί να χρησιμοποιηθεί ως αξιόπιστο «gate» για τη διόρθωση ασυνεπών προβλέψεων damage_type, προϋποθέτει ότι το gate task είναι το ίδιο αρκετά ακριβές ώστε οι διορθώσεις να είναι κατά μέσο όρο ωφέλιμες. Με Accuracy=85,75% στο damage_state, ένα μη αμελητέο ποσοστό λανθασμένων gate αποφάσεων μεταφράζεται σε λανθασμένες διορθώσεις damage_type, οι οποίες μπορούν να αντισταθμίσουν ή και να υπερκεράσουν το όφελος από τις σωστές διορθώσεις. Το εύρημα αυτό υποδεικνύει ότι η επιβολή λογικών περιορισμών μεταξύ tasks σε πολυκατηγορικά πλαίσια δεν αποδίδει αυτόματα όφελος και χρειάζεται να αξιολογείται εμπειρικά, όχι μόνο να αιτιολογείται θεωρητικά.

5.4 Περιορισμοί του Segmentation Track

Πέρα από το χαμηλό recall που συζητήθηκε στην ενότητα 4.3, η σύγκριση του segmentation ως classifier σεναρίου (ενότητα 4.4) με το classification track αναδεικνύει μια πρόσθετη μεθοδολογική δυσκολία: τα δύο σενάρια αξιολογούνται σε διαφορετικά υποσύνολα εικόνων (492 εικόνες με 4 κλάσεις για το classification, 221 εικόνες με 3 κλάσεις, χωρίς Undamaged, για το segmentation), και επομένως το γεγονός ότι το classification πέτυχε χαμηλότερο Accuracy (65,65% έναντι 68,78%) δεν επιτρέπει το συμπέρασμα ότι το segmentation track είναι εγγενώς ανώτερο στη διάκριση τύπου βλάβης. Η κλάση Undamaged, την οποία περιλαμβάνει μόνο το πρώτο σενάριο, ταξινομείται με σχετικά υψηλή ακρίβεια ($F1=0,81$) και «ανεβάζει» τεχνητά τον μέσο όρο του classification track προς τα κάτω όταν συγκρίνεται με ένα σενάριο 3 κλάσεων. Μια δίκαιη ποσοτική σύγκριση θα απαιτούσε αξιολόγηση του classification track στο ίδιο ακριβώς υποσύνολο damaged εικόνων που χρησιμοποιήθηκε για το segmentation ως classifier.

5.5 Περιορισμοί της Μελέτης

Η παρούσα εργασία παρουσιάζει τέσσερις κύριους περιορισμούς. Πρώτον, το επίπεδο παρουσίας της σοβαρότητας (ενότητα 3.10) είναι αμιγώς κανονιστικό (rule-based), με τιμές ορίων και στάθμιση ανά τύπο δομικού στοιχείου που ορίστηκαν ευρετικά και δεν έχουν επικυρωθεί έναντι πραγματικών ετικετών σοβαρότητας από ειδικό μηχανικό, καθώς τέτοιες ετικέτες δεν είναι διαθέσιμες στο φ-Net. Δεύτερον, η αξιολόγηση του ενοποιημένου pipeline (ενότητα 4.5) παραμένει ποιοτική, χωρίς συστηματική ποσοτική κατανομή Green/Yellow/Red σε ολόκληρο το σύνολο ελέγχου. Τρίτον, η αρχιτεκτονική δύο μοντέλων συνεπάγεται διάδοση σφάλματος (error propagation): ένα λανθασμένο component ή damage_type από το classification track μεταφέρεται απευθείας στο επίπεδο σύντηξης, χωρίς μηχανισμό διόρθωσης. Τέταρτον, δοκιμή Test-Time Augmentation για το segmentation track δεν κατέστη δυνατή λόγω περιορισμού υποστήριξης της συγκεκριμένης έκδοσης της βιβλιοθήκης Ultralytics (ενότητα 4.3). Τέλος, όπως αναδεικνύει το παράδειγμα 00026 (ενότητα 4.5), όταν το segmentation track δεν εντοπίζει καμία instance σε μια πραγματικά ρηγματωμένη εικόνα, η κάλυψη υπολογίζεται ως 0% και εμπίπτει από default στη ζώνη Minimal, με αποτέλεσμα πιθανή υποεκτίμηση της σοβαρότητας. Το pipeline δεν διαθέτει επί του παρόντος μηχανισμό διάκρισης ανάμεσα σε «πραγματική απουσία βλάβης στη μάσκα» και «αποτυχία ανίχνευσης».

Κεφάλαιο 6. Συμπεράσματα

6.1 Σύνοψη της Εργασίας

Η παρούσα εργασία ανέπτυξε ένα σύστημα αυτόματης ανίχνευσης και χαρακτηρισμού δομικής βλάβης σε δομικά στοιχεία από οπλισμένο σκυρόδεμα, βασισμένο σε δύο παράλληλα σκέλη: ένα πολυκαθικοντικό μοντέλο ταξινόμησης (MultiTaskYOLO), το οποίο αξιοποιεί backbone YOLO11m-cls με μηχανισμό προσοχής C2PSA για την ταυτόχρονη πρόβλεψη τεσσάρων χαρακτηριστικών (κατάσταση βλάβης, αποφλοίωση, τύπος δομικού στοιχείου, τύπος βλάβης) πάνω στο σύνολο δεδομένων φ-Net, και ένα συμπληρωματικό σκέλος εντοπισμού και μέτρησης ρωγμών, μέσω fine-tuning ενός μοντέλου YOLO11x-seg σε εξειδικευμένο σύνολο δεδομένων κατάτμησης τριών κλάσεων (Shear, Flexural, Combined).

6.2 Κύρια Ευρήματα

- Το μοντέλο ταξινόμησης πέτυχε μέση ακρίβεια (macro average) 76,85% στα τέσσερα tasks, με το damage_state να παρουσιάζει την υψηλότερη επίδοση (85,75%) και το damage_type τη χαμηλότερη (65,65%).
- Στο task damage_type, το μοντέλο υπερτερεί σημαντικά έναντι προσέγγισης βασισμένης σε μεγάλο γλωσσικό μοντέλο όρασης πάνω στο ίδιο σύνολο δεδομένων (65,65% έναντι 34,6% [13]), αναδεικνύοντας το πλεονέκτημα μιας εξειδικευμένης, συμπαγούς αρχιτεκτονικής CNN έναντι γενικού σκοπού μοντέλων μεγάλης κλίμακας σε στενά ορισμένα προβλήματα ταξινόμησης.
- Ο κανόνας HEA-style gating, παρότι λογικά αιτιολογημένος, δεν βελτίωσε τις καθαρές μετρικές ταξινόμησης του damage_type, καταδεικνύοντας ότι η επιβολή λογικών περιορισμών μεταξύ tasks σε ένα multi-task πλαίσιο δεν αποδίδει αυτόματα όφελος όταν το gate task δεν είναι απόλυτα αξιόπιστο.
- Αναπτύχθηκε και επιδείχθηκε ποιοτικά ένα ενοποιημένο two-model pipeline με επίπεδο σύντηξης σοβαρότητας (severity fusion), το οποίο συνδυάζει τις εξόδους ταξινόμησης και κατάτμησης σε μία ενιαία ετικέτα Green/Yellow/Red. Η επιβεβαίωση της ορθής ενεργοποίησης τόσο του βασικού πίνακα κανόνων όσο και των κανόνων παράκαμψης (spalling override, forced Green, Extensive→Red) σε πραγματικές εικόνες αποτελεί ένδειξη ορθής λειτουργίας, εκκρεμεί όμως η συστηματική ποσοτική αξιολόγηση σε ολόκληρο το σύνολο ελέγχου.

- Το segmentation track πέτυχε Box mAP50=0,338 και Mask mAP50=0,250 στο πραγματικό σύνολο ελέγχου, με κύριο περιορισμό τη χαμηλή ανάκληση μάσκας (28,6%). Η σάρωση confidence threshold δεν βελτίωσε περαιτέρω το recall. Η αξιολόγηση του ίδιου μοντέλου ως ταξινομητή σε επίπεδο εικόνας στο πραγματικό σύνολο ελέγχου πέτυχε Accuracy = 68,78% στη διάκριση Shear/Flexural/Combined.

6.3 Συνεισφορά της Εργασίας

Η κύρια συνεισφορά της εργασίας έγκειται στην ενοποίηση τεσσάρων καθηκόντων ταξινόμησης σε ένα ενιαίο multi-task μοντέλο με κοινό backbone, αντί της μεμονωμένης εκπαίδευσης ξεχωριστών μοντέλων ανά task, όπως συνηθίζεται στη σχετική βιβλιογραφία. Τέλος, η εργασία συνεισφέρει με τη συστηματική τεκμηρίωση της μηχανικής σημασίας των υπό μελέτη χαρακτηριστικών (Κεφάλαιο 3) και με την ανάδειξη ευρημάτων, που θέτουν σε προβληματισμό απλουστευτικές παραδοχές σχεδίασης πλαισίων fusion σε πολυκαθηκοντικά προβλήματα δομικής βλάβης.

6.4 Μελλοντικές Κατευθύνσεις

- Εξωτερική επικύρωση του μοντέλου ταξινόμησης σε ανεξάρτητο σύνολο δεδομένων από διαφορετικό σεισμικό γεγονός, για αξιόπιστη εκτίμηση της ικανότητας γενίκευσής του.
- Επέκταση και εμπλουτισμός του συνόλου δεδομένων κατάτμησης, με στόχο τη βελτίωση της ανάκλησης του segmentation track, ιδίως για την κλάση Flexural.
- Συστηματική ποσοτική αξιολόγηση του επιπέδου σύντηξης σοβαρότητας σε ολόκληρο το σύνολο ελέγχου, με καταγραφή της κατανομής Green/Yellow/Red και συχνότητας ενεργοποίησης κάθε κανόνα παράκαμψης, καθώς και επικύρωση των ορίων και της στάθμισης του πίνακα σύντηξης έναντι πραγματικών αξιολογήσεων ειδικού μηχανικού.
- Διερεύνηση εναλλακτικών τεχνικών βελτίωσης recall στο segmentation track πέραν των όσων δοκιμάστηκαν χωρίς ουσιαστικό όφελος, καθώς και επανάληψη της ανά κλάση αξιολόγησης σε μεγαλύτερο test set, ώστε να μειωθεί η στατιστική αβεβαιότητα ειδικά για την κλάση Flexural (μόλις 30 εικόνες ελέγχου).
- Διερεύνηση μαθησιακής (learned), αντί αμιγώς κανονιστικής (rule-based), προσέγγισης fusion μεταξύ των εξόδων ταξινόμησης και κατάτμησης, κατά το

πρότυπο μετα-ταξινομητών όπως το LightGBM που έχουν χρησιμοποιηθεί σε συναφείς εργασίες [11].

Αναφορές

- [1] Gao, Y., Mosalam, K.M. (2018). Deep transfer learning for image-based structural damage recognition. *Computer-Aided Civil and Infrastructure Engineering*, 33(9), 748-768.
- [2] Gao, Y., Mosalam, K.M. (2020). PEER Hub ImageNet: A large-scale multi-attribute benchmark data set of structural images. *Journal of Structural Engineering*, 146(10), 04020198.
- [3] Krizhevsky, A., Sutskever, I., Hinton, G.E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84-90.
- [4] Simonyan, K., Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- [5] He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770-778.
- [6] Jocher, G., Chaurasia, A., Qiu, J. (2023). Ultralytics YOLO (Version 8.0) [Software]. <https://github.com/ultralytics/ultralytics>.
- [7] Khanam, R., Hussain, M. (2024). YOLOv11: An overview of the key architectural enhancements. *arXiv preprint arXiv:2410.17725*.
- [8] Dondi, P., Gullotti, A., Inchingolo, M., Senaldi, I., Casarotti, C., Lombardi, L., Piastra, M. (2025). Post-earthquake structural damage detection with tunable semi-synthetic image generation. *Engineering Applications of Artificial Intelligence*, 147, 110302.
- [9] Mallepalli, A., Tuthill, L., Suthar, D., Sharma, A. (2025). Real-World Crack and Spalling Detection in Reinforced Concrete Structures via Dual YOLOv11 Segmentation. *SSRN preprint (μη κριτικά αξιολογημένο)*. <https://ssrn.com/abstract=5351666>.
- [10] Huang, S., Liu, Q., Chen, C., Chen, Y. (2025). A Real-time Concrete Crack Detection and Segmentation Model Based on YOLOv11. *arXiv preprint arXiv:2508.11517 (μη κριτικά αξιολογημένο)*.
- [11] Türier, A., Bai, Y., Sezen, H., Yilmaz, A. (2025). Deep learning-based automated damage detection in concrete structures using images from earthquake events. *The 2025 World Congress on Advances in Structural Engineering and Mechanics (ASEM25)*, Busan, Korea.
- [12] Türier, A., Bai, Y., Sezen, H., Yilmaz, A. (2026). Deep learning-based automated damage detection in concrete structures using images from earthquake events (εκτεταμένη έκδοση). *Journal of Infrastructure Intelligence and Resilience*, 5, 100208.
- [13] Jiang et al. (2025). Large language model for post-earthquake structural damage assessment. *Computer-Aided Civil and Infrastructure Engineering*, 40(31), 6325-6338.

- [14] Vijayakumar Kuttuva, G., Prawin, J. (2026). Structural defect segmentation using a semi-supervised algorithm integrating YOLO and the segment anything model. *Automation in Construction*, 182, 106709.
- [15] Cabral, R., Santos, R., Correia, J.A.F.O., Ribeiro, D. (2025). A Hybrid YOLO and Segment Anything Model Pipeline for Multi-Damage Segmentation in UAV Inspection Imagery. *Sensors*, 25(21), 6568.
- [16] Raushan, R., Singhal, V., Jha, R.K. (2025). Damage detection in concrete structures with multi-feature backgrounds using the YOLO network family. *Automation in Construction*, 170, 105887.
- [17] Hoskere, V., Narazaki, Y., Hoang, T., Spencer, B. (2020). MaDnet: multi-task semantic segmentation of multiple types of structural materials and damage in images of civil infrastructure. *Journal of Civil Structural Health Monitoring*, 10(5), 757-773.
- [18] Davoudi, R., Miller, G.R., Kutz, J.N. (2018). Structural load estimation using machine vision and surface crack patterns for shear-critical RC beams and slabs. *Journal of Computing in Civil Engineering*, 32(4), 04018024.
- [19] Βουθούνης, Π.Α. (2019). Αντοχή των Υλικών – Μηχανική του Παραμορφώσιμου Στερεού (4η έκδ.). Εκδόσεις Α. Βουθούνη, Αθήνα. ISBN 978-618-83280-3-7.
- [20] Καραγιάννης, Χ.Γ. (2019). Σχεδιασμός Κατασκευών από Ωπλισμένο Σκυρόδεμα και Σεισμικές Δράσεις. Εκδόσεις σοφία, Θεσσαλονίκη. ISBN 978-960-633-005-6.
- [21] Lin, T.-Y., Goyal, P., Girshick, R., He, K., Dollár, P. (2017). Focal loss for dense object detection. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2980-2988.