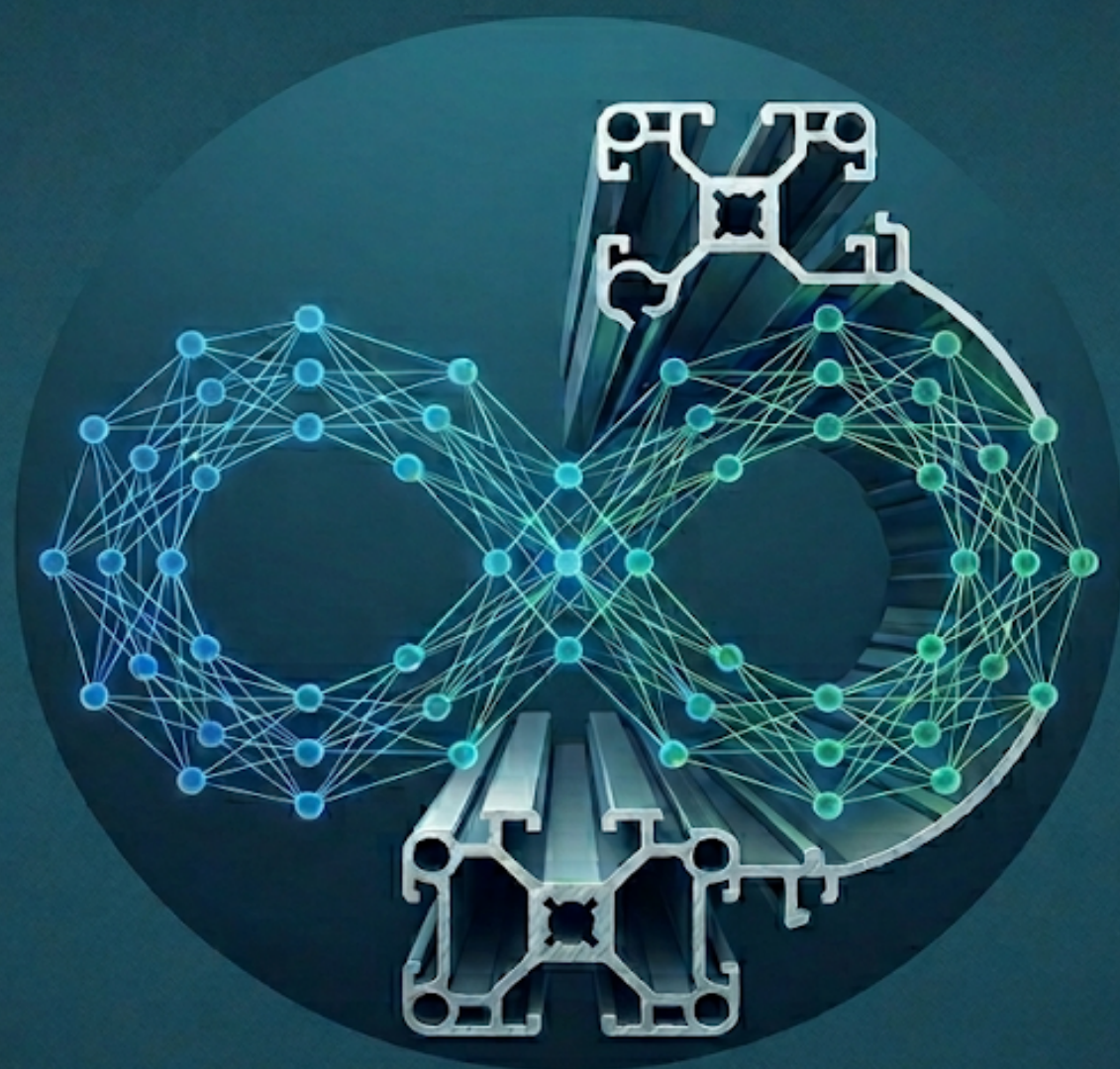


ΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΕΡΕΥΝΑ ΣΥΣΤΗΜΑΤΩΝ ΒΑΘΙΑΣ ΜΑΘΗΣΗΣ ΓΙΑ ΤΗΝ ΑΝΑΓΝΩΡΙΣΗ ΠΟΛΥΠΛΟΚΩΝ ΓΕΩΜΕΤΡΙΚΩΝ ΣΧΗΜΑΤΩΝ ΚΑΙ ΕΦΑΡΜΟΓΗ ΣΤΗ ΒΙΟΜΗΧΑΝΙΑ ΠΡΟΦΙΛ ΑΛΟΥΜΙΝΙΟΥ



ΚΑΦΕΣΤΙΔΗΣ ΣΤΥΛΙΑΝΟΣ

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ:
ΑΝΑΓΝΩΣΤΟΠΟΥΛΟΣ ΧΡΗΣΤΟΣ

“ Όταν η πληροφορία θα βρίσκεται παντού και θα είναι άμεσα διαθέσιμη μέσω έξυπνων συστημάτων, τότε η αξία δεν θα βρίσκεται στην απομνημόνευση αλλά στη φαντασία, στην κριτική σκέψη, στη σύνθεση ιδεών και στην ικανότητα να θέτεις τα σωστά ερωτήματα. Η δημιουργικότητα, η έμπνευση και η ανθρώπινη ευαισθησία θα γίνουν τα πραγματικά «εργαλεία επιβίωσης». “

Δημήτρης Νανόπουλος

*Ομότιμος καθηγητής Φυσικής Υψηλών Ενεργειών του Πανεπιστημίου Texas A&M,
τακτικό μέλος Ακαδημίας Αθηνών*

ΕΡΕΥΝΑ ΣΥΣΤΗΜΑΤΩΝ ΒΑΘΙΑΣ ΜΑΘΗΣΗΣ ΓΙΑ ΤΗΝ ΑΝΑΓΝΩΡΙΣΗ
ΠΟΛΥΠΛΟΚΩΝ ΓΕΩΜΕΤΡΙΚΩΝ ΣΧΗΜΑΤΩΝ
ΚΑΙ ΕΦΑΡΜΟΓΗ ΣΤΗ ΒΙΟΜΗΧΑΝΙΑ ΠΡΟΦΙΛ ΑΛΟΥΜΙΝΙΟΥ

«Θα ήθελα να εκφράσω τις θερμότερες ευχαριστίες μου στη σύζυγό μου, Ειρήνη και στον γιο μου Αλέξανδρο, καθώς και στην υπόλοιπη οικογένειά μου, για την αμέριστη συμπαράσταση και την πολύτιμη στήριξή τους καθ' όλη τη διάρκεια αυτού του ταξιδιού.

Επιπροσθέτως, οφείλω θερμές ευχαριστίες στον επιβλέποντα καθηγητή μου, κ. Χρήστο Αναγνωστόπουλο, καθώς και σε όλους τους συναδέλφους –ακαδημαϊκούς και μη– για την αστείρευτη υπομονή και την προθυμία τους να ανταποκριθούν στο πλήθος των ερωτημάτων μου.»

Στυλιανός Καφεστίδης

Περίληψη Διατριβής: Έρευνα Συστημάτων Βαθιάς Μάθησης για την Αναγνώριση Πολύπλοκων Γεωμετρικών Σχημάτων και Εφαρμογή στη Βιομηχανία Προφίλ Αλουμινίου

Εισαγωγή και Πρόβλημα

Η παρούσα εργασία αντιμετωπίζει την κρίσιμη ανάγκη για αυτοματοποιημένη και αξιόπιστη αναγνώριση πολύπλοκων γεωμετρικών σχημάτων διατομής στην παραγωγή προφίλ αλουμινίου, έναν τομέα που χαρακτηρίζεται από τεράστιο όγκο μοναδικών προϊόντων και υψηλή ευαισθησία σε σφάλματα χειροκίνητης διαχείρισης. Κεντρικός στόχος της έρευνας ήταν η ανάπτυξη και η αξιολόγηση μιας λύσης βασισμένης στην Τεχνητή Νοημοσύνη για την αναγνώριση αυτών των σχημάτων σε πραγματικό, «θορυβώδες» βιομηχανικό περιβάλλον, με έμφαση στη γεφύρωση του θεμελιώδους Χάσματος Προσομοίωσης-Πραγματικότητας (Sim-to-Real Gap).

Μεθοδολογία και Αρχιτεκτονική Συστήματος

Η μεθοδολογία υιοθέτησε μια διαδοχική αρχιτεκτονική (cascaded pipeline), συνδυάζοντας έναν Ανιχνευτή Αντικειμένων (Object Detector, YOLOv8n) για τον χωρικό εντοπισμό με έναν Ταξινομητή (Classifier, EfficientNetV2 ή MobileViT) για την ακριβή γεωμετρική ταυτοποίηση.

- **Μηχανική Συνθετικών Δεδομένων:** Για να ξεπεραστεί η σπανιότητα των πραγματικών δεδομένων, αναπτύχθηκε η αρχιτεκτονική "DataSilver", ένας αυτοματοποιημένος αγωγός που παρήγαγε 116.200 συνθετικές εικόνες από αρχεία CAD. Ο αγωγός αυτός ενσωμάτωσε τεχνικές όπως η προσομοίωση κατοπτρικής απόκρυψης, η έγχυση βαρέος θορύβου και το pipeline "Super-Grayscale" CLAHE, με σκοπό την υπολογιστική αναπαραγωγή των ατελειών του οπτικού αισθητήρα και την ανοσοποίηση του δικτύου έναντι της Προκατάληψης Υφής (Texture Bias).
- **Υλοποίηση Edge AI:** Η λύση ενσωματώθηκε πρακτικά στο «Κινητό Εργαστήριο» (Signal Profiler), μια διαπλατφορμική εφαρμογή Flutter/Dart, σχεδιασμένη για τη συλλογή δεδομένων και την εκτέλεση συμπερασμού (inference) σε πραγματικό χρόνο εντός κινητής συσκευής (TFLite).

Αποτελέσματα και Αξιολόγηση

Τα μοντέλα εκπαιδεύτηκαν με επιθετική επαύξηση δεδομένων, με το YOLOv8n να επιτυγχάνει σχεδόν τέλεια ακρίβεια (mAP@50-95: 0.995) στη συνθετική επικύρωση.

- **Αρχική Πρόκληση:** Η αρχική ανάπτυξη σε πραγματικές εικόνες αποκάλυψε σοβαρό Κενό Τομέα, με τους Ταξινομητές να αδυνατούν να γενικεύσουν σε νέα υλικά (π.χ., λευκή βαφή).
- **Προσαρμογή Πεδίου:** Το πρόβλημα αντιμετωπίστηκε με τη στρατηγική Εκμάθησης Μεταφοράς (Transfer Learning), όπου εφαρμόστηκε Αντικατάσταση του Τελικού Επιπέδου Ταξινόμησης και Μικρο-συντονισμός (Fine-Tuning) σε ένα μικρό σύνολο ημι-συνθετικών εικόνων.
- **Τελική Απόδοση:** Μετά την προσαρμογή, ο βέλτιστος συνδυασμός (YOLOv8n + EfficientNetV2) πέτυχε Συνολική Ακρίβεια Συστήματος 76.19% σε μη ιδωμένα πραγματικά δείγματα. Αυτός ο συνδυασμός επιβεβαιώθηκε ως η βέλτιστη ισορροπία ταχύτητας-ακρίβειας, με χρόνο συμπερασμού 717 ms ανά εικόνα, καθιστώντας το σύστημα πλήρως βιώσιμο για πρακτική εφαρμογή "Point-and-Shoot".

Συμπέρασμα

Η διατριβή επιβεβαιώνει την αποτελεσματικότητα της μεθοδολογίας "Sim-to-Real" όταν συνδυάζεται με προηγμένη Μηχανική Συνθετικών Δεδομένων και Εκμάθηση Μεταφοράς. Η προσέγγιση αυτή προσφέρει μια κλιμακώσιμη λύση για την ενσωμάτωση της Μηχανικής Όρασης στο πλαίσιο της Βιομηχανίας 4.0, μειώνοντας την εξάρτηση από την εκτεταμένη συλλογή και σχολιασμό πραγματικών δεδομένων. Μελλοντικές κατευθύνσεις περιλαμβάνουν την επέκταση για ανίχνευση ελαττωμάτων και τη διασύνδεση με συστήματα ERP, όπως και μια θεωρητική αρχιτεκτονική για την επικοινωνία και επανεκπαίδευση των μοντέλων.

Thesis Abstract: Research on Deep Learning Systems for Complex Geometric Shape Recognition and Application in the Aluminum Profile Industry

Introduction and Problem

The present work addresses the critical need for automated and reliable recognition of complex cross-sectional geometric shapes in aluminum profile manufacturing, a sector characterized by a vast volume of unique products and high susceptibility to manual handling errors. The primary objective of this research was to develop and evaluate an Artificial Intelligence-based solution for recognizing these shapes in a real-world, "noisy" industrial environment, with an emphasis on bridging the fundamental Sim-to-Real Gap.

Methodology and System Architecture

The methodology adopted a cascaded pipeline architecture, combining an Object Detector (YOLOv8n) for spatial localization with a Classifier (EfficientNetV2 or MobileViT) for precise geometric identification.

- **Synthetic Data Engineering:** To overcome the scarcity of real-world data, the "DataSilver" architecture was developed, an automated pipeline that generated 116,200 synthetic images from CAD files. This pipeline incorporated techniques such as specular occlusion simulation, heavy noise injection, and the "Super-Grayscale" CLAHE pipeline, aiming to computationally replicate optical sensor imperfections and immunize the network against Texture Bias.
- **Edge AI Implementation:** The solution was practically integrated into the "Mobile Laboratory" (Signal Profiler), a cross-platform Flutter/Dart application designed for data collection and real-time on-device inference using TFLite.

Results and Evaluation

The models were trained using aggressive data augmentation, with YOLOv8n achieving near-perfect accuracy (mAP@50-95: 0.995) on synthetic validation.

- **Initial Challenge:** Initial deployment on real images revealed a severe Domain Gap, with the Classifiers failing to generalize to new materials (e.g., white paint).
- **Domain Adaptation:** This issue was addressed through a Transfer Learning strategy, utilizing Classification Head Replacement and Fine-Tuning on a small set of semi-synthetic images.
- **Final Performance:** Post-adaptation, the optimal combination (YOLOv8n + EfficientNetV2) achieved a Total System Accuracy of 76.19% on unseen real-world samples. This combination was confirmed as the optimal balance between speed and accuracy, with an inference time of 717 ms per image, rendering the system fully viable for practical "Point-and-Shoot" application.

Conclusion

The thesis confirms the effectiveness of the "Sim-to-Real" methodology when combined with advanced Synthetic Data Engineering and Transfer Learning. This approach offers a scalable solution for integrating Machine Vision within the Industry 4.0 framework, reducing the reliance on extensive collection and annotation of real-world data. Future directions include expanding the system for defect detection and interfacing with ERP systems, as well as a theoretical architecture for the communication and retraining of the distributed models.

Περιεχόμενα

Κεφάλαιο 1: Εισαγωγή στην Αυτοματοποιημένη Αναγνώριση Σχημάτων	
1.1 Το Βιομηχανικό Πλαίσιο και η Ανάγκη για Αυτοματοποίηση.....	7
1.2 Παρουσίαση του Προβλήματος: Η Περίπτωση των Προφίλ Αλουμινίου.....	7
1.3 Στόχοι της Εργασίας.....	9
1.4 Σχετική έρευνα.....	9
Κεφάλαιο 2: Τεχνητά Νευρωνικά δίκτυα: Από τον αισθητήρα μέχρι τα συνελκτικά ΤΝΔ	
2.1 Τα πρώτα χρόνια των ΤΝΔ.....	18
2.2 Συνελκτικά Νευρωνικά Δίκτυα και σύγχρονες εξελίξεις.....	21
Κεφάλαιο 3 Μεθοδολογία	
3.1 Μεθοδολογία Σύνθεσης Δεδομένων.....	25
3.2 Εξαγωγή Διπλού Σχήματος: Ανιχνευτής και Ταξινομητής.....	27
3.3 Αρχιτεκτονικές Μοντέλων Νευρωνικών Δικτύων.....	37
3.3 Στρατηγική Εκπαίδευσης και Επιλογή Υπερπαραμέτρων.....	33
3.4. Σχεδιασμός και την ανάπτυξη του “Κινητού Έργαστηριου(“Signal Profiler”/Flutter-Dart)	41
3.5 Προκλήσεις Υλοποίησης και Εξέλιξη του Συστήματος.....	48
Κεφάλαιο 4: Αξιολόγηση Συστήματος σε Πραγματικές Συνθήκες	
4.1 Μεθοδολογία Αξιολόγησης και Μαθηματικός Ορισμός Μετρικών.....	52
4.2. Αποτελέσματα Χωρικού Εντοπισμού (Detectors Only).....	53
4.3. Αποτυχία Ταξινόμησης Λόγω Προκατάληψης Υφής (Texture Bias).....	54
4.4 Αποτυχία Ταξινόμησης Λόγω Προκατάληψης Υφής (Texture Bias).....	54
4.5 Αποτελέσματα Προσαρμογής (Fine-Tuning) και Τελική Αξιολόγηση.....	56
4.6 Τελικό Συμπέρασμα Μεθοδολογίας.....	57
4.7 Αξιολόγηση Υπολογιστικού Κόστους και Χρόνου Συμπερασμού.....	58
4.8 Πρακτική Εφαρμοσιμότητα (Deployment Viability).....	60
Κεφάλαιο 5: Συμπεράσματα και Μελλοντική Έρευνα	
5.1 Συνολική Αποτίμηση.....	61
5.2 Αξιοπιστία και Προστιθέμενη Αξία.....	62
5.3 Περιορισμοί του Συστήματος.....	62
5.4 Μελλοντικές Κατευθύνσεις.....	63
5.5 Επίλογος.....	67
Παράρτημα Α: Τεχνικά χαρακτηριστικά.....	68
Παράρτημα Β: Υλοποίηση (Github Link).....	68
Βιβλιογραφία.....	70

Κεφάλαιο 1: Εισαγωγή στην Αυτοματοποιημένη Αναγνώριση Σχημάτων

1.1 Το Βιομηχανικό Πλαίσιο και η Ανάγκη για Αυτοματοποίηση

Η βιομηχανία αποτελεί ένα ξεχωριστό οικοσύστημα με τεχνολογίες αιχμής, αλλά και ιδιαίτερες απαιτήσεις, όπως υψηλή ακρίβεια και ταχύτητα απόκρισης των εφαρμογών, κοστολογικά διαβαθμισμένες λύσεις, και όλα αυτά σε ένα περιβάλλον απαιτητικό και συχνά μη ελεγχόμενο. Η αυτοματοποιημένη αναγνώριση πολύπλοκων γεωμετρικών σχημάτων, συμπεριλαμβανομένων των πολυγωνικών δομών και των συνδυασμών βασικών μορφών, αποτελεί κρίσιμη απαίτηση για τις σύγχρονες βιομηχανικές εφαρμογές. Αυτή η ικανότητα μπορεί να ενισχύσει σημαντικά διαδικασίες όπως η αυτοματοποιημένη επιθεώρηση ποιότητας, η καθοδήγηση ρομποτικής συναρμολόγησης, η ανίχνευση ελαττωμάτων και η διαχείριση αποθεμάτων, οδηγώντας σε βελτιωμένη αποδοτικότητα, μειωμένα σφάλματα και εξοικονόμηση κόστους. Σε ένα βιομηχανικό πλαίσιο, η ικανότητα ταχείας και ακριβούς αναγνώρισης συγκεκριμένων γεωμετρικών σχημάτων είναι υψίστης σημασίας. Ένα τέτοιο σύστημα θα μπορούσε, για παράδειγμα, να επαληθεύει ότι το σωστό εξάρτημα χρησιμοποιείται στο σωστό στάδιο της συναρμολόγησης, να διασφαλίζει τον σωστό προσανατολισμό των μερών ή να ανιχνεύει αποκλίσεις από τυποποιημένες μορφές. Άλλωστε η μετάβαση προς τη Βιομηχανία 4.0, η οποία δίνει έμφαση στην αυτοματοποίηση και την ανταλλαγή δεδομένων, καθιστά την ανάπτυξη τέτοιων ευφυών συστημάτων όχι απλώς επιθυμητή, αλλά αναγκαία για τη διατήρηση της ανταγωνιστικότητας.

1.2 Παρουσίαση του Προβλήματος: Η Περίπτωση των Προφίλ Αλουμινίου

Η παρούσα πτυχιακή εργασία επικεντρώνεται στην αναγνώριση πολύπλοκων γεωμετρικών σχημάτων στον βιομηχανικό τομέα της παραγωγής προφίλ αλουμινίου, τα οποία αποτελούν συστήματα ανοιγόμενων και συρόμενων κουφωμάτων, είτε για οικιακή είτε για βιομηχανική χρήση. Οι απαιτήσεις της αγοράς σε θερμική και ηχητική μόνωση, αντιδιαρρηκτική ασφάλεια, πυρασφάλεια, ειδικές κατασκευές και κόστος, έχουν ωθήσει τη βιομηχανία στον σχεδιασμό εκατοντάδων διαφορετικών προφίλ, το καθένα με μια μοναδική και συχνά πολύπλοκη γεωμετρία διατομής. Αυτά τα σύγχρονα συστήματα προφίλ αλουμινίου και PVC αποτελούνται από διάφορων ειδών γεωμετρικά σχήματα, με αποτέλεσμα η εμπειρική διαχείρισή τους να αποτελεί πρόβλημα όχι μόνο για τις εταιρείες παραγωγής, αλλά και για τους εμπορικούς αντιπροσώπους και τους κατασκευαστές κουφωμάτων. Η λανθασμένη αναγνώριση ενός προφίλ μπορεί να οδηγήσει σε λάθη στη συναρμολόγηση, σπατάλη υλικών και καθυστερήσεις στην παραγωγή. Συνεπώς, η ανάπτυξη ενός αυτοματοποιημένου συστήματος για την ταχεία και αξιόπιστη αναγνώριση της γεωμετρίας των προφίλ είναι κρίσιμης σημασίας.



Παράδειγμα σύγχρονου συστήματος προφίλ αλουμινίου και PVC

1.3 Στόχοι της Εργασίας

Στόχος της παρούσας εργασίας αποτελεί την ανάπτυξη, υλοποίηση και αξιολόγηση μιας ολοκληρωμένης λύσης βασισμένης στην τεχνητή νοημοσύνη για την αυτοματοποιημένη αναγνώριση γεωμετρικών σχημάτων προφίλ αλουμινίου σε πραγματικό, «θορυβώδες» βιομηχανικό περιβάλλον. Κεντρικός πυρήνας της έρευνας είναι η γεφύρωση του «χάσματος της πραγματικότητας» (reality gap) — δηλαδή των εγγενών διαφορών μεταξύ των ιδανικών συνθηκών μιας προσομοίωσης και των αστάθμητων παραγόντων του εργοστασιακού χώρου.

Για την επίτευξη του σκοπού αυτού, η διαχείριση του προβλήματος μεταφέρεται σε μια κινητή συσκευή, η οποία λειτουργεί ως «κινητό εργαστήριο» συλλογής και επεξεργασίας. Οι επιμέρους ερευνητικοί και πρακτικοί στόχοι της εργασίας εξειδικεύονται ως εξής:

α) Ανάπτυξη Κινητής Εφαρμογής (Flutter/Dart): Σχεδιασμός και υλοποίηση μιας εφαρμογής για φορητές συσκευές με σκοπό τη συλλογή (φωτογράφιση) και την άμεση κατηγοριοποίηση των προφίλ σε πραγματικό χρόνο, δημιουργώντας δομημένα σύνολα δεδομένων εκπαίδευσης και επικύρωσης.

β) Cloud Υποδομή Διαχείρισης Δεδομένων (Google Cloud Storage): Ασφαλής αποθήκευση και συγχρονισμός των συλλεχθέντων δεδομένων σε μια απομακρυσμένη βάση δεδομένων στο

cloud, προσομοιώνοντας τις διαδικασίες διαχείρισης δεδομένων μεγάλου όγκου των σύγχρονων βιομηχανιών.

γ) Σχεδιασμός και Βελτιστοποίηση Μοντέλων (TensorFlow/Python): Κατασκευή, εκπαίδευση και αξιολόγηση αρχιτεκτονικών νευρωνικών δικτύων κατάλληλων για την αναγνώριση των σύνθετων διατομών, και η μετέπειτα μετατροπή τους σε ελαφριές εκδόσεις (Tensorflow Lite) για τη διασφάλιση της συμβατότητάς τους με το περιορισμένων πόρων περιβάλλον Android.

δ) Εγκατάσταση και Δοκιμές Πεδίου (Field Testing): Ενσωμάτωση του μοντέλου Lite στην κινητή εφαρμογή και εκτέλεση δοκιμών σε πραγματικές συνθήκες εργοστασίου (μεταβαλλόμενος φωτισμός, ξένων σωμάτων και αποκρύψεων) με σκοπό την ποσοτική και ποιοτική αξιολόγηση των αποτελεσμάτων.

1.4 Σχετική έρευνα

Η ανάπτυξη ενός συστήματος μηχανικής όρασης σε ένα εργαστηριακό περιβάλλον είναι σημαντικά διαφορετική από την ανάπτυξή του σε μια πραγματική βιομηχανική γραμμή παραγωγής. Τα βιομηχανικά περιβάλλοντα είναι εκ φύσεως «εχθρικά» για τα συστήματα όρασης, παρουσιάζοντας μια σειρά από προκλήσεις που μπορούν να υποβαθμίσουν δραματικά την απόδοση και την αξιοπιστία τους. Η επιτυχής υλοποίηση απαιτεί την κατανόηση αυτών των προκλήσεων και την εφαρμογή συνδυαστικών λύσεων, τόσο σε επίπεδο υλικού όσο και σε επίπεδο αλγορίθμων.

1.4.1 Μεταβλητός Φωτισμός (Variable Lighting)

Ο φωτισμός είναι ίσως ο πιο κρίσιμος παράγοντας για την επιτυχία μιας εφαρμογής μηχανικής όρασης. Ο ανεξέλεγκτος ή ακατάλληλος φωτισμός μπορεί να καταστήσει ακόμη και τον πιο προηγμένο αλγόριθμο αναποτελεσματικό. Το πρόβλημα εκδηλώνεται με διάφορους τρόπους, υπερβολικά έντονος φωτισμός μπορεί να προκαλέσει κορεσμό στους αισθητήρες της κάμερας, ενώ ανεπαρκής φωτισμός οδηγεί σε εικόνες με χαμηλή αντίθεση και αυξημένο θόρυβο. Ο ανομοιόμορφος φωτισμός δημιουργεί σκιές που μπορούν να αλλοιώσουν το αντιληπτό σχήμα ενός προφίλ. Η πρόκληση αυτή είναι ιδιαίτερα έντονη όταν πρόκειται για την επιθεώρηση μεταλλικών, ανακλαστικών επιφανειών, όπως αυτές του αλουμινίου. Σε αυτές τις περιπτώσεις, οι άμεσες πηγές φωτός δημιουργούν έντονη αντανάκλαση (glare) και λαμπερά σημεία (hot spots) που μπορούν να «τυφλώσουν» την κάμερα και να αποκρύψουν τις κρίσιμες ακμές που ορίζουν τη γεωμετρία του προφίλ.

Για την αντιμετώπιση αυτών των προβλημάτων, εφαρμόζονται δύο κατηγορίες λύσεων. Η πρώτη αφορά τον ενεργητικό έλεγχο του φωτισμού μέσω εξειδικευμένου υλικού. Για τις ανακλαστικές επιφάνειες, χρησιμοποιούνται τεχνικές όπως ο θολωτός φωτισμός (dome lighting), όπου το φως διαχέεται ομοιόμορφα από έναν ημισφαιρικό θόλο, ή ο ομοαξονικός φωτισμός (coaxial lighting), όπου το φως προβάλλεται παράλληλα με τον οπτικό άξονα της κάμερας. Και οι δύο τεχνικές ελαχιστοποιούν τις άμεσες αντανακλάσεις και παρέχουν ομοιόμορφη φωτεινότητα, επιτρέποντας την καθαρή αποτύπωση του σχήματος.

Η δεύτερη κατηγορία αφορά αλγοριθμικές τεχνικές. Αυτές περιλαμβάνουν μεθόδους προεπεξεργασίας της εικόνας για τη διόρθωση της φωτεινότητας και της αντίθεσης, καθώς και την εκπαίδευση των μοντέλων βαθιάς μάθησης με δεδομένα που έχουν υποστεί επαύξηση (data augmentation). Μέσω της επαύξησης, το μοντέλο εκτίθεται σε εικόνες με τεχνητά μεταβαλλόμενη φωτεινότητα, αντίθεση και έκθεση, μαθαίνοντας έτσι να είναι πιο ανθεκτικό και να γενικεύει καλύτερα στις πραγματικές, μεταβαλλόμενες συνθήκες του εργοστασίου. Στην παρούσα εργασία επικεντρωνόμαστε στην δεύτερη κατηγορία λύσεων.

1.3.2 Απόκρυψη Αντικειμένων (Object Occlusion)

Για την αντιμετώπιση της απόκρυψης (occlusion) και των αλληλοεπικαλύψεων (overlaps) που συναντώνται εγγενώς στα βιομηχανικά περιβάλλοντα (π.χ. στοιβαγμένα προφίλ σε παλέτες ή κάδους), η σύγχρονη έρευνα ξεφεύγει από την απλή ανίχνευση πλαισίου οριοθέτησης (bounding box). Αντιθέτως, προκρίνει μοντέλα κατάτμησης σε επίπεδο αντικειμένου (instance segmentation), με χαρακτηριστικότερο παράδειγμα την οικογένεια αλγορίθμων Mask R-CNN. Τα μοντέλα αυτά εξελίσσονται ραγδαία, εστιάζοντας σε τρεις βασικές ερευνητικές κατευθύνσεις προκειμένου να διακρίνουν με απόλυτη ακρίβεια τα κρυμμένα όρια κάθε μεμονωμένου προφίλ:

Αμοδική Κατάτμηση Στιγμιότυπου (Amodal Instance Segmentation): Η σύγχρονη έρευνα διαχωρίζει ρητά τις προβλέψεις της μάσκας κατάτμησης σε modal (το καθαρά ορατό τμήμα του αντικειμένου) και amodal (το πλήρες, πραγματικό περίγραμμα, συμπεριλαμβανομένου του κρυμμένου τμήματος). Για την ακριβή πρόβλεψη των μη ορατών ορίων σε επικαλυπτόμενα στοιχεία, αναπτύσσονται μοντέλα που ενσωματώνουν αλγορίθμους συσσωρευτικής μάθησης απόκρυψης (cumulative occlusion learning). Η προσέγγιση αυτή προσδίδει στο νευρωνικό δίκτυο την ικανότητα να «συμπληρώνει» νοητά το γεωμετρικό σχήμα του υποκείμενου αντικειμένου, αναδομώντας την πλήρη διατομή του προφίλ αλουμινίου παρά την απουσία οπτικής πληροφορίας.

Προσαρμοσμένες Αρχιτεκτονικές R-CNN για Βιομηχανικές Εφαρμογές: Προκειμένου να καλυφθούν οι αυστηρές απαιτήσεις της βιομηχανίας, έχουν αναπτυχθεί εξειδικευμένες παραλλαγές των τυπικών δικτύων, όπως το ROMS R-CNN (Rotation Occlusion Multi-Scale Region-CNN). Ο συγκεκριμένος αλγόριθμος έχει σχεδιαστεί ρητά για την ανίχνευση βιομηχανικού εξοπλισμού και εξαρτημάτων υπό εξαιρετικά πολύπλοκες συνθήκες, επιλύοντας ταυτόχρονα προβλήματα αλλαγής κλίμακας (multi-scale), τυχαίας περιστροφής και, πρωτίστως, σοβαρής αλληλοεπικάλυψης και απόκρυψης στη γραμμή παραγωγής.

Μοντελοποίηση Δύο Επιπέδων (Bilayer Modeling) και Δυναμική Βελτίωση Ακμών: Για την ελαχιστοποίηση της σύγχυσης στα σημεία τομής των προφίλ, προτείνονται αποκωδικοποιητές μάσκας με επίγνωση της απόκρυψης (occlusion-aware mask decoders). Χαρακτηριστικό παράδειγμα αποτελεί το Mask-Refined R-CNN (MR R-CNN), το οποίο ρυθμίζει δυναμικά το βήμα (dynamic stride) του επιπέδου ROIAlign. Η δυναμική αυτή προσαρμογή λειτουργεί ως ψηφιακός "μεγεθυντικός φακός", αυξάνοντας την πυκνότητα δειγματοληψίας στις περιοχές με υψηλή πολυπλοκότητα ακμών, όπως τα σημεία ένωσης των επικαλυπτόμενων προφίλ. Μέσω αυτής της διαδικασίας, το δίκτυο καταγράφει με μεγαλύτερη ευκρίνεια τις σχέσεις των εικονοστοιχείων (pixels), αποτρέποντας τη συγχώνευση των ορίων και τη συνακόλουθη λανθασμένη ταξινόμηση.

1.3.3 Ψηφιακός Θόρυβος

Ο ψηφιακός θόρυβος είναι μια αναπόφευκτη παράμετρος στις εικόνες που λαμβάνονται από ψηφιακούς αισθητήρες, ειδικά σε βιομηχανικά περιβάλλοντα όπου χρησιμοποιούνται αισθητήρες CMOS. Μπορεί να προκληθεί από διάφορους παράγοντες, όπως ηλεκτρομαγνητικές παρεμβολές από κοντινά μηχανήματα, εγγενείς περιορισμούς του ίδιου του υλικού της κάμερας, ή σφάλματα στην καταμέτρηση των φωτονίων (photon counting errors) κατά τη διαδικασία λήψης.

Ο θόρυβος υποβαθμίζει την ποιότητα της εικόνας, αλλοιώνοντας τις λεπτές ακμές και τα όρια που είναι κρίσιμα για την ακριβή αναγνώριση του γεωμετρικού σχήματος, και μπορεί να οδηγήσει σε σφάλματα στην ταξινόμηση. Σύμφωνα με σύγχρονες μελέτες πάνω στην απόδοση συστημάτων ανίχνευσης αντικειμένων (όπως τα μοντέλα YOLO και Faster-RCNN), οι πιο συνηθισμένοι τύποι θορύβου κατηγοριοποιούνται και επιδρούν ως εξής:

- **Θόρυβος Gaussian:** Προκαλείται συχνά από τον εγγενή θόρυβο των κυκλωμάτων (intrinsic circuit noise) και τις υψηλές θερμοκρασίες λειτουργίας των αισθητήρων. Η υποβάθμιση της απόδοσης των μοντέλων βαθιάς μάθησης είναι ευθέως ανάλογη με την τυπική απόκλιση (standard deviation) του εισαγόμενου θορύβου. Το φαινόμενο αυτό γίνεται ιδιαίτερα

καταστροφικό σε συνθήκες χαμηλού φωτισμού, καθώς η ποσότητα του θορύβου φτάνει να είναι συγκρίσιμη με τη μέγιστη τιμή των εικονοστοιχείων (pixels) της εικόνας.

- **Κρουστικός Θόρυβος (Salt-and-Pepper):** Έρευνες έχουν δείξει ότι ο κρουστικός θόρυβος έχει τεράστιο και δραστικό αντίκτυπο στην απόδοση των CNNs. Μόλις η πιθανότητα εμφάνισής του ξεπεράσει ένα συγκεκριμένο κατώφλι, η απόδοση των ανιχνευτών κατακρημνίζεται. Η αρνητική του επίδραση είναι τόσο ισχυρή που επισκιάζει πλήρως τις μεταβολές της φωτεινότητας, καθιστώντας τον παράγοντα του φωτισμού πρακτικά αμελητέο.
- **Θόρυβος Poisson (Shot Noise) και Uniform:** Ο θόρυβος Poisson είναι πολλαπλασιαστικός και αφορά τη φυσική συμπεριφορά του φωτός, ενώ ο ομοιόμορφος θόρυβος (Uniform) εμφανίζει παρόμοια, αλλά ελαφρώς ηπιότερη, συμπεριφορά με τον Gaussian.

Η Επίδραση της Κλίμακας (Object Size)

Ένας καθοριστικός παράγοντας που συνδυάζεται αρνητικά με τον θόρυβο είναι το μέγεθος των αντικειμένων προς ανίχνευση. Ενώ τα μεγάλα αντικείμενα μπορούν να εντοπιστούν ικανοποιητικά ακόμη και με υψηλά επίπεδα θορύβου Salt-and-Pepper υπό χαμηλό φωτισμό, η ανίχνευση μικρών αντικειμένων (όπως οι μικρότερες διατομές προφίλ) παρουσιάζει σοβαρές δυσκολίες. Στα μικρά αντικείμενα, τα μοντέλα συχνά αποτυγχάνουν ακόμη και υπό ιδανικές συνθήκες (καλός φωτισμός, απουσία θορύβου).

Στρατηγικές Αντιμετώπισης και το «Παράδοξο» του Θορύβου

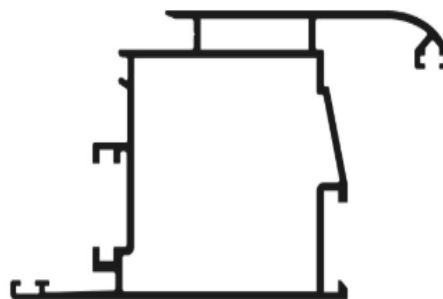
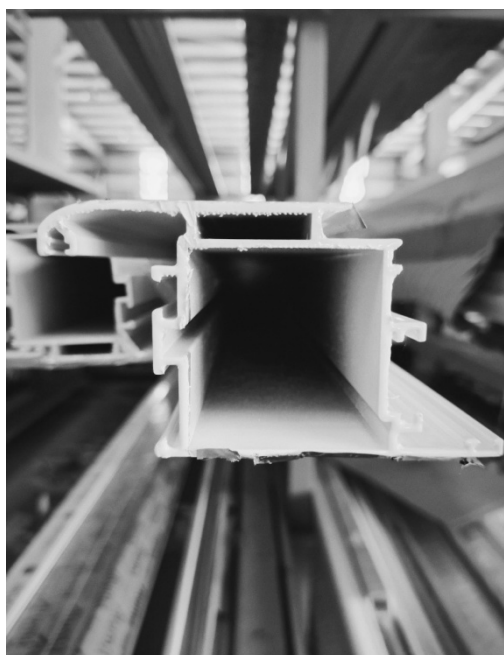
Οι λύσεις για την αντιμετώπιση του θορύβου περιλαμβάνουν την εφαρμογή αλγορίθμων αποθορυβοποίησης (denoising) στο στάδιο της προεπεξεργασίας. Επιπλέον, μια πολύ αποτελεσματική στρατηγική είναι η εκπαίδευση των μοντέλων σε σύνολα δεδομένων που περιέχουν τεχνητά προστιθέμενο θόρυβο. Αυτή η προσέγγιση λειτουργεί ως μια μορφή κανονικοποίησης, αναγκάζοντας το μοντέλο να μάθει να αγνοεί τον θόρυβο και να εστιάζει στα πραγματικά δομικά χαρακτηριστικά του σχήματος.

Τέλος, αξίζει να σημειωθεί ένα ενδιαφέρον πρακτικό εύρημα. Αν και ο γενικός κανόνας υπαγορεύει ότι η εισαγωγή θορύβου και η μείωση της φωτεινότητας έχουν αρνητικό αντίκτυπο στην απόδοση, υπάρχουν μεμονωμένες περιπτώσεις όπου η προσθήκη μιας μικρής ποσότητας θορύβου ή η ελαφριά μείωση του φωτισμού μπορεί να λειτουργήσει ευεργετικά, επιτρέποντας στο μοντέλο να ανιχνεύσει αντικείμενα που δεν εντοπίζονταν στην αρχική, "καθαρή" (raw) εικόνα.

1.3.4 Η Προσέγγιση "Sim-to-Real"

Μια από τις μεγαλύτερες προκλήσεις στην ανάπτυξη μοντέλων βαθιάς μάθησης για βιομηχανικές εφαρμογές είναι η συλλογή και επισημείωση ενός μεγάλου συνόλου δεδομένων. Η φωτογράφιση και η χειροκίνητη ταξινόμηση χιλιάδων δειγμάτων από εκατοντάδες διαφορετικά προφίλ είναι μια διαδικασία εξαιρετικά χρονοβόρα και δαπανηρή. Για την αντιμετώπιση αυτού του προβλήματος, μια καινοτόμος και αποτελεσματική προσέγγιση είναι η στρατηγική "Sim-to-Real" (Simulation-to-Reality).

Η κεντρική ιδέα αυτής της προσέγγισης είναι η χρήση συνθετικών δεδομένων που παράγονται από υπολογιστή για την εκπαίδευση του μοντέλου, και στη συνέχεια η εφαρμογή και η τελική ρύθμιση (fine-tuning) του μοντέλου σε ένα μικρότερο σύνολο πραγματικών δεδομένων. Στην περίπτωση των προφίλ αλουμινίου, τα συνθετικά δεδομένα μπορούν να παραχθούν απευθείας από τα υπάρχοντα αρχεία CAD. Από κάθε δισδιάστατο μοντέλο CAD, μπορούν να παραχθούν χιλιάδες δισδιάστατες εικόνες της διατομής του προφίλ, με τυχαίες μεταβολές στον προσανατολισμό, την κλίμακα και τον φωτισμό, δημιουργώντας έτσι ένα τεράστιο και ποικιλόμορφο σύνολο δεδομένων εκπαίδευσης χωρίς κανένα κόστος συλλογής.



Προφίλ αλουμινίου ανοιγόμενου συστήματος, πραγματική φωτογραφία και η σύγκριση με το αρχείο CAD

Αυτή ακριβώς την προσέγγιση ακολούθησε η έρευνα του Πανεπιστημίου του Salento το 2019, η οποία αποτέλεσε και την έμπνευση για την παρούσα εργασία. Οι ερευνητές χρησιμοποίησαν συνθετικά δεδομένα από αρχεία CAD για να εκπαιδεύσουν ένα Συνελικτικό Νευρωνικό Δίκτυο (CNN) για την ανίχνευση και την κατάτμηση προφίλ αλουμινίου. Η επικύρωση και ο έλεγχος του μοντέλου έγιναν με πραγματικές εικόνες που λήφθηκαν από κάμερες, αποδεικνύοντας την επιτυχία της μεταφοράς γνώσης από τον προσομοιωμένο στον πραγματικό κόσμο.

1.3.5 Η Πρόκληση της Μεταφοράς Γνώσης (The Reality Gap)

Παρά την επιτυχία της σε εργαστηριακό επίπεδο, η προσέγγιση "Sim-to-Real" ενέχει μια θεμελιώδη πρόκληση, γνωστή ως το «χάσμα της πραγματικότητας» (reality gap). Αυτό το χάσμα αναφέρεται στις εγγενείς διαφορές μεταξύ του ιδανικού, καθαρού περιβάλλοντος της προσομοίωσης και του πολύπλοκου, «θορυβώδους» πραγματικού κόσμου.

Ένα μοντέλο που εκπαιδεύεται αποκλειστικά σε τέλεια συνθετικά δεδομένα μπορεί να αποτύχει όταν έρθει αντιμέτωπο με τις ατέλειες του πραγματικού κόσμου, όπως ο θόρυβος των αισθητήρων, οι απρόβλεπτες αντανakλάσεις του φωτός, οι μικρές γρατζουνιές πάνω στα προφίλ, και οι μερικές αποκρύψεις από άλλα αντικείμενα. Η γεφύρωση αυτού του χάσματος είναι ο κεντρικός στόχος της παρούσας εργασίας. Η στρατηγική για την επίτευξη αυτού του στόχου περιλαμβάνει:

- **Domain Randomization:** Κατά τη δημιουργία των συνθετικών δεδομένων, εισάγονται σκόπιμα τυχαίες παραλλαγές που προσομοιώνουν τις ατέλειες του πραγματικού κόσμου. Αυτό περιλαμβάνει την προσθήκη τεχνητού θορύβου, τη μεταβολή της φωτεινότητας και της αντίθεσης, και την προσομοίωση μικρών αποκρύψεων.
- **Μάθηση Μεταφοράς (Transfer Learning):** Το μοντέλο που προ-εκπαιδεύεται στα συνθετικά δεδομένα δεν χρησιμοποιείται ως έχει, αλλά υποβάλλεται σε μια δεύτερη φάση εκπαίδευσης (fine-tuning) σε ένα μικρό, αλλά αντιπροσωπευτικό, σύνολο πραγματικών εικόνων που συλλέγονται από το βιομηχανικό περιβάλλον. Αυτό επιτρέπει στο μοντέλο να προσαρμόσει τη γνώση που απέκτησε από τα συνθετικά δεδομένα στις ιδιαιτερότητες του πραγματικού κόσμου.
- **Επαύξηση Δεδομένων (Data Augmentation):** Τόσο στα συνθετικά όσο και στα πραγματικά δεδομένα, εφαρμόζονται τεχνικές επαύξησης (περιστροφές, μετατοπίσεις κ.λπ.) για να αυξηθεί η ποικιλομορφία του συνόλου δεδομένων και να βελτιωθεί η ικανότητα γενίκευσης του μοντέλου. Η επιτυχής εφαρμογή αυτών των τεχνικών είναι κρίσιμη για τη μετάβαση από ένα μοντέλο που λειτουργεί σε ιδανικές συνθήκες σε ένα εύρωστο σύστημα ικανό να αποδίδει αξιόπιστα στο απαιτητικό περιβάλλον ενός σύγχρονου εργοστασίου.

1.3.5.1 Domain Randomization (Τυχαία Διαφοροποίηση Πεδίου)

Η μεταφορά μοντέλων Μηχανικής Μάθησης (και ειδικότερα Ενισχυτικής Μάθησης) από τα περιβάλλοντα προσομοίωσης στα φυσικά ρομπότ αποτελεί μια από τις μεγαλύτερες προκλήσεις στη σύγχρονη ρομποτική. Το λεγόμενο "Χάσμα Προσομοίωσης-Πραγματικότητας" (Sim2Real Gap) προκύπτει από τις ανακρίβειες των προσομοιωτών στην αποτύπωση των φυσικών παραμέτρων (τριβή, μάζα, αποσβέσεις) και την αδυναμία τους να αναπαράγουν πιστά τον οπτικό ρεαλισμό.

Για τη γεφύρωση αυτού του χάσματος, σε αντίθεση με την παραδοσιακή Ταυτοποίηση Συστήματος (System Identification) που απαιτεί ακριβή και δαπανηρή βαθμονόμηση, ή την Προσαρμογή Πεδίου (Domain Adaptation) που βασίζεται συχνά σε μεγάλο όγκο πραγματικών δεδομένων (π.χ. μέσω GANs), η βιβλιογραφία αναδεικνύει την Τυχαία Διαφοροποίηση Πεδίου (Domain Randomization - DR) ως μια εξαιρετικά αποδοτική, μη-επιβλεπόμενη (unsupervised) στρατηγική.

Η θεμελιώδης αρχή της DR είναι η δημιουργία πολλαπλών προσομοιωμένων περιβαλλόντων με έντονα τυχαιοποιημένες ιδιότητες (τόσο οπτικές, όπως φωτισμός, υφές και τοποθετήσεις καμερών, όσο και δυναμικές, όπως μάζα αντικειμένων και θόρυβος αισθητήρων). Εκθέτοντας το μοντέλο σε αυτή την τεράστια ποικιλία συνθηκών κατά την εκπαίδευση, ο πραγματικός κόσμος (target domain) καταλήγει να εκλαμβάνεται από το δίκτυο απλώς ως μία ακόμη τυχαία παραλλαγή της ευρείας κατανομής των δεδομένων εκπαίδευσης. Η επιτυχία της DR αποδίδεται μαθηματικά είτε ως μια μορφή δι-επίπεδης βελτιστοποίησης (bilevel optimization), είτε ως Μετα-Μάθηση (Meta-Learning), ειδικά όταν χρησιμοποιούνται επαναλαμβανόμενα δίκτυα (LSTMs) που προσαρμόζονται δυναμικά σε διαφορετικές φυσικές δυναμικές.

Από την Ομοιόμορφη στην Καθοδηγούμενη Τυχαιοποίηση (Guided Domain Randomization)

Η αρχική μορφή της μεθόδου (Uniform DR) βασίζεται στην ομοιόμορφη δειγματοληψία παραμέτρων μέσα σε προκαθορισμένα όρια. Ωστόσο, επειδή αυτή η προσέγγιση μπορεί να σπαταλήσει υπολογιστικούς πόρους σε εντελώς μη-ρεαλιστικά σενάρια, η σύγχρονη έρευνα επικεντρώνεται στην Καθοδηγούμενη Τυχαιοποίηση Πεδίου (Guided DR), η οποία χωρίζεται σε τρεις κύριες κατευθύνσεις:

1. **Καθοδήγηση βάσει Απόδοσης (Optimization for Task Performance):** Προσεγγίσεις όπως το AutoAugment ή το Learning to Simulate αντιμετωπίζουν την επιλογή των παραμέτρων DR ως πρόβλημα Ενισχυτικής Μάθησης ή εξελικτικών αλγορίθμων (π.χ. CMA-ES). Το σύστημα αξιολογεί πόσο καλά λειτουργεί μια διαμόρφωση προσομοίωσης στο τελικό έργο (validation set) και χρησιμοποιεί αυτό το σήμα ως «ανταμοιβή» για να δημιουργήσει πιο ωφέλιμα συνθετικά δεδομένα.

2. **Καθοδήγηση βάσει Πραγματικών Δεδομένων (Match Real Data):** Συστήματα όπως το SimOpt και το RCAN (Randomized-to-Canonical Adaptation Networks) λειτουργούν υβριδικά. Χρησιμοποιούν μια μικρή ποσότητα πραγματικών δεδομένων (τροχιές κίνησης ή πραγματικές εικόνες) για να ελαχιστοποιήσουν την απόκλιση μεταξύ του προσομοιωτή και του πραγματικού κόσμου, βελτιστοποιώντας στοχευμένα την κατανομή της τυχαιοποίησης.
3. **Καθοδήγηση εντός Προσομοιωτή (Guided by Sim Data):** Μέθοδοι όπως το DeceptionNet και το Active Domain Randomization (ADR) αποσκοπούν στη δημιουργία των πιο "δύσκολων" σεναρίων αποκλειστικά εντός της προσομοίωσης, χωρίς να χρειάζονται καθόλου πραγματικά δεδομένα. Το σύστημα αναζητά ενεργά τις παραλλαγές (π.χ. μέσω adversarial δικτύων) που μπερδεύουν περισσότερο την τρέχουσα πολιτική, αναγκάζοντας το μοντέλο να γίνει δραστικά πιο εύρωστο.

1.3.5.2 Μάθηση Μεταφοράς και Γεφύρωση μέσω Συνθετικών Δεδομένων (Bridged Transfer Learning & Fine-Tuning):

Το δίκτυο που έχει προ-εκπαιδευτεί στον τεράστιο όγκο των γενικών δεδομένων δεν χρησιμοποιείται αυτούσιο, αλλά υποβάλλεται σε μια δεύτερη φάση εκπαίδευσης (μικρορύθμιση) στο περιβάλλον-στόχο. Σύμφωνα με τη σύγχρονη βιβλιογραφία, και ειδικότερα τη μελέτη των Li et al. για την αξιοποίηση συνθετικών εικόνων στη Μάθηση Μεταφοράς, διαπιστώνεται ότι η απλή, ταυτόχρονη μίξη πραγματικών και συνθετικών δεδομένων κατά την εκπαίδευση υποβαθμίζει την τελική ακρίβεια του μοντέλου λόγω του εγγενούς χάσματος κατανομής (distribution gap).

Για την αντιμετώπιση αυτού του προβλήματος, η έρευνα προτείνει το πλαίσιο της Γεφυρωμένης Μεταφοράς (Bridged Transfer), το οποίο διαχωρίζει τη διαδικασία σε δύο αυστηρά βήματα:

- **Στάδιο 1 (Προσαρμογή Χαρακτηριστικών μέσω Συνθετικών Δεδομένων):** Το προ-εκπαιδευμένο μοντέλο εκπαιδεύεται αρχικά αποκλειστικά σε έναν μεγάλο όγκο συνθετικών εικόνων που παράγονται από μοντέλα τεχνητής νοημοσύνης (π.χ. Stable Diffusion). Σε αυτή τη φάση, ο εξαγωγέας χαρακτηριστικών (feature extractor) του δικτύου μαθαίνει τις νέες γεωμετρίες και καθίσταται πιο ευέλικτος.
- **Στάδιο 2 (Ταχεία Προσαρμογή σε Πραγματικά Δεδομένα):** Ο τελικός ταξινομητής (classifier) του μοντέλου επαναρχικοποιείται και το δίκτυο υποβάλλεται σε τελική μικρορύθμιση (fine-tuning) χρησιμοποιώντας αποκλειστικά το μικρό, αντιπροσωπευτικό σύνολο των πραγματικών εικόνων από τον φυσικό χώρο (π.χ. το εργοστάσιο).
- Για την περαιτέρω αύξηση της αποδοτικότητας, η έρευνα εισάγει την τεχνική της Αντιστροφής Στυλ Συνόλου Δεδομένων (Dataset Style Inversion - DSI). Αντί η παραγωγή συνθετικών

εικόνων να βασίζεται μόνο σε απλές λέξεις-κλειδιά, ο αλγόριθμος "μαθαίνει" ένα μοναδικό κρυπτογραφικό αποτύπωμα στυλ (style token) από τις πραγματικές φωτογραφίες του εργοστασίου. Αυτό το αποτύπωμα χρησιμοποιείται στη συνέχεια για να καθοδηγήσει τη δημιουργία των συνθετικών εικόνων (π.χ. δεδομένα CAD), διασφαλίζοντας ότι η υφή, ο φωτισμός και οι σκιές τους ευθυγραμμίζονται οπτικά με την πραγματικότητα του πεδίου-στόχου, αυξάνοντας δραματικά την ακρίβεια της μεταφοράς (έως και 30% σε σενάρια Few-Shot).

1.3.5.3 Επαύξηση Δεδομένων (Data Augmentation)

Εφαρμογή γεωμετρικών και επιφανειακών μετασχηματισμών (περιστροφές, μετατοπίσεις, καθρεφτισμοί) τόσο στα συνθετικά όσο και στα πραγματικά δεδομένα, ενισχύοντας την ικανότητα γενίκευσης του μοντέλου.

Η κρισιμότητα αυτής της στρατηγικής επιβεβαιώνεται και από τη σύγχρονη βιβλιογραφία, όπως στην έρευνα των Laskin et al. (2020) για την Ενισχυτική Μάθηση με Επαυξημένα Δεδομένα (RAD). Στη μελέτη αυτή αποδεικνύεται ότι η προσθήκη μιας απλής ενότητας (plug-and-play) επαύξησης δεδομένων μπορεί να αναβαθμίσει δραματικά την αποδοτικότητα των δεδομένων (data-efficiency) και την ικανότητα γενίκευσης των αλγορίθμων σε νέα, μη ιδωμένα περιβάλλοντα.

Ειδικότερα, η έρευνα ανέδειξε τις εξής βέλτιστες πρακτικές και συμπεράσματα:

Κυριαρχία Χωρικών Μετασχηματισμών: Από το σύνολο των διαθέσιμων τεχνικών (όπως η αλλαγή χρώματος, η προσθήκη θορύβου ή η αποκοπή τμημάτων), οι χωρικοί μετασχηματισμοί και κυρίως η τυχαία περικοπή (random crop) και η τυχαία μετατόπιση (random translate) είχαν τον μεγαλύτερο θετικό αντίκτυπο στην τελική απόδοση του συστήματος από οπτικά δεδομένα. Οι μετασχηματισμοί αυτοί επιτρέπουν στον αλγόριθμο να εστιάζει στο κύριο αντικείμενο δράσης, αγνοώντας άσχετες λεπτομέρειες του υποβάθρου.

Διατήρηση Χρονικής Πληροφορίας: Σε συστήματα που αναλύουν συνεχόμενα καρέ (frame stacking) για τον υπολογισμό ταχυτήτων και κινήσεων, η επαύξηση πρέπει να εφαρμόζεται τυχαία μεταξύ των διαφορετικών δειγμάτων μιας δέσμης (minibatch), αλλά απόλυτα σταθερά σε όλα τα διαδοχικά καρέ της ίδιας χρονικής ακολουθίας. Αυτό διασφαλίζει ότι η οπτική παραμόρφωση δεν καταστρέφει την αντίληψη του χρόνου από το μοντέλο.

Επαύξηση Δεδομένων Κατάστασης: Πέρα από τις εικόνες, η έρευνα εισήγαγε νέες τεχνικές για αριθμητικά δεδομένα κατάστασης (state-based inputs), όπως η τυχαία κλιμάκωση πλάτους (random amplitude scale). Η τεχνική αυτή πολλαπλασιάζει τα χαρακτηριστικά με μια τυχαία μεταβλητή,

αναγκάζοντας το μοντέλο να γίνει ανθεκτικό στον θόρυβο ενώ διατηρεί την εγγενή πληροφορία των δεδομένων (όπως το πρόσημο).

Υπεροχή έναντι Πολύπλοκων Αρχιτεκτονικών: Η ορθή εφαρμογή αυτών των επαυξήσεων επέτρεψε σε απλούς αλγορίθμους να ξεπεράσουν σε απόδοση (τόσο σε αποτελεσματικότητα όσο και σε ταχύτητα σύγκλισης) πολύ πιο περίπλοκες προσεγγίσεις αναπαράστασης της εποχής (όπως τα συστήματα CURL ή SLAC) στα ευρέως διαδεδομένα benchmarks DeepMind Control Suite και OpenAI Gym.

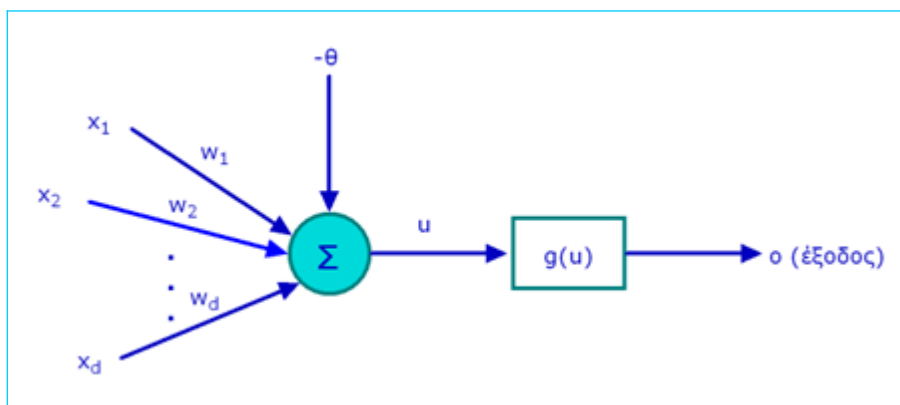
Κεφάλαιο 2: Τεχνητά Νευρωνικά δίκτυα: Από τον αισθητήρα μέχρι τα συνελικτικά ΤΝΔ

Η Τεχνητή Νοημοσύνη έχει εξελιχθεί σημαντικά τις τελευταίες δεκαετίες, περνώντας από απλά υπολογιστικά μοντέλα λήψης αποφάσεων σε σύνθετα συστήματα βαθιάς μάθησης που μπορούν να αναγνωρίζουν εικόνες, φωνή, κείμενο και πολύπλοκα πρότυπα δεδομένων. Κεντρικό ρόλο σε αυτή την εξέλιξη είχαν τα Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ), τα οποία εμπνέονται από τη λειτουργία του ανθρώπινου εγκεφάλου και αποτελούνται από διασυνδεδεμένες υπολογιστικές μονάδες, τους τεχνητούς νευρώνες.

2.1 Τα πρώτα χρόνια των ΤΝΔ

Τα πρώτα ΤΝΔ βασίζονταν σε σχετικά απλές αρχιτεκτονικές. Ένα από τα πρώτα σημαντικά μοντέλα ήταν ο αισθητήρας (perceptron) Rosenblatt (1962), το οποίο μπορούσε να εκτελεί βασικές ταξινομήσεις, διαχωρίζοντας δεδομένα σε κατηγορίες όταν αυτά ήταν γραμμικά διαχωρίσιμα. Συνιστά τη βάση, για όλες τις περεταίρω μελέτες, καθώς και για την κατασκευή πολυπλοκότερων νευρωνικών δικτύων. Αποτελείται από ένα μόνο νευρώνα και είναι το απλούστερο νευρωνικό δίκτυο που μπορεί να κατασκευαστεί. Το Perceptron διαθέτει έναν ορισμένο αριθμό εισόδων και παράγει μόνο μια έξοδο και απεικονίζεται στην Εικόνα 1.

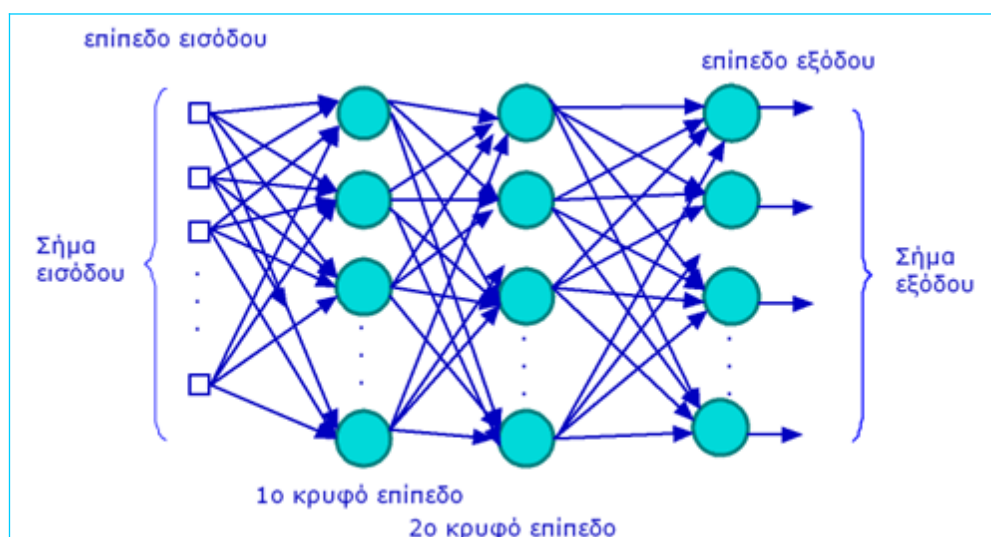
Παρότι το perceptron αποτέλεσε σημαντικό βήμα για την ανάπτυξη της μηχανικής μάθησης, είχε σοβαρούς περιορισμούς. Δεν μπορούσε, για παράδειγμα, να επιλύσει προβλήματα που απαιτούσαν μη γραμμικό διαχωρισμό, όπως το γνωστό πρόβλημα XOR.



Εικόνα 1: Perceptron.

Για την αντιμετώπιση αυτών των περιορισμών αναπτύχθηκαν τα πολυεπίπεδα νευρωνικά δίκτυα, γνωστά ως Multi-Layer Perceptrons (MLPs). Τα δίκτυα αυτά περιλαμβάνουν ένα επίπεδο εισόδου, ένα ή περισσότερα κρυφά επίπεδα και ένα επίπεδο εξόδου. Η εισαγωγή των κρυφών επιπέδων επέτρεψε στα δίκτυα να μαθαίνουν πιο σύνθετες σχέσεις μεταξύ εισόδων και εξόδων. Καθοριστικής σημασίας υπήρξε επίσης ο αλγόριθμος οπισθοδιάδοσης σφάλματος (backpropagation) [ref], μέσω του οποίου το δίκτυο προσαρμόζει τα βάρη των συνδέσεων του ώστε να μειώνει σταδιακά το σφάλμα πρόβλεψης.

Τα δίκτυα αυτά αποτελούνται από ένα σύνολο αισθητήρων και είναι μια γενίκευση του μοντέλου του απλού αισθητήρα. Συγκεκριμένα, αποτελείται από ένα επίπεδο εισόδου, ένα ή περισσότερα κρυφά επίπεδα (hidden layers) και ένα επίπεδο εξόδου. Αυτό που, κυρίως, χαρακτηρίζει τα δίκτυα τύπου MLP είναι ότι οι νευρώνες οποιουδήποτε επιπέδου τροφοδοτούν το αμέσως επόμενο και τροφοδοτούνται από το αμέσως προηγούμενο επίπεδό τους. Έτσι, έστω επίπεδο l , τότε αυτό τροφοδοτείται από το επίπεδο $l-1$ και τροφοδοτεί το επίπεδο $l+1$. Η αρχιτεκτονική ενός MLP με ένα επίπεδο εισόδου, δύο κρυφά επίπεδα και ένα επίπεδο εξόδου, είναι αυτή που παρουσιάζεται στην Εικόνα 2.



Εικόνα 2: MLP δύο κρυφών επιπέδων.

Το δίκτυο, που παρουσιάστηκε παραπάνω, είναι πλήρως διασυνδεδεμένο (fully interconnected), πράγμα που σημαίνει ότι ένας νευρώνας οποιουδήποτε επιπέδου συνδέεται με όλους τους νευρώνες του προηγούμενου επιπέδου. Η ροή του σήματος στο δίκτυο προχωρά σε μια προς τα εμπρός κατεύθυνση, από τα αριστερά προς τα δεξιά -από το επίπεδο εισόδου προς το επίπεδο εξόδου- από επίπεδο σε επίπεδο (Rumelhart et al., 1986).

Ωστόσο, παρά τη σημαντική πρόοδο, τα κλασικά πολυεπίπεδα δίκτυα αντιμετώπιζαν δυσκολίες όταν εφαρμόζονταν σε δεδομένα υψηλής διάστασης, όπως οι εικόνες. Μια εικόνα μπορεί να αποτελείται από χιλιάδες ή εκατομμύρια pixels, γεγονός που οδηγεί σε τεράστιο αριθμό παραμέτρων αν κάθε pixel συνδεθεί πλήρως με κάθε νευρώνα του επόμενου επιπέδου. Αυτό αυξάνει σημαντικά την υπολογιστική πολυπλοκότητα και τον κίνδυνο υπερπροσαρμογής.

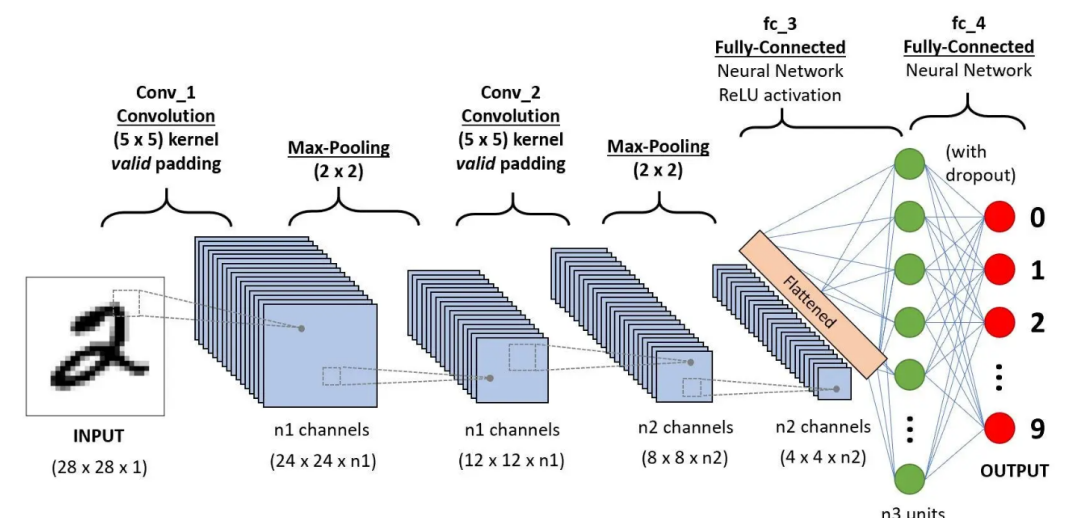
Η ανάγκη για πιο αποδοτική επεξεργασία εικόνων οδήγησε στην ανάπτυξη των Συνελικτικών Νευρωνικών Δικτύων, γνωστών ως Convolutional Neural Networks ή CNNs. Τα CNNs σχεδιάστηκαν ώστε να αξιοποιούν τη χωρική δομή των εικόνων. Αντί να αντιμετωπίζουν κάθε pixel ανεξάρτητα, χρησιμοποιούν μικρά φίλτρα ή πυρήνες συνέλιξης που «σαρώνουν» την εικόνα και εντοπίζουν τοπικά χαρακτηριστικά, όπως ακμές, γωνίες, υφές και σχήματα (LeCun et al., 1989).

2.2 Συνελικτικά Νευρωνικά Δίκτυα και σύγχρονες εξελίξεις

Τα Συνελικτικά Νευρωνικά Δίκτυα ή Convolutional Neural Networks αποτελούν μία από τις σημαντικότερες αρχιτεκτονικές της βαθιάς μάθησης, ιδιαίτερα για προβλήματα που σχετίζονται με εικόνες, βίντεο και γενικότερα δεδομένα με χωρική δομή. Η βασική τους διαφορά από τα κλασικά πλήρως συνδεδεμένα νευρωνικά δίκτυα είναι ότι δεν αντιμετωπίζουν κάθε είσοδο ανεξάρτητα, αλλά αξιοποιούν τη γειτονική σχέση μεταξύ των εικονοστοιχείων (pixels). Αυτό τα καθιστά ιδιαίτερα κατάλληλα για εφαρμογές υπολογιστικής όρασης, όπως ταξινόμηση εικόνων, ανίχνευση αντικειμένων, σημασιολογική κατάτμηση, αναγνώριση προσώπων, ιατρική απεικόνιση και αυτόνομη πλοήγηση.

Η βασική λειτουργία ενός CNN στηρίζεται στη συνέλιξη. Σε ένα συνελικτικό επίπεδο, μικρά φίλτρα, γνωστά και ως πυρήνες συνέλιξης, μετακινούνται πάνω στην εικόνα και υπολογίζουν τοπικές συσχετίσεις. Έτσι, το δίκτυο μπορεί να ανιχνεύει χαρακτηριστικά όπως ακμές, γωνίες, καμπύλες, υφές και απλά σχήματα. Στα πρώτα επίπεδα τα χαρακτηριστικά είναι συνήθως απλά και χαμηλού επιπέδου, ενώ στα βαθύτερα επίπεδα γίνονται πιο σύνθετα, συνδυάζοντας τις αρχικές πληροφορίες σε ανώτερες αναπαραστάσεις. Για παράδειγμα, ένα CNN που εκπαιδεύεται για αναγνώριση οχημάτων μπορεί αρχικά να ανιχνεύει περιγράμματα και τροχούς, και αργότερα να συνδυάζει αυτά τα στοιχεία για να αναγνωρίσει ολόκληρο το όχημα.

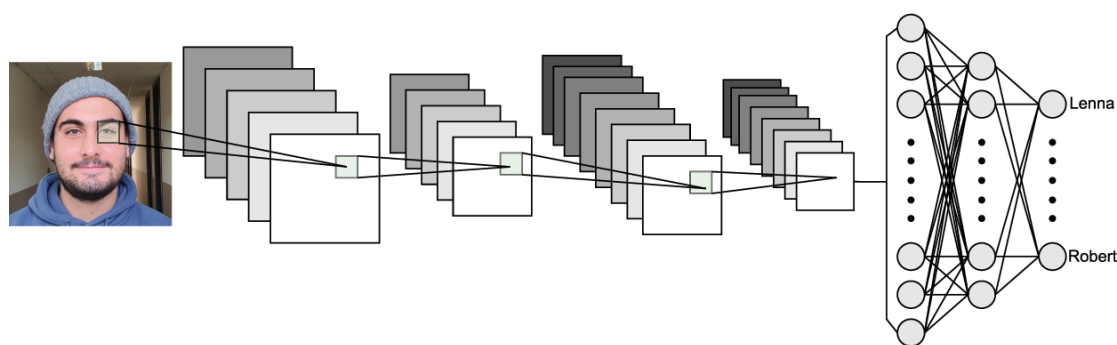
Ένα τυπικό CNN αποτελείται από διάφορα είδη επιπέδων. Τα convolutional layers εξάγουν χαρακτηριστικά, τα activation layers, όπως η ReLU, εισάγουν μη γραμμικότητα, ενώ τα pooling layers μειώνουν τη χωρική διάσταση των δεδομένων, διατηρώντας τις πιο σημαντικές πληροφορίες. Στο τέλος, συχνά υπάρχουν fully connected layers, τα οποία μετατρέπουν τα εξαγόμενα χαρακτηριστικά σε τελική απόφαση ταξινόμησης. Η χρήση κοινών φίλτρων σε ολόκληρη την εικόνα μειώνει σημαντικά τον αριθμό των παραμέτρων σε σχέση με ένα πλήρως συνδεδεμένο δίκτυο, γεγονός που βελτιώνει την αποδοτικότητα και περιορίζει τον κίνδυνο υπερπροσαρμογής. Ένα παράδειγμα αρχιτεκτονικής CNN για την αναγνώριση χαρακτήρων φαίνεται στην Εικόνα 3.



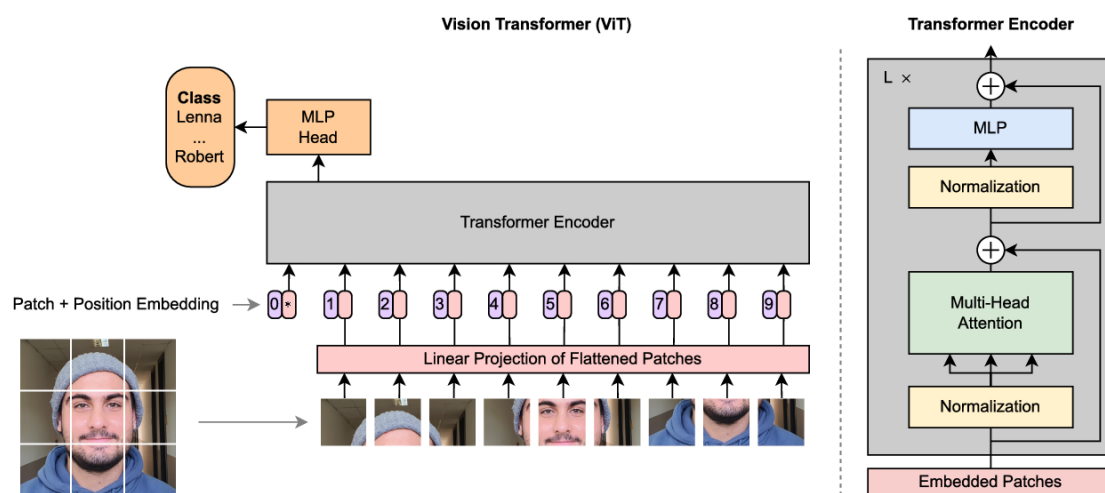
Εικόνα 3: Αρχιτεκτονική CNN για ταξινόμηση χειρόγραφων χαρακτήρων.

Η ιστορική εξέλιξη των CNNs συνδέεται με σημαντικές αρχιτεκτονικές. Τα πρώτα μοντέλα, όπως το LeNet (LeCun et al, 1998), χρησιμοποιήθηκαν κυρίως για αναγνώριση χειρόγραφων ψηφίων. Αργότερα, το AlexNet (Krizhevsky et al., 2012) έδειξε τη δύναμη των βαθιών CNNs σε μεγάλα σύνολα εικόνων, αξιοποιώντας GPU υπολογισμό και τεχνικές όπως dropout και ReLU. Στη συνέχεια εμφανίστηκαν αρχιτεκτονικές όπως τα VGGNet (Simonyan et al., 2015), GoogLeNet/Inception (Szegedy et al., 2015), ResNet (He et al., 2016), DenseNet (Huang et al., 2017), MobileNet (Howard et. al, 2017) και EfficientNet (Tan et al., 2019), καθεμία από τις οποίες εισήγαγε νέες ιδέες για βάθος, αποδοτικότητα, επαναχρησιμοποίηση χαρακτηριστικών ή μείωση υπολογιστικού κόστους. Ειδικά τα ResNets, με τις υπολειμματικές συνδέσεις, επέτρεψαν την εκπαίδευση πολύ βαθύτερων δικτύων, αντιμετωπίζοντας το πρόβλημα της υποβάθμισης της απόδοσης σε μεγάλα βάθη.

Τα τελευταία χρόνια, η κυριαρχία των CNNs αμφισβητήθηκε από τους Μετασχηματισμούς Όρασης ή όπως είναι πιο ευρέως γνωστοί ως Vision Transformers-ViT (Dosovitskiy et al., 2021). Οι ViT, που αρχικά έγιναν γνωστοί στην επεξεργασία φυσικής γλώσσας, εφαρμόστηκαν στην όραση μέσω μετασχηματισμών, όπου η εικόνα χωρίζεται σε υποεικόνες (patches) και επεξεργάζεται με μηχανισμούς προσοχής (attention mechanisms). Η προσέγγιση αυτή επέτρεψε στα μοντέλα να συλλαμβάνουν πιο εύκολα μακρινές σχέσεις μέσα στην εικόνα. Ωστόσο, οι καθαροί ViT (Εικόνα 4) συχνά απαιτούν πολύ μεγάλα σύνολα δεδομένων και υψηλό υπολογιστικό κόστος, ενώ δεν διαθέτουν εξ αρχής τις ισχυρές χωρικές προκαταλήψεις των CNNs, όπως η τοπικότητα και η μεταθετική ισοδυναμία. Για τον λόγο αυτό, η σύγχρονη έρευνα δεν εγκατέλειψε τα CNNs, αλλά άρχισε να τα επανασχεδιάζει και να τα συνδυάζει με τεχνικές από τους ViT.



(a) Common CNN architecture

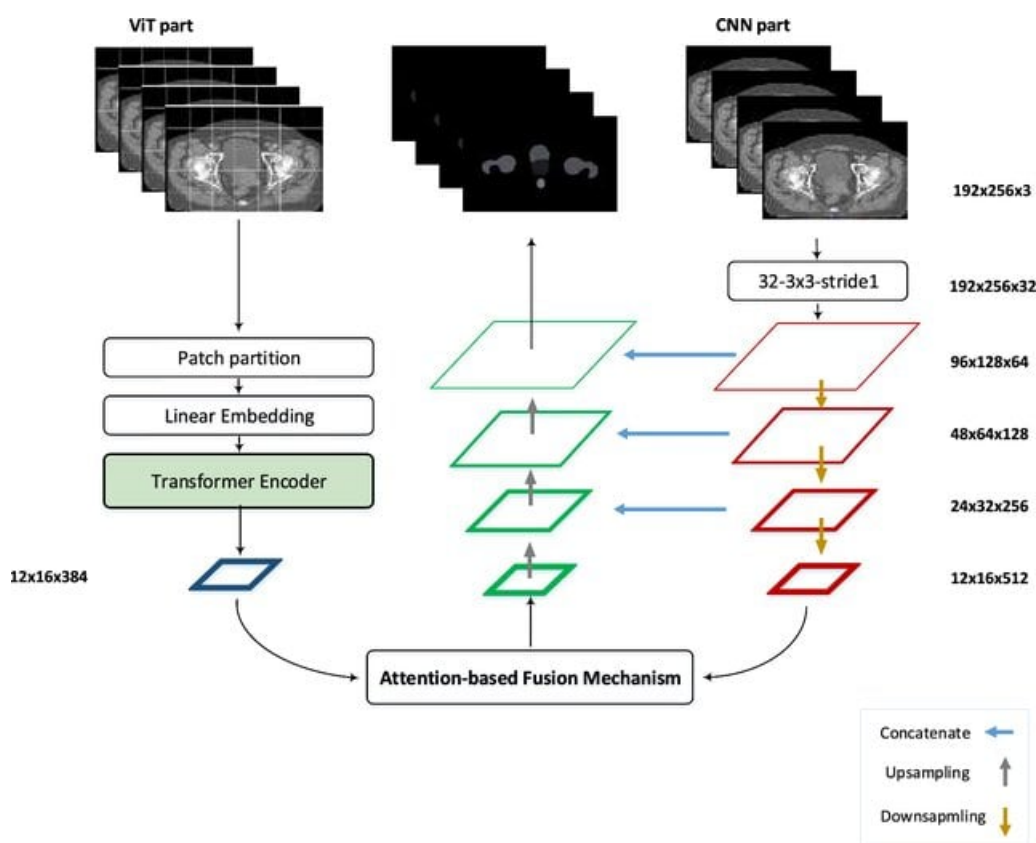


(b) Vision Transformer architecture

Εικόνα 4: Οπτική απεικόνιση των θεμελιωδών διαφορών μεταξύ των αρχιτεκτονικών των Vision Transformers (πάνω) και των CNNs (κάτω), καθώς και των μηχανισμών λειτουργίας τους (από <https://www.gti.ssr.upm.es/data/upm-gti-face-dataset>).

Μία από τις πιο χαρακτηριστικές πρόσφατες εξελίξεις είναι η αρχιτεκτονική ConvNeXt. Η εργασία “A ConvNet for the 2020s” από τους (Liu et. al, 2022) επανεξέτασε τον σχεδιασμό των κλασικών CNNs υπό το πρίσμα των Vision Transformers και έδειξε ότι ένα «εκσυγχρονισμένο» καθαρό συνελικτικό δίκτυο μπορεί να ανταγωνιστεί ισχυρά ViT μοντέλα. Το ConvNeXt βασίζεται σε τυπικά συνελικτικά στοιχεία, αλλά υιοθετεί σχεδιαστικές επιλογές εμπνευσμένες από σύγχρονες αρχιτεκτονικές, όπως μεγαλύτερους πυρήνες, διαφορετική οργάνωση blocks και βελτιωμένες πρακτικές εκπαίδευσης. Σύμφωνα με τους συγγραφείς, η οικογένεια ConvNeXt πέτυχε πολύ υψηλή ακρίβεια στο ImageNet και ανταγωνιστική απόδοση σε αναγνώριση αντικειμένων και σημασιολογική τμηματοποίηση εικόνας, διατηρώντας την απλότητα των ConvNets.

Μια δεύτερη σημαντική τάση είναι η χρήση μεγάλων πυρήνων συνέλιξης. Παραδοσιακά, πολλά CNNs χρησιμοποιούσαν μικρούς πυρήνες, όπως 3×3 , επειδή είναι αποδοτικά και μπορούν να συνδυαστούν σε βαθύτερα επίπεδα. Όμως, μεγαλύτεροι πυρήνες επιτρέπουν στο μοντέλο να βλέπει μεγαλύτερο τμήμα της εικόνας σε κάθε βήμα, προσεγγίζοντας σε κάποιο βαθμό την ικανότητα των ViT να μοντελοποιούν μακρινές εξαρτήσεις.



Εικόνα 5: Υβριδική αρχιτεκτονική CNN-ViT (Arjmandi et al., 2024).

Παράλληλα, έχουν αναπτυχθεί υβριδικές αρχιτεκτονικές CNN-ViT (Εικόνα 5). Σε αυτές, τα CNNs χρησιμοποιούνται συνήθως για την εξαγωγή τοπικών χαρακτηριστικών, ενώ οι Transformers αναλαμβάνουν τη μοντελοποίηση καθολικών σχέσεων. Αυτή η λογική αξιοποιεί τα πλεονεκτήματα και των δύο οικογενειών: τα CNNs είναι αποδοτικά και ισχυρά στη χωρική επεξεργασία, ενώ οι Transformers είναι ισχυροί στη σύνθεση πληροφορίας από απομακρυσμένες περιοχές της εικόνας. Διάφορες μελέτες αναφέρουν ότι τα υβριδικά CNN-ViT μοντέλα έχουν εξελιχθεί σε σημαντική ερευνητική κατεύθυνση για την υπολογιστική όραση, με εφαρμογές σε ταξινόμηση εικόνων, ανίχνευση αντικειμένων και παραγωγή εικόνων (Wu et al., 2021), (Yuan et al., 2021).

Συνολικά, οι τελευταίες εξελίξεις δείχνουν ότι τα CNNs δεν έχουν αντικατασταθεί πλήρως από τους Transformers. Αντίθετα, έχουν εισέλθει σε μια νέα φάση εξέλιξης. Τα σύγχρονα CNNs είναι

βαθύτερα, πιο αποδοτικά, συχνά σχεδιασμένα μέσω neural architecture search και εμπλουτισμένα με ιδέες από attention μηχανισμούς. Η τάση δεν είναι πλέον η απόλυτη διάκριση ανάμεσα σε CNNs και Transformers, αλλά η δημιουργία αρχιτεκτονικών που συνδυάζουν τοπική, χωρική και καθολική πληροφορία. Για τον λόγο αυτό, τα CNNs παραμένουν θεμελιώδη στην υπολογιστική όραση και συνεχίζουν να αποτελούν βασικό σημείο αναφοράς στη σύγχρονη Τεχνητή Νοημοσύνη.

Κεφάλαιο 3 Μεθοδολογία

3.1 Μεθοδολογία Σύνθεσης Δεδομένων

3.1.1 Εισαγωγή στο Σύνολο Δεδομένων "DataSilver"

Η μεγαλύτερη πρόκληση στην ανάπτυξη συστημάτων Μηχανικής Όρασης (Machine Vision) για βιομηχανικές εφαρμογές είναι η έλλειψη επαρκών, σχολιασμένων (annotated) πραγματικών δεδομένων. Για την αντιμετώπιση του χάσματος μεταξύ προσομοίωσης και πραγματικότητας (Sim-to-Real Gap), αναπτύχθηκε ένας αυτοματοποιημένος αγωγός δημιουργίας δεδομένων (Data Generation Pipeline), ικανός να μετατρέπει αμιγώς δισδιάστατα κατασκευαστικά σχέδια CAD (αρχεία DXF) σε ρεαλιστικές, συνθετικές σκηνές εκπαίδευσης.

Το παραγόμενο σύνολο δεδομένων, το οποίο ονομάστηκε "DataSilver", δημιουργήθηκε για να τροφοδοτήσει ταυτόχρονα και τα δύο στάδια της αρχιτεκτονικής του συστήματος (Ανιχνευτή και Ταξινομητή), εξάγοντας 116.200 εικόνες βελτιστοποιημένες μέσω μορφολογικών και φωτομετρικών αλγορίθμων.

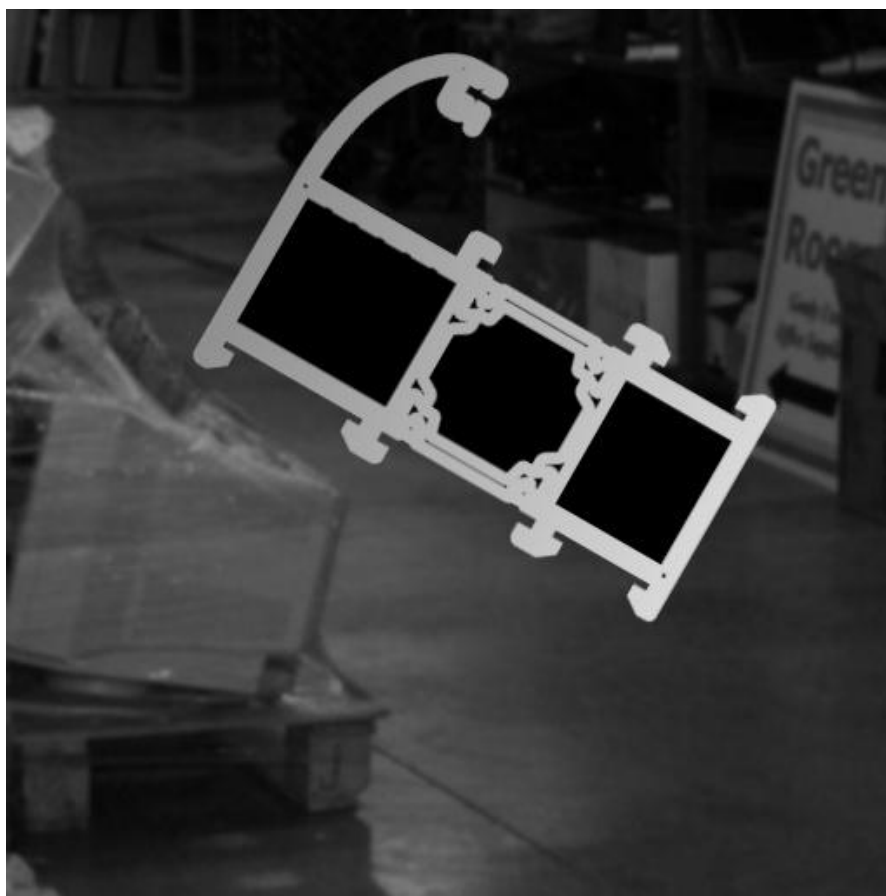
3.1.2 Από το CAD στη Γεωμετρική Αναπαράσταση

Η διαδικασία εκκινεί με την ανάγνωση των διανυσματικών αρχείων DXF. Δεδομένου ότι τα παραδοσιακά αρχεία CAD αποτελούνται από γραμμές και τόξα (vectors), το σύστημα αξιοποιεί βιβλιοθήκες απόδοσης (rendering) για να μεταφράσει τη γεωμετρία σε διαφανή γραφικά ψηφιδωτά (Raster PNGs).

Η κρισιμότερη καινοτομία σε αυτό το στάδιο εντοπίζεται στην Προηγμένη Διατήρηση Γεωμετρίας (Advanced Geometry Preservation). Αντί για μια απλή χρωματική επικάλυψη, ο αλγόριθμος εκτελεί Εσωτερική Ανάλυση Βασισμένη σε Κανόνες (Rule-Based Internal Analysis):

- **Σφράγιση Μικρο-διαρροών (Micro-Leak Sealing):** Εφαρμόζεται μορφολογικό κλείσιμο (Morphological Closing) με πυρήνα 3x3 για τη διόρθωση σχεδιαστικών ατελειών στο DXF, διασφαλίζοντας ότι τα εσωτερικά σχήματα είναι μαθηματικά συνεχή.

- **Αλγόριθμος Flood-Fill και Ανάλυση Θαλάμων:** Το σύστημα εντοπίζει τους κενούς χώρους εντός του προφίλ. Μέσω δυναμικού κατωφλίου (1.5% της κλίμακας του αντικειμένου), οι στενές εσοχές λαμβάνουν τη μεταλλική υφή, ενώ οι πραγματικοί κοίλοι θάλαμοι αποδίδονται με απόλυτο μαύρο χρώμα. Αυτό προσομοιώνει την πλήρη απουσία φωτός (Deep Shadows) που παρατηρείται στις πραγματικές κοιλότητες του αλουμινίου στο εργοστάσιο.
- **Σύνθεση Υφής:** Η εξωτερική επιφάνεια του μετάλλου παράγεται μαθηματικά μέσω της πρόσθεσης ενός βασικού γκρι τόνου, Γκαουσιανού θορύβου (Gaussian Noise), και γραμμικών διαβαθμίσεων (1D linear gradients), δημιουργώντας με αυτόν τον τρόπο την ψευδαίσθηση της μεταλλικής αντανάκλασης.



Σύνθεση Σκηνής και Επαύξηση Δεδομένων (Scene Generation)

Προκειμένου να αποτραπεί η απομνημόνευση του φόντου από το νευρωνικό δίκτυο (Background Bias), το παραγόμενο προφίλ ενσωματώνεται σε τυχαία, πραγματικά φόντα γραμμής παραγωγής. Τα φόντα αυτά μετατρέπονται προηγουμένως σε κλίμακα του γκρι, θολώνουν ομοιόμορφα (Gaussian Blur) και σκοτεινιάζουν (darkening blending). Η τεχνική αυτή αναπτύχθηκε για να κατευθύνει τον μηχανισμό προσοχής (attention mechanism) του δικτύου αποκλειστικά στο γεωμετρικό σχήμα του αντικειμένου ενδιαφέροντος. Κατά την τοποθέτηση, το αντικείμενο υφίσταται ακραίες γεωμετρικές μετασχηματίσεις, περιλαμβάνοντας **τυχαία κλιμάκωση από 40% έως 70% του καμβά** (YOLO_IMG_SIZE = 512x512), **περιστροφή 360 μοιρών** (Rotation Invariance), και αναστροφής (Flips) στον οριζόντιο και κάθετο άξονα.

3.2 Εξαγωγή Διπλού Σχήματος: Ανιχνευτής και Ταξινομητής

Ο αγωγός τερματίζει με τη δημιουργία δύο ξεχωριστών δομών δεδομένων για το σύστημα συμπεράσματος (inference), διαχωρισμένων αυστηρά σε σύνολα εκπαίδευσης και επικύρωσης (**Train/Val Split: 80/20**):

- **Δεδομένα Εντοπισμού (YOLO Detectors):** Το σύστημα υπολογίζει αυτόματα τις συντεταγμένες του πλαισίου οριοθέτησης (Bounding Box) βάσει του καναλιού Alpha (διαφάνεια) της τελικής εικόνας. Επιπλέον, για την καταστολή των Ψευδώς Θετικών προβλέψεων (False Positives), υλοποιήθηκε μια στρατηγική Έγχυσης Αρνητικών Δειγμάτων (10%). Εισήχθησαν εικόνες που περιέχουν μόνο το εργοστασιακό φόντο (χωρίς προφίλ), εξαναγκάζοντας τον ανιχνευτή να κατανοήσει την απουσία του αντικειμένου.
- **Δεδομένα Ταξινόμησης (Classifiers):** Για τα δίκτυα ταξινόμησης, εξάγεται μια περικομμένη εκδοχή (crop) του αντικειμένου. Προσομοιώνοντας τις εγγενείς ανακρίβειες των πραγματικών ανιχνευτών, προστίθεται χωρική διακύμανση (Bounding Box Jitter) 5% έως 20%. Τέλος, η εικόνα "γεμίζει" με ουδέτερο γκρι (Pad to Square) για να αποκτήσει απόλυτα τετράγωνο σχήμα (1:1 aspect ratio), αποτρέποντας την επιβλαβή παραμόρφωση των αναλογιών (aspect ratio distortion) κατά την εισαγωγή της στα δίκτυα.

3.3 Αρχιτεκτονικές Μοντέλων Νευρωνικών Δικτύων

Για την ανάπτυξη του διαδοχικού συστήματος (cascaded pipeline) αξιολογήθηκαν διαφορετικές αρχιτεκτονικές βαθιάς μάθησης, τόσο για το στάδιο του χωρικού εντοπισμού (Object Detection) όσο και για το στάδιο της ταξινόμησης (Classification). Η επιλογή των μοντέλων βασίστηκε στον κρίσιμο συμβιβασμό μεταξύ της ικανότητας γενίκευσης και του υπολογιστικού κόστους σε περιβάλλοντα περιορισμένων πόρων (Edge AI).

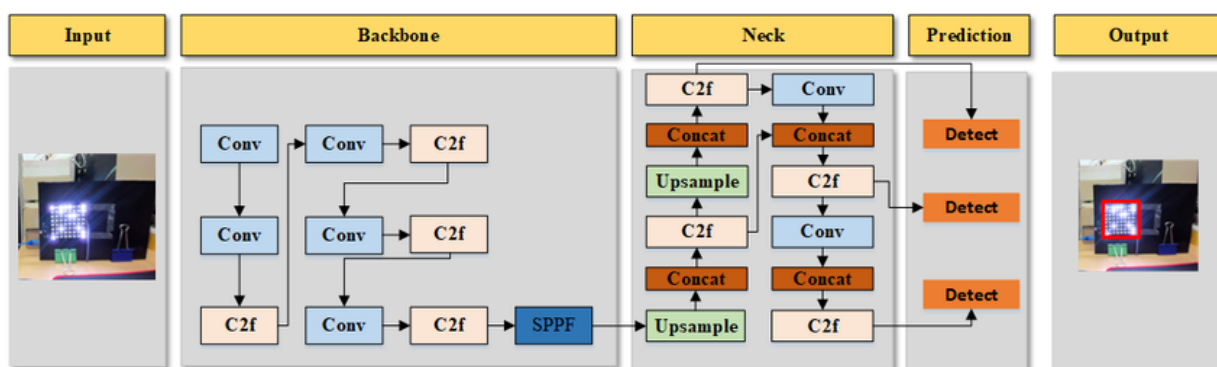
3.3.1 Αρχιτεκτονικές Μοντέλων Χωρικού Εντοπισμού (Object Detectors)

Η ανίχνευση του προφίλ αλουμινίου σε πραγματικό χρόνο απαιτεί τη χρήση αλγορίθμων εντοπισμού ενός σταδίου (Single-Stage Object Detectors). Στην παρούσα έρευνα επιλέχθηκαν αρχιτεκτονικές βασισμένες στην οικογένεια YOLO (You Only Look Once), οι οποίες αποτελούν το βιομηχανικό πρότυπο για ταχεία επεξεργασία.

Η Αρχιτεκτονική της Οικογένειας YOLOv8

- **Κορμός Εξαγωγής Χαρακτηριστικών (Backbone):** Αξιοποιεί μια τροποποιημένη μορφή του δικτύου CSPDarknet. Η σημαντικότερη αναβάθμιση είναι η αντικατάσταση των παραδοσιακών μπλοκ C3 από τα μπλοκ C2f (Cross Stage Partial with 2 Convolutions). Τα μπλοκ C2f συνδυάζουν χαρακτηριστικά υψηλού επιπέδου (high-level) με χαρακτηριστικά χαμηλού επιπέδου, βελτιώνοντας δραματικά τη ροή των κλίσεων (gradient flow) κατά την οπισθοδιάδοση (backpropagation) και διατηρώντας ταυτόχρονα το υπολογιστικό βάρος σε χαμηλά επίπεδα.

- **Λαιμός (Neck - Path Aggregation Network):** Το δίκτυο συγχώνευσης χαρακτηριστικών ακολουθεί τη δομή PAnet (Path Aggregation Network), επιτρέποντας την αμφίδρομη ροή (top-down και bottom-up) της σημασιολογικής πληροφορίας. Στην έκδοση v8, το στάδιο αυτό έχει απλοποιηθεί αφαιρώντας περιττές συνελίξεις 1x1 πριν την αναβάθμιση της ανάλυσης (upsampling), επιταχύνοντας έτσι την εκτέλεση.
- **Αποσυζευγμένη Κεφαλή χωρίς Άγκυρες (Decoupled Anchor-Free Head):** Σε αντίθεση με παλαιότερες εκδόσεις που βασίζονταν σε προκαθορισμένα κουτιά (anchor boxes), το YOLOv8 υιοθετεί μια anchor-free προσέγγιση. Επιπλέον, χρησιμοποιεί Αποσυζευγμένη Κεφαλή (Decoupled Head), διαχωρίζοντας τον υπολογισμό της πιθανότητας ύπαρξης αντικειμένου (objectness/classification) από την παλινδρόμηση των συντεταγμένων του πλαισίου (bounding box regression). Αυτός ο διαχωρισμός θεωρητικά επιτρέπει στο δίκτυο να εστιάζει καλύτερα στην ακριβή χωρική οριοθέτηση του προφίλ αλουμινίου.



Για τις ανάγκες της έρευνας δοκιμάστηκαν δύο παραλλαγές κλίμακας της αρχιτεκτονικής YOLOv8:

- **YOLOv8n (Nano):** Αποτελεί την ελαφρύτερη και ταχύτερη έκδοση, διαθέτοντας περίπου 3.2 εκατομμύρια παραμέτρους. Είναι βελτιστοποιημένη για εξαιρετικά γρήγορο συμπερασμό (inference) σε κινητές συσκευές.
- **YOLOv8s (Small):** Αποτελεί μια βαθύτερη και ευρύτερη εκδοχή (περίπου 11.2 εκατομμύρια παράμετροι). Η αύξηση του πλήθους των καναλιών (channels) προσφέρει θεωρητικά μεγαλύτερη χωρητικότητα αναπαράστασης (representational capacity) για την ανίχνευση πιο πολύπλοκων γεωμετριών, αλλά συνεπάγεται σημαντική αύξηση του υπολογιστικού χρόνου.

Η Αρχιτεκτονική YOLO26n

Το συγκεκριμένο μοντέλο ακολουθεί μια ακραία συμπιεσμένη φιλοσοφία σχεδιασμού. Στερείται του βάθους (depth) και του πλάτους (width) των παραδοσιακών δικτύων YOLO, ενσωματώνοντας δραστικά μειωμένο αριθμό συνελκτικών φίλτρων και παραμέτρων (υπολειπόμενο των 3 εκατομμυρίων). Ο σχεδιασμός αυτός στοχεύει στην επίτευξη μέγιστου ρυθμού καρέ (High FPS) σε μικροεπεξεργαστές IoT με ελάχιστη κατανάλωση μνήμης RAM, καθώς η δομή του είναι **NMS-Free** (δεν περιλαμβάνει μηχανισμό Καταστολής Μη Μέγιστων), παρακάμπτοντας έτσι ένα υπολογιστικά ακριβό βήμα μετα-επεξεργασίας.

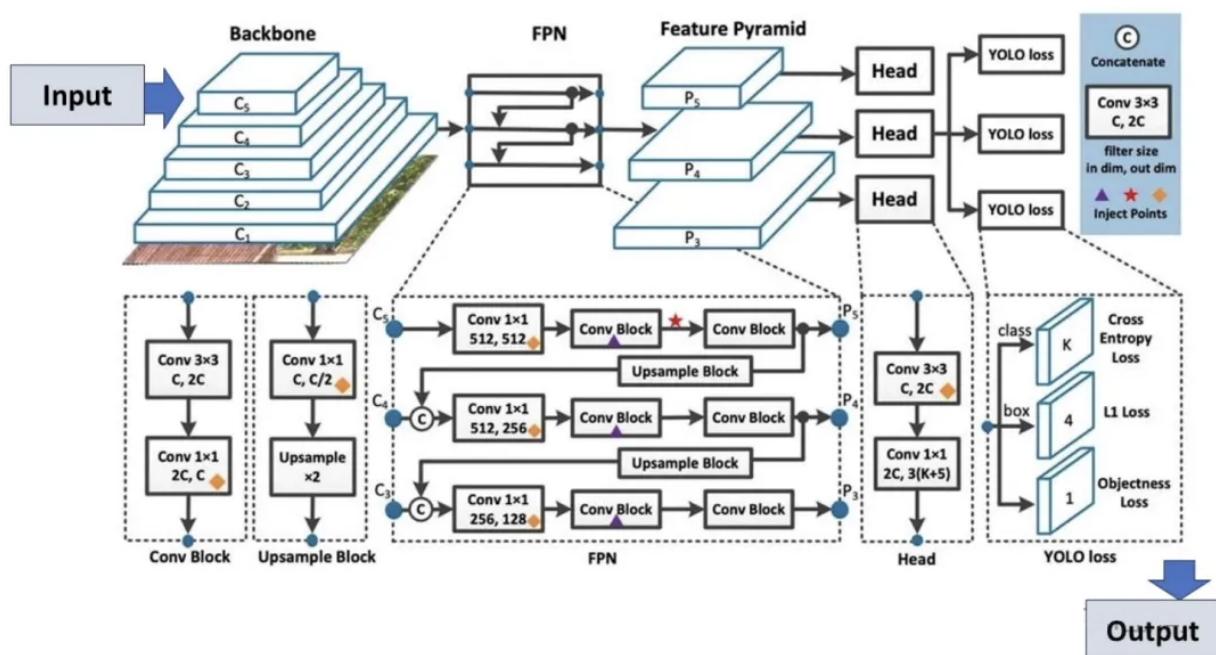
3.3.2 Αρχιτεκτονικές Μοντέλων Ταξινόμησης (Classification)

Αλγόριθμος YOLO-cls (Μοντέλο Ελαφριάς Ταξινόμησης)

Η έκδοση ταξινόμησης της οικογένειας YOLO (YOLO-cls) προτείνεται βιβλιογραφικά ως μια εξαιρετικά αποδοτική λύση για περιβάλλοντα αναπτυγμένα σε κινητές συσκευές (Android, TFLite pipelines) και μικροεπεξεργαστές IoT, όπου η διαθέσιμη μνήμη RAM και η ενεργειακή κατανάλωση υπόκεινται σε αυστηρούς περιορισμούς.

Δομική Αρχιτεκτονική: Η δομή του μοντέλου διαφοροποιείται αισθητά από τους τυπικούς ανιχνευτές αντικειμένων (object detectors) της ίδιας οικογένειας:

- **Εξαγωγή Χαρακτηριστικών (Backbone):** Διατηρείται ο βασικός κορμός εξαγωγής χαρακτηριστικών, αξιοποιώντας τα μπλοκ τύπου C2f και τα επίπεδα συνελίξεων για την αρχική ανάλυση της εικόνας και την ανίχνευση ακμών και υφών.
- **Αφαίρεση Επιπέδων Εντοπισμού:** Σε αντίθεση με την αρχιτεκτονική ανίχνευσης, ο κορμός του δικτύου και η πολύπλοκη κεφαλή εντοπισμού πλαισίων (bounding box head) παραλείπονται πλήρως.
- **Επίπεδο Ολικής Συμπύκνωσης (Global Average Pooling - GAP):** Οι παραγόμενοι χάρτες χαρακτηριστικών (feature maps) καταλήγουν σε ένα επίπεδο GAP 2D. Το επίπεδο αυτό αναλαμβάνει να συμπυκνώσει τη χωρική πληροφορία, εξάγοντας ένα ενιαίο, μονοδιάστατο διάνυσμα χαρακτηριστικών.
- **Κεφαλή Ταξινόμησης (Linear/Dense Layer):** Το δίκτυο καταλήγει σε ένα πλήρως συνδεδεμένο επίπεδο (fully connected layer). Στην παρούσα υλοποίηση, το επίπεδο αυτό διαστασιολογήθηκε στους 1.360 νευρώνες, αντιστοιχώντας ακριβώς στο πλήθος των διαθέσιμων κλάσεων του συνόλου δεδομένων εκπαίδευσης.



Αρχιτεκτονική EfficientNetV2-S (Αμιγώς Συνελικτικό Δίκτυο)

Το μοντέλο EfficientNetV2 αποτελεί μια αμιγώς συνελικτική αρχιτεκτονική (CNN), η οποία σχεδιάστηκε από την Google με στόχο τη βελτιστοποίηση της ισορροπίας μεταξύ ακρίβειας πρόβλεψης και χρόνου εκπαίδευσης. Η έκδοση Small (EfficientNetV2-S), η οποία αποτελείται από περίπου 21.4 εκατομμύρια παραμέτρους, παρουσιάζει συχνά ικανοποιητικές επιδόσεις σε απαιτητικές εφαρμογές οπτικής αναγνώρισης, όπως ο βιομηχανικός ποιοτικός έλεγχος.

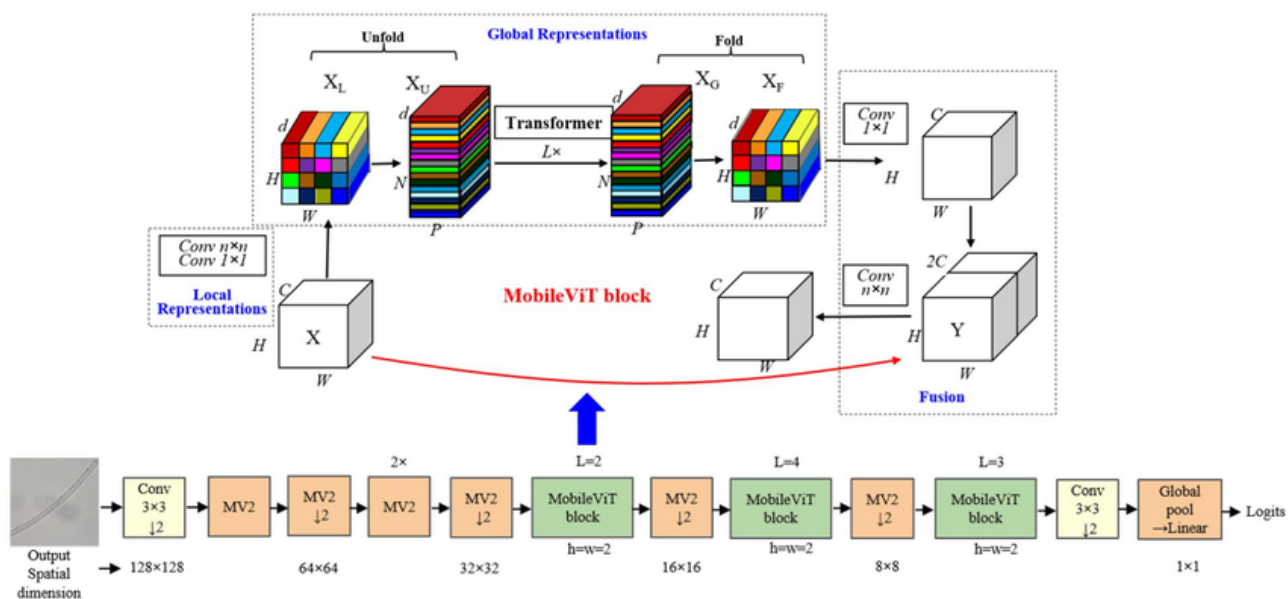
Καινοτομίες και Στάδια Αρχιτεκτονικής: Το δίκτυο οργανώνεται σε διαδοχικά στάδια, εισάγοντας κρίσιμες δομικές καινοτομίες:

- **Προοδευτική Εκμάθηση (Progressive Learning):** Η αρχιτεκτονική ενσωματώνει μηχανισμούς δυναμικής αύξησης του μεγέθους της εικόνας και της κανονικοποίησης (regularization) κατά τη διάρκεια της εκπαίδευσης, συμβάλλοντας στην ταχύτερη σύγκλιση του μοντέλου.
- **Επίπεδο Εκκίνησης (Stem):** Η επεξεργασία ξεκινά με μια βασική συνέλιξη 3×3 , συνοδευόμενη από τη συνάρτηση ενεργοποίησης Swish.
- **Μπλοκ Fused-MBConv (Αρχικά Στάδια):** Προκειμένου να αντιμετωπιστεί το φαινόμενο της υπολογιστικής συμφόρησης (memory bottleneck) που προκαλούσαν οι συνελίξεις βάθους (depthwise convolutions) στους σύγχρονους επιταχυντές, τα πρώιμα στάδια του δικτύου αξιοποιούν "συγχωνευμένες" (fused) τυπικές συνελίξεις 3×3 .
- **Μπλοκ MBConv και Μηχανισμός Προσοχής (Βαθύτερα Στάδια):** Στα βαθύτερα επίπεδα, η αρχιτεκτονική επανέρχεται στα Ανεστραμμένα Υπολειπόμενα Μπλοκ (Inverted Residual Blocks). Τα μπλοκ αυτά ενσωματώνουν έναν Μηχανισμό Συμπίεσης και Διέγερσης (Squeeze-and-Excitation - SE), ο οποίος λειτουργεί ως τοπικός μηχανισμός προσοχής (attention). Υπολογίζει δυναμικά τη σπουδαιότητα των διαφόρων καναλιών, ενισχύοντας τα κρίσιμα χαρακτηριστικά του σχήματος και καταστέλλοντας τον "θόρυβο" του υποβάθρου.
- **Τελικό Στάδιο:** Η αρχιτεκτονική ολοκληρώνεται με μια σημειακή συνέλιξη 1×1 (pointwise convolution), ακολουθούμενη από επίπεδα Pooling, Dropout και την τελική γραμμική ταξινόμηση.



3. **Fold & Fusion:** Τα τμήματα ανασυντίθενται (Fold) στην αρχική χωρική τους μορφή, συνενώνονται (Concatenation) με τα αρχικά συνελκτικά χαρακτηριστικά και υπόκεινται σε μια τελική συνέλιξη εξομάλυνσης 3×3 .

- **Κεφαλή Ταξινόμησης:** Όπως και στις προηγούμενες αρχιτεκτονικές, η πληροφορία καταλήγει σε ένα επίπεδο Global Average Pooling και στο γραμμικό επίπεδο (linear layer) για την εξαγωγή της πρόβλεψης.



3.3.3 Μετατροπή των Αμιγών CNN Μοντέλων (PyTorch σε TFLite). Βελτιστοποίηση και Μεταγλώττιση Μοντέλων για Edge Συσκευές (TFLite Compilation)

Σε αντίθεση με τα πειραματικά υβριδικά δίκτυα, η μετατροπή των αμιγώς Συνελκτικών Νευρωνικών Δικτύων (όπως η οικογένεια YOLO και το EfficientNetV2) στο πρότυπο TensorFlow Lite (TFLite) για τελική ανάπτυξη σε περιβάλλον Android, ακολούθησε μια τυποποιημένη διοχέτευση (pipeline). Δεδομένου ότι τα μοντέλα εκπαιδεύτηκαν στο οικοσύστημα του PyTorch, η άμεση μετατροπή τους σε TFLite δεν κατέστη εφικτή. Για την επίλυση του συγκεκριμένου ζητήματος, αξιοποιήθηκε το **ONNX** (Open Neural Network Exchange) ως ενδιάμεσος υπολογιστικός γράφος (Middleware Graph). Αρχικά, τα βάρη του PyTorch εξήχθησαν σε μορφή .onnx. Στη συνέχεια, μέσω του μεταγλωττιστή onnx2tf, ο γράφος μεταφράστηκε εγγενώς στο περιβάλλον **TFLite**.

Η πλέον κρίσιμη λειτουργία αυτού του σταδίου αφορούσε την αυτόματη αναδιάταξη (permutation) του σχήματος των τανυστών: ο μεταγλωττιστής μετέτρεψε επιτυχώς τη μορφή Channel-First του PyTorch (BCHW: [1, 3, 224, 224]) στη μορφή Channel-Last (NHWC: [1, 224, 224, 3]), η οποία απαιτείται αυστηρά από την αρχιτεκτονική του Android και των επεξεργαστών ARM. Εφόσον οι αρχιτεκτονικές των YOLO και EfficientNetV2 βασίζονται αποκλειστικά σε τυπικές χωρικές συνελίξεις (Standard/Depthwise Convolutions) και επίπεδα Pooling, ο χρόνος εκτέλεσης (runtime) του TFLite υποστηρίζει εγγενώς το σύνολο των μαθηματικών τους πράξεων. Η διαδικασία αυτή οδήγησε στην παραγωγή εξαιρετικά συμπίεσμων μοντέλων, απαλλαγμένων από τα περιττά βάρη των αλγορίθμων

βελτιστοποίησης (optimizer stripping), τα οποία καθίστανται ικανά να εκτελούν συμπερασμούς (inference) σε πραγματικό χρόνο στο Edge, παρουσιάζοντας ελάχιστο ενεργειακό και υπολογιστικό αποτύπωμα.

3.4 Στρατηγική Εκπαίδευσης και Επιλογή Υπερπαραμέτρων

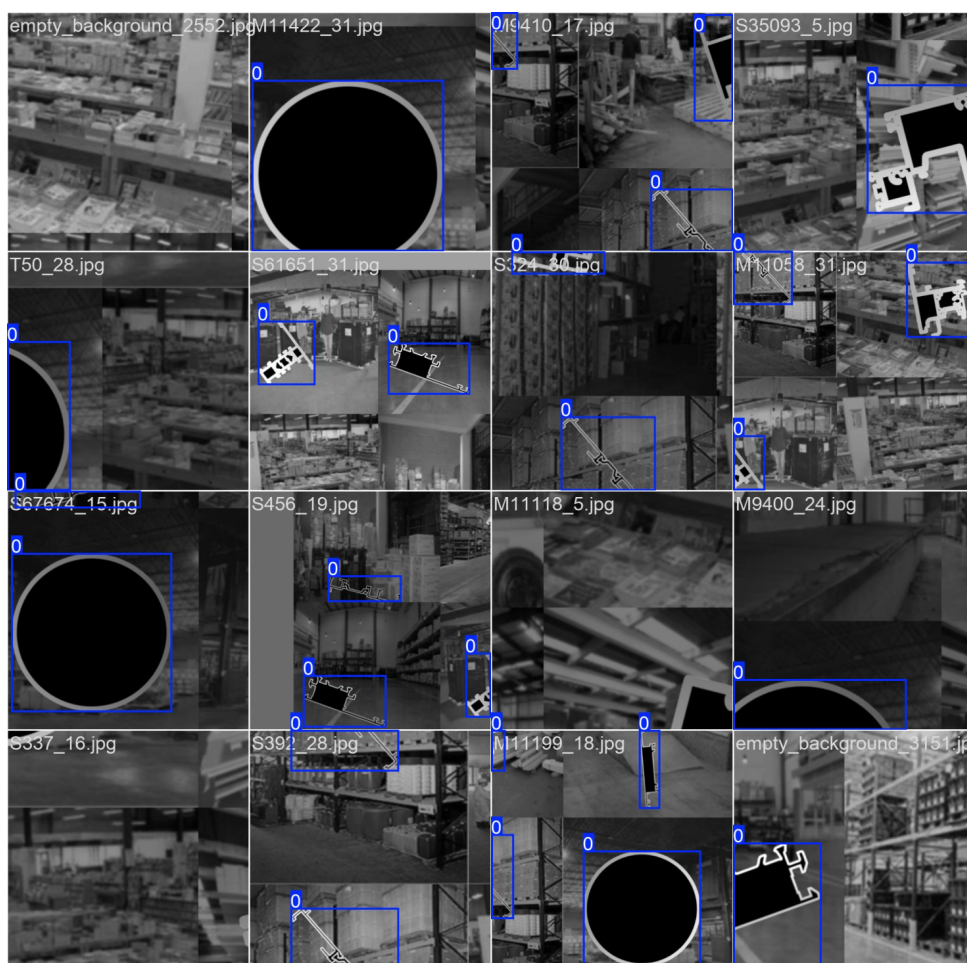
Το σύνολο των πειραμάτων εκπαίδευσης διεξήχθη επί του συνθετικού συνόλου δεδομένων "DataSetSilver", το οποίο αποτελείται από 116.200 εικόνες. Η αρχιτεκτονική της διοχέτευσης (pipeline) επέβαλε τον διαχωρισμό της εκπαίδευσης σε δύο διακριτές φάσεις: την εκπαίδευση των Ανιχνευτών (Detectors) για τον χωρικό εντοπισμό, και την εκπαίδευση των Ταξινομητών (Classifiers) για τη γεωμετρική αναγνώριση. Σε όλα τα μοντέλα τέθηκε μέγιστο όριο 100 εποχών (epochs) με μηχανισμό Πρόωρου Τερματισμού (Early Stopping, patience=20) με στόχο την αποτροπή της υπερεκπαίδευσης (overfitting) και τη διασφάλιση της βέλτιστης ικανότητας γενίκευσης. Η παρακολούθηση των μετρικών πραγματοποιήθηκε μέσω της πλατφόρμας **ClearML**.

3.4.1 Φάση 1η: Εκπαίδευση Ανιχνευτών (Object Detectors)

Αξιολογήθηκαν τρεις αρχιτεκτονικές εντοπισμού: το ελαφρύ YOLOv8n, το YOLO26n, και το βαθύτερο YOLOv8s. Η εκπαίδευση πραγματοποιήθηκε σε ανάλυση εικόνας $\text{imgsz}=512$, η οποία κρίθηκε ως η βέλτιστη τομή μεταξύ της διατήρησης χωρικής λεπτομέρειας και της ταχύτητας επεξεργασίας. Για τη διαχείριση της διαθέσιμης μνήμης γραφικών (VRAM), υιοθετήθηκε δυναμικό μέγεθος δέσμης (batch size). Ειδικότερα, τα μοντέλα YOLOv8n και YOLO26n εκπαιδεύτηκαν με $\text{batch}=128$, ενώ το YOLOv8s περιορίστηκε στο $\text{batch}=64$ προς αποφυγή σφαλμάτων εξάντλησης μνήμης (Out-Of-Memory).

Στρατηγική Επαύξησης Δεδομένων (Data Augmentation): Προκειμένου να ενισχυθεί η ανθεκτικότητα του ανιχνευτή σε πολύπλοκα βιομηχανικά περιβάλλοντα, εφαρμόστηκαν επιθετικές τεχνικές επαύξησης δεδομένων:

- **Mosaic (1.0):** Ενεργοποιήθηκε καθολικά, συνθέτοντας τέσσερις διαφορετικές εικόνες σε μία, εξαναγκάζοντας το μοντέλο να εντοπίζει προφίλ σε μη αναμενόμενες κλίμακες και πυκνές διατάξεις.
- **Copy-Paste (0.3) & Scale (0.5):** Ενίσχυσαν την ικανότητα του μοντέλου να διαχειρίζεται μερικές επικαλύψεις (occlusions) και οπτικές παραμορφώσεις κλίμακας.
- **Φωτομετρική Αλλοίωση (HSV):** Εισήχθησαν δυναμικές διακυμάνσεις στην απόχρωση ($\text{hsv_h}=0.015$), τον κορεσμό ($\text{hsv_s}=0.7$) και τη φωτεινότητα ($\text{hsv_v}=0.4$), προσομοιώνοντας τις ενδεχόμενες μεταβολές φωτισμού στο εργοστασιακό περιβάλλον.



3.4.2 Φάση 2η: Εκπαίδευση Ταξινομητών (Classifiers)

Η δεύτερη φάση αφορούσε την εκπαίδευση των εξειδικευμένων ταξινομητών (YOLO26n-cls, EfficientNetV2, MobileViT). Δεδομένου ότι τα μοντέλα αυτά λαμβάνουν ως είσοδο τις ήδη περικομμένες εικόνες (crops) από τον ανιχνευτή, η ανάλυση εισόδου μειώθηκε βελτιστοποιημένα στο $\text{imgsz}=256$, αυξάνοντας την ταχύτητα συμπερασμού.

Κρίσιμες Επιλογές Υπερπαραμέτρων:

- **Εξομάλυνση Ετικετών (Label Smoothing = 0.1):** Λόγω της παρουσίας τεχνητού θορύβου στο σύνολο "DataSetSilver", υφίστατο ο κίνδυνος υπερβολικής βεβαιότητας (overconfidence) του δικτύου. Η εφαρμογή label smoothing (10%) αποτρέπει θεωρητικά το δίκτυο από το να αποδίδει απόλυτη πιθανότητα (1.0) σε μια κλάση, ωθώντας το να καταναίμει την πιθανότητα μαθαίνοντας πιο ανθεκτικά και γενικευμένα χαρακτηριστικά.
- **Αναπροσαρμογή Ρυθμού Εκμάθησης (Learning Rate - LR):** Το EfficientNetV2, ως σταθερό συνελκτικό δίκτυο, εκπαιδεύτηκε με ρυθμό $\text{lr}=1\text{e-}3$, επιτυγχάνοντας ταχεία σύγκλιση. Αντιθέτως, στα μοντέλα YOLO26n-cls και MobileViT, ο ρυθμός μειώθηκε αυστηρά στο $\text{lr}=1\text{e-}4$. Η επιλογή αυτή (σε συνδυασμό με τη χρήση ακρίβειας FP32 αντί για AMP) κρίθηκε

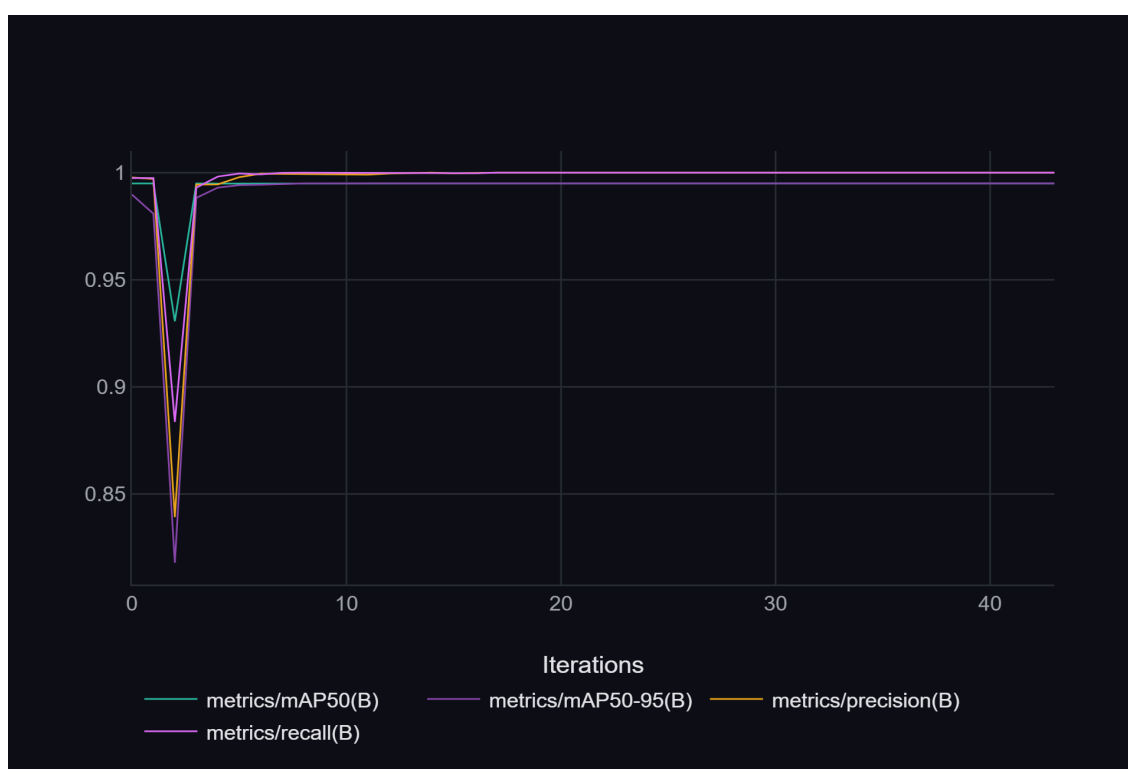
απαραίτητη για την αποτροπή της κατάρρευσης των κλίσεων (gradient explosion), αντιμετωπίζοντας την εγγενή αστάθεια των αρχιτεκτονικών Transformer στα πρώιμα στάδια της εκπαίδευσης.

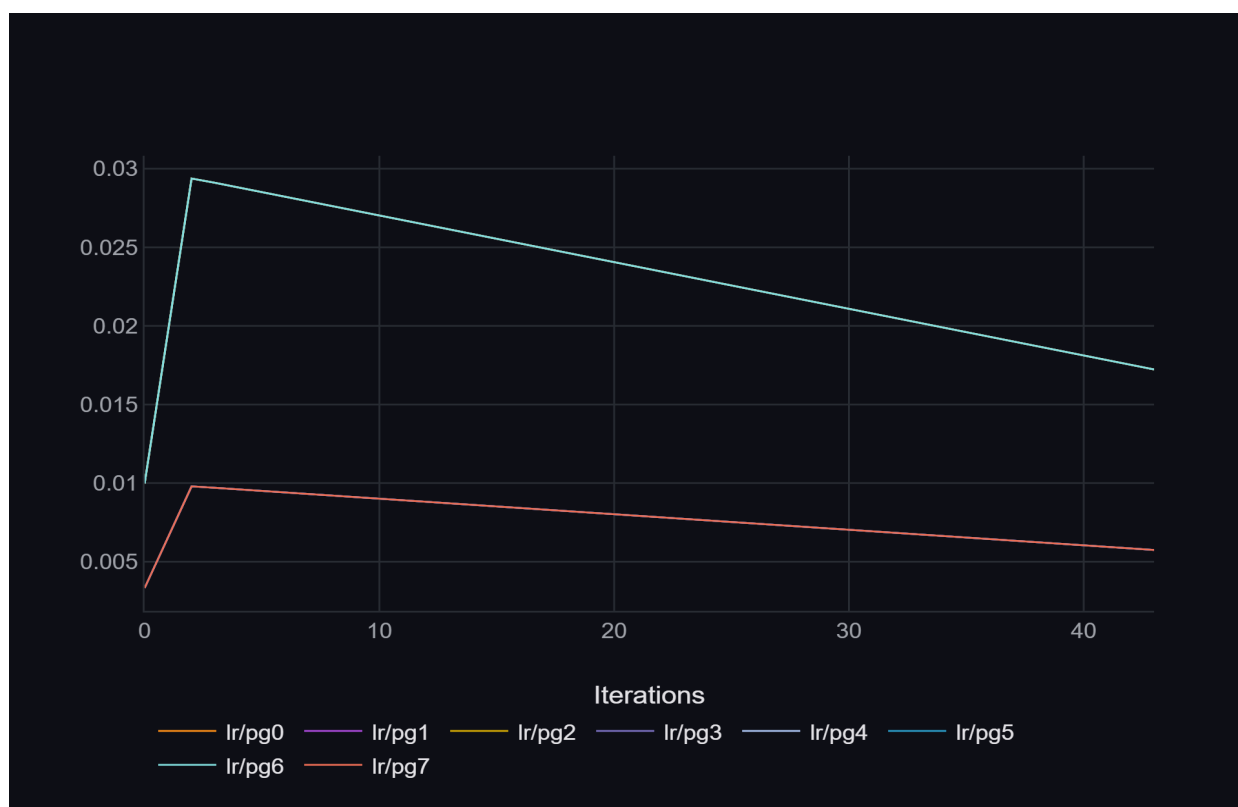
3.4.3 Αξιολόγηση και Αποτελέσματα Εκπαίδευσης Ανιχνευτών

Η εκπαίδευση πραγματοποιήθηκε σε υπολογιστική πλατφόρμα εξοπλισμένη με NVIDIA GeForce RTX 3090 (24GB VRAM). Τα αρχεία καταγραφής (logs) επιβεβαίωσαν τη μαθηματική σύγκλιση των μοντέλων.

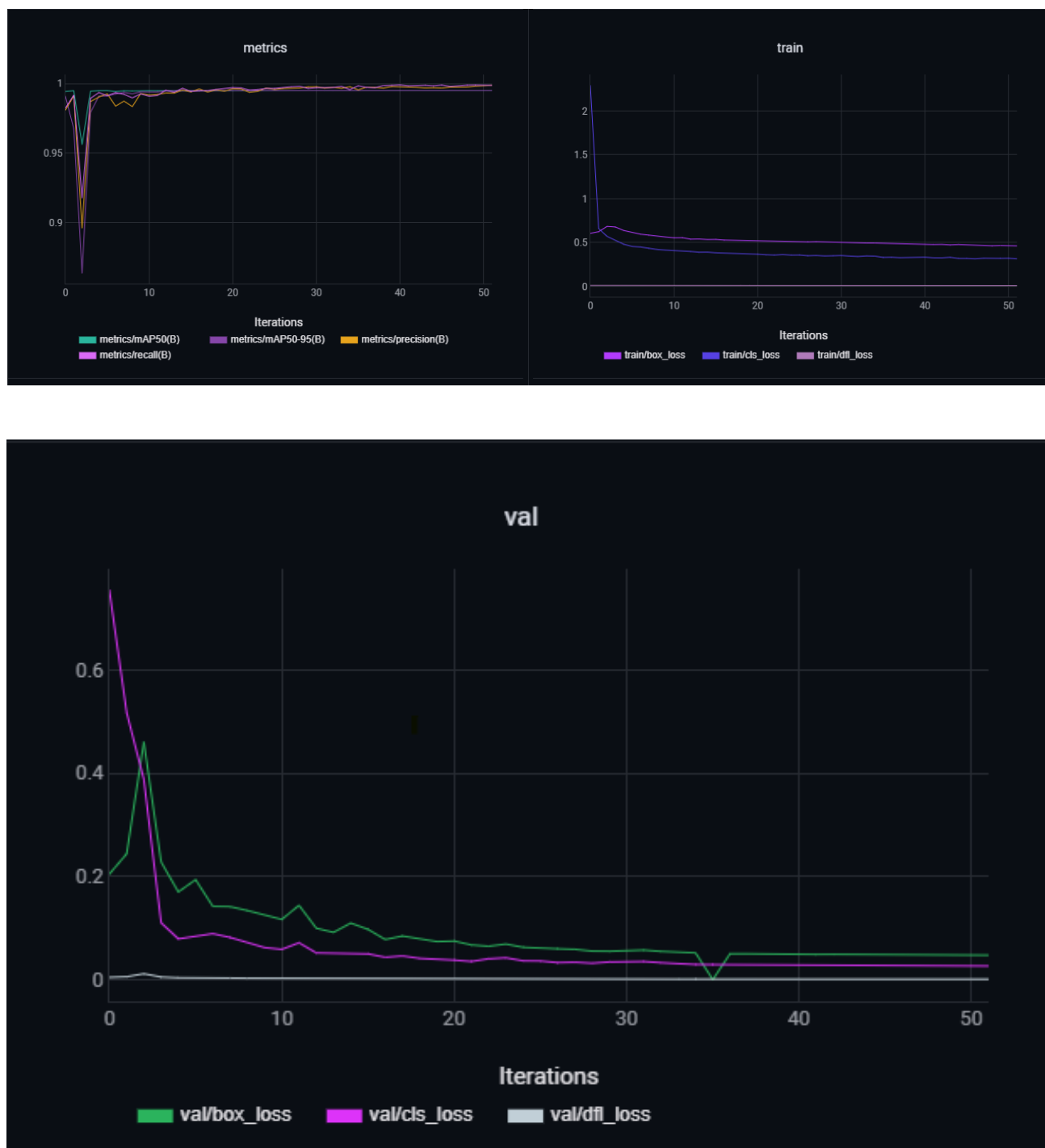
Ανάλυση Σύγκλισης YOLOv8n (Nano):

Η εκπαίδευση ολοκληρώθηκε σε 2,14 ώρες. Ο μηχανισμός Early Stopping διέκοψε τη διαδικασία στην Εποχή 44, καθώς το δίκτυο είχε ήδη επιτύχει τη βέλτιστη ικανότητα γενίκευσης στην Εποχή 24 (best.pt). Από την Εποχή 25 και έπειτα, παρά την οριακή μείωση των σφαλμάτων εκπαίδευσης (box_loss, cls_loss), οι μετρικές επικύρωσης (Validation Metrics) σταθεροποιήθηκαν, αποτρέποντας την απομνημόνευση (overfitting) του συνθετικού θορύβου. Στο σύνολο επικύρωσης (12.039 εικόνες), το μοντέλο επέτυχε απόλυτες βαθμολογίες (Precision: 1.0, Recall: 1.0, mAP@50: 0.995, mAP@50-95: 0.995). Η ταχύτητα εξαγωγής συμπεράσματος καταγράφηκε στα 0.4ms ανά εικόνα (στο περιβάλλον GPU), υποδεικνύοντας ότι το συμπιεσμένο μοντέλο (6.2MB) διαθέτει τις προδιαγραφές για ανάπτυξη σε επίπεδα πραγματικού χρόνου (>30 FPS) σε συσκευές Edge.



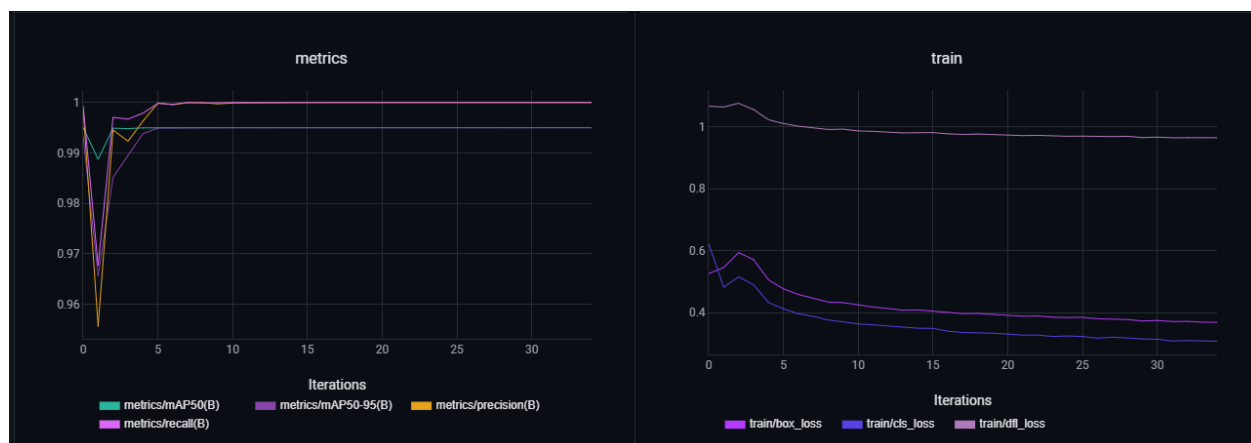
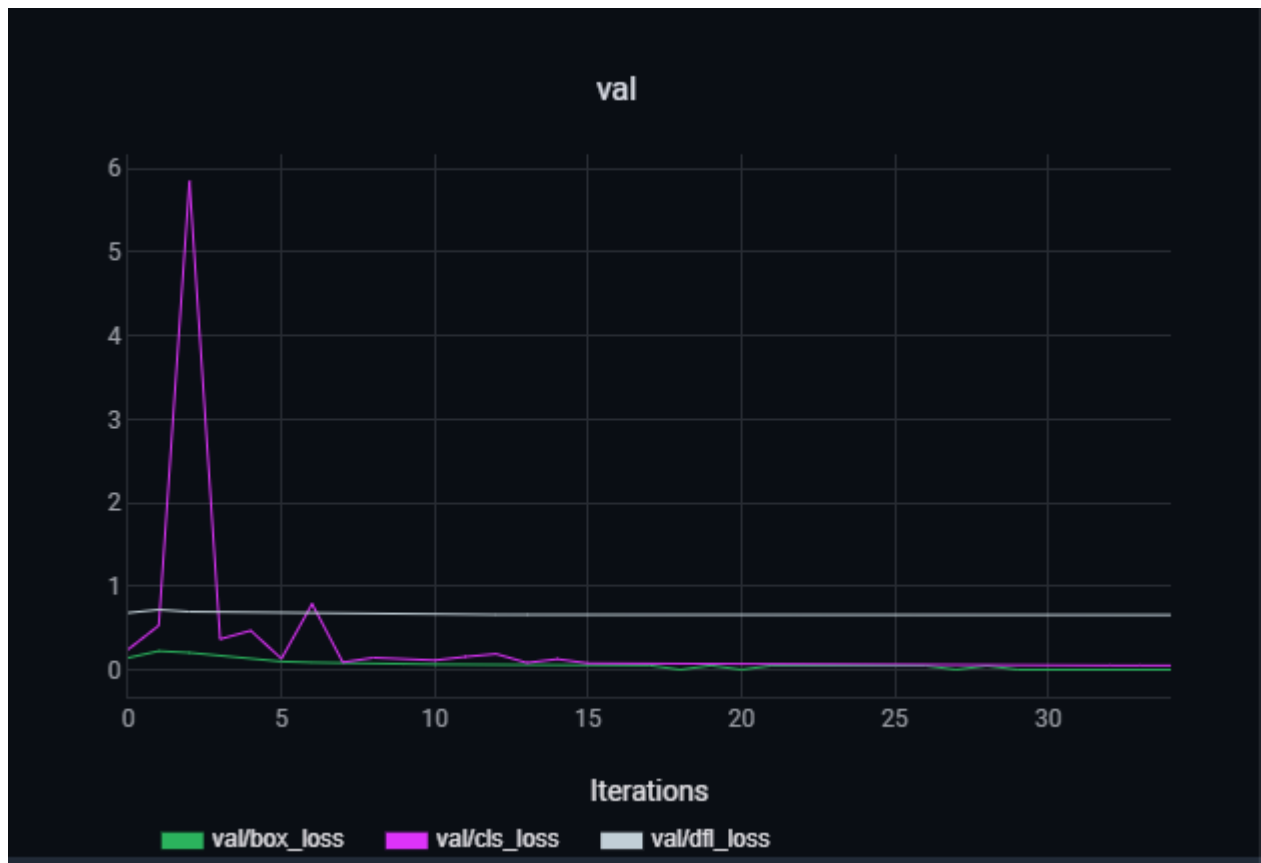


Ανάλυση Εναλλακτικών Ανιχνευτών:YOLO26n: Ολοκλήρωσε 52 εποχές σε 3,4 ώρες. Παρά το μικρότερο υπολογιστικό του βάρος, απαίτησε περισσότερες εποχές για να επιτύχει παρόμοια επίπεδα γενίκευσης (mAP50-95: 0.995), καταγράφοντας ωστόσο αντίστοιχα υψηλή ταχύτητα συμπερασμού (0.4ms).



YOLOv8s: Ως βαθύτερο δίκτυο, εμφάνισε ταχύτερη απομνημόνευση της γεωμετρίας. Η εκπαίδευση διακόπηκε στην 35η εποχή (βέλτιστα βάρη στην Εποχή 15) διατηρώντας το mAP στο 0.995.

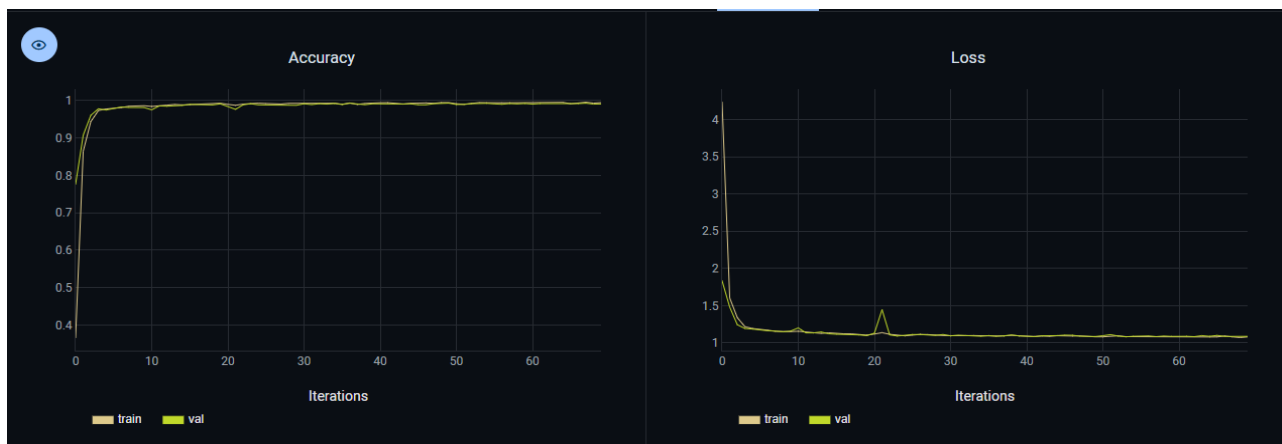
Διαπιστώνεται ότι για το συγκεκριμένο πρόβλημα, η αύξηση της πολυπλοκότητας του ανιχνευτή προσφέρει φθίνουσα απόδοση (diminishing returns) εις βάρος του υπολογιστικού χρόνου (0.8ms).



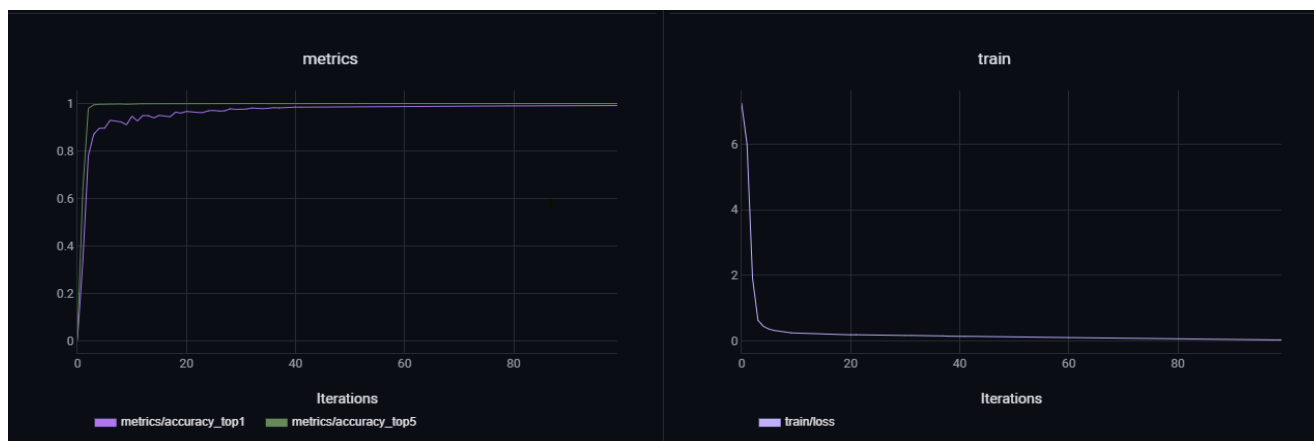
3.4.4 Αξιολόγηση Σταδίου Ταξινόμησης (Classifiers)

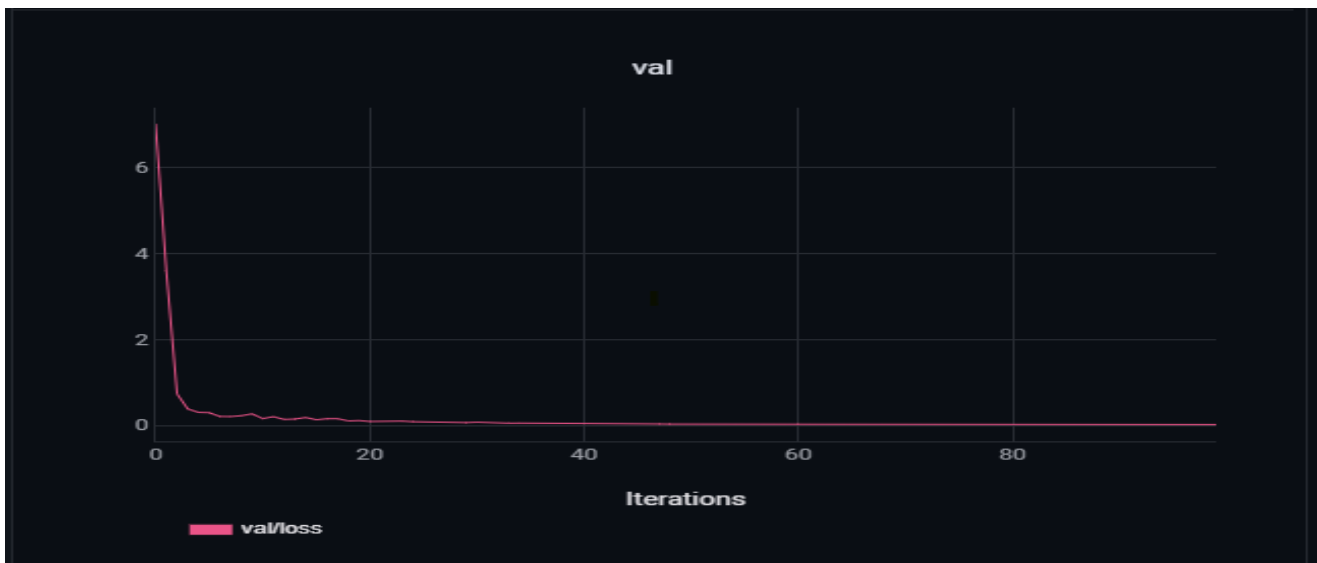
Οι τρεις ταξινομητές αξιολογήθηκαν στο σύνολο "DataSilver" (1.360 κλάσεις).

1. **YOLO26n-cls:** Ολοκλήρωσε 100 εποχές επιτυγχάνοντας ακρίβεια (top1_acc) 0.992, υποδεικνύοντας ότι η υψηλή ταχύτητα συμπερασμού (0.1ms) δύναται να συνοδεύεται από υψηλή ακρίβεια σε κανονικοποιημένες περικοπές (crops).

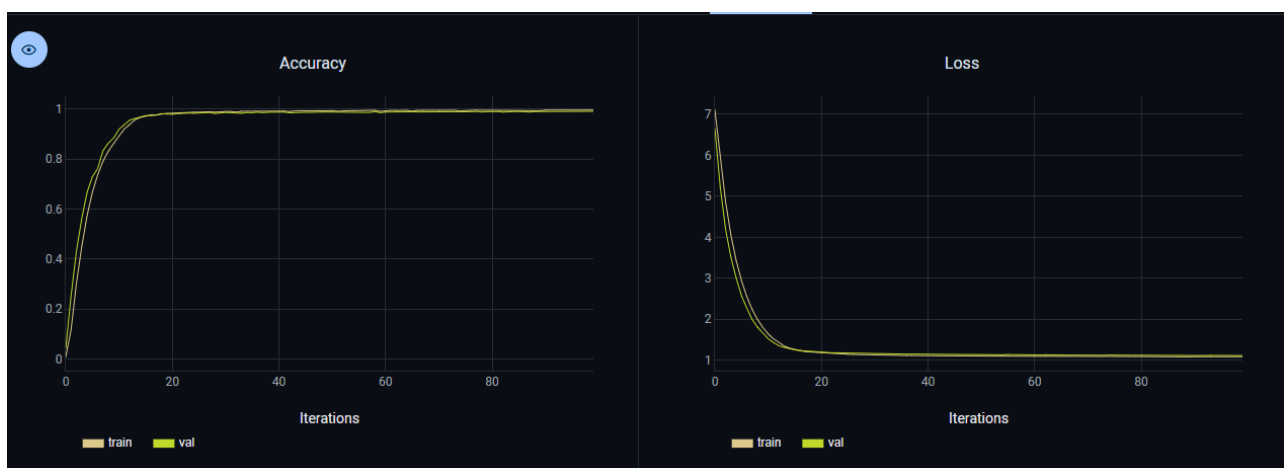


2. **EfficientNetV2:** Παρουσίασε την υψηλότερη ακρίβεια επικύρωσης (Val Acc: 0.9926). Η χρήση Label Smoothing διατήρησε το σφάλμα εκπαίδευσης σε ελεγχόμενα επίπεδα (Loss $\approx$ 1.08), προστατεύοντας το δίκτυο από υπερβολική βεβαιότητα.





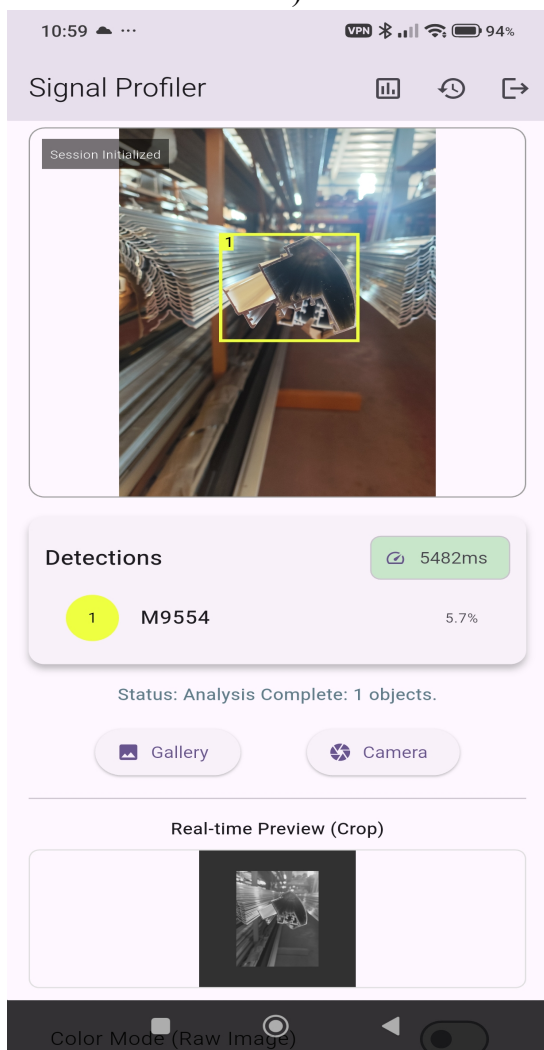
3 MobileViT: Παρά την αναγκαία χρήση χαμηλού ρυθμού εκμάθησης και την απουσία AMP, κατόρθωσε να επιτύχει ικανοποιητική ακρίβεια (Val Acc: 0.9909) μέσω του μηχανισμού Self-Attention. Ωστόσο, ο αυξημένος χρόνος εκπαίδευσης και οι απαιτήσεις υλοποίησης στο Edge καθιστούν τα CNNs πιο πρακτικές επιλογές για τη δεδομένη εφαρμογή.



3.5 Σχεδιασμός και την ανάπτυξη του “Κινητού Έργαστηριου(“Signal Profiler”/Flutter-Dart)

Όπως περιγράφεται στους ερευνητικούς στόχους, η μετάβαση από ένα θεωρητικό μοντέλο μηχανικής όρασης σε μια πρακτική βιομηχανική εφαρμογή απαιτεί έναν ισχυρό μηχανισμό συλλογής, δοκιμής και εκτέλεσης. Το παρόν κεφάλαιο αναλύει τον σχεδιασμό και την ανάπτυξη του "Signal Profiler", μιας ειδικά κατασκευασμένης διαπλατφορμικής (cross-platform) εφαρμογής Flutter. Η εφαρμογή αυτή, η οποία αναφέρεται στην παρούσα έρευνα και ως «Κινητό Έργαστήριο», γεφυρώνει το χάσμα μεταξύ των φυσικών βιομηχανικών πόρων και των ψηφιακών μοντέλων τεχνητής νοημοσύνης (Sim2Real). Η αναγκαιότητα του συστήματος προκύπτει από τις εγγενείς προκλήσεις του βιομηχανικού περιβάλλοντος. Εξυπηρετεί δύο πρωταρχικούς ρόλους:

1. **Μηχανή Συμπερασμού Εντός Συσκευής (On-device Inference Engine):** Εκτελεί ταχύτατη ανίχνευση και ταξινόμηση αντικειμένων δύο σταδίων (two-stage object detection and classification) σε τοπικό επίπεδο.



2. **Συλλέκτης Δεδομένων (Data Harvester):** Συλλέγει δεδομένα εκπαίδευσης υψηλής πιστότητας, αξιοποιώντας την αλληλεπίδραση «ανθρώπου στον βρόχο» (human-in-the-loop) για τη δημιουργία συνόλων δεδομένων επόμενης γενιάς.

3.5.1 Αρχιτεκτονικός Σχεδιασμός και Ταυτοχρονισμός (Concurrency)

Οι κινητές συσκευές (Edge Devices) διαθέτουν περιορισμένο εύρος ζώνης στο κύριο νήμα εκτέλεσης (main thread). Για τη διατήρηση της ομαλότητας της Διεπαφής Χρήστη (UI) κατά τη διάρκεια βαρέων υπολογιστικών εργασιών, η εφαρμογή υλοποιεί ένα αυστηρό Μοτίβο Απομονωμένου Εργάτη (Isolate-Worker Pattern), διαχωρίζοντας τις διεργασίες:

- **Απομόνωση Προεπισκόπησης (Real-time Preview Isolate):** Δημιουργεί ένα αποκλειστικό νήμα για τα φίλτρα της βιβλιοθήκης OpenCV (π.χ. αλγόριθμος

Otsu, εξομάλυνση υφής), διασφαλίζοντας ότι η «ζωντανή» ανατροφοδότηση της κάμερας δεν προκαλεί καθυστερήσεις (jank) στη διεπαφή.

- **Απομόνωση Συμπερασμού (Inference Isolate):** Μεταφέρει τη βαριά εκτέλεση των μοντέλων TensorFlow Lite (TFLite) και τους εγγενείς (native) χειρισμούς μνήμης C++ σε ένα νήμα παρασκηνίου.
- **Απομόνωση Διατήρησης (Persistence Isolate):** Διαχειρίζεται την απαιτητική σειριοποίηση/αποσειριοποίηση JSON για την καταγραφή ιστορικού και επιδόσεων (benchmark logging).

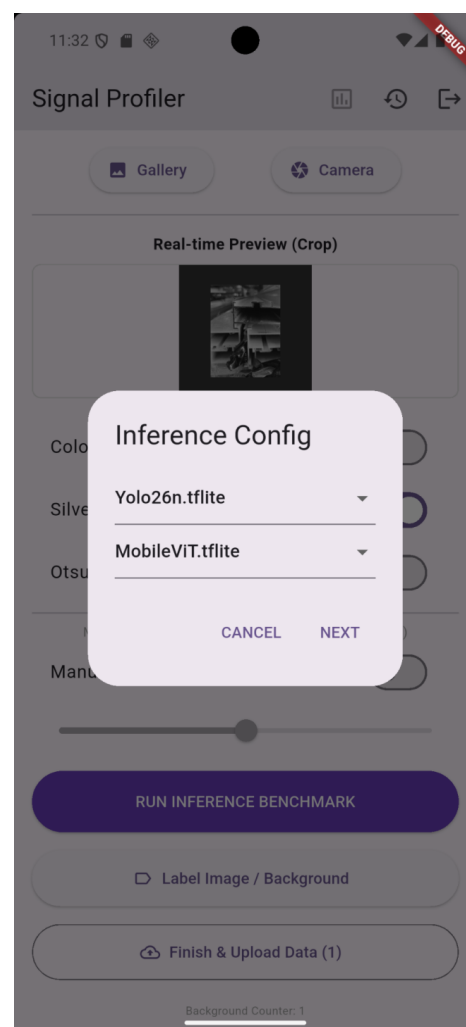
Σε επίπεδο οργάνωσης κώδικα, υιοθετείται μια Αρχιτεκτονική Προσανατολισμένη στις Υπηρεσίες (Service-Oriented Architecture - SOA), διαχωρίζοντας τις λειτουργίες σε ανεξάρτητες υπηρεσίες, όπως η **AuthService** για την ασφαλή διαχείριση ταυτοτήτων, η **FirestoreStorageService** για τις αποστολές στο νέφος, και η **ImageUtils**, η οποία λειτουργεί ως γέφυρα υψηλής απόδοσης μεταξύ της μνήμης της γλώσσας Dart και της εγγενούς μνήμης C++.

3.5.2 Η Διοχέτευση Μηχανικής Όρασης (The Computer Vision Pipeline)

Η καρδιά της εφαρμογής είναι μια μαθηματικά συγχρονισμένη διοχέτευση δεδομένων (pipeline) η οποία συμμορφώνεται αυστηρά με τέσσερις κρίσιμους μηχανικούς περιορισμούς, κατανεμημένους στα δύο στάδια ανάλυσης.

Στάδιο 1: Χωρικός Εντοπισμός (Object Detector)

- **Συμβατότητα Μοντέλων (Model Compatibility):** Η εφαρμογή ενσωματώνει έναν καθολικό αποκωδικοποιητή (Universal Decoder), ικανό να υποστηρίζει τόσο τις τυπικές εξόδους του YOLOv8 (ακατέργαστες συντεταγμένες cx/cy/w/h) όσο και προσαρμοσμένες υπερ-ελαφριές αρχιτεκτονικές (όπως το YOLO26n) με ενσωματωμένο μηχανισμό Καταστολής Μη Μέγιστων (Baked NMS).
- **Γεωμετρική Διατήρηση (Scale-Preserving Padding):** Σε αντίθεση με τις προσεγγίσεις που παραμορφώνουν την

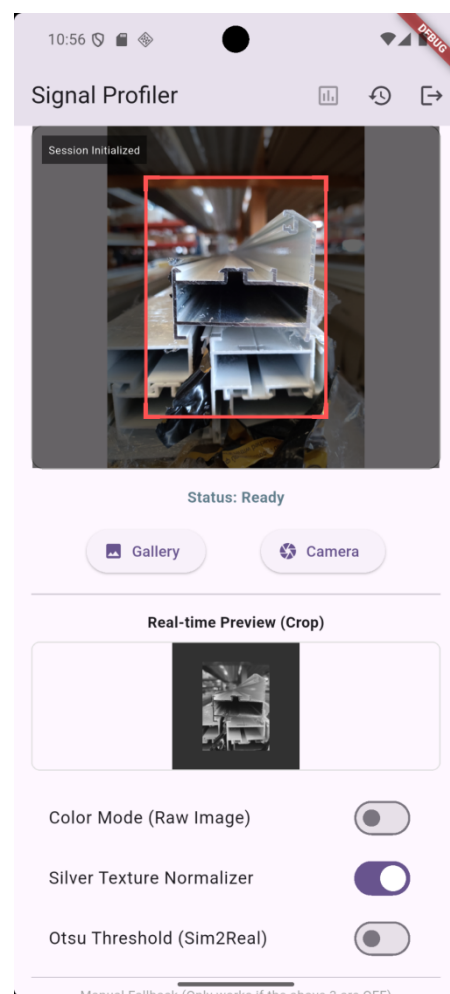


εικόνα για να ταιριάζει στις διαστάσεις του τανυστή εισόδου (tensor stretching), το σύστημα υπολογίζει έναν τετράγωνο καμβά "Παγκόσμιας Κλίμακας" (World-Scale). Αυτή η προσέγγιση διατηρεί ακέραιη τη γεωμετρική ταυτότητα των προφίλ, διασφαλίζοντας ότι τα κυκλικά χαρακτηριστικά δεν παραμορφώνονται σε οβάλ κατά την ανίχνευση.

- **Ευθυγράμμιση Χρωματικού Χώρου (BGR/RGB Alignment):** Οι αισθητήρες και η βιβλιοθήκη OpenCV τείνουν να επεξεργάζονται τα κανάλια χρώματος με διάταξη BGR. Το Pipeline διορθώνει αυτή τη σειρά εγγενώς σε RGB πριν η εικόνα εισέλθει στο μοντέλο, αποτρέποντας την καταστροφική μετατόπιση χρωμάτων η οποία θα μηδένιζε την ικανότητα αναγνώρισης του συστήματος.

Στάδιο 2: Εκλέπτυνση Χαρακτηριστικών και Ταξινόμηση (Classifier)

- **Ακριβής Περικοπή (Tight Cropping):** Όταν ο ανιχνευτής (Detector) εντοπίσει ένα προφίλ, η τελική περικοπή (crop) δεν εξάγεται από τη συμπιεσμένη εικόνα εισόδου του πρώτου μοντέλου, αλλά επαναπροβάλλεται μαθηματικά ώστε να εξαχθεί απευθείας από τα αρχικά δεδομένα υψηλής ανάλυσης του αισθητήρα της κάμερας.
- **Πυκνότητα Πληροφορίας (Information Density):** Το εξαγόμενο τμήμα προσαρμόζεται (padding) με τέτοιο τρόπο ώστε το αντικείμενο ενδιαφέροντος να καταλαμβάνει τουλάχιστον το 80% του τανυστή εισόδου. Αυτό μεγιστοποιεί τη χωρική ανάλυση των χαρακτηριστικών που τροφοδοτούν στους ταξινομητές (όπως το EfficientNetV2 ή το MobileViT).
- **Κανονικοποίηση (ImageNet Normalization):** Εφαρμόζεται αυστηρά η τυπική μαθηματική φόρμουλα κανονικοποίησης $(\text{Pixel}/255.0 - \text{Mean})/\text{Std}$ σε κάθε κανάλι ξεχωριστά, εξασφαλίζοντας απόλυτη ταύτιση της κινητής συσκευής με το περιβάλλον στο οποίο εκπαιδεύτηκαν τα μοντέλα.



3.5.3 Συλλογή Δεδομένων και Ακεραιότητα Συνόλων (Data Harvesting & Dataset Integrity)

Διεπαφή Λήψης και Προ-επεξεργασίας

Σε αντίθεση με μια απλή εφαρμογή κάμερας, η οθόνη αυτή λειτουργεί ως το "φορητό εργαστήριο" εικόνας. Παρέχει στον χρήστη εργαλεία για να επεξεργαστεί την εικόνα πριν την αποθήκευση, προσομοιώνοντας τον τρόπο με τον οποίο θα την

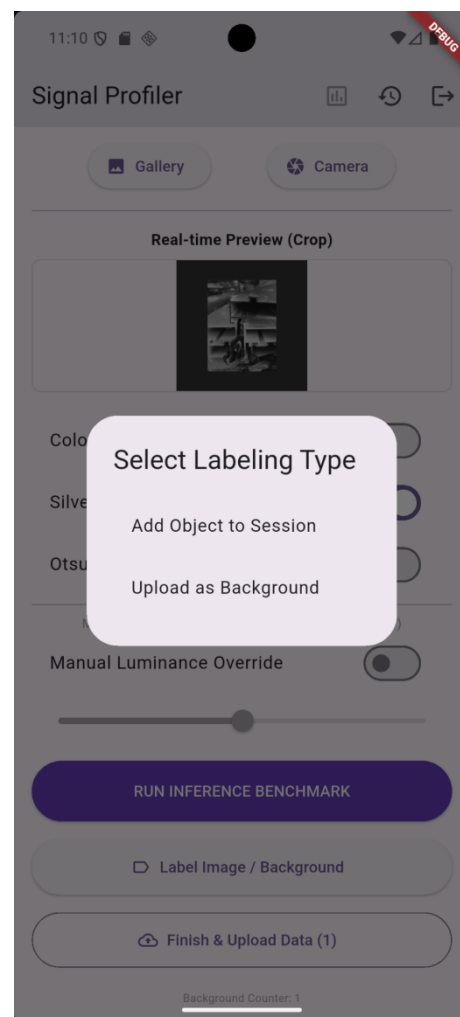
"έβλεπε" ένας αλγόριθμος.

Η διαδικασία λήψης ακολουθεί τα εξής βήματα:

1. **Σκόπευση και Λήψη:** Ο χρήστης χρησιμοποιεί την κάμερα ώστε να τραβήξει μια φωτογραφία του προφίλ αλουμινίου.
2. **Προ επισκόπηση:** Η εικόνα "παγώνει" στην οθόνη, επιτρέποντας τον ποιοτικό έλεγχο (ευκρίνεια, φωτισμός).
3. **Εφαρμογή Φίλτρων:** Ο χρήστης μπορεί να επιλέξει μεταξύ τριών καταστάσεων προβολής:
 - **Κανονικοποιητής Ασημένιας Υφής (Silver Texture Normalizer):** Εφαρμόζει τον αλγόριθμο CLAHE για την ενίσχυση της τοπικής αντίθεσης, προσομοιώνοντας τις συνθήκες υπό τις οποίες εκπαιδεύτηκε το μοντέλο DataSilver.
 - **Κατώφλι Otsu (Otsu Thresholding):** Υπολογίζει αυτόματα το βέλτιστο επίπεδο δυαδικοποίησης για τον διαχωρισμό του προφίλ από το φόντο, εξαλείφοντας τις διακυμάνσεις του φωτισμού.
 - **Χειροκίνητη Φωτεινότητα & Κατώφλι:** Παρέχει στον χρήστη πλήρη έλεγχο μέσω sliders για την προσαρμογή της εικόνας σε ακραίες βιομηχανικές συνθήκες.

Η ενσωμάτωση των φίλτρων αυτών στην εφαρμογή επιτρέπει την ευθυγράμμιση των δεδομένων εισόδου με τις μαθηματικές απαιτήσεις των μοντέλων βαθιάς μάθησης, μειώνοντας το χάσμα Sim2Real. Η εφαρμογή είναι ειδικά βελτιστοποιημένη για τη συλλογή βιομηχανικών δεδομένων:

- **Σημασιολογική Ονοματολογία Αρχείων (Semantic Filename Generation):** Το σύστημα ονομάζει αυτόματα τα αρχεία με βάση μια ταξινομημένη λίστα των μοναδικών κλάσεων που εντοπίστηκαν κατά τη συνεδρία, προσθέτοντας μια χρονοσφραγίδα (π.χ., ClassA_ClassB_1714912345.jpg). Αυτό καθιστά τα παραγόμενα σύνολα δεδομένων αυτο-περιγραφικά (self-describing).
- **Κανονικοποίηση Συντεταγμένων:** Οι σχολιασμοί (annotations) αποθηκεύονται σε μορφή JSON χρησιμοποιώντας κανονικοποιημένες συντεταγμένες (από 0.0 έως 1.0). Με αυτόν τον τρόπο, τα παραγόμενα σύνολα δεδομένων καθίστανται



ανεξάρτητα από την ανάλυση (resolution-independent) και είναι άμεσα έτοιμα για εκπαίδευση σε οποιοδήποτε υλικό (hardware).

- **Απόδοση "Κουτιών Φάντασμα" (Ghost Box Rendering):** Υλοποιείται ένα "Οπτικό Ίχνος Εργασίας" (Visual Paper Trail) όπου τα ήδη επισημασμένα αντικείμενα παραμένουν ορατά στην οθόνη ως ημιδιαφανείς επικαλύψεις, αποτρέποντας διπλοεγγραφές και πλεονάζουσα εργασία από τον χειριστή.

3.5.4 Βέλτιστες Μηχανικές Πρακτικές και Βελτιστοποιήσεις

Η μετάβαση σε περιβάλλοντα Edge Computing απαιτεί βαθιά κατανόηση της διαχείρισης πόρων:

- **Εγγενής Διαχείριση Πόρων (Native Resource Management)**

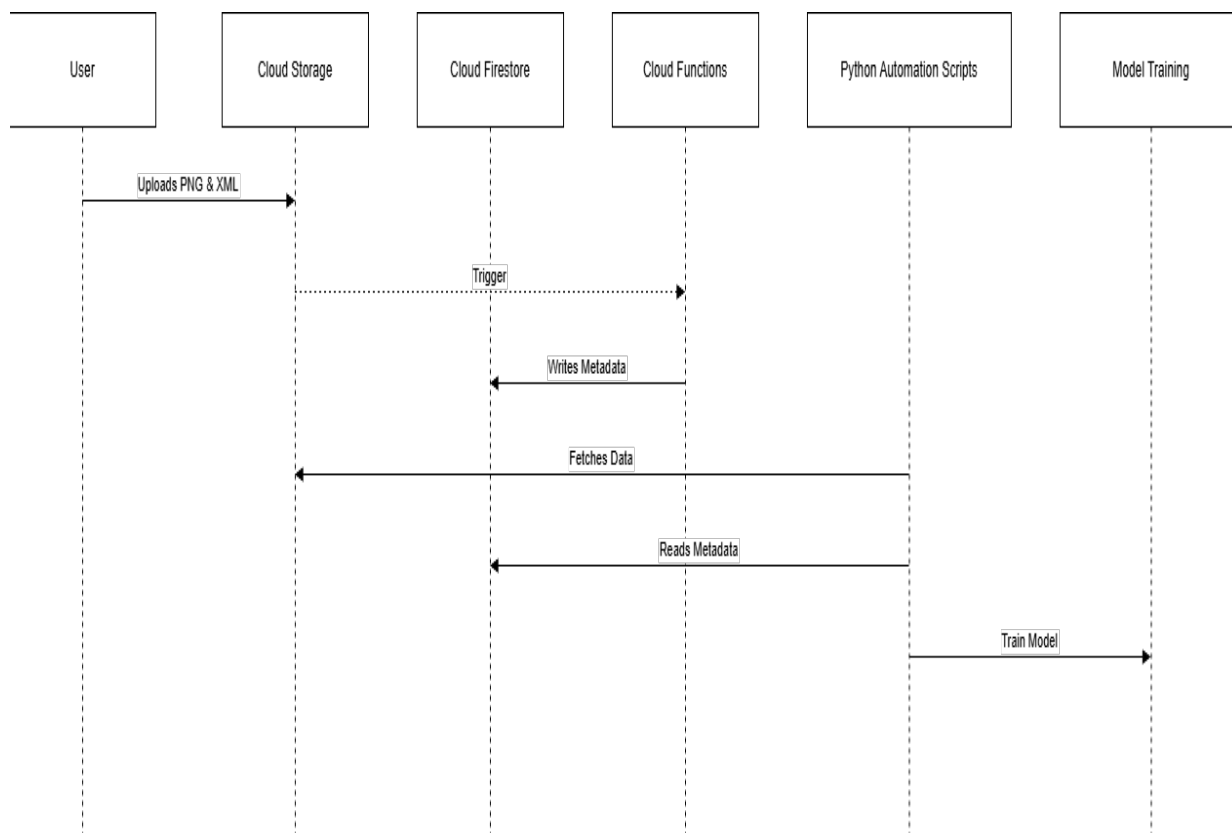
Η εφαρμογή αξιοποιεί τις συνδέσεις FFI της βιβλιοθήκης `opencv_dart`. Προκειμένου να αποφευχθούν διαρροές εγγενούς μνήμης (native memory leaks), οι οποίες είναι αόρατες στον Συλλέκτη Απορριμμάτων (Garbage Collector) της γλώσσας Dart, η υλοποίηση ακολουθεί ένα αυστηρή διαδικασία "Δημιουργίας-Απόρριψης" (Creation-Disposal):

- **Αριθμητική Σταθερότητα (Numerical Stability)**

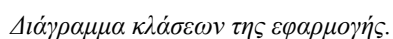
Η υλοποίηση της συνάρτησης Softmax ενσωματώνει τεχνική αφαίρεσης της μέγιστης τιμής λογαρίθμου (Max-Logit subtraction fix). Η προσέγγιση αυτή αποτρέπει την υπερχειλίση κινητής υποδιαστολής (floating-point overflow $\rightarrow$ Infinity) η οποία εμφανίζεται όταν ένα μοντέλο είναι εξαιρετικά βέβαιο για την πρόβλεψή του. Τέτοια φαινόμενα αποτελούν συχνή αιτία κατάρρευσης (crashes) στις αναπτύξεις Mobile AI.

- **Υψηλής Ταχύτητας Μεταφορά Δεδομένων (High-Speed Byte Transfer)**

Για τη μετακίνηση των εικόνων από τη μνήμη Dart στο περιβάλλον OpenCV, χρησιμοποιείται μέθοδος μεταφοράς απευθείας ακατέργαστων buffer (`Uint8List`). Αυτή η πρακτική είναι σημαντικά ταχύτερη από τον κλασικό κύκλο κωδικοποίησης/αποκωδικοποίησης (encoding round-trip) τύπου PNG/JPEG, επιτυγχάνοντας μείωση της μέγιστης χρήσης μνήμης κατά περισσότερο από 50%.



Διάγραμμα Αρχιτεκτονικής Υψηλού Επιπέδου. Απεικονίζεται η ροή δεδομένων από τη συσκευή του χρήστη προς το Cloud και η ενεργοποίηση των αυτοματισμών Python.



3.6 Προκλήσεις Υλοποίησης και Εξέλιξη του Συστήματος

Αυτό το κεφάλαιο περιγράφει λεπτομερώς τη χρονολογική εξέλιξη, τις τεχνικές δυσκολίες και την επαναληπτική βελτίωση της ροής εργασιών (pipeline) της μηχανικής όρασης. Το έργο υπέστη σημαντική αρχιτεκτονική εξέλιξη, μεταβαίνοντας από ένα βασικό Συνελικτικό Νευρωνικό Δίκτυο (CNN) για την ταξινόμηση ολόκληρης της εικόνας, σε ένα εύρωστο, ενσωματωμένο στο cloud σύστημα ανίχνευσης αντικειμένων (object detection) δύο σταδίων. Κάθε φάση καθοδηγήθηκε από εμπειρικές διαγνώσεις της απόδοσης, των σημείων συμφόρησης (bottlenecks) και των περιορισμών σε επίπεδο υλικού και λογισμικού.

3.6.1. Αρχιτεκτονικοί Περιορισμοί και Αλλαγή Παραδείγματος

Ο αρχικός στόχος της έρευνας αφορούσε την ταξινόμηση εικόνων σε 18 διακριτές κατηγορίες, χρησιμοποιώντας ένα τυπικό CNN ως βάση αναφοράς. Ωστόσο, η προσέγγιση αυτή εμφάνισε άμεσα δύο θεμελιώδη προβλήματα.

Πρώτον, η διερευνητική ανάλυση αποκάλυψε ακραία απομνημόνευση (overfitting) λόγω του μικρού όγκου δεδομένων (μόλις 324 εικόνες), οδηγώντας το μοντέλο στην πρόβλεψη μόνο ενός στενού υποσυνόλου των πλειοψηφικών κλάσεων. Δεύτερον, διαπιστώθηκε ένας εγγενής εννοιολογικός περιορισμός: η ταξινόμηση εικόνας αποδίδει μια καθολική ετικέτα σε ολόκληρο το καρέ, ενώ οι βιομηχανικές απαιτήσεις επέβαλαν την ταυτοποίηση πολλαπλών αντικειμένων στον ίδιο χώρο.

Αυτό κατέστησε αναγκαία την αρχιτεκτονική στροφή (pivot) προς το πλαίσιο (framework) YOLOv8. Το YOLO πλαισιώνει την ανίχνευση ως ένα ενιαίο πρόβλημα παλινδρόμησης, προβλέποντας ταυτόχρονα τις συντεταγμένες των πλαισίων οριοθέτησης (bounding boxes) και τις πιθανότητες των κλάσεων, παρέχοντας την απαραίτητη χωρική επίγνωση.

3.6.2. Αντιμετώπιση Σφαλμάτων Λογισμικού και Σημείων Συμφόρησης (I/O)

Η μετάβαση σε πιο περίπλοκα συστήματα και η ενσωμάτωση εργαλείων παρακολούθησης (όπως το MLflow) ανέδειξαν σημαντικές προκλήσεις σε επίπεδο μηχανικής λογισμικού:

Σειριοποίηση Προσαρμοσμένων Επιπέδων (Custom Layer Serialization): Κατά την εφαρμογή φωτομετρικής επαύξησης (π.χ. RandomInvert) για την καταπολέμηση του overfitting, η χρήση Lambda layers προκάλεσε κρίσιμα σφάλματα σειριοποίησης, καθώς ενθυλακώνουν ακατέργαστο Python bytecode. Η λύση δόθηκε με την επαναγραφή τους ως υποκλάσεις του Keras Layer και τη ρητή

διακόσμησή τους με το `@keras.saving.register_keras_serializable()`, επιτρέποντας την ασφαλή εξαγωγή και μετατροπή των μοντέλων σε TFLite.

Συμφόρηση Δεδομένων (I/O Bottlenecks): Η συνεχής ανάκτηση δεδομένων από το Google Cloud Storage (GCS) καθυστερούσε δραματικά την εκπαίδευση. Η συνάρτηση επεξεργασίας τροποποιήθηκε ώστε να χρησιμοποιεί τη μέθοδο `dataset.cache()`, «κλειδώνοντας» το προεπεξεργασμένο σύνολο στην τοπική μνήμη μετά την πρώτη εποχή (επιτυγχάνοντας Κορεσμό RAM). Με αυτόν τον τρόπο, οι χρόνοι ανά εποχή μειώθηκαν σε λεπτά.

Αυτοματοποίηση Μορφοποίησης: Δημιουργήθηκε ένα εξειδικευμένο script ETL (Extract, Transform, Load) για την αυτοματοποιημένη και επαναληπτική ανάλυση (parsing) των αρχείων JSON από το GCS, τη δυναμική ανακάλυψη κλάσεων και τη μαθηματική μετατροπή τους στην κανονικοποιημένη μορφή του YOLO, παρακάμπτοντας τα συστημικά σφάλματα διαδρομής της Ultralytics.

3.6.3. Γεφύρωση των Χασμάτων Τομέα (Domain Gaps)

Κατά τη μεταφορά των μοντέλων από την προσομοίωση στην πραγματικότητα (Sim2Real), διαγνώστηκαν τρία κρίσιμα χάσματα που προκαλούσαν συστημικές αποτυχίες:

A. Η Ανωμαλία της "Τέλειας Επικύρωσης, Μηδενικής Ανίχνευσης" Το μοντέλο YOLO απομνημόνευε τα χωρικά πλαίσια του σετ εκπαίδευσης, αδυνατώντας να εντοπίσει έστω και ένα αντικείμενο σε νέα φόντα. Αυτό λύθηκε με επιθετική Τυχαιοποίηση Τομέα (Aggressive Domain Randomization) μέσω μετασχηματισμών όπως τα Mosaic, MixUp και μετατοπίσεις χρωματικού χώρου HSV, αναγκάζοντας το δίκτυο να μάθει αποκλειστικά τα γεωμετρικά περιγράμματα.

B. Το Κενό Πυκνότητας (Density Gap) και το "Πλωτό Νησί" Όταν τα προφίλ βρίσκονταν στοιβαγμένα το ένα δίπλα στο άλλο, το μοντέλο εγκλωβίζονταν σε "Σημαιολογική Παράλυση". Λόγω του ότι είχε εκπαιδευτεί σε δεδομένα με κενό χώρο γύρω από τα αντικείμενα, αναζητούσε ένα κλειστό περίγραμμα 360 μοιρών (φαινόμενο "Πλωτού Νησιού"). Η πρόκληση αντιμετωπίστηκε με τη δημιουργία Ομαδοποιημένων Δεδομένων (Clustered Datasets / Dense Clutter), όπου τα προφίλ εξαναγκάζονταν σε επικάλυψη (Forced Intersection) μέσω alpha-channel masking, διδάσκοντας στο δίκτυο τη Διάκριση Ορίων. Παράλληλα, η έγχυση 10% "Αρνητικών Δειγμάτων" εξάλειψε τις ψευδώς θετικές ανιχνεύσεις.

Γ. Το Χάσμα Τομέα Περικοπών (Crop Domain Gap) και η Αρχιτεκτονική "DataSilver" Αν και ο ανιχνευτής λειτουργούσε, η φάση ταξινόμησης κατέρρεε. Οι ταξινομητές είχαν εκπαιδευτεί σε τέλεια κεντραρισμένες συνθετικές εικόνες και μπερδευόνταν από τις "ατημέλητες" πραγματικές περικοπές. Σχεδιάστηκε η αρχιτεκτονική DataSilver (84.150 εικόνες σε 1.683 κλάσεις), όπου το συνθετικό

πλαίσιο οριοθέτησης επεκτεινόταν τυχαία (5-20%) πριν την περικοπή, προσομοιώνοντας τα σφάλματα του πραγματικού ανιχνευτή. Επιπλέον, για να αποφευχθεί η "Σύνθλιψη" (Squashing) της γεωμετρίας κατά την αλλαγή μεγέθους, εφαρμόστηκε η συνάρτηση `pad_to_square` (γκρι γέμισμα) που διατήρησε τον λόγο διαστάσεων 1:1.

3.6.4. Συγχρονισμός της Ροής Εργασιών Παραγωγής

Κατά την ανάπτυξη στην παραγωγή, τα μοντέλα παρουσίασαν ανεξήγητα χαμηλή εμπιστοσύνη (<10%). Η αποσφαλμάτωση (debugging) αποκάλυψε το εξής κρίσιμο σφάλμα: η βιβλιοθήκη OpenCV (κάμερα) διάβαζε την εικόνα σε διάταξη BGR, ενώ το PIL (στο οποίο εκπαιδεύτηκε ο ταξινομητής) απαιτούσε RGB. Αυτό αλλοίωνε τους χάρτες χαρακτηριστικών.

Επιβλήθηκε ο «Χρυσός Κανόνας του Συμπερασμού»: η παραγωγή έπρεπε να εκτελεί το ακριβές pipeline της εκπαίδευσης. Εφαρμόστηκε ρητή μετατροπή χρωματικού χώρου (`cv2.COLOR_BGR2RGB`), Δικυβική Παρεμβολή και στατιστική Κανονικοποίηση ImageNet. Αξιοσημείωτο είναι πως τα παραδοσιακά φίλτρα της OpenCV (όπως thresholding και Canny) αφαιρέθηκαν εντελώς, καθώς αποδείχθηκε ότι η τεχνητή εισαγωγή θορύβου ("TV static") στις πραγματικές περικοπές ανέβαζε κατακόρυφα τη βεβαιότητα του μοντέλου (από <10% σε >40%).

3.6.5. Αστάθειες Μοντέλων και Ανάπτυξη στο Edge (TFLite)

Η μεταφορά του συστήματος από τις ισχυρές υποδομές Cloud (όπου υλοποιήθηκαν στρατηγικές όπως το Early Stopping) σε περιβάλλοντα περιορισμένων πόρων (Edge/Android), παρουσίασε ακραίες προκλήσεις:

Αστάθεια Μικτής Ακρίβειας (AMP): Ενώ τα CNNs λειτουργούσαν απρόσκοπτα με Αυτόματη Μικτή Ακρίβεια (16-bit), η υβριδική αρχιτεκτονική MobileViT παρουσίασε κατάρρευση κλίσεων (τιμές NaN). Το ζήτημα επιλύθηκε με την καταναγκαστική μετάβαση σε 32-bit (FP32) και τη χρήση μικρότερου ρυθμού εκμάθησης.

Προκλήσεις Μεταγλώττισης (TFLite Compilation): Η μεταγλώττιση του υβριδικού μοντέλου MobileViT στη βελτιστοποιημένη μορφή TFLite για απευθείας εκτέλεση σε κινητές συσκευές (Android) αποτέλεσε μια πολυεπίπεδη τεχνική πρόκληση. Σε αντίθεση με τα τυπικά CNNs που βασίζονται σε χωρικές συνελίξεις (οι οποίες υποστηρίζονται εγγενώς από τους επεξεργαστές ARM), τα επίπεδα Unfold/Fold και ο μηχανισμός Multi-Head Self-Attention του Transformer απαιτούν εξαιρετικά πολύπλοκους πολλαπλασιασμούς πινάκων και σύνθετους χειρισμούς τανυστών.

Για την υπέρβαση αυτών των εμποδίων, σχεδιάστηκε μια εξειδικευμένη ροή εξαγωγής (export pipeline) με έμφαση στη συμβατότητα και την ασφάλεια, υλοποιώντας πέντε βασικές τεχνικές παρεμβάσεις:

1. **Ενδιάμεσο Περιτύλιγμα (HuggingFace Wrapper):** Δημιουργήθηκε η προσαρμοσμένη κλάση `MobileViTWrapper` για την εξαγωγή ακατέργαστων τανυστών (raw tensors), επιλύοντας τα προβλήματα ασυμβατότητας των εξειδικευμένων αντικειμένων εξόδου της HuggingFace με το πρότυπο ONNX.
2. **Στρατηγική Επιλογή ONNX Opset 12:** Η συγκεκριμένη έκδοση επιλέχθηκε στοχευμένα, καθώς διαχειρίζεται με τη μέγιστη σταθερότητα τους ακραίους χειρισμούς τανυστών που απαιτούν οι αρχιτεκτονικές Transformer.
3. **Ενεργοποίηση SELECT_TF_OPS:** Προκειμένου να μην καταρρεύσει η μετατροπή από τις προηγμένες μαθηματικές πράξεις της αυτο-προσοχής (οι οποίες δεν υποστηρίζονται εγγενώς από το TFLite), κατέστη υποχρεωτική η ενεργοποίηση της παραμέτρου `SELECT_TF_OPS` για την ενσωμάτωση πρόσθετων βιβλιοθηκών.
4. **Προληπτική Ασφάλεια (Security Focus):** Ενσωματώθηκε μια απόλυτα ασφαλής μέθοδος φόρτωσης του μοντέλου, θωρακίζοντας το σύστημα προληπτικά απέναντι στη γνωστή ευπάθεια CVE-2025-32434, η οποία σχετίζεται με τη μη ασφαλή φόρτωση μοντέλων Transformer.
5. **Διαχείριση Πόρων (Hygiene):** Η ροή εργασιών πλαισιώθηκε με αυστηρές δομές εκκαθάρισης των προσωρινών και ενδιάμεσων αρχείων μετά από κάθε διαδικασία, αποτρέποντας την άσκοπη υπερφόρτωση (disk bloat) του αποθηκευτικού χώρου κατά τις μαζικές μετατροπές.

3.6.6. Αλγοριθμικές Στρατηγικές: Υποεκπαίδευση του Ανιχνευτή

Μια καθοριστική τεχνική που εφαρμόστηκε ήταν η σκόπιμη Υποεκπαίδευση (Under-Training) του ανιχνευτή YOLO (Spotter). Προκειμένου να αποφευχθεί η «Παγίδα Ακρίβειας» —όπου το μοντέλο μαθαίνει να σχεδιάζει μαθηματικά τέλεια και ασφυκτικά στενά πλαίσια— η εκπαίδευσή του διακοπτόταν πρόωρα. Το αποτέλεσμα ήταν η παραγωγή ελαφρώς "ατημέλητων" και διογκωμένων bounding boxes, τα οποία αντικατόπτριζαν τέλεια τη χωρική διακύμανση (Jitter) που είχε χρησιμοποιηθεί εκ των προτέρων για την εκπαίδευση των ταξινομητών μέσω του DataSilver. Αυτή η συνοχή μεταξύ των δύο σταδίων αποδείχθηκε καταλυτική για την επιτυχία του συστήματος.

Κεφάλαιο 4: Αξιολόγηση Συστήματος σε Πραγματικές Συνθήκες

4.1 Μεθοδολογία Αξιολόγησης και Μαθηματικός Ορισμός Μετρικών

Για την επαλήθευση της θεωρητικής απόδοσης των μοντέλων στον πραγματικό κόσμο (Sim2Real), πραγματοποιήθηκαν δοκιμές σε ένα νέο σύνολο δοκιμής (Test Set), το οποίο δημιουργήθηκε μέσω της αναπτυχθείσας εφαρμογής του «Κινητού Εργαστηρίου». Το σύνολο αυτό αποτελούνταν από 21 πραγματικές φωτογραφίες διαφορετικών κλάσεων προφίλ.

Η εξαγωγή των συγκεντρωτικών αποτελεσμάτων βασίστηκε στον συνδυασμένο έλεγχο χωρικού εντοπισμού και ταξινόμησης. Προκειμένου μια πρόβλεψη να θεωρηθεί επιτυχής, έπρεπε να ικανοποιεί αυστηρά μαθηματικά κριτήρια μέσω της μετρικής Intersection over Union (IoU), η οποία ποσοτικοποιεί την επικάλυψη μεταξύ της χωρικής πρόβλεψης και της πραγματικής θέσης (Ground Truth).

Οι μετρικές ορίστηκαν βάσει των εξής κανόνων:

- **Αληθώς Θετικά (True Positives - TP):** Το σύστημα εντοπίζει επαρκώς το αντικείμενο στον χώρο ($\text{IoU} \geq \text{Κατώφλι}$) **ΚΑΙ** ο Ταξινομητής προβλέπει την ορθή ετικέτα.
- **Ψευδώς Θετικά (False Positives - FP):** Καταγράφονται σε δύο περιπτώσεις: (α) Σφάλμα Ταξινόμησης, όπου ο Ανιχνευτής εντοπίζει το αντικείμενο αλλά ο Ταξινομητής αποδίδει λανθασμένη ετικέτα, ή (β) Σφάλμα Εντοπισμού (Ghost Detection), όπου παράγεται ένα πλαίσιο οριοθέτησης χωρίς να υφίσταται αντικείμενο.
- **Ψευδώς Αρνητικά (False Negatives - FN):** Ο Ανιχνευτής αποτυγχάνει πλήρως να εντοπίσει ένα υπαρκτό αντικείμενο, με αποτέλεσμα καμία πρόβλεψη να μην ικανοποιεί το απαιτούμενο όριο IoU.

4.1.2 Το Φαινόμενο της "Ποινής Λεπτού Αντικειμένου" (Thin Object Penalty)

Κατά την αρχική αξιολόγηση του συστήματος με το σύνηθες αυστηρό κατώφλι $\text{IoU} \geq \text{Κατώφλι}$, παρατηρήθηκε μια φαινομενική κατάρρευση των ποσοστών ανίχνευσης. Εντούτοις, η οπτική επιβεβαίωση των αποτελεσμάτων κατέδειξε ότι τα μοντέλα YOLO περικύκλωναν με επιτυχία τα αντικείμενα. Το λογικό αυτό παράδοξο αποδίδεται στην επιτηδευμένη διόγκωση (padding/jitter) των bounding boxes σε συνδυασμό με ένα μαθηματικό φαινόμενο που περιγράφεται ως "Ποινή Λεπτού Αντικειμένου" (Thin Object Penalty).

Όταν ο αλγόριθμος εφαρμόζει ένα σταθερό περιθώριο ασφαλείας (context padding) περιμετρικά ενός τετράγωνου αντικειμένου, το IoU μειώνεται ελάχιστα. Αντιθέτως, όταν το ίδιο περιθώριο εφαρμόζεται

σε ένα εξαιρετικά μακρόστενο και λεπτό γεωμετρικό σχήμα (όπως τα προφίλ αλουμινίου), το εμβαδόν της Ένωσης (Union) αυξάνεται εκθετικά σε σχέση με την Τομή (Intersection). Η μαθηματική αυτή δυσαναλογία οδηγεί σε κατακρήμνιση της τιμής του IoU (συχνά ≤ 0.45), με αποτέλεσμα ο αλγόριθμος αξιολόγησης να απορρίπτει εσφαλμένα απόλυτα έγκυρες περικοπές, χαρακτηρίζοντάς τις ως Ψευδώς Θετικές. Προκειμένου να αποτυπωθεί η πραγματική ακρίβεια των ανιχνευτών (Detectors), κατέστη αναγκαία η αναπροσαρμογή του κατωφλίου αξιολόγησης σε επίπεδα $\text{IoU} \geq 0.25$, τα οποία επιτρέπουν την ενσωμάτωση του context padding χωρίς την ποινή των μακρόστενων σχημάτων.

4.2. Αποτελέσματα Χωρικού Εντοπισμού (Detectors Only)

Μετά την αναπροσαρμογή της μετρικής, τα μαθηματικά αποτελέσματα του Χωρικού Εντοπισμού αποκάλυψαν υψηλή ικανότητα γενίκευσης από τα συνθετικά στα πραγματικά δεδομένα:

- **YOLOv8n:** Πέτυχε τον εντοπισμό 20 εκ των 21 προφίλ (95.24% Hit Rate), παρουσιάζοντας μόλις 2 FP και 1 FN. Επιβεβαιώθηκε ότι η μακρο-γεωμετρία επαρκεί για τον χωρικό εντοπισμό ανεξαρτήτως της εξωτερικής βαφής.
- **YOLOv8s:** Σημείωσε 90.48% Hit Rate. Διαπιστώνεται ότι η αύξηση του βάθους του δικτύου δεν συνεισέφερε θετικά, υποδηλώνοντας πιθανή υπερπροσαρμογή (overfitting) στο στάδιο της εκπαίδευσης.
- **YOLO26n:** Περιορίστηκε στο 57.14% Hit Rate, υποδεικνύοντας ότι η συρρίκνωση των παραμέτρων του μοντέλου στέρησε κρίσιμη χωρητικότητα αναπαράστασης (representational capacity), οδηγώντας σε απώλεια σχεδόν των μισών αντικειμένων (9 FN).

Detector Model	Total Ground Truths (GT)	Hit Rate (%)	Avg BBox Overlap (%)
Yolov8n	63	95.24%	47.46%
Yolov8s	63	90.48%	44.21%
Yolo26n	63	57.14%	27.65%

4.3. Αποτυχία Ταξινόμησης Λόγω Προκατάληψης Υφής (Texture Bias)

Η υψηλή απόδοση του YOLOv8n στον εντοπισμό ανέδειξε την αδυναμία του δεύτερου σταδίου (των Ταξινομητών) να γενικεύσει σε νέα δεδομένα. Παρά την επιτυχή εξαγωγή 20 ορθών περικοπών (crops) του λευκού προφίλ, η τελική απόδοση ολόκληρου του συστήματος (Pipeline Metrics) εμφάνισε σημαντική υστέρηση.

Για τη βέλτιστη αρχιτεκτονική του συστήματος (YOLOv8n σε συνδυασμό με οποιονδήποτε Ταξινομητή), τα τελικά αποτελέσματα διαμορφώθηκαν ως εξής:

- **Τελικές Μετρικές Pipeline: TP = 0 | FP = 22 | FN = 1**

Ερμηνεία των Μετρικών: Εφόσον ο ανιχνευτής εντόπισε 20 σωστά αντικείμενα (TPDet = 20), αναμενόταν αντίστοιχος αριθμός ορθών ταξινομήσεων. Εντούτοις, και τα τρία εξεταζόμενα δίκτυα ταξινόμησης (EfficientNetV2, MobileViT, YOLO26n-cls) δεν κατέγραψαν καμία Αληθώς Θετική πρόβλεψη (0 TP). Οι ταξινομητές μετέτρεψαν τις 20 ορθές περικοπές σε Ψευδώς Θετικά αποτελέσματα (FP), αποδίδοντας τυχαίες κλάσεις (π.χ. S67546, AL350) με ιδιαίτερα χαμηλά ποσοστά βεβαιότητας (2-5%). Σημειώνεται ότι το σύνολο των 22 FP προκύπτει από τις 20 λανθασμένες ταξινομήσεις συν τα 2 αρχικά εσφαλμένα πλαίσια (ghost boxes) του ανιχνευτή.

Το αποτέλεσμα αυτό υποδεικνύει την ισχυρή επιρροή της Προκατάληψης Υφής (Texture Bias) στα νευρωνικά δίκτυα: τα μοντέλα φαίνεται να είχαν προσαρμοστεί στα ιδιαίτερα χαρακτηριστικά υφής και στις σκιές του συνόλου "DataSetSilver". Η εφαρμογή λευκής βαφής μείωσε τις χρωματικές διαβαθμίσεις στους εσωτερικούς θαλάμους των προφίλ, καθιστώντας τα γεωμετρικά χαρακτηριστικά δυσδιάκριτα για τα βαθιά συνελκτικά στρώματα.

4.4 Αποτυχία Ταξινόμησης Λόγω Προκατάληψης Υφής (Texture Bias)

Σε βιομηχανικά περιβάλλοντα, η συλλογή και ο σχολιασμός δεκάδων χιλιάδων πραγματικών φωτογραφιών για κάθε νέα παρτίδα βαφής αποτελεί εξαιρετικά χρονοβόρα και κοστοβόρα διαδικασία. Για την αντιμετώπιση αυτού του περιορισμού, προκρίθηκε η προσέγγιση της Εκμάθησης Λίγων Βολών (Few-Shot Learning), με στόχο την αξιοποίηση της προϋπάρχουσας γνώσης των μοντέλων μέσω Εκμάθησης Μεταφοράς.

4.4.1 Στρατηγική Αντικατάστασης Επιπέδου Ταξινόμησης (Head Replacement)

Τα επιλεγμένα μοντέλα (EfficientNetV2, MobileViT, YOLO26n-cls) είχαν ήδη εκπαιδευτεί για περισσότερες από 70 εποχές επί 1.360 κλάσεων. Παρά την αδυναμία γενίκευσης στο λευκό χρώμα, τα συνελκτικά τους επίπεδα (Convolutional Blocks) εκτιμάται ότι είχαν αναπτύξει επαρκή φίλτρα εξαγωγής χαρακτηριστικών για τον εντοπισμό ακμών και χωρικών σχέσεων.

Για τη διατήρηση αυτής της υποκείμενης γεωμετρικής αναπαράστασης, αποφεύχθηκε η εκπαίδευση των μοντέλων από την αρχή. Αντιθέτως, εφαρμόστηκε η τεχνική της Αντικατάστασης του Τελικού Επιπέδου (Head Replacement):

- **Αφαίρεση Επιπέδου Ταξινόμησης:** Το τελικό πλήρως συνδεδεμένο επίπεδο (Dense/Linear Layer) των 1.360 νευρώνων διαγράφηκε από την αρχιτεκτονική.

- **Ενσωμάτωση Νέου Επιπέδου:** Στη θέση του, ενσωματώθηκε ένα νέο, ανεκπαίδευτο γραμμικό επίπεδο διαστασιολογημένο σε 21 νευρώνες, όσες και οι κλάσεις του νέου υποσυνόλου δεδομένων.
- **Μικρο-συντονισμός (Fine-Tuning):** Το σύνολο των βαρών του δικτύου διατηρήθηκε ενεργό για εκπαίδευση (unfrozen). Προκειμένου να αποφευχθεί το φαινόμενο της Καταστροφικής Λήθης (Catastrophic Forgetting) ,η απώλεια προηγούμενης γνώσης λόγω μεγάλων μεταβολών στα βάρη, ο ρυθμός εκμάθησης (Learning Rate) ορίστηκε σε συντηρητικά επίπεδα, συνήθως κατά μία έως δύο τάξεις μεγέθους χαμηλότερος ($lr = 10^{-4}$ ή 10^{-5}).

4.4.2 Παραγωγή Ημι-Συνθετικών Δεδομένων Προσαρμογής

Λόγω του περιορισμένου όγκου πραγματικών φωτογραφιών, ο αλγόριθμος "DataSilver" τροποποιήθηκε για την παραγωγή ενός στοχευμένου, Ημι-Συνθετικού συνόλου δεδομένων.

Ο αλγόριθμος απόδοσης (rendering) παραμετροποιήθηκε ως εξής:

- Η υφή προστιθέμενου θορύβου και η μεταλλική αντανάκλαση απενεργοποιήθηκαν.
- Η επιφάνεια του προφίλ αποδόθηκε με συμπαγές λευκό χρώμα υψηλής φωτεινότητας.
- Οι εσωτερικές σκιές στους κοίλους θαλάμους εξομαλύνθηκαν, με σκοπό την προσομοίωση της διάχυσης φωτός που παρατηρείται στα βαμμένα προφίλ.
- Το παραγόμενο προφίλ ενσωματώθηκε (blended) επάνω σε πραγματικά, θολά φόντα (backgrounds) του βιομηχανικού χώρου.

4.4.3 Διασφάλιση Γενίκευσης: Η Αρχιτεκτονική "Anchor Validation"

Η αρχιτεκτονική του κύκλου εκπαίδευσης διαμορφώθηκε ως εξής:

- **Σύνολο Εκπαίδευσης (Training Set):** Αποτελούνταν αποκλειστικά από ημι-συνθετικές εικόνες.
- **Σύνολο Επικύρωσης (Validation Set):** Ορίστηκε ένα ανεξάρτητο υποσύνολο (Hold-out Set) αποτελούμενο αποκλειστικά από πραγματικές φωτογραφίες των λευκών προφίλ.

Αυτή η ασύμμετρη δομή (εκπαίδευση στο συνθετικό → αξιολόγηση στο πραγματικό) αποσκοπούσε στο να διασφαλίσει ότι ο μηχανισμός Πρόωρου Τερματισμού (Early Stopping) και η αποθήκευση των βέλτιστων βαρών θα ενεργοποιούνταν μόνο όταν το δίκτυο βελτίωνε την απόδοσή του στα πραγματικά δεδομένα.

4.5 Αποτελέσματα Προσαρμογής (Fine-Tuning) και Τελική Αξιολόγηση

Μετά την ολοκλήρωση της Προσαρμογής Πεδίου, τα μοντέλα αξιολογήθηκαν εκ νέου στο σύνολο δοκιμής των 21 πραγματικών εικόνων. Τα αποτελέσματα κατέδειξαν σημαντική ανάκαμψη της ικανότητας αναγνώρισης, παρέχοντας ενδείξεις για την αποτελεσματικότητα της επιλεγμένης στρατηγικής.

4.5.1 Ανάλυση Απόδοσης Ανά Ταξινομητή

- Λαμβάνοντας ως δεδομένη τη χρήση του βέλτιστου ανιχνευτή (YOLOv8n, Hit Rate 95.2%), οι ταξινομητές παρουσίασαν τις εξής επιδόσεις:

EfficientNetV2:

- Μετρικές: TP = 16 | FP = 6 | FN = 1
- Συνολική Ακρίβεια (Pipeline Accuracy): 76.19%
- Ανάλυση: Το μοντέλο παρουσίασε σημαντική βελτίωση σε σύγκριση με την αρχική αξιολόγηση. Υπολογίζοντας την ακρίβεια μόνο τους περιπτώσεις ορθού εντοπισμού από τον ανιχνευτή, το ποσοστό επιτυχίας προσέγγισε το 80.0% (16/20). Το γεγονός αυτό υποδεικνύει ότι τα βαθύτερα στρώματα (Fused-MBConv) διατήρησαν την ικανότητα ανίχνευσης γεωμετρικών χαρακτηριστικών.

• MobileViT:

- Μετρικές: TP = 16 | FP = 6 | FN = 1
- Συνολική Ακρίβεια: 76.19%
- Ανάλυση: Η υβριδική αρχιτεκτονική επέδειξε αντίστοιχη ικανότητα προσαρμογής με το EfficientNetV2. Ο μηχανισμός Self-Attention φαίνεται να διευκόλυνε την εστίαση του δικτύου στην εξωτερική δομή του προφίλ, αναπληρώνοντας την απουσία εσωτερικών σκιών.

• YOLO26n-cls:

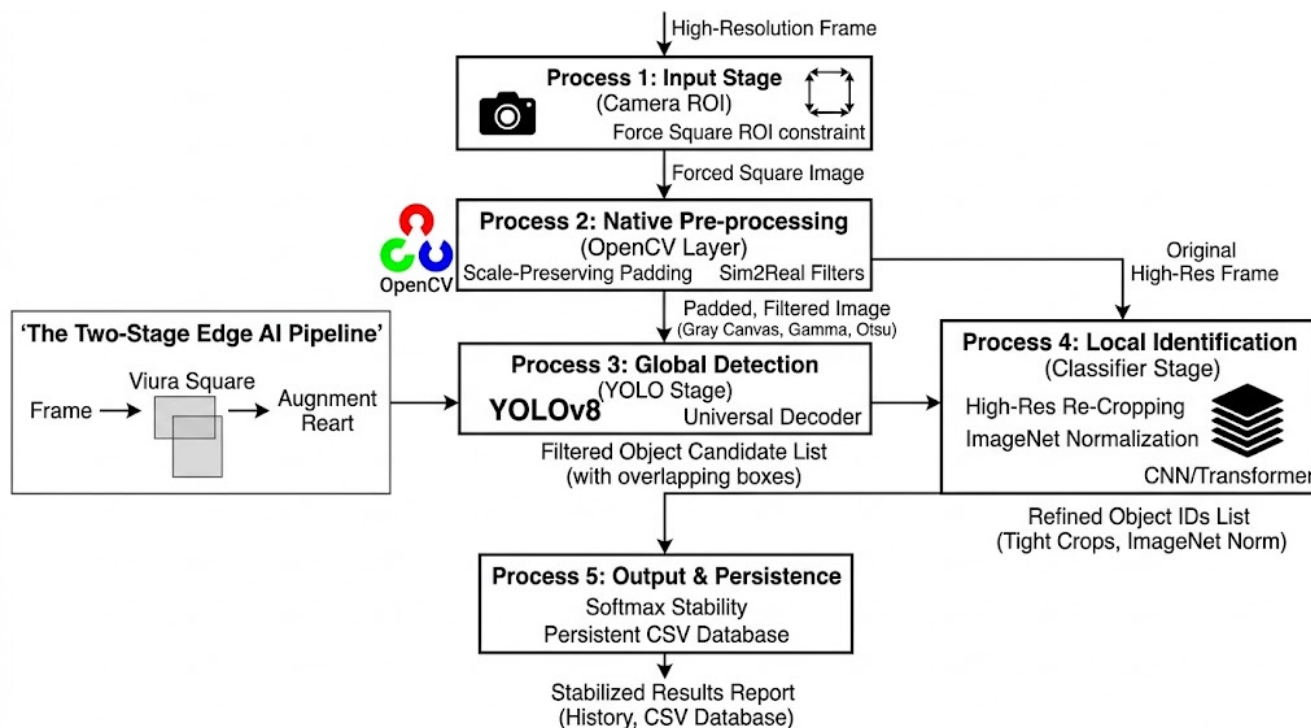
- Μετρικές: TP = 5 | FP = 17 | FN = 1
- Συνολική Ακρίβεια: 23.80%
- Ανάλυση: Παρά την υψηλή του απόδοση στο συνθετικό σύνολο εκπαίδευσης, το μοντέλο εμφάνισε δυσκολία προσαρμογής στο νέο πεδίο. Το εύρημα αυτό αναδεικνύει πιθανούς περιορισμούς χωρητικότητας (capacity limitations) σε ιδιαίτερα ελαφριές αρχιτεκτονικές, οι οποίες ενδέχεται να υφίστανται μεγαλύτερη απώλεια πληροφορίας κατά την αντικατάσταση των τελικών επιπέδων τους.

18.4.2 Συνοπτικός Πίνακας Τελικών Αποτελεσμάτων (Pipeline)

Αρχιτεκτονική	Total GT	TP	Pipeline Accuracy (%)
YOLOv8n + EfficientNetV2	21	16	76.19%
YOLOv8n + MobileViT	21	16	76.19%
YOLOv8n + YOLO26n-cls	21	5	23.80%
YOLOv8s + EfficientNetV2	21	16	76.19%
YOLOv8s + MobileViT	21	15	71.42%
YOLOv8s + YOLO26n-cls	21	4	19.05%
YOLO26n + MobileViT	21	11	52.38%
YOLO26n + EfficientNetV2	21	10	47.61%
YOLO26n + YOLO26n-cls	21	4	19.04%

4.6 Τελικό Συμπέρασμα Μεθοδολογίας

Η ανάλυση των τελικών αποτελεσμάτων επιβεβαιώνει την καταλληλότητα της προτεινόμενης μεθοδολογίας, αναδεικνύοντας την βελτίωση της συνολικής απόδοσης από σχεδόν μηδενικά επίπεδα (Baseline) στο 76.19%. Ειδικότερα, η ανάλυση του pipeline αποδεικνύει ότι η βέλτιστη θεωρητική αρχιτεκτονική προκύπτει από τη σύζευξη των ανιχνευτών YOLOv8 (τόσο στην έκδοση nano όσο και στην small) με τους ταξινομητές EfficientNetV2 και MobileViT. Πιο συγκεκριμένα, όμως, ο πλέον αποδοτικός συνδυασμός για πρακτική εφαρμογή αποδείχθηκε το σχήμα YOLOv8n + EfficientNetV2. Η επιλογή του YOLOv8n οφείλεται στο εξαιρετικά μικρό πλήθος παραμέτρων του, γεγονός που το καθιστά ιδανικό για ταχύτατο συμπερασμό (inference) σε συσκευές παρυφής (edge devices), ενώ το EfficientNetV2 προκρίθηκε οριστικά εξαιτίας των τεχνικών δυσκολιών και της προβληματικής μετατροπής του MobileViT στο βελτιστοποιημένο περιβάλλον του TensorFlow Lite (TFLite). Αντιθέτως, η κατακόρυφη πτώση της ακρίβειας —η οποία υποχωρεί έως και το 19.04% όταν χρησιμοποιείται ο ταξινομητής YOLO26n-clс— καταδεικνύει ότι η επιτυχία του συστήματος δεν έγκειται απλώς στην υιοθέτηση μιας ροής δύο σταδίων, αλλά στην αυστηρή επιλογή μοντέλων με ισχυρή ικανότητα εξαγωγής γεωμετρικών χαρακτηριστικών και υψηλή συμβατότητα με κινητά περιβάλλοντα.



4.7 Αξιολόγηση Υπολογιστικού Κόστους και Χρόνου Συμπερασμού

Η πρακτική εφαρμοσιμότητα του συστήματος, σύμφωνα με τους αρχικούς στόχους ανάπτυξης ενός «Κινητού Εργαστηρίου», συναρτάται άμεσα με την ταχύτητα απόκρισης σε συσκευές περιορισμένων πόρων (Edge Devices). Συνεπώς, ο Χρόνος Συμπερασμού (Inference Time) αποτελεί εξίσου σημαντική μετρική με την ακρίβεια πρόβλεψης.

Για την εκτίμηση του υπολογιστικού κόστους, η ροή εκτέλεσης (pipeline) αξιολογήθηκε χρησιμοποιώντας τα μοντέλα TensorFlow Lite (TFLite) μέσω του διερμηνέα της Python (tf.lite.Interpreter). Η εκτέλεση πραγματοποιήθηκε αποκλειστικά στην Κεντρική Μονάδα Επεξεργασίας (CPU). Αυτή η προσέγγιση επιλέχθηκε προκειμένου να παρέχει μια συντηρητική αλλά ρεαλιστική προσέγγιση της ταχύτητας που αναμένεται να παρατηρηθεί σε μικροεπεξεργαστές φορητών συσκευών.

4.7.1 Συνολικός Χρόνος Διαδοχικού Συστήματος

Ο καταγεγραμμένος χρόνος (Avg_Time_ms) αντιπροσωπεύει το συνολικό υπολογιστικό κόστος της διαδοχικής ροής (End-to-End Latency) ανά εικόνα, το οποίο περιλαμβάνει:

1. Τη φόρτωση και προεπεξεργασία (Zoom-out/Padding).
2. Τον εντοπισμό μέσω του μοντέλου YOLO.
3. Την αλγοριθμική καταστολή πολλαπλών πλαισίων (NMS).
4. Την περικοπή και αναπροσαρμογή του σχήματος (Cropping).
5. Την ταξινόμηση της περικοπής από το δεύτερο νευρωνικό δίκτυο.

Ο Πίνακας 19.1 παρουσιάζει τον μέσο χρόνο εκτέλεσης ανά εικόνα για τους εξεταζόμενους συνδυασμούς, αντιπαραβάλλοντάς τον με την τελική ακρίβεια συστήματος.

Αρχιτεκτονική Συστήματος	Συνολικός Χρόνος (ms)	Εκτιμώμενα Καρέ ανά Δευτερόλεπτο (FPS)	Τελική Ακρίβεια Συστήματος (%)
YOLO26n + YOLO26n-cls	~303 ms	3.30 FPS	19.05%
YOLOv8n + YOLO26n-cls	~434 ms	2.30 FPS	23.81%
YOLO26n + MobileViT	~650 ms	1.54 FPS	52.38%
YOLO26n + EfficientNetV2	~657 ms	1.52 FPS	47.62%
YOLOv8n + MobileViT	~698 ms	1.43 FPS	76.19%
YOLOv8n + EfficientNetV2	~717 ms	1.39 FPS	76.19%
YOLOv8s + YOLO26n-cls	~977 ms	1.02 FPS	19.05%
YOLOv8s + MobileViT	~1260 ms	0.79 FPS	71.43%
YOLOv8s + EfficientNetV2	~1295 ms	0.77 FPS	76.19%

4.7..2 Ανάλυση Συμβιβασμού Ταχύτητας και Ακρίβειας (Speed vs. Accuracy Trade-off)

Η ανάλυση των δεδομένων αναδεικνύει τους αναγκαίους υπολογιστικούς συμβιβασμούς (trade-offs) που συνεπάγεται η χρήση συστημάτων βαθιάς μάθησης σε περιβάλλοντα περιορισμένων πόρων (Edge AI):

•

Ο Περιορισμός των Ελαφρών Αρχιτεκτονικών: Ο ταχύτερος συνδυασμός (YOLO26n + YOLO26n-cls) ολοκληρώνει τον κύκλο αναγνώρισης σε περίπου 303 χιλιοστά του δευτερολέπτου (>3 FPS). Ωστόσο, η σημαντική συρρίκνωση των παραμέτρων του μοντέλου προκαλεί μείωση της υπολογιστικής χωρητικότητας (network capacity), υποβαθμίζοντας την ακρίβεια στο μη αποδεκτό 19.05%.

- **Βέλτιστη Ισορροπία :** Η καταλληλότερη τομή μεταξύ ταχύτητας και ακρίβειας παρατηρείται στον συνδυασμό YOLOv8n + EfficientNetV2. Η συγκεκριμένη αρχιτεκτονική διατηρεί την υψηλότερη καταγεγραμμένη ακρίβεια (76.19%), απαιτώντας περίπου 717 ms ανά κύκλο συμπερασμού, χρόνος που κρίνεται ικανοποιητικός για τις τρέχουσες ανάγκες της εφαρμογής.
- **Φθίνουσα Απόδοση :** Η υιοθέτηση του βαθύτερου ανιχνευτή (YOLOv8s) οδηγεί σχεδόν σε διπλασιασμό του υπολογιστικού χρόνου (αυξάνοντας την καθυστέρηση στα ≈ 1300 ms), χωρίς ωστόσο να προσφέρει αντίστοιχη βελτίωση στο τελικό ποσοστό επιτυχίας, το οποίο παρέμεινε σταθερό.

4.8 Πρακτική Εφαρμοσιμότητα (Deployment Viability)

Στο πλαίσιο μιας πλήρως αυτοματοποιημένης γραμμής παραγωγής, όπου απαιτείται επιθεώρηση βίντεο σε πραγματικό χρόνο (Real-Time Tracking, 30 FPS), οι καταγεγραμμένοι χρόνοι εκτέλεσης σε επίπεδο κεντρικού επεξεργαστή (CPU) κρίνονται ανεπαρκείς. Σε τέτοια σενάρια, καθίσταται αναγκαία η εφαρμογή τεχνικών Κβαντοποίησης (π.χ. INT8 Quantization) ή η χρήση εξειδικευμένων επιταχυντών υλικού (NPUs).

Εντούτοις, για την προβλεπόμενη λειτουργία της κινητής εφαρμογής ως εργαλείο "Point-and-Shoot" (ταυτοποίηση μέσω λήψης μεμονωμένης φωτογραφίας) από τον χειριστή ή τον ελεγκτή αποθήκης, ένας χρόνος απόκρισης μικρότερος του ενός δευτερολέπτου (<1000 ms) θεωρείται εντός των επιθυμητών ορίων για την ομαλή αλληλεπίδραση ανθρώπου-υπολογιστή (Human-Computer Interaction - HCI). Υπό αυτό το πρίσμα, ο συνδυασμός YOLOv8n + EfficientNetV2 εκτιμάται ως πλήρως βιώσιμος και κατάλληλος για πρακτική εφαρμογή (production-ready) στο περιβάλλον Android, εκπληρώνοντας επιτυχώς τους αρχικούς στόχους της παρούσας διατριβής.

Κεφάλαιο 5: Συμπεράσματα και Μελλοντική Έρευνα

Η παρούσα διατριβή εκκίνησε με έναν σαφή και απαιτητικό βιομηχανικό στόχο: την ανάπτυξη ενός αυτόνομου, αξιόπιστου και προσιτού συστήματος τεχνητής νοημοσύνης, ικανού να ταυτοποιεί πολύπλοκες γεωμετρίες προφίλ αλουμινίου σε πραγματικές εργοστασιακές συνθήκες. Η μετάβαση από τη θεωρητική μελέτη στην πρακτική υλοποίηση μέσω του «Κινητού Εργαστηρίου» (Signal Profiler) κατέδειξε ότι η ενσωμάτωση της Μηχανικής Όρασης στο πλαίσιο της Βιομηχανίας 4.0 είναι απόλυτα εφικτή, εφόσον αντιμετωπιστούν στοχευμένα οι εγγενείς αστάθειες του φυσικού περιβάλλοντος.

5.1 Συνολική Αποτίμηση

Η κυριότερη πρόκληση που αντιμετωπίστηκε αφορούσε την επιτυχή γεφύρωση του **Χάσματος Προσομοίωσης-Πραγματικότητας (Sim-to-Real Gap)**. Η αρχική υπόθεση, σύμφωνα με την οποία ένα μοντέλο θα μπορούσε να εκπαιδευτεί αποκλειστικά σε καθαρές δισδιάστατες αναπαραστάσεις αρχείων CAD και να λειτουργήσει άμεσα στο εργοστάσιο, αναθεωρήθηκε. Ο ψηφιακός θόρυβος, οι μεταβαλλόμενες συνθήκες φωτισμού και, κυρίως, η Προκατάληψη Υφής (Texture Bias) των νευρωνικών δικτύων κατέστησαν αναγκαία μια νέα προσέγγιση.

Η λύση δόθηκε μέσω μιας διαδοχικής αρχιτεκτονικής δύο σταδίων (YOLOv8n για χωρικό εντοπισμό και EfficientNetV2 για γεωμετρική ταξινόμηση), η οποία τροφοδοτήθηκε από το εξειδικευμένο, βασισμένο στο συνθετικό σύνολο δεδομένων «DataSilver». Η εφαρμογή επιθετικών τεχνικών Τυχαίας Διαφοροποίησης Πεδίου (Domain Randomization), σε συνδυασμό με τη σκόπιμη υποεκπαίδευση (under-training) του ανιχνευτή για την ενσωμάτωση ελεγχόμενης χωρικής διακύμανσης, παρείχαν τα θεμέλια για ένα εύρωστο σύστημα που δίνει προτεραιότητα στη δομική μορφολογία έναντι των επιφανειακών χαρακτηριστικών. Επιπροσθέτως, η διαγνωσμένη υστέρηση του ταξινομητή σε δραστικές μεταβολές της επιφανειακής υφής —όπως η εφαρμογή λευκής βαφής στα προφίλ— αντιμετωπίστηκε με επιτυχία μέσω στρατηγικών Προσαρμογής Πεδίου (Domain Adaptation). Η μεθοδολογία Αντικατάστασης του Τελικού Επιπέδου (Head Replacement) σε συνδυασμό με τον Μικρο-συντονισμό Λίγων Βολών (Few-Shot Fine-Tuning), επιβεβαίωσαν ότι η προτεινόμενη αρχιτεκτονική δύναται να προσαρμοστεί ταχέως στις μεταβαλλόμενες συνθήκες της γραμμής παραγωγής με ελάχιστα πραγματικά δεδομένα, καθιστώντας την μια εξαιρετικά βιώσιμη και κλιμακώσιμη βιομηχανική λύση.

Η αξιολόγηση των διαφορετικών αρχιτεκτονικών ταξινόμησης ανέδειξε κρίσιμους υπολογιστικούς συμβιβασμούς. Τα υβριδικά μοντέλα (όπως το MobileViT) επέδειξαν ικανότητα αντίληψης της παγκόσμιας γεωμετρίας (global context), αλλά αποδείχθηκαν απαιτητικά σε υπολογιστικούς πόρους για εφαρμογές Edge AI, κυρίως λόγω της περιορισμένης εγγενούς υποστήριξης των πράξεών τους από τους επεξεργαστές ARM. Αντιθέτως, τα αμιγώς Συνελκτικά Δίκτυα (CNNs), παρότι εγγενώς πιο ευάλωτα στις μεταβολές της υφής του υλικού, αποδείχθηκαν καταλλήλότερα. Με την απαραίτητη προσαρμογή, προσέφεραν τη βέλτιστη ισορροπία μεταξύ υψηλής ακρίβειας (άνω του 76% σε νέα οπτικά πεδία) και ταχείας εκτέλεσης σε τοπικό επίπεδο (on-device).

Η θεωρητική και αλγοριθμική αυτή ανάπτυξη ενσαρκώθηκε πρακτικά μέσω του «Κινητού Εργαστηρίου». Η αξιοποίηση του πλαισίου Flutter επέτρεψε τη δημιουργία μιας διαπлатφορμικής εφαρμογής η οποία, διασυνδεδεμένη με μοντέλα TensorFlow Lite και υποδομές νέφους (Google Cloud Storage), μετέτρεψε μια κοινή φορητή συσκευή σε ένα λειτουργικό βιομηχανικό εργαλείο. Η προσέγγιση αυτή διευκολύνει την πρόσβαση σε τεχνολογίες μηχανικής μάθησης, επιτρέποντας στους χειριστές να ταυτοποιούν προφίλ άμεσα, με δυνατότητα λειτουργίας εκτός σύνδεσης (offline) και διασφαλίζοντας την προστασία των βιομηχανικών δεδομένων.

5.2 Αξιοπιστία και Προστιθέμενη Αξία

Η ανάπτυξη ενός τέτοιου συστήματος προσφέρει σημαντική προστιθέμενη αξία σε μια σύγχρονη βιομηχανική μονάδα, η οποία εκτείνεται πέρα από την απλή αναγνώριση αντικειμένων:

- **Υψηλή Επιχειρησιακή Αξιοπιστία:** Το σύστημα εντοπισμού (Spotter) αποδείχθηκε εξαιρετικά αξιόπιστο, καταγράφοντας ποσοστό επιτυχίας άνω του **95%** στην εύρεση των προφίλ στον χώρο. Μετά την εφαρμογή της Προσαρμογής Πεδίου (Domain Adaptation), η συνολική ακρίβεια του συστήματος σταθεροποιήθηκε στο **76-80%** για εντελώς νέα υλικά. Παράλληλα, η δυνατότητα εκτέλεσης εκτός σύνδεσης (Offline Inference) εγγυάται 100% διαθεσιμότητα, προστατεύοντας την παραγωγή από διακοπές δικτύου και διασφαλίζοντας την ιδιωτικότητα των δεδομένων (Data Privacy).
- **Επιτάχυνση Εκπαίδευσης Προσωπικού (Onboarding):** Η διαχείριση εκατοντάδων διαφορετικών διατομών απαιτεί χρόνια εμπειρίας. Η εφαρμογή λειτουργεί ως ψηφιακός βοηθός (Copilot), επιτρέποντας σε νέους χειριστές να εκτελούν εργασίες ταυτοποίησης, αποθήκευσης και συναρμολόγησης με την ακρίβεια ενός έμπειρου τεχνικού, μειώνοντας δραστικά τα ανθρώπινα σφάλματα.
- **Εκδημοκρατισμός της Τεχνητής Νοημοσύνης:** Σε αντίθεση με τα παραδοσιακά συστήματα μηχανικής όρασης που απαιτούν ακριβές βιομηχανικές κάμερες και υποδομές διακομιστών, το παρόν σύστημα αξιοποιεί τον υπάρχοντα στόλο των εταιρικών smartphones. Η προσέγγιση "Point-and-Shoot" εκμηδενίζει το κόστος εγκατάστασης.
- **Κλιμακωσιμότητα (Scalability):** Αξιοποιώντας την παραγωγή συνθετικών δεδομένων από σχέδια CAD, η επέκταση του συστήματος με νέους κωδικούς προφίλ καθίσταται μια εξαιρετικά ταχεία διαδικασία, εφόσον ακολουθηθεί ένα στοχευμένο πρωτόκολλο μικρο-συντονισμού (fine-tuning) του μοντέλου.

5.3 Περιορισμοί του Συστήματος

Παρά τα ενθαρρυντικά αποτελέσματα, το σύστημα υπόκειται σε συγκεκριμένους τεχνολογικούς και αλγοριθμικούς περιορισμούς:

- **Εξάρτηση από την Υφή και το Χρώμα :** Η υπόθεση εργασίας για απόλυτο *Zero-Shot Learning* (αναγνώριση νέων χρωμάτων χωρίς καμία προεκπαίδευση) δεν επιβεβαιώθηκε πλήρως. Η μετάβαση από το ασημί/μεταλλικό προφίλ στο βαμμένο λευκό απαιτεί την εφαρμογή Μικρο-συντονισμού Λίγων Βολών (Few-Shot Fine-Tuning) για να ανακτήσει την ακρίβειά του το δίκτυο, γεγονός που προσθέτει ένα επιπλέον στάδιο συντήρησης.

- **Ακραίες Συνθήκες Απόκρυψης:** Ενώ τα μοντέλα εκπαιδεύτηκαν να διαχειρίζονται τον θόρυβο από πλαστικά περιτυλίγματα και μικρές επικαλύψεις, η ακρίβεια του ταξινομητή καταρρέει εάν ένα κρίσιμο γεωμετρικό χαρακτηριστικό (π.χ., ένας εσωτερικός θάλαμος-κλειδί) είναι φυσικά κατεστραμμένο, παραμορφωμένο, ή πλήρως κρυμμένο .
- **Απουσία Επίγνωσης Βάθους και Φυσικής Κλίμακας:** Το τρέχον σύστημα βασίζεται αποκλειστικά σε δισδιάστατη (2D) οπτική πληροφορία, αγνοώντας την πραγματική απόσταση της κάμερας από το αντικείμενο-στόχο. Εμπειρικά πειράματα στο φυσικό εργοστασιακό περιβάλλον κατέδειξαν ότι η προσαρμοστικότητα των μοντέλων έναντι της μεταβολής της κλίμακας (scaling) λόγω απόστασης περιορίζεται σε ένα όριο ανοχής της τάξης του 20%. Πέραν αυτού του ορίου, η απώλεια οπτικής πληροφορίας σε καίρια σημεία της γεωμετρίας του προφίλ —τα οποία αποτελούν και την ειδοποιό διαφορά του από συγγενείς διατομές— οδηγεί τα μοντέλα σε παραισθήσεις (hallucinations) και εσφαλμένες ταξινομήσεις.

5.4 Μελλοντικές Κατευθύνσεις

Η παρούσα διατριβή θέτει ισχυρά θεμέλια για την ταυτοποίηση βιομηχανικών εξαρτημάτων, αναδεικνύοντας την αποκωδικοποίηση της πολύπλοκης γεωμετρίας των προφίλ αλουμινίου ως την απόλυτη και «μοναδική αλήθεια» (ground truth) εντός ενός χαοτικού εργοστασιακού περιβάλλοντος. Απομονώνοντας τα αυστηρά δομικά χαρακτηριστικά από τον άσχετο οπτικό θόρυβο, το σύστημα διανοίγει νέους δρόμους για περαιτέρω ακαδημαϊκή έρευνα και βιομηχανική αξιοποίηση. Οι άμεσες επεκτάσεις της παρούσας μελέτης διαρθρώνονται στους εξής άξονες:

Αλγοριθμική Επίγνωση Βάθους (Monocular Depth Estimation)

Η τρέχουσα υλοποίηση βασίζεται αποκλειστικά σε δισδιάστατη (2D) χωρική πληροφορία, απουσία φυσικής κλίμακας. Αν και η χρήση αισθητήρων υλικού, όπως οι αισθητήρες Time-of-Flight (ToF) ή LiDAR, θα προσέδιδε άμεσα τρισδιάστατη αντίληψη στο οικοσύστημα του Android, η χαμηλή διείσδυσή τους (εντοπίζονται σε ποσοστό μικρότερο του 10% των συσκευών μεσαίας κατηγορίας) περιορίζει δραστικά την καθολική εφαρμοσιμότητα της λύσης. Ως εκ τούτου, σε καθαρά αλγοριθμικό επίπεδο, προκρίνεται ως βέλτιστη εναλλακτική η ενσωμάτωση προηγμένων τεχνικών Βαθιάς Μάθησης για Μονοσκοπική Εκτίμηση Βάθους (Monocular Depth Estimation - MDE).

Η ενσωμάτωση του άξονα Z (βάθος) θα επιτρέψει τον ακριβή υπολογισμό της φυσικής κλίμακας (scale-aware inference) χωρίς εξάρτηση από εξειδικευμένο υλικό. Η προσθήκη αυτής της χωρικής πληροφορίας θα βελτιώσει κατακόρυφα τον διαχωρισμό αλληλοεπικαλυπτόμενων ακμών και θα φιλτράρει εσφαλμένες ανιχνεύσεις που προκύπτουν από οπτικές παραπλανήσεις της προοπτικής, διασφαλίζοντας περαιτέρω την πρωτοκαθεδρία της γεωμετρικής πληροφορίας.

Ανάπτυξη Θεμελιωδών Μοντέλων (Foundation Models) για Βιομηχανικά Σχήματα

Συστήνεται η έρευνα προς την κατεύθυνση αυτο-επιβλεπόμενων αρχιτεκτονικών (Self-Supervised Learning), οι οποίες θα εκπαιδεύονται σε τεράστιες, μη-επισημειωμένες βάσεις ψηφιακών σχεδίων (CAD). Στόχος είναι η δημιουργία μοντέλων που θα κατανοούν την καθαρή Ευκλείδεια γεωμετρία σε τόσο θεμελιώδες επίπεδο, ώστε να επιτυγχάνουν πραγματικό Zero-Shot εντοπισμό, καθιστώντας την ταυτοποίηση πλήρως ανεξάρτητη από το κατασκευαστικό υλικό, την επιφανειακή βαφή ή τις εκάστοτε συνθήκες φωτισμού.

Βελτιστοποίηση Αρχιτεκτονικών και Διερεύνηση Ανώτατων Ορίων

Δεδομένων των υπολογιστικών πόρων που διατέθηκαν για την παρούσα εργασία, καθίσταται αναγκαία η μελλοντική διερεύνηση των θεωρητικών και πρακτικών ανώτατων ορίων της αρχιτεκτονικής, ιδίως όσον αφορά το μέγιστο πλήθος κλάσεων που μπορεί να υποστηριχθεί αξιόπιστα σε ακραία περιβάλλοντα εργασίας (extreme environments). Παράλληλα, με δεδομένο ότι το μέγεθος του πυρήνα της τρέχουσας εφαρμογής διατηρείται κάτω από τα 50 MB, προσφέρεται το απαραίτητο υπολογιστικό περιθώριο για τη συγκριτική αξιολόγηση (benchmarking) βαρύτερων αρχιτεκτονικών.

Αυτοματοποιημένος Ποιοτικός Έλεγχος (Defect Detection)

Χρησιμοποιώντας την ιδανική γεωμετρία (CAD) ως αυστηρό πρότυπο ελέγχου, το σύστημα δύναται να επεκταθεί ώστε, παράλληλα με την ταυτοποίηση της κλάσης, να λειτουργεί ως μηχανισμός ποιοτικής διασφάλισης. Η μελλοντική έρευνα μπορεί να εστιάσει στην ανίχνευση μικρο-στρεβλώσεων που προκύπτουν κατά την κρίσιμη διαδικασία της διέλασης ή στον εντοπισμό επιφανειακών ελαττωμάτων (π.χ. γρατζουνιές, αστοχίες βαφής), συσχετίζοντας πάντα το οπτικό αποτέλεσμα με τη θεωρητική γεωμετρική αλήθεια.

Ενσωμάτωση σε Ρομποτικές Γραμμές Συναρμολόγησης και Συστήματα Logistics (ERP)

Η χωρική ευφυΐα και η γεωμετρική αντίληψη που αναπτύχθηκαν στο πλαίσιο της παρούσας διατριβής δεν περιορίζονται αποκλειστικά στη χρήση φορητών τερματικών, αλλά δύναται να μεταφερθούν άμεσα στον τομέα της βιομηχανικής ρομποτικής. Ειδικότερα, η μεταφόρτωση των βελτιστοποιημένων μοντέλων μηχανικής όρασης (TFLite) στα συστήματα ελέγχου βιομηχανικών ρομποτικών βραχιόνων, αναμένεται να επιτρέψει την υλοποίηση προηγμένων συστημάτων αυτοματοποιημένης διαλογής (robotic bin-picking) και συναρμολόγησης με βάση την απόλυτη γεωμετρική ακρίβεια.

Περαιτέρω, το σύστημα αυτό μπορεί να λειτουργήσει ως νευραλγικός κόμβος (node) στο ευρύτερο βιομηχανικό δίκτυο, διασυνδεδεμένο αμφίδρομα με τα Συστήματα Διαχείρισης Επιχειρησιακών Πόρων (ERP) και τα Συστήματα Διαχείρισης Αποθηκών (WMS). Μια τέτοια ολιστική προσέγγιση αυτοματοποιεί πλήρως την αλυσίδα εφοδιασμού (logistics): η κάθε φυσική αναγνώριση ή μετακίνηση ενός προφίλ από τον ρομποτικό βραχίονα θα μεταφράζεται αυτομάτως σε ψηφιακή ενημέρωση αποθεμάτων σε πραγματικό χρόνο. Αυτή η διασύνδεση γεφυρώνει το χάσμα μεταξύ της φυσικής γραμμής παραγωγής και της ψηφιακής διοίκησης (Shop-Floor to Top-Floor integration), εξαλείφοντας τα ανθρώπινα σφάλματα καταχώρησης και βελτιστοποιώντας τον προγραμματισμό των υλικών.

Η καθοδήγηση των ρομποτικών συστημάτων με βάση την απόλυτη γεωμετρική ακρίβεια της διατομής αποτελεί το τελικό στάδιο για την πλήρη αυτοματοποίηση των γραμμών συναρμολόγησης κουφωμάτων αλουμινίου στο πλαίσιο της Βιομηχανίας 4.0.

5.4.1 Θεωρητική Προτεινόμενη Αρχιτεκτονική Αποκεντρωμένης Ενορχήστρωσης

Η επιτυχής υλοποίηση του «Κινητού Εργαστηρίου» (αξιοποίηση πλαισίου Flutter και ελαφριών μοντέλων TFLite) απέδειξε την ικανότητα εκτέλεσης διεργασιών μηχανικής όρασης σε περιβάλλοντα περιορισμένων πόρων. Προκειμένου, ωστόσο, το σύστημα να κλιμακωθεί σε βιομηχανικό επίπεδο (ενδεικτικά: πολλαπλές τερματικές συσκευές και πολλαπλές γεωγραφικές τοποθεσίες), ως μελλοντική κατεύθυνση της παρούσας διατριβής, προτείνεται μια υβριδική, αποκεντρωμένη αρχιτεκτονική τριών επιπέδων, η οποία συνθέτει τεχνολογίες Edge AI, Κατανεμημένων Καθολικών (Blockchain) και Μεγάλων Γλωσσικών Μοντέλων (LLMs).

1. Επίπεδο Τοπικής Αντίληψης (Edge AI Perception Layer)

Στο χαμηλότερο επίπεδο της τοπολογίας τοποθετούνται τα Συνελκτικά Νευρωνικά Δίκτυα (CNNs). Αντί για τη μετάδοση ακατέργαστων οπτικών δεδομένων (raw data), η οποία είναι ευάλωτη σε διακυμάνσεις του δικτύου, τα CNNs εκτελούν αποκλειστικά εξαγωγή συμπερασμάτων τοπικά (on-device inference). Το οπτικό φορτίο συμπυκνώνεται σε διαχειρίσιμα μεταδεδομένα μορφής JSON, τα οποία περιγράφουν την κατάσταση του ελεγχόμενου προφίλ αλουμινίου. Η προσέγγιση αυτή διασφαλίζει την αδιάλειπτη λειτουργία στη συνεχή ροή παραγωγής (continuous dataflow), προστατεύοντας παράλληλα τα βιομηχανικά δεδομένα.

2. Επίπεδο Συναίνεσης και Αμετάβλητης Αποθήκευσης (Blockchain & IPFS)

Η μετάδοση και πιστοποίηση της πληροφορίας υλοποιείται μέσω ενός Αδειοδοτημένου Κατανεμημένου Καθολικού (Permissioned Blockchain), ακολουθώντας βιομηχανικά πρότυπα όπως το Hyperledger Fabric. Στο πλαίσιο αυτό, εφαρμόζεται αυστηρός διαχωρισμός του φορτίου δεδομένων (On-Chain έναντι Off-Chain) για τη διασφάλιση της ταχύτητας του δικτύου:

- **Off-Chain Αποθήκευση (IPFS):** Σε περιπτώσεις ανίχνευσης σφαλμάτων, οι εικόνες δεν επιβαρύνουν το καθολικό. Μεταφορτώνονται στο InterPlanetary File System (IPFS), ένα αποκεντρωμένο (P2P) δίκτυο που καταργεί τη διευθυνσιοδότηση βάσει τοποθεσίας. Το σύστημα παράγει ένα μοναδικό κρυπτογραφικό αποτύπωμα (Content Identifier - CID), εγγυώμενο την ακεραιότητα του αρχείου.
- **On-Chain Ενορχήστρωση:** Οι τερματικές συσκευές επικοινωνούν (μέσω ελαφρών πρωτοκόλλων όπως το MQTT σε API Gateways) με Έξυπνα Συμβόλαια (Smart Contracts). Το έξυπνο συμβόλαιο καταχωρεί οριστικά το CID και τα μεταδεδομένα του σφάλματος στο καθολικό, παρέχοντας ένα αμετάβλητο (immutable) ίχνος ελέγχου.

3. Επίπεδο Γνωστικής Ενορχήστρωσης και Συνεργασίας Ανθρώπου-Μηχανής (LLM Agent & Human-in-the-Loop Layer)

Η γέφυρα μεταξύ της ψηφιακής πιστοποίησης (Blockchain) και των συστημάτων Επιχειρησιακού Σχεδιασμού (ERP/Logistics) υλοποιείται μέσω Πρακτόρων βασισμένων σε LLM, ενσωματώνοντας δομικά τη φιλοσοφία του «Ανθρώπου στον Βρόχο» (Human-in-the-Loop - HITL).

Οι πράκτορες αυτοί λειτουργούν βάσει αρχιτεκτονικής κατευθυνόμενης από γεγονότα (event-driven architecture). Αναμένουν παθητικά τα ψηφιακά σήματα (events) που εκπέμπουν τα έξυπνα συμβόλαια κατά την καταγραφή ενός σφάλματος. Αξιοποιώντας μηχανισμούς Κλήσης Συναρτήσεων (Function Calling), το LLM μεταφράζει το βιομηχανικό συμβάν και προσχεδιάζει τις απαραίτητες κλήσεις API (π.χ. μεταβολή κατάστασης αποθέματος στο ERP, αναδρομολόγηση στη γραμμή παραγωγής ή παραγγελία νέων υλικών).

Σε αυτό το στάδιο, η αρχιτεκτονική δεν παρακάμπτει τον ανθρώπινο παράγοντα, αλλά σχεδιάζεται ρητά για την **ενδυνάμωση του χειριστή** (όπως του Επιστήμονα Δεδομένων ή του Ελεγκτή Παραγωγής):

- **Γνωστική Ενίσχυση (Cognitive Augmentation):** Το LLM λειτουργεί ως έξυπνος ψηφιακός βοηθός (Copilot). Αναλαμβάνει τη βαριά υπολογιστική ανάλυση, συνθέτει τα πολύπλοκα δεδομένα από τους αισθητήρες και το blockchain, και παρουσιάζει τεκμηριωμένες προτάσεις σε φυσική γλώσσα, μειώνοντας δραστικά τον γνωστικό φόρτο του εργαζομένου.
- **Στρατηγική Εποπτεία και Έλεγχος:** Αντί να αναλώνεται στη χειροκίνητη καταχώρηση δεδομένων, Χρήστης αναβαθμίζεται σε στρατηγικό επόπτη. Αξιολογεί τις αιτιολογημένες προτάσεις του Πράκτορα και με ένα κλικ εγκρίνει (authorize), παραμετροποιεί ή απορρίπτει την τελική εκτέλεση του API στο ERP.

4. Βρόχος Συνεχούς Εκπαίδευσης (Decentralized MLOps)

Η διασφάλιση της συνεχούς προσαρμοστικότητας του συστήματος εδράζεται στον πυρήνα της Επιβλεπόμενης Μάθησης (Supervised Learning) —της μεθοδολογίας, δηλαδή, κατά την οποία ένα μοντέλο εκπαιδεύεται να παράγει ακριβείς προβλέψεις βασισμένο σε ένα σύνολο δεδομένων που φέρουν ορθές, ρητά καθορισμένες ετικέτες (ground-truth). Προκειμένου να υλοποιηθεί αυτή η εξελικτική πορεία στο αναπτυσσόμενο βιομηχανικό δίκτυο, η αρχιτεκτονική συνδυάζει τεχνικές Ενεργού (Active) και Ομοσπονδιακής (Federated) Μάθησης. Η διαδικασία δομείται σε άμεση συνέργεια μεταξύ των Αυτόνομων Πρακτόρων και του Μηχανικού Βαθιάς Μάθησης (ML/DL Engineer), διατηρώντας ενεργή την αναγκαιότητα της ανθρώπινης εποπτείας σε κάθε κρίσιμο στάδιο:

- **Ανίχνευση Εκτός Κατανομής (OOD Detection):** Όταν το CNN εντοπίσει ανωμαλία με δείκτη βεβαιότητας χαμηλότερο του επιτρεπτού ορίου, κατατάσσει το εύρημα ως άγνωστο (UNKNOWN_DEFECT), μεταφορτώνει την εικόνα στο IPFS και εγγράφει το CID στο Blockchain, ειδοποιώντας ταυτόχρονα το σύστημα ενορχήστρωσης.
- **Συνεργατική Επανεκπαίδευση (Human-Agent Synergy):** Ένας κεντρικός MLOps ενορχηστρωτής (Agent) ανακτά το CID, πραγματοποιεί προκαταρκτική ανάλυση (π.χ.

ομαδοποίηση παρόμοιων σφαλμάτων) και προετοιμάζει τα δεδομένα. Αντί να προχωρήσει σε «τυφλή» και πλήρως αυτοματοποιημένη επανεκπαίδευση, ο Πράκτορας λειτουργεί ως βοηθός (Copilot) και παρουσιάζει τα ευρήματα στον Μηχανικό ML/DL. Ο Μηχανικός αναλαμβάνει τον κεντρικό, στρατηγικό έλεγχο: αξιολογεί τα οπτικά δεδομένα, επαληθεύει ή διορθώνει τις προτεινόμενες ετικέτες (annotations) και καθοδηγεί τη διαδικασία μικρο-ρύθμισης (fine-tuning). Αυτή η συνέργεια διασφαλίζει ότι το μοντέλο μαθαίνει την πραγματική γεωμετρική διαφοροποίηση, αποτρέποντας την αφομοίωση λανθασμένων μοτίβων.

- **Ασφαλής Ενημέρωση (Consensus):** Μετά την τελική έγκριση του επικαιροποιημένου μοντέλου από τον Μηχανικό, τα νέα υπολογιστικά βάρη (weights) του δικτύου διανέμονται πίσω στις Edge συσκευές. Η διανομή πραγματοποιείται αποκλειστικά κατόπιν κρυπτογραφικής συναίνεσης μέσω του Blockchain, αποτρέποντας επιθέσεις "δηλητηρίασης" των δεδομένων (data poisoning) και διασφαλίζοντας την ακεραιότητα του βιομηχανικού δικτύου.

Ουσιαστικά, η αποκεντρωμένη αρχιτεκτονική Blockchain λειτουργεί ως το αποκλειστικό και αδιάβλητο κανάλι επικοινωνίας μεταξύ των κατανεμημένων μοντέλων. Διαμέσου αυτού του δικτύου, το σύνολο της ανταλλασσόμενης πληροφορίας διατηρείται αυστηρά κρυπτογραφημένο, προσφέροντας ταυτόχρονα ένα αμετάβλητο μητρώο ιχνηλασιμότητας (backtracking log). Αυτή η υποδομή εγγυάται την απόλυτη διαφάνεια και την πλήρη ιστορική αναδρομή για κάθε μεταβολή, ενημέρωση ή επανεκπαίδευση του συστήματος, θωρακίζοντας πλήρως τον κύκλο ζωής της τεχνητής νοημοσύνης.

5.5 Επίλογος

Η αφετηρία της παρούσας διατριβής χαρακτηρίστηκε από μια σημαντική επιστημονική πρόκληση: την οξεία έλλειψη διαθέσιμων δεδομένων και την απουσία εκτεταμένης προγενέστερης βιβλιογραφίας στον τομέα της αναγνώρισης βιομηχανικών προφίλ αλουμινίου —με μόλις μία καταγεγραμμένη σχετική έρευνα στο παρελθόν. Παρά τον αρχικό αυτό περιορισμό, η μελέτη ολοκλήρωσε την εκτενή αποτίμηση των αρχιτεκτονικών και των μεθοδολογιών, καταλήγοντας στην επιτυχή υλοποίηση ενός αυτόνομου και προσιτού συστήματος μηχανικής όρασης, προσαρμοσμένου στο απαιτητικό βιομηχανικό περιβάλλον.

Η θεωρητική και αλγοριθμική ανάπτυξη ενσαρκώθηκε πρακτικά μέσω του «Κινητού Εργαστηρίου» (αξιοποιώντας το πλαίσιο Flutter, μοντέλα TensorFlow Lite και υποδομές Google Cloud Storage). Η συγκριτική αξιολόγηση των αποτελεσμάτων ανέδειξε ξεκάθαρα πως η βέλτιστη ισορροπία μεταξύ υπολογιστικής ταχύτητας και διαγνωστικής ακρίβειας επιτυγχάνεται μέσω της διαδοχικής αρχιτεκτονικής YOLOv8n + EfficientNetV2. Ο συγκεκριμένος συνδυασμός κατέγραψε την υψηλότερη συνολική ακρίβεια συστήματος (76.19%), απαιτώντας ταυτόχρονα μόλις 717 ms ανά κύκλο συμπερασμού (ρυθμός ανανέωσης 1.39 FPS). Για την προβλεπόμενη λειτουργία της κινητής εφαρμογής ως εργαλείο "Point-and-Shoot" (ταυτοποίηση μέσω λήψης μεμονωμένης φωτογραφίας) από τον χειριστή ή τον ελεγκτή αποθήκης, αυτός ο χρόνος απόκρισης —σταθερά μικρότερος του ενός δευτερολέπτου (<1000 ms)— κρίνεται απολύτως ικανοποιητικός. Εμπίπτει στα βέλτιστα όρια για την ομαλή αλληλεπίδραση ανθρώπου-υπολογιστή (Human-Computer Interaction - HCI), καθιστώντας τον

προτεινόμενο συνδυασμό πλήρως βιώσιμο και έτοιμο για παραγωγή (production-ready) στο περιβάλλον Android.

Κεντρικός πυλώνας αυτής της επιτυχίας αποτέλεσε η στρατηγική αναγνώριση της γεωμετρίας του προφίλ ως της «μοναδικής αλήθειας» (ground truth). Παραμερίζοντας επιφανειακά χαρακτηριστικά όπως το χρώμα ή η υφή —τα οποία αλλοιώνονται εύκολα στον πραγματικό κόσμο— το σύστημα στηρίχθηκε αποκλειστικά στα αυστηρά δομικά όρια που ορίζονται από τα αρχεία CAD. Αυτή η αυστηρή προσήλωση στη γεωμετρική ακεραιότητα ανοίγει νέους, πολλά υποσχόμενους δρόμους στην έρευνα, ειδικά όταν συνδυαστεί με την κλιμάκωση του συστήματος.

Κοιτώντας προς το μέλλον, τα αυτόνομα συστήματα όρασης της Βιομηχανίας 4.0 καλούνται να διαχειριστούν τεράστιους όγκους δεδομένων από πολλαπλές γεωγραφικές τοποθεσίες, γεγονός που δημιουργεί προκλήσεις διαχείρισης και ασφάλειας. Η προτεινόμενη σύζευξη της μηχανικής όρασης (Edge AI) με Κατανεμημένα Καθολικά (Blockchain) και Αυτόνομους Πράκτορες (LLM Agents) προσφέρει μια κρίσιμη διέξοδο. Μια τέτοια αποκεντρωμένη αρχιτεκτονική δεν εξασφαλίζει απλώς την κρυπτογραφική ασφάλεια των δεδομένων εκπαίδευσης. Επιτρέπει την αποτελεσματική διαχείριση και κατανεμημένη επεξεργασία του μαζικού όγκου ανωμαλιών που καταγράφονται καθημερινά, αποτρέποντας τη συμφόρηση των κεντρικών διακομιστών και επιτρέποντας την ασφαλή Ομοσπονδιακή Μάθηση με απόλυτη ιχνηλασιμότητα.

Εν κατακλείδι, η επιτακτική ανάγκη για την επίλυση απτών, καθημερινών προβλημάτων —τόσο στο αυστηρό πλαίσιο της βιομηχανικής παραγωγής όσο και στον ευρύτερο κοινωνικό ιστό— καθιστά την εφαρμοσμένη τεχνολογική έρευνα πιο επίκαιρη και αναγκαία από ποτέ. Παρά τον σύγχρονο απόηχο που περιβάλλει τα ογκώδη, παραμετρικά Μεγάλα Γλωσσικά Μοντέλα (LLMs) —τα οποία συχνά λειτουργούν ως αδιαφανή «μαύρα κουτιά» (black boxes), διαχειρίσιμα σχεδόν αποκλειστικά από ένα εξαιρετικά περιορισμένο ολιγοπώλιο τεχνολογικών κολοσσών— η παρούσα μελέτη καταδεικνύει μια διαφορετική πραγματικότητα. Η τεχνολογία των αμιγώς Συνελικτικών Νευρωνικών Δικτύων (CNNs) διαθέτει ακόμη τεράστιο, ανεκμετάλλευτο χώρο για έρευνα και ουσιαστική βιομηχανική αξιοποίηση. Ως μια αλγοριθμικά πιο ώριμη, υπολογιστικά ελαφριά και βαθιά παραμετροποιήσιμη λύση, προσφέρει εφαρμόσιμες απαντήσεις στο άκρο του δικτύου (Edge), εκεί όπου τα συγκεντρωτικά, βαριά μοντέλα αδυνατούν να επιχειρήσουν.

Μέσα σε αυτό το τεχνολογικό τοπίο, η αρμονική συνεργασία μεταξύ ανθρώπου και μηχανής, όπως αναδείχθηκε από την παρούσα προσέγγιση, αποδεικνύει πως η τεχνητή νοημοσύνη δεν αποσκοπεί στην αντικατάσταση του ανθρώπινου παράγοντα, αλλά στην ουσιαστική γνωστική και επιχειρησιακή του ενδυνάμωση. Σε αυτή τη συνέργεια, το προσαρμοσμένο σύστημα αναλαμβάνει τον βαρύ και επαναλαμβανόμενο υπολογιστικό φόρτο, απελευθερώνοντας τον χειριστή ώστε να διατηρεί τον αδιαπραγμάτευτο στρατηγικό έλεγχο και την τελική λήψη αποφάσεων. Εν τέλει, η σύζευξη αυτή καταδεικνύει ότι η μηχανική όραση αποτελεί απλώς τον εκτελεστικό βραχίονα, θέτοντας ως μοναδικό και απόλυτο άνω όριο στην εξέλιξη και εφαρμογή αυτών των συστημάτων την ίδια την ανθρώπινη φαντασία και δημιουργικότητα. Αυτή ακριβώς η αλληλεπίδραση αποτελεί τον θεμελιώδη πυλώνα για την ομαλή μετάβαση στις έξυπνες, ευέλικτες και πρωτοπόρες βιομηχανικές μονάδες του μέλλοντος.

Παράρτημα Α: Η πειραματική ανάπτυξη και η αξιολόγηση του συστήματος πραγματοποιήθηκαν σε τρία διακριτά υπολογιστικά περιβάλλοντα, προκειμένου να καλυφθούν οι ποικίλες απαιτήσεις των φάσεων της έρευνας.

Η αρχική υλοποίηση, η συγγραφή του κώδικα και οι προκαταρκτικές δοκιμές εκτελέστηκαν σε τοπικό σταθμό εργασίας, εξοπλισμένο με επεξεργαστή AMD Ryzen 7 3750H (2.30 GHz), 16 GB κύριας μνήμης RAM και διακριτή μονάδα επεξεργασίας γραφικών NVIDIA GeForce GTX 1650 (4 GB VRAM).

Δεδομένων των αυξημένων υπολογιστικών απαιτήσεων για την εκπαίδευση των νευρωνικών δικτύων, η κύρια φάση της εκμάθησης μεταφέρθηκε σε περιβάλλον υπολογιστικού νέφους (cloud computing). Η συγκεκριμένη υποδομή αξιοποίησε 9 εικονικούς πυρήνες (vCPUs), 50 GB μνήμης RAM, 40 GB αποθηκευτικού χώρου, καθώς και έναν ισχυρό επιταχυντή γραφικών NVIDIA RTX 3090 με 24 GB αποκλειστικής μνήμης βίντεο (VRAM), ελαχιστοποιώντας δραστικά τον χρόνο εκπαίδευσης.

Τέλος, η πρακτική εφαρμογή και η τελική αξιολόγηση του συστήματος σε συνθήκες πραγματικού χρόνου (inference) πραγματοποιήθηκαν μέσω του «κινητού εργαστηρίου», αξιοποιώντας τη φορητή συσκευή POCO X6 5G, επιβεβαιώνοντας έτσι την ομαλή και αποδοτική λειτουργία των μοντέλων σε αρχιτεκτονικές Edge AI.

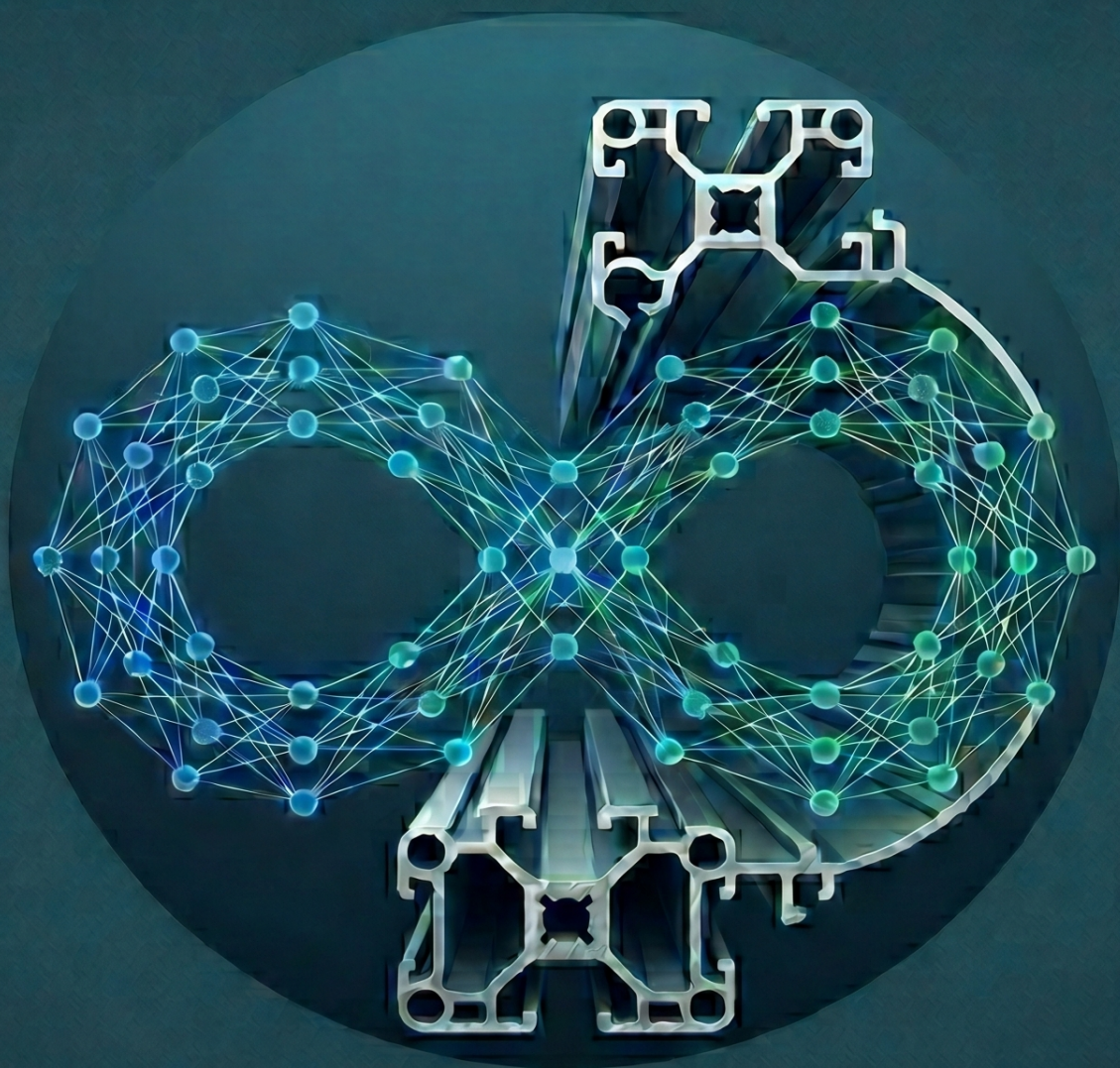
Παράρτημα Β: https://github.com/KafestidisStelios/signal_profiler

Βιβλιογραφία

- **Βιβλιογραφία**
- **Andreas, J.** (2019). *Good-Enough Compositional Data Augmentation*. arXiv. <https://arxiv.org/pdf/1904.09545>
- **Deep learning based defect detection in aluminium castings using radiographic imaging.** (2022). *Computers in Industry*, 137, 103612. <https://doi.org/10.1016/j.compind.2022.103612>
- **Flutter Gems.** (n.d.). *Top Flutter packages for machine learning*. <https://fluttergems.dev/machine-learning/>
- **Flutter Official Documentation.** (n.d.). *Take a picture using the camera*. <https://docs.flutter.dev/cookbook/plugins/picture-using-camera>
- **Laskin, M., Lee, K., Stooke, A., Pinto, L., Abbeel, P., & Srinivas, A.** (2020). *Reinforcement Learning with Augmented Data*. arXiv. <https://arxiv.org/pdf/2004.14990>
- **Li, S., & Ré, C.** (2020, February 26). *Automating the Art of Data Augmentation: Part I Overview*. Hazy Research. <https://hazyresearch.stanford.edu/blog/2020-02-26-data-augmentation-part1>
- **Li, Q., & Chen, G.** (2021). Recognition of industrial machine parts based on transfer learning with convolutional neural network. *PLOS ONE*, 16(1), e0245735. <https://www.researchgate.net/publication/348852652>
- **Li, Y., Dong, X., Chen, C., Li, J., Wen, Y., Spranger, M., & Lyu, L.** (2024). *Is Synthetic Image Useful for Transfer Learning? An Investigation into Data Generation, Volume, and Utilization*. arXiv:2403.19866.
- **Lightly.ai.** (n.d.). *An Introduction to Transfer Learning*. <https://www.lightly.ai/blog/transfer-learning>
- **Liu, X., He, W., Zhang, Y., Yao, S., & Cui, Z.** (2021). Effect of dual-convolutional neural network model fusion for Aluminum profile surface defects classification and recognition. *Mathematical Biosciences and Engineering*, 19(1), 997-1025. <https://doi.org/10.3934/mbe.2022046>
- **Mazzeo, P.-L., Argentieri, A., & De Luca, F.** (2019). Convolutional neural networks for recognition and segmentation of aluminum profiles. *14th International Conference on Computer Vision Theory and Applications (VISAPP)*.
- **Occlusion in machine vision systems: Challenges & solutions.** (n.d.). *ScienceDirect Topics*. <https://www.sciencedirect.com/topics/computer-science/occlusion-problem>

- **Rodríguez-Rodríguez, J. A., López-Rubio, E., Ángel-Ruiz, J. A., & Molina-Cabello, M. A.** (2024). On the impact of noise and brightness on the performance of object detection methods. *Sensors*, 24(3), 821.
<https://pmc.ncbi.nlm.nih.gov/articles/PMC10856852/pdf/sensors-24-00821.pdf>
- **Shorten, C., & Khoshgoftaar, T. M.** (2019). A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, 6(60).
- **TensorFlow Official Documentation.** (n.d.). *TensorFlow Lite: Bringing AI to Mobile and Embedded Devices*. <https://www.tensorflow.org/lite>
- **Tobin, J., Fong, R., Ray, A., Schneider, J., Zaremba, W., & Abbeel, P.** (2017). Domain Randomization for Transferring Deep Neural Networks from Simulation to the Real World. *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.
- **UnitX.** (n.d.). *Exploring the Concept of Occlusion in Vision Systems*.
- **Wang, L., & Cui, L.** (2022). Automatic recognition and detection system based on machine vision. *Journal of Control Science and Engineering*, 2022.
<https://www.researchgate.net/publication/361814743>
- **Weng, L.** (2019, May 5). *Domain Randomization for Sim2Real Transfer*. Lil'Log.
<https://lilianweng.github.io/posts/2019-05-05-domain-randomization/>
-

«Δηλώνω ρητά ότι, σύμφωνα με το άρθρο 8 του Ν. 1599/1986 η παρούσα εργασία αποτελεί αποκλειστικά προϊόν προσωπικής εργασίας και δεν προσβάλλει κάθε μορφής δικαιώματα διανοητικής ιδιοκτησίας, προσωπικότητας και προσωπικών δεδομένων τρίτων, δεν περιέχει έργα/εισφορές τρίτων για τα οποία απαιτείται άδεια των δημιουργών/ δικαιούχων και δεν είναι προϊόν μερικής ή ολικής αντιγραφής, οι πηγές δε που χρησιμοποιήθηκαν περιορίζονται στις βιβλιογραφικές αναφορές και μόνον και πληρούν τους κανόνες της επιστημονικής παράθεσης.»



Email: kafestidistelios@gmail.com
GitHub: <https://github.com/KafestidisStelios>

