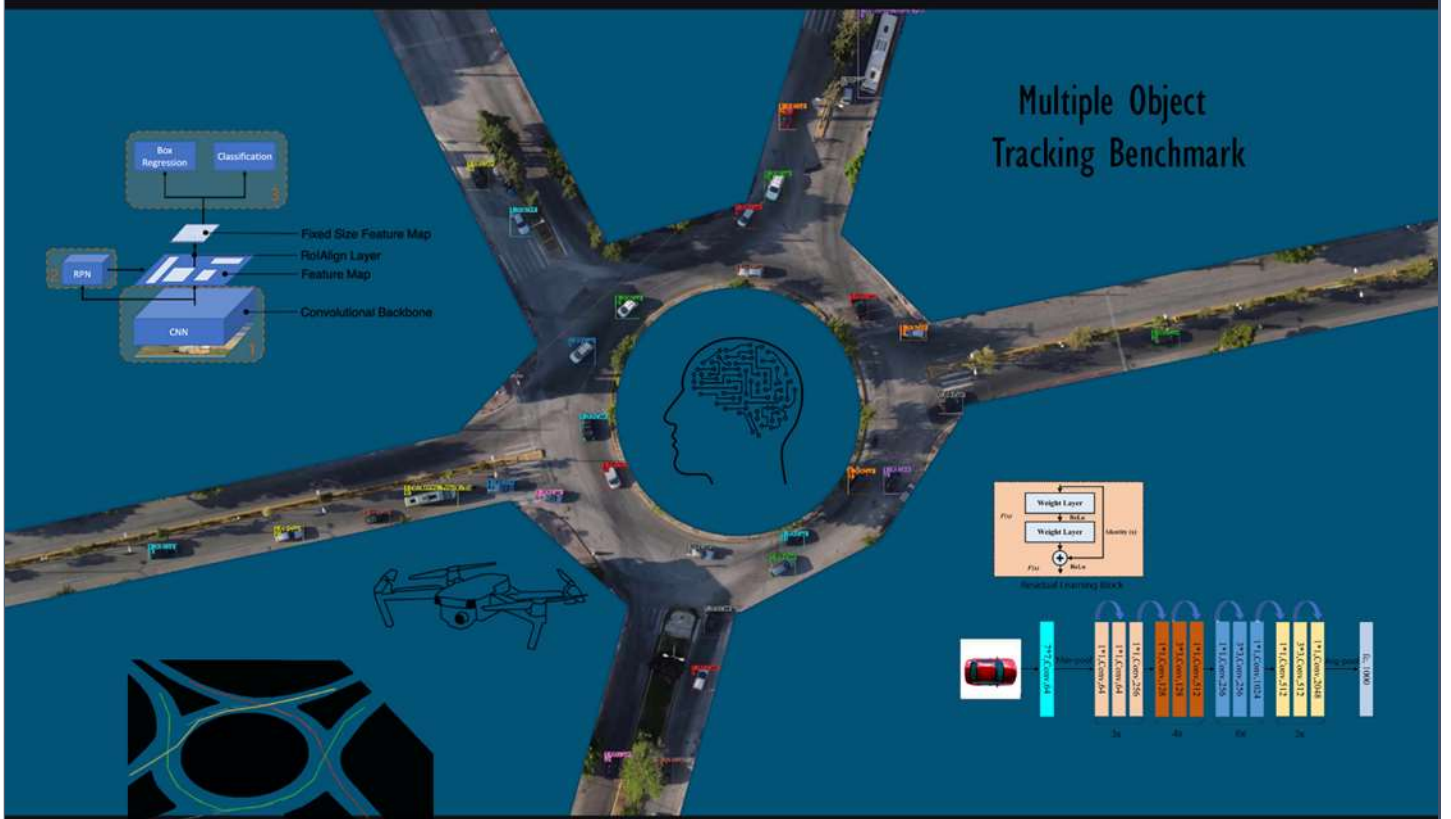




Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Αγρονόμων και Τοπογράφων Μηχανικών –
Μηχανικών Γεωπληροφορικής
Τομέας Τοπογραφίας



Διπλωματική Εργασία:

«Ανάπτυξη μεθοδολογίας εξαγωγής κυκλοφοριακών δεδομένων μέσω τεχνικών Παρακολούθησης Πολλαπλών Αντικειμένων από UAV»

Συντάξας: Φλωρόπουλος Δημήτριος (Α.Μ.: rs18007)

Επιβλέπων: Καραντζαλος Κωνσταντίνος, Καθηγητής

Συνεπιβλέπουσα: Σπυροπούλου Ιωάννα, Αναπληρώτρια Καθηγήτρια

Αθήνα, Σεπτέμβριος 2023

Τριμελής Εξεταστική Επιτροπή

Καράντζαλος Κωνσταντίνος

Καθηγητής

Σπυροπούλου Ιωάννα

Αναπληρώτρια Καθηγήτρια

Κεπαπτσόγλου Κωνσταντίνος

Αναπληρωτής Καθηγητής

Three-member Committee

Karantzas Konstantinos

Professor

Spyropoulou Ioanna

Associate Professor

Kepatsoglou Konstantinos

Associate Professor

ΔΗΛΩΣΗ ΑΝΑΓΝΩΡΙΣΗΣ ΠΝΕΥΜΑΤΙΚΩΝ ΔΙΚΑΙΩΜΑΤΩΝ

Δηλώνω ότι η διπλωματική αυτή εργασία αποτελεί στο σύνολό της δική μου εργασία, και κανένα τμήμα της δεν έχει χρησιμοποιηθεί για την κτήση άλλου τίτλου σπουδών. Όπου έχει χρησιμοποιηθεί υλικό από άλλες πηγές, αυτές έχουν αναφερθεί με ακρίβεια και πληρότητα.

Υπογραφή Φοιτητή,



Φλωρόπουλος Δημήτριος

Ευχαριστίες

Θα ήθελα να ευχαριστήσω θερμά τους επιβλέποντες καθηγητές κ. Καράντζαλο Κωνσταντίνο και κ. Σπυροπούλου Ιωάννα, ο τρόπος διδασκαλίας των οποίων με προέτρεψε έντονα να μελετήσω ένα ενδιαφέρον project που συνδυάζει τεχνογνωσία της Ροής Τοπογραφίας II στο πεδίο εφαρμογής Συγκοινωνιακής Τεχνικής, ενώ από το Εργαστήριο Τηλεπισκόπησης, θα ήθελα επίσης να ευχαριστήσω τους υποψήφιους διδάκτορες Τσιρώνη Βασίλη και Ψαλλά Αθηνά, για τη συνεχή καθοδήγηση στα τεχνικά ζητήματα της εργασίας. Η ανταπόκριση και επικοινωνία υπήρξε άριστη εκ μέρους τους, και με βοήθησε σημαντικά καθ' όλο το διάστημα εκπόνησης της εργασίας.

Θερμές ευχαριστίες οφείλω προς τους στενούς φίλους και συναδέλφους μου Γκετσόπουλο Πέτρο, Παπαδοπούλου Παρασκευή, και Σπανό Σωτήρη, με τους οποίους συζητούσα συχνά για την εργασία και μου παρείχαν πολύτιμες ιδέες και συμβουλές σε ζητήματα οπτικοποίησης/γραφιστικής, προγραμματισμού, και γενικότερης διαχείρισης.

Τέλος, ένα ιδιαίτερο ευχαριστώ προς τα μέλη της οικογένειάς μου, καθώς υπήρξαν μέντορές μου στο επάγγελμα του Τοπογράφου Μηχανικού, και στέκονται πλάι μου ως στήριγμα σε κάθε βήμα της ζωής μου. Ευχαριστώ γενικότερα φίλους και συγγενείς για την αμέριστη υποστήριξή τους τα τελευταία χρόνια.

Φλωρόπουλος Δημήτριος,

Αθήνα, Σεπτέμβριος 2023

Περίληψη

Εν έτει 2023 η Τεχνητή Νοημοσύνη έχει σημειώσει σημαντική πρόοδο ώστε να καταφέρει να συμβάλει αξιόλογα σε πεδία εφαρμογών κάθε επιστημονικού τομέα, συμπεριλαμβανομένης της Μηχανικής Μεταφορών. Στην παρούσα Διπλωματική Εργασία αναπτύχθηκε μια δυσκολής μεθοδολογία, με τελικό στόχο την αξιολόγηση του επιπέδου εξυπηρέτησης αστικών ισόπεδων οδικών κόμβων. Στο πρώτο της στάδιο, η εν λόγω μεθοδολογία αποτελούνταν από ένα σύστημα τεχνικών Βαθιάς Μάθησης με υλοποιήσεις ανοιχτού κώδικα του οργανισμού OpenMMLab, το οποίο έφερε σε πέρας την εφαρμογή του αλγόριθμου Παρακολούθησης Πολλαπλών Αντικειμένων «DeepSORT» σε κυκλοφοριακές λήψεις βίντεο από UAVs, ορίζοντας τα κινούμενα οχήματα ως στόχους ενδιαφέροντος, υπό ταξινόμηση σε τέσσερις κατηγορίες. Έπειτα, στο δεύτερο της στάδιο, οι εξαγόμενες τροχιές των οχημάτων αξιοποιήθηκαν από προσαρμοσμένους προγραμματιστικούς μηχανισμούς οι οποίοι καταμετρούσαν τις διελεύσεις με πληροφορία προέλευσης – προορισμού, ώστε να προκύπτει ανάθεση κυκλοφοριακού φόρτου ζήτησης στα εκάστοτε ρεύματα κυκλοφορίας. Με διαθέσιμη την «κατανεμημένη» εκτίμηση της ζήτησης, ακολουθήθηκε εκτέλεση του τυπολογίου λειτουργικής ανάλυσης κατά Highway Capacity Manual 2010, για τον τελικό χαρακτηρισμό των επιπέδων εξυπηρέτησης. Τα παραπάνω εφαρμόστηκαν στα UAV πλάνα του Suncity Dataset, προερχόμενα από περιοχές πολυσύχναστων οδικών κόμβων στη συνοικία Ηλιούπολη Αττικής. Συμπερασματικά, τα ενθαρρυντικά αποτελέσματα στις ποσοτικές μετρικές του MOT Benchmark, στους συντιθέμενους δείκτες κυκλοφοριακού φόρτου, αλλά και στους ποιοτικούς ελέγχους που πραγματοποιήθηκαν, επικύρωσαν την τεχνική βιωσιμότητα της χρήσης UAV μέσω του πεδίου των μετρήσεων κοινωνικών μελετών. Μάλιστα, διακρίθηκαν σημαντικά συγκριτικά πλεονεκτήματα που την καθιστούν ιδιαίτερα ανταγωνιστική μεταξύ άλλων καθιερωμένων μεθόδων.

Λέξεις Κλειδιά: Τεχνητή Νοημοσύνη, Μηχανική Μάθηση, Βαθιά Μάθηση, Όραση Υπολογιστών, Παρακολούθηση Πολλαπλών Αντικειμένων, Ανίχνευση Αντικειμένων, Συγκοινωνιακή Τεχνική, επίπεδο εξυπηρέτησης, Μη Επανδρωμένα Ιπτάμενα Οχήματα (UAVs)

Abstract

In the year 2023, Artificial Intelligence has made significant progress to achieve remarkable contributions to fields of application in every scientific field, including Transportation Engineering. In this Diploma Thesis a two-stage methodology was developed, with the final objective of evaluating the level of service at urban level crossings. At its first stage, the methodology was consisted of a Deep Learning techniques system with open-source implementations by the OpenMMLab organization, which carried out the application of the "DeepSORT" Multiple Object Tracking algorithm in traffic video captures from UAVs, defining moving vehicles as targets of interest, classified into four categories. Then, at its second stage, the extracted vehicle trajectories were exploited by customized programming mechanisms that counted the crossings with origin-destination information to derive traffic demand assignments to the respective traffic streams. With the assessment of the "allocated" demand available, the operational analysis template of the Highway Capacity Manual 2010 was executed, for the final designations of the levels of service. The above were applied to the UAV footages of the Suncity Dataset, derived from areas of busy road junctions in the Ilioupolis district of Attica. Concluding, the encouraging results in the quantitative metrics of the MOT Benchmark, in the synthesized traffic volume indicators, as well as in the qualitative supervision carried out, validated the technical viability of the use of UAVs in the field of transportation survey measurements. In fact, significant comparative advantages were distinguished that make it highly competitive amongst other established methods.

Keywords: Artificial Intelligence, Machine Learning, Deep Learning, Computer Vision, Multiple Object Tracking, Object Detection, Transportation Engineering, Level of Service, Unmanned Aerial Vehicles (UAVs)

Πίνακας Περιεχομένων

Ευχαριστίες	9
Περίληψη.....	10
Abstract	11
Κεφάλαιο 1 ^ο : Εισαγωγή	20
1.1: Αφορμές προσέγγισης θέματος.....	20
1.2: Συνεισφορά του πεδίου επιστήμης υπολογιστών	20
1.3: Παράθεση γενικών αναμενόμενων οφελών και προκλήσεων	23
1.4: Διάρθρωση διπλωματικής εργασίας.....	24
Κεφάλαιο 2 ^ο : Επιστημονικό υπόβαθρο	25
2.1: Θεωρητικό σκέλος πληροφορικής.....	25
2.1.1: Δομή τεχνητών νευρωνικών δικτύων	25
2.1.2: Συνελκτικά νευρωνικά δίκτυα	29
2.1.3: Ανίχνευση αντικειμένων.....	34
2.1.4: Παρακολούθηση αντικειμένων	38
2.2: Θεωρητικό σκέλος κυκλοφοριακής τεχνικής	43
2.2.1: Βασικά μεγέθη κυκλοφοριακής ροής	43
2.2.2: Θεμελιώδεις κι εμπειρικές σχέσεις μεταξύ κυκλοφοριακών μεγεθών.....	45
2.2.3: Μελέτη κυκλοφοριακής ικανότητας κι επιπέδου εξυπηρέτησης.....	47
2.3: Σχετικές έρευνες μελέτης κυκλοφορίας με χρήση UAV και DL.....	56
Κεφάλαιο 3 ^ο : Περιγραφή δεδομένων μελέτης.....	58
3.1: Τεχνικά χαρακτηριστικά	58
3.2: Ποιοτικά χαρακτηριστικά, προκλήσεις και ιδιομορφίες	61
Κεφάλαιο 4 ^ο : Μεθοδολογία.....	64
4.1: Διαθέσιμα υλικά μέσα και προγραμματιστικό πλαίσιο	64
4.2: Φάση εκπαίδευσης ανιχνευτή (Detector)	66
4.3: Φάση εκπαίδευσης επαναταυτοποίησης (ReID).....	76
4.4: Παραμετροποίηση μοντέλου παρακολούθησης (Tracker).....	81
4.5: Προσαρμοσμένοι αλγόριθμοι κυκλοφοριακής τεχνικής	85
4.5.1: Συνάρτηση αξιοποίησης των δεδομένων εξόδου του tracker.....	85
4.5.2: Μονάδες εκτίμησης κυκλοφοριακών μεγεθών	88
4.5.3: Μονάδα χαρακτηρισμού επιπέδου εξυπηρέτησης.....	97
4.6: Μετρικές ποσοτικής αξιολόγησης	100
4.6.1: Μετρικές ανίχνευσης αντικειμένων	100
4.6.2: Μετρικές παρακολούθησης πολλαπλών αντικειμένων	102
4.6.3: Πρόσθετες μετρικές κυκλοφοριακού φόρτου	105
Κεφάλαιο 5 ^ο : Τελικά αποτελέσματα – Συμπεράσματα.....	106

5.1: Αποτελέσματα και αξιολόγηση μεθοδολογίας Βαθιάς Μάθησης	106
5.1.1: Επιδόσεις συνιστώσας ανίχνευσης	106
5.1.2: Επιδόσεις μοντέλου Παρακολούθησης Πολλαπλών Αντικειμένων	112
5.1.3: Περιπτώσεις αστοχίας	116
5.2: Αποτελέσματα και αξιολόγηση κυκλοφοριακής προσέγγισης	119
5.2.1: Αποτελέσματα κυκλοφοριακών μεγεθών και σχετικών μετρικών	120
5.2.2: Τελική κατάταξη σε στάθμη εξυπηρέτησης	124
5.3: Αποτίμηση: προοπτικές βελτίωσης και συμπεράσματα	127
Βιβλιογραφικές αναφορές	131

Ευρετήριο Σχημάτων

Σχήμα 1: Η τυπική μορφή μιας δομής πολυεπίπεδων νευρώνων (multi-layer perceptron) με 4 συνολικές στρώσεις, 1 εισόδου, 2 κρυφές και 1 εξόδου. (πηγή: https://medium.com/codex/introduction-to-how-an-multilayer-perceptron-works-but-without-complicated-math-a423979897ac , προσπελάστηκε Αυγ. 2023).....	25
Σχήμα 2: Παράδειγμα εφαρμογών (2x2) max, average και sum pooling σε έναν απλό πίνακα ακέραιων τιμών. (πηγή: https://www.shiksha.com/online-courses/articles/a-comprehensive-guide-to-convolutional-neural-networks/ , προσπελάστηκε Αυγ. 2023).....	29
Σχήμα 3: Η τυπική δομή ενός Συνελκτικού Νευρωνικού Δικτύου (Convolutional Neural Network, CNN). (πηγή: https://nafizshahriar.medium.com/what-is-convolutional-neural-network-cnn-deep-learning-b3921bdd82d5 , προσπελάστηκε Αυγ. 2023).....	30
Σχήμα 4: Σχεδιάγραμμα της αρχιτεκτονικής LeNet-5. (πηγή: [9]).....	31
Σχήμα 5: Σχεδιάγραμμα της αρχιτεκτονικής AlexNet. (πηγή: [10]).....	31
Σχήμα 6: Σχεδιαγράμματα των παραλλαγών της αρχιτεκτονικής VGG Net, VGG-16 και VGG-19. (πηγή: https://datahacker.rs/deep-learning-vgg-16-vs-vgg-19/ , προσπελάστηκε Αυγ. 2023).....	32
Σχήμα 7: Σχεδιάγραμμα της αρχιτεκτονικής GoogLeNet/Inception-v1. (πηγή: https://medium.com/analytics-vidhya/cnns-architectures-lexnet-alexnet-vgg-googlenet-resnet-and-more-666091488df5 , προσπελάστηκε Αυγ. 2023).....	32
Σχήμα 8: Σχεδιάγραμμα της αρχιτεκτονικής ResNet-50 (κάτω), καθώς και του Residual Learning Blog (άνω). (πηγή: [42]).....	33
Σχήμα 9: Παραδείγματα εξαγωγής «up-right» bounding box σε αυτοκίνητα ως αντικείμενα ενδιαφέροντος. (πηγή: https://www.precisebposolution.com/bounding-box.html , προσπελάστηκε Αυγ. 2023).....	34
Σχήμα 10: Παραδείγματα oriented object detection (αριστερά) και instance segmentation (δεξιά). (πηγές: https://developer.nvidia.com/blog/detecting-rotated-objects-using-the-odtk/ , https://datascience.eu/computer-vision/mask-r-cnn/ , προσπελάστηκαν Αυγ. 2023).....	34
Σχήμα 11: Σχεδιάγραμμα της συνολικής αρχιτεκτονικής Faster R-CNN, που σχηματίζεται με τη συμβολή των ξεχωριστών «παράλληλων» υπο-δικτύων RPN και Fast R-CNN detector. (πηγή: [42]).....	37
Σχήμα 12: Η αρχιτεκτονική του CNN που χρησιμοποιείται κατά τον πρωτότυπο αλγόριθμο DeepSORT για την απόκτηση των περιγραφών εμφάνισης των αντικειμένων. (πηγή: [16]).	41
Σχήμα 13: Σύνοψη της πρωτότυπης αλγοριθμικής αλληλουχίας αντιστοίχισης μεταξύ ανιχνεύσεων και υπαρχόντων τροχιών στον DeepSORT tracker. (πηγή: [16]).....	42
Σχήμα 14: Στο διάγραμμα φθίνουσας κατάταξης ωρών αιχμής ως προς το μέγεθος ωριαίου φόρτου, η χαρακτηριστική ώρα σχεδιασμού επιλέγεται μεταξύ της 30 ^{ης} και 100 ^{ης} εξ αυτών. Εντός αυτής της περιοχής (με κόκκινο στο σχήμα) παρατηρείται ότι οι καμπύλες παρουσιάζουν έντονη κυρτότητα, οπότε και αποκλείονται του σχεδιασμού οι ακραίοι αλλά ιδιαίτερα σπάνιοι φόρτοι (~0.3-1% της διάρκειας του έτους). Η «EMHK» που εμφανίζεται στον άξονα y αφορά την Ετήσια Μέση Ημερήσια Κυκλοφορία, δηλαδή το συνολικό αριθμό διερχόμενων οχημάτων σε ένα έτος διά τις μέρες του έτους. (πηγή: [23]).....	43
Σχήμα 15: Σχηματική απεικόνιση του χωρικού διαχωρισμού μεταξύ δύο οχημάτων. (πηγή: [23]).....	45
Σχήμα 16: Διαγράμματα σχέσεων κυκλοφοριακών μεγεθών (ταχύτητας-πυκνότητας, ταχύτητας-φόρτου, φόρτου-πυκνότητας), όπως αυτά προκύπτουν από τις θεωρητικές εξισώσεις κυκλοφοριακής ροής. (πηγή: [23]).....	46
Σχήμα 17: Αντιπαραβολή διαγραμμάτων σχέσεων κυκλοφοριακών μεγεθών (ταχύτητας-πυκνότητας, ταχύτητας-φόρτου, φόρτου-πυκνότητας), όπως αυτά προκύπτουν από εμπειρικά αποτελέσματα κυκλοφοριακών παρατηρήσεων, έναντι των αντίστοιχων θεωρητικών διαγραμμάτων. (πηγή: [23]).....	47

Σχήμα 18: Σχεδιάγραμμα χωροθέτησης αφικνούμενου φόρτου (entry flow), και εμπλεκόμενου φόρτου (circulating flow) σε έναν κυκλικό κόμβο 2 λωρίδων εντός του δακτυλίου, και 2 λωρίδων επί μιας πρόσβασης. (πηγή: [24]).....	54
Σχήμα 19: Δείγματα μερικών καρτέ από το Suncity Dataset, καλύπτοντας τις 6 διαφορετικές εναλλακτικές μεταξύ φωτισμού/ώρας λήψης και τοποθεσίας. Η αριστερή στήλη περιλαμβάνει τις λήψεις των 8 π.μ., ενώ η δεξιά των 11 π.μ.. Στην πρώτη σειρά απεικονίζεται ο κόμβος Ηρώων Πολυτεχνείου – Παύλου Πεντάρη, στη δεύτερη ο κόμβος Πρωτόπαπα – Σοφοκλή Βενιζέλου, ενώ στην τρίτη ο κυκλικός κόμβος Πλατείας 25ης Μαρτίου.....	61
Σχήμα 20: Τα περιεχόμενα του οργανισμού ανοιχτού κώδικα OpenMMLab (πηγή: [32]).....	64
Σχήμα 21: Παραδείγματα εμφάνισης ενός δίκυκλου (κίτρινο bounding box) κι ενός αυτοκινήτου (τυρκουάζ bounding box) στο Suncity Dataset.	66
Σχήμα 22: Παραδείγματα εμφάνισης πολλαπλών αυτοκινήτων στο Suncity Dataset.	66
Σχήμα 23: Διάκριση εμφάνισης μεταξύ της κατηγορίας Semi Truck/ Van (ματζέντα bounding box) και Heavy Truck/Bus (κόκκινο bounding box).....	66
Σχήμα 24: Απόσπασμα από το επιλεγθέν υποσύνολο του Aerial Multi-Vehicle Detection Dataset [29], το οποίο χρησιμοποιήθηκε για έλεγχο επιδόσεων γενίκευσης του παραχθέντος μοντέλου ανίχνευσης. Παρατηρούνται πιο πυκνές συστοιχίες οχημάτων, ειδικά σε περιοχές στάθμευσης, καθώς και σύνθετες συνθήκες φωτισμού. Γενικότερα, τα αντικείμενα εμφανίζονται με ελαφρώς μικρότερο ποσοστό του μεγέθους της εικόνας, απ' ό,τι στο Suncity Dataset.	67
Σχήμα 25: Υποδείγματα εγγραφών περιγραφών στη μορφή του Common Objects in Context (COCO) Dataset [33].....	68
Σχήμα 26: Απεικόνιση του περιβάλλοντος εργασίας του Computer Vision Annotation Tool (CVAT), κατά τη δημιουργία των δεδομένων αλήθειας στο Suncity Dataset.	69
Σχήμα 27: Απόσπασμα από annotation file σε MOT 1.1 format. Από αριστερά προς τα δεξιά διακρίνονται τα εξής: <frame_id>, <instance_id>, <x1>, <y1>, <w>, <h>, <conf> (=1 στα ground truth annotations), <class_id>, <visibility>. Το συγκεκριμένο αρχείο «gt.txt» βρίσκεται εντός τοποθεσίας τύπου «./seq_name/gt/». Ο φάκελος «seq_name» περιέχει επίσης τον υποφάκελο «img1», ο οποίος διαθέτει τα καρτέ στα οποία αντιστοιχούν οι περιγραφές, με ονομασίες που παραπέμπουν στα frame IDs. Ακόμη, στην ίδια τοποθεσία υπάρχει το αρχείο ορισμού ακολουθίας «seqinfo.ini».	76
Σχήμα 28: Οι τομές των bounding boxes ενός αυτοκινήτου στα διαδοχικά καρτέ στα οποία εμφανίζεται σε ένα βίντεο.	77
Σχήμα 29: Υποδείγματα εγγραφών περιγραφών στη μορφή COCOVideo.	83
Σχήμα 30: Το διάγραμμα της αποσυμπιεσμένης έκδοσης του αρχείου εξόδου μορφής pickle. Η τιμή x πράττει προσπέλαση στα καρτέ της ακολουθίας ελέγχου, η τιμή y διαχωρίζει τις εγγραφές - διανύσματα διαφορετικών κατηγοριών εντός ίδιου καρτέ, και τέλος η τιμή w αποτελεί δείκτη για τα στοιχεία στιγμιαίας ανίχνευσης εντός κάθε εγγραφής της μορφής «(instance_id, xmin, ymin, xmax, ymax, confidence)».	85
Σχήμα 31: Περιγραφή της επιθυμητής προγραμματιστικής επεξεργασίας η οποία επιτυγχάνει τη μετατροπή του λεξικού στο Σχήμα 30, σε 4 παράγωγα λεξικά (1 για κάθε τύπο οχήματος) που θα φέρουν εγγραφές στοιχείων τροχιάς της μορφής «(xcenter, ycenter, frame_id)» με πρωτεύον κλειδί το instance_id.....	86
Σχήμα 32: Παρουσίαση των διατομών καταμέτρησης στο κυκλικό κόμβο της Πλατείας 25 ^{ης} Μαρτίου. Με ματζέντα αποτυπώνεται η αθροιστική οπτικοποίηση των «έγκυρων» τροχιών που προήλθαν από το DJI_0073.mp4, οι οποίες φαίνεται ότι τέμνονται ορθά με τις διατομές.	88
Σχήμα 33: Παρουσίαση των διατομών καταμέτρησης στους κόμβους Ηρώων Πολυτεχνείου – Παύλου Πεντάρη (άνω) και Πρωτόπαπα – Σοφοκλή Βενιζέλου (κάτω). Αντίστοιχα, με ματζέντα αποτυπώνεται η αθροιστική οπτικοποίηση των «έγκυρων» τροχιών που προήλθαν από τα DJI_0107.mp4 και DJI_0560.mov.	89

Σχήμα 34: Περιγραφή της μετρικής IoU. (πηγή: https://idiotdeveloper.com/what-is-intersection-over-union-iou/ , προσπελάστηκε Αυγ. 2023).....	100
Σχήμα 35: Καμπύλες Precision – Recall, για διαφορετικά κατώφλια IoU. (πηγή: https://hasty.ai/docs/mp-wiki/metrics/map-mean-average-precision , προσπελάστηκε Αυγ. 2023).....	101
Σχήμα 36: Σύγκριση πληροφορίας αλήθειας (αριστερά), με τις εξαγόμενες ανιχνεύσεις (δεξιά) στον κυκλικό κόμβο Πλατείας 25 ^{ης} Μαρτίου.....	106
Σχήμα 37: Σύγκριση πληροφορίας αλήθειας (αριστερά), με τις εξαγόμενες ανιχνεύσεις (δεξιά) στον κόμβο Πρωτόπαπα – Σοφοκλή Βενιζέλου.....	107
Σχήμα 38: Σύγκριση πληροφορίας αλήθειας (αριστερά), με τις εξαγόμενες ανιχνεύσεις (δεξιά) στον κόμβο Ηρώων Πολυτεχνείου – Παύλου Πεντάρη.....	107
Σχήμα 39: Σύγκριση πληροφορίας αλήθειας (άνω), με τις εξαγόμενες ανιχνεύσεις (κάτω) σε ένα από τα καρέ του υποσυνόλου του Aerial Multi-Vehicle Detection Dataset.....	110
Σχήμα 40: Καμπύλες κόστους ανά εποχή εκπαίδευσης του παραχθέντος μοντέλου ανίχνευσης.....	111
Σχήμα 41: Παραδείγματα οπτικοποίησης εξαγόμενων τροχιών οχημάτων από το βίντεο DJI_0073.mp4.....	112
Σχήμα 42: Παραδείγματα οπτικοποίησης εξαγόμενων τροχιών οχημάτων από το βίντεο DJI_0107.mp4.....	112
Σχήμα 43: Επιτυχής παρακολούθηση της πορείας δικύκλων και μεγάλου φορτηγού στην πρώτη και δεύτερη εικόνα, αντίστοιχα. Στο τρίτο απόσπασμα ο tracker αποτυπώνει άριστα την τροχιά ενός ημιφορτηγού που κάνει ασυνήθιστη αλλαγή κατεύθυνσης κατά την πραγματοποίηση ελιγμού στάθμευσης.....	113
Σχήμα 44: Παραδείγματα οπτικοποίησης εξαγόμενων τροχιών οχημάτων από το βίντεο DJI_0560.mov.....	113
Σχήμα 45: Παραδείγματα αστοχίας με εμφάνιση «στιγμαίων» ανιχνύσεων ID.....	116
Σχήμα 46: Περίπτωση διαχωρισμού τροχιάς υπό διαφορετικά IDs. Επιπλέον έχει μεταβληθεί η πρόβλεψη κατηγορίας οχήματος του ανιχνευτή.....	117
Σχήμα 47: Περίπτωση μεταφοράς του ID = 34 από τον αρχικό στόχο που εξέρχεται των ορίων βιντεοσκόπησης, σε νέο-εισερχόμενο στόχο.....	117
Σχήμα 48: Διάκριση ρευμάτων στον κόμβο Ηρώων Πολυτεχνείου – Παύλου Πεντάρη, όπως ορίζει το υφιστάμενο πρόγραμμα φωτεινής σηματοδότησης. (γεωχωρικό υποβαθρο Google Earth).....	119
Σχήμα 49: Διάκριση ρευμάτων στον κόμβο Πρωτόπαπα – Σοφοκλή Βενιζέλου, όπως ορίζει το υφιστάμενο πρόγραμμα φωτεινής σηματοδότησης. (γεωχωρικό υποβαθρο Google Earth).....	119
Σχήμα 50: Ο κυκλικός κόμβος της Πλατείας 25 ^{ης} Μαρτίου φέρει 5 κλάδους, ο καθένας με 2 λωρίδες στην πρόσβασή του, ενώ υπάρχουν 2 λωρίδες εντός του δακτυλίου. Επομένως κάθε λωρίδα πρόσβασης διαθέτει διακριτή κυκλοφοριακή ικανότητα και αναλύεται ξεχωριστά. Όπως εξηγείται στην επόμενη σελίδα, θεωρείται η κατανομή φόρτου της μορφής «Left-Through» κινήσεων για την αριστερή λωρίδα και «Through-Right» κινήσεων για τη δεξιά. Για το διαμοιρασμό του φόρτου σε 2 λωρίδες προς κοινή «Through» κίνηση, εφαρμόζονται αναλογίες 50-50. Για παράδειγμα, η ζήτηση των οχημάτων που θέλουν να μετακινηθούν από τον Κλάδο 2 στον Κλάδο 4, θεωρείται ισομοιρασμένη στη δεξιά και αριστερή λωρίδα της πρόσβασης. (γεωχωρικό υποβαθρο Google Earth).....	120
Σχήμα 51: Περίπτωση κατά την οποία ένα όχημα δεν κατάφερε να διέλθει του κόμβου Πρωτόπαπα – Σοφοκλή Βενιζέλου, παρότι αντίκρισε 3 κόκκινες ενδείξεις στο Ρεύμα 5 (οι 2 ολοκληρωμένες διάρκειας + 1 περιορισμένης κατά την άφιξή του).....	125
Σχήμα 52: Παρατήρηση συνθηκών κυκλοφοριακής συμφόρησης στον Κλάδο 2, τόσο στον κυκλικό κόμβο (άνω), όσο και στον κόμβο Πρωτόπαπα – Σοφοκλή Βενιζέλου (κάτω).....	125
Σχήμα 53: Περίπτωση «αποτυχημένου» κύκλου στο Ρεύμα 3 του κόμβου Ηρώων Πολυτεχνείου – Παύλου Πεντάρη.....	126

Ευρετήριο Πινάκων

Πίνακας 1: Ισοδυναμία κατηγοριών οχημάτων σε Μονάδες Επιβατικών Αυτοκινήτων (ΜΕΑ). (πηγή: ΥΠΕΧΩΔΕ)	48
Πίνακας 2: Τιμές συντελεστή m ανάλογα με τον ποιοτικό χαρακτηρισμό των αφίξεων οχημάτων στην κυκλοφορία. (πηγή: [24])	51
Πίνακας 3: Στοιχεία υπολογισμού συντελεστή DF. (πηγή: [24]).....	52
Πίνακας 4: Αντιστοιχία στάθμης εξυπηρέτησης με το διάστημα μέσης αναμονής ανά όχημα σε ισόπεδο κόμβο φωτεινής σηματοδότησης. (πηγή: [24]).....	53
Πίνακας 5: Αντιστοιχία στάθμης εξυπηρέτησης με το διάστημα μέσης αναμονής ανά όχημα σε κυκλικό κόμβο. (πηγή: [24]).....	55
Πίνακας 6: Συγκεντρωτική παράθεση των τεχνικών χαρακτηριστικών του Suncity Dataset, ανά αρχείο βίντεο.	58
Πίνακας 7: Αποτελέσματα ποσοτικής αξιολόγησης με τις μετρικές mAP και mAR @ [0.30:0.80].....	108
Πίνακας 8: Αποτελέσματα ποσοτικής αξιολόγησης με τις μετρικές mAP@0.50 και mAP@0.75.	108
Πίνακας 9: Αποτελέσματα ποσοτικής αξιολόγησης σε detection-by-frame, βάσει Precision και Recall για IoU > 0.5.	108
Πίνακας 10: Αποτελέσματα μετρικών ποσοτικής αξιολόγησης του MOT Benchmark (I). ...	114
Πίνακας 11: Αποτελέσματα μετρικών ποσοτικής αξιολόγησης του MOT Benchmark (II)...	114
Πίνακας 12: Αποτελέσματα ανώτερων μετρικών ποσοτικής αξιολόγησης του MOT Benchmark.....	115
Πίνακας 13: Αποτελέσματα προσδιορισμού ζήτησης σχεδιασμού υπό τη μορφή κυκλοφοριακού φόρτου [ΜΕΑ/ώρα], κατανεμημένη σε κελιά προέλευσης-προορισμού μεταξύ των κλάδων στον κόμβο Ηρώων Πολυτεχνείου – Παύλου Πεντάρη.	120
Πίνακας 14: Αποτελέσματα προσδιορισμού ζήτησης σχεδιασμού υπό τη μορφή κυκλοφοριακού φόρτου [ΜΕΑ/ώρα], κατανεμημένη σε κελιά προέλευσης-προορισμού μεταξύ των κλάδων στον κόμβο Πρωτόπαπα – Σοφοκλή Βενιζέλου.....	121
Πίνακας 15: Αποτελέσματα προσδιορισμού ζήτησης σχεδιασμού υπό τη μορφή κυκλοφοριακού φόρτου [ΜΕΑ/ώρα], κατανεμημένη σε κελιά προέλευσης-προορισμού μεταξύ των κλάδων στον κυκλικό κόμβο Πλατείας 25ης Μαρτίου.	121
Πίνακας 16: Αποτελέσματα μετρικών αξιολόγησης εκτίμησης κυκλοφοριακού φόρτου.....	122
Πίνακας 17: Αποτελέσματα προσδιορισμού δευτερευόντων κυκλοφοριακών μεγεθών (δε συμμετέχουν στην εκτίμηση επιπέδων εξυπηρέτησης).....	123
Πίνακας 18: Συγκεντρωτικά αποτελέσματα των εξαγόμενων χαρακτηρισμών επιπέδου εξυπηρέτησης.	124

Ευρετήριο Αποσπασμάτων Κώδικα

Απόσπασμα Κώδικα 1: Αρχείο παραμετροποίησης της δομής MMDetection, για το παραχθέν μοντέλο ανίχνευσης Faster R-CNN. Σημαντικότερες ρυθμίσεις αποτέλεσαν: α) η μείωση Anchor scale από την προεπιλογή 8 σε 3, ώστε να προσεγγίζεται το μέγεθος των αντικειμένων, β) η αύξηση RoI layer feature size από την προεπιλογή 7 σε 18, για να αποτυπώνεται η μέγιστη δυνατή πληροφορία χαρακτηριστικών που επιτρέπουν οι περιορισμοί VRAM, καθώς και γ) η ελάττωση του ορίου IoU θετικών δειγμάτων στο RoI τμήμα από την προεπιλογή 0.5 σε 0.3, για ανοχή σε λιγότερο ακριβές «localization» έναντι των μικροσκοπικών αντικειμένων. (πηγή πρωτότυπου: [32]).....	69
Απόσπασμα Κώδικα 2: Αρχείο παραμετροποίησης της δομής MMTracking, για το ανεξάρτητο παραχθέν μοντέλο επαναταυτοποίησης BaseReID, βασισμένο στο ResNet-50. Οι προεπιλεγμένες τιμές μεταβλητών ήταν αρκετές ώστε να επιλυθεί άριστα το πρόβλημα κατηγοριοποίησης ~34000 εικονιδίων σε 78 οχήματα, μετά από 10 εποχές (mAP = 100%). Ωστόσο η εκπαίδευση απεδείχθη ιδιαίτερα χρονοβόρα, συγκριτικά με την αντίστοιχη για το detection. (πηγή πρωτότυπου: [32]).....	77
Απόσπασμα Κώδικα 3: Αρχείο παραμετροποίησης της δομής MMTracking, το οποίο υλοποιεί την τελική υλοποίηση παρακολούθησης DeepSORT, βασισμένο στα υπο-μοντέλα Faster R-CNN και BaseReID. Το μοντέλο διατίθεται για εξαγωγή συμπερασμάτων (inference) και ποσοτική αξιολόγηση στα πρότυπα του MOT Challenge. Στις υπερπαραμέτρους επαναταυτοποίησης έχει δοθεί επιπλέον βάρος ευαισθησίας στην τοποθεσία/θέση των στόχων, έναντι της εμφάνισης/μεγέθους τους (0.8 προς 0.2), καθώς πρόκειται για κινούμενα οχήματα. (πηγή πρωτότυπου: [32]).....	81
Απόσπασμα Κώδικα 4: Ανάγνωση ενός αρχείου «.pkl» με την εισαγωγή της βιβλιοθήκης Pickle.....	86
Απόσπασμα Κώδικα 5: Διαχωρισμός των δεδομένων ανά κατηγορία οχήματος.....	86
Απόσπασμα Κώδικα 6: Η σύνθετη δομή επανάληψης για την εξαγωγή του συνόλου στοιχείων τροχίας για την κατηγορία «Car».....	87
Απόσπασμα Κώδικα 7: Επιστροφές τιμών μεταβλητής της συνάρτησης «pkl2tracks».....	87
Απόσπασμα Κώδικα 8: Ορισμός των διατομών καταμέτρησης σε μορφή LineString της βιβλιοθήκης Shapely.....	90
Απόσπασμα Κώδικα 9: Ο βασικός μηχανισμός υλοποίησης τροχιών και καταμέτρησης διελεύσεων.....	90
Απόσπασμα Κώδικα 10: Η ανάθεση των καταμετρήσεων διελεύσεων σε πίνακα προέλευσης – προορισμού.....	91
Απόσπασμα Κώδικα 11: Η ανάθεση των καταμετρήσεων U-turn στην κύρια διαγώνιο πίνακα προέλευσης – προορισμού.....	92
Απόσπασμα Κώδικα 12: Διαχείριση των συνθηκών «ιδιόρρυθμων» τροχιών.....	93
Απόσπασμα Κώδικα 13: Παράμετροι τεχνικών συνθηκών λήψης, και υπολογισμός μεγέθους εδαφοψηφίδας.....	94
Απόσπασμα Κώδικα 14: Προσθήκη των απαραίτητων ρυθμίσεων στο βασικό μηχανισμό υλοποίησης τροχιών που επεξηγήθηκε προηγουμένως, ώστε να καθίσταται εφικτός ο υπολογισμός ταχυτήτων διαδρομής, και διάνυσης μήκους συναρτήσει καρέ.....	94
Απόσπασμα Κώδικα 15: Εύρεση μέσης πυκνότητας, ενδεικτικά από το σύνολο τροχιών κατηγορίας «Car».....	95
Απόσπασμα Κώδικα 16: Διαφοροποίηση στην εκτίμηση της μέσης πυκνότητας εντός του δακτυλίου κυκλικού κόμβου (ενδεικτικά για την κατηγορία «Car»).....	96
Απόσπασμα Κώδικα 17: Κλήση των απαραίτητων συναρτήσεων και αναγωγή σε MEA και ωριαία βάση αναφοράς, ώστε να σχηματιστεί ο πίνακας προέλευσης-προορισμού που δηλώνει τον ωριαίο κυκλοφοριακό φόρτο κάθε διαδρομής.....	97

Απόσπασμα Κώδικα 18: Κατανομή της ζήτησης στα διακριθέντα ρεύματα (ή λωρίδες) κυκλοφορίας. Στους κόμβους φωτεινής σηματοδότησης γίνεται ορισμός των στοιχείων του υφιστάμενου προγράμματος.....	98
Απόσπασμα Κώδικα 19: Πραγματοποίηση των απαραίτητων υπολογισμών κατά HCM 2010 για τον προσδιορισμό των μέσων καθυστερήσεων (προφανώς, για τον κυκλικό κόμβο ακολουθούνται διαφορετικοί υπολογισμοί, που αναφέρονται στην αντίστοιχη σχετική μεθοδολογία λειτουργικής ανάλυσης).....	98
Απόσπασμα Κώδικα 20: Συνάρτηση χαρακτηρισμού επιπέδου εξυπηρέτησης.....	99

Κεφάλαιο 1^ο: Εισαγωγή

1.1: Αφορμές προσέγγισης θέματος

Η προηγούμενη δεκαετία αποτελεί ορόσημο για την επιστήμη των υπολογιστών, καθώς κατά τη διάρκειά της κατατέθηκαν πρωτότυπες ερευνητικές συνεισφορές, οι οποίες οδήγησαν, εν τέλει, στη ραγδαία ανάπτυξη της τεχνητής νοημοσύνης που παρατηρείται σήμερα. Θα μπορούσε να ειπωθεί πλέον με ασφάλεια, ότι η πληροφορική αλληλεπιδρά με κάθε επιστημονικό κλάδο, και επιτυγχάνει να δίνει λύσεις σε πολυποίκιλα ζητήματα της καθημερινότητας.

Ειδικά για την ελληνική επικαιρότητα, τοποθετείται στο προσκήνιο η περίπτωση της Αθήνας, μιας μεγαλούπολης με συνεχή κυκλοφοριακά προβλήματα σχεδόν σε ολόκληρη την έκτασή της. Η ανάγκη για μελέτη και βελτίωση των υφιστάμενων συνθηκών δημιουργεί έντονο ενδιαφέρον για ενασχόληση με το αντικείμενο της κυκλοφοριακής τεχνικής. Όπως είναι αναμενόμενο, από αυτό το πεδίο εφαρμογής δε θα μπορούσε να απουσιάζει η προοπτική συνδρομής της πληροφορικής.

Συνεπώς, θέτοντας ως όραμα τον πλήρη έλεγχο της οδικής κυκλοφορίας μέσω ψηφιακών-οπτικών μεθόδων, επιλέχθηκε στην παρούσα διπλωματική εργασία η μελέτη αξιολόγησης της συνεισφοράς ενός μοντέλου βαθιάς μάθησης στο εξής θεμελιώδες κυκλοφοριακό ζήτημα: την **παρακολούθηση (tracking) πολλαπλών οχημάτων**, από την επεξεργασία δεδομένων εικόνας και βίντεο. Μάλιστα, εκτός από τον έλεγχο της ικανότητας του συνιστάμενου μοντέλου, από τις εξαγόμενες τροχιές αποσκοπείται η εκτίμηση των απαραίτητων κυκλοφοριακών μεγεθών, ώστε με την κατάλληλη εφαρμογή λειτουργικής ανάλυσης να πραγματοποιηθεί κατάταξη αξιολόγησης των οδικών τμημάτων κόμβων σε **επίπεδα εξυπηρέτησης**.

1.2: Συνεισφορά του πεδίου επιστήμης υπολογιστών

[1] Οι αλγόριθμοι **Μηχανικής Μάθησης (Machine Learning, ML)** συμβάλλουν καταλυτικά στην εξαγωγή συμπερασμάτων από την διαχείριση δεδομένων εικόνας. Όπως διατυπώθηκε αρχικά από τον Arthur Samuel (1959), πρόκειται για *«πεδίο μελέτης που δίνει στους υπολογιστές την ικανότητα να μαθαίνουν, χωρίς να έχουν ρητά προγραμματιστεί»*. Ανάλογα με τον τρόπο και το βαθμό κατά τον οποίο ο χρήστης προετοιμάζει ένα υπολογιστικό σύστημα για εκπαίδευση, διακρίνονται οι παρακάτω τύποι μάθησης:

- **Επιβλεπόμενη (Supervised):** Αξιοποιείται η μάθηση των ιδιοτήτων από τα πρότυπα ενός υποσυνόλου εκπαίδευσης (training subset) στην πραγματοποίηση προβλέψεων επί ενός υποσυνόλου ελέγχου (test subset). Για τη σταδιακή

προσαρμογή - σύγκλιση (adaptation - convergence) και τη στατιστική αξιολόγηση (evaluation) της απόδοσης του μοντέλου, για τα δύο υποσύνολα είναι γνωστή η πληροφορία των επιθυμητών αποτελεσμάτων, συνιστώντας το σύνολο δεδομένων αλήθειας (ground truth dataset).

- **Μη επιβλεπόμενη (Unsupervised):** Εφαρμόζεται αποκλειστικά σε μη χαρακτηρισμένα δεδομένα, με στόχο την αναζήτηση βέλτιστων ομαδοποιήσεων ή συσσωρεύσεων σε διανυσματικούς χώρους προτύπων. Εφόσον δεν διατίθεται στα δεδομένα εισόδου η πληροφορία αλήθειας, δεν είναι δυνατή η ποσοτική αξιολόγηση ενός τέτοιου μοντέλου.
- **Ημι-επιβλεπόμενη / Ενισχυτική (Reinforcement):** Διαδικασία μάθησης κατά την οποία το υπολογιστικό σύστημα προσαρμόζεται μέσω των αλληλεπιδράσεων του με το περιβάλλον, μεγιστοποιώντας -χωρίς εξωτερική καθοδήγηση- μια συνάρτηση «κέρδους» από κάθε ενέργειά του.

Η **Βαθιά Μάθηση (Deep Learning, DL)** διαφοροποιείται ελαφρώς από τη Μηχανική Μάθηση, αποτελώντας ταυτόχρονα υποχώρο της, καθώς αξιοποιεί τις δυνατότητες των δομών **Τεχνητών Νευρωνικών Δικτύων (Artificial Neural Networks, ANNs)** στην «αφομοίωση» και εξαγωγή αναπαραστάσεων υψηλού επιπέδου (deep features). Συγκριτικά με τις υπόλοιπες οικογένειες αλγορίθμων Μηχανικής Μάθησης, επιτυγχάνονται καλύτερες επιδόσεις στη διάκριση μοτίβων μεταξύ των δεδομένων μελέτης, ακόμη και μη γραμμικά διαχωρίσιμων στους διανυσματικούς χώρους προτύπων. Επιπλέον, υποστηρίζεται κι ενθαρρύνεται η διαλειτουργικότητα μεταξύ διαφορετικών μοντέλων εφαρμογών, μέσω της προοπτικής μεταφοράς μάθησης (transfer learning) και επανασυντονισμού (fine-tuning). Στον αντίποδα, τέτοιες δομές, εκτός της σαφώς δυσκολότερης ερμηνείας τους, χρειάζονται υπέρογκα δεδομένα εκπαίδευσης, και συνεπώς φέρουν αυξημένες απαιτήσεις, τόσο σε υπολογιστική ισχύ, όσο και σε χρονική διάρκεια εκτέλεσης. Ειδικά για το πεδίο της **Υπολογιστικής Όρασης (Computer Vision, CV)**, όπου ζητείται η δημιουργία «τεχνητής αντίληψης» (machine perception) μέσω της αξιοποίησης ψηφιακών δεδομένων εικόνων, η Βαθιά Μάθηση βρίσκει ιδανική εφαρμογή με τη συνδρομή μιας ειδικής στην ανάλυση οπτικού υλικού κλάσης νευρωνικών δικτύων, των **Συνελκτικών Νευρωνικών Δικτύων (Convolutional Neural Networks, CNNs)**.

Η παρακολούθηση πολλαπλών αντικειμένων σε δομές βίντεο, η οποία εξετάζεται στην παρούσα διπλωματική εργασία, αποτελεί ένα παράγωγο εγχείρημα των προαναφερθέντων τεχνικών, και συγκεκριμένα πρόκειται για εφαρμογή Επιβλεπόμενης Βαθιάς Μάθησης. Οι χρησιμοποιούμενοι αλγόριθμοι δέχονται ως δεδομένα εισόδου εικόνες – καρέ από ακολουθίες βίντεο, για τις οποίες υπάρχουν τα αντίστοιχα αρχεία σχολιασμών για τα στοιχεία των αντικειμένων ενδιαφέροντος, καθιστώντας δυνατή την ποσοτική αξιολόγηση.

Η ερευνητική κοινότητα από τα μέσα της περασμένης δεκαετίας έχει συνεισφέρει έμπρακτα στην πρόκληση της παρακολούθησης πολλαπλών αντικειμένων με μια πληθώρα μοντέλων και αρχιτεκτονικών. Για την αξιολόγηση τους υπό το φάσμα ενός ενιαίου συγκριτικού σημείου αναφοράς (benchmark), δημιουργήθηκε το **Multiple Object Tracking (MOT) Benchmark** [2], το οποίο παρέχει τα εξής:

- Μια ευρεία ποικιλία δομών δεδομένων οπτικού υλικού, μαζί με τα αρχεία σχολιασμού για την πληροφορία αλήθειας στην παρακολούθηση πολλαπλών αντικειμένων. Οι στόχοι προς παρακολούθηση είναι μόνο πεζοί. Έχουν σχηματιστεί τα εξής σετ δεδομένων:
 - MOT15 Challenge [3]: Σύνολο αποτελούμενο από 22 ακολουθίες βίντεο, συγκεντρωμένων μέσω «third-party» πηγών, από στατικές και κινούμενες κάμερες.
 - MOT16 Challenge [4]: Σύνολο αποτελούμενο από 14 ακολουθίες βίντεο από στατικές και κινούμενες κάμερες, με πιο αυστηρό πρωτόκολλο σχολιασμών (αντικείμενα ενδιαφέροντος θεωρούνται οι απευθείας ορατοί κινούμενοι πεζοί, οι ακίνητοι ή οι ορατοί πεζοί μέσω ανακλάσεων θεωρούνται ασαφείς στόχοι, ενώ οι υπόλοιπες κατηγορίες αντικειμένων εξαιρούνται από την αξιολόγηση).
 - MOT17 Challenge [4]: Οι ίδιες 14 ακολουθίες βίντεο με το MOT16, παρόμοιου πρωτοκόλλου σχολιασμών, με ανανεωμένα και πιο ακριβή δεδομένα αλήθειας, και επιπλέον διαθέσιμα αποτελέσματα ανίχνευσης από τους ανιχνευτές Faster R-CNN, SDP και DPM για χρήση κατά την αξιολόγηση.
 - MOT20 Challenge [5]: Σύνολο 8 ακολουθιών βίντεο στο οποίο απεικονίζονται στατικές σκηνές υψηλού συνωστισμού, ακολουθώντας το πρωτόκολλο σχολιασμού των προηγούμενων εκδόσεων.
 - MOTS Challenge [6]: Σύνολο 8 ακολουθιών βίντεο από στατικές και κινούμενες κάμερες, στοχεύοντας επιπρόσθετα την αξιολόγηση παρακολούθησης πεζών σε επίπεδο εικονοστοιχείων, ακολουθώντας το πρωτόκολλο σχολιασμού των προηγούμενων εκδόσεων.
- Ένα εκτενές σύστημα μετρικών αξιολόγησης του μοντέλου στις προβλέψεις θέσης, τροχιάς και ταυτοποίησης των αντικειμένων.

Η παρούσα διπλωματική εργασία επικεντρώνεται αποκλειστικά σε διερχόμενα οχήματα ως αντικείμενα ενδιαφέροντος, και μάλιστα ταξινομημένα ανά κατηγορίες κατά την παρακολούθησή τους, και όχι τους πεζούς, όπως συμβαίνει στα MOT Challenge σετ δεδομένων. Ωστόσο, όπως θα αιτιολογηθεί σαφώς σε επόμενη ενότητα, η ερμηνευτική

σημασία των μετρικών αξιολόγησης του MOT Benchmark δεν αποτρέπει την συμβατότητα τους προς διαφορετικής υπόστασης κινούμενα αντικείμενα – στόχους, για αυτό και χρησιμοποιούνται απρόσκοπτα.

1.3: Παράθεση γενικών αναμενόμενων οφελών και προκλήσεων

Η ενσωμάτωση αλγορίθμων παρακολούθησης οχημάτων στα ψηφιακά συστήματα διαχείρισης κυκλοφορίας υπόσχεται σημαντικά οφέλη στον τομέα των μεταφορών. Με την εφαρμογή τους σε πραγματικό χρόνο, μειώνεται στο ελάχιστο ο χρόνος ανταπόκρισης των αρμόδιων κέντρων διαχείρισης απέναντι σε απρόβλεπτα συμβάντα, όπως ατυχήματα ή γενικότερα ακινητοποιήσεις οχημάτων, με την ανίχνευση μηδενικών ταχυτήτων σε καίρια οδικά τμήματα που συνήθως χαρακτηρίζονται από συνεχή ροή ή και υψηλές ταχύτητες. Ακόμη, με την υποστήριξη από πρόσθετες πλατφόρμες λογισμικών πλοήγησης, καθίσταται δυνατή η διαρκής ενημέρωση των κυκλοφοριακών συνθηκών σε οδικές αρτηρίες, ώστε οι χρήστες να μπορούν δυναμικά να προσαρμόζουν τη διαδρομή τους αποφεύγοντας τμήματα κυκλοφοριακής συμφόρησης. Κατά την ασύγχρονη ενσωμάτωσή τους, μάλιστα, το ενδιαφέρον στρέφεται στην προοπτική αντικατάστασης των συμβατικών μεθόδων παρατηρήσεων συγκοινωνιακών μελετών (π.χ. τοποθέτηση ανιχνευτών επαγωγικών βρόχων, σωλήνων πίεσης αέρα, ανιχνευτών υπερύθρων/υπερήχων/laser, ή και μετρήσεις προσωπικού επί του πεδίου).

Εκτός των αμιγώς κυκλοφοριακών πλεονεκτημάτων, οι αλγόριθμοι παρακολούθησης οχημάτων, εφόσον συνδυαστούν με επαρκή κάλυψη από οπτικά μέσα, θα μπορούσαν να μειώσουν σημαντικά τις βάρδιες μετακινήσεων των οχημάτων περιπολίας στους αυτοκινητοδρόμους, εξοικονομώντας καθημερινά ένα σεβαστό όγκο καυσίμου και ταυτόχρονα συμβάλλοντας στον περιορισμό των εκπομπών διοξειδίου του άνθρακα. Τονίζεται πως κατά την υφιστάμενη κατάσταση, τα συγκεκριμένα οχήματα σαρώνουν ολόκληρη την έκταση των ελληνικών αυτοκινητοδρόμων καθ' όλη τη διάρκεια της ημέρας.

Παρ' όλα τα παραπάνω οφέλη, η ψηφιακή παρακολούθηση πολλαπλών οχημάτων εγείρει μια σημαντική αμφιβολία: τη χωρική αποδοτικότητά της. Οι όψεις των οχημάτων ξεχωρίζουν βέλτιστα από το αδιάφορο υπόβαθρο, όταν το πρόβλημα υποβαθμίζεται στις 2 διαστάσεις, δηλαδή διατίθενται λήψεις βίντεο σε προσέγγιση ορθής προβολής («ορθές» λήψεις). Η καταλληλότητα τέτοιων λήψεων αποτελεί μια διαδικασία εξισορρόπησης μεταξύ πολυάριθμων αλληλένδετων παραγόντων:

- Με ορθές λήψεις ελαχιστοποιούνται οι δυνατότητες χωρικής κάλυψης (σε σύγκριση με διαφορετικούς προσανατολισμούς της κάμερας), αλλά ωστόσο εξασφαλίζεται η διευκόλυνση της ανίχνευσης αντικειμένων.

- Αν μια κάμερα τοποθετηθεί σε αυξημένο ύψος, τότε για την διασφάλιση της ποιότητας ορθών λήψεων θα πρέπει η παρεχόμενη ανάλυση να είναι υψηλή.
- Σε περίπτωση που η διαθέσιμη κάμερα είναι χαμηλής ανάλυσης, τότε για την «ορθή» κάλυψη μεγάλης έκτασης τίθενται προς αντιμετώπιση οι προκλήσεις ανίχνευσης μικροσκοπικών αντικειμένων.

Συνήθως καθίσταται ιδιαίτερα δύσκολη η τοποθέτηση σταθερής κάμερας σε μεγάλο ύψος για τη διεξαγωγή ορθών λήψεων υψηλής χωρικής κάλυψης, οπότε καταληκτικά **εξετάζεται η συνδρομή εναέριων μέσων**. Για την ανάπτυξη της δομής παρακολούθησης πολλαπλών αντικειμένων στην παρούσα διπλωματική εργασία, το σετ δεδομένων μελέτης αποτελείται από ακολουθίες βίντεο προερχόμενες από ενσωματωμένη κάμερες σε **μη επανδρωμένα ιπτάμενα οχήματα (UAVs / drones)**.

1.4: Διάρθρωση διπλωματικής εργασίας

Το κυρίως μέρος της διπλωματικής εργασίας δομείται από την παρακάτω ακολουθία κεφαλαίων:

- Στο 2^ο κεφάλαιο τοποθετείται το διαθέσιμο επιστημονικό υπόβαθρο τόσο για το σκέλος της υπολογιστικής νοημοσύνης, όσο και για τα στοιχεία θεωρίας κυκλοφοριακής τεχνικής.
- Το χαρακτηριστικά του σετ δεδομένων πάνω στο οποίο εφαρμόστηκε το προγραμματιστικό πλαίσιο, καθώς και η περιγραφή των προκλήσεων και ιδιομορφιών του, περιγράφονται αναλυτικά στο 3^ο κεφάλαιο.
- Το 4^ο κεφάλαιο περιλαμβάνει τη διαδικασία ανάπτυξης μεθοδολογίας για την παρακολούθηση πολλαπλών αντικειμένων στο σετ κυκλοφοριακών δεδομένων, αλλά και τις κατάλληλες προγραμματιστικές τροποποιήσεις ώστε να επιτευχθεί η σωστή αξιοποίηση των δεδομένων εξόδου του μοντέλου.
- Τέλος, στο 5^ο κεφάλαιο παρατίθεται μια γενική αποτίμηση της επίδοσης του μοντέλου, πραγματοποιείται ποσοτική αξιολόγησή του κι εμφανίζονται οπτικά αποτελέσματα, ενώ αναφέρονται και περιπτώσεις αστοχίας, προοπτικές βελτίωσης και γενικά συμπεράσματα.

Κεφάλαιο 2^ο: Επιστημονικό υπόβαθρο

2.1: Θεωρητικό σκέλος πληροφορικής

Στις ακόλουθες υποενότητες παρουσιάζεται το σύνολο του εγγύτερα προσκείμενου επιστημονικού υποβάθρου στη δημιουργία μοντέλων παρακολούθησης πολλαπλών αντικειμένων, θέτοντας ως αφετηρία την πρωταρχική έννοια της Βαθιάς Μάθησης, τον απλό νευρώνα.

2.1.1: Δομή τεχνητών νευρωνικών δικτύων

Σε παραλληλισμό με ένα βιολογικό σύστημα, ο απλός **τεχνητός νευρώνας (perceptron)** αποτελεί το «κύτταρο» μιας δομής τεχνητών νευρωνικών δικτύων. Καθένας από αυτούς δέχεται σταθμισμένα κατά βάρη δεδομένα εισόδου, και αν μετά το άθροισμά (S) τους υπερβαίνεται μια τιμή κατωφλίου (T), ο νευρώνας παράγει σήμα εξάγοντας τιμή $y = 1$, ειδάλτως όταν δεν παράγεται σήμα εξάγεται τιμή $y = 0$ [7]:

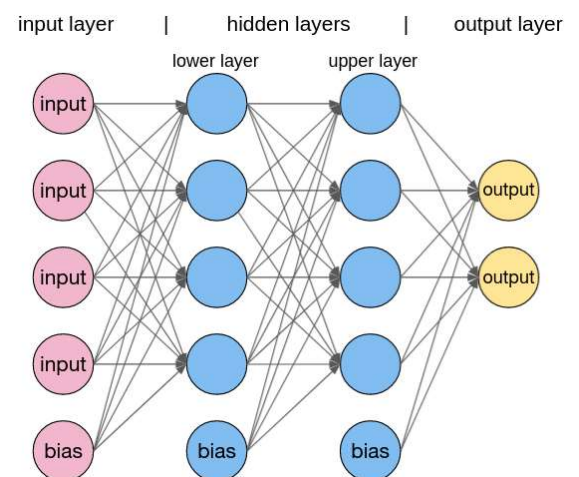
$$S = w_1 \cdot x_1 + \dots + w_n \cdot x_n = \sum_{i=1}^n w_i \cdot x_i ,$$

όπου x_i = οι τιμές δεδομένων εισόδου, και w_i = οι τιμές των βαρών τους.

Αν $S > T$, τότε $y = 1$, αλλιώς $y = 0$.

Επομένως ο απλός νευρώνας πραγματοποιεί μια δυαδική ταξινόμηση, έχοντας επιτυχία όταν τα δεδομένα εισόδου είναι γραμμικά διαχωρίσιμα.

Με την τοποθέτηση πολλαπλών νευρώνων-κόμβων σε περισσότερα από τρία επίπεδα πλήρως συνδεδεμένων (fully connected: κάθε κόμβος ενός επιπέδου συνδέεται σταθμισμένα κατά βάρος w_{ij} , με κάθε κόμβο του επόμενου επιπέδου) μεταξύ τους, σχηματίζεται δομή **πολυεπίπεδων νευρώνων (multi-layer perceptron)** [8]. Συνηθίζεται σε κάθε επίπεδο να προστίθεται επιπλέον ένας «always-on» κόμβος μεροληψίας - «bias», ο οποίος δεν έχει σύνδεση με κόμβους προηγούμενων στρώσεων, και φυσικά εξάγει πάντοτε τιμή $y = 1$, ώστε να δίνεται με βεβαιότητα στους επόμενους κόμβους μη μηδενική πληροφορία. Τα πρώτα και τελευταία επίπεδα καλούνται αντίστοιχα «επίπεδα εισόδου» κι «επίπεδα εξόδου», ενώ τα υπόλοιπα ενδιάμεσα ονομάζονται «κρυφά επίπεδα» (hidden layers).



Σχήμα 1: Η τυπική μορφή μιας δομής πολυεπίπεδων νευρώνων (multi-layer perceptron) με 4 συνολικές στρώσεις, 1 εισόδου, 2 κρυφές και 1 εξόδου. (πηγή: <https://medium.com/codex/introduction-to-how-an-multilayer-perceptron-works-but-without-complicated-math-a423979897ac>, προσηλεάστηκε Αυγ. 2023)

Οι τιμές εξόδου των κόμβων, καθορίζονται κατά το προς-τα-εμπρός πέρασμα των δεδομένων (**forward pass**) από το δίκτυο, με την επίδραση κάποιας **συνάρτησης ενεργοποίησης (activation function)**. Όταν οι συναρτήσεις αυτές είναι μη γραμμικές, καθίστανται υπεύθυνες για την εισαγωγή «μη γραμμικότητας» σε ένα νευρωνικό δίκτυο, δίνοντας του τη δυνατότητα να προσαρμόσει τα σήματα των κόμβων σε πιο περίπλοκα πρότυπα στα στοιχεία εισόδου. Η συνθήκη κατωφλίου στην περίπτωση του απλού νευρώνα αποτελεί γραμμική συνάρτηση ενεργοποίησης. Σε γενικευμένη μορφή εξίσωσης, η έξοδος ενός νευρώνα z με συνάρτηση ενεργοποίησης f , έχει ως εξής [1]:

$$z(x_1, \dots, x_n) = f(\sum_{i=1}^n w_i \cdot x_i + b),$$

όπου x_i = οι τιμές δεδομένων εισόδου, w_i = οι τιμές των βαρών τους, και b = ο δείκτης bias.

Μερικές από τις πιο κοινές μη γραμμικές συναρτήσεις ενεργοποίησης είναι οι παρακάτω [1]:

- Η συνάρτηση **ReLU (Rectify Linear Unit)**: $f(x) = \begin{cases} 0, & \text{για } x < 0 \\ x, & \text{για } x \geq 0 \end{cases}$
- Η **Σιγμοειδής (Sigmoid)** συνάρτηση: $f(x) = \frac{1}{1+e^{-x}}$
- Η συνάρτηση **Υπερβολικής Εφαπτομένης**: $f(x) = \tanh(x) = \frac{2}{1+e^{-2 \cdot x}} - 1$
- Η συνάρτηση **Softmax**, που συνήθως εφαρμόζεται στο επίπεδο εξόδου ενός νευρωνικού δικτύου για δημιουργία κατανομής πιθανοτήτων μεταξύ των εξαγόμενων τιμών:

$$\sigma(z)_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}},$$

όπου $K \geq 1$ το πλήθος τιμών εξόδου, z = το διάνυσμα του νευρώνα εξόδου

Τα τεχνητά νευρωνικά δίκτυα εκπαιδεύονται στη διάκριση προτύπων από μια κατανομή δεδομένων ακολουθώντας έναν γενικό σκοπό: τον προσδιορισμό των βέλτιστων παραμέτρων – βαρών του δικτύου, οι οποίες οδηγούν στην ελαχιστοποίηση μιας **συνάρτησης κόστους (Loss Function)**. Για το σύνολο των δεδομένων ελέγχου ορίζεται η **εμπειρική συνάρτηση κόστους (Empirical Loss)**: $J(w) = \frac{1}{n} \sum_{i=0}^n L(f(x_i, w), y_i)$, όπου L = η συνάρτηση κόστους. Αναλόγως του είδους εφαρμογής του νευρωνικού δικτύου, επιλέγεται και η κατάλληλη συνάρτηση κόστους L . Ενδεικτικά [1]:

- Η συνάρτηση **L1 Loss ή Least Absolute Deviations (LAD)**, για προβλήματα παλινδρόμησης με έμφαση στην επίδραση των «outlier» τιμών:

$$L1 = \sum_{i=0}^n |Y_i - \hat{Y}_i|$$

- Η συνάρτηση **Μέσου Τετραγωνικού Σφάλματος (Mean Squared Error, MSE)**, για προβλήματα παλινδρόμησης συνεχών μεταβλητών:

$$MSE = \frac{1}{n} \cdot \sum_{i=0}^n (Y_i - \hat{Y}_i)^2$$

- Η συνάρτηση **Binary Cross Entropy Loss (BCE Loss)**, για προβλήματα δυαδικής ταξινόμησης:

$$BCE = -\frac{1}{N} \cdot \sum_{i=0}^N y_i \cdot \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)$$

- Η συνάρτηση **Categorical Cross Entropy Empirical Loss (CCE Loss)**, για προβλήματα ταξινόμησης σε πολλαπλές κατηγορίες:

$$CCE = -\frac{1}{n} \cdot \sum_{i=0}^n \sum_{j=0}^m y_j \cdot \log(\hat{y}_j) + (1 - y_j) \cdot \log(1 - \hat{y}_j)$$

- Η συνάρτηση **Focal Loss (FL)**, για προβλήματα ανισοβαρούς ταξινόμησης με ιδιαίτερα δύσκολα - αταξινομήτα παραδείγματα:

$$FL = -\sum_{i=0}^n a_i \cdot (i - p_i)^{\gamma} \cdot \log_b(p_i)$$

Η προσαρμογή των βαρών ως το σημείο το οποίο έχει επιτευχθεί ελαχιστοποίηση του κόστους επιτυγχάνεται κατά το **προς-τα-πίσω πέρασμα (backward pass)**, με εφαρμογή των τεχνικών **βελτιστοποίησης (optimization)**. Η μέθοδος Καταβιβασμού Κλίσης (Gradient Descent) αποτελεί το βασικό αλγόριθμο βελτιστοποίησης, κατά τον οποίο οι παράμετροι (w, b) αρχικοποιούνται τυχαία, και πραγματοποιείται ανανέωση τους σε κάθε επανάληψη βάσει του επιλεγμένου ρυθμού εκμάθησης (learning rate) και των υπολογισμένων μερικών παραγώγων (έως 1^{ης} τάξης) της εμπειρικής συνάρτησης κόστους ως προς (w, b), έως ότου θεωρηθεί σύγκλιση, δηλαδή [1]:

Έστω οι μερικές παράγωγοι $\frac{\partial}{\partial w} J(w) = \nabla_w J$, $\frac{\partial}{\partial b} J(w) = \nabla_b J$, όπου $J = \eta$ εμπειρική συνάρτηση κόστους, και $\alpha = \rho$ ρυθμός εκμάθησης. Τότε:

$$\begin{cases} w_{new} = w - \alpha \cdot \nabla_w J \\ b_{new} = b - \alpha \cdot \nabla_b J \end{cases}$$

Οι μερικές παράγωγοι της εμπειρικής συνάρτησης κόστους υπολογίζονται με τη μέθοδο **Back-Propagation**, που βασίζεται στον κανόνα αλυσίδας απειροστικού λογισμού ($\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}$). Με αυτό τον τρόπο αντιμετωπίζεται η ιδιομορφία υπολογισμού των μερικών παραγώγων των παραμέτρων στους κόμβους «μεταγενέστερων» επιπέδων, εφόσον αποτελούν σύνθετες συναρτήσεις με μεταβλητές τις παραμέτρους κόμβων «προγενέστερων» επιπέδων.

Επιπρόσθετα, επειδή η διαχείριση της συνολικής παραγωγού της εμπειρικής συνάρτησης κόστους φέρει μεγάλες υπολογιστικές απαιτήσεις, θεωρούνται «σημεία» εντός των δεδομένων όπου στα οποία η εμπειρική ταυτίζεται με τη συνάρτηση κόστους. Οπότε, κατά τον καταβιβασμό κλίσης χρησιμοποιείται η μερική παράγωγος της εμπειρικής ανά batch (= πλήθος σημείων) συνάρτησης $J_B(w)$, η οποία αποδεικνύεται υπολογιστικά αποδοτικότερη [1]:

$$\frac{\partial}{\partial w} J_B(w) = \frac{1}{B} \sum_{i=1}^B \frac{\partial}{\partial w} J_i(w), \text{ όπου } B = \text{το μέγεθος batch.}$$

Σε αυτή την περίπτωση πραγματοποιείται **Στοχαστικός Καταβιβασμός Κλίσης (Stochastic Gradient Descent, SGD)**. Πρόσθετοι αλγόριθμοι βελτιστοποίησης, βασισμένοι στον SGD, είναι οι εξής [1]:

- **Momentum SGD:** Επέκταση του SGD, με ανανέωση των παραμέτρων βάσει του κινούμενου μέσου των παραγώγων προηγούμενων επαναλήψεων (momentum).
- **AdaGrad:** Καταβιβασμός κλίσης με δυναμικά φθίνοντα ρυθμό εκμάθησης (ανά παράμετρο) όσο γίνεται καλύτερη προσέγγιση της σύγκλισης.
- **RMSProp:** Καταβιβασμός κλίσης με μεταβαλλόμενο ρυθμό εκμάθησης (ανά παράμετρο) βάσει των τετραγώνων παραγώγων προηγούμενων επαναλήψεων.
- **Adam:** Αρκετά συχνά χρησιμοποιούμενος βελτιστοποιητής, κατά τον οποίο ο ρυθμός εκμάθησης (ανά παράμετρο) μεταβάλλεται συναρτήσει ενός φθίνοντος εκθετικού μέσου παραγώγων προηγούμενων επαναλήψεων.

Για να επιτευχθεί «έξυπνη» ελάφρυνση ενός νευρωνικού δικτύου από ατελέσφορα σύνθετα μέρη, αποτρέποντάς το έτσι από την προσαρμογή του στις αποκλειστικές ιδιαιτερότητες ενός σετ δεδομένων (overfitting), εφαρμόζονται οι τεχνικές **εξομάλυνσης (regularization)**. Οι σημαντικότερες εξ' αυτών είναι οι εξής [1]:

- **Dropout:** Αγνόηση κόμβων – νευρώνων σε κάθε επανάληψη, βάσει μιας επιλεγμένης πιθανότητας p , με σκοπό την αποφυγή overfitting και εξάρτησης του δικτύου από μεμονωμένους νευρώνες. Η τεχνική αυτή αυξάνει το χρόνο επίτευξης σύγκλισης, αλλά μειώνει το χρόνο εκπαίδευσης ανά εποχή.
- **Weight Normalization:** Κανονικοποίηση των βαρών του δικτύου, εισάγοντας στην εμπειρική συνάρτηση κόστους έναν όρο βασισμένο σε κάποια νόρμα των βαρών, ώστε αυτά να ανανεώνονται αντίστοιχα κατά τον καταβιβασμό κλίσης. Αποφυγή υψηλών τιμών στα βάρη, που θα δημιουργούσε αριθμητική αστάθεια.
- **Batch Normalization:** Κανονικοποίηση ανά χαρακτηριστικό σε ένα batch, βάσει των στατιστικών μεγεθών του (μέση τιμή, τυπική απόκλιση, κ.α.), ώστε να ακολουθεί την κανονική κατανομή. Συμβάλλει στην ταχύτερη σύγκλιση.

- **Data Augmentation:** Ενίσχυση του σετ εκπαίδευσης με επιπλέον δεδομένα που προέρχονται από μετασχηματισμούς (μετάθεση, στροφή, εισαγωγή gaussian noise, random crop, κ.α.) των προϋπάρχοντων δεδομένων.

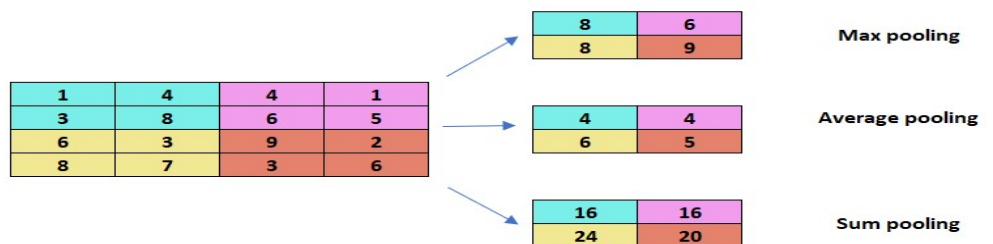
2.1.2: Συνελικτικά νευρωνικά δίκτυα

Στις εφαρμογές κατά τις οποίες τα δεδομένα εισόδου αποτελούν εικόνες, το κατάλληλο είδος δικτύου για τη διάκριση χρήσιμων χαρακτηριστικών είναι τα **Συνελικτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks, CNNs)**. Οι **συνελικτικές (convolutional)** στρώσεις των CNNs που βρίσκονται μεταξύ των hidden layers, δέχονται ως είσοδο εικόνες υπό την μορφή τανυστών (tensors) $2^{ns} / 3^{ns}$ τάξης (greyscale/RGB), και πραγματοποιώντας πράξεις συνέλιξης εξάγουν πίνακες χαρακτηριστικών (feature maps) που διαθέτουν την απαραίτητη πληροφορία υψηλού επιπέδου. Ακόμη, στις τιμές του προϊόντος μιας συνελικτικής στρώσης επιδράει μια συνάρτηση ενεργοποίησης (ουσιαστικά κάθε εικονοστοιχείο του feature map αποτελεί κόμβο), πριν το σήμα διέλθει σε επόμενη στρώση. Πριν την εφαρμογή της συνέλιξης μιας εικόνας f , με ένα δισδιάστατο πυρήνα (kernel) g , ο πυρήνας «ανακλάται», αλλά για απλοποίηση υπολογισμών η ενέργεια αυτή συνήθως παραλείπεται, οπότε η πράξη που πραγματοποιείται είναι η **αυτοσυσχέτιση (cross-correlation)**:

$$(f \otimes g)_{ij} = \sum_{m=0}^{k_1-1} \sum_{n=0}^{k_2-1} f(i+m, j+n) \cdot g(m, n)$$

, όπου i, j = οι διαστάσεις της εικόνας, και m, n = οι διαστάσεις του πυρήνα

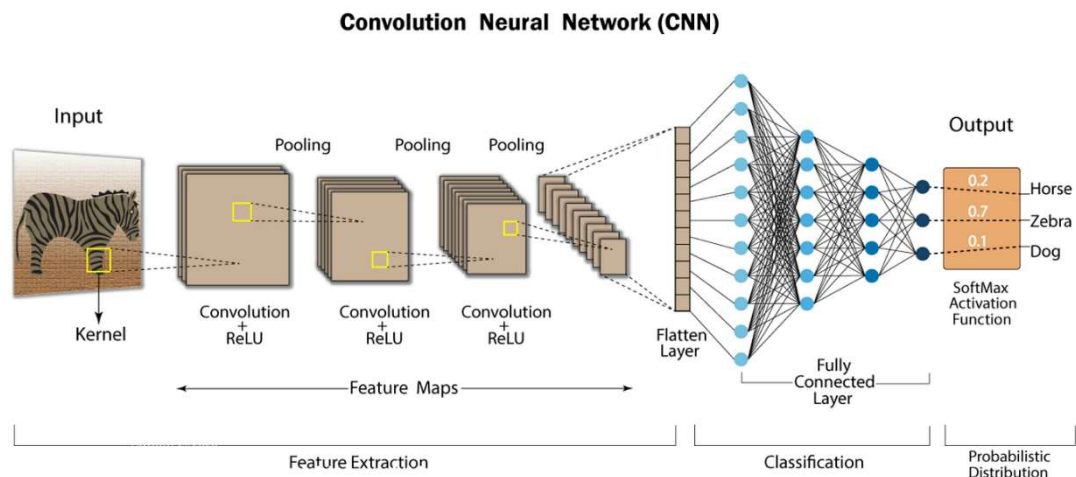
Εκτός των συνελικτικών στρώσεων, στα CNNs συμμετέχουν και επίπεδα **συγκέντρωσης (pooling)** τα οποία στοχεύουν στη μείωση των διαστάσεων των feature maps (κατά μήκος και πλάτος, όχι κατά κανάλια) για υπολογιστική διευκόλυνση κι ευρωστία, διατηρώντας παράλληλα τη χρήσιμη σημασιολογική πληροφορία. Σημαντικοί παράμετροι αποτελούν το βήμα (stride) και το «γέμισμα» (padding). Με τον τρόπο που φαίνεται στο παρακάτω σχήμα, το pooling μπορεί να συμβεί βάσει του μέσου όρου (average pooling), της μέγιστης τιμής (max pooling), και του αθροίσματος (sum pooling) μιας περιοχής του αρχικού πίνακα χαρακτηριστικών.



Σχήμα 2: Παράδειγμα εφαρμογών (2x2) max, average και sum pooling σε έναν απλό πίνακα ακέραιων τιμών. (πηγή: <https://www.shiksha.com/online-courses/articles/a-comprehensive-guide-to-convolutional-neural-networks/>, προσπελάστηκε Αυγ. 2023)

Στην περίπτωση που οι χωρικές διαστάσεις του τελικού τανυστή χρειάζεται να είναι προκαθορισμένες, συνήθως μετά την τελευταία κατά σειρά συνελκτική στρώση, εφαρμόζεται **προσαρμοσμένη συγκέντρωση (adaptive pooling)**, είτε κατά average, είτε κατά max.

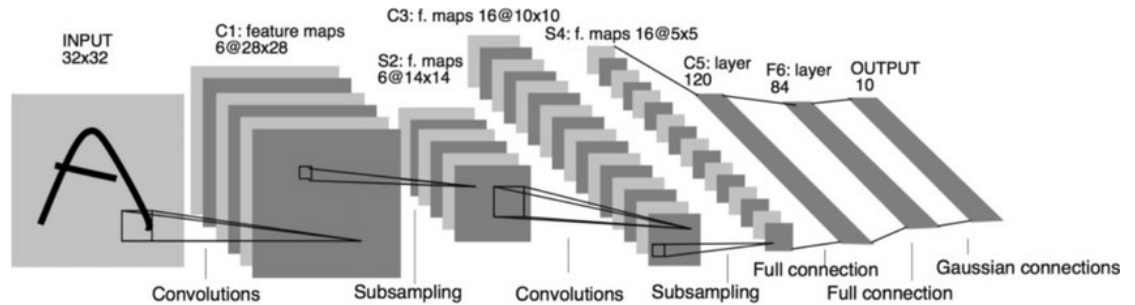
Η αλληλουχία των συνελκτικών και συγκεντρωτικών στρώσεων συναποτελεί το **τμήμα εξαγωγής χαρακτηριστικών (feature extraction head)**. Σε ένα τυπικό CNN, ακολουθεί το **τμήμα ταξινόμησης (classification head)**, το οποίο απαρτίζεται από ένα ή περισσότερα επίπεδα fully-connected νευρώνων. Σε αυτό εισάγεται η τελική πληροφορία του τμήματος εξαγωγής χαρακτηριστικών, εφόσον έχει υποστεί «ισοπέδωση» (flattening), δηλαδή μετατροπή σε τανυστή 1^{ης} τάξης. Περισσότερες fully-connected στρώσεις συνεπάγονται περαιτέρω δυνατότητες εκμάθησης μη γραμμικού διαχωρισμού των δεδομένων για την τελική ταξινόμησή τους. Το πλήθος κλάσεων ισούται με αυτό των κόμβων στη στρώση εξόδου, οι οποίοι τοποθετούνται προς κατάταξη σε κάποια κατηγορία στο τελευταίο **τμήμα πιθανοτικής κατανομής (probabilistic distribution head)**, με εφαρμογή συνάρτησης ενεργοποίησης Softmax.



Σχήμα 3: Η τυπική δομή ενός Συνελκτικού Νευρωνικού Δικτύου (Convolutional Neural Network, CNN). (πηγή: <https://nafizshahriar.medium.com/what-is-convolutional-neural-network-cnn-deep-learning-b3921bdd82d5>, προσπελάστηκε Αυγ. 2023)

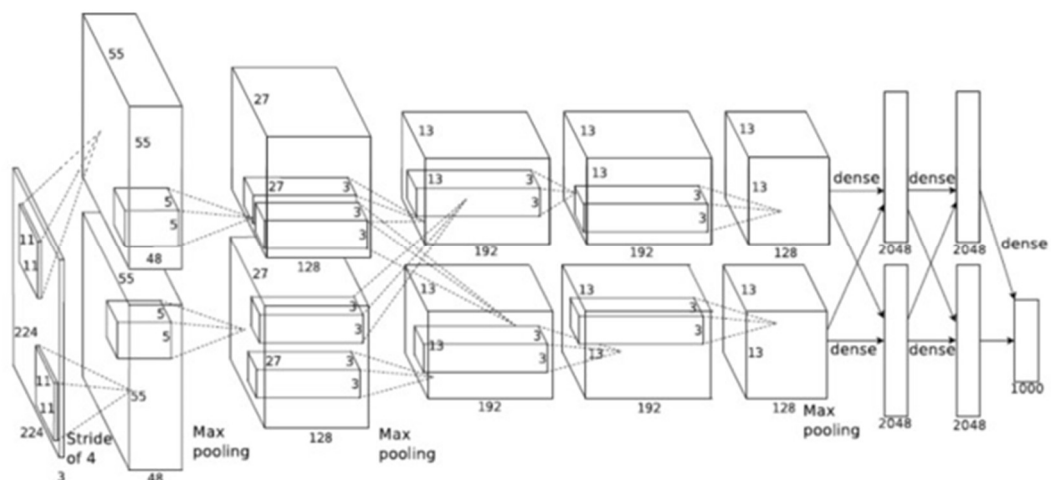
Ο τρόπος σχεδιασμού ενός νευρωνικού δικτύου, δηλαδή ο αριθμός, το είδος, οι παράμετροι καθώς η ακολουθία των στρώσεων του, αποτελεί την **αρχιτεκτονική (architecture)** του νευρωνικού δικτύου. Παρακάτω παρατίθενται μερικές από τις πιο γνωστές αρχιτεκτονικές νευρωνικών δικτύων:

- **LeNet-5 (1998)** [9]: Μια από τις πρώτες αρχιτεκτονικές CNN, που δημιουργήθηκε για την αυτόματη ταξινόμηση χειρόγραφων αριθμών και γραμμάτων. Αποτελείται από 7 στρώσεις, όπως φαίνεται παρακάτω στο παρακάτω διάγραμμα. Αντίθετα με πιο μοντέρνες αρχιτεκτονικές, χρησιμοποιεί συναρτήσεις ενεργοποίησης Sigmoid, και όχι ReLU. Διαθέτει ~60 χιλιάδες εκπαιδύσιμες παραμέτρους.



Σχήμα 4: Σχεδιάγραμμα της αρχιτεκτονικής LeNet-5. (πηγή: [9])

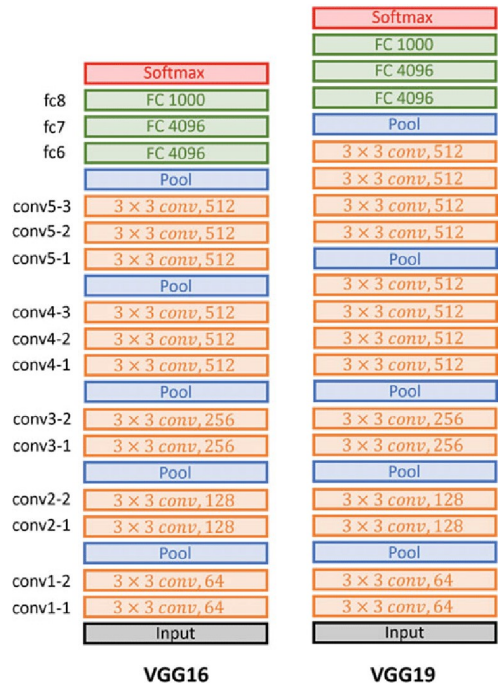
- **AlexNet (2012)** [10]: Πρόκειται για το δίκτυο CNN που σήμανε την αρχή για μια ραγδαία ανάπτυξη της Βαθιάς Μάθησης. Σχεδιάστηκε για το διαγωνισμό ImageNet Large Scale Visual Recognition Challenge (ILSVRC) του 2012, και εκπαιδεύτηκε ταυτόχρονα σε δύο κάρτες γραφικών NVIDIA GTX 580 3GB επί 5-6 ημέρες. Για αυτό το λόγο διακρίνονται 2 διαφορετικές ακολουθίες εκπαίδευσης (training pipelines), με «χωρισμένα» features και fully-connected νευρώνες. Χρησιμοποιούνται ReLU, dropouts, momentum SGD και data augmentation. Αποτελείται από 8 στρώσεις με ~60 εκατομμύρια εκπαιδύσιμες παραμέτρους.



Σχήμα 5: Σχεδιάγραμμα της αρχιτεκτονικής AlexNet. (πηγή: [10])

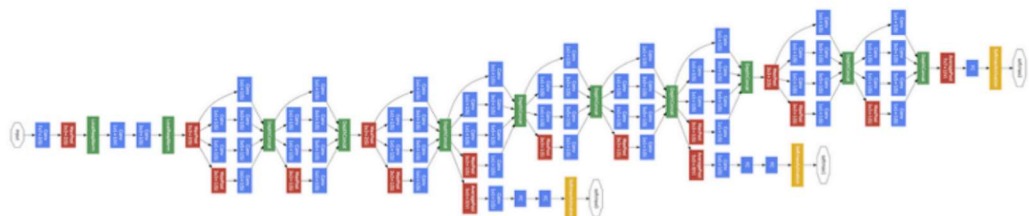
- **VGG (Visual Geometry Group) Net (2015)**

[11]: Όπως το AlexNet, σχεδιάστηκε για συμμετοχή στο διαγωνισμό ImageNet Large Scale Visual Recognition Challenge (ILSVRC), όμως για το έτος 2014. Ανάλογα το πλήθος των συνελκτικών στρώσεων, υπάρχουν 2 παραλλαγές του, το VGG-16 και το VGG-19. Χρησιμοποιούνται ReLU, dropouts, SGD, ενώ οι εκπαιδευσιμες παράμετροι είναι ~138 εκατομμύρια για την έκδοση VGG-16 και ~143 εκατομμύρια για την VGG-19. Το δίκτυο αποδίδει ικανοποιητικά σε σκοπούς **μεταφοράς μάθησης (transfer learning)**, όπου η προεκπαιδευμένη του κατάσταση στο ImageNet dataset επανασυντονίζεται σε διαφορετικά δεδομένα για ταξινόμηση εικόνων.

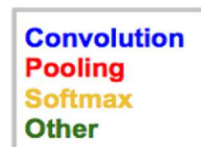


Σχήμα 6: Σχεδιαγράμματα των παραλλαγών της αρχιτεκτονικής VGG Net, VGG-16 και VGG-19. (πηγή: <https://datahacker.rs/deep-learning-vgg-16-vs-vgg-19/>, προσπελάστηκε Αυγ. 2023)

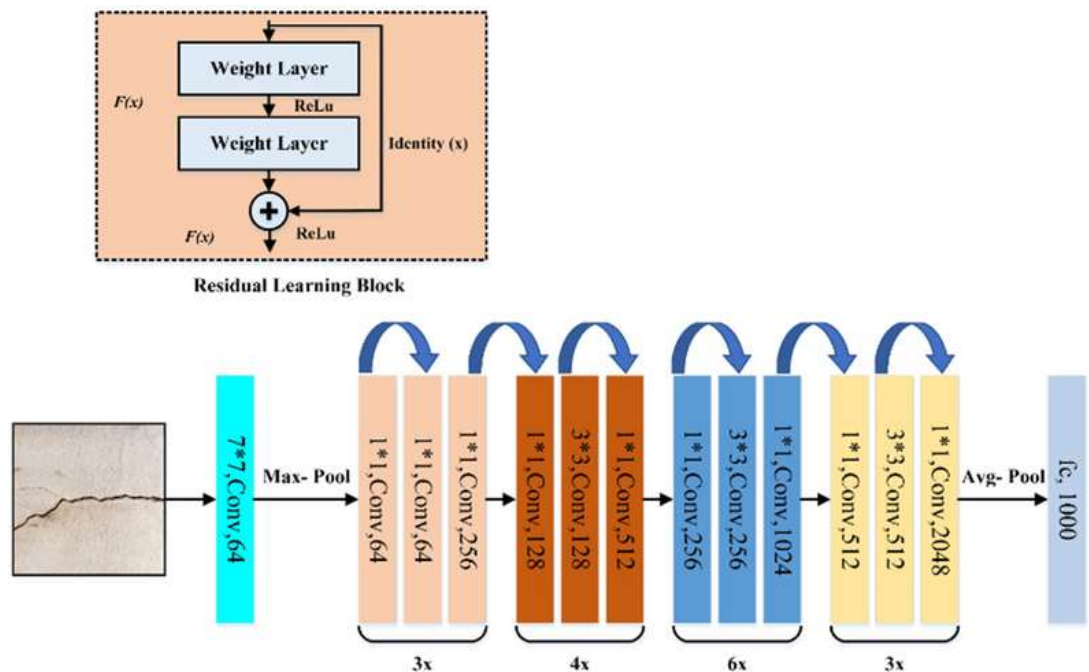
- **GoogLeNet/Inception-v1 (2014)** [12]: Πρόκειται για ένα CNN που καινοτομεί με την εισαγωγή του «inception module», δηλαδή την ταυτόχρονη πράξη πολλαπλών συνελκτικών φίλτρων και συνενώσεων στην ίδια στρώση. Με αυτό τον τρόπο προσφέρονται τα πλεονεκτήματα εφαρμογής ποικιλίας διαστάσεων φίλτρων (τοπικές λεπτομέρειες + καθολική πληροφορία), χωρίς να «βαθαίνει» το δίκτυο με περισσότερα επίπεδα που απαιτούν περισσότερη υπολογιστική ισχύ. Ακόμη, χρησιμοποιεί τμήματα ταξινόμησης «ενδιάμεσα» στο δίκτυο για αποφυγή εξαφανιζόμενων παραγώγων. Κατέκτησε τη νίκη στο διαγωνισμό ImageNet Large Scale Visual Recognition Challenge (ILSVRC) του 2014. Παρά τις 22 στρώσεις του, φέρει μονάχα ~6.6 εκατομμύρια παραμέτρους.



Σχήμα 7: Σχεδιάγραμμα της αρχιτεκτονικής GoogLeNet/Inception-v1. (πηγή: <https://medium.com/analytics-vidhya/cnns-architectures-lenet-alexnet-vgg-googlenet-resnet-and-more-666091488df5>, προσπελάστηκε Αυγ. 2023)



- **ResNet (2014)** [13]: Το συγκεκριμένο δίκτυο υπάρχει σε πολλαπλές εκδόσεις που κάποιες ξεπερνούν τις 100 στρώσεις: ResNet-50 ~25.6 εκατομμύρια παράμετροι, ResNet-101 ~44.5 εκατομμύρια παράμετροι, ResNet-152 ~60.2 εκατομμύρια παράμετροι. Παρόλο που πρόκειται για ιδιαίτερα «βαθύ» CNN, το πρόβλημα των εξαφανιζόμενων παραγώγων αντιμετωπίζεται με την εφαρμογή **προσπέρασης συνδέσεων (skip connections)**. Με αυτό το μηχανισμό, το μοντέλο προωθεί δεδομένα εισόδου στο αποτέλεσμα εξόδου ενός σετ συνελκτικών στρώσεων, οπότε η συμπεριφορά μιας βαθύτερης στρώσης αναμένεται σίγουρα μη χειρότερη (**Residual Learning, εκμάθηση identity function**). Κατά το διαγωνισμό ImageNet Large Scale Visual Recognition Challenge (ILSVRC) του 2015, ξεπέρασε τις επιδόσεις παλαιότερων αρχιτεκτονικών και κατέκτησε την πρώτη θέση. Πρόκειται για ένα CNN που αποδίδει εξαιρετικά στα προβλήματα του τομέα της Όρασης Υπολογιστών.



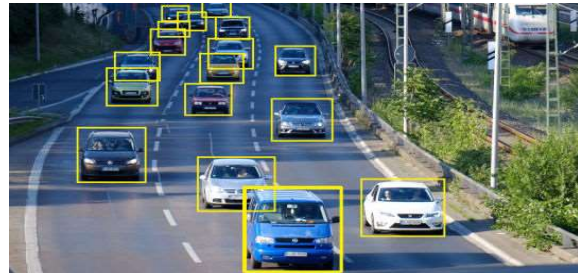
Σχήμα 8: Σχεδιάγραμμα της αρχιτεκτονικής ResNet-50 (κάτω), καθώς και του Residual Learning Blog (άνω). (πηγή: [42])

2.1.3: Ανίχνευση αντικειμένων

Η **ανίχνευση αντικειμένων (Object Detection)** αποτελεί μια τεχνική του πεδίου της Όρασης Υπολογιστών, που αφορά την εύρεση της παρουσίας αντικειμένων σε δεδομένα εικόνας, και κατάταξη τους σε κάποια προκαθορισμένη κατηγορία. Το εγχείρημα αναλύεται σε δύο προβλήματα επιβλεπόμενης μάθησης: ένα πρόβλημα ταξινόμησης (classification) για την αναζήτηση μιας κατηγορίας αντικειμένου, κι ένα πρόβλημα παλινδρόμησης (regression) για το χωρικό εντοπισμό (localization) του αντικειμένου της εν λόγω κατηγορίας στην εικόνα.

Στη δισδιάστατη ανίχνευση αντικειμένων, η επίλυση του χωρικού εντοπισμού ενός αντικειμένου ορίζεται με ένα **κουτί οριοθέτησης (bounding box)**. Πρόκειται για το ορθογώνιο παραλληλόγραμμο με το ελάχιστο εμβαδόν που περιγράφει την απεικόνιση του αντικειμένου σε μια εικόνα, και κατά γενική περίπτωση οι πλευρές του είναι παράλληλες στους άξονες x και y (up-right bounding box). Ένα up-right bounding box, επομένως, συνήθως προσδιορίζεται υπό τα παρακάτω πρότυπα:

- **Pascal VOC format:**
(xmin, ymin, xmax, ymax)
- **COCO format:**
(xmin, ymin, width, height)
- **YOLO format:**
(xcenter, ycenter, width, height)



Σχήμα 9: Παραδείγματα εξαγωγής «up-right» bounding box σε αυτοκίνητα ως αντικείμενα ενδιαφέροντος. (πηγή: <https://www.precisebposolution.com/bounding-box.html>, προσπελάστηκε Αυγ. 2023)

Ωστόσο, σε πιο εξειδικευμένες εφαρμογές της ανίχνευσης αντικειμένων, όπως την προσανατολισμένη ανίχνευση αντικειμένων (oriented object detection), ή την κατάτμηση δειγμάτων (instance segmentation), ζητείται το προσανατολισμένο («μη παράλληλο κατά x,y») bounding box, και η μάσκα περιγράμματος σε επίπεδο εικονοστοιχείου, αντίστοιχα.



Σχήμα 10: Παραδείγματα oriented object detection (αριστερά) και instance segmentation (δεξιά). (πηγές: <https://developer.nvidia.com/blog/detecting-rotated-objects-using-the-odtk/>, <https://datascience.eu/computer-vision/mask-r-cnn/>, προσπελάστηκαν Αυγ. 2023)

Τα μοντέλα ανίχνευσης αντικειμένων που βασίζονται στη Βαθιά Μάθηση διακρίνονται σε δύο γενικές κατηγορίες βάσει της μορφής αρχιτεκτονικής τους:

- **Ανιχνευτές δύο σταδίων (2 – stage detectors):** Αντιμετωπίζουν το πρόβλημα σε δύο στάδια. Πρώτα, από τον εξαγόμενο πίνακα χαρακτηριστικών από ένα «backbone» CNN δημιουργείται ένα **δίκτυο προτάσεων περιοχής (Region Proposal Network, RPN)** υπολογίζοντας ένα μέτρο ύπαρξης ενός αντικειμένου (objectness) –οποιασδήποτε κατηγορίας– σε μια περιοχή της εικόνας, όπου για κάθε προτασόμενη περιοχή εξάγεται κι ο αντίστοιχος νέος πίνακας χαρακτηριστικών, σχηματίζοντας τις **περιοχές ενδιαφέροντος (Regions Of Interest, ROI)**. Κατά το δεύτερο στάδιο, γίνεται ανίχνευση ανά περιοχή ενδιαφέροντος προσδιορίζοντας την ταξινόμηση και το χωρικό εντοπισμό (detection network). Πιο πολύπλοκοι και πιο «αργοί» ανιχνευτές, αλλά συνήθως πιο εύρωστοι και αποδοτικοί.
- **Ανιχνευτές ενός σταδίου (1 – stage detectors):** Πραγματοποιούν ανίχνευση αντικειμένων σε ένα στάδιο με τη διέλευση των δεδομένων από ένα συνεχές και, συνήθως, πλήρως συνελκτικό δίκτυο.

Στη μετ' έπειτα ενότητα της μεθοδολογίας, χρησιμοποιείται ο πασίγνωστος ανιχνευτής δύο σταδίων **Faster Region-based Convolutional Neural Network (Faster R-CNN)** [14]. Ο συγκεκριμένος detector καινοτομεί με τη χρήση των «αγκυρώσεων» (**anchors**), δηλαδή ομάδων από παραλληλόγραμμα διαφορετικών κλιμάκων και αναλογιών πλευρών που χρησιμοποιούνται για την πρόταση περιοχών ύπαρξης αντικειμένων (κατά «κυλιόμενα παράθυρα»). Αναλυτικά, η πρωτότυπη ροή εκτέλεσής του για «διέλευση» μιας εικόνας:

- Η αρχική εικόνα διέρχεται μέσω ενός κοινού (shared) backbone CNN (κατά την πρωτότυπη έκδοση, χρησιμοποιείται το VGG-16) για τα τμήματα RPN και detection network της παλαιότερης αρχιτεκτονικής Fast R-CNN, κι εξάγονται πίνακες χαρακτηριστικών «ταυτόχρονα» και για τα δύο μέρη.
- Αναφορικά με το τμήμα RPN:
 - Τοποθετούνται anchors σε θέσεις της αρχικής εικόνας (anchor locations). Κάθε εικονοστοιχείο feature map αντιστοιχίζεται σε ένα anchor location και άρα σε μια ομάδα anchors (9 συγκεκριμένα by default) επί της εικόνας.
 - Ο πίνακας χαρακτηριστικών φιλτράρεται από συνελκτική στρώση (3x3), εξάγοντας χαρακτηριστικά για κάθε anchor, κι έπειτα διακρίνονται δύο κλάδοι:
 - Ένας κλάδος δυαδικής ταξινόμησης σε background ή foreground που πραγματοποιεί (1x1) συνέλιξη, από το οποίο εξάγεται η

πιθανότητα κάθε ένα από τα 9 anchors να περιέχει ένα αντικείμενο ($2 \times 9 = 18$ bands).

- Ένας κλάδος παλινδρόμησης bounding box που εφαρμόζει στρώση (1x1) συνέλιξης, και προσαρμόζει κάθε μια από τις 4 συντεταγμένες του προτεινόμενου bounding box ($4 \times 9 = 36$ bands).
- Στη δυαδική ταξινόμηση το μοντέλο εξετάζει την ύπαρξη αντικειμένων στην αρχική εικόνα, ελέγχοντας την ύπαρξη πληροφορίας αλήθειας σε κάθε σημείο στον feature map, το οποίο αντιστοιχίζεται σε μια ομάδα anchors που αφορά την αρχική εικόνα. Ο έλεγχος αυτός πραγματοποιείται με τον υπολογισμό του μέτρου objectness σε σύγκριση με την πληροφορία αλήθειας, βάσει της μετρικής **Intersection Over Union (IoU)**:

$$Objectness_{score} = \begin{cases} \text{Θετικό αν } IoU > 0.5 \\ \text{Μη αρνητικό - Θετικό αν } IoU \in [0.3, 0.5] \\ \text{Αρνητικό αν } IoU < 0.3 \end{cases}$$

$$\text{με } IoU = \frac{bbox_{gt} \cap anchor}{bbox_{gt} \cup anchor}$$

- Χρησιμοποιείται η εξής πολυδιεργατική συνάρτηση κόστους:

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \cdot \sum_i L_{cls}(p_i, p_i^*) + \lambda \cdot \frac{1}{N_{reg}} \cdot \sum_i p_i^* \cdot L_{reg}(t_i, t_i^*)$$

$$\text{με } L_{cls}(p_i, p_i^*) = BCE(p_i, p_i^*), L_{reg}(t_i, t_i^*) = L1_{loss}(t_i, t_i^*)$$

όπου $i = 0$ δείκτης anchor σε ένα mini-batch, $p_i = \eta$ πιθανότητα ενός anchor να περιγράφει τη θέση ενός αντικειμένου, $p_i^* = \eta$ ετικέτα πληροφορίας αλήθειας η οποία είναι 1 αν το anchor θεωρείται θετικό και 0 αν θεωρείται αρνητικό (βάσει $objectness_{score}$), $t_i = \text{το δiάνυσμα 4 παραμέτρων ενός προτεινόμενου bounding box}$, $t_i^* = \text{το δiάνυσμα 4 παραμέτρων του bounding box αλήθειας που συνδέεται στο θετικό anchor (βάσει } objectness_{score})$, $N_{cls} = \text{το μέγεθος mini-batch}$, $N_{reg} = \text{το πλήθος anchor θέσεων}$.

- Στην παλινδρόμηση – διόρθωση των bounding boxes χρησιμοποιούνται οι παρακάτω παραμετροποιήσεις:

$$t_x = \frac{x - x_a}{w_a}, t_y = \frac{y - y_a}{h_a}$$

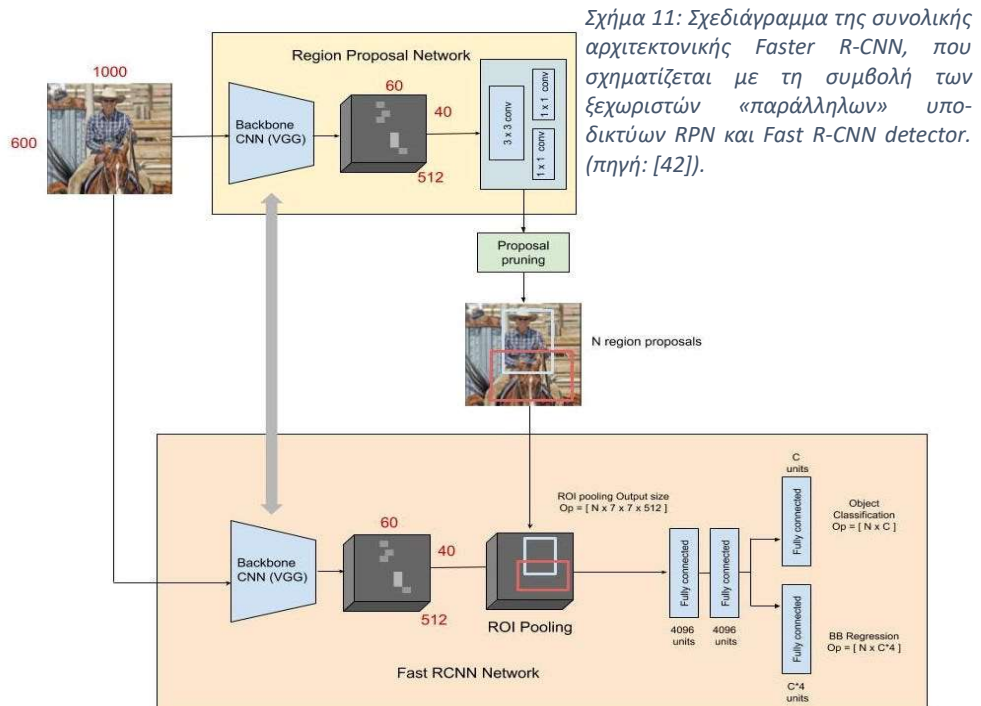
$$t_w = \log\left(\frac{w}{w_a}\right), t_h = \log\left(\frac{h}{h_a}\right)$$

$$t_x^* = \frac{x^* - x_a}{w_a}, t_y^* = \frac{y^* - y_a}{h_a}$$

$$t_w^* = \log\left(\frac{w^*}{w_a}\right), t_h^* = \log\left(\frac{h^*}{h_a}\right)$$

όπου (x, y, w, h) = οι παράμετροι προσδιορισμού του προτεινόμενου bounding box, (x_a, y_a, w_a, h_a) = οι παράμετροι προσδιορισμού του anchor, (x^*, y^*, w^*, h^*) = οι παράμετροι προσδιορισμού του bounding box αλήθειας.

- Επειδή το RPN παράγει πολλαπλές προτάσεις bounding box, και μάλιστα επικαλυπτόμενες, επιλέγονται μόνο αυτές με την μεγαλύτερη **πιθανότητα πρόβλεψης (confidence)** από όσες επικαλύπτονται (**Non Maximum Suppression, NMS**).
- Σχετικά με το τμήμα detection network του Fast R-CNN:
 - Οι προτάσεις προβάλλονται στους πίνακες χαρακτηριστικών που εξήχθησαν από το backbone και η περιοχή τους απομονώνεται δημιουργώντας τις **περιοχές ενδιαφέροντος (Region Of interest, RoI)**.
 - Οι περιοχές ενδιαφέροντος χωρίζονται σε κελιά τιμών και με pooling μετασχηματίζονται σε πίνακες χαρακτηριστικών προκαθορισμένου μεγέθους (**RoI pooling layers**).
 - Τέλος, ακολουθούν 2 fully-connected στρώσεις, και έπειτα γίνεται διαχωρισμός σε δύο κλάδους:
 - Έναν fully-connected κλάδο ταξινόμησης σε C πλήθος κατηγοριών.
 - Έναν fully-connected κλάδο για παλινδρόμηση των προβλεπόμενων bounding box (C παλινδρομητές, ένας για κάθε κατηγορία).



Επίσης, όσον αφορά την ακολουθία εκπαίδευσης, οι δημιουργοί του Faster R-CNN προτείνουν μια μέθοδο **εναλλασσόμενης προσαρμογής RPN κι Fast R-CNN**. Δηλαδή, εκπαίδευση του RPN (το οποίο έχει αρχικοποιηθεί στο ImageNet) και χρήση των εξαγόμενων προτάσεων για εκπαίδευση και του Fast R-CNN detector. Έπειτα, ο συντονισμός του Fast R-CNN detector αρχικοποιεί το RPN, και η διαδικασία αυτή επαναλαμβάνεται.

2.1.4: Παρακολούθηση αντικειμένων

Το εγχείρημα της **παρακολούθησης αντικειμένων (Object Tracking)** «διαδέχεται» αυτό της ανίχνευσης αντικειμένων. Αφορά τον χωρικό εντοπισμό και την μονοσήμαντη ταυτοποίηση ενός αντικειμένου – στόχου, σε κάθε καρέ μιας ακολουθίας βίντεο. Πρόκειται για ένα σύνθετο πρόβλημα της Όρασης Υπολογιστών, καθώς απαιτεί την αντιμετώπιση των προκλήσεων αλλαγής των στοιχείων εμφάνισης του στόχου κατά την κίνηση του με την πάροδο του χρόνου.

Η **παρακολούθηση ενός μοναδικού αντικειμένου (Single Object Tracking, SOT)** σε μια αλληλουχία σκηνών αποτελεί το απλούστερο πρόβλημα tracking, καθώς δεν απαιτεί απαραίτητα ως συνιστώσα μια μονάδα ανίχνευσης αντικειμένων. Ο χρήστης αρχικοποιεί ένα μοντέλο με τη χειροκίνητη εισαγωγή ενός bounding box που περιέχει το στόχο στο αρχικό καρέ μιας ακολουθίας (target region), κι έπειτα με τη συνεργασία ενός **μοντέλου κίνησης (motion model)**, ενός **μοντέλου εμφάνισης (appearance model)**, κι ενός **μοντέλου συσχέτισης δεδομένων (data association model)**, επιτυγχάνεται η συνεχής πρόβλεψη της θέσης του στόχου (με bounding boxes) σε διαδοχικά καρέ, μέχρις ότου αυτή τερματιστεί. Ο εν λόγω τερματισμός (termination) μπορεί να επέλθει «επιθυμητά», όπως με την έξοδο του αντικειμένου από την περιοχή ενδιαφέροντος, αλλά και «ανεπιθύμητα», όπως από απότομες αλλαγές κίνησης (sudden motion changes) και συγκαλύψεις (occlusions). Πάντως, στη δεύτερη περίπτωση η χρήση ενός object detector μπορεί να συμβάλει σημαντικά στον επανενοτοπισμό.

Στην **παρακολούθηση πολλαπλών αντικειμένων (Multiple Object Tracking, MOT)**, η αποφυγή χρήσης detector, αν και εφικτή, φέρει πολλά μειονεκτήματα, όπως η ανάγκη χειροκίνητης αρχικοποίησης πολλαπλών target regions, η απώλεια ευρωστίας σε μεταβολές εμφάνισης και κίνησης σε πολλαπλά αντικείμενα, καθώς και η αδυναμία διαχείρισης αντικειμένων που εισέρχονται στο πλάνο μελέτης σε μεταγενέστερο χρόνο. Αυτομάτως, γίνεται εμφανές ότι οι εφαρμογές MOT είναι αρκετά πολυπλοκότερες από τις SOT, αλλά ως προς τη σημασιολογία τους δε διαφέρουν ιδιαίτερα. Πρόκειται για την επέκταση του προβλήματος εντοπισμού και ταυτοποίησης σε περισσότερους στόχους ταυτόχρονα, σε τη διάρκεια μιας ακολουθίας βίντεο, ενώ ενδέχεται να ανήκουν σε μια ποικιλία διαφορετικών κατηγοριών.

Παρακάτω παρατίθενται οι δύο γενικότερες οικογένειες αλγορίθμων πολλαπλής παρακολούθησης αντικειμένων που χρησιμοποιούν τη συμβολή μονάδας ανίχνευσης αντικειμένων:

- **Tracking-by-Detection (TbD):** Τα συγκεκριμένα μοντέλα χρησιμοποιούν μια προσέγγιση που περιλαμβάνει τη συνεργασία δύο μελών. Το ένα μέλος αφορά την ανίχνευση αντικειμένων, όπου χρησιμοποιείται ένας ανιχνευτής ώστε να εξάγει προβλέψεις μεμονωμένα σε κάθε καρέ της ακολουθίας βίντεο. Κατά το δεύτερο μέλος, οι ανιχνεύσεις από πολλαπλά διαδοχικά καρέ συνδέονται μεταξύ τους με τη χρήση μοντέλων κίνησης, συσχέτισης δεδομένων, ή ενσωμάτωσης Βαθιάς Μάθησης (DL-based embeddings, παραπέμπει σε DL-based TbD), οπότε και εκτελείται η παρακολούθηση. Παραδείγματα τέτοιων μοντέλων αποτελούν οι Tractor [15], DeepSORT [16], FairMOT [17], CenterTrack [18], κ.α..
- **Joint Detection and Tracking (JDT):** Σε αυτή την κατηγορία ανήκουν δομές που πραγματοποιούν ταυτόχρονα ανίχνευση και παρακολούθηση πολλαπλών αντικειμένων σε μια ενιαία end-to-end προγραμματιστική ακολουθία. Τέτοιοι trackers είναι οι ByteTrack [19], QDTrack [20], RetinaTrack [21] κ.α..

Ο tracker που χρησιμοποιείται για τα πειράματα της ανάπτυξης μεθοδολογίας, **Deep Simple Online and Real Time Tracking** ή **DeepSORT** [16], ανήκει στην οικογένεια μοντέλων DL-based TbD, και αποτελεί μια προσέγγιση βελτίωσης των επιδόσεων του προγενέστερου αλγορίθμου **Simple Online and Real Time Tracking (SORT)** [22], χρησιμοποιώντας ως καινοτομία την ενσωμάτωση DL-based χαρακτηριστικών εμφάνισης. Αναλυτικά ο πρωτότυπος τρόπος λειτουργίας του:

- Αρχικά, χρησιμοποιείται ένας ανεξάρτητα εκπαιδευμένος ανιχνευτής, προεπιλεγμένα ο **Faster R-CNN**, για να εξάγει τις ανιχνεύσεις των αντικειμένων που απεικονίζονται στο καρέ t της ακολουθίας, σε μορφή bounding box.
- Οι ανιχνεύσεις αυτές προβάλλονται στο καρέ $t+1$ ως παρατηρήσεις αρχικοποίησης ενός **μοντέλου κίνησης φίλτρου Kalman** (σταθερών ταχυτήτων και γραμμικής πρόβλεψης), το οποίο δύναται να προβλέψει θέσεις bounding boxes σε επόμενα frame με ομαλή διαδοχή, και άρα να συνάψει τροχιές ξεχωριστών id.
- Η **συσχέτιση μεταξύ στόχων (association)** σε διαφορετικά καρέ γίνεται με συνεργασία των στοιχείων κίνησης και εμφάνισης των αντικειμένων:
 - Σε επόμενα καρέ ο Faster R-CNN εξάγει νέες ανιχνεύσεις, οπότε υπάρχουν δύο είδη bounding box: αυτά που έχουν προκύψει από πρόβλεψη του detector, και αυτά που έχουν «διαδοθεί» από τις καταστάσεις του μοντέλου κίνησης. Για την αφομοίωση της πληροφορίας κίνησης των στόχων,

χρησιμοποιείται πρώτα η τετραγωνική απόσταση Mahalanobis μεταξύ των δύο αυτών χωρικών ενδείξεων ως εξής:

$$d^{(1)}(i, j) = (d_j - y_i)^T \cdot S_i^{-1} \cdot (d_j - y_i)$$

όπου y_i = το διάνυσμα των στοιχείων του bounding box που «διαδίδεται» από το μοντέλο κίνησης, S_i = ο πίνακας συνδιακύμανσης που ορίζει το επίπεδο αξιοπιστίας της «διάδοσης», και d_j = το διάνυσμα των στοιχείων bounding box που προβλέπεται από τον detector

Βάσει μόνο της προηγούμενης μετρικής, απορρίπτονται με ποσοστό εμπιστοσύνης 95% ακατάλληλες συσχετίσεις με τη βοήθεια της παρακάτω δυαδικής κατωφλίωσης:

$$b_{i,j}^{(1)} = \mathbb{1}[d^{(1)}(i, j) \leq t^{(1)}]$$

όπου $t^{(1)} = 9.4877$ η πρωτότυπη τιμή κατωφλίου απόστασης

Η απόσταση Mahalanobis γενικά είναι επαρκής για τη συσχέτιση αντικειμένων βραχυπρόθεσμα όταν οι διαδόσεις του μοντέλου κίνησης είναι αξιόπιστες, αλλά στα προβλήματα όπου υπάρχουν μετατοπίσεις της κάμερας, και μη σταθερές ταχύτητες των αντικειμένων, αυτό δε συμβαίνει. Επομένως, δημιουργείται μια δεύτερη μετρική, ανθεκτική απέναντι σε μακροπρόθεσμα occlusions, η οποία μετράει τη μικρότερη γωνιακή απόσταση (cosine distance) μεταξύ του state i της διαδιδόμενης τροχιάς, και της j ανίχνευσης του detector:

$$d^{(2)}(i, j) = \min \{1 - r_j^T \cdot r_k^{(i)} | r_k^{(i)} \in R_i\}$$

όπου r_j = ο descriptor εμφάνισης της j ανίχνευσης bounding box ($\|r_j\| = 1$),

$R_k = \{r_k^{(i)}\}_{k=1}^{L_k}$ το σύνολο των τελευταίων $L_k = 100$ συσχετισμένων descriptor εμφάνισης της τροχιάς k

Ομοίως με πριν, είναι εφικτή η απόρριψη ακατάλληλων συσχετίσεων ως εξής:

$$b_{i,j}^{(2)} = \mathbb{1}[d^{(2)}(i, j) \leq t^{(2)}]$$

Οι δύο μετρικές ενσωματώνονται σταθμισμένα σε μια νέα μετρική, ενώ προκύπτει ένα ακόμη δυαδικό κριτήριο καταλληλότητας συσχετίσεων από το γινόμενο των προηγούμενων κριτηρίων:

$$c_{i,j} = \lambda \cdot d^{(1)}(i, j) + (1 - \lambda) \cdot d^{(2)}(i, j)$$

$$b_{i,j} = \prod_{m=1}^2 b_{i,j}^{(m)}$$

όπου λ = υπερπαραμέτρος ($\lambda = 0$ όταν υπάρχει αρκετή μετατόπιση κάμερας, ώστε να μη συμμετέχει η μετρική απόστασης Mahalanobis)

- Οι **descriptors εμφάνισης** που υπεισέρχονται στη δεύτερη μετρική ως embeddings αποτελούν και την βασική καινοτομία του DeepSORT tracker. Για την απόκτησή τους χρησιμοποιείται ένα CNN (εξ' ου και ο χαρακτηρισμός «DL-based» embeddings) το οποίο εκπαιδεύεται ανεξάρτητα στην επαναταυτοποίηση (reidentification) αντικειμένων, δηλαδή στο πρόβλημα ταξινόμησης του ίδιου αντικειμένου όπως απεικονίζεται στην τροχιά του σε διαδοχικά καρέ ενός βίντεο. Το πρωτότυπο δίκτυο εκπαιδεύτηκε σε reidentification dataset που περιλάμβανε πάνω από 1.1 εκατομμύρια εικόνες από 1261 πεζούς.

Name	Patch Size/Stride	Output Size
Conv 1	3 × 3/1	32 × 128 × 64
Conv 2	3 × 3/1	32 × 128 × 64
Max Pool 3	3 × 3/2	32 × 64 × 32
Residual 4	3 × 3/1	32 × 64 × 32
Residual 5	3 × 3/1	32 × 64 × 32
Residual 6	3 × 3/2	64 × 32 × 16
Residual 7	3 × 3/1	64 × 32 × 16
Residual 8	3 × 3/2	128 × 16 × 8
Residual 9	3 × 3/1	128 × 16 × 8
Dense 10		128
Batch and ℓ_2 normalization		128

Σχήμα 12: Η αρχιτεκτονική του CNN που χρησιμοποιείται κατά τον πρωτότυπο αλγόριθμο DeepSORT για την απόκτηση των περιγραφών εμφάνισης των αντικειμένων. (πηγή: [16])

- Το τελικό matching ανιχνεύσεων - τροχιών δρομολογείται ως εξής:
 - Ορίζονται ως είσοδος οι δείκτες των τροχιών $T = \{1, \dots, N\}$ και των ανιχνεύσεων $D = \{1, \dots, M\}$, ενώ εισάγεται η υπερπαραμέτρος $A_{\max} = \eta$ μέγιστη διάρκεια κατά την οποία διατηρείται «ζωντανή» μια τροχιά ενώ δεν έχουν υπάρξει συσχετίσεις με κάποια ανίχνευση.
 - Υπολογίζεται ο πίνακας κόστους $C = [c_{i,j}]$ των συσχετίσεων, καθώς και ο πίνακας με τις αποδεκτές συσχετίσεις $B = [b_{i,j}]$.
 - Αφού αρχικοποιηθεί το σύνολο «unmatched detections» $U \leftarrow D$, και «matches» $M \leftarrow \emptyset$, ξεκινάει μια δομή επανάληψης «για n εντός μιας διάρκειας A_{\max} », όπου εκχωρούνται σε ένα σύνολο T_n στοιχεία του συνόλου T τα οποία δεν έχουν συσχετιστεί με κάποια ανίχνευση στα τελευταία n καρέ.
 - Με τη βοήθεια του Hungarian αλγορίθμου αντιστοιχούνται με ελαχιστοποίηση του κόστους C τα στοιχεία των συνόλων T_n, U , εξάγοντας τις πιθανές συσχετίσεις $[x_{i,j}]$.
 - Τα σύνολα matches M και unmatched detections U ανανεώνονται βάσει των έγκυρων συσχετίσεων $[x_{i,j}]$, όπως αυτές φιλτράρονται από το συνολικό κριτήριο $b_{i,j}$. Η δομή επανάληψης τελειώνει κι επιστρέφονται τα ανανεωμένα αυτά σύνολα ως έξοδος.
 - Με αυτή τη μέθοδο αντιστοιχούνται με προτεραιότητα τροχιές του μοντέλου κίνησης οι οποίες έχουν πιο συχνές συσχετίσεις με ανιχνεύσεις.

- Τέλος, εφαρμόζεται μια συσχέτιση βάσει IoU μεταξύ των μη αντιστοιχημένων ή ακατάλληλων τροχιών μόνο για $n = 1$, ώστε να ληφθούν έγκυρες περιπτώσεις ξαφνικής αλλαγής εμφάνισης, και προβληματικής αρχικοποίησης του μοντέλου κίνησης φίλτρου Kalman.

Listing 1 Matching Cascade

Input: Track indices $\mathcal{T} = \{1, \dots, N\}$, Detection indices $\mathcal{D} = \{1, \dots, M\}$, Maximum age A_{\max}

- 1: Compute cost matrix $\mathbf{C} = [c_{i,j}]$ using Eq. 5
- 2: Compute gate matrix $\mathbf{B} = [b_{i,j}]$ using Eq. 6
- 3: Initialize set of matches $\mathcal{M} \leftarrow \emptyset$
- 4: Initialize set of unmatched detections $\mathcal{U} \leftarrow \mathcal{D}$
- 5: **for** $n \in \{1, \dots, A_{\max}\}$ **do**
- 6: Select tracks by age $\mathcal{T}_n \leftarrow \{i \in \mathcal{T} \mid a_i = n\}$
- 7: $[x_{i,j}] \leftarrow \text{min_cost_matching}(\mathbf{C}, \mathcal{T}_n, \mathcal{U})$
- 8: $\mathcal{M} \leftarrow \mathcal{M} \cup \{(i, j) \mid b_{i,j} \cdot x_{i,j} > 0\}$
- 9: $\mathcal{U} \leftarrow \mathcal{U} \setminus \{j \mid \sum_i b_{i,j} \cdot x_{i,j} > 0\}$
- 10: **end for**
- 11: **return** \mathcal{M}, \mathcal{U}

Σχήμα 13: Σύνοψη της πρωτότυπης αλγοριθμικής αλληλουχίας αντιστοίχισης μεταξύ ανιχνεύσεων και υπαρχόντων τροχιών στον DeepSORT tracker. (πηγή: [16])

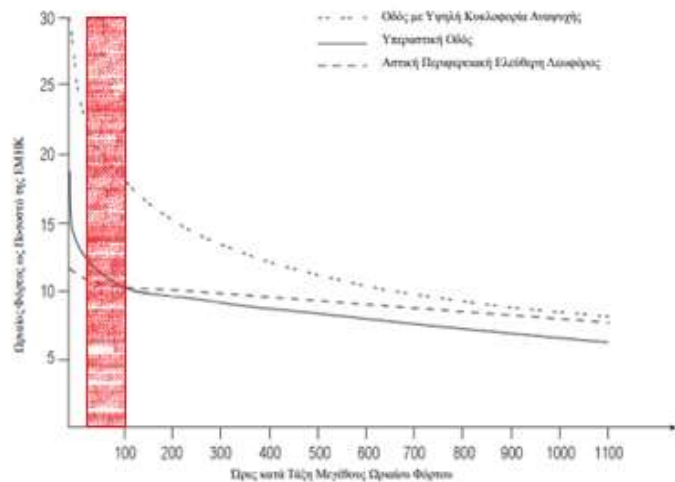
2.2: Θεωρητικό σκέλος κυκλοφοριακής τεχνικής

2.2.1: Βασικά μεγέθη κυκλοφοριακής ροής

Η κυκλοφοριακή ροή (traffic flow) αποτελεί την κεντρική έννοια στον τομέα της κυκλοφοριακής τεχνικής, καθώς αφορά γενικά την κίνηση των οχημάτων σε ένα οδικό τμήμα. Περιγράφεται ανά επιμέρους τμήματα από τα **μεγέθη κυκλοφοριακής ροής (traffic flow variables)**.

[23] Πρωταρχικό κυκλοφοριακό μέγεθος αποτελεί ο **κυκλοφοριακός φόρτος (traffic volume)**, ο οποίος ορίζεται ως ο *συνολικός αριθμός οχημάτων που διέρχονται από ένα σημείο η μια διατομή οδού στη μονάδα του χρόνου*, και συνήθως εκφράζεται σε οχήματα ανά ημέρα, ή ώρα κ.ο.κ.. Συνήθως δίνεται βαρύτητα στον **ισοδύναμο ωριαίο ρυθμό ροής (equivalent hour flow rate)**, ο οποίος αποτελεί το *συνολικό αριθμό οχημάτων που διέρχονται από ένα σημείο ή μια διατομή οδού κατά τη διάρκεια μιας χρονικής περιόδου μικρότερης της ώρας (συνήθως 15'), ανηγμένο στην ώρα*, και προφανώς εκφράζεται πάντοτε σε «οχήματα ανά ώρα». Όσον αφορά

την τιμή φόρτου που χρησιμοποιείται στο σχεδιασμό, λαμβάνεται υπόψη η **ωριαία διακύμανση** του φόρτου κατά τη διάρκεια μιας ημέρας. Συγκεκριμένα, οι οδοί σχεδιάζονται ώστε να μπορούν να εξυπηρετήσουν κυκλοφοριακούς φόρτους που κυμαίνονται μεταξύ των θέσεων 30 και 100 (ανάλογα την κατηγορία οδού) στην κατάταξη κυκλοφοριακών ωριαίων **φόρτων αιχμής** κατά τη διάρκεια ενός έτους. Επιπλέον, υπεισέρχεται και μια αύξηση λόγω της πιθανής ανομοιόμορφης κατανομής του φόρτου εντός της ώρας, με τη χρήση του **Συντελεστή Ώρας Αιχμής, ΣΩΑ (Peak Hour Factor, PHF)**, ο οποίος ορίζεται ως εξής:



Σχήμα 14: Στο διάγραμμα φθίνουσας κατάταξης ωρών αιχμής ως προς το μέγεθος ωριαίου φόρτου, η χαρακτηριστική ώρα σχεδιασμού επιλέγεται μεταξύ της 30^{ης} και 100^{ης} εξ αυτών. Εντός αυτής της περιοχής (με κόκκινο στο σχήμα) παρατηρείται ότι οι καμπύλες παρουσιάζουν έντονη κυρτότητα, οπότε και αποκλείονται του σχεδιασμού οι ακραίοι αλλά ιδιαίτερα σπάνιοι φόρτοι (~0.3-1% της διάρκειας του έτους). Η «ΕΜΗΚ» που εμφανίζεται στον άξονα γ αφορά την Ετήσια Μέση Ημερήσια Κυκλοφορία, δηλαδή το συνολικό αριθμό διερχόμενων οχημάτων σε ένα έτος διά τις μέρες του έτους. (πηγή: [23])

$$\Sigma\Omega A = \frac{V}{4 \cdot V_{max15}}$$

όπου V = ο ωριαίος φόρτος (οχημ./ώρα), και V_{max15} = ο μέγιστος φόρτος 15 συνεχόμενων λεπτών εντός της ώρας αιχμής.

Στην κυκλοφοριακή τεχνική ορίζονται πολυάριθμα είδη ταχυτήτων, το καθένα με διαφορετική συγκοινωνιακή εφαρμογή. Παρακάτω παρατίθενται τα βασικότερα εξ αυτών [23]:

- **Ταχύτητα σημείου (spot speed):** Η ταχύτητα που έχει ένα όχημα όταν περνά από δεδομένο σημείο.
- **Ταχύτητα διαδρομής (travel speed):** Η μέση ταχύτητα με την οποία κινείται ένα όχημα από το σημείο προέλευσης στο σημείο προορισμού, υπολογιζόμενων και των καθυστερήσεών του. Χωρίς τις καθυστερήσεις, γίνεται αναφορά στην **ταχύτητα πορείας (running speed)**.
- **Ταχύτητα λειτουργίας (operating speed):** Η μέγιστη ταχύτητα διαδρομής με την οποία μπορεί να κινηθεί ένα όχημα σε δεδομένο τμήμα του οδικού δικτύου με την υπάρχουσα κυκλοφορία, χωρίς να ξεπεράσει τη μέγιστη ταχύτητα ασφαλείας.
- **Ταχύτητα ελεύθερης ροής (free flow speed):** Η ταχύτητα λειτουργίας σε τμήμα μιας αρτηρίας, σε συνθήκες ελεύθερης ροής.
- **Μέση ταχύτητα χρόνου (time mean speed):** Ο αριθμητικός μέσος των ταχυτήτων σημείου των οχημάτων που περνούν μπροστά από μια διατομή οδού σε μια δεδομένη χρονική περίοδο:

$$u_t = \frac{1}{n(x)} \cdot \sum_{i=1}^{n(x)} u_i(x)$$

όπου $u_i(x)$ = η ταχύτητα σημείου του οχήματος i που διέρχεται από μια διατομή x ,
 $n(x)$ = ο αριθμός των οχημάτων που διέρχονται από τη διατομή x .

- **Μέση ταχύτητα χώρου (space mean speed):** Ο αριθμητικός μέσος των ταχυτήτων σημείου που έχουν, σε μια ορισμένη χρονική στιγμή, όλα οχήματα που βρίσκονται σε ένα δεδομένο τμήμα της οδού:

$$u_s = \frac{L}{\sum_{j=1}^{N(T)} \frac{t_j}{N(T)}}$$

όπου t_j = είναι ο χρόνος που χρειάζεται το όχημα j για να διανύσει το οδικό τμήμα μήκους L , εκ των $N(T)$ οχημάτων που διένυσαν το εν λόγω οδικό τμήμα.

[23] Ένα ακόμη χρήσιμο μακροσκοπικά και ταυτόχρονα εύληπτο οπτικά κυκλοφοριακό μέγεθος αποτελεί η **πυκνότητα κυκλοφορίας (traffic density)**. Ορίζεται ως ο αριθμός των οχημάτων που κινούνται σε μια δεδομένη χρονική στιγμή στη μονάδα μήκους της οδού (συνήθως εκφράζεται με οχήματα ανά χιλιόμετρο), ενώ αλγεβρικά δίνεται ως εξής:

$$k(t) = \frac{n(t)}{L}$$

Συγγενές σημασιολογικά μέγεθος της πυκνότητας αποτελεί η **κατάληψη (occupancy)**, η οποία καθίσταται ως το ποσοστό του χρόνου που μια διατομή καταλαμβάνεται από όχημα:

$$O(x) = \frac{1}{T} \cdot \sum_i^{n(x)} \frac{L_i + d}{u_i(x)}$$

[23] Ακόμη, στην κυκλοφορία ο **διαχωρισμός (headway)**, δηλαδή η απόσταση μεταξύ δύο διαδοχικών οχημάτων σε μια λωρίδα, ως προς ένα σταθερό σημείο αναφοράς τους, εκφράζεται με δύο μεθόδους:

- Με μονάδα μήκους, ως **χωρικός διαχωρισμός (space headway)**. Ο μέσος χωρικός διαχωρισμός είναι υπολογίσιμος ως εξής:

$$\bar{d}(t) = \frac{\sum_{j=1}^{n(x)} d_j(t)}{n(t)}, \text{ ή } \bar{d}(t) \approx \frac{1}{k(t)} \text{ αν } L = 1 \text{ χλμ.}$$

όπου d_j = ο χωρικός διαχωρισμός του οχήματος j από το προπορευόμενο του.

- Με μονάδα χρόνου, ως **χρονικός διαχωρισμός (time headway)**. Αντίστοιχα για το μέσο χρονικό διαχωρισμό:

$$\bar{h}(x) = \frac{\sum_{i=1}^{n(x)} h_i(x)}{n(x)}$$

όπου h_i = ο χρονικός διαχωρισμός του οχήματος i από το προπορευόμενο του.



Σχήμα 15: Σχηματική απεικόνιση του χωρικού διαχωρισμού μεταξύ δύο οχημάτων. (πηγή: [23])

2.2.2: Θεμελιώδεις κι εμπειρικές σχέσεις μεταξύ κυκλοφοριακών μεγεθών

[23] Η θεωρία κυκλοφοριακής ροής βασίζεται στη θεμελιώδη σχέση που συνδέει τα τρία μακροσκοπικά μεγέθη της (φόρτος, πυκνότητα, μέση ταχύτητα χώρου):

$$q = k \cdot u_s$$

όπου q = ο κυκλοφοριακός φόρτος, k = η μέση πυκνότητα σε χρονικό διάστημα T , και u_s = η μέση ταχύτητα χώρου.

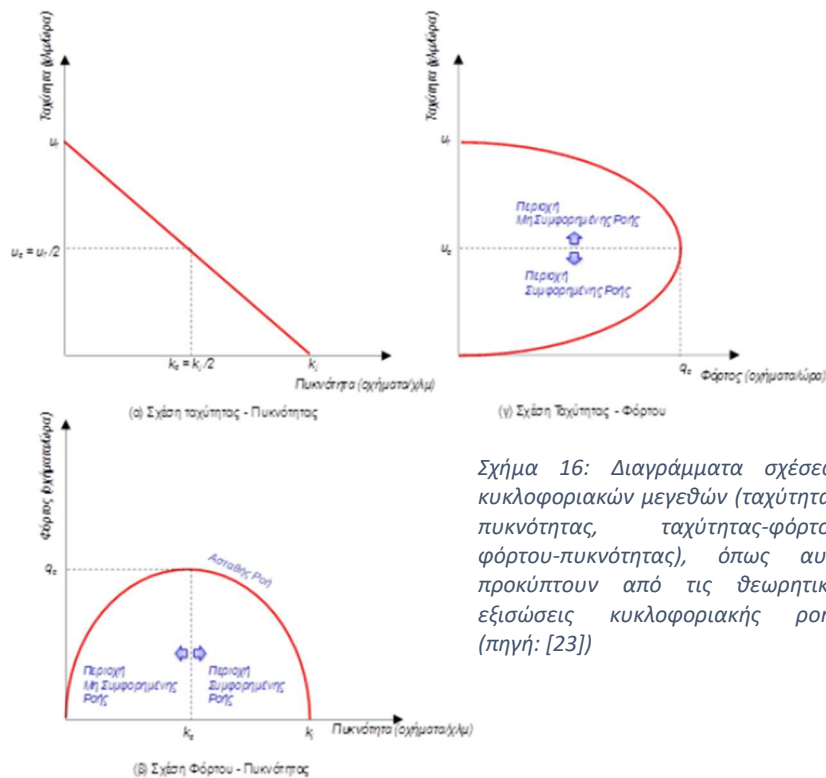
Περαιτέρω, με την ιδανική θεώρηση ότι η ταχύτητα χώρου και η πυκνότητα συνδέονται γραμμικά με μια σχέση της μορφής $u_s = u_f - \left(\frac{u_f}{k_j}\right) \cdot k$ (1) (όπου u_f = η ταχύτητα ελεύθερης ροής, k_j = η πυκνότητα κορεσμού), και πολλαπλασιάζοντας και τα δύο μέλη με k , προκύπτει η εξίσωση:

$$q = u_f \cdot k - \left(\frac{u_f}{k_j}\right) \cdot k^2 \quad (2)$$

Για $k \rightarrow 0$ (ανυπαρξία οχημάτων) ή $k \rightarrow k_j$ (ακινητοποιημένα οχήματα) προκύπτει $q \rightarrow 0$, ενώ η μέγιστη τιμή φόρτου q_c εμφανίζεται όταν $\frac{dq}{dk} = 0 \Leftrightarrow k = k_c = \frac{k_j}{2}$ (κρίσιμο σημείο συμφόρησης), ενώ αντικαθιστώντας υπολογίζεται $q_c = \frac{u_f \cdot k_j}{2}$. Με επίλυση της (1) ως προς k , και πολλαπλασιάζοντας και τα δύο μέλη με u_s , προκύπτει η εξίσωση:

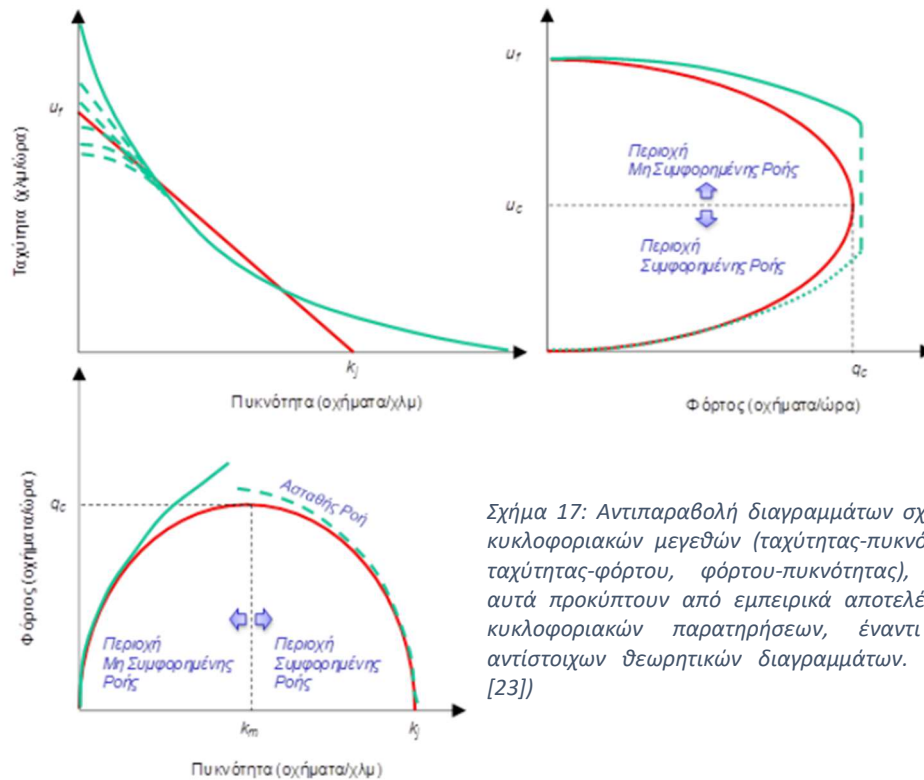
$$q = k_j \cdot u_s - \left(\frac{k_j}{u_f}\right) \cdot u_s^2 \quad (3)$$

Για $q \uparrow$ ισχύει ότι $u_s \downarrow$, όταν γίνεται αναφορά σε συνθήκες ελεύθερης ροής, με οριοθέτηση τη μέγιστη τιμή φόρτου $q = q_c$. Από την τιμή αυτή κι έπειτα ισχύει $k > k_c$ (μετάβαση σε συνθήκες συμφόρησης), οπότε ο φόρτος αρχίζει να μειώνεται, όπως και η ταχύτητα. Με την παραπάνω πληροφορία προκύπτουν τα αντίστοιχα θεμελιώδη διαγράμματα για τις σχέσεις ταχύτητας χώρου – πυκνότητας, φόρτου – ταχύτητας χώρου, και φόρτου - πυκνότητας.



Σχήμα 16: Διαγράμματα σχέσεων κυκλοφοριακών μεγεθών (ταχύτητας-πυκνότητας, ταχύτητας-φόρτου, φόρτου-πυκνότητας), όπως αυτά προκύπτουν από τις θεωρητικές εξισώσεις κυκλοφοριακής ροής. (πηγή: [23])

Πραγματοποιώντας στατιστική επεξεργασία από κυκλοφοριακές μετρήσεις, παρατηρήθηκε ότι τα θεμελιώδη διαγράμματα διαφέρουν ελαφρώς από την πραγματικότητα, διότι βασίζονται στην ιδανική θεώρηση ομογενούς και στάσιμης ροής. Η γραφική μοντελοποίηση των εμπειρικών αποτελεσμάτων των σχετικών ερευνών φαίνεται με αντιπαραβολή στα παρακάτω διαγράμματα:



Σχήμα 17: Αντιπαραβολή διαγραμμάτων σχέσεων κυκλοφοριακών μεγεθών (ταχύτητας-πυκνότητας, ταχύτητας-φόρτου, φόρτου-πυκνότητας), όπως αυτά προκύπτουν από εμπειρικά αποτελέσματα κυκλοφοριακών παρατηρήσεων, έναντι των αντίστοιχων θεωρητικών διαγραμμάτων. (πηγή: [23])

2.2.3: Μελέτη κυκλοφοριακής ικανότητας κι επιπέδου εξυπηρέτησης

Βασικός στόχος του τομέα κυκλοφοριακής τεχνικής αποτελεί ο προσδιορισμός της **κυκλοφοριακής ικανότητας (traffic capacity)** ενός οδικού συστήματος, δηλαδή *το μέγιστο οριαίο ρυθμό ροής οχημάτων υπό τις επικρατούσες οδικές/κυκλοφοριακές συνθήκες*, και η εξέταση των επιλογών που μπορούν να την αναβαθμίσουν. Ταυτόχρονα, ιδιαίτερο ενδιαφέρον αποδίδει η **στάθμη εξυπηρέτησης (level of service)** ως ένα άμεσο παράγωγο της κυκλοφοριακής ικανότητας, το οποίο αποτελεί ένα ποιοτικό μέγεθος αξιολόγησης ενός οδικού τμήματος, εύκολα κατανοητό και από το γενικό πληθυσμό. Πρόκειται για μια απλή κατάταξη από A (ελεύθερη ροή, free flow) έως F (κορεσμός, saturation), με την πρακτική εφαρμογής μιας σχετικής λειτουργικής ανάλυσης (operational analysis) να διαφέρει ανάλογα με τον τύπο συστήματος. Ωστόσο από τη σκοπιά της παρούσας διπλωματικής εργασίας, και σύμφωνα με το διαθέσιμο υλικό δεδομένων μελέτης, θα γίνει αναφορά στη μελέτη επιπέδου εξυπηρέτησης κυκλικών και ισόπεδων κόμβων φωτεινής σηματοδότησης, ακολουθώντας τα πρότυπα του **Highway Capacity Manual (HCM 2010)** [24].

Τονίζεται επίσης, ότι για την εξαγωγή επιπέδου εξυπηρέτησης με αξιοπιστία, είναι απαραίτητη η ανάλυση της σύνθεσης της κυκλοφορίας. Οχήματα διαφορετικών κατηγοριών συμπεριφέρονται με τις αντίστοιχες ιδιομορφίες στο σύνολο της κυκλοφορίας, δηλαδή μπορεί να χρησιμοποιούν διαφορετικά εύρη ταχυτήτων, να απαιτούν περισσότερο/λιγότερο χρόνο για ευέλικτους ελιγμούς, και φυσικά να καταλαμβάνουν περισσότερο/λιγότερο χώρο στο οδόστρωμα. Για αυτό το λόγο, στη λειτουργική ανάλυση κατά HCM 2010 χρησιμοποιούνται **Μονάδες Επιβατικών Αυτοκινήτων, MEA (Passenger Car Units, PCU)**, επιτυγχάνοντας την αναγωγή διαφορετικών οχημάτων σε συγκρίσιμες οντότητες. Οι σχετικές αντιστοιχίες εμφανίζονται στον παρακάτω πίνακα του ΥΠΕΧΩΔΕ:

Πίνακας 1: Ισοδυναμία κατηγοριών οχημάτων σε Μονάδες Επιβατικών Αυτοκινήτων (MEA). (πηγή: ΥΠΕΧΩΔΕ)

Κατηγορίες Οχημάτων		MEA
Κατηγορία I:	σκούτερ, μοτοποδήλατα, μοτοσυκλέτες	0.5
Κατηγορία II:	επιβατικά ΙΧ, ταξί, ημιφορτηγά, μικρά φορτηγά	1.0
Κατηγορία III:	επιβατικά με τροχόσπιτα, μικρο-λεωφορεία, φορτηγά, λεωφορεία και πούλμαν	2.0
Κατηγορία IV:	φορτηγά επικαθήμενα, φορτηγά ειδικού φορτίου, τρόλλεϋ, αρθρωτά λεωφορεία	3.0
Κατηγορία V:	φορτηγά με ρυμουλκούμενο	4.0
Κατηγορία VI:	συρμός φορτηγών	5.0
Κατηγορία VII:	άλλα οχήματα (τρακτέρ, κατασκευαστικά οχήματα κ.ά.)	-

2.2.3.1: Λειτουργική ανάλυση ισόπεδου κόμβου φωτεινής σημάτωσης (HCM 2010)

Σύμφωνα με τη μεθοδολογία του HCM 2010 [24], βασικό μέγεθος για το χαρακτηρισμό της κυκλοφοριακής ικανότητας ισόπεδων κόμβων φωτεινής σηματοδότησης, αποτελεί η **ροή κορεσμού (saturation flow)**. Πρόκειται για το μέγιστο ωριαίο ρυθμό ροής οχημάτων που μπορεί να υποστηρίξει ένα οδικό σύστημα (στην προκειμένη περίπτωση ένα ρεύμα, δηλαδή λωρίδες κυκλοφορίας που ελέγχονται από την ίδια ένδειξη φωτεινού σηματοδότη) υπό τις επικρατούσες οδικές/κυκλοφοριακές συνθήκες, αλλά με την παραδοχή 100% χρησιμοποιημένου χρόνου πράσινης ένδειξης. Ο υπολογισμός της απαιτεί τον προσδιορισμό πολλαπλών κυκλοφοριακών συντελεστών προσαρμογής, και δίνεται από την παρακάτω εξίσωση:

$$S = S_o \cdot N \cdot f_w \cdot f_{HV} \cdot f_g \cdot f_p \cdot f_{bb} \cdot f_a \cdot f_{RT} \cdot f_{LT} \left[\frac{MEA}{\omega\rho\alpha} \right]$$

όπου $S_o = 1900$ MEA/ώρα πρασίνου, $N = 0$ αριθμός λωρίδων

Ειδικά για κάθε συντελεστή προσαρμογής f :

- **Συντελεστής προσαρμογής λόγω πλάτους λωρίδας (f_w):**

$$f_w = 1 + \frac{w - 3.60}{900}$$

όπου $w \leq 4.80$ το πλάτος λωρίδας (αν $w > 4.80$ γίνεται ανάλυση σε 2 λωρίδες)

- Συντελεστής προσαρμογής λόγω βαρέων οχημάτων (f_{HV}):

$$f_{HV} = \frac{100}{100 + \%_{HV}}$$

όπου $\%_{HV}$ = το ποσοστό βαρέων οχημάτων

- Συντελεστής προσαρμογής κλίσης (f_g):

$$f_g = 1 - \frac{\%_G}{200}$$

όπου $-6\% \leq \%_G \leq +6\%$ το ποσοστό κλίσης (εκτός ορίου ισχύουν οι οριακές τιμές)

- Συντελεστής προσαρμογής λόγω στάθμευσης οχημάτων (f_p):

$$f_p = \begin{cases} 1 - \frac{0.1 + \frac{18}{3600} \cdot N_p}{N}, & \text{αν επιτρέπεται η στάθμευση} \\ 1, & \text{αν δεν επιτρέπεται η στάθμευση} \end{cases}$$

όπου N_p = ο αριθμός αναμενόμενων ελιγμών για στάθμευση (ανά ώρα), N = ο αριθμός λωρίδων

- Συντελεστής προσαρμογής τύπου περιοχής (f_a):

$$f_a = \begin{cases} 0.90, & \text{για κεντρικές εμπορικές περιοχές} \\ 1.00, & \text{για άλλες περιοχές} \end{cases}$$

- Συντελεστής προσαρμογής δεξιών στρεφουσών κινήσεων (f_{RT}):

- Αποκλειστική λωρίδα δεξιάς στροφής – προστατευόμενη φάση:

$$f_{RT} = 0.85$$

Σημειώνεται ότι κατά την προστατευόμενη φάση δεν επιτρέπεται η ταυτόχρονη διέλευση κυκλοφοριακών ρευμάτων που διασταυρώνονται (εδώ διασταύρωση οδικού ρεύματος με ρεύμα πεζών). Αντίθετα, κάτι τέτοιο επιτρέπεται κατά την επιτρεπόμενη φάση.

- Αποκλειστική λωρίδα δεξιάς στροφής – επιτρεπόμενη φάση:

$$f_{RT} = 0.85 - \left(\frac{ped}{2100}\right)$$

όπου $0 \leq ped \leq 1700$ οι διερχόμενοι πεζοί (ανά ώρα)

- Αποκλειστική λωρίδα δεξιάς στροφής – προστατευόμενη και επιτρεπόμενη φάση:

$$f_{RT} = 0.85 - \left(\frac{ped}{2100}\right) \cdot (1 - P_{RTA})$$

όπου P_{RTA} = η αναλογία δεξιών στροφών που χρησιμοποιεί την προστατευόμενη

- **Μοιραζόμενη λωρίδα δεξιάς στροφής – προστατευόμενη φάση:**

$$f_{RT} = 1.00 - 0.15 \cdot P_{RT}$$

όπου P_{RT} = η αναλογία δεξιών στροφών στη λωρίδα

- **Μοιραζόμενη λωρίδα δεξιάς στροφής – επιτρεπόμενη φάση:**

$$f_{RT} = 1.00 - P_{RT} \cdot \left[0.15 + \left(\frac{ped}{2100} \right) \right]$$

- **Μοιραζόμενη λωρίδα δεξιάς στροφής – προστατευόμενη και επιτρεπόμενη φάση:**

$$f_{RT} = 1.00 - P_{RT} \cdot \left[0.15 + \left(\frac{ped}{2100} \right) \right] \cdot (1 - P_{RTA})$$

- **Πρόσβαση μιας λωρίδας:**

$$f_{RT} = 1.00 - P_{RT} \cdot \left[0.135 + \left(\frac{ped}{2100} \right) \right]$$

- **Συντελεστής προσαρμογής αριστερών στρεφουσών κινήσεων (f_{LT}):**

- **Αποκλειστική λωρίδα αριστερής στροφής – προστατευόμενη φάση:**

$$f_{LT} = 0.95$$

- **Αποκλειστική λωρίδα αριστερής στροφής – επιτρεπόμενη φάση:**

Απαιτεί ειδικό υπολογισμό, διότι συνήθως δεν ενδείκνυται.

- **Αποκλειστική λωρίδα αριστερής στροφής – προστατευόμενη και επιτρεπόμενη φάση:**

Απαιτεί ειδικό υπολογισμό, διότι συνήθως δεν ενδείκνυται.

- **Μοιραζόμενη λωρίδα αριστερής στροφής – προστατευόμενη φάση:**

$$f_{LT} = \frac{1.00}{1.00 + 0.05 \cdot P_{LT}}$$

όπου P_{LT} = η αναλογία αριστερών στροφών στη λωρίδα

- **Μοιραζόμενη λωρίδα αριστερής στροφής – επιτρεπόμενη φάση:**

Απαιτεί ειδικό υπολογισμό, διότι συνήθως δεν ενδείκνυται.

- **Μοιραζόμενη λωρίδα αριστερής στροφής – προστατευόμενη και επιτρεπόμενη φάση:**

$$f_{LT} = \begin{cases} \frac{1400 - V_o}{[(1400 - V_o) + (235 + 0.435 \cdot V_o) \cdot P_{LT}]}, & \text{για } V_o \leq 1220 \frac{MEA}{\text{ώρα}} \text{ (φόρτος τεμνόμενης κατεύθυνσης)} \\ \frac{1}{1 + 4.525 \cdot P_{LT}}, & \text{για } V_o > 1220 \frac{MEA}{\text{ώρα}} \end{cases}$$

Δεδομένου ότι έχει πραγματοποιηθεί ο καθορισμός των σταδίων (κινήσεις ανά εναλλαγή ενδείξεων) σε ένα πρόγραμμα φωτεινής σηματοδότησης ενός ισόπεδου κόμβου, οι διάρκειες ενεργών πράσινων ενδείξεων των ρευμάτων, αλλά και η συνολική διάρκεια του κύκλου του, η κυκλοφοριακή ικανότητα δίνεται από τη σχέση:

$$c_i = s_i \cdot \frac{g_i}{C}$$

όπου s_i = η ροή κορεσμού στο ρεύμα i , g_i = η διάρκεια πράσινων ενδείξεων στο ρεύμα i κατά ένα κύκλο συνολικής διάρκειας C

Για τον χαρακτηρισμό του επιπέδου εξυπηρέτησης απαιτείται ο προσδιορισμός της μέσης καθυστέρησης, δηλαδή του αριθμητικού μέσου των διαφορών μεταξύ των πραγματικών χρόνων διάνυσης ενός οδικού τμήματος, με τους χρόνους διάνυσης του ίδιου οδικού τμήματος εάν εξέλειπε ο λόγος που εμποδίζει τις απρόσκοπτες κινήσεις των οχημάτων. Για ένα ρεύμα υπολογίζεται ως εξής:

$$d = d_1 \cdot DF + d_2 \left[\frac{sec}{\text{όχημα}} \right]$$

$$\text{με } d_1 = \frac{0.38 \cdot C \cdot \left(1 - \frac{g}{C}\right)^2}{1 - X_{\min} \cdot \frac{g}{C}} \text{ και } d_2 = 173 \cdot X^2 \cdot \left[(X - 1) + \sqrt{(X - 1)^2 + m \cdot \left(\frac{X}{C}\right)} \right]$$

όπου $X = \frac{V}{c}$ ο βαθμός κορεσμού (το πηλίκο κυκλοφοριακού φόρτου ζήτησης (**traffic demand**) με την κυκλοφοριακή ικανότητα), $X_{\min} = \min\{X, 1\}$, $m =$ ο συντελεστής προσαρμογής λόγω τυχαίων αφίξεων, $DF =$ ο συντελεστής προσαρμογής λόγω συντονισμού η τύπου κυκλοφοριακού ελέγχου

Οι συντελεστές m και DF επιλέγονται σύμφωνα με όσα ορίζουν οι παρακάτω πίνακες:

Πίνακας 2: Τιμές συντελεστή m ανάλογα με τον ποιοτικό χαρακτηρισμό των αφίξεων οχημάτων στην κυκλοφορία. (πηγή: [24])

	Τύπος αφίξεων	m
1	άφιξη 80% του φόρτου υπό μορφή φάλαγγας στην έναρξη της κόκκινης ένδειξης	8
2	άφιξη 40% - 80% του φόρτου σε πυκνή φάλαγγα στο μέσο της κόκκινης ένδειξης ή μικρότερες διάσπαρτες φάλαγγες καθ' όλη την διάρκεια της κόκκινης ένδειξης	12
3	πλήρως τυχαίες αφίξεις, ώστε η κύρια φάλαγγα περιλαμβάνει λιγότερο από 40% του φόρτου	16
4	άφιξη 40% - 80% του φόρτου σε πυκνή φάλαγγα στο μέσο της πράσινης ένδειξης ή μικρότερες διάσπαρτες φάλαγγες καθ' όλη την διάρκεια της πράσινης ένδειξης	12
5	άφιξη 80% του φόρτου υπό μορφή φάλαγγας στην έναρξη της πράσινης ένδειξης	8
6	άφιξη όλων των οχημάτων με οργανωμένο τρόπο κατά την διάρκεια της πράσινης ένδειξης	4

Πίνακας 3: Στοιχεία υπολογισμού συντελεστή DF. (πηγή: [24])

Τύπος σηματοδότησης	Μη συντονισμένος κόμβος	Συντονισμένος κόμβος
Σταθερού χρόνου	1.00	$(1 - R_p g / C) f_p / (1 - gC)$
Επενεργούμενη (χωρίς βελτιστοποίηση)		
Επενεργούμενη ομάδα λωρίδων	0.85	1.0
Μη-επενεργούμενη ομάδα λωρίδων	0.85	$(1 - R_p g / C) f_p / (1 - gC)$
Επενεργούμενη (με βελτιστοποίηση)	0.85	1.0

	Τύπος αφίξεων	f_p	R_p
1	άφιξη 80% του φόρτου υπό μορφή φάλαγγας στην έναρξη της κόκκινης ένδειξης	1.00	0.333
2	άφιξη 40% - 80% του φόρτου σε πυκνή φάλαγγα στο μέσο της κόκκινης ένδειξης ή μικρότερες διάσπαρτες φάλαγγες καθ' όλη την διάρκεια της κόκκινης ένδειξης	0.93	0.667
3	Πλήρως τυχαίες αφίξεις, ώστε η κύρια φάλαγγα περιλαμβάνει λιγότερο από 40% του φόρτου	1.00	1.000
4	άφιξη 40% - 80% του φόρτου σε πυκνή φάλαγγα στο μέσο της πράσινης ένδειξης ή μικρότερες διάσπαρτες φάλαγγες καθ' όλη την διάρκεια της πράσινης ένδειξης	1.15	1.333
5	άφιξη 80% του φόρτου υπό μορφή φάλαγγας στην έναρξη της πράσινης ένδειξης	1.00	1.667
6	άφιξη όλων των οχημάτων με οργανωμένο τρόπο κατά την διάρκεια της πράσινης ένδειξης	1.00	2.000

Μπορεί να γίνει η αναγωγή της μέσης καθυστέρησης σε επίπεδο μιας πρόσβασης k ως εξής:

$$d_k = \frac{\sum_j d_j \cdot V_j}{\sum_j V_j}$$

όπου d_j = η μέση καθυστέρηση του ρεύματος j , V_j = ο κυκλοφοριακός φόρτος ζήτησης του ρεύματος j

Τελικά πραγματοποιείται η αναγωγή της μέσης καθυστέρησης σε επίπεδο κόμβου:

$$d_s = \frac{\sum_k d_k \cdot V_k}{\sum_k V_k}$$

όπου d_j = η μέση καθυστέρηση της πρόσβασης k , V_j = ο κυκλοφοριακός φόρτος ζήτησης της πρόσβασης k

Οι αντιστοιχίες μεταξύ μέσης καθυστέρησης και κατάταξης σε στάθμη εξυπηρέτησης παρατίθενται στον παρακάτω πίνακα:

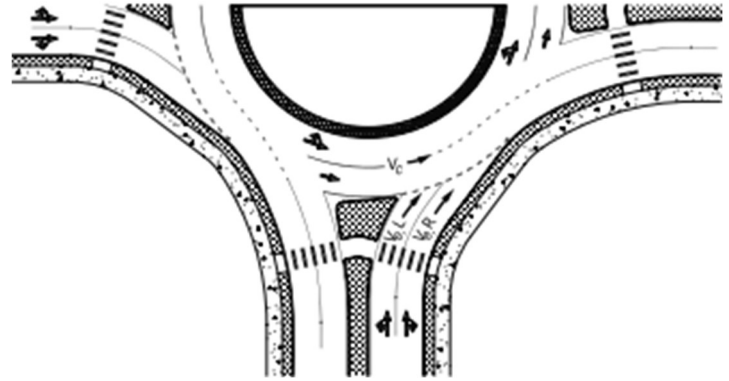
Πίνακας 4: Αντιστοιχία στάθμης εξυπηρέτησης με το διάστημα μέσης αναμονής ανά όχημα σε ισόπεδο κόμβο φωτεινής σηματοδότησης. (πηγή: [24])

Στάθμη εξυπηρέτησης	μέση αναμονή ανά όχημα (δλ)
A	$d \leq 5.0$
B	$5.0 < d \leq 15.0$
C	$15.0 < d \leq 25.0$
D	$25.0 < d \leq 40.0$
E	$40.0 < d \leq 60.0$
F	$60.0 < d$

Αναλυτικότερα, σε επίπεδο κόμβου, η κατάταξή του σε στάθμη A σημαίνει ότι η πλειοψηφία των οχημάτων διέρχονται από τον κόμβο με ύπαρξη πράσινης ένδειξης, χωρίς να πραγματοποιήσουν στάσεις. Στη στάθμη B τα περισσότερα οχήματα συνεχίζουν να κινούνται απρόσκοπτα, αν και αυξάνεται το πλήθος εκείνων που αναγκάζονται σε στάση. Από τη στάθμη Γ παρατηρούνται περιπτώσεις οχημάτων που χρειάζονται παραπάνω από έναν κύκλο για να διέλθουν του κόμβου («αποτυχημένος» κύκλος), αλλά παρ' όλα αυτά αρκετά οχήματα περνούν χωρίς στάσεις. Ο αριθμός των οχημάτων που διέρχονται ελεύθερα υπό πράσινη ένδειξη μειώνεται περαιτέρω στη στάθμη Δ. Στη στάθμη E οι «αποτυχημένοι» κύκλοι θεωρούνται ιδιαίτερα συχνοί, ενώ ο κυκλοφοριακός φόρτος προσεγγίζει την κυκλοφοριακή ικανότητα. Τέλος, στη στάθμη F υπάρχουν, πλέον, συνθήκες συμφόρησης, όπου στην πλειοψηφία των κύκλων οι σχηματισμένες ουρές αναμονής δεν «αδειάζουν» κατά την πράσινη ένδειξη.

2.2.3.2: Λειτουργική ανάλυση κυκλικού κόμβου (HCM 2010)

Η διαδικασία προσδιορισμού του επιπέδου εξυπηρέτησης ενός κυκλικού κόμβου αποτελεί ένα ζήτημα αλληλεπίδρασης των κυκλοφοριακών φόρτων ζήτησης που αφικνούνται σε κάθε πρόσβαση (entry flow), με τον εμπλεκόμενο φόρτο που κατά παραδοχή βρίσκεται σε προτεραιότητα εντός του δακτυλίου (circulating flow).



Σχήμα 18: Σχεδιάγραμμα χωροθέτησης αφικνούμενου φόρτου (entry flow), και εμπλεκόμενου φόρτου (circulating flow) σε έναν κυκλικό κόμβο 2 λωρίδων εντός του δακτυλίου, και 2 λωρίδων επί μιας πρόσβασης. (πηγή: [24])

Θεωρώντας πως έχει πραγματοποιηθεί αναγωγή της σύνθεσης της κυκλοφορίας σε

ΜΕΑ, δεν υπεισέρχεται στους υπολογισμούς ο συντελεστής προσαρμογής λόγω βαρέων οχημάτων f_{HV} , όπως ορίζει η σχετική βιβλιογραφία. Επομένως, αρχικά υπολογίζεται σε επίπεδο πρόσβασης ένας συντελεστής εμπόδισης λόγω κυκλοφορίας πεζών, σύμφωνα με το παρακάτω τυπολόγιο:

$$f_{ped} = \begin{cases} 1 - 0.0001397 \cdot n_{ped}, & \text{για } v_{i,pce} \leq 881 \text{ και } n_{ped} \leq 101 \\ \frac{1119.5 - 0.715 \cdot v_{c,pce} - 0.644 \cdot n_{ped} + 0.00073 \cdot v_{c,pce} \cdot n_{ped}}{1068.6 - 0.654 \cdot v_{c,pce}}, & \text{για } v_{i,pce} \leq 881 \text{ και } n_{ped} > 101 \\ 1, & \text{για } v_{i,pce} > 881 \text{ όταν και οι πεζοί διέρχονται μεταξύ ακινητοποιημένων σε ουρά οχημάτων} \end{cases}$$

όπου n_{ped} = το πλήθος πεζών που διέρχονται μπροστά από την πρόσβαση, $v_{i,pce}$ = ο κυκλοφοριακός φόρτος ζήτησης που αφικνείται στην πρόσβαση [ΜΕΑ/ώρα], $v_{c,pce}$ = ο εμπλεκόμενος έναντι της πρόσβασης κυκλοφοριακός φόρτος [ΜΕΑ/ώρα].

Ακολουθεί ο υπολογισμός της κυκλοφοριακής ικανότητας ανά πρόσβαση (ή λωρίδα πρόσβασης), για τον οποίο διακρίνονται διαφορετικοί τρόποι υπολογισμού, βάσει τον λωρίδων εντός κι εκτός του δακτυλίου:

- Για κυκλικούς κόμβους **μιας λωρίδας κυκλοφορίας, εντός κι εκτός του δακτυλίου:**

$$c_l = 1130 \cdot f_{ped} \cdot e^{(-1.0 \cdot 10^{-3}) \cdot v_{c,pce}}$$

- Για κυκλικούς κόμβους **μιας λωρίδας κυκλοφορίας εντός του δακτυλίου, και δύο στην πρόσβαση:**

$$c_l = 1130 \cdot f_{ped} \cdot e^{(-1.0 \cdot 10^{-3}) \cdot v_{c,pce}}$$

- Για κυκλικούς κόμβους **δύο λωρίδων κυκλοφορίας εντός του δακτυλίου, και μιας στην πρόσβαση:**

$$c_l = 1130 \cdot f_{ped} \cdot e^{(-0.7 \cdot 10^{-3}) \cdot v_{c,pce}}$$

- Για κυκλικούς κόμβους δύο λωρίδων κυκλοφορίας, εντός κι εκτός του δακτυλίου:

$$c_{l,R} = 1130 \cdot f_{ped} \cdot e^{(-0.7 \cdot 10^{-3}) \cdot v_{c,pce}} \text{ (δεξιά λωρίδα)}$$

$$c_{l,L} = 1130 \cdot f_{ped} \cdot e^{(-0.75 \cdot 10^{-3}) \cdot v_{c,pce}} \text{ (αριστερή λωρίδα)}$$

Στη συνέχεια υπολογίζεται η μέση καθυστέρηση ανά πρόσβαση (ή λωρίδα πρόσβασης) ως εξής:

$$d = \frac{3600}{c_{m,x}} + 900 \cdot T \cdot \left[\frac{V_x}{c_{m,x}} - 1 + \sqrt{\left(\frac{V_x}{c_{m,x}} - 1\right)^2 + \frac{\left(\frac{3600}{c_{m,x}}\right) \cdot \left(\frac{V_x}{c_{m,x}}\right)}{450 \cdot T}} \right] + 5$$

όπου $c_{m,x}$ = η κυκλοφοριακή ικανότητα της πρόσβασης x (ή μιας λωρίδας πρόσβασης x), V_x = ο κυκλοφοριακός φόρτος ζήτησης που αφικνείται στην πρόσβαση/λωρίδα x, T = η χρονική περίοδος ανάλυσης σε ώρες.

Τελικά απομένει η αναγωγή των επιμέρους μέσων καθυστερήσεων σε επίπεδο ολόκληρου του κυκλικού κόμβου:

$$d_I = \frac{\sum_x d_x \cdot V_x}{\sum_x V_x}$$

όπου d_x = η μέση καθυστέρηση στην πρόσβαση x (ή σε μια λωρίδα πρόσβασης x).

Σημειώνεται πως οι ερμηνείες των επιμέρους σταθμών εξυπηρέτησης κυκλικού κόμβου δεν διαφέρουν σηματοδοτικά από τις αντίστοιχες των ισόπεδων κόμβων φωτεινής σηματοδότησης, αν εξαιρεθεί, βέβαια, το γεγονός ότι πλέον τα οχήματα καθυστερούν λόγω εμπλεκόμενου δυναμικού φόρτου στις διελεύσεις, και όχι λόγω των ενδείξεων ενός «στατικού» προγράμματος παραχώρησης προτεραιότητας. Ωστόσο, τα διαστήματα τιμών διαφέρουν ελαφρώς, όπως φαίνεται στο διπλανό σχετικό πίνακα:

Πίνακας 5: Αντιστοιχία στάθμης εξυπηρέτησης με το διάστημα μέσης αναμονής ανά όχημα σε κυκλικό κόμβο. (πηγή: [24])

Στάθμη εξυπηρέτησης	μέση αναμονή ανά όχημα (δλ)
A	$d \leq 10$
B	$10 < d \leq 15$
C	$15 < d \leq 25$
D	$25 < d \leq 35$
E	$35 < d \leq 50$
F	$50 < d$

2.3: Σχετικές έρευνες μελέτης κυκλοφορίας με χρήση UAV και DL

Η εισαγωγή των UAVs στην παρακολούθηση της κυκλοφορίας αποτελεί σχετικά πρόσφατο φαινόμενο στην ερευνητική κοινότητα, αλλά προηγείται περίπου κατά μια δεκαετία της ανάπτυξης των σύγχρονων καινοτομιών στον τομέα της Βαθιάς Μάθησης.

Μια από τις πρώτες γνωστές σχετικές δημοσιεύσεις πραγματοποιήθηκε από τους B. Coifman et al. (2006) [25], αναφερόμενοι γενικά στα πλεονεκτήματα των UAVs, και εξετάζοντας κινούμενα εναέρια πλάνα 2 ωρών από τον οδικό άξονα «Ohio State Route 315». Σκοπός ήταν η μελέτη του επιπέδου εξυπηρέτησής του (βάσει πυκνότητας), της μέσης ετήσιας ημερήσιας κυκλοφορίας, της λειτουργίας διασταυρώσεων, της μέτρησης ροών προέλευσης – προορισμού, και της εύρεσης θέσεων σε εγγύτερο χώρο στάθμευσης. Οι λήψεις διενεργήθηκαν με UAV τύπου BAT III με 2 ενσωματωμένες κάμερες, από ύψος πτήσης 150 m, ενώ κινούνταν με ταχύτητα 50 χλμ./ώρα. Πραγματοποιήθηκε η εξαγωγή κυκλοφοριακών στοιχείων (πυκνότητα, ταχύτητα, κυκλοφοριακός φόρτος) τόσο από μεμονωμένα καρέ, όσο και από μια σειρά ακολουθιών τους, ώστε να προκύψουν τα επιθυμητά τελικά προϊόντα της έρευνας, αξιοποιώντας την καταμέτρηση των απεικονιζόμενων οχημάτων σε ένα δεδομένο ορατό μήκος οδού. Ωστόσο οι μέθοδοί τους βασίζονταν αποκλειστικά στη γενίκευση πληροφορίας προερχόμενη από «χειροκίνητη» επιτήρηση πλάνων, και όχι σε κάποια αυτοματοποιημένη μεθοδολογία με τη συνδρομή της πληροφορικής.

Μετά από σχεδόν 10 χρόνια, ξεκινά η αξιοποίηση των μεθόδων υπολογιστικής νοημοσύνης, και δη της Βαθιάς Μάθησης, σε βίντεο προερχόμενα από UAV για την ανάλυση της κυκλοφορίας, όπως φάνηκε στην έρευνα των N. Kim και M. Chervonenkis (2015) [26]. Ο στόχος τους ήταν ο αυτοματοποιημένος εντοπισμός ιδιορρυθμιών στην οδική κυκλοφορία, βάσει των κριτηρίων Bayes και Neyman-Pearson, ώστε με τη λήψη των κατάλληλων αποφάσεων να επιτυγχάνεται βελτίωση της χωρητικότητας και της οδικής ασφάλειας. Οπότε για αυτό το σκοπό, εκτός από την ανάλυση των προϋπάρχοντων συστημάτων τεχνητής όρασης (AVS), προτάθηκε η χρήση κατάτμησης εικόνας σε περιοχές ενδιαφέροντος (Regions of Interest, RoI), ώστε η αναγνώριση-ανίχνευση των τυχόν αφύσικων καταστάσεων που επικρατούν στο συγκεκριμένο τμήμα να πράττεται με τη συνδρομή πολυεπίπεδου νευρωνικού δικτύου, χωρίς, βέβαια, να αναφέρεται η χρήση κάποιας συγκεκριμένης αρχιτεκτονικής.

Μεταβαίνοντας σε πιο πρόσφατες μελέτες, οι J. Wang et al. (2019) [27] διατύπωσαν μια ολοκληρωμένη μεθοδολογία παρακολούθησης πολλαπλών οχημάτων από μη κατακόρυφα πλάνα UAV, σχετικά αραιής κυκλοφορίας. Χρησιμοποιήθηκε μέθοδος tracking-by-detection, με μοντέλο ανίχνευσης YOLO v3 εκπαιδευμένο σε λήψεις από DJI Phantom 4 Pro ανάλυσης 2720x1530 pixels στα 30 fps, και με εφαρμογή CNN για εξαγωγή χαρακτηριστικών εμφάνισης

επαναυτοποίησης εκπαιδευμένο στο VeRi dataset. Σημειώθηκαν υψηλές επιδόσεις, έως και 88.7 % στην μετρική MOTA του δικτύου MOT Challenge. Η έρευνα, όμως, δεν πραγματοποίησε διάκριση των οχημάτων σε διαφορετικές κατηγορίες, ενώ μάλιστα δεν αναφέρθηκε στο πεδίο εφαρμογής κυκλοφοριακής τεχνικής. Πάντως, λόγω των ομοιοτήτων με τον αλγόριθμο DeepSORT και του παρόμοιου πεδίου εφαρμογής, τίθεται συμβατή για ποσοτική σύγκριση με τα αποτελέσματα της μεθοδολογίας Βαθιάς Μάθησης από την παρούσα διπλωματική εργασία.

Ακόμη, την τελευταία τριετία έχουν αναδειχθεί και 2 σπουδαίες «εγχώριες» επιστημονικές συνεισφορές:

- Στα πλαίσια του project μεγάλης κλίμακας «pNEUMA», οι Barmprounakis και Gerolimimis (2020) [28] χρησιμοποίησαν μια ομάδα αποτελούμενη από 10 UAVs για τη συλλογή κυκλοφοριακών πλάνων συνολικής διάρκειας 59 ωρών, σε μια έντονα συμφορημένη έκταση 1.3 km² εντός της Αθήνας, αποτελούμενη από 100 κόμβους και 100 χιλιόμετρα οδικού δικτύου. Με εφαρμογή τεχνικών ανίχνευσης και παρακολούθησης πολλαπλών αντικειμένων, αλλά και τη συνδρομή της δομής DataFromSky, καταγράφηκαν ροές κυκλοφορίας με 98.8% ακρίβεια εντοπισμού, εκτιμήθηκαν ταχύτητες με σφάλμα 2.97 χλμ./ώρα, ενώ σχηματίστηκε μακροσκοπικά πίνακας προέλευσης προορισμού και διαγράμματα κυκλοφοριακών μεγεθών.
- Οι R. Makrigiorgis et al. (2022) [29] στην έρευνά τους σχημάτισαν ένα σύνολο δεδομένων 9048 εικόνων, ανάλυσης από 1080p έως και 4K, προερχόμενο από λήψεις UAV σε περιοχές της Κύπρου. Απεικονίζονται περισσότερα από 250000 οχήματα, με διάκριση στις κατηγορίες «Car», «Bus», and «Truck». Το συγκεκριμένο dataset διατίθεται για την εκπαίδευση και τον έλεγχο μοντέλων ανίχνευσης. Σημειώνεται πως χρησιμοποιήθηκε ένα υποσύνολό του για τον έλεγχο ευρωστίας/γενίκευσης του παραχθέντος detector στην παρούσα διπλωματική εργασία, όπως θα παρατεθεί σε επόμενη ενότητα.

Κεφάλαιο 3^ο: Περιγραφή δεδομένων μελέτης

Η ανάπτυξη της μεθοδολογίας παρακολούθησης οχημάτων επιλέχθηκε να εφαρμοστεί στο Suncity Dataset, ένα σύνολο ακολουθιών βίντεο προερχόμενων από ενσωματωμένες κάμερες σε μη επανδρωμένα ιπτάμενα οχήματα (UAVs / drones). Τα συγκεκριμένα δεδομένα μελετήθηκαν στο παρελθόν στη διπλωματική εργασία του συναδέλφου Μιχάλα Μιχάλη (2022) [30], με περισσότερη εμβάθυνση στις δυνατότητες των UAVs στον έλεγχο της κυκλοφορίας. Το παρόν project διαφοροποιείται, καθώς δίνεται έμφαση στην λειτουργία του μοντέλου παρακολούθησης πολλαπλών αντικειμένων και στην επίδοσή του στις μετρικές του MOT Benchmark, καθώς και στην μετα-επεξεργασία των δεδομένων εξόδου του στην αξιολόγηση των οδικών κόμβων.

3.1: Τεχνικά χαρακτηριστικά

Το Suncity Dataset, όπως άλλωστε μαρτυρά η ονομασία του, περιέχει πλάνα από τη συνοικία της Ηλιούπολης στην Αθήνα, τα οποία αφορούν 3 διαφορετικούς κυκλοφοριακούς κόμβους της. Πρόκειται για τον κυκλικό κόμβο της Πλατείας 25^{ης} Μαρτίου, και τους ισόπεδους κόμβους φωτεινής σηματοδότησης στις συμβολές των οδών Ηρώων Πολυτεχνείου – Παύλου Πεντάρη και των λεωφόρων Πρωτόπαπα – Σοφοκλή Βενιζέλου.

Το πλήρες αρχείο δεδομένων περιλαμβάνει τις καταγραφές που παρατίθενται στον παρακάτω πίνακα:

Πίνακας 6: Συγκεντρωτική παράθεση των τεχνικών χαρακτηριστικών του Suncity Dataset, ανά αρχείο βίντεο.

Βίντεο	Τοποθεσία	Ύψος λήψης	Πεδίο όρασης (FOV)	Ανάλυση	Μέγεθος	Διάρκεια	UAV μοντέλο	FPS
DJI_0073.mp4	κυκλικός κόμβος	163 m	84°	3840	3.50 GB	05'00"	Mavic Air 2	25
	Πλατείας 25 ^{ης} Μαρτίου (8 π.μ.)			2160				
DJI_0074.mp4	κυκλικός κόμβος	163 m	84°	3840	3.50 GB	05'00"	Mavic Air 2	25
	Πλατείας 25 ^{ης} Μαρτίου (8 π.μ.)			2160				
DJI_0075.mp4	κυκλικός κόμβος	163 m	84°	3840	1.59 GB	02'17"	Mavic Air 2	25
	Πλατείας 25 ^{ης} Μαρτίου (8 π.μ.)			2160				
DJI_0076.mp4	κυκλικός κόμβος	163 m	84°	3840	82.9 MB	00'06"	Mavic Air 2	25
	Πλατείας 25 ^{ης} Μαρτίου (8 π.μ.)			2160				

	κυκλικός κόμβος			3840				
DJI_0077.mp4	Πλατείας 25 ^{ης} Μαρτίου (8 π.μ.)	163 m	84°	x	3.50 GB	05'00"	Mavic Air 2	25
				2160				
DJI_0078.mp4	Πλατείας 25 ^{ης} Μαρτίου (8 π.μ.)	163 m	84°	x	2.17 GB	03'07"	Mavic Air 2	25
				2160				
DJI_0079.mp4	Πλατείας 25 ^{ης} Μαρτίου (8 π.μ.)	163 m	84°	x	23.6 MB	00'01"	Mavic Air 2	25
				2160				
DJI_0095.mp4	Ηρώων Πολυτεχνείου – Παύλου Πεντάρη (11 π.μ.)	146 m – 156 m	84°	x	3.50 GB	14'18"	Mavic Air 2	25
				1920 1080				
DJI_0096.mp4	Ηρώων Πολυτεχνείου – Παύλου Πεντάρη (11 π.μ.)	146 m – 156 m	84°	x	1.44 GB	05'54"	Mavic Air 2	25
				1920 1080				
DJI_0107.mp4	Ηρώων Πολυτεχνείου – Παύλου Πεντάρη (8 π.μ.)	146 m – 156 m	84°	x	3.50 GB	14'18"	Mavic Air 2	25
				1920 1080				
DJI_0108.mp4	Ηρώων Πολυτεχνείου – Παύλου Πεντάρη (8 π.μ.)	146 m – 156 m	84°	x	1.40 GB	05'44"	Mavic Air 2	25
				1920 1080				
DJI_0560.mov	Πρωτόπαπα – Σοφοκλή Βενιζέλου (8 π.μ.)	200 m – 205 m	72°	x	3.50 GB	14'19"	Mavic 2 Pro	25
				1920 1080				
DJI_0561.mov	Πρωτόπαπα – Σοφοκλή Βενιζέλου (8 π.μ.)	200 m – 205 m	72°	x	177 MB	00'42"	Mavic 2 Pro	25
				1920 1080				
DJI_0562.mov	Πρωτόπαπα – Σοφοκλή Βενιζέλου (11 π.μ.)	200 m – 205 m	72°	x	88.4 MB	00'21"	Mavic 2 Pro	25
				1920 1080				
DJI_0563.mov	Πρωτόπαπα – Σοφοκλή Βενιζέλου (11 π.μ.)	200 m – 205 m	72°	x	3.50 GB	14'19"	Mavic 2 Pro	25
				1920 1080				
DJI_0564.mov	Πρωτόπαπα – Σοφοκλή Βενιζέλου (11 π.μ.)	200 m – 205 m	72°	x	192 MB	00'46"	Mavic 2 Pro	25
				1920 1080				

	κυκλικός κόμβος								
DJI_0909.mp4	Πλατείας 25 ^{ης} Μαρτίου (8 π.μ.)	163 m	84°	x	3.50 GB	05'00"	Mavic Air 2	25	
	κυκλικός κόμβος								
DJI_0910.mp4	Πλατείας 25 ^{ης} Μαρτίου (8 π.μ.)	163 m	84°	x	127 MB	00'10"	Mavic Air 2	25	
	κυκλικός κόμβος								
DJI_0911.mp4	Πλατείας 25 ^{ης} Μαρτίου (11 π.μ.)	163 m	84°	x	3.50 GB	05'00"	Mavic Air 2	25	
	κυκλικός κόμβος								
DJI_0912.mp4	Πλατείας 25 ^{ης} Μαρτίου (11 π.μ.)	163 m	84°	x	691 MB	00'58"	Mavic Air 2	25	

Σημειώνεται ότι τα βίντεο DJI_0076.mp4, DJI_0079.mp4, DJI_0910.mp4 και DJI_0562.mον (υπογραμμισμένα με κόκκινο στον πίνακα), λόγω της μικρής διάρκειάς τους, δε χρησιμοποιήθηκαν σε καμία εκ των μεταγενέστερων εφαρμογών. Επίσης, κατά τη διάρκεια της ακολουθίας DJI_0095.mp4 (υπογραμμισμένη με κίτρινο στον πίνακα) μεταβλήθηκε απότομα κι αισθητά ο προσανατολισμός και το ύψος πτήσης του drone, οπότε πραγματοποιήθηκε διαχωρισμός της σε DJI_0095_1.mp4 και DJI_0095_2.mp4, αφαιρώντας παράλληλα το σύντομο χρονικό διάστημα προσαρμογής.

3.2: Ποιοτικά χαρακτηριστικά, προκλήσεις και ιδιομορφίες

Στο Suncity Dataset απεικονίζεται μεγάλο πλήθος οχημάτων ανά καρτέ (κάθε κατηγορίας), τόσο κινούμενων, όσο και σταθμευμένων, και μάλιστα σε μια ποικιλία συνδυασμών σκηνικού. Αυτό επιτεύχθηκε προγραμματίζοντας τις πτήσεις των UAVs στις 3 επιλεγμένες τοποθεσίες επιπρόσθετα και για διαφορετικές χρονικές περιόδους, 8 π.μ. και 11 π.μ., οι οποίες αποτελούσαν ώρες αυξημένων διελεύσεων, εφόσον αφορούσαν εργάσιμη ημέρα. Στο παρακάτω σχήμα τοποθετούνται στιγμιότυπα στα οποία φαίνονται οι 6 ουσιαστικές διαφοροποιήσεις φωτισμού – τοποθεσίας:



Σχήμα 19: Δείγματα μερικών καρτέ από το Suncity Dataset, καλύπτοντας τις 6 διαφορετικές εναλλακτικές μεταξύ φωτισμού/ώρας λήψης και τοποθεσίας. Η αριστερή στήλη περιλαμβάνει τις λήψεις των 8 π.μ., ενώ η δεξιά των 11 π.μ.. Στην πρώτη σειρά απεικονίζεται ο κόμβος Ηρώων Πολυτεχνείου – Παύλου Πεντάρη, στη δεύτερη ο κόμβος Πρωτόπαπα – Σοφοκλή Βενιζέλου, ενώ στην τρίτη ο κυκλικός κόμβος Πλατείας 25ης Μαρτίου.

Περαιτέρω, με μερικές πρόχειρες γραφικές μετρήσεις υπολογίστηκε πως ένα επιβατικό αυτοκίνητο απεικονιζόταν με μήκος ~70-80 εικονοστοιχείων στις 4K λήψεις, και ~30-40 εικονοστοιχείων στις 1080p, μεγέθη που αντιστοιχούν περίπου στο 2% της μεγάλης διάστασης των καρτέ. Έγινε αντιληπτό ότι στο Suncity Dataset τα οχήματα θα μπορούσαν να χαρακτηριστούν ως «μικροσκοπικά αντικείμενα» (tiny objects). Η ανίχνευση, και κατ' επέκταση η παρακολούθηση αντικειμένων μικροσκοπικής κλίμακας απαιτούσε την αντιμετώπιση μιας σειράς προκλήσεων:

- Η διαθέσιμη χωρική ανάλυση (spatial resolution) ενδέχεται να αδυνατεί να αναπαραγάγει αρκετές από τις οπτικές λεπτομέρειες του αντικειμένου. Με αυτό τον τρόπο δυσχεραίνεται η αναγνώριση/κατηγοριοποίησή του. Αλλά ταυτόχρονα, καθώς η ανάλυση βελτιώνεται, απαιτείται αισθητά περισσότερη υπολογιστική ισχύς για την ανάλυση «μεγαλύτερων» εικόνων.
- Η εκπαίδευση ενός μοντέλου στην ανίχνευση μικροσκοπικών αντικειμένων πιθανότατα δε θα φέρει αποτέλεσμα στην ανίχνευση των ίδιων αντικειμένων απεικονιζόμενων σε μεγαλύτερη -πλουσιότερη σε οπτικές λεπτομέρειες- κλίμακα, και αντιστρόφως.
- Τα όρια μικροσκοπικών αντικειμένων συνήθως εμφανίζονται «θολωμένα» και ασαφή, διότι το πλήθος των pixel των οποίων η δειγματοληψία τιμών φωτεινότητας επηρεάζεται από την εναλλαγή σκηνής στην εγγύτερη περιοχή του ορίου, συνιστά υψηλό ποσοστό επί των συνολικών pixel του αντικειμένου (μειωμένη αντίθεση αντικειμένου – υποβάθρου). Το γεγονός αυτό μειώνει την αξιοπιστία οριοθέτησης από ένα bounding box.
- Στις «bird's eye» (δηλ. κατακόρυφες) λήψεις, τα μικροσκοπικά αντικείμενα είναι περισσότερο επιρρεπή σε συγκαλύψεις (occlusions), είτε λόγω των περιοχών σκίασης, είτε λόγω παρεμβολής τρισδιάστατων ουδέτερων αντικειμένων, όπως π.χ. βλάστηση, κτήρια, στύλοι, κ.α. . Επιπρόσθετα, προς στα άκρα των καρέ και ειδικότερα για ευρύτερες γωνίες πεδίου όρασης, η ποσοστιαία μεταβολή της φαινομενικής κλίμακας ενός αντικειμένου είναι μεγαλύτερη.
- Ειδικά για το πρόβλημα της ανίχνευσης μικροσκοπικών αντικειμένων δημιουργούνται και σημαντικά αλγοριθμικά ζητήματα:
 - Στον Faster R-CNN detector, η επιλογή των κατάλληλων μεγεθών των anchors καθίσταται ιδιαίτερα δύσκολη: Με ορισμό μεγαλύτερων διαστάσεων τους, μειώνεται η πιθανότητα IoU – αντιστοίχισης με κάποιο ground truth bounding box (σαφέστατα μικρότερου μεγέθους), οπότε τα αντικείμενα συχνά δεν είναι «ορατά» από το μοντέλο. Αντίστοιχα, μικρά μεγέθη anchors αποδίδουν επιπλέον ασταθείς («exploding»/«vanishing») παραγώγους κατά τη βελτιστοποίηση.
 - Έστω ένα αντικείμενο διαστάσεων 30 x 30 pixels, για το οποίο η διαδικασία παλινδρόμησης ενός μοντέλου εξάγει ως βέλτιστη πρόβλεψη ένα bounding – box μετατοπισμένο κατά μόλις 2 pixels τόσο ως προς X, όσο και ως προς Y. Στην περίπτωση αυτή
$$IoU = \frac{28 \times 28}{32 \times 32 - 2 \times 2 \times 2} = 0.77$$
 Παρατηρείται πως για μικρά σφάλματα οριοθέτησης, η IoU δέχεται σημαντικές μειώσεις, οπότε και απαιτείται χαμηλότερο κατώφλι για τα θετικά δείγματα.

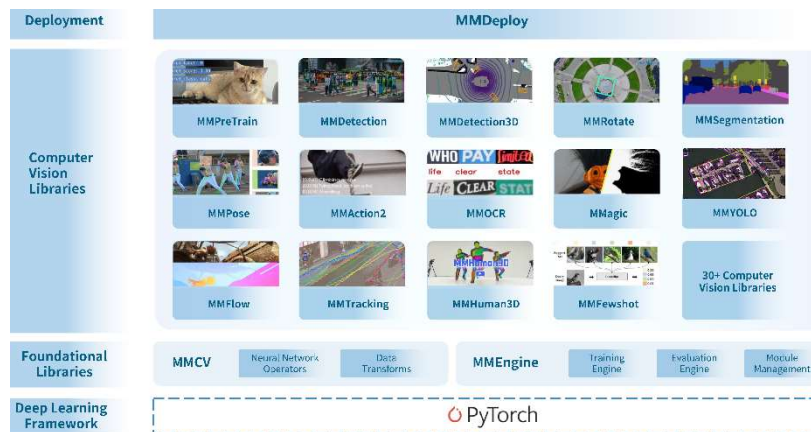
Ακόμη, με πιο προσεκτικό έλεγχο των βίντεο του dataset παρατηρήθηκε η αναμενόμενη ύπαρξη μικρο-μετακινήσεων στρέψης και μετατόπισης των λήψεων, λόγω της φυσιολογικής αδυναμίας επίτευξης απόλυτης σταθερότητας του UAV. Αυτή η συμπεριφορά κρίθηκε πως θα μπορούσε δυνητικά να βλάψει τη διαδικασία επαναταυτοποίησης οχημάτων, καταργώντας την ιδανική «1 προς 1» αντιστοιχία συντεταγμένων εικόνας και πραγματικής θέσης. Ως τετριμμένο παράδειγμα, ένα προσωρινά ακινητοποιημένο αυτοκίνητο, λόγω της σχετικής μετακίνησης της κάμερας μεταβάλλει τις συντεταγμένες εικόνας που καταλαμβάνει από καρέ σε καρέ, και άρα διαγράφει μια πλασματική τροχιά ανάλογη της εν λόγω σχετικής μετακίνησης. Για το λόγο αυτό, κάθε βίντεο δέχτηκε μερική σταθεροποίηση μέσω της κλάσης vidstab [31].

Κεφάλαιο 4^ο: Μεθοδολογία

Στο παρόν κεφάλαιο της διπλωματικής εργασίας παρατίθενται αναλυτικά τα βήματα ανάπτυξης μεθοδολογίας για την εφαρμογή του αλγόριθμου παρακολούθησης πολλαπλών αντικειμένων DeepSORT στα διερχόμενα οχήματα που απεικονίζονται στο Suncity Dataset. Αρχικά αναφέρονται τα μέσα υλικού και λογισμικού που χρησιμοποιήθηκαν για τη διενέργεια του αλγορίθμου, ενώ έπειτα περιγράφονται τα επιμέρους βήματά του (detection → association/reidentification → tracking). Τέλος, παρουσιάζονται αλγόριθμοι διαχείρισης και αξιοποίησης των δεδομένων εξόδου στο πεδίο της κυκλοφοριακής τεχνικής.

4.1: Διαθέσιμα υλικά μέσα και προγραμματιστικό πλαίσιο

Η εφαρμογή των τεχνικών Βαθιάς Μάθησης που συνιστούν το σύνολο της μεθοδολογίας πραγματοποιήθηκε με τη συνδρομή του προγραμματιστικού πλαισίου του οργανισμού ανοιχτού κώδικα OpenMMLab [32], ο οποίος έχει συνδράμει σημαντικά σε εργασίες της ερευνητικής κοινότητας, παρέχοντας τα εξής:



Σχήμα 20: Τα περιεχόμενα του οργανισμού ανοιχτού κώδικα OpenMMLab (πηγή: [32])

- Μια πληθώρα από προγραμματιστικές δομές (frameworks) εξειδικευμένες σε επιμέρους εφαρμογές Υπολογιστικής Όρασης, από τις οποίες χρησιμοποιούνται οι εξής:
 - **MMDetection** για τη φάση εκπαίδευσης μοντέλου ανίχνευσης.
 - **MMTracking**, η οποία «χτίζεται» πάνω σε υπόβαθρο MMDetection, για τη σύσταση του μοντέλου παρακολούθησης.

Κάθε δομή περιέχει υλοποιήσεις αλγορίθμων τελευταίας τεχνολογίας, οι οποίοι ρυθμίζονται εύκολα μέσω τροποποίησης αρχείων παραμετροποίησης (configuration ‘config’ files), ενώ υπάρχουν δυνατότητες διαλειτουργικότητας μεταξύ τους.

- Τις θεμελιώδεις βιβλιοθήκες:
 - **MMCV**: Διαθέτει ποικιλία τελεστών νευρωνικών δικτύων, καθώς και εργαλεία μετασχηματισμών δεδομένων.
 - **MMEEngine**: Περιλαμβάνει μηχανισμούς εκπαίδευσης, αξιολόγησης, και γενικότερης διαχείρισης μοντέλων.

Το σύνολο των προαναφερθέντων βασίζεται σε γλώσσα προγραμματισμού **Python**, και συγκεκριμένα στην κατεξοχήν βιβλιοθήκη Βαθιάς Μάθησης **PyTorch**.

Για την παρούσα μεθοδολογία χρησιμοποιήθηκαν οι συγκεκριμένες εκδόσεις προγραμματιστικών πακέτων:

- Για το σκέλος εκπαίδευσης του μοντέλου ανίχνευσης:
 - Python v. 3.8.15
 - PyTorch v. 1.12.0
 - CUDA toolkit v. 11.6.0
 - MMDetection v. 3.0.0
 - MMCV v. 2.0.0
 - MMEEngine v. 0.7.4
- Για το σκέλος επαναταυτοποίησης και παρακολούθησης:
 - Python v. 3.8.15
 - PyTorch v. 1.12.0
 - CUDA toolkit v. 11.6.0
 - MMDetection v. 2.25.0
 - MMTracking v. 0.14.0
 - MMCV v. 1.6.1
 - MMEEngine v. 0.7.4
- Σημειώνεται ότι, επειδή υπάρχει αναντιστοιχία μεταξύ των πιο πρόσφατων εκδόσεων MMDetection, MMTracking, καθώς και των υπόλοιπων εξαρτήσεων από βιβλιοθήκες, για την υλοποίηση της δομής παρακολούθησης χρησιμοποιήθηκε ενσωμάτωση checkpoint μοντέλου ανίχνευσης από το MMDetection 3.0.0, το οποίο είναι πλήρως συμβατό για χρήση από παλαιότερες εκδόσεις MMDetection (όπως η 2.25.0), πάνω στις οποίες μπορούσε να «χτιστεί» η δομή MMTracking λόγω συμβατότητας.

Όσον αφορά το διαθέσιμο υλικό μέρος, το σύνολο των διεργασιών εκτελέστηκε σε υπολογιστικό σύστημα λογισμικού Windows 11 με τα παρακάτω τεχνικά χαρακτηριστικά:

- CPU: Intel Core i5 – 13600KF 2.6GHz
- GPU: NVIDIA RTX 3070 (8 GB GDDR6 VRAM, 5888 CUDA cores)
- RAM: 2x16 GB Kingston Fury 5200 MHz DDR5 CL 40
- Motherboard: ASRock Z690 Phantom Gaming 4/D5 ATX
- Hard Drive: Adata Legend SSD 1 TB M.2 NVMe PCI Express 4.0
- PSU: Corsair TX-M Series TX750M 750W

4.2: Φάση εκπαίδευσης ανιχνευτή (Detector)

Όπως αναφέρθηκε κατά το κεφάλαιο επιστημονικού υποβάθρου, η λειτουργία ενός «Tracking-by-Detection» μοντέλου παρακολούθησης απαιτεί την ανεξάρτητη/«offline» εκπαίδευση ενός μοντέλου ανίχνευσης. Ξεκινώντας λοιπόν από το συγκεκριμένο στάδιο, επιλέχθηκε η χρησιμοποίηση του Faster R-CNN, ενός end-to-end detector που, αν και προγενέστερος, παρέχει ανιχνεύσεις συγκρίσιμης ακρίβειας με τους υπόλοιπους «state-of-the-art» αλγορίθμους, ενώ φέρει χαμηλότερες απαιτήσεις σε εικονική μνήμη (VRAM). Μάλιστα, ήταν ο ανιχνευτής ο οποίος χρησιμοποιήθηκε κατά την πρωτότυπη έκδοση του DeepSORT tracker που αναπτύσσεται σε αυτή τη μεθοδολογία. Στόχος αποτελεί η εξαγωγή ανιχνεύσεων των οχημάτων, υπό την ταξινόμηση σε τέσσερις κατηγορίες:

- **Motorcycle – Δίκυκλο (0.5 MEA):**

Αναφέρεται στα δίκυκλα οχήματα οποιουδήποτε κυβισμού και μεγέθους.



Σχήμα 21: Παραδείγματα εμφάνισης ενός δίκυκλου (κίτρινο bounding box) κι ενός αυτοκινήτου (κυανό bounding box) στο Suncity Dataset.

- **Car – Αυτοκίνητο (1.0 MEA):**

Συμπεριλαμβάνει κοινά επιβατικά αυτοκίνητα, αλλά και οχήματα τροφοδοσίας/μεταφορών περίπου ίσου μεγέθους.

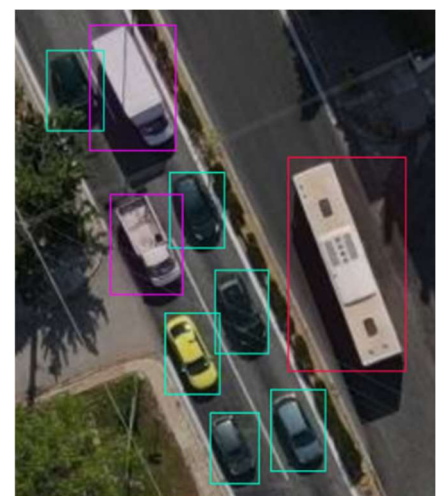


Σχήμα 22: Παραδείγματα εμφάνισης πολλαπλών αυτοκινήτων στο Suncity Dataset.

- **Semi Truck/Van – Ημιφορτηγό/Βαν (1.5 MEA):**

Αποτελείται από φορτηγά/ημιφορτηγά έως και 2 φορές μεγαλύτερου μήκους από επιβατικά αυτοκίνητα, καθώς και από «αγροτικά» αυτοκίνητα. Επειδή φέρει στοιχεία από τις κατηγορίες II και III από ΥΠΕΧΩΔΕ, η ισοδυναμία της θεωρήθηκε 1.5 MEA, κατά άτυπη σύμβαση.

- **Heavy Truck/Bus – Βαρύ φορτηγό/Λεωφορείο (2.0 MEA):** Πρόκειται για τα μεγαλύτερα σε μέγεθος φορτηγά και λεωφορεία, τουλάχιστον 2 φορές στο μήκος των επιβατικών αυτοκινήτων.



Σχήμα 23: Διάκριση εμφάνισης μεταξύ της κατηγορίας Semi Truck/Van (ματζέντα bounding box) και Heavy Truck/Bus (κόκκινο bounding box)

Με τη χρήση της βιβλιοθήκης Python «OpenCV», τα βίντεο του Suncity Dataset χωρίστηκαν σε ένα σύνολο μεμονωμένων καρτέ. Από αυτό δομήθηκε το σετ δεδομένων αλήθειας, το οποίο αποτελούνταν από τα εξής δομικά μέρη:

- **Δεδομένα εκπαίδευσης:** Ένα πολύ μικρό υποσύνολο, μόλις 300 καρτέ από 6 διαφορετικά αρχεία βίντεο (για να υπάρχει πληροφορία από κάθε διαφορετικό σενάριο φωτισμού – τοποθεσίας), τα οποία προέκυψαν με συστηματική δειγματοληψία ισοδιάστασης μερικών εκατοντάδων καρτέ, ώστε να παρατηρείται διαφοροποίηση σκηνικού – αντικειμένων σε καθένα από αυτά (δεδομένου ότι 25 καρτέ = 1 δευτερόλεπτο στο βίντεο). Φυσικά, οι εικόνες συνδυάζονται με ένα αρχείο «.json» που περιλαμβάνει το σύνολο των απαραίτητων περιγραφών (annotations) σχετικά με τις θέσεις των bounding boxes και το χαρακτηρισμό κατηγορίας. Η μορφή του αρχείου περιγραφής των δεδομένων ακολουθεί αυτή του Common Objects in Context (COCO) Dataset [33] (βλ. επόμενο σχήμα). Συνολικά, τα δεδομένα εκπαίδευσης περιλαμβάνουν ~36000 οχήματα, εκ των οποίων ~1200 αφορούν δίκυκλα, ~400 βαρέα οχήματα, ~32600 αυτοκίνητα, και ~1800 ημιφορτηγά, συμπεριλαμβάνοντας και ακινητοποιημένα/σταθμευμένα οχήματα. Από πλευράς κατανομής δειγμάτων πρόκειται για ένα αρκετά μη ισορροπημένο σετ δεδομένων, αλλά ωστόσο η χαρακτηριστική εμφάνιση και κλίμακα των αντικειμένων μειονοτικών κατηγοριών αποζημιώνει τις συνέπειες αυτής της ανισορροπίας.
- **Δεδομένα ελέγχου:** Εφαρμόστηκε ποσοτικός έλεγχος στα ίδια τα δεδομένα εκπαίδευσης, αλλά και στα διαδοχικά καρτέ (βίντεο) ελέγχου του αλγόριθμου παρακολούθησης (όπως αναφέρονται στο κεφ. 4.4), ως «detection-by-frame». Επιπλέον, για να εξεταστούν οι δυνατότητες γενίκευσης του μοντέλου σε εντελώς διαφορετικά σετ δεδομένων, ελέγχθηκαν 184 καρτέ από το Aerial Multi-Vehicle Detection Dataset [29] (~45000 οχήματα), με προσαρμοσμένο αρχείο περιγραφών για να υπάρχει μερική συμφωνία μεταξύ των κατηγοριών. Οι κατηγορίες του Aerial Multi-Vehicle Detection Dataset ήταν 3, Car, Truck και Bus, οπότε αντιστοιχίστηκαν ως εξής: Motorcycle → \emptyset , Car → Car, Heavy_Truck/Bus → Bus, Semi_Truck/Van → Truck.

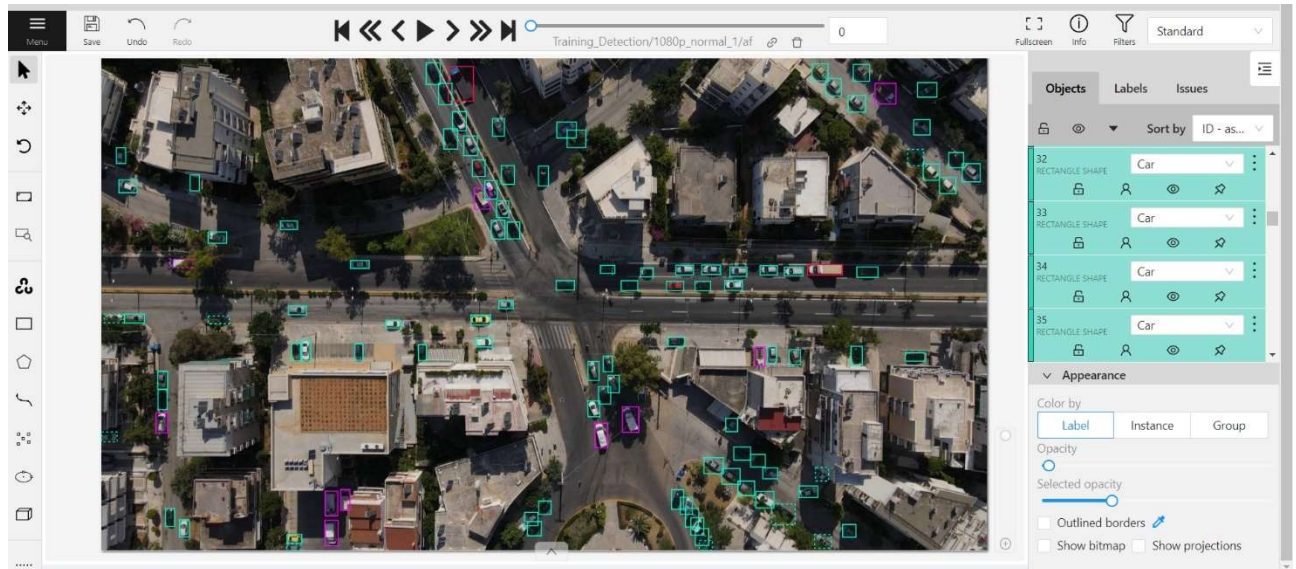
Σχήμα 24: Απόσπασμα από το επιλεγθέν υποσύνολο του Aerial Multi-Vehicle Detection Dataset [29], το οποίο χρησιμοποιήθηκε για έλεγχο επίδοσεων γενίκευσης του παραχθέντος μοντέλου ανίχνευσης. Παρατηρούνται πιο πυκνές συστοιχίες οχημάτων, ειδικά σε περιοχές στάθμευσης, καθώς και σύνθετες συνθήκες φωτισμού. Γενικότερα, τα αντικείμενα εμφανίζονται με ελαφρώς μικρότερο ποσοστό του μεγέθους της εικόνας, απ' ό,τι στο Suncity Dataset.



```
"categories": [
  {
    "id": 1,
    "name": "Motorcycle",
    "supercategory": ""
  },
  {
    "id": 2,
    "name": "Heavy_Truck/Bus",
    "supercategory": ""
  },
  {
    "id": 3,
    "name": "Car",
    "supercategory": ""
  },
  {
    "id": 4,
    "name": "Semi_Truck/Van",
    "supercategory": ""
  }
],
"images": [
  {
    "id": 1,
    "width": 1920,
    "height": 1080,
    "file_name": "Training_Dataset/frame0-00-00.00_1080p_normal_1_afternoon_1.jpg",
    "license": 0,
    "flickr_url": "",
    "coco_url": "",
    "date_captured": 0
  },
  {
    "id": 300,
    "width": 1920,
    "height": 1080,
    "file_name": "Training_Dataset/frame0-12-15.00_1080p_normal_2_midday.jpg",
    "license": 0,
    "flickr_url": "",
    "coco_url": "",
    "date_captured": 0
  }
],
"annotations": [
  {
    "id": 1,
    "image_id": 1,
    "category_id": 3,
    "segmentation": [],
    "area": 998.2499999999999,
    "bbox": [
      1054.14,
      755.5,
      26.62,
      37.5
    ],
    "iscrowd": 0,
    "attributes": {
      "occluded": false,
      "rotation": 0.0
    }
  },
  {
    "id": 36060,
    "image_id": 300,
    "category_id": 3,
    "segmentation": [],
    "area": 564.1127999999973,
    "bbox": [
      949.96,
      860.63,
      33.34,
      16.92
    ],
    "iscrowd": 0,
    "attributes": {
      "occluded": true,
      "rotation": 0.0
    }
  }
]
```

Σχήμα 25: Υποδείγματα εγγραφών περιγραφών στη μορφή του Common Objects in Context (COCO) Dataset [33].

Η δημιουργία των περιγραφών των δεδομένων αλήθειας πραγματοποιήθηκε γραφικά με «χειροκίνητο» τρόπο χρησιμοποιώντας το Computer Vision Annotation Tool (CVAT) [34]. Το εν λόγω εργαλείο λειτουργεί σε περιβάλλον Docker Engine, κι επιτρέπει την ανάρτηση των δεδομένων μελέτης ως project σε ένα τοπικό διακομιστή, αφού οριστούν οι απαραίτητες ετικέτες χαρακτηριστικών. Στην καρτέλα «Jobs» βρίσκεται το κύριο περιβάλλον εργασίας, στο οποίο σχεδιάστηκαν με ορισμό 2 σημείων τα «up-right» bounding boxes που περικλείουν τα οχήματα. Φυσικά υπάρχει σχετική επιλογή ορισμού ή αλλαγής κατηγορίας των σχεδιασμένων σχημάτων σε οποιοδήποτε στάδιο εργασίας. Με την ολοκλήρωση του συνόλου περιγραφών, η σχεδιασμένη πληροφορία εξάγεται σε αρχείο κειμένου, υποστηρίζοντας πολυάριθμες μορφές (στην προκειμένη περίπτωση επιλέχθηκε COCO 1.1 format). Σημειώνεται επίσης ότι απεικονίσεις οχημάτων με πάνω από ~60% ποσοστό συγκάλυψης αποκλείστηκαν, ενώ όσα θεωρήθηκαν καλυμμένα από ~10 έως ~60% συμπεριλήφθηκαν με το χαρακτηρισμό «occluded».



Σχήμα 26: Απεικόνιση του περιβάλλοντος εργασίας του Computer Vision Annotation Tool (CVAT), κατά τη δημιουργία των δεδομένων αλήθειας στο Suncity Dataset.

Με προετοιμασμένο το σύνολο δεδομένων αλήθειας, ακολουθήθηκε ο σχηματισμός ενός αρχείου παραμετροποίησης από τη δομή MMDetection σύμφωνα με το οποίο δρομολογήθηκε ο ανιχνευτής Faster R-CNN. Το μοντέλο εκπαιδεύτηκε για 60 εποχές, αποθηκεύοντας την κατάσταση του κάθε 10 εποχές, και η εκπαίδευση διήρκεσε περίπου 6.5 ώρες, καταναλώνοντας τα 7-7.1 GB από τα πραγματικά διαθέσιμα 7.32 GB VRAM.

Απόσπασμα Κώδικα 1: Αρχείο παραμετροποίησης της δομής MMDetection, για το παραχθέν μοντέλο ανίχνευσης Faster R-CNN. Σημαντικότερες ρυθμίσεις αποτέλεσαν: α) η μείωση Anchor scale από την προεπιλογή 8 σε 3, ώστε να προσεγγίζεται το μέγεθος των αντικειμένων, β) η αύξηση Roi layer feature size από την προεπιλογή 7 σε 18, για να αποτυπώνεται η μέγιστη δυνατή πληροφορία χαρακτηριστικών που επιτρέπουν οι περιορισμοί VRAM, καθώς και γ) η ελάττωση του ορίου IoU θετικών δειγμάτων στο Roi τμήμα από την προεπιλογή 0.5 σε 0.3, για ανοχή σε λιγότερο ακριβές «localization» έναντι των μικροσκοπικών αντικειμένων. (πηγή πρωτότυπου: [32])

```
1 #Παραμετροποίηση του Faster R-CNN μοντέλου.
2 model = dict(
3     type='FasterRCNN', #Ορισμός της υλοποίησης αλγορίθμου Faster R-CNN.
4     #Εφαρμογή μιας δομής προεπεξεργασίας εικόνων, η οποία πραγματοποιεί κανονικοποίηση
5     #και padding για την εφαρμογή των συνελίξεων.
6     data_preprocessor=dict(
7         type='DetDataPreprocessor', #Ορισμός υλοποίησης δομής προεπεξεργασίας.
8         mean=[123.675, 116.28, 103.53], #Ορισμός μέσης τιμής [R,G,B] (από ImageNet).
9         std=[58.395, 57.12, 57.375], #Ορισμός τυπικής απόκλισης [R,G,B] (από ImageNet).
10        bgr_to_rgb=True, #Μετατροπή "διαβάσματος" καναλιών εικόνας από BGR σε RGB.
11        pad_size_divisor=32), #Ορισμός διαιρέτη της padded εικόνας.
12    #Παραμετροποίηση του νευρωνικού δικτύου "backbone" πάνω στο οποίο βασίζεται ο detector.
13    backbone=dict(
14        type='ResNet', #Επιλογή της αρχιτεκτονικής ResNet.
15        depth=50, #Βάθος επιπέδων του ResNet. Εδώ ουσιαστικά επιλέγεται η παραλλαγή ResNet-50.
16        num_stages=4, #Ορισμός 4 σταδίων-μπλοκ συνελικτικών επιπέδων στα οποία
17        #κατανέμονται τα συνολικά επίπεδα.
18        out_indices=(0, 1, 2, 3), #Οι δείκτες των πινάκων χαρακτηριστικών που
19        #παράγονται από κάθε στάδιο.
```

```
20     frozen_stages=1, #Παγιοποίηση των παραμέτρων του πρώτου σταδίου.
21     norm_cfg=dict(type='BN', requires_grad=True), #Παραμετροποίηση εκπαιδευσιμου
22     #Batch Normalization
23     norm_eval=True, #Παγιοποίηση στατιστικών στο Batch Normalization
24     style='pytorch', #Ορισμός μορφής backbone για pytorch
25     #(τα συνελκτικικά επίπεδα 2ου σταδίου είναι 3x3).
26     init_cfg=dict(type='Pretrained', checkpoint='torchvision://resnet50'),
27     #Αρχικοποίηση με προεκπαιδευμένο μοντέλο.
28 #Ορισμός του τμήματος "neck" για την άντληση πληροφοριών από διαφορετικά στάδια του backbone,
29 #αποσκοπώντας στη διαχείριση χαρακτηριστικών διαφορετικών κλιμάκων.
30     neck=dict(
31         type='FPN', #Επιλογή χρήσης Feature Pyramid Network, αν και δεν υπάρχει
32         #στον πρωτότυπο Faster R-CNN.
33         in_channels=[256, 512, 1024, 2048], #Κανάλια εισόδου, προερχόμενα από το backbone.
34         out_channels=256, #Κανάλια εξόδου
35         num_outs=5), #Το FPN θα παράγει 5 διαφορετικές κλίμακες πινάκων χαρακτηριστικών.
36 #Παραμετροποίηση του τμήματος RPN.
37     rpn_head=dict(
38         type='RPNHead', #Επιλογή της κλασσικής υλοποίησης τμήματος RPN.
39         in_channels=256, #Κανάλια εισόδου, προερχόμενα από το FPN.
40         feat_channels=256, #Κανάλια χαρακτηριστικών.
41         #Παραμετροποίηση γεννήτριας anchors.
42         anchor_generator=dict(
43             type='AnchorGenerator', #Επιλογή κλασσικής υλοποίησης γεννήτριας anchors.
44             scales=[3], #Ορισμός της βασικής κλίμακας anchors που θα εφαρμόζονται
45             #σε διαφορετικά επίπεδα (εδω 3x3).
46             base_sizes=[4, 8, 16, 32, 64],
47             #Μεγέθη των anchors που πολλαπλασιάζονται με την κλίμακα σε κάθε επίπεδο,
48             #για απόδοση αντιστοιχίας στην έκταση της αρχικής εικόνας.
49             #Διακρίνονται 5 επίπεδα, όσες και οι 5 διαφορετικές κλίμακες πινάκων
50             #χαρακτηριστικών που εξάγει το FPN, ενώ αντιστοιχεί στα strides τους.
51             #Δηλαδή με 1080p αρχικές εικόνες, οι διαστάσεις feature map είναι (w/stride, h/stride):
52             #(1920/4,1080/4)=(480,270), (240,135), (120,68), (60,34), (30,17).
53             #Οπότε anchor με βασικό μέγεθος 3x3 "κυλιέται" σε κάθε μια από αυτές τις διαστάσεις.
54             #Στο επίπεδο (640,360) το 3x3 anchor στοχεύει στην ανίχνευση μικρών αντικειμένων.
55             #Αντίστοιχα στο επίπεδο (30,17) στοχεύει στην ανίχνευση μεγαλύτερων.
56             #Επιλέχθηκαν κοινοί διαρέτες των 1920 και 1080.
57             ratios=[0.5, 1.0, 2.0], #Οι αναλογίες πλευρών των anchors.
58             strides=[4, 8, 16, 32, 64]), #Τα strides στους πίνακες χαρακτηριστικών
59             #που αναφέρθηκαν παραπάνω.
60         #Κωδικοποίηση bounding box σε επίπεδο RPN.
61         bbox_coder=dict(
62             type='DeltaXYWHBBoxCoder', #Επιλογή κλασσικής υλοποίησης κωδικοποίησης bounding box.
63             target_means=[0.0, 0.0, 0.0, 0.0], #Εξουδετέρωση μέσης τιμής,
64             #για μη εφαρμογή κανονικοποίησης συντεταγμένων.
65             target_stds=[1.0, 1.0, 1.0, 1.0]), #Εξουδετέρωση τυπικής απόκλισης,
66             #για μη εφαρμογή κανονικοποίησης συντεταγμένων.
67         #Ρύθμιση συνάρτησης κόστους ταξινόμησης σε επίπεδο RPN.
68         loss_cls=dict(
69             type='CrossEntropyLoss', use_sigmoid=True, loss_weight=1.0),
70         #Χρήση Binary Cross-Entropy Loss (sigmoid αντί softmax),
71         #για διάκριση background-foreground.
72         #Ρύθμιση συνάρτησης κόστους για την παλινδρόμηση bounding box σε επίπεδο RPN.
73         loss_bbox=dict(type='L1Loss', loss_weight=1.0)), #Χρήση L1Loss.
74 #Παραμετροποίηση του τμήματος RoI.
75     roi_head=dict(
```

```
76     type='StandardRoIHead', #Επιλογή κλασσικής υλοποίησης τμήματος RoI.
77     #Ρύθμιση εξαγωγή RoI.
78     bbox_roi_extractor=dict(
79         type='SingleRoIExtractor', #Επιλογή τύπου εξαγωγή ενός RoI.
80         roi_layer=dict(type='RoIAlign', output_size=18, sampling_ratio=0),
81         #Ορισμός RoI Align στρώσης με μέγεθος εξόδου 18x18
82         #και adaptive ratio στην εξαγωγή RoI(sampling ratio).
83         out_channels=256, #Κανάλια χαρακτηριστικών εξόδου.
84         featmap_strides=[4, 8, 16, 32]), #Strides των πινάκων χαρακτηριστικών,
85         #σε καθένα από τα στάδια του backbone.
86     #Ρύθμιση τμήματος bounding box προβλέψεων.
87     bbox_head=dict(
88         type='Shared2FCBBoxHead', #Επιλογή κλασσικού τύπου τμήματος bounding box προβλέψεων
89         # (διαβίβαση των χαρακτηριστικών σε
90         #fully connected στρώσεις νευρωνικού δικτύου).
91         in_channels=256, #Κανάλια εισόδου, προερχόμενων από τον εξαγωγή RoI.
92         fc_out_channels=1024, #Κανάλια εξόδου στις fully connected στρώσεις.
93         roi_feat_size=18, #Μέγεθος χαρακτηριστικών RoI.
94         num_classes=4, #Πλήθος foreground κατηγοριών ταξινόμησης.
95         #Κωδικοποίηση bounding box σε επίπεδο τελικών προβλέψεων.
96         bbox_coder=dict(
97             type='DeltaXYWHBBoxCoder',
98             target_means=[0.0, 0.0, 0.0, 0.0],
99             target_stds=[0.1, 0.1, 0.2, 0.2]),
100         reg_class_agnostic=False, #Η παλινδρόμηση συμβαίνει εν-γνώσει της κατηγορίας.
101         #Ρύθμιση συνάρτησης κόστους ταξινόμησης σε επίπεδο τελικών προβλέψεων.
102         loss_cls=dict(
103             type='CrossEntropyLoss', use_sigmoid=False, loss_weight=1.0),
104         #Χρήση Categorical Cross-Entropy Loss (softmax αντί sigmoid),
105         #για διάκριση πολλαπλών κατηγοριών.
106         #Ρύθμιση συνάρτησης κόστους για την παλινδρόμηση bounding box σε επίπεδο
107         #τελικών προβλέψεων.
108         loss_bbox=dict(type='LlLoss', loss_weight=1.0)), #Χρήση LlLoss.
109     #Παραμετροποίηση χαρακτηριστικών τιμών κατά την εκπαίδευση.
110     train_cfg=dict(
111         #Χαρακτηριστικές τιμές εκπαίδευσης του RPN.
112         rpn=dict(
113             #Εκχώρητης δειγμάτων RPN.
114             assigner=dict(
115                 type='MaxIoUAssigner', #Επιλογή εκχώρησης βάσει IoU.
116                 pos_iou_thr=0.7, #Κατώφλι θετικών δειγμάτων.
117                 neg_iou_thr=0.3, #Κατώφλι αρνητικών δειγμάτων.
118                 min_pos_iou=0.3, #Κατώφλι μη αρνητικών δειγμάτων = θετικών.
119                 match_low_quality=True, #Εκχώρηση δειγμάτων χαμηλής ποιότητας.
120                 ignore_iof_thr=-1), #Αποφυγή χρήσης κατωφλίου τιμής Intersection
121                 #over Foreground για αγνόηση δειγμάτων.
122             #Δειγματολήπτης RPN.
123             sampler=dict(
124                 type='RandomSampler', #Επιλογή τυχαίου δειγματολήπτη.
125                 num=256, #Αριθμός δειγμάτων.
126                 pos_fraction=0.5, #Αναλογία θετικών δειγμάτων.
127                 neg_pos_ub=-1, #Μη οριοθέτηση αρνητικών δειγμάτων, βάσει πλήθους θετικών.
128                 add_gt_as_proposals=False), #Μη προσθήκη δεδομένων αλήθειας ως προτάσεις.
129                 allowed_border=-1, #Οι προτάσεις δεν περιορίζονται από την εγγύτητά τους
130                 #στο όριο της εικόνας.
131                 pos_weight=-1, #Αυτόματη εκχώρηση βαρών στα θετικά δείγματα.
```



```
132     debug=False),
133     #Παραμετροποίηση δημιουργίας προτάσεων.
134     rpn_proposal=dict(
135         nms_pre=2000, #Πλήθος προτάσεων πριν από NMS.
136         max_per_img=1000, #Μέγιστο πλήθος προτάσεων μετά από NMS.
137         nms=dict(type='nms', iou_threshold=0.7), #Ορισμός κατωφλιού IoU για NMS.
138         min_bbox_size=0), #Ορισμός ελάχιστου επιτρεπόμενου μεγέθους bounding box.
139     #Χαρακτηριστικές τιμές εκπαίδευσης του R-CNN στο τμήμα RoI.
140     rcnn=dict(
141         #Εκχωρητής δειγμάτων RoI.
142         assigner=dict(
143             type='MaxIoUAssigner', #Επιλογή εκχώρησης βάσει IoU.
144             pos_iou_thr=0.3, #Κατώφλι θετικών δειγμάτων.
145             neg_iou_thr=0.3, #Κατώφλι αρνητικών δειγμάτων.
146             min_pos_iou=0.3, #Κατώφλι μη αρνητικών δειγμάτων = θετικών.
147             match_low_quality=True, #Εκχώρηση δειγμάτων χαμηλής ποιότητας.
148             ignore_iof_thr=-1), #Αποφυγή χρήσης κατωφλιού τιμής Intersection
149             #over Foreground για αγνόηση δειγμάτων.
150         #Δειγματολήπτης RoI.
151         sampler=dict(
152             type='RandomSampler', #Επιλογή τυχαίου δειγματολήπτη.
153             num=512, #Αριθμός δειγμάτων.
154             pos_fraction=0.25, #Αναλογία θετικών δειγμάτων.
155             neg_pos_ub=-1, #Μη οριοθέτηση αρνητικών δειγμάτων, βάσει πλήθους θετικών.
156             add_gt_as_proposals=True), #Προσθήκη δεδομένων αλήθειας ως προτάσεις.
157             pos_weight=-1, #Αυτόματη εκχώρηση βαρών στα θετικά δείγματα.
158             debug=False)),
159     #Παραμετροποίηση χαρακτηριστικών τιμών κατά τον έλεγχο.
160     test_cfg=dict(
161         #Χαρακτηριστικές τιμές του RPN κατά τον έλεγχο.
162         rpn=dict(
163             nms_pre=1000, #Πλήθος προτάσεων πριν από NMS.
164             max_per_img=1000, #Μέγιστο πλήθος προτάσεων μετά από NMS.
165             nms=dict(type='nms', iou_threshold=0.7), #Ορισμός κατωφλιού IoU για NMS.
166             min_bbox_size=0), #Ορισμός ελάχιστου επιτρεπόμενου μεγέθους bounding box.
167             #Χαρακτηριστικές τιμές του R-CNN στο τμήμα RoI κατά τον έλεγχο.
168             rcnn=dict(
169                 score_thr=0.05, #Εμφάνιση προβλέψεων με confidence > 5%
170                 nms=dict(type='nms', iou_threshold=0.5), #Ορισμός κατωφλιού IoU για NMS.
171                 max_per_img=150)) #Ορισμός μεγίστους αριθμού προβλέψεων ανα εικόνα.
172
173 #Ορισμός format του dataset, και κύριας τοποθεσίας φακέλου.
174 dataset_type = 'CocoDataset' #Επιλογή COCO format.
175 data_root = 'path/to/data/root'
176 backend_args = None
177
178 #Πληροφορίες των κατηγοριών.
179 meta_info=dict(
180     classes=('Motorcycle', 'Heavy_Truck/Bus', 'Car', 'Semi_Truck/Van'),
181     palette=[(255, 20, 20), (20, 255, 20), (20, 20, 255),
182             (220, 60, 130)])
183
184 #Ορισμός ακολουθίας εκπαίδευσης.
185 train_pipeline = [
186     dict(type='LoadImageFromFile', backend_args=None), #Φόρτωση εικόνας από αρχείο.
187     dict(type='LoadAnnotations', with_bbox=True), #Φόρτωση annotations μαζί με bounding boxes.
```

```
188 dict(type='Resize', scale=(1920, 1080), keep_ratio=True), #Resize των εικόνων σε (1920x1080).
189 dict(type='RandomFlip', prob=0.5), #Τυχαία ανάκλαση εικόνας.
190 dict(type='PackDetInputs') #Μορφοποίηση των annotations για απόδοση κλειδιών στα δείγματα.
191 ]
192
193 #Ορισμός ακολουθίας ελέγχου.
194 test_pipeline = [
195     dict(type='LoadImageFromFile', backend_args=None), #Φόρτωση εικόνας από αρχείο.
196     dict(type='LoadAnnotations', with_bbox=True), #Φόρτωση annotations μαζί με bounding boxes.
197     dict(type='Resize', scale=(1920, 1080), keep_ratio=True), #Resize των εικόνων σε (1920x1080).
198     dict(
199         type='PackDetInputs',
200         meta_keys=('img_id', 'img_path', 'ori_shape', 'img_shape',
201                   'scale_factor')) #Μορφοποίηση των annotations για απόδοση κλειδιών στα δείγματα.
202 ]
203
204 #Παραμετροποίηση φορτωτή εκπαίδευσης.
205 train_dataloader = dict(
206     batch_size=1, #Ορισμός μεγέθους batch.
207     num_workers=2,
208     persistent_workers=True,
209     sampler=dict(type='DefaultSampler', shuffle=True),
210     batch_sampler=dict(type='AspectRatioBatchSampler'),
211     #Χαρακτηριστικά dataset.
212     dataset=dict(
213         type=dataset_type,
214         data_root=data_root,
215         ann_file=
216         'path/to/annotations.json',
217         data_prefix=dict(
218             img='path/to/images/'),
219         filter_cfg=dict(filter_empty_gt=True, min_size=32),
220         pipeline=train_pipeline,
221         backend_args=None,
222         meta_info=meta_info))
223
224 #Παραμετροποίηση φορτωτή επικύρωσης.
225 val_dataloader = dict(
226     batch_size=1, #Ορισμός μεγέθους batch.
227     num_workers=2,
228     persistent_workers=True,
229     drop_last=False,
230     sampler=dict(type='DefaultSampler', shuffle=False),
231     #Χαρακτηριστικά dataset.
232     dataset=dict(
233         type=dataset_type,
234         data_root=data_root,
235         ann_file=
236         'path/to/annotations.json',
237         data_prefix=dict(
238             img='path/to/images/'),
239         test_mode=True,
240         pipeline=test_pipeline,
241         backend_args=None,
242         meta_info=meta_info))
243
```

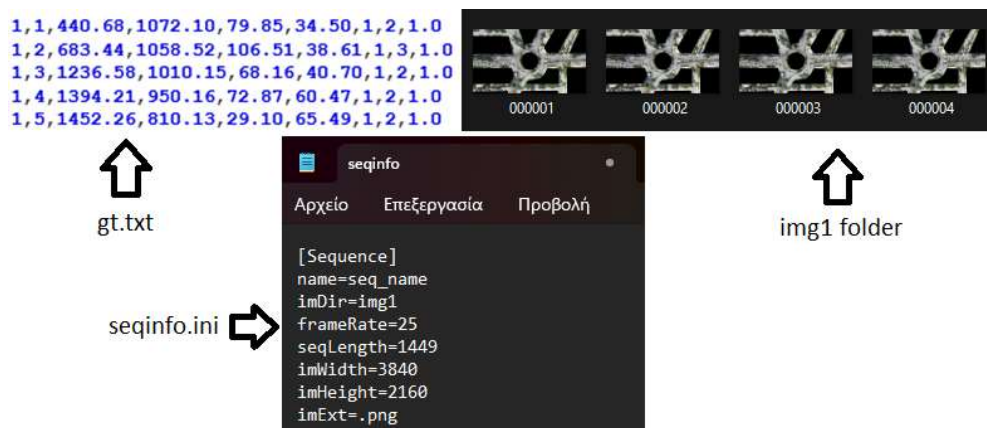
```
244 #Παραμετροποίηση φορτωτή ελέγχου.
245 test_dataloader = dict(
246     batch_size=1, #Ορισμός μεγέθους batch.
247     num_workers=2,
248     persistent_workers=True,
249     drop_last=False,
250     sampler=dict(type='DefaultSampler', shuffle=False),
251     #Χαρακτηριστικά dataset.
252     dataset=dict(
253         type=dataset_type,
254         data_root=data_root,
255         ann_file=
256         'path/to/annotations.json',
257         data_prefix=dict(
258             img='path/to/images/'),
259         test_mode=True,
260         pipeline=test_pipeline,
261         backend_args=None,
262         meta_info=meta_info)
263
264 #Παραμετροποίηση αξιολόγησης στα δεδομένα επικύρωσης.
265 val_evaluator = dict(
266     type='CocoMetric',
267     ann_file=
268     'path/to/annotations.json',
269     metric='bbox',
270     format_only=False,
271     backend_args=None)
272
273 #Παραμετροποίηση αξιολόγησης στα δεδομένα ελέγχου.
274 test_evaluator = dict(
275     type='CocoMetric',
276     ann_file=
277     'path/to/annotations.json',
278     metric='bbox',
279     format_only=False,
280     backend_args=None)
281
282 #Παραμετροποίηση επαναλήψεων εκπαίδευσης, επικύρωσης, κι ελέγχου.
283 train_cfg = dict(type='EpochBasedTrainLoop', max_epochs=120, val_interval=10)
284 #Ορισμός εποχών εκπαίδευσης, και ισοδιάστασης αξιολόγησης.
285 val_cfg = dict(type='ValLoop')
286 test_cfg = dict(type='TestLoop')
287
288 #Παραμετροποίηση βελτιστοποιητή.
289 optim_wrapper = dict(
290     type='OptimWrapper',
291     optimizer=dict(type='SGD', lr=0.02, momentum=0.9, weight_decay=0.0001)) #Επιλογή SGD,
292     #βασικού βαθμού εκμάθησης, με momentum και weight decay.
293
294 #Ρυθμίσεις εναλλαγής βαθμού εκμάθησης.
295 param_scheduler = [
296     #Προθέρμανση του μοντέλου με χρήση αρχικά μικρότερων ρυθμών εκμάθησης,
297     #γραμμικά αυξανόμενων για 600 διελεύσεις εικόνων (2 εποχές).
298     dict(
299         type='LinearLR', start_factor=0.0001, by_epoch=False, begin=0, end=600),
```

```
300 #Σύμφωνα με την παρακάτω ρύθμιση ο ρυθμός εκμάθησης μικραίνει μια τάξη μεγέθους
301 #σε κάθε εποχή-milestone.
302 dict(
303     type='MultiStepLR',
304     begin=0,
305     end=120,
306     by_epoch=True,
307     milestones=[4, 22, 60],
308     gamma=0.1)
309 ]
310
311
312 #Μη αυτοματοποιημένη διαχείριση ρυθμού εκμάθησης.
313 auto_scale_lr = dict(enable=False, base_batch_size=16)
314
315 default_scope = 'mmdet'
316
317 #Hooks για πραγματοποίηση ενεργειών.
318 default_hooks = dict(
319     timer=dict(type='IterTimerHook'),
320     logger=dict(type='LoggerHook', interval=300),
321     param_scheduler=dict(type='ParamSchedulerHook'),
322     checkpoint=dict(type='CheckpointHook', interval=10),
323     sampler_seed=dict(type='DistSamplerSeedHook'),
324     visualization=dict(type='DetVisualizationHook'))
325
326 #Ρύθμιση περιβάλλοντος.
327 env_cfg = dict(
328     cudnn_benchmark=False,
329     mp_cfg=dict(mp_start_method='fork', opencv_num_threads=0),
330     dist_cfg=dict(backend='nccl'))
331
332 #Ρύθμιση οπτικοποιήσεων.
333 vis_backends = [dict(type='LocalVisBackend')]
334 visualizer = dict(
335     type='DetLocalVisualizer',
336     vis_backends=[dict(type='LocalVisBackend')],
337     name='visualizer')
338
339 #Ρύθμιση αρχείου καταγραφής.
340 log_processor = dict(type='LogProcessor', window_size=50, by_epoch=True)
341 log_level = 'INFO'
342
343 #Επιλογή συνέχισης/φόρτωσης.
344 load_from = None
345 resume = False
346
347 #Επιλογή launcher.
348 launcher = 'none'
349
350 #Επιλογή τοποθεσίας φακέλου εργασίας.
351 work_dir = './work_dirs\\faster-rcnn_r50_fpn_1x_coco'
```

4.3: Φάση εκπαίδευσης επαναταυτοποίησης (ReID)

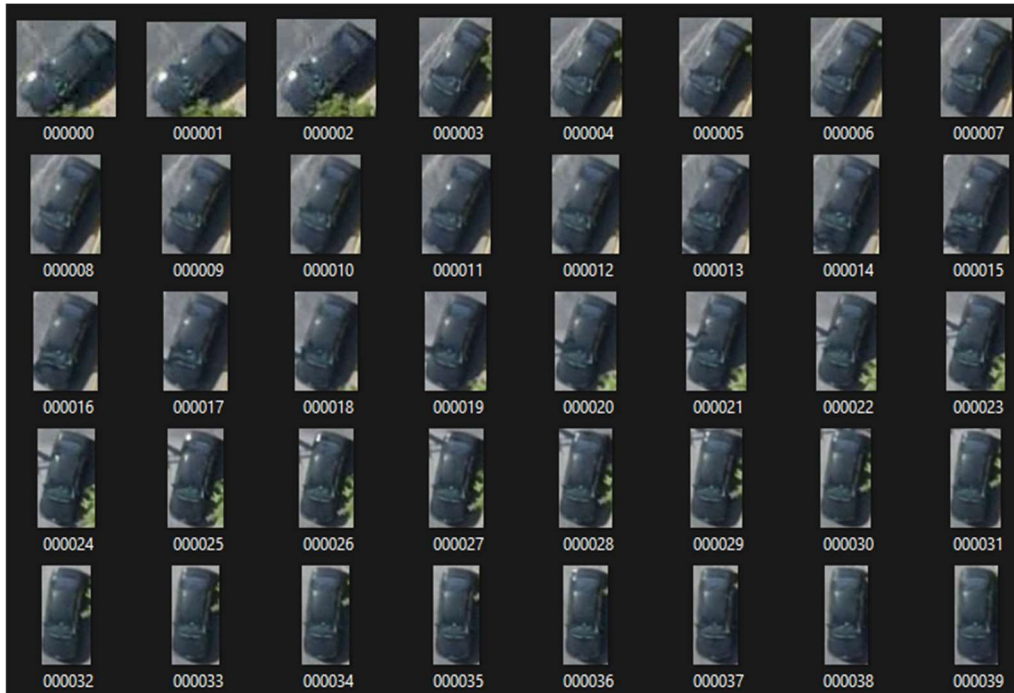
Την εκπαίδευση του detector διαδέχθηκε η «offline» εκπαίδευση του μοντέλου CNN που βοηθά στην απόκτηση των descriptors εμφάνισης των οχημάτων, που άλλωστε αποτελεί και τη βασική καινοτομία του DeepSORT tracker. Η ενέργεια αυτή αποσκοπεί στην εξομάλυνση του προβλήματος αντιστοίχισης στόχων διαδοχικών καρέ, ώστε να προσεγγίζεται διαρκής επαναταυτοποίηση των οχημάτων υπό μοναδικά και μη μεταβαλλόμενα IDs καθ' όλη τη διάρκεια μιας ακολουθίας βίντεο. Χρησιμοποιήθηκε το μοντέλο «BaseReID» της δομής MMTracking, το οποίο βασίζεται στο ResNet-50 CNN.

Το μοντέλο επαναταυτοποίησης εκπαιδεύτηκε στα 78 οχήματα που απεικονίζονται στο βίντεο DJI_0912.mp4 διάρκειας ~58" (1449 καρέ). Σε κάθε καρέ σχεδιάστηκαν (με το CVAT) τα bounding boxes όλων των κινούμενων οχημάτων, διακρινόμενων σε κατηγορίες, και επιλέχθηκε η εξαγωγή του αρχείου κειμένου περιγραφών (.txt) με τη μορφή MOT 1.1 [4], η οποία επεξηγείται παρακάτω:



Σχήμα 27: Απόσπασμα από annotation file σε MOT 1.1 format. Από αριστερά προς τα δεξιά διακρίνονται τα εξής: <frame_id>, <instance_id>, <x1>, <y1>, <w>, <h>, <conf> (=1 στα ground truth annotations), <class_id>, <visibility>. Το συγκεκριμένο αρχείο «gt.txt» βρίσκεται εντός τοποθεσίας τύπου «../seq_name/gt/». Ο φάκελος «seq_name» περιέχει επίσης τον υποφάκελο «img1», ο οποίος διαθέτει τα καρέ στα οποία αντιστοιχούν οι περιγραφές, με ονομασίες που παραπέμπουν στα frame IDs. Ακόμη, στην ίδια τοποθεσία υπάρχει το αρχείο ορισμού ακολουθίας «seqinfo.ini».

Το BaseReID μοντέλο δέχεται ως δεδομένα εκπαίδευσης τις τομές των bounding boxes με τα καρέ, όπου βρίσκεται αποκλειστικά η απεικόνιση του αντικειμένου ενδιαφέροντος. Στην προκειμένη περίπτωση, υπήρξε ένας φάκελος για κάθε όχημα που περιείχε τις εν λόγω τομές. Το format αυτό αναφέρεται ως «ReID», και διαφέρει σαφώς από το MOT 1.1 πρότυπο, αλλά επιλέχθηκε η εξαγωγή σε MOT 1.1 διότι περιλαμβανόταν αρχείο κώδικα «mot2reid.py» στη δομή MMTracking, που εκτελούσε εύκολα τη μετατροπή σε ReID.



Σχήμα 28: Οι τομές των bounding boxes ενός αυτοκινήτου στα διαδοχικά καρέ στα οποία εμφανίζεται σε ένα βίντεο.

Παρατίθεται παρακάτω το αντίστοιχο αρχείο παραμετροποίησης του MMTracking για τη λειτουργία του μοντέλου BaseReID, το οποίο εκπαιδεύτηκε για 10 εποχές συνολικής διάρκειας περίπου 16 ωρών.

Απόσπασμα Κώδικα 2: Αρχείο παραμετροποίησης της δομής MMTracking, για το ανεξάρτητο παραχθέν μοντέλο επαναταυτοποίησης BaseReID, βασισμένο στο ResNet-50. Οι προεπιλεγμένες τιμές μεταβλητών ήταν αρκετές ώστε να επιλυθεί άριστα το πρόβλημα κατηγοριοποίησης ~34000 εικονιδίων σε 78 οχήματα, μετά από 10 εποχές (mAP = 100%). Ωστόσο η εκπαίδευση απεδείχθη ιδιαίτερα χρονοβόρα, συγκριτικά με την αντίστοιχη για το detection. (πηγή πρωτότυπου: [32])

```
1 #Παραμετροποίηση του μοντέλου επαναταυτοποίησης.
2 TRAIN_REID = True
3 model = dict(
4     type='BaseReID', #Ορισμός της υλοποίησης BaseReID
5     #Παραμετροποίηση του νευρικού δικτύου "backbone" που πράττει την ταξινόμηση
6     backbone=dict(
7         type='ResNet', #Επιλογή της αρχιτεκτονικής ResNet.
8         depth=50, #Βάθος επιπέδων του ResNet. Εδώ ουσιαστικά επιλέγεται η παραλλαγή ResNet-50.
9         num_stages=4, #Ορισμός 4 σταδίων-μπλοκ συνελκτικών επιπέδων στα οποία
10        #κατανέμονται τα συνολικά επίπεδα.
11        out_indices=(3, ), #Οι δείκτες των πινάκων χαρακτηριστικών που παράγονται από κάθε στάδιο.
12        style='pytorch', #Ορισμός μορφής backbone για pytorch
13        #(τα συνελκτικά επίπεδα 2ου σταδίου είναι 3x3).
14        #Ορισμός του τμήματος "neck" εφαρμόζοντας Global Average Pooling πριν την διαδοχή
15        #fully-connected στρώσεων.
16        neck=dict(type='GlobalAveragePooling', kernel_size=(8, 4), stride=1),
17        #Ορισμός του τμήματος κατηγοριοποίησης
18        head=dict(
19            type='LinearReIDHead', #Επιλογή του τύπου 'LinearReIDHead' τμήματος κατηγοριοποίησης.
20            num_fcs=2, #Χρήση 2 επιπέδων fully-connected στρώσεων
```

```
21     in_channels=2048, #Κανάλια εισόδου.
22     fc_channels=1024, #Κανάλια fully-connected στρώσεων.
23     out_channels=128, #Κανάλια εξόδου.
24     num_classes=78, #Αριθμός κλάσεων κατηγοριοποίησης, ίσος με τον αριθμό των
25     #διαφορετικών οχημάτων που απεικονίζονται στο χρησιμοποιηθέν ReID dataset.
26     loss=dict(type='CrossEntropyLoss', loss_weight=1.0), #Ορισμός συνάρτησης
27     #κόστους κατηγοριοποίησης (cross-entropy loss).
28     loss_pairwise=dict(type='TripletLoss', margin=0.3, loss_weight=1.0), #Προσθήκη
29     #συνάρτησης κόστους TripletLoss, η οποία προσαρμόζει το μοντέλο ώστε να
30     #υπάρχει μεγάλη "απόσταση" από ένα εκπαιδευμένο δείγμα αναφοράς προς ένα αρνητικό δείγμα,
31     #και μικρή προς ένα θετικό.
32     norm_cfg=dict(type='BN1d'), #Εφαρμογή Batch Normalization
33     act_cfg=dict(type='ReLU'), #Επιλογή της ReLU ως συνάρτηση ενεργοποίησης.
34     #Αρχικοποίηση των βαρών του CNN βάσει προεκπεδευμένου checkpoint από ImageNET.
35     init_cfg=dict(
36         type='Pretrained',
37         checkpoint=
38         'https://download.openmmlab.com/./resnet50_batch256_imagenet_20200708-cfb998bf.pth'
39     ))
40
41 #Ορισμός του 'ReIDDataset' ως φορμάτ του dataset.
42 dataset_type = 'ReIDDataset'
43
44 #Ορισμός ακολουθίας εκπαίδευσης.
45 train_pipeline = [
46     dict(type='LoadMultiImagesFromFile', to_float32=True), #Φόρτωση πολλαπλών εικόνων
47     #(τομών bounding boxes) από αρχείο, σε dtype float32.
48     #Resize τομών bounding boxes με τις εικόνες σε (128x256).
49     dict(
50         type='SeqResize',
51         img_scale=(128, 256),
52         share_params=False,
53         keep_ratio=False,
54         bbox_clip_border=False,
55         override=False),
56     #Τυχαία ανάκλιση.
57     dict(
58         type='SeqRandomFlip',
59         share_params=False,
60         flip_ratio=0.5,
61         direction='horizontal'),
62     #Κανονικοποίηση βάσει μέσης τιμής και τυπικής απόκλισης, σε RGB.
63     dict(
64         type='SeqNormalize',
65         mean=[123.675, 116.28, 103.53],
66         std=[58.395, 57.12, 57.375],
67         to_rgb=True),
68     #Φόρτωση κλειδιών εικόνας και ετικετών από την ακολουθία βίντεο.
69     dict(type='VideoCollect', keys=['img', 'gt_label']),
70     dict(type='ReIDFormatBundle')
71 ]
72
73 #Ορισμός ακολουθίας ελέγχου.
74 test_pipeline = [
75     dict(type='LoadImageFromFile'),
76     dict(type='Resize', img_scale=(128, 256), keep_ratio=False),
```

```
77     dict(  
78         type='Normalize',  
79         mean=[123.675, 116.28, 103.53],  
80         std=[58.395, 57.12, 57.375],  
81         to_rgb=True),  
82     dict(type='ImageToTensor', keys=['img']),  
83     dict(type='Collect', keys=['img'], meta_keys=[])  
84 ]  
85  
86 #Ορισμός χαρακτηριστικών dataset, για training, validation και test.  
87 data_root = 'path/to/reid/train/data/'  
88 data = dict(  
89     samples_per_gpu=1,  
90     workers_per_gpu=2,  
91  
92     train=dict(  
93         type='ReIDDataset',  
94         triplet_sampler=dict(num_ids=8, ins_per_id=4),  
95         data_prefix=  
96         'path/to/reid/train/data/images',  
97         ann_file=  
98         'path/to/reid/train/data/annotations.txt',  
99         pipeline=train_pipeline),  
100  
101     val=dict(  
102         type='ReIDDataset',  
103         triplet_sampler=None,  
104         data_prefix=  
105         'path/to/reid/train/data/images',  
106         ann_file=  
107         'path/to/reid/train/data/annotations.txt',  
108         pipeline=test_pipeline),  
109  
110     test=dict(  
111         type='ReIDDataset',  
112         triplet_sampler=None,  
113         data_prefix=  
114         'path/to/reid/train/data/images',  
115         ann_file=  
116         'path/to/reid/train/data/annotations.txt',  
117         pipeline=test_pipeline))  
118  
119 #Αξιολόγηση με mAP.  
120 evaluation = dict(interval=1, metric='mAP')  
121  
122 #Ρύθμιση βελτιστοποίησης και ρυθμού εκμάθησης.  
123 optimizer = dict(type='SGD', lr=0.01, momentum=0.9, weight_decay=0.0001)  
124 optimizer_config = dict(grad_clip=None)  
125 lr_config = dict(  
126     policy='step',  
127     warmup='linear',  
128     warmup_iters=34070,  
129     warmup_ratio=0.001,  
130     step=[5])  
131 total_epochs = 10  
132
```



```
133 #Ρύθμιση συχνότητας αποθήκευσης και καταγραφής αναφορών.
134 checkpoint_config = dict(interval=1)
135 log_config = dict(interval=50, hooks=[dict(type='TextLoggerHook')])
136 dist_params = dict(backend='nccl')
137 log_level = 'INFO'
138
139 #Μη ορισμός συνέχισης/φόρτισης από προγενέστερο μοντέλο.
140 load_from = None
141 resume_from = None
142
143 workflow = [('train', 1)]
144 opencv_num_threads = 0
145 mp_start_method = 'fork'
146
147 #Φάκελος εργασίας.
148 work_dir = './work_dirs/resnet50_b32x8_MOT17'
149
150 gpu_ids = [0]
```

4.4: Παραμετροποίηση μοντέλου παρακολούθησης (Tracker)

Με ολοκληρωμένη την ανεξάρτητη εκπαίδευση των μοντέλων ανίχνευσης και επαναταυτοποίησης, απέμεινε μονάχα η δημιουργία ενός νέου αρχείου παραμετροποίησης MMTracking, το οποίο υλοποιούσε το τελικό μοντέλο DeepSORT, θέτοντας ως υπο-μοντέλα (sub-models) τον Faster R-CNN detector, και το BaseReID, φορτώνοντας τις αποθηκευμένες εκπαιδευμένες καταστάσεις τους. Το εν λόγω παραχθέν μοντέλο, δεν υποστήριζε κάποια εκ νέου «end-to-end» εκπαίδευση, αλλά ήταν άμεσα διαθέσιμο για την εξαγωγή συμπερασμάτων (inference), καθώς και για την ποσοτική αξιολόγησή του, δεδομένης της ύπαρξης δεδομένων ελέγχου. Παρακάτω παρουσιάζεται η παραμετροποίηση:

Απόσπασμα Κώδικα 3: Αρχείο παραμετροποίησης της δομής MMTracking, το οποίο υλοποιεί την τελική υλοποίηση παρακολούθησης DeepSORT, βασισμένο στα υπο-μοντέλα Faster R-CNN και BaseReID. Το μοντέλο διατίθεται για εξαγωγή συμπερασμάτων (inference) και ποσοτική αξιολόγηση στα πρότυπα του MOT Challenge. Στις υπερπαραμέτρους επαναταυτοποίησης έχει δοθεί επιπλέον βάρος ευαισθησίας στην τοποθεσία/θέση των στόχων, έναντι της εμφάνισης/μεγέθους τους (0.8 προς 0.2), καθώς πρόκειται για κινούμενα οχήματα. (πηγή πρωτότυπου: [32])

```
1 #Παραμετροποίηση του μοντέλου DeepSORT.
2 model = dict(
3     #Ορισμός υπο-μοντέλου ανίχνευσης.
4     detector=dict(
5         #####
6         #Εδώ τοποθετείται η παραμετροποίηση του Faster R-CNN μοντέλου που περιγράφηκε#
7         #νωρίτερα για την παροχή ανιχνύσεων.                                     #
8         #####
9
10        #Πλέον χρησιμοποιείται το checkpoint από την ανεξάρτητη εκπαίδευση του Faster R-CNN.
11        init_cfg=dict(
12            type='Pretrained',
13            checkpoint=
14                'path/to/faster-rcnn/epoch_30.pth'
15        )),
16        type='DeepSORT', #Επιλογή της υλοποίησης DeepSORT.
17        #Ορισμός υπο-μοντέλου κίνησης.
18        motion=dict(type='KalmanFilter', center_only=False), #Επιλογή του φίλτρου Kalman ως μοντέλο
19        κίνησης.
20        #Ορισμός υπο-μοντέλου επαναταυτοποίησης.
21        reid=dict(
22            #####
23            #Εδώ τοποθετείται η παραμετροποίηση του BaseReID μοντέλου που περιγράφηκε#
24            #νωρίτερα για την επαναταυτοποίηση των στόχων.                                     #
25            #####
26
27            #Πλέον χρησιμοποιείται το checkpoint από την ανεξάρτητη εκπαίδευση του BaseReID.
28            init_cfg=dict(
29                type='Pretrained',
30                checkpoint=
31                    'path/to/resnet-50-reid/epoch_10.pth'
32            )),
33        #Ορισμός υπερπαραμέτρων του tracker.
34        tracker=dict(
```

```
35     type='SortTracker', #Επιλογή του tracker τύπου SORT.
36     obj_score_thr=0.90, #Κατώφλι θετικών δειγμάτων.
37     #Ρυθμίσεις υπερπαραμέτρων αντιστοίχισης δειγμάτων (επαναταυτοποίησης).
38     reid=dict(
39         num_samples=150, #Αριθμός δειγμάτων.
40         img_scale=(256, 128), #Αλλαγή διαστάσεων δειγμάτων.
41         img_norm_cfg=None, #Μη κανονικοποίηση τιμών εικόνων.
42         match_score_thr=2.0), #Κατώφλι αντιστοίχισης.
43     match_iou_thr=0.2, #Τελικό κατώφλι αντιστοίχισης βάσει IoU.
44     #Ορισμός υπερπαραμέτρων ώστε να καθοριστεί η συμμετοχή της θέσης και
45     #της εμφάνισης στην αντιστοίχιση.
46     momentums=dict(
47         position=0.8, #Υπερπαραμέτρος θέσης.
48         size=0.2, #Υπερπαραμέτρος εμφάνισης.
49     ),
50     num_tentatives=4, #Ορισμός αριθμού δοκιμών/επαναλήψεων.
51     num_frames_retain=75) #Ορισμός μέγιστης διάρκειας σε πλήθος καρτέ (Αmax),
52     #χωρίς αντιστοίχιση τροχιάς στόχου και ανίχνευσης,
53     #κατά την οποία μια τροχιά διατηρείται "ζωντανή".
54
55 #Ορισμός format του dataset, και κύριας τοποθεσίας φακέλου.
56 dataset_type = 'MOTChallengeDataset' #Επιλέγεται το format του MOTChallenge dataset.
57 data_root = 'C:/Thesis/Datasets/tracking_annotations_cocoformat/'
58 img_norm_cfg = dict(
59     mean=[123.675, 116.28, 103.53], std=[58.395, 57.12, 57.375], to_rgb=True)
60     #Κανονικοποίηση και μετατροπή BGR-->RGB.
61
62 #Ορισμός ακολουθίας ελέγχου. Δε χρειάζεται ακολουθία εκπαίδευσης καθώς τα συστατικά
63 #υπο-μοντέλα εκπαιδεύονται ανεξάρτητα.
64 test_pipeline = [
65     dict(type='LoadImageFromFile'),
66     dict(
67         type='MultiScaleFlipAug',
68         img_scale=(1920, 1080),
69         flip=False,
70         transforms=[
71             dict(type='Resize', keep_ratio=True),
72             dict(type='RandomFlip'),
73             dict(
74                 type='Normalize',
75                 mean=[123.675, 116.28, 103.53],
76                 std=[58.395, 57.12, 57.375],
77                 to_rgb=True),
78             dict(type='Pad', size_divisor=32),
79             dict(type='ImageToTensor', keys=['img']),
80             dict(type='VideoCollect', keys=['img'])
81         ]
82     ]
83
84 #Παραμετροποίηση τροφοδότησης αλγορίθμου με τα δεδομένα προς επικύρωση κι έλεγχο.
85 data = dict(
86     samples_per_gpu=2,
87     workers_per_gpu=2,
88     val=dict(
89         type='MOTChallengeDataset',
90         ann_file=
```

```
91     'path/to/annotations.json',
92     img_prefix='path/to/images',
93     ref_img_sampler=None,
94     pipeline=test_pipeline),
95     test=dict(
96         type='MOTChallengeDataset',
97         ann_file=
98         'path/to/annotations.json',
99         img_prefix='path/to/images',
100        ref_img_sampler=None,
101        pipeline=test_pipeline))
102
103 #Επιλογή μετρικών αξιολόγησης.
104 evaluation = dict(metric=['bbox', 'track'], interval=1)
105 search_metrics = ['MOTA', 'IDF1', 'FN', 'FP', 'IDs', 'MT', 'ML']
106 #Ορισμός τοποθεσίας φακέλου εργασίας.
107 work_dir = './work_dirs/deepsort_faster-rcnn_fpn_4e_mot17-private-half'
```

Για να πραγματοποιηθεί inference σε μια ακολουθία βίντεο, ήταν απαραίτητη πρώτα η μετατροπή της σε COCOVideo format. Δηλαδή, δημιουργία ενός φακέλου με τοποθετημένα όλα τα καρέ σε χρονική σειρά, κι ενός αρχείου κειμένου «.json» με εγγραφές περιγραφών της μορφής:

```
{
  "videos": [
    {
      "id": 1,
      "name": "DJI_0073",
      "fps": 25,
      "width": 3840,
      "height": 2160
    }
  ],
  "categories": [
    {
      "id": 1,
      "name": "Motorcycle"
    },
    {
      "id": 2,
      "name": "Car"
    },
    {
      "id": 3,
      "name": "Semi_Truck/Van"
    },
    {
      "id": 4,
      "name": "Heavy_Truck/Bus"
    }
  ],
  "images": [
    {
      "id": 0,
      "video_id": 1,
      "file_name": "E:/SUNCITY_MOT_17_INFERENCE/DJI_0073/test/DJI_0073/img1\\000001.png",
      "height": 2160,
      "width": 3840,
      "frame_id": 0,
      "mot_frame_id": 1
    },
    . . .
    {
      "id": 7516,
      "video_id": 1,
      "file_name": "E:/SUNCITY_MOT_17_INFERENCE/DJI_0073/test/DJI_0073/img1\\007514.png",
      "height": 2160,
      "width": 3840,
      "frame_id": 7516,
      "mot_frame_id": 7517
    }
  ]
}
```

Σχήμα 29: Υποδείγματα εγγραφών περιγραφών στη μορφή COCOVideo.

Τα αποτελέσματα εξάγονταν σε αρχείο «pickle» format («.pkl»), και απαιτούσαν περαιτέρω επεξεργασία για οποιαδήποτε αξιοποίησή τους. Τα βίντεο που εξετάστηκαν για την απόκτηση των απαραίτητων στοιχείων για τη λειτουργική ανάλυση σε κάθε κόμβο, υπέστησαν εφαρμογή μασκών απομόνωσης των κύριων οδικών τμημάτων (πρακτικά «μαύρισμα» των περιοχών των οικοδομικών τετραγώνων και δευτερεύοντος οδικού δικτύου), προς διευκόλυνση των διεργασιών Υπολογιστικής Όρασης. Πρόκειται για τα παρακάτω:

- **Κυκλικός κόμβος Πλατείας 25ης Μαρτίου: DJI_0073.mp4 + DJI_0074.mp4 + DJI_0075.mp4**, διαδοχικά βίντεο συνολικής διάρκειας 12' 17".
- **Ηρώων Πολυτεχνείου – Παύλου Πεντάρη: DJI_0107.mp4**, διάρκειας 14' 18".
- **Πρωτόπαπα – Σοφοκλή Βενιζέλου: DJI_0560.mov**, διάρκειας 14' 19".

Οι συγκεκριμένες λήψεις συνέβησαν 8 π.μ., και επιλέχθηκαν γιατί η κυκλοφορία παρουσίαζε υψηλότερη αιχμή τη συγκεκριμένη ώρα, αν και βέβαια ο φωτισμός κρίνεται ελαφρώς δυσμενέστερος.

Ακόμη, για την ποσοτική αξιολόγηση του μοντέλου εξετάστηκαν οι εξής ακολουθίες:

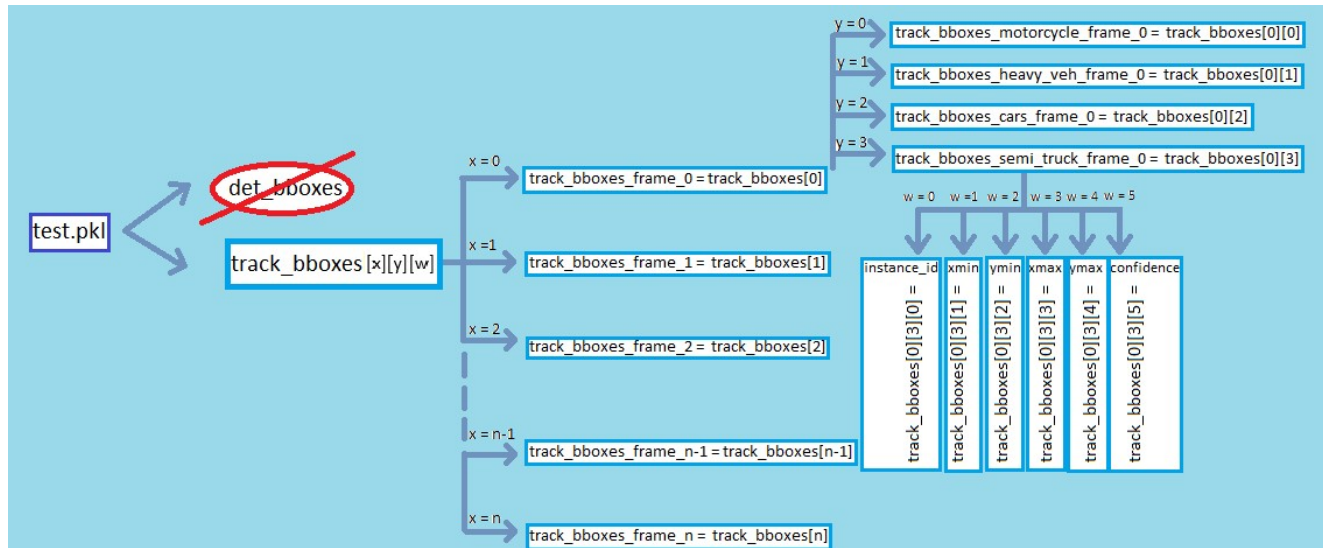
- **Κυκλικός κόμβος Πλατείας 25ης Μαρτίου: 30"** από το **DJI_0073.mp4**.
- **Ηρώων Πολυτεχνείου – Παύλου Πεντάρη: 30"** από **DJI_0095_1.mp4**.
- **Πρωτόπαπα – Σοφοκλή Βενιζέλου: Πλήρες DJI_0564.mp4**, διάρκειας **46"**.

4.5: Προσαρμοσμένοι αλγόριθμοι κυκλοφοριακής τεχνικής

Το μοντέλο παρακολούθησης πολλαπλών αντικειμένων προσέφερε ως τελικό «προϊόν» τα στοιχεία τροχιάς των οχημάτων, και μάλιστα όχι σε άμεσα αξιοποιήσιμη μορφή. Εκτός αυτού, για την κυκλοφοριακή προσέγγιση απαιτούνταν η σύσταση ενός αλγοριθμικού συστήματος το οποίο εισάγοντας τροχιές, θα μπορούσε να υπολογίσει κρίσιμα κυκλοφοριακά μεγέθη ώστε να αξιολογηθεί το επίπεδο εξυπηρέτησης των κόμβων. Για τις εφαρμογές λειτουργικής ανάλυσης αρκούσε μονάχα η εκτίμηση του κυκλοφοριακού φόρτου, αλλά ωστόσο εξετάστηκε και ο υπολογισμός ταχυτήτων διαδρομής, μέσης πυκνότητας και διαχωρισμών. Παρακάτω, περιγράφονται τα απαραίτητα προγράμματα κώδικα Python, τα οποία δημιουργήθηκαν για να επιτύχουν τους προαναφερθέντες σκοπούς και να ολοκληρώσουν το «ενεργό» σκέλος της μεθοδολογίας.

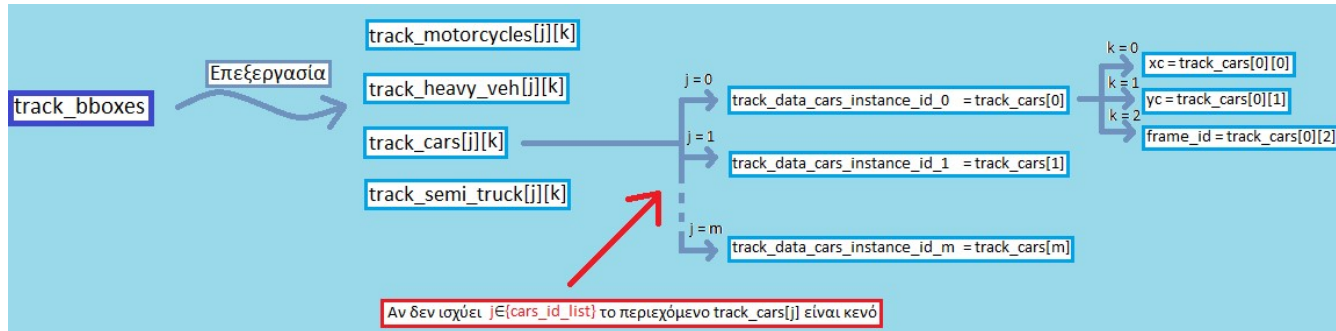
4.5.1: Συνάρτηση αξιοποίησης των δεδομένων εξόδου του tracker

Η μορφή «pickle» στην οποία εξήχθησαν τα αποτελέσματα του μοντέλου DeepSORT αποτελούσε μια σειριοποίηση (serialization) ενός συνόλου αντικειμένων, που για τη συγκεκριμένη περίπτωση πρόκειται για ένα σύνολο αντιστοιχίσεων κλειδιών – τιμών (dictionary). Η αρχική σύσταση του αρχείου διαγραμματικά είχε ως εξής:



Σχήμα 30: Το διάγραμμα της αποσυμπιεσμένης έκδοσης του αρχείου εξόδου μορφής pickle. Η τιμή x πράττει προσέλαση στα καρέ της ακολουθίας ελέγχου, η τιμή y διαχωρίζει τις εγγραφές - διανύσματα διαφορετικών κατηγοριών εντός ίδιου καρέ, και τέλος η τιμή w αποτελεί δείκτη για τα στοιχεία στιγμιαίας ανίχνευσης εντός κάθε εγγραφής της μορφής «(instance_id, xmin, ymin, xmax, ymax, confidence)».

Για να γίνει αξιοποιήσιμη η πληροφορία εξόδου στην εφαρμογή της κυκλοφοριακής τεχνικής, με «απομονωμένες» τις ζητούμενες τροχιές ανά id οχήματος, κρίθηκε απαραίτητη η επεξεργασία της ώστε να αποδίδονται τέσσερα (όσα και οι κατηγορίες οχημάτων) σύνολα με τις συντεταγμένες (xc,yc) των οχημάτων στα αντίστοιχα καρέ εμφάνισής τους (frame_id). Η επιθυμητή μετατροπή υποδεικνύεται σχηματικά παρακάτω:



Σχήμα 31: Περιγραφή της επιτυμητής προγραμματιστικής επεξεργασίας η οποία επιτυγχάνει τη μετατροπή του λεξικού στο Σχήμα 30, σε 4 παράγωγα λεξικά (1 για κάθε τύπο οχήματος) που θα φέρουν εγγραφές στοιχείων τροχιάς της μορφής «(xcenter, ycenter, frame_id)» με πρωτεύον κλειδί το instance_id.

Δημιουργήθηκε μια συνάρτηση «pkl2tracks» η οποία ανταποκρίνεται στην επεξεργασία που δηλώνεται παραπάνω για οποιοδήποτε δοθέν «.pkl» αρχείο. Αρχικά, πράττει ανάγνωση του «.pkl» αρχείου με χρήση της βιβλιοθήκης Pickle, και οι υποχώροι του «det_bboxes» και «track_bboxes» εκχωρούνται ως μεταβλητές:

Απόσπασμα Κώδικα 4: Ανάγνωση ενός αρχείου «.pkl» με την εισαγωγή της βιβλιοθήκης Pickle.

```
1 import pickle as pickle
2
3 def pkl2tracks(pkl_path):
4     with open(pkl_path, 'rb') as f:
5         data = pickle.load(f)
6
7     det_bboxes_data = data['det_bboxes']
8     track_bboxes_data = data['track_bboxes']
```

Σύμφωνα με το format του συνόλου «track_bboxes», με μια δομή επανάληψης (για κάθε καρτέ της εξεταζόμενης ακολουθίας) τα δεδομένα διαχωρίζονται σε 4 υποσύνολα ανταποκρινόμενων στις 4 διαφορετικές κατηγορίες οχημάτων:

Απόσπασμα Κώδικα 5: Διαχωρισμός των δεδομένων ανά κατηγορία οχήματος.

```
1 data_motorcycle = []
2 data_heavy_veh = []
3 data_car = []
4 data_semi_truck = []
5
6 i=0
7 for keys in track_bboxes_data:
8     data_motorcycle.append(track_bboxes_data[i][0])
9     data_heavy_veh.append(track_bboxes_data[i][1])
10    data_car.append(track_bboxes_data[i][2])
11    data_semi_truck.append(track_bboxes_data[i][3])
12    i+=1
```

Ύστερα, ακολουθείται μια μέθοδος που πραγματοποιείται κοινά για κάθε κατηγορία. Χρησιμοποιείται μια τριπλή δομή επανάληψης, όπου η «εξωτερική» αφορά την προσπέλαση

των καρτέ (frame_id), η «αμέσως ενδότερη» την εναλλαγή μεταξύ των id των στόχων (instance_id), ενώ η «εσωτερική» τη διαδοχή των διαφορετικών εγγραφών – ανιχνεύσεων που αφορούν έναν μοναδικό στόχο σε διαφορετικά καρτέ. Από τα στοιχεία εγγραφής (xmin, xmax, ymin, ymax, confidence) που αναφέρονται στο bounding box μιας ανίχνευσης, γίνεται μετασχηματισμός σε συντεταγμένες τροχιάς (xc, yc) που αντιστοιχούν στο κεντρικό σημείο του, ενώ γίνεται προσθήκη του frame_id το οποίο ισούται με τον αριθμό της εξωτερικής επανάληψης. Τα παραπάνω ισχύουν σε φιλτραρισμένες εγγραφές, βάσει μιας τιμής κατωφλίου confidence. Σχηματίζεται τελικά ένα dictionary λιστών με κλειδί το instance_id, με την εφαρμογή ενός κριτηρίου «if-else», ώστε να δημιουργείται νέα υπο-λίστα (sublist) κάθε φορά που αναλύεται ένα «πρωτοφανές» instance_id, και να ανανεώνει την υπο-λίστα «προϋπάρχοντος» instance_id. Ακόμη, πραγματοποιούνται ενέργειες ώστε το περιεχόμενο του dictionary να ταξινομείται κατά αύξουσα σειρά βάσει instance_id, ενώ το περιεχόμενο των υπο-λισταίων του κατά αύξουσα σειρά βάσει frame_id (δηλαδή χρονική σειρά). Παρακάτω φαίνεται ο φαίνεται ο σχετικός κώδικας με παράδειγμα την κατηγορία οχημάτων «Car» (εξαγωγή συνόλου track_cars):

Απόσπασμα Κώδικα 6: Η σύνθετη δομή επανάληψης για την εξαγωγή του συνόλου στοιχείων τροχιάς για την κατηγορία «Car»

```
1 track_cars = {}
2
3 for p in range(0, len(track_bboxes_data)):
4     for n in range(len(data_car[p])):
5         for entry in data_car[p]:
6             instance_id = int(entry[0])
7             xmin, ymin, xmax, ymax, confidence = entry[1:]
8             if confidence > 0.80:
9                 if instance_id in track_cars:
10                    track_cars[instance_id].append(((xmin+xmax)/2,
11                                                    (ymin+ymax)/2,
12                                                    p))
13                else:
14                    track_cars[instance_id] = [((xmin+xmax)/2,
15                                                (ymin+ymax)/2,
16                                                p)]
17
18 for keys in track_cars:
19     track_cars[keys] = list(set(track_cars[keys]))
20
21 for instance_id, track_data in track_cars.items():
22     sorted_list = sorted(track_data, key=lambda entry: entry[2])
23     track_cars[instance_id] = sorted_list
24
```

Η συνάρτηση επιστρέφει τα σύνολα τροχιών για κάθε διαφορετική κατηγορία οχημάτων:

Απόσπασμα Κώδικα 7: Επιστροφές τιμών μεταβλητής της συνάρτησης «rk12tracks».

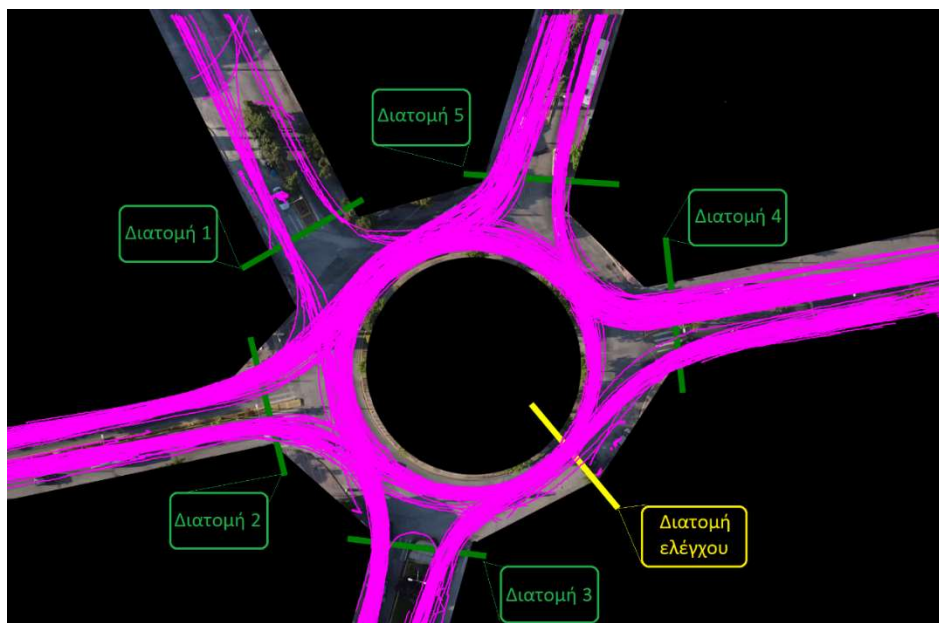
```
1 return track_motorcycle, track_heavy_veh, track_cars, track_semi_truck
```


4.5.2: Μονάδες εκτίμησης κυκλοφοριακών μεγεθών

Στη συνέχεια πραγματοποιήθηκε η δημιουργία του σημαντικότερου προγραμματιστικού τμήματος της εφαρμοσμένης μεθοδολογίας, καθώς πρόκειται για τις μονάδες συναρτήσεων που αφορούν τον υπολογισμό των πρωτογενών κυκλοφοριακών μεγεθών. Εφόσον κάθε οδικός κόμβος δύναται να αποτελείται από απρόβλεπτες ιδιομορφίες τοπολογίας, σε κάθε μονάδα σχηματίστηκαν διαφορετικές συναρτήσεις όμοιου αλγοριθμικού συλλογισμού, κάθε μια προσαρμοσμένη αποκλειστικά στον αντίστοιχο κόμβο μελέτης. Παρακάτω επεξηγούνται όλοι οι μηχανισμοί που αξιοποιούν τα στοιχεία τροχιάς των οχημάτων για τον προαναφερθέντα σκοπό, ενώ παρεμβάλλονται καίρια αποσπάσματα από τους σχετικούς κώδικες.

4.5.2.1: Καταμέτρηση διελεύσεων - κυκλοφοριακός φόρτος

Όσον αφορά την εκτίμηση του απαραίτητου για τη λειτουργική ανάλυση μεγέθους, του κυκλοφοριακού φόρτου, αυτός πρόκυπτε μέσα από την καταμέτρηση των διερχόμενων οχημάτων, η οποία εξαρτιόταν από την τομή των τροχιών τους με διατομές προσβάσεων κι εξόδου. Επομένως, αρχικά ορίστηκε από ένα σταθερό ευθύγραμμο τμήμα στην απεικόνιση κάθε κλάδου, το οποίο λειτουργούσε ως διατομή αναφοράς αποσκοπώντας στην συμπερίληψη τόσο των εισερχόμενων όσο και των εξερχόμενων οχημάτων στον κόμβο. Προτιμήθηκε έναντι ορισμού ξεχωριστών διατομών εισόδου και εξόδου, καθώς τα πλάνα λήψης των δεδομένων μελέτης δεν ήταν απόλυτα στατικά, και ενδέχεται να δημιουργούνταν συγχύσεις μεταξύ τους εξαιτίας των μετατοπίσεων. Επιπρόσθετα, ορίστηκε κι ένα ακόμη σταθερό ευθύγραμμο τμήμα ελέγχου εντός της φυσικής περιοχής του κόμβου.



Σχήμα 32: Παρουσίαση των διατομών καταμέτρησης στο κυκλικό κόμβο της Πλατείας 25^{ης} Μαρτίου. Με ματζέντα αποτυπώνεται η αθροιστική οπτικοποίηση των «έγκυρων» τροχιών που προήλθαν από το DJI_0073.mp4, οι οποίες φαίνεται ότι τέμνονται ορθά με τις διατομές.



Σχήμα 33: Παρουσίαση των διατομών καταμέτρησης στους κόμβους Ηρώων Πολυτεχνείου – Παύλου Πεντάρη (άνω) και Πρωτόπαπα – Σοφοκλή Βενιζέλου (κάτω). Αντίστοιχα, με ματζέντα αποτυπώνεται η αθροιστική οπτικοποίηση των «έγκυρων» τροχιών που προήλθαν από τα DJI_0107.mp4 και DJI_0560.mov.

Στην υλοποίηση κώδικα, τα ευθύγραμμα τμήματα διατομών συντάχθηκαν στην αρχή κάθε συνάρτησης στη μορφή LineString της βιβλιοθήκης Shapely, με ορίσματα 2 σημείων.

Απόσπασμα Κώδικα 8: Ορισμός των διατομών καταμέτρησης σε μορφή LineString της βιβλιοθήκης Shapely.

```
1 Line_0 = LineString([(2150, 1200), (2400, 1500)])
2 Line_1 = LineString([(1280, 780), (1630, 580)])
3 Line_2 = LineString([(1300, 1000), (1400, 1400)])
4 Line_3 = LineString([(1600, 1600), (2000, 1650)])
5 Line_4 = LineString([(2550, 700), (2600, 1150)])
6 Line_5 = LineString([(1950, 500), (2400, 530)])
```

Ως «Line_0» ορίστηκε η διατομή ελέγχου, ο ρόλος της οποίας είναι βοηθητικός απέναντι στην αντιμετώπιση ειδικών περιπτώσεων σφαλμάτων που επεξηγούνται παρακάτω. Ακολουθεί ο βασικός μηχανισμός υλοποίησης τροχιών και καταμέτρησης διελεύσεων, ο οποίος χρησιμοποιείται σε κάθε λεξικό διαφορετικού τύπου οχήματος:

Απόσπασμα Κώδικα 9: Ο βασικός μηχανισμός υλοποίησης τροχιών και καταμέτρησης διελεύσεων.

```
1 veh_intersections=[]
2
3 for p in track_veh:
4
5     new_instance_intersections=[]
6     instance_intersections=[]
7
8     for i in range(len(track_veh[p])-1):
9         if len(track_veh[p]) > 10:
10            track_miniline = LineString([(track_veh[p][i][0],
11                                           track_veh[p][i][1]),
12                                          (track_veh[p][i+1][0],
13                                           track_veh[p][i+1][1])])
14            if track_miniline.length < 300:
15                checks = [track_miniline.intersects(Line_0),
16                          track_miniline.intersects(Line_1),
17                          track_miniline.intersects(Line_2),
18                          track_miniline.intersects(Line_3),
19                          track_miniline.intersects(Line_4),
20                          track_miniline.intersects(Line_5)]
21
22            for j in range(len(checks)):
23                if checks[j] is True:
24                    instance_intersections.append(j)
25
26            new_instance_intersections=[]
27            for value in instance_intersections:
28                if not new_instance_intersections or (value !=
29                                                       new_instance_intersections[-1]):
30                    new_instance_intersections.append(value)
31            veh_intersections.append(new_instance_intersections)
32
33 veh_final_intersections = [sublist for sublist in veh_intersections if sublist]
```

Δημιουργείται μια δομή επανάληψης που αναλαμβάνει την προσπέλαση όλων των τροχιών οχημάτων. Σε κάθε τροχιά εφαρμόζεται μια ακόμη εσωτερική δομή επανάληψης που σε κάθε βήμα της δημιουργεί ένα ευθύγραμμο τμήμα μεταξύ δύο διαδοχικών σημείων της, το οποίο διέρχεται μέσω boolean τοπολογικού ελέγχου τομής («intersects()») με κάθε μια από τις ορισθείσες διατομές. Η ύπαρξη «True» σε έναν τέτοιο έλεγχο αποδίδει τον κωδικό της διατομής (0 για την ελέγχου, και 1 έως 4 ή 5 για τις βασικές) σε μια λίστα που περιγράφει όλες τις τομές μιας τροχιάς. Επειδή το μήκος των ευθυγράμμων τμημάτων είναι αρκετά μικρό και ευαίσθητο απέναντι στις μετακινήσεις της κάμερας, υπάρχει περίπτωση «κοντινά» ευθύγραμμα τμήματα να τέμνουν την ίδια στατική διατομή πολλαπλές φορές. Για το λόγο αυτό, πραγματοποιείται διαγραφή των διαδοχικών διπλότυπων κωδικών σε μια φιλτραρισμένη νέα λίστα.

Εφόσον τα στοιχεία τροχιάς ενός στόχου είναι ταξινομημένα βάσει του άξοντος `frame_id`, η σειρά δημιουργίας των εν λόγω ευθυγράμμων τμημάτων ταυτίζεται με τη χρονική σειρά της πορείας του, οπότε και οι έλεγχοι τομής προκύπτουν με σαφή χρονική ακολουθία. Η ανάγκη για προσδιορισμό τομών προορισμού και προέλευσης οδήγησε στη «διαφορική» αντιμετώπιση των τροχιών με διακριτά ευθύγραμμα τμήματα, και όχι στη μετατροπή τους σε ενιαία `LineString`. Επίσης, τονίζεται ότι αποκλείονται του παραπάνω μηχανισμού τροχιές αποτελούμενες από λιγότερα από 10 σημεία, αλλά και ευθύγραμμα τμήματα μήκους μεγαλύτερου από 300 εικονοστοιχεία, καθώς αυτές οι περιπτώσεις αποτελούν εκφάνσεις «θορύβου» του μοντέλου παρακολούθησης. Συνεπώς, η τελική λίστα για τις τομές μιας τροχιάς μπορεί να είναι της επιθυμητής τυπικής μορφής «[1,2], [2,0,1], [4], [3,0,5]», κ.ο.κ., αλλά φυσικά ενδέχεται να είναι και κενή. Το σύνολο των μη κενών λιστών τοποθετούνται ως υπο-λίστες στην τελική «veh_final_intersections» λίστα.

Έπειτα, ανάλογα με όλους τους πιθανούς αποδεκτούς συνδυασμούς προστίθεται μια καταμέτρηση διέλευσης στο αντίστοιχο κελί ενός (NxN) πίνακα προέλευσης – προορισμού, όπου N ο αριθμός των κλάδων. Η συγκεκριμένη ανάθεση πραγματοποιείται με τη χρήση πολλαπλών `if` προτάσεων:

Απόσπασμα Κώδικα 10: Η ανάθεση των καταμετρήσεων διελεύσεων σε πίνακα προέλευσης – προορισμού.

```
1     veh_volumes = np.zeros((5,5))
2
3     for sublist in veh_final_intersections:
4
5         #Με προέλευση την πρόσβαση 1
6         if sublist == [1, 2]:
7             veh_volumes[1,0]+=1
8         if sublist == [1, 3]:
9             veh_volumes[2,0]+=1
10        if sublist == [1, 0, 4]:
11            veh_volumes[3,0]+=1
12        if sublist == [1, 0, 5]:
13            veh_volumes[4,0]+=1
```

Έτσι, κατά μήκος μιας γραμμής εμφανίζονται τα οχήματα που καταλήγουν σε έναν κλάδο, ενώ κατά μήκος μια στήλης τα οχήματα που ξεκινούν από μια πρόσβαση κλάδου. Στην κύρια διαγώνιο απαριθμούνται οχήματα που εισέρχονται κι εξέρχονται του ίδιου κλάδου, συνθήκη που κανονικά δύναται να συμβεί μόνο σε κυκλικό κόμβο. Η προσθήκη της διατομής ελέγχου μεταξύ δύο οποιονδήποτε προσβάσεων μπορεί εύκολα να αναγνωρίσει τις «U-turns», καθώς αποδίδονται υπο-λίστες τομής της μορφής «[X,0,X]».

Απόσπασμα Κώδικα 11: Η ανάθεση των καταμετρήσεων U-turn στην κύρια διαγώνιο πίνακα προέλευσης – προορισμού.

```
1      #Για ανίχνευση U-turns
2      if sublist == [1, 0, 1]:
3          veh_volumes[0,0]+=1
4      if sublist == [2, 0, 2]:
5          veh_volumes[1,1]+=1
6      if sublist == [3, 0, 3]:
7          veh_volumes[2,2]+=1
8      if sublist == [4, 0, 4]:
9          veh_volumes[3,3]+=1
10     if sublist == [5, 0, 5]:
11         veh_volumes[4,4]+=1
```

Επιπρόσθετα, η διατομή ελέγχου βοηθούσε στη διαχείριση «ιδιόρρυθμων» τροχιών, λόγω κατακερματισμών αλλά και «μεταφερόμενων» IDs που αποτελούν σφάλματα του μοντέλου παρακολούθησης. Στην παρούσα μεθοδολογία, μια τροχιά χαρακτηρίστηκε ως ιδιόρρυθμη εφόσον δεν τεμόταν με 2 βασικές διατομές, και όπου απαιτείται και με τη διατομή ελέγχου, οπότε για αυτές διακρίθηκαν ειδικές περιπτώσεις:

- Αν μια τροχιά δεν τέμνει κάποια από τις διατομές δεν χρειάζεται κάποια περαιτέρω ενέργεια, καθώς δε συμμετέχει στο μηχανισμό.
- Αν μια τροχιά τέμνει μια βασική διατομή (υπο-λίστα τομής [X], ή [0,X]/[X,0]), τότε θεωρείται ότι θα υπάρχει άλλη μια τροχιά, διαφορετικού ID, η οποία θα τέμνει μια άλλη βασική διατομή (υπο-λίστα τομής [Y], ή [0,Y]/[Y,0]) και θα αναφέρεται στην πραγματικότητα στον ίδιο στόχο. Το «άθροισμα» των δύο αυτών εγγραφών ουσιαστικά ισοδυναμεί με μια κανονική καταμέτρηση διέλευσης οχήματος, οπότε κάθε τομή μιας βασικής διατομής θεωρείται «μισή» διέλευση.
- Αν μια τροχιά τέμνει 2 βασικές διατομές και την τομή ελέγχου, αλλά το 0 δεν είναι το ενδιάμεσο στοιχείο της υπο-λίστας:
 - [X,Y,0]: Θεωρείται ότι μετά από μια ολοκληρωμένη διέλευση, κατά την έξοδο του οχήματος από το πλάνο έχει πραγματοποιηθεί μεταφορά ID σε ένα αντίθετα εισερχόμενο όχημα, οπότε και διανύεται μια επιπλέον «μισή» διέλευση μέχρι την τομή με τη διατομή ελέγχου.

- [0,X,Y]: Αντίστροφα, μετά από «μισή» διέλευση και μεταφορά ID, διανύεται μια πλήρης διέλευση μέχρι την τομή της τροχιάς με μια διαφορετική διατομή.
- Αν μια τροχιά αποδίδει υπο-λίστα τομής [X,Y,Z] (ή και [X,Y,X]) σημαίνει ότι λόγω μεταφοράς ID έχει πραγματοποιήσει 2 πλήρεις διελεύσεις.
- Αν μια υπο-λίστα τομής φέρει παραπάνω από 3 κωδικούς (εξαιρετικά απίθανο, συνήθως 4), τότε κατά σύμβαση θεωρείται πως η τροχιά έχει διανύσει το ισοδύναμο 2 διελεύσεων.

Οι συγκεκριμένες περιπτώσεις που προέκυψαν από τα σφάλματα ταυτοποίησης του μοντέλου ανίχνευσης αποτελούν μειοψηφία μεν, μη αμελητέα δε. Σε αρκετές εξ αυτών δεν ήταν δυνατόν να γίνει ορθολογικά προσμέτρησή τους σε κάποιο ζεύγος προέλευσης – προορισμού, καθώς απουσιάζει η απαραίτητη πληροφορία ή καθίσταται ασαφής. Πρόκειται κυρίως για τα συμβάντα όπου απέμεναν «μισές» διελεύσεις, αλλά και για τη σπάνια περίπτωση υπο-λίστας τομής > 3 στοιχείων, αντίστοιχα. Χρησιμοποιώντας μερικά επιπλέον if statements, οι ειδικές περιπτώσεις αντιμετωπίστηκαν κατάλληλα, ενώ οπουδήποτε πρόκλυπε αοριστία καταχωρούνταν σε μια μεταβλητή το ισοδύναμο ποσό ασαφούς κυκλοφοριακού φόρτου (ambiguous volume), το οποίο τελικά μοιράστηκε σταθμισμένα στα κελιά του πίνακα προέλευσης προορισμού, «ευνοώντας» τις συχνότερες διαδρομές (από τις ορθολογικές διελεύσεις).

Απόσπασμα Κώδικα 12: Διαχείριση των συνθηκών «ιδιόρρυθμων» τροχιών.

```
1     if len(sublist) == 1:
2         ambiguous_volume+=0.5
3     if len(sublist) == 2 and 0 in sublist:
4         ambiguous_volume+=0.5
5     if len(sublist) == 3 and sublist[0] == 0 and sublist[2] != 0:
6         veh_volumes[(sublist[2]-1), (sublist[1]-1)]+=1
7         ambiguous_volume+=0.5
8     if len(sublist) == 3 and sublist[2] == 0 and sublist[0] != 0:
9         veh_volumes[(sublist[1]-1), (sublist[0]-1)]+=1
10        ambiguous_volume+=0.5
11    if len(sublist) == 3 and 0 not in sublist:
12        veh_volumes[(sublist[1]-1), (sublist[0]-1)]+=1
13        veh_volumes[(sublist[2]-1), (sublist[1]-1)]+=1
14    if len(sublist) > 3:
15        ambiguous_volume+=2.0
16
17    weights = (veh_volumes/(1+sum(sum(veh_volumes))))
18    final_veh_volumes=(weights*ambiguous_volume)+veh_volumes).round()
```

Τονίζεται πως θα μπορούσε στους κόμβους φωτεινής σηματοδότησης να οριστούν δύο ή περισσότερες διατομές ελέγχου, οι οποίες θα απέδιδαν κοινά την τιμή 0 στην υπο-λίστα τομών, αλλά ο ορισμός μονάχα μιας που θα τέμνει χωρικά τη δεσπόζουσα ροή είναι υπεραρκετός για τη συμπερίληψη της συντριπτικής πλειοψηφίας των σφαλμάτων μεταφερόμενων IDs.

4.5.2.2: Ταχύτητα διαδρομής και διάνυση μήκους συναρτήσει καρτέ

Στη μονάδα καταμέτρησης διελεύσεων ενσωματώθηκαν υπολογισμοί οι οποίοι επιτελούσαν τον προσδιορισμό ταχυτήτων διαδρομής. Αρχικά, προστίθενται παράμετροι που αφορούν τις συνθήκες και τον εξοπλισμό λήψης, και υπολογίζονται στοιχεία που χαρακτηρίζουν το μέγεθος εδαφοψηφίδας, δηλαδή την αντιστοιχία εικονοστοιχείου σε μονάδα πραγματικού μήκους στο απεικονιζόμενο περιβάλλον.

Απόσπασμα Κώδικα 13: Παράμετροι τεχνικών συνθηκών λήψης, και υπολογισμός μεγέθους εδαφοψηφίδας

```
1 half_diagon = np.sqrt(pow(3840,2)+pow(2160,2))/2
2 fov_half_diagonal = (84/2)*np.pi/180
3 frame_rate = 25
4 altitude = 163
5 half_diagon_meters = altitude*np.tan(fov_half_diagonal)
6 pix2meter = half_diagon_meters/half_diagon
7 meter2pix = 1/pix2meter
```

Εντός της «εσωτερικής» δομής επανάληψης που αναφέρεται σε κάθε τροχιά, πραγματοποιείται ο υπολογισμός της μέσης ταχύτητας του στόχου για τη διάνυση ενός ευθυγράμμου τμήματος, διαιρώντας το μήκος του με τη διαφορά των frame_id, με μετατροπή της σε χλμ./ώρα. Όλες αυτές οι ταχύτητες τοποθετούνται σε ένα dictionary αναφερόμενες σε συγκεκριμένο instance_id, ενώ τελικά από το μέσο όρο τους προκύπτει για κάθε στόχο η ταχύτητα διαδρομής.

Απόσπασμα Κώδικα 14: Προσθήκη των απαραίτητων ρυθμίσεων στο βασικό μηχανισμό υλοποίησης τροχιών που εξηγήθηκε προηγουμένως, ώστε να καθίσταται εφικτός ο υπολογισμός ταχυτήτων διαδρομής, και διάνυσης μήκους συναρτήσει καρτέ.

```
1 instance_miniline_velocities = {}
2 mean_instance_velocities = {}
3 track_length = {}
4
5 for p in track_veh:
6
7     . . .
8
9     for i in range(len(track_veh[p])-1):
10
11         . . .
12
13             time_frame = track_veh[p][i][2]
14             miniline_velocity = 3.6*((track_miniline.length*pix2meter)/
15                                     ((track_veh[p][i+1][2]-time_frame)/
16                                     frame_rate))
17
18             if p in instance_miniline_velocities:
19                 instance_miniline_velocities[p].append((miniline_velocity))
20             else:
21                 instance_miniline_velocities[p] = [(miniline_velocity)]
```

```
22         if p in track_length:
23             track_length[p].append((track_miniline.length*pix2meter,
24                                     time_frame))
25         else:
26             track_length[p] = [(track_miniline.length*pix2meter,
27                                 time_frame)]
28     . . .
29
30
31     if len(track_veh[p]) > 10:
32
33         if p in mean_instance_velocities:
34             mean_instance_velocities[p].append((sum(instance_miniline_velocities[p])/
35                                                    len(instance_miniline_velocities[p])))
36         else:
37             mean_instance_velocities[p] = (sum(instance_miniline_velocities[p])/
38                                             len(instance_miniline_velocities[p]))
```

Το σύνολο «track_length» περιέχει την πληροφορία για το μήκος που διανύεται από τον στόχο με την πάροδο των καρτέ της ακολουθίας. Μπορεί να χρησιμοποιηθεί για τον υπολογισμό του συνολικού μήκους διαδρομής κάθε τροχιάς, και για το σχηματισμό διαγραμμάτων θέσης – χρόνου (x,t). Τελικώς, κάθε μια προσαρμοσμένη στον κόμβο συνάρτηση, επιστρέφει (για κάθε κατηγορία οχημάτων) τον πίνακα προέλευσης προορισμού, τον ασαφή κυκλοφοριακό φόρτο, τις «μέσες στιγμιαίες» και ταχύτητες διαδρομής των στόχων, καθώς και το σύνολο διάνυσης απόστασης συναρτήσει των ανάλογων καρτέ.

4.5.2.3: Μέση πυκνότητα, χωρικός και χρονικός διαχωρισμός

Περαιτέρω, σε ξεχωριστή μονάδα αρχείου κώδικα υλοποιήθηκαν συναρτήσεις εκτίμησης μέσης πυκνότητας, με αναγωγή σε ΜΕΑ ανά μήκος ενός χιλιομέτρου. Πρόκειται για μια απλή καταμέτρηση των συνολικών ανιχνεύσεων σε κάθε καρτέ, και εξαγωγή του μέσου όρου τους καθ' όλη την ακολουθία. Η είσοδος σε κάθε συνάρτηση είναι το «.pkl» αρχείο τροχιών, το οποίο διαβάζεται και γίνεται ο διαχωρισμός των δεδομένων σε κατηγορίες οχημάτων (data_car, data_motorcycle, κ.ο.κ.), όπως στη συνάρτηση pkl2tracks().

Απόσπασμα Κώδικα 15: Εύρεση μέσης πυκνότητας, ενδεικτικά από το σύνολο τροχιών κατηγορίας «Car».

```
1     densities_per_frame_car = []
2
3     for p in range(len(data_car)):
4         vehicles_per_frame_car = data_car[p].shape[0]
5         density_car = np.ceil(vehicles_per_frame_car/L)
6         densities_per_frame_car.append(density_car)
7
8     mean_density_car = sum(densities_per_frame_car)/len(densities_per_frame_car)
```


Ενώ χρησιμοποιήθηκε η ίδια συνάρτηση για τους ισόπεδους κόμβους Φ/Σ, για την περίπτωση του κυκλικού κόμβου δημιουργήθηκε μια ξεχωριστή συνάρτηση που προσαρμόζεται στην εύρεση της μέσης πυκνότητας εντός του δακτυλίου, αποκλείοντας τις ανιχνεύσεις εκτός μιας buffer ζώνης από το κέντρο του, χρησιμοποιώντας τον τοπολογικό έλεγχο «.contains» όπως φαίνεται παρακάτω:

Απόσπασμα Κώδικα 16: Διαφοροποίηση στην εκτίμηση της μέσης πυκνότητας εντός του δακτυλίου κυκλικού κόμβου (ενδεικτικά για την κατηγορία «Car»).

```
1     x_c = 1975
2     y_c = 1050
3     center = Point(x_c, y_c)
4     radius = 600
5     roundabout = center.buffer(radius)
6
7     densities_per_frame_car = []
8
9     for p in range(len(data_car)):
10        coord_list = []
11        for k in range(len(data_car[p])):
12            coords = (data_car[p][k][1]+
13                    data_car[p][k][2])/2, (data_car[p][k][2]+
14                    data_car[p][k][4])/2
15            if roundabout.contains(Point(coords)):
16                coord_list.append(coords)
17        vehicles_per_frame_car = len(coord_list)
18        density_car = np.ceil(vehicles_per_frame_car/L)
19        densities_per_frame_car.append(density_car)
20
21    mean_density_car = sum(densities_per_frame_car)/len(densities_per_frame_car)
```

Η συνολική μέση πυκνότητα προκύπτει από το άθροισμα των επιμέρους μέσων πυκνοτήτων για τις διαφορετικές κατηγορίες οχημάτων. Από τη συνολική μέση πυκνότητα, εφόσον το μήκος αναφοράς είναι το 1 χιλιόμετρο, δύναται να υπολογιστεί ο μέσος χωρικός διαχωρισμός, ως το αντίστροφό της. Με γνωστή τη μέση ταχύτητα διαδρομής ως διαιρέτης του μέσου χωρικού διαχωρισμού, μπορεί να γίνει εκτίμηση για το μέσο χρονικό διαχωρισμό.

4.5.3: Μονάδα χαρακτηρισμού επιπέδου εξυπηρέτησης

Στη μονάδα χαρακτηρισμού επιπέδου εξυπηρέτησης κλήθηκαν οι συναρτήσεις των μονάδων εκτίμησης κυκλοφοριακών μεγεθών, ώστε να αποδοθούν τα απαραίτητα δεδομένα για την εκτέλεση των υπολογισμών λειτουργικής ανάλυσης στους κυκλοφοριακούς κόμβους.

Αρχικά, κρίνεται απαραίτητο να σχηματιστεί ένας συνολικός πίνακας προέλευσης – προορισμού ο οποίος θα περιέχει τη ζήτηση υπό τη μορφή ωριαίου κυκλοφοριακού φόρτου από και προς τους διαφορετικούς κλάδους. Μετά την κλήση των συναρτήσεων, αθροίζονται τα αντίστοιχα κελιά από τους επιμέρους πίνακες εξόδου από διαφορετικές κατηγορίες οχημάτων, αφού έχει γίνει αναγωγή των τιμών τους στην ενιαία Μονάδα Επιβατικών Αυτοκινήτων (ΜΕΑ) με τον πολλαπλασιασμό τους με τον κατάλληλο συντελεστή ισοδυναμίας. Μετά την προβολή της ζήτησης σε ωριαία βάση, δεν έχει νόημα η χρήση του ΣΩΑ, διότι οι παρατηρήσεις σε κάθε κόμβο διαρκούν ελαφρώς λιγότερο από 15 λεπτά. Ενδεικτικά, παρατίθενται τα αποσπάσματα κώδικα από τη διαχείριση των τροχιών της ακολουθίας DJI_0107.mp4:

Απόσπασμα Κώδικα 17: Κλήση των απαραίτητων συναρτήσεων και αναγωγή σε ΜΕΑ και ωριαία βάση αναφοράς, ώστε να σχηματιστεί ο πίνακας προέλευσης-προορισμού που δηλώνει τον ωριαίο κυκλοφοριακό φόρτο κάθε διαδρομής.

```
1   pkl_path_0107 = 'E:/SUNCITY_INFERENCE_videos/test/evaluation/test_DJI_0107.pkl'
2   (track_motorcycle_0107, track_heavy_veh_0107, track_cars_0107,
3     track_semi_truck_0107) = pkl2tracks(pkl_path_0107)
4
5   (car_volumes_0107, ambiguous_car_0107,
6     _, _, _) = traffic_volume_normal_1(track_cars_0107)
7   (motorcycle_volumes_0107, ambiguous_motorcycle_0107,
8     _, _, _) = traffic_volume_normal_1(track_motorcycle_0107)
9   (heavy_veh_volumes_0107, ambiguous_heavy_veh_0107,
10    _, _, _) = traffic_volume_normal_1(track_heavy_veh_0107)
11  (semi_truck_volumes_0107, ambiguous_semi_truck_0107,
12    _, _, _) = traffic_volume_normal_1(track_semi_truck_0107)
13
14  car_pcu_0107 = car_volumes_0107
15  motorcycle_pcu_0107 = 0.5*motorcycle_volumes_0107
16  semi_truck_pcu_0107 = 1.5*semi_truck_volumes_0107
17  heavy_veh_pcu_0107 = 2.0*heavy_veh_volumes_0107
18  pcu_0107 = (car_pcu_0107 + motorcycle_pcu_0107 + semi_truck_pcu_0107 +
19             heavy_veh_pcu_0107)
20
21  Volume_pcu_hour_normal_1 = np.ceil((60/(14+18/60))*(pcu_0107))
```

Ο «ασαφής» κυκλοφοριακός φόρτος καταχωρείται σε μεταβλητές επειδή χρησιμοποιείται για την εξαγωγή μετρικής αξιολόγησης, όπως θα εξηγηθεί στο επόμενο υποκεφάλαιο.

Ύστερα, αφού έχει πραγματοποιηθεί η διάκριση των επιμέρους κυκλοφοριακών ρευμάτων (ή λωρίδων στον κυκλικό κόμβο), γίνεται η εκχώρηση της ζήτησης που τους αντιστοιχεί από

το συνολικό πίνακα προέλευσης – προορισμού. Ακόμη, όσον αφορά τους κόμβους φωτεινής σηματοδότησης εισάγονται ως ανεξάρτητες παράμετροι οι διάρκειες κύκλου και ενεργού πρασίνου, όπως αυτές προσδιορίζονται από την παρατήρηση της κυκλοφορίας στα βίντεο.

Απόσπασμα Κώδικα 18: Κατανομή της ζήτησης στα διακριθέντα ρεύματα (ή λωρίδες) κυκλοφορίας. Στους κόμβους φωτεινής σηματοδότησης γίνεται ορισμός των στοιχείων του υφιστάμενου προγράμματος.

```
1 Vol1_1 = (Volume_pcu_hour_normal_1[1,0] + Volume_pcu_hour_normal_1[2,0] +
2         Volume_pcu_hour_normal_1[3,0])
3 Vol1_2 = (Volume_pcu_hour_normal_1[2,1] + Volume_pcu_hour_normal_1[3,1])
4 Vol1_3 = (Volume_pcu_hour_normal_1[0,2] + Volume_pcu_hour_normal_1[1,2] +
5         Volume_pcu_hour_normal_1[3,2])
6 Vol1_4 = (Volume_pcu_hour_normal_1[1,3] + Volume_pcu_hour_normal_1[2,3])
7 Vol1_5 = (Volume_pcu_hour_normal_1[0,3])
8
9 C1 = 90
10 g1_1 = 54
11 g1_2 = 22
12 g1_3 = 14
13 g1_4 = 22
14 g1_5 = 76
```

Ακολουθούν οι τυπικοί υπολογισμοί που περιγράφονται στο κεφάλαιο 2.2.3 για τον προσδιορισμό των μέσων καθυστερήσεων των ρευμάτων (ή λωρίδων), ώστε να γίνει αναγωγή τους σε επίπεδο πρόσβασης, και τελικά σε επίπεδο ολόκληρου του κόμβου.

Απόσπασμα Κώδικα 19: Πραγματοποίηση των απαραίτητων υπολογισμών κατά HCM 2010 για τον προσδιορισμό των μέσων καθυστερήσεων (προφανώς, για τον κυκλικό κόμβο ακολουθούνται διαφορετικοί υπολογισμοί, που αναφέρονται στην αντίστοιχη σχετική μεθοδολογία λειτουργικής ανάλυσης).

```
1 So = 1900
2 N = 2
3 w = 3.50
4 G = -6
5 peds = 50
6 fw = (1 + ((w - 3.60) / 900))
7 fHV = 100 / (100 + (semi_truck_pcu_0107[1,0] + heavy_veh_pcu_0107[1,0] +
8         semi_truck_pcu_0107[2,0] + heavy_veh_pcu_0107[2,0] +
9         semi_truck_pcu_0107[3,0] + heavy_veh_pcu_0107[3,0]) / Vol1_1)
10 fG = 1 - (G / 200)
11 fp = 1
12 fa = 1
13 fRT = 1 - (Volume_pcu_hour_normal_1[1,0] / Vol1_1) * (0.135 + (peds / 2100))
14 fLT = 1 / (1 + 0.05 * (Volume_pcu_hour_normal_1[3,0] / Vol1_1))
15 S1_1 = So * N * fw * fHV * fG * fp * fa * fRT * fLT
16 c1_1 = S1_1 * g1_1 / C1
17 X1_1 = Vol1_1 / c1_1
18 Xmin1_1 = min({X1_1, 1})
19 m = 12
20 DF = 1
21 d1 = (0.38 * C1 * pow((1 - g1_1 / C1), 2)) / (1 - Xmin1_1 * g1_1 / C1)
22 d2 = 173 * pow(X1_1, 2) * ((X1_1 - 1) + pow((pow((X1_1 - 1), 2) + m * (X1_1 / C1)), 0.5))
23 d1_1 = d1 * DF + d2
```

```
24
25 #. . .
26 #Ομοίως για κάθε ρεύμα (ή λωρίδα)
27 #. . .
28
29 d1_access_1 = d1_1
30 d1_access_2 = d1_2
31 d1_access_3 = d1_3
32 d1_access_4 = (d1_4*Vol1_4 + d1_5*Vol1_5)/(Vol1_4 + Vol1_5)
33
34 d_normal_1 = (d1_access_1*Vol1_1 + d1_access_2*Vol1_2 + d1_access_3*Vol1_3 +
35              d1_access_4*(Vol1_4 + Vol1_5))/(Vol1_1 + Vol1_2 + Vol1_3 + Vol1_4 + Vol1_5)
```

Τέλος, το αποτέλεσμα χαρακτηρισμού σε στάθμη εξυπηρέτησης εξάγεται από μια κοινή συνάρτηση που δέχεται ως είσοδο μια λίστα τιμών μέσης καθυστέρησης, και μια μεταβλητή Boolean για την υπόδειξη του τύπου του κόμβου (ισόπεδος Φ/Σ ή κυκλικός), ώστε να εφαρμοστούν τα κατάλληλα όρια τιμών.

Απόσπασμα Κώδικα 20: Συνάρτηση χαρακτηρισμού επιπέδου εξυπηρέτησης.

```
1 def level_of_service(d_list, traffic_light_junction):
2
3     result_los=[]
4     if traffic_light_junction is True:
5         for i in range(len(d_list)):
6             if d_list[i] <= 5.0:
7                 result_los.append('Level of Service is A')
8             if d_list[i] > 5.0 and d_list[i]<=15.0:
9                 result_los.append('Level of Service is B')
10            if d_list[i] > 15.0 and d_list[i]<= 25.0:
11                result_los.append('Level of Service is C')
12            if d_list[i] > 25.0 and d_list[i]<= 40.0:
13                result_los.append('Level of Service is D')
14            if d_list[i] > 40.0 and d_list[i]<= 60.0:
15                result_los.append('Level of Service is E')
16            if d_list[i] > 60.0:
17                result_los.append('Level of Service is F')
18        else:
19            for i in range(len(d_list)):
20                if d_list[i] <= 10.0:
21                    result_los.append('Level of Service is A')
22                if d_list[i] > 10.0 and d_list[i]<=15.0:
23                    result_los.append('Level of Service is B')
24                if d_list[i] > 15.0 and d_list[i]<= 25.0:
25                    result_los.append('Level of Service is C')
26                if d_list[i] > 25.0 and d_list[i]<= 35.0:
27                    result_los.append('Level of Service is D')
28                if d_list[i] > 35.0 and d_list[i]<= 50.0:
29                    result_los.append('Level of Service is E')
30                if d_list[i] > 50.0:
31                    result_los.append('Level of Service is F')
32
33    return result_los
```

4.6: Μετρικές ποσοτικής αξιολόγησης

Η επίδοση των επιμέρους μοντέλων, αλλά και του τελικού σκοπού της μεθοδολογίας, αξιολογήθηκε ποσοτικά μέσω ενός εκτεταμένου συνόλου μετρικών επί των εξαγόμενων προβλέψεων. Η εν λόγω αξιολόγηση ξεκινά από τον πρωτογενή χαρακτηρισμό των αποτελεσμάτων δυαδικών λογικών προβλημάτων:

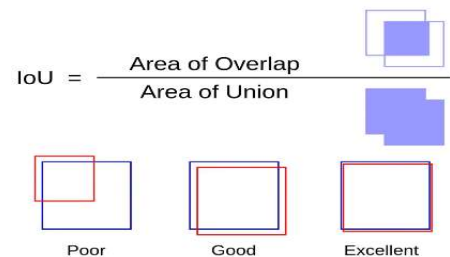
- **Ορθώς Θετικό (True Positive, TP):** Καλείται το αποτέλεσμα ενός αλγορίθμου που προβλέπει ως θετικό ένα δείγμα, το οποίο, βάσει δεδομένων αλήθειας, είναι πράγματι θετικό.
- **Ορθώς Αρνητικό (True Negative, TN):** Καλείται το αποτέλεσμα ενός αλγορίθμου που προβλέπει ως αρνητικό ένα δείγμα, το οποίο, βάσει δεδομένων αλήθειας, είναι πράγματι αρνητικό.
- **Λανθασμένα Θετικό (False Positive, FP):** Καλείται το αποτέλεσμα ενός αλγορίθμου που προβλέπει ως θετικό ένα δείγμα, το οποίο όμως, βάσει δεδομένων αλήθειας, θα έπρεπε να προβλεφθεί ως αρνητικό.
- **Λανθασμένα Αρνητικό (False Negative, FN):** Καλείται το αποτέλεσμα ενός αλγορίθμου που προβλέπει ως αρνητικό ένα δείγμα, το οποίο όμως, βάσει δεδομένων αλήθειας, θα έπρεπε να προβλεφθεί ως θετικό.

4.6.1: Μετρικές ανίχνευσης αντικειμένων

Αρχικά, το πρόβλημα συνεχών μεταβλητών της (χωρικής) ανίχνευσης ενός αντικειμένου, μετατρέπεται σε πρόβλημα δυαδικού αποτελέσματος με τη χρήση του κριτηρίου μετρικής **Intersection over Union (IoU)**:

$$IoU = \frac{bbox_{gt} \cap bbox_{pred}}{bbox_{gt} \cup bbox_{pred}}$$

Η συγκεκριμένη μετρική εκφράζει το ποσοστό (0 έως 1) κατά το οποίο το προβλεπόμενο bounding box τέμνει την πληροφορία αλήθειας, και άρα μεγαλύτερη τιμή υποδεικνύει ακριβέστερη πρόβλεψη. Το αποτέλεσμα της ανίχνευσης κατηγοριοποιείται δυαδικά με την επιβολή μιας τιμής κατωφλίου, π.χ. 0.7, της οποίας η υπέρβαση ή όχι σημαίνει αποδεκτή/θετική ή απορριπτόμενη/αρνητική ανίχνευση, αντίστοιχα. Όπως αναφέρθηκε σε προηγούμενο κεφάλαιο, η IoU συμμετέχει ως κριτήριο και σε διεργασίες εκτός των σκοπών αξιολόγησης,



Σχήμα 34: Περιγραφή της μετρικής IoU. (πηγή: <https://idiotdeveloper.com/what-is-intersection-over-union-iou/>, προσπελάστηκε Αυγ. 2023)

όπως στο «φιλτράρισμα» προτάσεων (NMS) ενός RPN, και στη συσχέτιση δεδομένων (data association) κατά την παρακολούθηση πολλαπλών αντικειμένων.

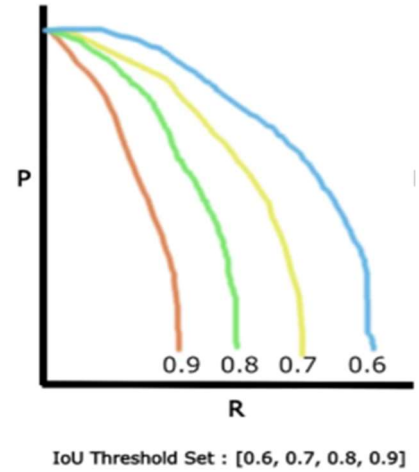
[35] Σύμφωνα με τα παραπάνω ορίζονται δύο βασικοί δείκτες ακρίβειας:

- **Precision (ακρίβεια)** = $\frac{TP}{TP+FP}$, $\in [0, 1]$, \uparrow better:

Ειδικά για την ανίχνευση αντικειμένων, η αναλογία των ορθών ανιχνεύσεων (ανά κατηγορία) έναντι των συνολικών ανιχνεύσεων που προβλέφθηκαν (TP + FP).

- **Recall (ανάκτηση)** = $\frac{TP}{TP+FN}$, $\in [0, 1]$, \uparrow better:

Ειδικά για την ανίχνευση αντικειμένων, η αναλογία των ορθών ανιχνεύσεων (ανά κατηγορία) έναντι του πλήθους των αναμενόμενων ανιχνεύσεων που ορίζουν τα δεδομένα αλήθειας (TP + FN).



Σχήμα 35: Καμπύλες Precision – Recall, για διαφορετικά κατώφλια IoU. (πηγή: <https://hasty.ai/docs/mp-wiki/metrics/map-mean-average-precision>, προσπελάστηκε Αυγ. 2023)

Τα μεγέθη αυτά οπτικοποιούνται (ανά κατηγορία) στην καμπύλη Precision/Recall για διάφορες πιθανές τιμές κατώφλιου IoU, δείχνοντας την ακρίβεια που διατηρεί το μοντέλο καθώς η ανάκτησή του αυξάνεται. Σε ένα τυπικό μοντέλο η αύξηση της ανάκτησης συνεπάγεται την αύξηση των λανθασμένων θετικών δειγμάτων και άρα τη μείωση της ακρίβειας, για αυτό και η καμπύλη που προκύπτει είναι φθίνουσα. Το εμβαδόν μεταξύ της καμπύλης και του οριζόντιου άξονα αποτελεί τη **μέση ακρίβεια (Average Precision, AP)** του μοντέλου για δοθέν κατώφλι IoU και κατηγορία, και υπολογίζεται ως ο αριθμητικός μέσος των ακριβειών που επιτυγχάνονται σε ίσα χωρισμένα διαστήματα ανάκτησης. Φυσικά θέτοντας λιγότερο αυστηρό κριτήριο (χαμηλότερο κατώφλι IoU), η απόδοση εμφανίζεται βελτιωμένη, έως ότου προσεγγίζεται η ιδανική κατάσταση του μοντέλου (τέλειες προβλέψεις), που αποδίδει καμπύλη με τη μορφή οριζόντιας ευθείας, όπου ισχύει Precision = 1 για κάθε τιμή Recall. Υπολογίζοντας, εν τέλει, το **μέσο όρο των μέσων ακριβειών (Mean Average Precision, mAP)** που προκύπτουν από κάθε κατηγορία για δεδομένο κατώφλι IoU, σχηματίζεται μια συνολική ποσοτική αξιολόγηση για το μοντέλο ανίχνευσης αντικειμένων.

$$mAP = \frac{1}{Num_{classes}} \cdot \sum_{c \in classes} \frac{TP_c}{TP_c + FP_c}$$

[33] Στο COCO Dataset Challenge χρησιμοποιείται η μετρική **mAP@0.50:0.95**, η οποία αποτελεί παραλλαγή της mAP με κριτήριο ένα σύνολο κατωφλίων IoU από 0.50 έως 0.95 με ισοδιάσταση 0.05. Δηλαδή, υπολογίζεται διαφορετική mAP για κατώφλια IoU = 0.50, IoU = 0.55, IoU = 0.60 ... , IoU = 0.95, και τελικά η mAP@0.50:0.95 αποτελεί το μέσο όρο αυτών. Παρατηρείται πως το διάστημα (0.95, 1.00] αποκλείεται της κατανομής, καθώς θεωρείται

στατιστικά σπάνια η επίτευξη τόσο ιδανικών χωρικών εντοπισμών. Για τον ίδιο λόγο, στην παρούσα μεθοδολογία η μετρική προσαρμόζεται στις προκλήσεις της ανίχνευσης μικροσκοπικών αντικειμένων αποκλείοντας το διάστημα (0.80, 1.00), ενώ παράλληλα θεωρώντας αποδεκτό το [0.30,0.50]. Τελικώς, εφαρμόζεται η μετρική αξιολόγησης **mAP@0.30:0.80**, επίσης ισοδιάστασης 0.05, με επιπρόσθετη διάκριση σε small (Area < 1024 pixels), medium (Area ∈ [1024,9216] pixels) και large (Area > 9216 pixels). Επίσης εξετάζονται οι **mAP@0.50 (PASCAL VOC metric, mAP για κατώφλι IoU = 0.50)** και **mAP@0.75 (strict metric, mAP για κατώφλι IoU = 0.75)**.

Με παρόμοιους υπολογισμούς στο διάγραμμα Precision/Recall, προκύπτουν οι μετρικές της μέσης ανάκτησης (**Average Recall, AR**), και του μέσου όρου των μέσων ανακτήσεων (**mean Average Recall, mAR**). Χρησιμοποιούνται οι αντίστοιχες παραλλαγές **mAR@0.30:0.80, mAR@0.50** και **mAR@0.75**.

4.6.2: Μετρικές παρακολούθησης πολλαπλών αντικειμένων

Προχωρώντας στο τελικό στάδιο της μεθοδολογίας Βαθιάς Μάθησης, η παρακολούθηση πολλαπλών αντικειμένων αξιολογείται ποσοτικά σύμφωνα με το **σύστημα μετρικών του MOT Benchmark** (στην πλειοψηφία τους μέρος των **Clear MOT Metrics** [36]):

- **MOTA (Multiple Object Tracking Accuracy), ∈ [0, 1], ↑ better:** Αποτελεί μια από τις πρώτες (2008) μετρικές αξιολόγησης, η οποία λαμβάνει υπόψη τα προβληματικά δείγματα ανίχνευσης (FN+FP) σε κάθε καρέ, και τις μη επιθυμητές αλλαγές ID στους στόχους (IDSW), αλλά όχι την ποιότητα χωρικού εντοπισμού (localization). Δίνεται από τον τύπο:

$$MOTA = 1 - \frac{FN + FP + IDSW}{gtDets},$$

όπου gtDets = το σύνολο των ανιχνεύσεων που υποδεικνύουν τα δεδομένα αλήθειας.

- **MOTP (Multiple Object Tracking Precision), ∈ [0, 1], ↓ better:** Αφορά το μέσο σφάλμα χωρικού εντοπισμού σε μορφή IoU των στόχων που παρακολουθούνται, δίχως να αξιολογείται η συνέχεια της τροχιάς τους. Προσδιορίζεται από την παρακάτω εξίσωση:

$$MOTP = \frac{\sum_t d_{t,i}}{\sum_t c_t},$$

όπου $d_{t,i}$ = η επικάλυψη του προβλεπόμενου bounding box του στόχου i με το αντίστοιχο ground truth bounding box στο frame t , και c_t = το πλήθος των αντιστοιχίσεων στόχων και ανιχνεύσεων στο frame t .

- **HOTA (Higher Order Tracking Accuracy)** [37], $\in [0, 1]$, **↑ better**: Πρόκειται για μια νεότερη (2020) σύνθετη μετρική που μετράει πόσο καλά συντονίζονται οι τροχιές των αντιστοιχισμένων ανιχνεύσεων, σταθμίζοντάς το σε όλες τις ανιχνεύσεις που αντιστοιχίζονται, ενώ παράλληλα «τιμωρεί» για τις ανιχνεύσεις που δεν αντιστοιχίζονται. Πραγματοποιεί μια εξισορροπημένη αξιολόγηση μεταξύ ανίχνευσης και παρακολούθησης των στόχων. Υπολογίζεται ως εξής:

$$HOTA = \int_{0 < a \leq 1} HOTA_a \approx \frac{1}{19} \cdot \sum_{a=0.05, a+=0.05}^{0.95} HOTA_a, \text{ όπου:}$$

$$HOTA_a = \sqrt{DetA_a \cdot AssA_a},$$

$$DetA_a = Det_{IoU} = \frac{TP}{TP+FN+FP} \text{ (Detection Accuracy, Ορθότητα Ανίχνευσης),}$$

$$AssA_a = \frac{1}{TP} \cdot \sum_{c \in TP} Ass_{IoU}(C) = \frac{1}{TP} \cdot \sum_{c \in TP} \frac{TPA(c)}{TPA(c) + FNA(c) + FPA(c)}$$

(Association Accuracy, Ορθότητα Συσχέτισης),

με TPA, FNA, FPA = True Positive, False Negative και False Positive Associations,

αντίστοιχα, και $a =$ κατώφλι IoU για Hungarian matching

Για $a = \alpha_{\min} = 0.05$ (threshold 0) προκύπτει η παραλλαγή HOTA(0) η οποία ουσιαστικά προσαρμόζει ελάχιστο φιλτράρισμα δειγμάτων βάσει χωρικού εντοπισμού (θεωρώντας ότι ο χωρικός εντοπισμός είναι ιδανικός και δεν διαβάλλει τη μετρική). Ακόμη, υπάρχει η επιπλέον παραλλαγή της, $HOTALocA(0) = HOTA(0) \cdot LocA(0)$, με $LocA(0) = \frac{1}{TP} \cdot \sum_{c \in TP} Loc_{IoU}$ (ορθότητα στο threshold 0, η οποία εκφράζει τις συνέπειες της ακρίβειας χωρικού εντοπισμού των ελάχιστα φιλτραρισμένων αντιστοιχίσεων στη μετρική HOTA).

- **Open-World Tracking Accuracy (OWTA)** [38], $\in [0, 1]$, **↑ better**:: Αποτελεί μια γενικευμένη μορφή της HOTA σε εφαρμογές open-world tracking (παρακολούθηση μη προκαθορισμένων, αλλά δυναμικά μεταβαλλόμενων στόχων ενδιαφέροντος), αξιοποιώντας την ανάκτηση των ανιχνεύσεων έναντι της ορθότητας (accuracy) τους. Δίνεται από τον τύπο:

$$OWTA_a = \sqrt{DetRe_a \cdot AssA_a},$$

$$\text{με } DetRe_a = \frac{TP}{TP+FN}$$

- Ακόμη, εξετάστηκαν και λιγότερο σύνθετα, αλλά εξίσου σημαντικά μεγέθη αξιολόγησης όπως:

- **ID Precision (IDP)**, $\in [0, 1]$, **↑ better**: Η ακρίβεια στο λογικό πρόβλημα ορθής ανάθεσης ID σε έναν ανιχνευόμενο στόχο βάσει των κριτηρίων ομοιότητας. Υπολογίζεται ως:

$$IDP = \frac{IDTP}{IDTP + IDFP}$$

- **ID Recall (IDR)**, $\in [0, 1]$, **↑ better**: Η ανάκτηση στο λογικό πρόβλημα ορθής ανάθεσης ID σε έναν ανιχνευόμενο στόχο βάσει των κριτηρίων ομοιότητας. Υπολογίζεται ως:

$$IDR = \frac{IDTP}{IDTP + IDFN}$$

- **ID F1 Score (IDF1)**, $\in [0, 1]$, **↑ better**: Πρόκειται για το F1 score της ορθής ανάθεσης ID, το οποίο φανερώνει την ορθότητα του μοντέλου αποδίδοντας ενισχυμένη βαρύτητα στα ορθώς θετικά δείγματα:

$$IDF1 = \frac{IDTP}{IDTP + 0.5 \cdot IDFN + 0.5 \cdot IDFP}$$

- **Mostly Tracked (MT)**, $\in \mathbb{N}$, **↑ better**: Το πλήθος των στόχων που παρακολουθήθηκαν ορθώς για περισσότερο από το 80% της διάρκειάς τους.
- **Mostly Lost (ML)**, $\in \mathbb{N}$, **↓ better**: Το πλήθος των στόχων που παρακολουθήθηκαν ορθώς για λιγότερο από το 20% της διάρκειάς τους.
- **Partially Tracked (PT)**, $\in \mathbb{N}$, **↓ better vs MT, ↑ better vs ML**: Το πλήθος των στόχων που παρακολουθήθηκαν ορθώς κατά το 20 έως 80 % της διάρκειάς τους.
- **Fragmentations (FM)**, $\in \mathbb{N}$, **↓ better**: Το πλήθος περιπτώσεων όπου ένας ground truth στόχος παρακολουθείται ορθώς, χάνεται προσωρινά, κι έπειτα επανασταυτοποιείται. Σε περιβάλλον πολλαπλών συγκαλύψεων λαμβάνεται θετικά υπόψη στην αξιολόγηση.
- **ID Switch (IDSw)**, $\in \mathbb{N}$, **↓ better**: Το πλήθος των περιπτώσεων όπου προβλέπεται η αλλαγή του ID ενός ground truth στόχου.
- **ID Transfer (IDt)**, $\in \mathbb{N}$, **↓ better**: Το πλήθος των περιπτώσεων κατά τις οποίες ένα ID έχει μεταφερθεί σε διαφορετικούς στόχους.
- **ID Ascend (IDa)**, $\in \mathbb{N}$, **↓ better**: Το πλήθος των περιστατικών όπου εκχωρείται ένα νέο (πρωτοφανές) διαφορετικό ID σε έναν προϋπάρχων ground truth στόχο. Αποτελεί υποπερίπτωση IDSw.

- **ID Migrate (IDm)**, $\in \mathbb{N}$, \downarrow **better**: Το πλήθος των περιστατικών όπου ένα προϋπάρχον ID αντιστοιχίζεται σε έναν νέο (πρωτοφανή) ground truth στόχο. Αποτελεί υποπερίπτωση IDt.

4.6.3: Πρόσθετες μετρικές κυκλοφοριακού φόρτου

Το εξαγόμενο πρωτογενές μέγεθος το οποίο χρησιμοποιείται στον κυκλοφοριακό σχεδιασμό και την ανάλυση, ο κυκλοφοριακός φόρτος, αξιολογείται ποσοτικά από δύο νέες και προσαρμοσμένες στην μεθοδολογία μετρικές:

- **Ποσοστιαία απόκλιση εκτίμησης συνολικού κυκλοφοριακού φόρτου** $\in [0, 1]$, \downarrow **better**: Αποτελεί μια μετρική απόλυτης αξιολόγησης της μεθοδολογίας στον τελικό της σκοπό, συγκρίνοντας τον συνολικό εκτιμώμενο φόρτο, με τον πραγματικό που μετρήθηκε «χειροκίνητα» παρακολουθώντας το βίντεο:

$$OverallVolume_{Dev} = \left| 1 - \frac{OverallVolume_{pred}}{OverallVolume_{real}} \right|$$

- **Δείκτης αξιόπιστου υπολογισμού κυκλοφοριακού φόρτου** $\in [0, 1]$, \uparrow **better**: Πρόκειται για το ποσοστό κατά το οποίο ο αλγόριθμος υπολογίζει με βέβαιο χαρακτηρισμό τον κυκλοφοριακό φόρτο, «χωρικά» κατανεμημένα. Δηλαδή με προσμέτρηση τροχιών που τέμνουν δύο διατομές προσβάσεων, και όπου απαιτείται, με επιπλέον τμήση στη διατομή ελέγχου, ώστε να προκύπτουν σαφώς ζεύγη προέλευσης-προορισμού. Αναδεικνύεται, ουσιαστικά, η ικανότητα διατήρησης κοινής ταυτότητας στα οχήματα κατά τη διέλευσή τους από τις κρίσιμες περιοχές του κόμβου. Δίνεται από τον τύπο:

$$AssuranceIndex = 1 - \frac{OverallVolume_{ambiguous}}{OverallVolume_{pred}},$$

όπου $OverallVolume_{ambiguous}$ = ο συνολικός μη κατανεμημένος χωρικά κυκλοφοριακός φόρτος που καταμετρήθηκε.

Κεφάλαιο 5^ο: Τελικά αποτελέσματα – Συμπεράσματα

Στο τελευταίο κεφάλαιο της διπλωματικής εργασίας παρατίθενται τα αποτελέσματα και οι αξιολογήσεις – ποιοτικές και ποσοτικές- όλων των επιμέρους συνιστωσών της μεθοδολογίας, καθώς και των τελικών προβλέψεων στο πεδίο εφαρμογής κυκλοφοριακής τεχνικής. Ακόμη, εκφράζονται τα συμπεράσματα που προκύπτουν από το σύνολο του εγχειρήματος, σχετικά με τις αστοχίες του, τις προοπτικές βελτίωσης και τα οφέλη της καθιέρωσής του.

5.1: Αποτελέσματα και αξιολόγηση μεθοδολογίας Βαθιάς Μάθησης

Η επίδοση της μεθοδολογίας Βαθιάς Μάθησης που υλοποιήθηκε, εκτός από την τελική ικανότητά της στην παρακολούθηση των τροχιών των ανιχνευμένων στόχων με την πάροδο των καρέ ενός βίντεο, υπό απαράλλακτη ταυτότητα για καθένα εξ αυτών, εξαρτάται πρωτίστως και από την αξιοπιστία των παρεχόμενων ανιχνεύσεων. Η αποτίμηση της «detection συνιστώσας», όπως και της καταληκτικής παρακολούθησης πολλαπλών αντικειμένων, παρουσιάζονται στο παρόν υποκεφάλαιο.

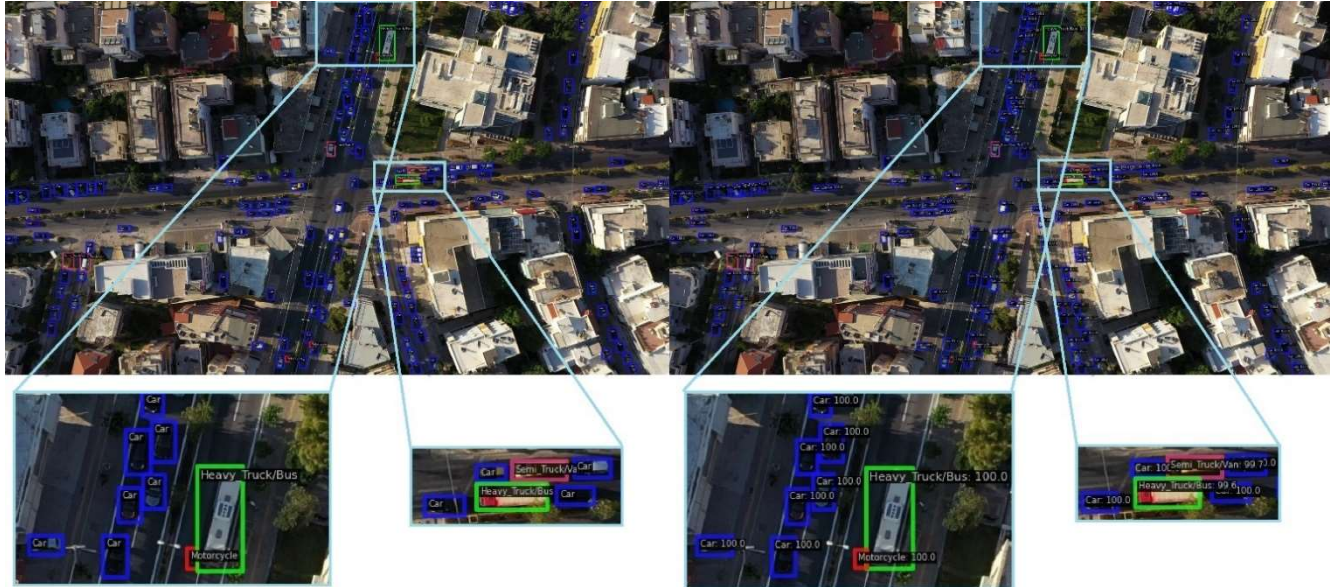
5.1.1: Επιδόσεις συνιστώσας ανίχνευσης

Ο ανιχνευτής που εκπαιδεύτηκε κατάφερε εντυπωσιακά αποτελέσματα στον εντοπισμό και την κατηγοριοποίηση των οχημάτων, σε κάθε είδος λήψης, και σε πυκνές φάλαγγες, ακόμη και για την κατηγορία των δίκυκλων που εμφανίζονται ως εξαιρετικά μικροσκοπικά αντικείμενα.



Σχήμα 36: Σύγκριση πληροφορίας αλήθειας (αριστερά), με τις εξαγόμενες ανιχνεύσεις (δεξιά) στον κυκλικό κόμβο Πλατείας 25^{ης} Μαρτίου.

Τα λιγοστά σφάλματα περιορίστηκαν σε κοινότυπες λανθασμένες κατηγοριοποιήσεις οχήματος, σε αβλεπίες λόγω συγκαλύψεων φωτοσκίασης ή «αναγλύφου», και σε ελαφρώς «άστοχα» bounding boxes, κυρίως σε μεγαλύτερα ή/και σπάνιου παρουσιαστικού οχήματα.



Σχήμα 37: Σύγκριση πληροφορίας αλήθειας (αριστερά), με τις εξαγόμενες ανιχνεύσεις (δεξιά) στον κόμβο Πρωτόπαπα – Σοφοκλή Βενιζέλου.



Σχήμα 38: Σύγκριση πληροφορίας αλήθειας (αριστερά), με τις εξαγόμενες ανιχνεύσεις (δεξιά) στον κόμβο Ηρώων Πολυτεχνείου – Παύλου Πεντάρη.

Όσον αφορά την ποσοτική αξιολόγηση, εφαρμόστηκαν οι mAP/mAR μετρικές στον έλεγχο επί του σετ εκπαίδευσης, αλλά κι επί των 184 καρέ του Aerial Multi-Vehicle Detection Dataset. Για τα διαδοχικά καρέ ελέγχου του αλγορίθμου παρακολούθησης παρατίθεται η απόλυτη «ποτ class-wise» ακρίβεια και ανάκληση για $IoU > 0.5$.

Πίνακας 7: Αποτελέσματα ποσοτικής αξιολόγησης με τις μετρικές mAP και mAR @[0.30:0.80].

Δεδομένα ελέγχου	Τύπος μετρικής	Μικρά αντικείμενα	Μεσαία αντικείμενα	Μεγάλα αντικείμενα	Όλα τα αντικείμενα
Σετ εκπαίδευσης	mAP@[0.30:0.80]	97.2 % (για έως 150 ανιχνεύσεις/εικόνα)	97.7 % (για έως 150 ανιχνεύσεις/εικόνα)	96.8 % (για έως 150 ανιχνεύσεις/εικόνα)	97.1 % (για έως 150 ανιχνεύσεις/εικόνα)
Σετ εκπαίδευσης	mAR@[0.30:0.80]	98.4 % (για έως 150 ανιχνεύσεις/εικόνα)	98.9 % (για έως 150 ανιχνεύσεις/εικόνα)	98.5 % (για έως 150 ανιχνεύσεις/εικόνα)	98.4 % (για έως 150 ανιχνεύσεις/εικόνα)
Υποσύνολο Aerial Multi-Vehicle Detection Dataset	mAP@[0.30:0.80]	20.9 % (για έως 300 ανιχνεύσεις/εικόνα)	47.2 % (για έως 300 ανιχνεύσεις/εικόνα)	0	41.0 % (για έως 300 ανιχνεύσεις/εικόνα)
Υποσύνολο Aerial Multi-Vehicle Detection Dataset	mAR@[0.30:0.80]	25.4 % (για έως 300 ανιχνεύσεις/εικόνα)	57.0 % (για έως 300 ανιχνεύσεις/εικόνα)	0	48.7 % (για έως 300 ανιχνεύσεις/εικόνα)

Πίνακας 8: Αποτελέσματα ποσοτικής αξιολόγησης με τις μετρικές mAP@0.50 και mAP@0.75.

Δεδομένα ελέγχου	Τύπος μετρικής	Όλα τα αντικείμενα
Σετ εκπαίδευσης	mAP@0.50	98.7 % (για έως 150 ανιχνεύσεις/εικόνα)
Σετ εκπαίδευσης	mAP@0.75	94.6 % (για έως 150 ανιχνεύσεις/εικόνα)
Υποσύνολο Aerial Multi-Vehicle Detection Dataset	mAP@0.50	49.6 % (για έως 300 ανιχνεύσεις/εικόνα)
Υποσύνολο Aerial Multi-Vehicle Detection Dataset	mAP@0.75	23.5 % (για έως 300 ανιχνεύσεις/εικόνα)

Πίνακας 9: Αποτελέσματα ποσοτικής αξιολόγησης σε detection-by-frame, βάσει Precision και Recall για $IoU > 0.5$.

Δεδομένα ελέγχου	GT Dets	Dets	FP ↓	FN ↓	Precision (IoU > 0.5) ↑	Recall (IoU > 0.5) ↑
30" από DJI_0073.mp4 (Κυκλικός κόμβος Πλατείας 25ης Μαρτίου)	28550	29071	1775	1254	93.9 %	95.6 %
30" από DJI_0095_1.mp4 (Ηρώων Πολυτεχνείου – Παύλου Πεντάρη)	24222	26370	3076	928	88.3 %	96.2 %
Πλήρες DJI_0564.mp4 (Πρωτόπαπα – Σοφοκλή Βενιζέλου)	36262	44013	11654	3903	73.5 %	89.2 %
Σύνολο	89034	99454	16505	6085	83.4 %	93.2 %

Στα δεδομένα που προέρχονται από το ίδιο το Suncity Dataset, από ένα πολύ μικρό ποσοστό του οποίου πραγματοποιήθηκε η εκπαίδευση του παρόντος μοντέλου ανίχνευσης, προέκυψαν σχεδόν άριστες επιδόσεις σε όλες τις μετρικές αξιολόγησης. Ακόμη και στα ~2600 καρέ του Πίνακα 9, που η συντριπτική τους πλειοψηφία δε συμμετείχε στην εκπαίδευση, τα αποτελέσματα ακρίβειας και ανάκλησης ήταν εξαιρετικά. Υπήρξε μια πτώση της ακρίβειας στο βίντεο του κόμβου Πρωτόπαπα – Σοφοκλή Βενιζέλου, αλλά πρόκειται για τη λήψη που αντιστοιχούσε στο μεγαλύτερο ύψος πτήσης UAV, και στην κατώτερη ανάλυση μεταξύ του υπόλοιπου dataset (1080p). Επομένως, η παρατηρούμενη «έκπτωση» στην ακρίβεια κρίθηκε δικαιολογημένη, ενώ μάλιστα διατηρήθηκε εντός ικανοποιητικών επιπέδων.

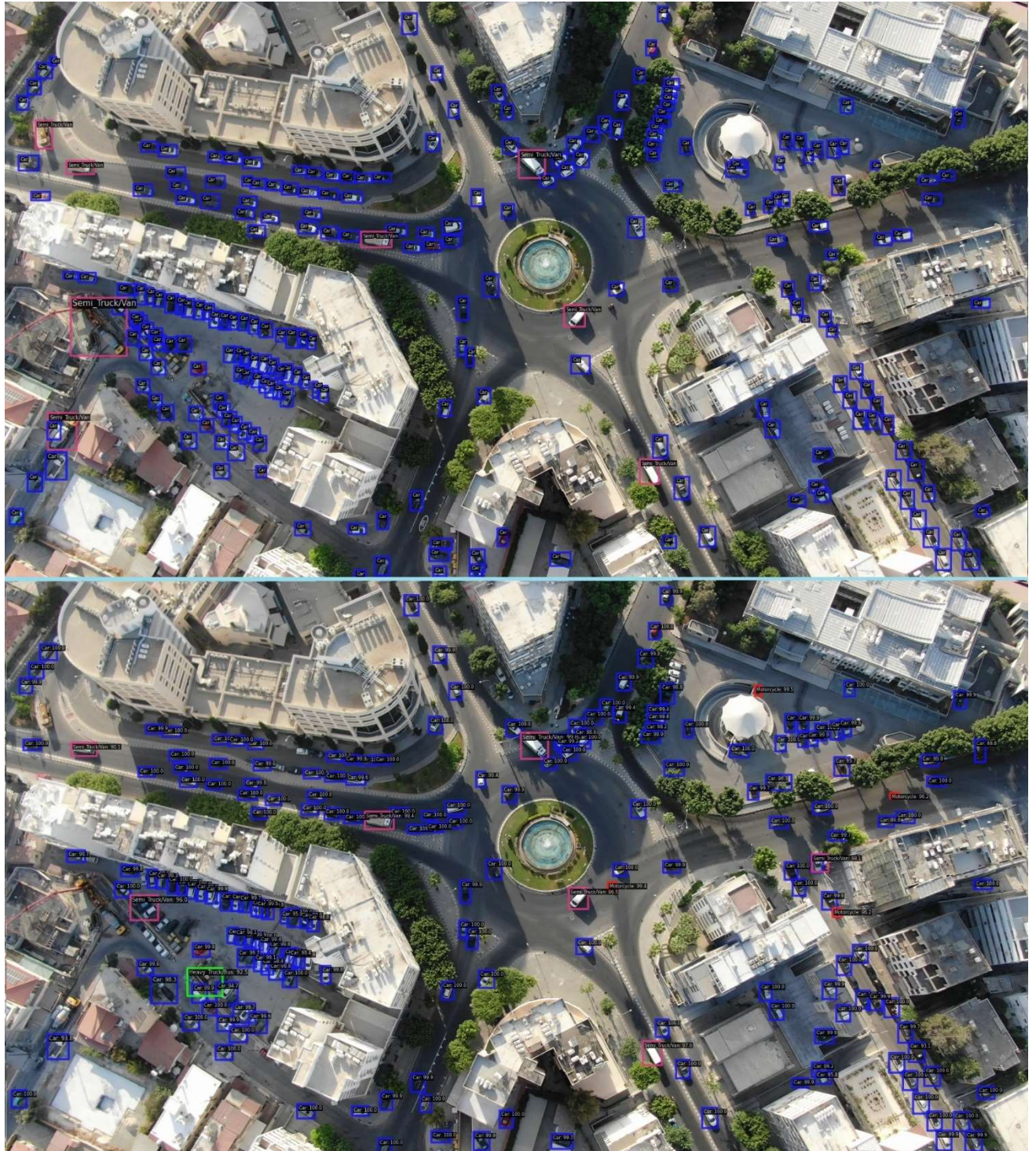
Όσον αφορά την εφαρμογή του μοντέλου στο υποσύνολο του Aerial Multi-Vehicle Detection Dataset, οι μετρικές αξιολόγησης παρουσιάστηκαν σαφώς μειωμένες:

- Στα αντικείμενα μεσαίου μεγέθους τού εμφανίστηκαν οι καλύτερες τιμές των mAP και mAR @[0.30:0.80], στη γειτονιά του 50%.
- Στα λιγιστά large objects (~150 εκ των συνολικών 45000) οι mAP και mAR @[0.30:0.80] ήταν μηδενικές, διότι αναφέρονταν σε ένα μοναδικό «ιδιόρρυθμο» στόχο ο οποίος δεν κατέστη εφικτό να ανιχνευτεί στα πολλαπλά καρέ που εμφανιζόταν. Συνεπώς, δε λαμβάνεται ιδιαίτερα υπόψιν.
- Η πλειοψηφία των αντικειμένων ανήκει στην τάξη μεγέθους «small», και απέδωσε μέτριες τιμές, εγγύτερα του 20 και 25 % (mAP και mAR @[0.30:0.80] αντίστοιχα).
- Τελικά, για όλα τα μεγέθη οχημάτων αποδόθηκε περίπου 40 και 50 % mAP και mAR @[0.30:0.80] αντίστοιχα.
- Σε παραπλήσια επίπεδα κυμάνθηκαν οι αυστηρότερες μετρικές mAP@0.50 και 0.70.

Η αναντιστοιχία μεταξύ των πρωτότυπων κατηγοριών του Aerial Multi-Vehicle Detection Dataset με τη διάκριση που ακολουθήθηκε κατά την εκπαίδευση στο Suncity Dataset δυσχέρυνε περαιτέρω την ποσοτική άποψη προσαρμογής του μοντέλου. Για παράδειγμα, οι διαφορετικές κατηγορίες «Bus» και «Truck» σαν οπτική πληροφορία χαρακτηριστικών συμπεριλήφθηκαν στην κοινή κατηγορία «Heavy_Truck/Bus» του Suncity, ενώ η «Semi_Truck/Van» αναφερόταν σε μικρότερα φορτηγά και ημιφορτηγά, τα οποία στο πρωτότυπο αρχείο περιγραφών του Aerial Multi-Vehicle Detection Dataset συχνά κατατάσσονταν στην κατηγορία «Car». Δημιουργήθηκε μια σύγχυση πληροφορίας που προκάλεσε σφάλματα ορθής ταξινόμησης οχημάτων, και άρα απώλεια θετικών δειγμάτων που έπληξε και τις μετρικές αξιολόγησης.

Ακόμη, η απουσία κατηγορίας «Motorcycle» δεν επηρέασε τον υπολογισμό των mAP/mAR. Γενικότερα η πλήρης έλλειψη πληροφορίας αλήθειας σε κάποια κατηγορία δεν

λαμβάνεται υπόψη στη στάθμιση των AP/AR μεταξύ όλων των κατηγοριών, δηλαδή δεν συνεπάγεται προσθήκη μιας μηδενικής τιμής για το μέσο όρο. Πράγματι σημειώθηκαν κάποιες επιτυχείς ανιχνεύσεις των δικύκλων, αλλά επειδή απουσίαζαν από το πρωτότυπο αρχείο περιγραφών δεδομένων αλήθειας δεν απέδωσαν AP/AR και δεν επηρέασαν τις μετρικές. Μονάχα με την λανθασμένη κατηγοριοποίηση ως αυτοκίνητα λογίστηκαν ως False Positive και «επιβάρυναν» την κατηγορία «Car».

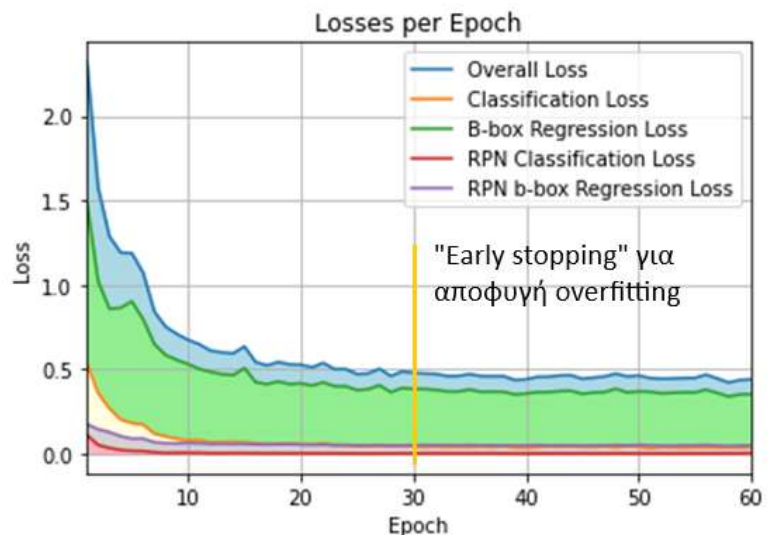


Σχήμα 39: Σύγκριση πληροφορίας αλήθειας (άνω), με τις εξαγόμενες ανιχνεύσεις (κάτω) σε ένα από τα καρέ του υποσυνόλου του Aerial Multi-Vehicle Detection Dataset.

Παρ' όλα αυτά, τονίζεται πως δεν αποτέλεσε ρεαλιστικό στόχο η επίτευξη υψηλού βαθμού στην ποσοτική αξιολόγηση του παραχθέντος μοντέλου επί ενός σετ δεδομένων που μοιάζει μεν οπτικά με το σετ εκπαίδευσης, αλλά ωστόσο ενδέχεται να διαφέρει στατιστικά στις τιμές των εικονοστοιχείων, ή κυριότερα στις απεικονιζόμενες κλίμακες και οπτικές γωνίες των αντικειμένων κάθε κατηγορίας. Ένα δείγμα οπτικής εποπτείας των αποτελεσμάτων του ελέγχου που παρατίθεται παραπάνω «διαψεύδει» τη σχετικά χαμηλή ποσοτική αξιολόγηση, και δείχνει ότι το μοντέλο προσαρμόζεται επαρκώς στον πεδίο (domain) του «ξένου» dataset.

Σε γενικές γραμμές κρίθηκε ως μια αποδεκτή εφαρμογή ανίχνευσης, αφήνοντας αισιόδοξες προσδοκίες για την αποδοτικότητα του παραχθέντος μοντέλου σε εφαρμογές μεταφοράς μάθησης σε προσαρμοσμένα μοντέλα για κάθε περιοχή. Η ποιότητά της όμως αστόχησε στις περιοχές όπου βρίσκονται πυκνά σταθμευμένα οχήματα, στις οποίες το μοντέλο είτε προέβλεπε την ύπαρξη ενός μεγαλύτερου οχήματος στη θέση πολλαπλών μικρότερων, είτε λόγω σύγχυσης της οπτικής πληροφορίας αδυνάτησε να εξάγει θετικό δείγμα. Στο σετ εκπαίδευσης του μοντέλου σπάνια υπήρχε σκηνικό τόσο υπερβολικά πυκνής συστοιχίας οχημάτων. Επίσης, σε μερικές περιπτώσεις δημιουργήθηκαν προβλήματα προσαρμογής κλίμακας στις κατηγορίες βαρέων οχημάτων και φορτηγών, καθώς πλέον εμφανίζονταν ακόμη και ως ~25x50 στόχοι, δηλαδή κοντά στη συνήθη κλίμακα που διαθέτουν τα επιβατικά αυτοκίνητα στο σετ εκπαίδευσης, οπότε ταξινομήθηκαν ως «Car».

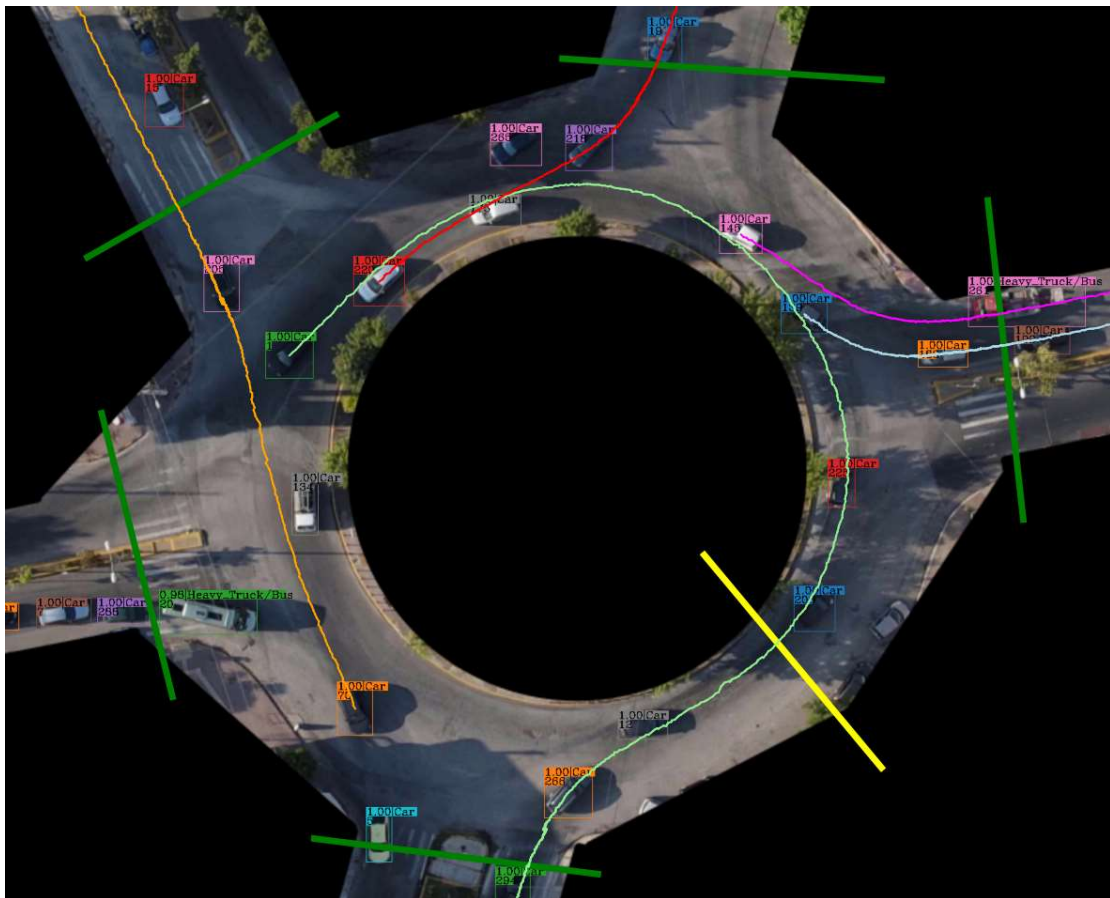
Τέλος, στο παρακάτω γράφημα παρουσιάζεται η πρόοδος των συναρτήσεων κόστους κατά την εκπαίδευση, συναρτήσει της παρόδου των εποχών. Ως κριτήριο πρόωρου τερματισμού για διάκριση μιας λιγότερο «overfit» κατάστασης του μοντέλου, ορίστηκε η διαφορά μεταξύ του συνολικού κόστους διαδοχικών checkpoints να είναι λιγότερο από 0.05. Εφόσον η ισοδιάσταση σημείων αποθήκευσης είναι ίση με 10 εποχές, η 30^η εποχή εκπλήρωσε αυτό το κριτήριο, όπως προέκυψε από τη σύγκριση του κόστους της με αυτό της 20^{ης} εποχής.



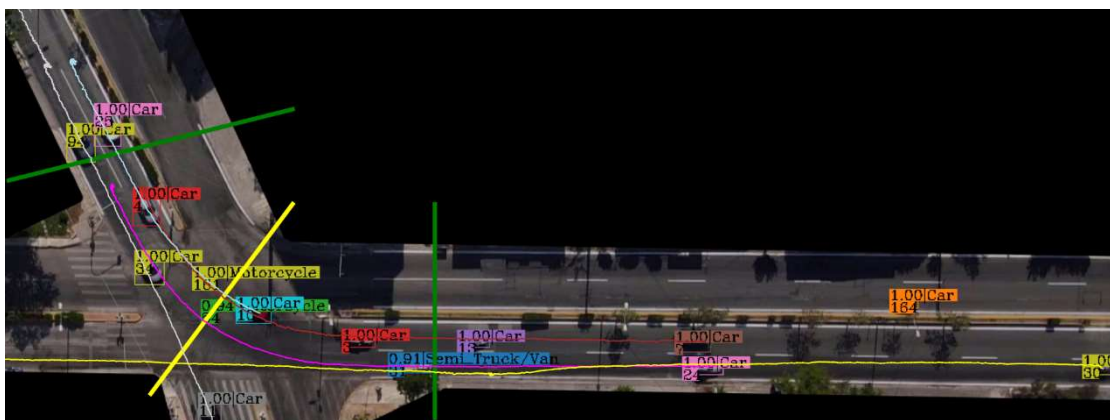
Σχήμα 40: Καμπύλες κόστους ανά εποχή εκπαίδευσης του παραχθέντος μοντέλου ανίχνευσης.

5.1.2: Επιδόσεις μοντέλου Παρακολούθησης Πολλαπλών Αντικειμένων

Με δεδομένη την παροχή ιδιαίτερα ποιοτικών ανιχνεύσεων σε κάθε ξεχωριστό καρέ ενός εξαταζόμενου βίντεο, η παραμετροποίηση κι εφαρμογή του DeepSORT αλγορίθμου απέδωσε εξαιρετικά αποτελέσματα, εξάγοντας ομαλές και ακέραιες τροχιές για μια ισχυρή πλειοψηφία των οχημάτων, ακόμη και σε αρκετά απαιτητικές συνθήκες εναλλαγής φωτισμού, πυκνών ανιχνεύσεων, και συγκαλύψεων. Παρακάτω τοποθετούνται αποσπάσματα δειγμάτων οπτικοποίησης από τις εξαγόμενες τροχιές 5 αυτοκινήτων (για αποφυγή οπτικής σύγχυσης) σε καθεμία από τις 3 διαφορετικές τοποθεσίες μελέτης.



Σχήμα 41: Παραδείγματα οπτικοποίησης εξαγόμενων τροχιών οχημάτων από το βίντεο DJI_0073.mp4.

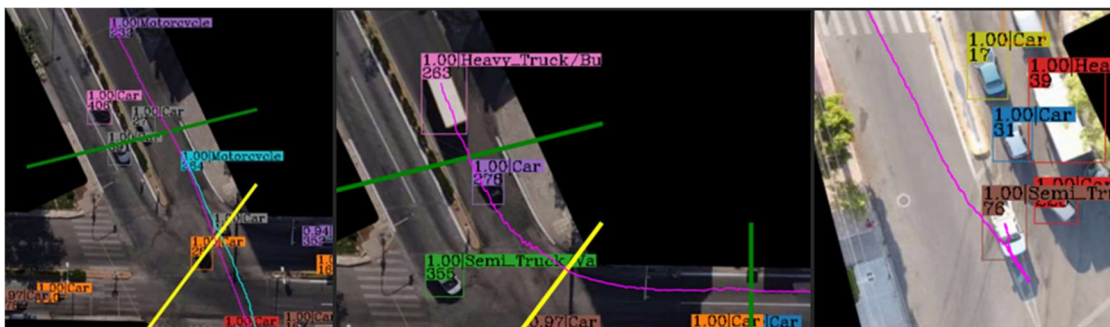


Σχήμα 42: Παραδείγματα οπτικοποίησης εξαγόμενων τροχιών οχημάτων από το βίντεο DJI_0107.mp4.



Σχήμα 44: Παραδείγματα οπτικοποίησης εξαγόμενων τροχιών οχημάτων από το βίντεο DJI_0560.mov.

Οι τροχιές αποτυπώθηκαν με ενθαρρυντικά αποτελέσματα ακόμη και στα αρκετά μικρότερα σε μέγεθος δίκυκλα, αλλά και στα μεγαλύτερα βαρέα οχήματα, τα οποία όμως συχνά έφεραν λιγότερο ομαλές τροχιές λόγω της απόκλισης του κεντροειδούς στα ευρεία και ορθογώνια bounding boxes. Αξιόλογη κρίθηκε, επίσης, η επίδοση του μοντέλου σε απροσδόκητες και «αφύσικες» αλλαγές στην κατεύθυνση των οχημάτων.



Σχήμα 43: Επιτυχής παρακολούθηση της πορείας δίκυκλων και μεγάλου φορτηγού στην πρώτη και δεύτερη εικόνα, αντίστοιχα. Στο τρίτο απόσπασμα ο tracker αποτυπώνει άριστα την τροχιά ενός ημιφορτηγού που κάνει ασυνήθιστη αλλαγή κατεύθυνσης κατά την πραγματοποίηση ελιγμού στάθμευσης.

Το θετικό ποιοτικό έλεγχο πιστοποιήσαν οι υψηλές αποδόσεις στις μετρικές ποσοτικής αξιολόγησης του MOT Benchmark:

Πίνακας 10: Αποτελέσματα μετρικών ποσοτικής αξιολόγησης του MOT Benchmark (I).

Δεδομένα ελέγχου	DetA ↑	DetPr ↑	DetRe ↑	AssA ↑	AssPr ↑	AssRe ↑	IDP ↑	IDR ↑	IDF1 ↑	LocA ↑
30" από DJI 0073.mp4 (Κυκλικός κόμβος Πλατείας 25ης Μαρτίου)	79.2 %	84.5 %	86.1 %	81.6 %	86.9 %	86.9 %	89.9 %	91.5 %	90.7 %	88.9 %
30" από DJI 0095 1.mp4 (Ηρώων Πολυτεχνείου – Παύλου Πεντάρη)	76.0 %	80.6 %	87.8 %	81.7 %	88.2 %	87.2 %	83.6 %	91.0 %	87.1 %	90.0 %
Πλήρες DJI 0564.mp4 (Πρωτόπαπα – Σοφοκλή Βενιζέλου)	57.4 %	64.1 %	77.8 %	79.0 %	87.1 %	82.3 %	71.7 %	87.0 %	78.6 %	86.5 %
Σύνολο	68.4 %	74.4 %	83.1 %	80.9 %	87.7 %	85.6 %	80.1 %	89.5 %	84.6 %	88.3 %

Πίνακας 11: Αποτελέσματα μετρικών ποσοτικής αξιολόγησης του MOT Benchmark (II).

Δεδομένα ελέγχου	GT Tracks/ IDs	MT ↑	PT ↓/↑	ML ↓	FM ↓	Tracks/ IDs	IDSw ↓	IDa ↓	IDt ↓	IDm ↓
30" από DJI 0073.mp4 (Κυκλικός κόμβος Πλατείας 25ης Μαρτίου)	68	63 (92.6 %)	3 (4.5 %)	2 (2.9 %)	169	240	54	40	7	4
30" από DJI 0095 1.mp4 (Ηρώων Πολυτεχνείου – Παύλου Πεντάρη)	77	72 (93.5 %)	2 (2.6 %)	3 (3.9 %)	254	258	103	78	8	6
Πλήρες DJI 0564.mp4 (Πρωτόπαπα – Σοφοκλή Βενιζέλου)	77	63 (81.8 %)	10 (13.0 %)	4 (5.2%)	405	282	171	66	13	2
Σύνολο	222	198 (89.2 %)	15 (6.8 %)	9 (4.0%)	828	780	328	184	28	12

Αρχικά, στον Πίνακα 10 εμφανίζονται τα πρωτογενή στοιχεία αξιοπιστίας (accuracy, precision, recall, F1 score) για τις «1-προς-1» αντιστοιχίσεις μεταξύ προβλεπόμενων και ground truth στόχων, βάσει των κριτηρίων ομοιότητας. Όσον αφορά το βίντεο στον κόμβο Πρωτόπαπα – Σοφοκλή Βενιζέλου, προέκυψε καθολικά χειρότερη αξιολόγηση, ιδιαίτερα εμφανώς στο πεδίο ορθούς ανίχνευσης (περίπου 5-10 % κατώτερα από το μέσο όρο), όπου εκεί απέδωσε καλύτερα η εφαρμογή στο βίντεο του κυκλικού κόμβου, ομοίως με το σκέλος της ταυτοποίησης. Το βίντεο του κόμβου Ηρώων Πολυτεχνείου – Παύλου Πεντάρη εξήγαγε τις βέλτιστες τιμές των μετρικών συσχέτισης δεδομένων και χωρικού εντοπισμού, ως απόρροια του μικρότερου ύψους λήψης.

Στον Πίνακα 11 παρατίθενται ορισμένες απλές -αλλά ταυτόχρονα κρίσιμες για την αξιολόγηση- καταμετρήσεις φαινομένων τροχιάς. Όπως ήταν αναμενόμενο, οι τροχιές του βίντεο ελέγχου στον κυκλικό κόμβο απέδωσαν σημαντικά λιγότερες περιπτώσεις σφαλμάτων ταυτοποίησης (μεταβολές/ανταλλαγές/μεταφορές ID), όπως και κατακερματισμούς. Ταυτόχρονα, επιτεύχθηκε η κυρίαρχη (>80%) παρακολούθηση για το 92.6% των εξεταζόμενων τροχιών, μόλις 0.9% πίσω από την αντίστοιχη επίδοση της εφαρμογής στον κόμβο Ηρώων Πολυτεχνείου – Παύλου Πεντάρη. Σαφώς μειωμένες οι τιμές των μετρικών στα πλάνα που προήλθαν από τον κόμβο Πρωτόπαπα – Σοφοκλή Βενιζέλου, όπου παρατηρήθηκαν σχεδόν οι μισοί κατακερματισμοί και τα μισά σφάλματα ταυτοποίησης, εκ του συνόλου ελέγχου.

Τελικά, στον παρακάτω πίνακα παρουσιάζονται οι ανώτερες μετρικές ποσοτικής αξιολόγησης του MOT Benchmark:

Πίνακας 12: Αποτελέσματα ανώτερων μετρικών ποσοτικής αξιολόγησης του MOT Benchmark.

Δεδομένα ελέγχου	MOTA ↑	HOTA ↑	OWTA ↑	HOTA(0) ↑	LocA(0) ↑	MOTP ↓
30" από DJI_0073.mp4 (Κυκλικός κόμβος Πλατείας 25ης Μαρτίου)	89.2 %	80.4 %	83.8 %	90.8 %	87.9 %	0.12
30" από DJI_0095_1.mp4 (Ηρώων Πολυτεχνείου – Παύλου Πεντάρη)	83.0 %	78.8 %	84.7 %	87.4 %	89.1 %	0.11
Πλήρες DJI_0564.mp4 (Πρωτόπαπα – Σοφοκλή Βενιζέλου)	56.6 %	67.3 %	78.3 %	78.3 %	84.5 %	0.15
Σύνολο	74.3 %	74.3 %	82.0 %	84.5 %	86.9 %	0.13

Οι επιφανέστερες εξ' αυτών, MOTA και HOTA, εμφανίστηκαν ιδιαίτερα υψηλές στο βίντεο ελέγχου του κυκλικού κόμβου (~ 90 και 80 % αντίστοιχα), αρκετά πιο πάνω από τις επίσημες υποβολές στα αποτελέσματα των MOT Challenge. Σε παραπλήσια υψηλά επίπεδα διατηρήθηκε συνολικά ο έλεγχος στο βίντεο του κόμβου Ηρώων Πολυτεχνείου – Παύλου Πεντάρη, όπου αναδείχθηκε καλύτερος κατά OWTA και LocA(0), ενώ αισθητά δυσμενέστερη παρουσιάστηκε, ως συνήθως, η ποσοτική αξιολόγηση παρακολούθησης στον κόμβο Πρωτόπαπα – Σοφοκλή Βενιζέλου. Ομοιόμορφα για κάθε δοκιμή, το σφάλμα χωρικού εντοπισμού MOTP παρέμεινε χαμηλό, μόλις μεταξύ των τιμών μορφής IoU 0.10 και 0.15.

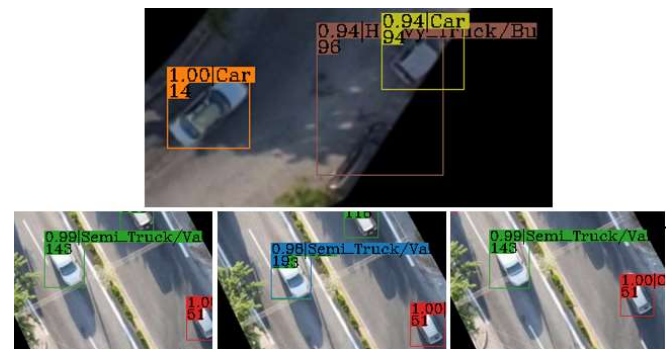
Βέβαια, τονίζεται πως δεν τέθηκε ζήτημα ορθούς σύγκρισης με τις επίσημες υποβολές στο MOT Challenge, καθώς αυτές αφορούν εφαρμογές πάνω σε ένα διαφορετικό είδος στόχων, τους διερχόμενους πεζούς, με ξεχωριστές ιδιομορφίες και προκλήσεις να διέπουν την παρακολούθησή τους. Αξιοπίστο «baseline» μπορούν να αποτελέσουν τα αποτελέσματα της μεθοδολογίας των J. Wang et al. (2019) [27], που εφαρμόζεται σε UAV πλάνα με στόχους

ενδιαφέροντος τα διερχόμενα οχήματα, ως μια μοναδική γενική κατηγορία. Η επίτευξη 89.2 % στην μετρική MOTA για το βίντεο βέλτιστης διαθέσιμης ποιότητας ανάλυσης, ξεπέρασε οριακά το υψηλότερο αντίστοιχο score της εν λόγω έρευνας (88.7 %) για το διακριτό βίντεο ελέγχου «M0101», το οποίο ωστόσο έχει σημαντικά μεγαλύτερη διάρκεια από τα 30” της διάρκειας ελέγχου στο DJI_0073.mp4. Γενικότερα, η συνολική επίτευξη 74.3 % σε MOTA και HOTA, θεωρήθηκε ένα άκρως πειστικό αποτέλεσμα για τις δυνατότητες της μεθοδολογίας, ενώ επιπρόσθετα πλήχθηκε σημαντικά από τη συμμετοχή ενός δείγματος ελέγχου κατώτερης ποιότητας οπτικής πληροφορίας, από τον κόμβο Πρωτόπαπα – Σοφοκλή Βενιζέλου.

5.1.3: Περιπτώσεις αστοχίας

Φυσικά από τη μεθοδολογία Βαθιάς Μάθησης για την Παρακολούθηση Πολλαπλών Αντικειμένων δεν απουσίασαν ορισμένες περιπτώσεις αστοχίας. Παρακάτω αναφέρονται οι συχνότερες/κυριότερες εξ αυτών:

- **Προσωρινή απώλεια στόχου:** Διακοπή παρακολούθησης της τροχιάς για λιγοστά καρέ, και άμεση επανέναρξή της. Δεν αποτέλεσε σημαντική αστοχία, έπληξε σχεδόν αμελητέα την ποσοτική αξιολόγηση, ενώ συνήθως δεν γινόταν αντιληπτή κατά την παρακολούθηση ενός βίντεο αποτελεσμάτων στα 25 fps.
- **«Στιγμιαίες» τροχιές (ή στιγμιαίο ID Switch):** Πρόκειται για ενδείξεις ανιχνεύσεων που συγκροτούν μια τροχιά με διάρκεια ζωής μόλις λιγοστών καρέ, είτε μαζί με μια υπάρχουσα ζωντανή τροχιά, είτε σε κάποια περιοχή εντελώς εκτός ground truth πληροφορίας. Ειδικά για τη δεύτερη περίπτωση, αυτό συνέβαινε



Σχήμα 45: Παραδείγματα αστοχίας με εμφάνιση «στιγμιαίων» ανιχνεύσεων ID.

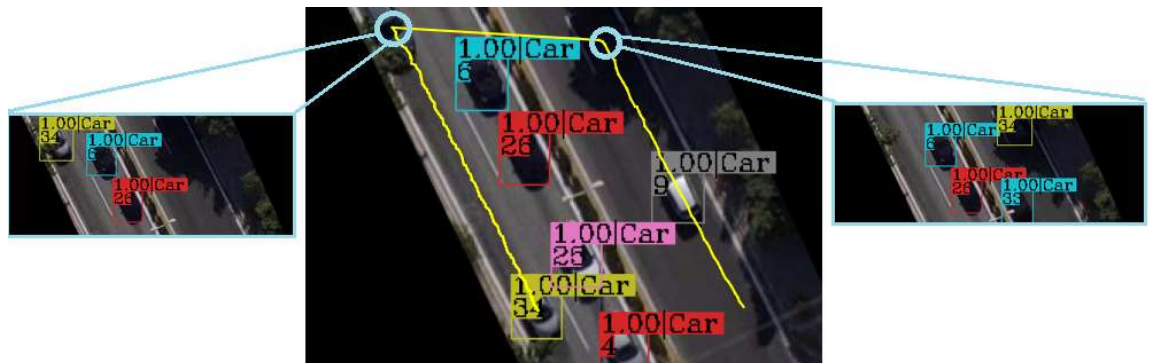
συχνά σε περιπτώσεις όπου η αλληλεπίδραση της μάσκας απομόνωσης με το απεικονιζόμενο περιβάλλον δημιουργούσε χαρακτηριστικά που προσέγγιζαν εμφανισιακά ένα όχημα. Πέρα από τη δυσχέρεια της οπτικής εποπτείας, και την αρνητική επιρροή στις μετρικές ποσοτικής αξιολόγησης, δεν επηρεάστηκε υπολογίσιμα η μεθοδολογία στο πεδίο της κυκλοφοριακής τεχνικής. Συχνότερα συνέβαινε σε στατικά οχήματα.

- **Περιπτώσεις ID Switch στον πλήρη διαχωρισμό τροχιάς:** Αποτέλεσαν αδυναμία του μοντέλου απέναντι στην απόδοση ενός μοναδικού ID σε ένα όχημα καθ' όλη τη διάρκεια της τροχιάς του, και γενικότερα η απουσία τους ήταν ιδιαίτερα επιθυμητή. Λαμβάνονται υπόψη στη μετρική MOA, οπότε δυσχέραναν την ακρίβεια της κατανομής διελεύσεων σε ζεύγη προέλευσης – προορισμού. Στην εκτίμηση της κυκλοφοριακής ζήτησης, αντιμετωπίστηκαν συμβατικά με το διαμοιρασμό των «αόριστων» διελεύσεων ευνοώντας τις συχνότερες διαδρομές, όπως προέκυπταν από τις ορθολογικές διελεύσεις. Οφειλόταν κυρίως στη μεταβολή της εμφάνισης του οχήματος – στόχου, λόγω αλλαγής φωτισμού ή διεύθυνσης κίνησης, συγκαλύψεων κ.α..



Σχήμα 46: Περίπτωση διαχωρισμού τροχιάς υπό διαφορετικά IDs. Επιπλέον έχει μεταβληθεί η πρόβλεψη κατηγορίας οχήματος του ανιχνευτή.

- **Περιπτώσεις ID Transfer:** Σημαντικά σφάλματα ταυτοποίησης, τα οποία αντιστοιχίζαν το ισοδύναμο 2 (ή -εξαιρετικά σπάνια- περισσότερων) διελεύσεων σε μόλις 1 ID στόχου, δυσχεραίνοντας τη σχετική καταμέτρηση. Κατά την παρουσίαση των μετρικών ποσοτικής αξιολόγησης, διαπιστώθηκε πως σε 222 τροχιές ελέγχου παρατηρήθηκαν 28 συμβάντα μεταφοράς ID, οπότε κατά προσέγγιση 10% των τροχιών υπέκυψε στο εν λόγω σφάλμα. Συνήθως συνέβαιναν όταν ένα όχημα εξερχόταν της περιοχής του πλάνου, αλλά εντός σύντομου χρονικού διαστήματος ένα παρόμοιο όχημα εισερχόταν στη σκηνή από το εγγύτερο αντίθετο ρεύμα. Όπως εξηγήθηκε σε προηγούμενο κεφάλαιο, η χρήση της διατομής ελέγχου βοήθησε στη συμβατική αντιμετώπιση αρκετών εξ αυτών των περιπτώσεων. Ακόμη, η ελαχιστοποίηση της παραμέτρου A_{max} θα μπορούσε να περιορίσει περαιτέρω τα μεταφερόμενα ID, αλλά ταυτόχρονα να δυσκολέψει την επιθυμητή επαναταυτοποίηση έπειτα από συγκαλύψεις.

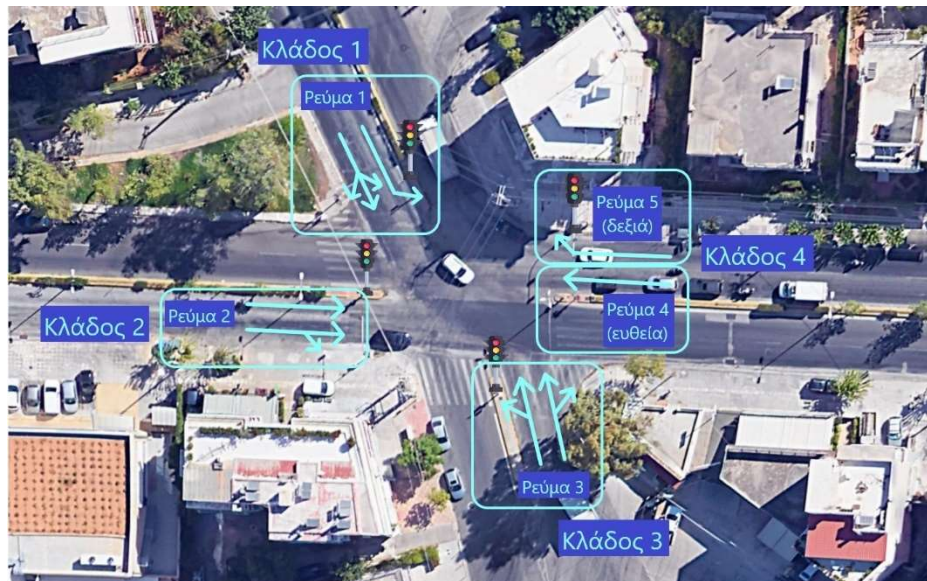


Σχήμα 47: Περίπτωση μεταφοράς του ID = 34 από τον αρχικό στόχο που εξέρχεται των ορίων βιντεοσκόπησης, σε νέο-εισερχόμενο στόχο.

- **Δυσκολία παρακολούθησης δικύκλων:** Σε αρκετά πλάνα διέρχονταν δίκυκλα οχήματα τα οποία παρουσιάζονται ως μικροσκοπικά αντικείμενα και με ιδιαίτερα δυσδιάκριτη όψη. Ενώ το ανεξάρτητο μοντέλο ανίχνευσης καταφέρνει τη διάκρισή τους ακόμη και σε αρκετά απαιτητικές εκφάνσεις τους, το μοντέλο παρακολούθησης σε τέτοιες περιπτώσεις αποδίδει συχνά διακοπτόμενες τροχιές, ή και με παρεμβολή στιγμιαίων/μόνιμων ID Switches. Γενικότερα πρόκειται για μια μερική αστοχία εντός αποδεκτών ορίων, δεδομένης της δυσκολίας της πρόκλησης.

5.2: Αποτελέσματα και αξιολόγηση κυκλοφοριακής προσέγγισης

Έπεται η παρουσίαση των αποτελεσμάτων των προγραμματιστικών μηχανισμών λειτουργικής ανάλυσης στους κόμβους μελέτης, ως παράγωγη διαδικασία της μεθοδολογίας Βαθιάς Μάθησης για την «εξόρυξη» των στοιχείων τροχιών των οχημάτων. Ωστόσο πρώτα, κρίνεται αναγκαία η υπόδειξη της τυπικής ονομασίας των κλάδων που εφαρμόστηκε, καθώς και των διαχωρισμών τους σε επιμέρους ρεύματα (ή λωρίδες):



Σχήμα 48: Διάκριση ρευμάτων στον κόμβο Ηρώων Πολυτεχνείου – Παύλου Πεντάρη, όπως ορίζει το υφιστάμενο πρόγραμμα φωτεινής σηματοδότησης. (γεωχωρικό υποβαθρο Google Earth)



Σχήμα 49: Διάκριση ρευμάτων στον κόμβο Πρωτόπαπα – Σοφοκλή Βενιζέλου, όπως ορίζει το υφιστάμενο πρόγραμμα φωτεινής σηματοδότησης. (γεωχωρικό υποβαθρο Google Earth)



Σχήμα 50: Ο κυκλικός κόμβος της Πλατείας 25^{ης} Μαρτίου φέρει 5 κλάδους, ο καθένας με 2 λωρίδες στην πρόσβασή του, ενώ υπάρχουν 2 λωρίδες εντός του δακτυλίου. Επομένως κάθε λωρίδα πρόσβασης διαθέτει διακριτή κυκλοφοριακή ικανότητα και αναλύεται ξεχωριστά. Όπως εξηγείται στην επόμενη σελίδα, θεωρείται η κατανομή φόρτου της μορφής «Left-Through» κινήσεων για την αριστερή λωρίδα και «Through-Right» κινήσεων για τη δεξιά. Για το διαμοιρασμό του φόρτου σε 2 λωρίδες προς κοινή «Through» κίνηση, εφαρμόζονται αναλογίες 50-50. Για παράδειγμα, η ζήτηση των οχημάτων που θέλουν να μετακινηθούν από τον Κλάδο 2 στον Κλάδο 4, θεωρείται ισομοιρασμένη στη δεξιά και αριστερή λωρίδα της πρόσβασης. (γεωχωρικό υποβαθρο Google Earth)

5.2.1: Αποτελέσματα κυκλοφοριακών μεγεθών και σχετικών μετρικών

Σύμφωνα με όσα περιγράφονται κατά τη μεθοδολογία, παρατίθενται οι συνολικοί (συνδυάζοντας όλες τις κατηγορίες οχημάτων) πίνακες προέλευσης – προορισμού που προκύπτουν για κάθε κυκλοφοριακό κόμβο. Αυτοί δηλώνουν τη ζήτηση σχεδιασμού βάσει της οποίας υλοποιείται η λειτουργική ανάλυση για το χαρακτηρισμό του επιπέδου εξυπηρέτησης. Σημειώνεται πως στους τελικούς υπολογισμούς έχει εφαρμοστεί στρογγυλοποίηση προς τον αμέσως μεγαλύτερο ακέραιο (np.ceil()).

Πίνακας 13: Αποτελέσματα προσδιορισμού ζήτησης σχεδιασμού υπό τη μορφή κυκλοφοριακού φόρτου [ΜΕΑ/ώρα], κατανεμημένη σε κελιά προέλευσης-προορισμού μεταξύ των κλάδων στον κόμβο Ηρώων Πολυτεχνείου – Παύλου Πεντάρη.

Ζήτηση σχεδιασμού υπό τη μορφή κυκλοφοριακού φόρτου [ΜΕΑ/ώρα]		Προέλευση			
		Κλάδος 1	Κλάδος 2	Κλάδος 3	Κλάδος 4
Προορισμός	Κλάδος 1	0	0	317	928
	Κλάδος 2	5	0	53	342
	Κλάδος 3	47	30	0	5
	Κλάδος 4	1089	219	36	0

Πίνακας 14: Αποτελέσματα προσδιορισμού ζήτησης σχεδιασμού υπό τη μορφή κυκλοφοριακού φόρτου [ΜΕΑ/ώρα], κατανεμημένη σε κελιά προέλευσης-προορισμού μεταξύ των κλάδων στον κόμβο Πρωτόπαπα – Σοφοκλή Βενιζέλου.

Ζήτηση σχεδιασμού υπό τη μορφή κυκλοφοριακού φόρτου [ΜΕΑ/ώρα]		Προέλευση			
		Κλάδος 1	Κλάδος 2	Κλάδος 3	Κλάδος 4
Προορισμός	Κλάδος 1	0	158	1236	47
	Κλάδος 2	473	0	0	303
	Κλάδος 3	936	17	0	202
	Κλάδος 4	168	319	30	0

Πίνακας 15: Αποτελέσματα προσδιορισμού ζήτησης σχεδιασμού υπό τη μορφή κυκλοφοριακού φόρτου [ΜΕΑ/ώρα], κατανεμημένη σε κελιά προέλευσης-προορισμού μεταξύ των κλάδων στον κυκλικό κόμβο Πλατείας 25ης Μαρτίου.

Ζήτηση σχεδιασμού υπό τη μορφή κυκλοφοριακού φόρτου [ΜΕΑ/ώρα]		Προέλευση				
		Κλάδος 1	Κλάδος 2	Κλάδος 3	Κλάδος 4	Κλάδος 5
Προορισμός	Κλάδος 1	0	5	0	10	0
	Κλάδος 2	49	25	69	616	123
	Κλάδος 3	142	79	64	211	726
	Κλάδος 4	79	367	88	10	115
	Κλάδος 5	0	159	340	59	0

Γνωρίζοντας την κατανομή της ζήτησης στους κόμβους εξήχθησαν ορισμένα ποιοτικά συμπεράσματα για την κυκλοφορία:

- **Στον Κυκλικό κόμβο Πλατείας 25ης Μαρτίου:**
 - Παρατηρήθηκε σημαντικά μικρότερη ζήτηση για διελεύσεις που ξεκινούν από μια πρόσβαση και καταλήγουν στην αμέσως επόμενη, οπότε διαμορφώσεις παράκαμψης του κόμβου δε θα εξυπηρετούσαν ιδιαίτερα αποδοτικά.
 - Οι πιο «επιβαρυνόμενες» διαδρομές που ξεκινούν από κάθε κλάδο, ήταν αυτές που κατέληγαν «αντιδιαμετρικά» του δακτυλίου. Οπότε, σύμφωνα με HCM 2010, θα μπορούσε να θεωρηθεί ότι στην πράξη η ανάθεση φόρτου γίνεται ως «Left-Through» για την αριστερή λωρίδα και ως «Through-Right» για τη δεξιά, καθώς οι οδηγοί επιλέγοντας την πρόσκαιρα ταχύτερη λωρίδα προκαλούν «μοίρασμα» της αυξημένης «Through» ζήτησης και στις δύο λωρίδες.

- **Στον κόμβο Φ/Σ Ηρώων Πολυτεχνείου – Παύλου Πεντάρη:**
 - Παρατηρήθηκε ότι η μεγαλύτερη ζήτηση σχηματιζόταν από τις μετακινήσεις «Κλάδος 1 → Κλάδος 4» και αντίστροφα. Έγινε εμφανές πως στο συγκεκριμένο κόμβο δεν υπήρχε η τυπική μορφή διασταύρωσης ενός κύριου άξονα ανώτερης ζήτησης, κι ενός δευτερεύοντα κατώτερης. Το κυρτό οδικό τμήμα «Κλάδος 1 – Κλάδος 4» αποτελούσε την αρτηρία δεσπόζουσας ροής, πάνω στην οποία προσαρμόζονταν δύο «λανθάνοντες» κλάδοι, οι 2 και 3.
 - Εφόσον από τον Κλάδο 4 απαγορευόταν βάσει σήμανσης η αριστερή στρέφουσα κίνηση προς τον Κλάδο 3, θα έπρεπε το αντίστοιχο κελί στον πίνακα να είναι μηδενικό κι όχι ίσο με 5. Το μη μηδενικό κελί, επομένως, δήλωνε παράνομες διελεύσεις. Πράγματι, εντός των περίπου 14' του σχετικού βίντεο πραγματοποιήθηκε μια τέτοια κίνηση, και κατέστη εφικτή καθώς το «εχθρικό» Ρεύμα 2 δε φέρει ιδιαίτερη ζήτηση. Οπότε κατά την αναγωγή σε ωριαία βάση και μαζί με την προς-τα-άνω στρογγυλοποίηση προέκυψε το ποσό 5.
- **Στον κόμβο Φ/Σ Πρωτόπαπα – Σοφοκλή Βενιζέλου:**
 - Σε αντιπαραβολή με τον κόμβο Ηρώων Πολυτεχνείου – Παύλου Πεντάρη, στον παρόντα παρατηρήθηκε αυξημένη ζήτηση για τη μετακίνηση «Κλάδος 1 → Κλάδος 3» και αντίστροφα, διακρίνοντας τη Λεωφ. Πρωτόπαπα ως κύριο άξονα.
 - Δεν παρατηρήθηκαν στη συγκεκριμένη περίπτωση παράνομες στρέφουσες κινήσεις στο Ρεύμα 4, γεγονός που αιτιολογείται από το ότι θα ήταν πρακτικά αδύνατο να βρεθεί ο χρονικός διαχωρισμός για ασφαλή διέλευση ανάμεσα στην αρκετά υψηλή ζήτηση του αντίθετου Ρεύματος 2.

Η αξιοπιστία του ποσού εκτιμώμενης ζήτησης χαρακτηρίζεται από τις μετρικές ορθολογικού υπολογισμού και απόλυτης ποσοστιαίας απόκλισης που ορίστηκαν στο υποκεφάλαιο 4.6.3. Τα αποτελέσματά τους στην παρούσα εφαρμογή ήταν τα εξής:

Πίνακας 16: Αποτελέσματα μετρικών αξιολόγησης εκτίμησης κυκλοφοριακού φόρτου.

Κόμβος	Assurance Index ↑	Overall Volume Deviation ↓
Κυκλικός κόμβος Πλατείας 25ης Μαρτίου (εντός του δακτυλίου)	74.3 %	4.7 %
Πολυτεχνείου – Παύλου Πεντάρη	79.6 %	3.6 %
Πρωτόπαπα – Σοφοκλή Βενιζέλου	79.4 %	0.8 %

Η εφαρμογή του μοντέλου παρακολούθησης στο βίντεο του κυκλικού κόμβου μπορεί να διαθέτει την καλύτερη επίδοση στις ανώτερες μετρικές ανίχνευσης – παρακολούθησης του MOT Benchmark, αλλά εδώ φάνηκε πως αποδίδει ελαφρώς χειρότερα από τις άλλες δύο ακολουθίες σε μια αξιολόγηση ικανότητας καταμέτρησης συνεχών τροχιών. Αυτό αιτιολογείται από το γεγονός ότι για να οριστεί μια ορθολογική καταμέτρηση διέλευσης στην περίπτωση του κυκλικού κόμβου, η τροχιά του στόχου έπρεπε να παραμείνει ζωντανή για σημαντικά μεγαλύτερο μήκος επί της εικόνας, με ανάλογη αυξημένη διάρκεια σε καρέ, από ότι στις περιπτώσεις των άλλων δύο κόμβων. Παρ’ όλα αυτά, ο μηχανισμός προσμέτρησης της μη κατανεμημένης ζήτησης φαίνεται πως συνέβαλε καταλυτικά ώστε ο συνολικός κυκλοφοριακός φόρτος που προβλέφθηκε να διαφέρει κατά λιγότερο από 5% με την πραγματικότητα, σε όλες τις περιπτώσεις. Μη συμπερίληψή του αποδεικνύεται πως θα συνεπαγόταν αποκλίσεις της τάξεως του 20%.

Ακόμη, παρουσιάζονται τα αποτελέσματα εκτίμησης των δευτερευόντων κυκλοφοριακών μεγεθών, τα οποία δε συμμετέχουν στον προσδιορισμό του επιπέδου εξυπηρέτησης, αλλά ωστόσο η προοπτική υπολογισμού τους εξακολουθεί να αποτελεί ενδιαφέρον.

Πίνακας 17: Αποτελέσματα προσδιορισμού δευτερευόντων κυκλοφοριακών μεγεθών (δε συμμετέχουν στην εκτίμηση επιπέδων εξυπηρέτησης).

Κόμβος	Μέση πυκνότητα [ΜΕΑ/γλμ.]	Μέσος χωρικός διαχωρισμός [μέτρα]	Μέση ταχύτητα διαδρομής [γλμ./ώρα]	Εκτίμηση μέσου χρονικού διαχωρισμού (μέσος χωρικός διαχωρισμός / μέση ταχύτητα διαδρομής) [δευτ.]
Κυκλικός κόμβος Πλατείας 25ης Μαρτίου (εγγύτερα του δακτύλιου)	45	22.2	19.7	4.1
Πολυτεχνείου – Παύλου Πεντάρη	27	37.0	28.6	4.7
Πρωτόπαπα – Σοφοκλή Βενιζέλου	39	25.6	21.5	4.3

Όπως ήταν αναμενόμενο, η μέση ταχύτητα διαδρομής των οχημάτων στον κυκλικό κόμβο εκτιμήθηκε χαμηλότερη, καθώς η κίνησή τους γινόταν σε καμπύλο οδικό τμήμα, και στην πλειοψηφία των περιπτώσεων εφόσον είχε προηγηθεί μερική ακινητοποίηση αναμένοντας χωρικό διάκενο για την είσοδο στο δακτύλιο. Στους κόμβους Φ/Σ τα οχήματα μπορούσαν να διέλθουν με αρκετά υψηλή ταχύτητα αν δεν αναγκαζόντουσαν σε στάση λόγω κόκκινης ένδειξης, αλλά το γεγονός αυτό ισοσταθμίζεται από τη διάρκεια που παρέμεναν εντελώς ακινητοποιημένα εντός ενός κύκλου, οπότε η διαφορά δεν απεδείχθη μεγάλη. Επίσης, η ύπαρξη μεγαλύτερης μέσης πυκνότητας και μικρότερης μέσης ταχύτητας διαδρομής στον κόμβο Πρωτόπαπα – Σοφοκλή Βενιζέλου, αποτέλεσε μια πρώιμη ένδειξη ότι ο συγκεκριμένος κόμβος θα αποδίδει κατώτερο επίπεδο εξυπηρέτησης από τον ομόλογό του στη συμβολή Ηρώων Πολυτεχνείου – Παύλου Πεντάρη, γνωρίζοντας ότι οι συνολικές διελεύσεις δε διέφεραν υπερβολικά μεταξύ τους.

5.2.2: Τελική κατάταξη σε στάθμη εξυπηρέτησης

Τα αποτελέσματα της λειτουργικής ανάλυσης (μέση καθυστέρηση, στάθμη εξυπηρέτησης) εμφανίζονται συγκεντρωτικά στον παρακάτω πίνακα, σε όλα τα διαφορετικά επίπεδα αναγωγής (ρεύμα/λωρίδα, κλάδος/πρόσβαση, κόμβος συνολικά).

Πίνακας 18: Συγκεντρωτικά αποτελέσματα των εξαγόμενων χαρακτηρισμών επιπέδου εξυπηρέτησης.

Κόμβος	Μέση καθυστέρηση [sec/MEA]	Στάθμη (Α έως F)	Κλάδος/ Πρόσβαση	Μέση καθυστέρηση [sec/MEA]	Στάθμη (Α έως F)	Ρεύμα/ Λωρίδα	Βαθμός κορεσμού [V/c]	Μέση καθυστέρηση [sec/MEA]	Στάθμη (Α έως F)
Κυκλικός κόμβος Πλατείας 25ης Μαρτίου	39.4	E	1	37.0	E	δεξιά λωρίδα	0.67	41.9	E
						αριστερή λωρίδα	0.30	24.9	C
			2	55.3	F	δεξιά λωρίδα	0.60	25.2	D
						αριστερή λωρίδα	0.91	76.5	F
			3	23.5	C	δεξιά λωρίδα	0.52	26.5	D
						αριστερή λωρίδα	0.65	20.0	C
			4	22.3	C	δεξιά λωρίδα	0.53	15.7	C
						αριστερή λωρίδα	0.77	27.0	D
			5	55.0	F	δεξιά λωρίδα	0.87	48.8	E
						αριστερή λωρίδα	0.90	60.1	F
Ηρώων Πολυτεχνείου – Παύλου Πεντάρη	18.7	C	1	10.8	B	1	0.51	10.8	B
			2	21.3	C	2	0.28	21.3	C
			3	40.1	E	3	0.72	40.1	E
			4	18.4	C	4	0.74	37.8	D
			5	0.67	11.1	B			
Πρωτόπαπα – Σοφοκλή Βενιζέλου	118.0	F	1	134.3	F	1	0.69	39.9	D
			2	72.3	F	2	1.16	145.6	F
			3	17.8	C	3	0.93	72.3	F
			4	167.4	F	4	0.61	17.8	C
						5	1.17	167.4	F

Γενικότερα τα αποτελέσματα που προέκυψαν ανταποκρίνονται ρεαλιστικά στην ερμηνεία των επιπέδων για την κατάσταση της κυκλοφορίας, οπότε μπορεί να ειπωθεί πως η εξαγόμενη κατανομή ζήτησης δεν απεδείχθη άστοχη. Όπως μαρτυρά η παράθεση του βαθμού κορεσμού ανά ρεύμα, δηλώνεται η αιτιότητα μεταξύ προσέγγισης συνθηκών συμφόρησης (δηλαδή όταν η ζήτηση πλησιάζει και υπερβαίνει την κυκλοφοριακή ικανότητα) και χαρακτηρισμού «F» στο επίπεδο εξυπηρέτησης. Οι αριστερές λωρίδες των κλάδων 2 και 5 στον κυκλικό κόμβο, και τα ρεύματα 2,3 και 5 στον κόμβο Πρωτόπαπα – Σοφοκλή Βενιζέλου απέδιδαν μεγάλες ουρές οχημάτων καθώς αδυνατούσαν να εκκενώσουν έγκαιρα τα οχήματα. Στη δεύτερη περίπτωση το φαινόμενο ξεχώριζε περισσότερο με την ύπαρξη επανειλημμένων «αποτυχημένων» κύκλων από το πρόγραμμα φωτεινής σηματοδότησης.

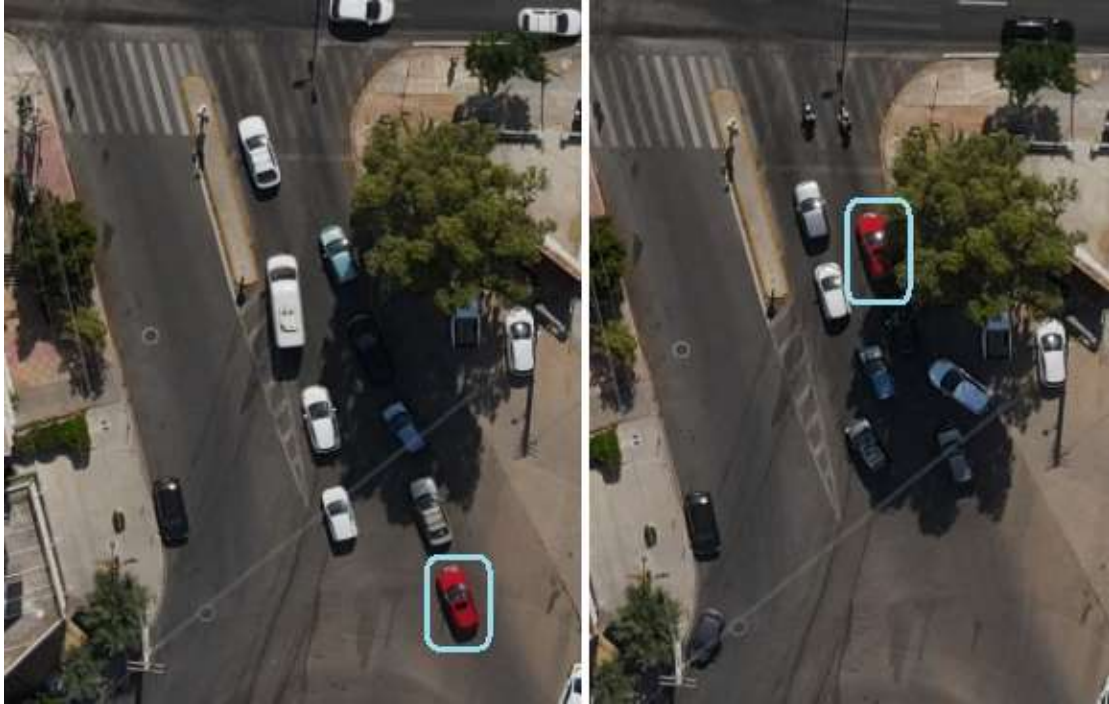


Σχήμα 52: Παρατήρηση συνθηκών κυκλοφοριακής συμφοράς στον Κλάδο 2, τόσο στον κυκλικό κόμβο (άνω), όσο και στον κόμβο Πρωτόπαπα – Σοφοκλή Βενιζέλου (κάτω).



Σχήμα 51: Περίπτωση κατά την οποία ένα όχημα δεν κατάφερε να διέλθει του κόμβου Πρωτόπαπα – Σοφοκλή Βενιζέλου, παρότι αντίκρισε 3 κόκκινες ενδείξεις στο Ρεύμα 5 (οι 2 ολοκληρωμένες διάρκειας + 1 περιορισμένης κατά την άφιξή του).

Αντίστοιχα, στον κόμβο Πολυτεχνείου – Παύλου Πεντάρη ένα σημαντικό ποσοστό των οχημάτων διερχόταν δίχως να αναγκαστεί σε στάση από κόκκινη ένδειξη φωτεινού σηματοδότη. «Αποτυχημένοι» κύκλοι συνέβησαν ιδιαίτερα σπάνια, και μόνο στο Ρεύμα 3 που άλλωστε απέδωσε τη χειρότερη στάθμη εξυπηρέτησης («E») στο συγκεκριμένο κόμβο, λόγω της μικρής διάρκειας πράσινης ένδειξης που μείωσε σημαντικά την κυκλοφοριακή ικανότητα.



Σχήμα 53: Περίπτωση «αποτυχημένου» κύκλου στο Ρεύμα 3 του κόμβου Ηρώων Πολυτεχνείου – Παύλου Πεντάρη.

5.3: Αποτίμηση: προοπτικές βελτίωσης και συμπεράσματα

Οι προοπτικές για τη βελτίωση των επιδόσεων της παρούσας μεθοδολογίας βασίζονται κυρίως στη διαθεσιμότητα επαρκούς υπολογιστικής ισχύος για ενίσχυση της αποδοτικότητας των τεχνικών Βαθιάς Μάθησης. Σε ένα «mid-range» υπολογιστικό σύστημα οικιακής χρήσης, με κάρτα γραφικών μόλις 8 GB χωρητικότητας VRAM έγιναν σαφέστατα κάποιες εκπτώσεις στην ποιότητα εκτέλεσης των επιμέρους διεργασιών, αλλά παρ' όλα αυτά αποδείχθηκε πως το εγχείρημα παρακολούθησης πολλαπλών οχημάτων ολοκληρώνεται επιτυχώς, και μάλιστα με ικανοποιητικά αποτελέσματα. Ωστόσο, με παραπάνω διαθέσιμη VRAM (12-16 GB) καθίστανται δυνατά τα εξής:

- Εκπαίδευση του μοντέλου ανίχνευσης με resize των εικόνων στο μεγαλύτερο μέγεθος του dataset (4K) ανά αυξημένο batch size, με προσθήκη των επιπέδων μεγέθους αγκυρώσεων 2 και 4, και αύξηση του RoI feature size σε 49 (από 18).
- Πειραματισμός με νεότερες αρχιτεκτονικές ανιχνευτών μεγαλύτερων απαιτήσεων μνήμης, όπως η YOLOX – Large.

Ακόμη, με μοντέλο GPU ανώτερης αρχιτεκτονικής και περισσότερων πυρήνων CUDA:

- Χρονικά αποδοτική εκπαίδευση επαναταυτοποίησης σε σημαντικά μεγαλύτερο σετ δεδομένων τροχιών, ώστε να χρησιμοποιείται από το μοντέλο παρακολούθησης ένα CNN με βέλτιστα αφομοιωμένα χαρακτηριστικά εμφάνισης.
- Χρονικά αποδοτική εκπαίδευση ανιχνευτή με πολλαπλούς τυχαίους σχηματισμούς ενίσχυσης δεδομένων, για τη σύσταση πιο εύρωστου και ικανοποιητικά εφαρμόσιμου μοντέλου σε «ξένο» dataset.
- Ταχύτερη εξαγωγή προβλέψεων παρακολούθησης, με την προοπτική να προκύπτει αξιολόγηση οδικού συστήματος ακόμη και σε πραγματικό χρόνο (στο παρόν υπολογιστικό σύστημα μόλις ~3 καρέ/δευτερόλεπτο για tracking inference).

Ακόμη και χωρίς τα παραπάνω, τόσο η ποιοτική και ποσοτική αξιολόγηση του μοντέλου παρακολούθησης πολλαπλών αντικειμένων, όσο και η πιστότητα στην εκτίμηση κυκλοφοριακών μεγεθών (ιδίως της ζήτησης), και κατά συνέπεια του επιπέδου εξυπηρέτησης των οδικών συστημάτων, ανέδειξαν τη συνολική αποτελεσματικότητα της εφαρμογής στο πεδίο της κυκλοφοριακής τεχνικής. Μπορεί με βεβαιότητα να ειπωθεί πως η παρούσα μεθοδολογία με τη χρήση δεδομένων βίντεο από UAV μέσα, αξίζει να καθιερωθεί στο δυναμικό των πιθανών μεθόδων παρατηρήσεων συγκοινωνιακών μελετών, με ορισμένα σημαντικά συγκριτικά πλεονεκτήματα έναντι των συχνά χρησιμοποιούμενων σχηματισμών επαγωγικών βρόγχων και σωλήνων πίεσης:

- **Οικονομικο-τεχνικά αποδοτική μέθοδος**, μόλις περίπου 600 € αξία για ένα DJI Mavic Air 2 UAV, το οποίο μπορεί να επαναχρησιμοποιηθεί για μεγάλο αριθμό πτήσεων, και να καλύψει μια ευρεία περιοχή σε ένα οδικό σύστημα από εναέρια προοπτική. Ένας ηλεκτρονικός υπολογιστής οικιακής χρήσης στο εύρος κόστους 1000 – 1500 € είναι υπεραρκετός για την ασύγχρονη εκτέλεση των παρόντων μοντέλων στο στάδιο της μετα-επεξεργασίας. Σαφώς, οι απαιτήσεις υπολογιστικής ισχύος, ταχύτητας τηλεπικοινωνιών, και κατά συνέπεια κόστους, αυξάνονται σε περίπτωση που υπάρχει ανάγκη για ανάλυση σε πραγματικό χρόνο, η οποία πάντως καθίσταται εφικτή. Από την άλλη, η εγκατάσταση ανιχνευτών επαγωγικών βρόγχων κοστολογείται στη «γειτονιά» των 2000 € ανά τεμάχιο (ζεύγος), ενώ αντίστοιχα για σωλήνες πίεσης λίγο κάτω από τα 1000 €, διενεργώντας το πολύ σε μια διατομή [39]. Οπότε, το κόστος πολλαπλασιάζεται ανάλογα με την πολυπλοκότητα του οδικού συστήματος, και ειδικά για την περίπτωση των κόμβων μελέτης, θα ήταν σημαντικά αυξημένο για την κάλυψη όλων των ρευμάτων.
- Πρόκειται για μια **ευέλικτη και φορητή λύση**, καθώς μπορεί να προγραμματιστεί εύκολα για πτήση σε οποιαδήποτε περιοχή επιτρέπει η σχετική νομοθεσία, απασχολώντας μονάχα ένα χειριστή, δίχως να χρειάζεται οποιαδήποτε μόνιμη εγκατάσταση ή επέμβαση στο περιβάλλοντα χώρο και την κυκλοφορία. Επιπλέον, η επανάληψη ενός πειράματος παρατηρήσεων χαρακτηρίζεται ιδιαίτερα απλή, καθώς δεν απαιτείται «πεισματική» ακρίβεια στην αναπαράσταση των συνθηκών λήψης (φωτισμός, ύψος πτήσης). Το «παρελκόμενο» μοντέλο Βαθιάς Μάθησης μπορεί να εξάγει ικανοποιητικά αποτελέσματα σε ελαφρώς διαφορετικά σκηνικά, ή -προτιμότερα- δύναται να προσαρμοστεί στα νέα δεδομένα με μερικές εποχές εκπαίδευσης συντονισμού του detector, με αρχικοποίηση σε υπάρχον σχετικό checkpoint.
- Η προσφερόμενη ευελιξία συνδυάζεται με την **παροχή σχεδόν στατικών κατακόρυφων πλάνων**, μεγιστοποιώντας τη **χωρική αποδοτικότητα** με την υποκατάσταση των σταθερά τοποθετημένων υπερυψωμένων καμερών κλειστού κυκλώματος.
- Η **απόδοση οπτικής πληροφορίας** χρησιμεύει πάντοτε για την ποιοτική αξιολόγηση της αξιοπιστίας των παρατηρήσεων. Δηλαδή, καθίσταται δυνατή η διασταύρωση των ποσοτικών αποτελεσμάτων, με την «εικόνα» των πραγματικών συνθηκών του κόμβου, ώστε να εξακριβώνεται η ορθή διεξαγωγή της μεθοδολογίας.

- Παρέχεται **υψηλή ορθότητα ταξινόμησης οχημάτων σε κατηγορίες βάσει των χαρακτηριστικών εμφάνισής τους**, κι όχι σύμφωνα με περιστασιακά διαφορούμενα στοιχεία, όπως η επαγωγική απόκριση και το βάρος τους.
- Η **συνεισφορά στον τομέα των συγκοινωνιακών** δεν περιορίζεται μονάχα στην αξιολόγηση ισόπεδων οδικών κόμβων. Οι εξαγόμενες τροχιές και η ζήτηση, όπως και τα «δευτερεύοντα» κυκλοφοριακά μεγέθη που υπολογίστηκαν ενδεικτικά (μέση πυκνότητα, ταχύτητα διαδρομής, διαχωρισμοί), θα μπορούσαν να αξιοποιηθούν σε εφαρμογές αξιολόγησης αυτοκινητοδρόμων/υπεραστικών οδών, ή στην ανίχνευση οδικών παραβάσεων, με κατάλληλη προσαρμογή του σκέλους υπολογισμών.

Στον αντίποδα, τα UAVs διακατέχονται από ορισμένα ζητήματα λειτουργικής αξιοπιστίας [40]:

- Η **χαμηλή ενεργειακή αυτονομία**, η οποία δεν ξεπερνά τα 30 με 35 λεπτά σε ιδανικές συνθήκες, δυσχεραίνει σημαντικά το εγχείρημα παροχής «ζωντανής» μετάδοσης πλάνων σε συνεχείς χρονικές περιόδους για ανάλυση των κυκλοφοριακών συνθηκών σε πραγματικό χρόνο. Απαιτείται προγραμματισμός ταυτόχρονων πτήσεων πολλαπλών UAVs, ώστε να υπάρχει διαθέσιμη εφεδρεία μετάδοσης, όσο γίνεται ανεφοδιασμός των μπαταριών κάποιων εκ των συσκευών. Ωστόσο, η αυτονομία ενός UAV είναι υπεραρκετή για εργασίες που απαιτούν σύντομες περιόδους παρατηρήσεων, όπως στην παρούσα διπλωματική εργασία, όπου δύναται τα παρατηρούμενα μεγέθη να προβληθούν σε ευρύτερη χρονική βάση (π.χ. από μετρήσεις 10-15', αναγωγή σε 1 ώρα).
- Τίθεται **περιορισμός βεληνεκούς επικοινωνίας μερικών χιλιομέτρων** (8-10 χιλιόμετρα για τα UAVs που χρησιμοποιήθηκαν για το dataset μελέτης) μεταξύ χειριστή και συσκευής, αλλά στην πράξη πολύ δύσκολα θα κριθεί μη επαρκές.
- **Εξαρτώνται άμεσα από τις καιρικές συνθήκες**, καθώς η ύπαρξη ισχυρών ανέμων και ατμοσφαιρικών κατακρημνίσεων ή πυκνής ομίχλης, μπορούν είτε να επηρεάσουν την ποιότητα καταγραφής των βίντεο, είτε να αναγκάσουν ματαίωση πτήσεων.
- Η **ολική εξάρτηση από την εναέρια διαθέσιμη οπτική πληροφορία**, περιορίζει τη νυχτερινή (ή υπό σκίαση) βιντεοσκόπηση και αποτρέπει την παρακολούθηση στεγασμένων οδικών τμημάτων. Ωστόσο σε τέτοιες περιπτώσεις θα μπορούσε να εξεταστεί η συνδρομή από ενεργητικούς δέκτες εκτός του ορατού φάσματος, ή από θερμικές κάμερες.

Συμπερασματικά, τα UAVs ανταλλάσσουν τις επιδόσεις γενικής αυτονομίας χάριν της ευελιξίας και οικονομικο-τεχνικής αποδοτικότητας. Αποτελούν σχετικά πρόσφατη μέθοδο καταγραφής κυκλοφοριακών δεδομένων, με νομοθετικό πλαίσιο που ακόμη εξελίσσεται στις περισσότερες χώρες, ενώ η αποδοχή τους από σεβαστή μερίδα του πληθυσμού χαρακτηρίζεται από μερική δυσπιστία, ιδίως για λόγους προστασίας της ιδιωτικής ζωής [41]. Θα ήταν εύλογο οι κατασκευαστικές καινοτομίες να επικεντρωθούν στην προοπτική ενίσχυσης της λειτουργικής αξιοπιστίας των UAV, ώστε να εξελιχθούν σταδιακά σε δεσπόζοντα μέσα εξαγωγής κυκλοφοριακών δεδομένων.

Όσον αφορά τη συμβατή υλοποιημένη μεθοδολογία Βαθιάς Μάθησης, **δύναται να λειτουργήσει ικανοποιητικά σε κατακόρυφες λήψεις** από μελλοντικές έρευνες οποιασδήποτε τοποθεσίας, είτε εναέριες είτε από υπερυψωμένο σημείο θέασης, αξιοποιώντας τη δυνατότητα μεταφοράς μάθησης σε νέα – προσαρμοσμένα για κάθε περιοχή – μοντέλα ανίχνευσης (ή και επαναταυτοποίησης), με αρχικοποίηση στα checkpoints που δημιουργήθηκαν από το παρόν project. Για αύξηση των πιθανοτήτων αποδοτικής γενίκευσης εκ του παραχθέντος μοντέλου ανίχνευσης, ενδείκνυται η εκτέλεση πτήσεων λήψης μεταξύ 100 και 200 μέτρων, ώστε το επίπεδο λεπτομέρειας και η κλίμακα των αντικειμένων να είναι παραπλήσια. Ωστόσο, κρίνεται αδύνατη η εφαρμογή σε πλήρως επίγεια πλάνα, προφανώς λόγω ασυμβατότητας των χαρακτηριστικών των οχημάτων από το πεδίο όρασης ενός πεζού, σε σύγκριση με την εναέρια «bird's eye» όψη τους. Θα αποτελούσε ενδιαφέρον μελλοντικό project η υλοποίηση ενός μοντέλου ανίχνευσης το οποίο θα περιλαμβάνει στη λίστα ανιχνεύσιμων κατηγοριών του πολλαπλές εκδοχές ενός τύπου οχήματος, ανάλογα τις διαφορετικές προοπτικές λήψης (π.χ. Αυτοκίνητο – κατακόρυφη λήψη, Αυτοκίνητο – πλάγια λήψη, Αυτοκίνητο – εμπρόσθια λήψη, Αυτοκίνητο – οπίσθια λήψη, κ.ο.κ.), ώστε να δύναται να εφαρμοστεί σε συνδυασμό διαφορετικών ειδών πλάνων. Κατ' επέκταση, για τη μετάβαση σε παρόμοιο αλγόριθμο παρακολούθησης, θα χρειαζόταν εκπαίδευση επαναταυτοποίησης μεταξύ εικόνων όπου ένα όχημα εμφανίζεται προοδευτικά υπό διαφορετική γωνία θέασης, ή και κλίμακα.

Βιβλιογραφικές αναφορές

- [1] "Outline of Artificial Intellinence," Wikipedia, <https://w.wiki/7bjv> (προσπελάστηκε Αυγ. 2023).
- [2] MOT Challenge, <https://motchallenge.net/> (προσπελάστηκε Αυγ. 2023).
- [3] (PDF) L. Leal-Taixé, A. Milan, I. Reid, S. Roth, and K. Schindler, "MOTChallenge 2015: Towards a benchmark for multi-target tracking," arXiv.org, <https://arxiv.org/abs/1504.01942> (προσπελάστηκε Αυγ. 2023).
- [4] (PDF) A. Milan, L. Leal-Taixe, I. Reid, S. Roth, and K. Schindler, "MOT16: A Benchmark for Multi-Object Tracking," arXiv.org, <https://arxiv.org/abs/1603.00831> (προσπελάστηκε Αυγ. 2023).
- [5] (PDF) P. Dendorfer et al., "MOT20: A benchmark for multi object tracking in crowded scenes," arXiv.org, <https://arxiv.org/abs/2003.09003> (προσπελάστηκε Αυγ. 2023).
- [6] (PDF) P. Voigtlaender et al., "MOTS: Multi-object tracking and segmentation," arXiv.org, <https://arxiv.org/abs/1902.03604> (προσπελάστηκε Αυγ. 2023).
- [7] B. A. Forouzan and F. Mosharraf, "18," in *Foundations of Computer Science*, Australia: Thomson Learning, 2012.
- [8] "Multilayer Perceptron," ScienceDirect Topics, <https://www.sciencedirect.com/topics/computer-science/multilayer-perceptron> (προσπελάστηκε Αυγ. 2023).
- [9] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998. doi:10.1109/5.726791.
- [10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional Neural Networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017. doi:10.1145/3065386.
- [11] (PDF) K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv.org, <https://doi.org/10.48550/arXiv.1409.1556> (προσπελάστηκε Αυγ. 2023).
- [12] C. Szegedy et al., "Going deeper with convolutions," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015. doi:10.1109/cvpr.2015.7298594.
- [13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016. doi:10.1109/cvpr.2016.90.
- [14] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and*

- Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.
doi:10.1109/tpami.2016.2577031.
- [15] P. Bergmann, T. Meinhardt, and L. Leal-Taixe, “Tracking without bells and whistles,” *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019.
doi:10.1109/iccv.2019.00103.
- [16] (PDF) N. Wojke, A. Bewley, and D. Paulus, “Simple online and realtime tracking with a Deep Association metric,” arXiv.org, <https://arxiv.org/abs/1703.07402> (προσπελάστηκε Αυγ. 2023).
- [17] Y. Zhang, C. Wang, X. Wang, W. Zeng, and W. Liu, “Fairmot: On the fairness of detection and re-identification in multiple object tracking,” *International Journal of Computer Vision*, vol. 129, no. 11, pp. 3069–3087, 2021. doi:10.1007/s11263-021-01513-4.
- [18] X. Zhou, V. Koltun, and P. Krähenbühl, “Tracking objects as points,” *Computer Vision – ECCV 2020*, pp. 474–490, 2020. doi:10.1007/978-3-030-58548-8_28.
- [19] Y. Zhang *et al.*, “ByteTrack: Multi-object tracking by associating every detection box,” *Lecture Notes in Computer Science*, pp. 1–21, 2022. doi:10.1007/978-3-031-20047-2_1.
- [20] T. Fischer *et al.*, “QDTrack: Quasi-dense similarity learning for appearance-only multiple object tracking,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–13, 2023. doi:10.1109/tpami.2023.3301975.
- [21] Z. Lu, V. Rathod, R. Votel, and J. Huang, “RetinaTrack: Online Single Stage Joint Detection and tracking,” *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020. doi:10.1109/cvpr42600.2020.01468.
- [22] A. Bewley, Z. Ge, L. Ott, F. Ramos, and B. Upcroft, “Simple online and realtime tracking,” *2016 IEEE International Conference on Image Processing (ICIP)*, 2016.
doi:10.1109/icip.2016.7533003.
- [23] Ι. Φραντζεσκάκης, Ι. Γκόλιας, Μ. Πιτσιάβα-Λατινοπούλου, *Κυκλοφοριακή Τεχνική*. Αθήνα: Α. ΠΑΠΑΣΩΤΗΡΙΟΥ & ΣΙΑ Ι.Κ.Ε., 2009.
- [24] Transportation Research Board 2010 Executive Committee, *Highway Capacity Manual 2010*, vol. 4, 4 Vols. Washington, D.C.: Transportation Research Board of the National Academies, 2010.
- [25] B. Coifman, M. McCord, R. G. Mishalani, M. Iswalt, and Y. Ji, “Roadway traffic monitoring from an unmanned aerial vehicle,” *IEE Proceedings - Intelligent Transport Systems*, vol. 153, no. 1, p. 11, 2006. doi:10.1049/ip-its:20055014.
- [26] N. V. Kim and M. A. Chervonenkis, “Situation control of unmanned aerial vehicles for Road Traffic Monitoring,” *Modern Applied Science*, vol. 9, no. 5, 2015.
doi:10.5539/mas.v9n5p1.
- [27] J. Wang, S. Simeonova, and M. Shahbazi, “Orientation- and scale-invariant multi-vehicle detection and tracking from unmanned aerial videos,” *Remote Sensing*, vol. 11, no. 18, p. 2155, 2019. doi:10.3390/rs11182155.

- [28] E. Barmounakis and N. Geroliminis, “On the new era of urban traffic monitoring with massive drone data: The pneuma large-scale field experiment,” *Transportation Research Part C: Emerging Technologies*, vol. 111, pp. 50–71, 2020. doi:10.1016/j.trc.2019.11.023.
- [29] R. Makrigiorgis, P. Kolios, and C. Kyrkou, “Aerial Multi-vehicle detection dataset,” Zenodo, <https://zenodo.org/record/7053442> (προσπελάστηκε Αυγ. 2023).
- [30] (PDF) M. Michalas, Δ.Ε. : “Ανάπτυξη μεθοδολογίας αξιοποίησης κυκλοφοριακών δεδομένων από Μη Επανδρωμένα Οχήματα,” DSpace - Κεντρική Βιβλιοθήκη Ε.Μ.Π., <https://dspace.lib.ntua.gr/xmlui/handle/123456789/55554> (προσπελάστηκε Αυγ. 2023).
- [31] A. Spannauer, “PYTHON_VIDEO_STAB: A python package to stabilize videos using opencv,” GitHub, https://github.com/AdamSpannbauer/python_video_stab (προσπελάστηκε Αυγ. 2023).
- [32] OpenMMLab Contributors, “OpenMMLab Toolbox,” GitHub, <https://github.com/open-mmlab> (προσπελάστηκε Άνοιξη 2023).
- [33] T.-Y. Lin *et al.*, “Microsoft Coco: Common Objects in Context,” *Computer Vision – ECCV 2014*, pp. 740–755, 2014. doi:10.1007/978-3-319-10602-1_48.
- [34] CVAT, <https://www.cvat.ai/> (προσπελάστηκε Άνοιξη. 2023).
- [35] J. Solawetz, “Mean average precision (MAP) in object detection,” Roboflow Blog, <https://blog.roboflow.com/mean-average-precision/> (προσπελάστηκε Αυγ. 2023).
- [36] K. Bernardin and R. Stiefelagen, “Evaluating multiple object tracking performance: The clear Mot Metrics,” *EURASIP Journal on Image and Video Processing*, vol. 2008, pp. 1–10, 2008. doi:10.1155/2008/246309.
- [37] J. Luiten *et al.*, “Hota: A higher order metric for evaluating multi-object tracking,” *International Journal of Computer Vision*, vol. 129, no. 2, pp. 548–578, 2020. doi:10.1007/s11263-020-01375-2.
- [38] Y. Liu *et al.*, “Opening up open world tracking,” *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. doi:10.1109/cvpr52688.2022.01846.
- [39] (PDF) “ΠΡΟΜΗΘΕΙΑ ΕΓΚΑΤΑΣΤΑΣΗ ΚΑΙ ΣΥΝΤΗΡΗΣΗ ΕΞΟΠΛΙΣΜΟΥ ΚΥΚΛΟΦΟΡΙΑΚΩΝ ΜΕΤΡΗΣΕΩΝ,” Εγνατία Οδός, https://egnatia.eu/wp-content/uploads/2019/11/5375_timologio.pdf (προσπελάστηκε Αυγ. 2023).
- [40] “DJI - official website,” DJI Official, <https://www.dji.com/> (προσπελάστηκε Αυγ. 2023).
- [41] A. Smith *et al.*, “Public acceptance of the use of drones for Logistics: The State of play and moving towards more informed debate,” *Technology in Society*, vol. 68, p. 101883, 2022. doi:10.1016/j.techsoc.2022.101883.

- [42] L. Ali *et al.*, “Performance evaluation of deep CNN-based crack detection and localization techniques for concrete structures,” *Sensors*, vol. 21, no. 5, p. 1688, 2021. doi:10.3390/s21051688.