



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΠΟΛΙΤΙΚΩΝ ΜΗΧΑΝΙΚΩΝ

ΤΟΜΕΑΣ ΜΕΤΑΦΟΡΩΝ ΚΑΙ ΣΥΓΚΟΙΝΩΝΙΑΚΗΣ ΥΠΟΔΟΜΗΣ

Ανίχνευση κυκλοφοριακών ουρών σε προσβάσεις σηματοδοτούμενου κόμβου με χρήση αλγόριθμων υπολογιστικής όρασης

Διπλωματική Εργασία



Παναγιώτης - Δημήτριος Χριστόπουλος

Επιβλέπουσα: Ελένη Ι. Βλαχογιάννη, Αναπληρώτρια Καθηγήτρια Ε.Μ.Π.

Αθήνα, Μάρτιος 2021

ΕΥΧΑΡΙΣΤΙΕΣ

Με την ολοκλήρωση της Διπλωματικής μου Εργασίας, θα ήθελα να ευχαριστήσω θερμά την κ. Ελένη Βλαχογιάννη, Αναπληρώτρια Καθηγήτρια της Σχολής Πολιτικών Μηχανικών Ε.Μ.Π., για την επιλογή ενός τόσο ιδιαίτερου θέματος και για την υποστήριξη και συνεργασία της σε όλα τα στάδια υλοποίησής του.

Θα ήθελα επίσης να ευχαριστήσω τον υποψήφιο Διδάκτωρ - Ερευνητή Παναγιώτη Φαφουτέλλη, για την διάθεση χρόνου και την πολύτιμη βοήθειά του για την καθοδήγηση πάνω σε κρίσιμα θέματα της Διπλωματικής Εργασίας.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου που με στήριξε σε κάθε βήμα των σπουδών μου και τους κοντινούς μου ανθρώπους που ήταν δίπλα μου σε αυτόν τον σημαντικό κύκλο της ζωής μου.

Αθήνα, Μάρτιος 2021

Παναγιώτης - Δημήτριος Χριστόπουλος

Ανίχνευση κυκλοφοριακών ουρών σε προσβάσεις σηματοδοτούμενου κόμβου με χρήση αλγόριθμων υπολογιστικής όρασης

Παναγιώτης - Δημήτριος Χριστόπουλος

Επιβλέπουσα: Ελένη Ι. Βλαχογιάννη, Αναπληρώτρια Καθηγήτρια Ε.Μ.Π.

Σύνοψη

Τα τελευταία χρόνια, η μετάβαση από τις κλασσικές μεθόδους στις σύγχρονες τεχνολογίες έχει επιτρέψει την ακριβέστερη ανάλυση κυκλοφοριακών φαινομένων και, ταυτόχρονα, την αντιμετώπιση κυκλοφοριακών προβλημάτων. Στόχο της παρούσας Διπλωματικής Εργασίας αποτελεί η δημιουργία ενός αλγόριθμου υπολογιστικής όρασης, ο οποίος θα πραγματοποιεί ανίχνευση κυκλοφοριακών ουρών σε προσβάσεις σηματοδοτούμενων κόμβων, και κατηγοριοποίησή τους ανάλογα με το είδος τους, με σκοπό την αντιμετώπιση κυκλοφοριακών συμφορήσεων. Τα δεδομένα για την εκπαίδευση του αλγόριθμου ελήφθησαν από πείραμα βιντεοσκόπησης και καταγραφής της κυκλοφορίας οχημάτων, με χρήση ΣμηΕΑ σε κεντρικούς οδικούς άξονες της Αθήνας. Για τη διαμόρφωση του τελικού αλγόριθμου, τα δεδομένα υπέστησαν επεξεργασία και επισημάνθηκαν βάσει εμφάνισης ή μη ουρών πέντε διαφορετικών κατηγοριών, και στη συνέχεια εκπαιδεύτηκε μοντέλο συνελκτικού νευρωνικού δικτύου βαθιάς μάθησης. Ο αλγόριθμος που αναπτύχθηκε έχει τη δυνατότητα να πραγματοποιεί αξιολογες ποσοστιαίες προβλέψεις κατηγοριοποίησης των σχηματιζόμενων ουρών, και δύναται να χρησιμοποιηθεί, με μικρή τροποποίηση, σε μεθοδολογίες μεταφοράς μάθησης. Η περαιτέρω αξιολόγηση της ποιότητας των αποτελεσμάτων, κατέδειξε την ανάγκη για συμπληρωματική εκπαίδευση του αλγόριθμου με τη χρήση δείγματος διαφορετικών δεδομένων μικτών συνθηκών, με σκοπό την ανίχνευση ουρών σε πραγματικούς χρόνους.

Λέξεις κλειδιά: ανίχνευση κατηγορίας, κυκλοφοριακή ουρά, ΣμηΕΑ, υπολογιστική όραση, βαθιά μάθηση, συνελκτικό νευρωνικό δίκτυο, μεταφορά μάθησης

Traffic queue detection at signalized junctions using computer vision algorithms

Panagiotis - Dimitrios Christopoulos

Supervisor: Eleni I. Vlahogianni, Associate Professor NTUA

Abstract

In recent years, the transition from traditional methods to modern technologies has provided a more accurate way of analyzing traffic conditions and addressing traffic issues. The objective of this Diploma Thesis is to propose a computer vision algorithm for the detection of traffic queues at signalized junctions and their classification, in order to deal with traffic congestions. The data for the training of the algorithm were sourced from a study that was conducted for the recording of the traffic in an extended part of the downtown Athens area, with the use of a swarm of UAVs. For the configuration of the algorithm, the data were processed and labelled based on the appearance or non-appearance of five different queue classes, and subsequently a deep convolutional neural network model was trained. The developed algorithm is able to make reliable percentage classification predictions of the formed queues, and can be used for transfer learning, with slight adjustments. The evaluation of the quality of the results indicate the need for supplementary training of the algorithm for the real-time queue detection, using a sample of different mixed conditions data.

Keywords: class detection, traffic queue, UAVs, computer vision, deep learning, convolutional neural network, transfer learning

ΠΕΡΙΛΗΨΗ

Η ραγδαία ανάπτυξη της τεχνολογίας των μη επανδρωμένων αεροσκαφών τα τελευταία χρόνια έχει οδηγήσει σε σημαντικές προόδους στη συλλογή και επεξεργασία κυκλοφοριακών δεδομένων. Οι τροχιές των οχημάτων που προκύπτουν από την καταγραφή της κυκλοφορίας μπορούν να χρησιμοποιηθούν για την μικροσκοπική ανάλυση διαφόρων κυκλοφοριακών φαινομένων και την εξαγωγή συμπερασμάτων σχετικά με την αντιμετώπιση κυκλοφοριακών προβλημάτων. Στόχο της παρούσας Διπλωματικής Εργασίας αποτελεί η δημιουργία ενός αλγόριθμου υπολογιστικής όρασης για την ανίχνευση και κατηγοριοποίηση κυκλοφοριακών ουρών, με σκοπό την αντιμετώπιση κυκλοφοριακών συμφορήσεων. Συγκεκριμένα ο αλγόριθμος έχει τη δυνατότητα να εντοπίζει το σχηματισμό ουρών από καταγεγραμμένες εικόνες βιντεοσκόπησης, και να πραγματοποιεί ποσοστιαία εκτίμηση των κατηγοριών οι οποίες αντιστοιχούν στις εικόνες.

Κατά τη βιβλιογραφική ανασκόπηση αναδείχθηκαν τα πλεονεκτήματα και οι περιορισμοί της επιλεχθείσας μεθοδολογίας χρήσης μη επανδρωμένων αεροσκαφών, καθώς και τα αποτελέσματα των ερευνών με σχετικό αντικείμενο, βάσει των οποίων ορίστηκε το τελικό θέμα της Διπλωματικής Εργασίας. Από την ανάλυση των συναφών ερευνών προέκυψε ότι τα πειράματα βιντεοσκόπησης με ΣμηΕΑ αποτελούν έναν αποδοτικό τρόπο συλλογής δεδομένων υψηλής ποιότητας, συγκριτικά με τη χρήση εγκατεστημένων καμερών παρακολούθησης. Επιπλέον, τα μοντέλα αλγόριθμων υπολογιστικής όρασης τα οποία βασίζονται σε πραγματικές βάσεις δεδομένων μεγάλου πλήθους οδηγούν σε ακριβέστερα αποτελέσματα. Ακόμη, η ανίχνευση των κυκλοφοριακών ουρών η οποία βασίζεται σε ανάλυση εικόνων, αντιμετωπίζεται ικανοποιητικά με μοντέλα συνελκτικών νευρωνικών δικτύων και την εφαρμογή βαθιάς μάθησης καθώς έχουν τη δυνατότητα εκπαίδευσης ιδιαίτερων περιγραφικών χαρακτηριστικών υψηλής υπολογιστικής ισχύος, λόγω της χρήσης ισχυρών γραφικών μονάδων επεξεργασίας (GPUs).

Η συλλογή των δεδομένων πραγματοποιήθηκε σε μία ευρύτερη περιοχή στο κέντρο της Αθήνας στο πλαίσιο μίας έρευνας συλλογής δεδομένων σε ένα εκτεταμένο δίκτυο προκειμένου να αναδειχθούν τα ιδιαίτερα χαρακτηριστικά των ΣμηΕΑ, καθώς και η δυνατότητα να ξεπεραστούν υφιστάμενοι περιορισμοί στην παρακολούθηση και καταγραφή κυκλοφοριακών ροών. Κατά τη διάρκεια του πειράματος βιντεοσκόπησης, σμήνος αεροσκαφών αποτελούμενο από 10 ΣμηΕΑ τοποθετήθηκε σε καθορισμένες θέσεις όπου και καταγράφηκε η κυκλοφορία. Από τα βίντεο χρησιμοποιήθηκαν τα συλλεγμένα δεδομένα από τις τρεις περιοχές όπου παρατηρήθηκαν οι περισσότεροι κυκλοφοριακοί φόρτοι σε σηματοδοτούμενους κόμβους, τα οποία στη συνέχεια επισημάνθηκαν με το είδος ουράς που παρατηρήθηκε. Έπειτα από μία μικρή επεξεργασία, τα δεδομένα συνολικού αριθμού 38668 καρέ εικόνων και επισημάνσεων αντίστοιχα, εισήχθησαν στον αλγόριθμο από τα οποία το 25% χρησιμοποιήθηκε για την αξιολόγηση των μοντέλων και τα υπόλοιπα για την εκπαίδευση αυτών, για κάθε περιοχή καταγραφής ξεχωριστά.

Από το εκπαιδευόμενο μοντέλο του αλγόριθμου, μέσω του συνελκτικού νευρωνικού δικτύου εξήχθησαν οι πληροφορίες από τα μητρώα των δεδομένων οι οποίες περιέχουν τα εικονοκύτταρα των εικόνων, όπου σε συνδυασμό με τα επισημασμένα δεδομένα και την εισαγωγή και εξαγωγή των δεδομένων από διάφορα στρώματα εκπαίδευσης, προέκυψε ο ταξινομητής κλάσεων κυκλοφοριακών ουρών. Η ποιότητα του ταξινομητή αξιολογήθηκε με τον υπολογισμό των τυπικών μετρικών ακρίβειας, των καμπυλών ROC και των τιμών περιοχών AUC. Όλα τα μοντέλα των διαφορετικών περιοχών καταγραφής παρουσίασαν πολύ ικανοποιητικά αποτελέσματα υψηλής ακρίβειας και αξιόλογες προβλέψεις, ενώ το μοντέλο ενιαίας βάσης λόγω μεγάλου όγκου πολύπλοκων δεδομένων δεν παρουσίασε αναμενόμενα αποτελέσματα. Στον τελικό αλγόριθμο εφαρμόστηκε μεταφορά μάθησης της εκπαιδευμένης γνώσης στα δεδομένα εικόνων των άλλων περιοχών παρόμοιων και διαφορετικών χαρακτηριστικών.

Από τα διάφορα στάδια εκπόνησης της παρούσας Διπλωματικής Εργασίας προέκυψαν αποτελέσματα άμεσα συνδεδεμένα με τον κύριο στόχο που είχε οριστεί. Παρακάτω παρουσιάζονται τα σημαντικότερα συμπεράσματα τα οποία προέκυψαν.

Η συγκεκριμένη Διπλωματική Εργασία αποτελεί μία από τις πρώτες έρευνες που **πραγματοποιεί προβλέψεις κατηγοριοποίησης κυκλοφοριακών ουρών** για την αντιμετώπιση και αποφυγή κυκλοφοριακών συμφορήσεων, με χρήση υπολογιστικής όρασης και καταγεγραμμένων δεδομένων από ΣμηΕΑ. Οι περισσότερες έρευνες πάνω στο αντικείμενο αντιμετώπισης συμφορήσεων εστιάζουν περισσότερο στην εξαγωγή και ανάλυση των κυκλοφοριακών μεταβλητών των οχημάτων με χρήση παρόμοιας μεθοδολογίας και βιντεοσκοπημένων δεδομένων. Όμως, για την αποτελεσματική αντιμετώπιση του προβλήματος και τη δημιουργία ενός αυτόνομου συστήματος παρακολούθησης της κυκλοφορίας, απαιτείται η μελέτη ποικίλων κυκλοφοριακών συνθηκών σε βάθος.

Τα συνελκτικά νευρωνικά δίκτυα βαθιάς μάθησης αποτελούν **κατάλληλη μέθοδο για τη διαμόρφωση των αλγόριθμων ταξινόμησης ουρών** καθώς είναι σε θέση να διαχειριστούν πολύπλοκα επισημασμένα δεδομένα μεγάλου πλήθους. Η αναγνώριση τέτοιων κυκλοφοριακών συνθηκών σε πραγματικούς χρόνους αποτελεί τεράστιο πλεονέκτημα για κυκλοφοριακές αναλύσεις, η οποία μπορεί να επιτευχθεί με το συνδυασμό νευρωνικών δικτύων CNN και ισχυρής υπολογιστική ισχύος.

Ο **αλγόριθμος είναι εξαιρετικής ακρίβειας για δεδομένα ίδιων χαρακτηριστικών**, όπως προκύπτει από τις επαναλήψεις των μοντέλων και τις τελικές τιμές των δεδομένων εκπαίδευσης και επικύρωσης των διαγραμμάτων ακρίβειας και απώλειας οι οποίες συγκλίνουν. Αντιθέτως, για πολλαπλά δεδομένα διαφορετικών χαρακτηριστικών, ο αλγόριθμος δεν είναι καθόλου ικανοποιητικός καθώς το μοντέλο υπερφορτώνεται με πληροφορίες και διακόπτει τη διαδικασία εκπαίδευσης.

Ο **πολύ ικανοποιητικός υπολογιστικός χρόνος** των μοντέλων είναι ένα από τα βασικά πλεονεκτήματα του αλγόριθμου καθώς επιτυγχάνεται μεγάλη ακρίβεια προβλέψεων με **χρήση βάσεων δεδομένων μεγάλου πλήθους**. Με τη χρήση

μεγάλων βάσεων, ο αλγόριθμος δεν περιορίζεται αποκλειστικά στην αναγνώριση της ύπαρξης ή μη κυκλοφοριακής ουράς, αλλά στην ανίχνευση του είδους κάθε ουράς. Βάσει αυτής της πολυπλοκότητας ο αλγόριθμος απαιτεί περισσότερο υπολογιστικό χρόνο, χωρίς αυτός να υπερβαίνει την αναμενόμενη διάρκεια εκπαίδευσης του μοντέλου.

Η **υψηλή ποιότητα των προβλέψεων** του αλγόριθμου παραμένει ακόμη και στις περιπτώσεις χρήσης μικρού δείγματος δοκιμών. Ο αλγόριθμος είναι ικανός να αναγνωρίσει όλες τις κατηγορίες ουρών που έχουν οριστεί και να προσφέρει έως και **μηδενικές λανθασμένες προβλέψεις**, για κατηγορίες ουρών μεγάλου ποσοστού εμφάνισης στα δεδομένα δοκιμών. Ειδικότερα, οι ποσοστιαίες εκτιμήσεις που πραγματοποιεί ο αλγόριθμος είναι αξιόπιστες καθώς ακόμη και στις λανθασμένες ταξινομήσεις ουρών, παρουσιάζει ελάχιστες αποκλίσεις από την πραγματικότητα.

Η χρήση των γνώσεων του μοντέλου του αλγόριθμου σε έναν αλγόριθμο διαφορετικής βάσης δεδομένων με **μεταφορά μάθησης παρουσιάζει μη ικανοποιητικά, αλλά ενθαρρυντικά αποτελέσματα**. Η ποιότητα των προβλέψεων είναι αναμενόμενη καθώς το επαναχρησιμοποιούμενο μοντέλο έχει εκπαιδευτεί με διαφορετικά καρέ εικόνων, όμως ο αλγόριθμος μπορεί και ανιχνεύει την πλειοψηφία των κοινών κατηγοριών ουρών που εμφανίζονται και στα δύο βίντεο. Αντιθέτως, για διαφορετικά καρέ εικόνων και ποσοστών εμφάνισης συγκεκριμένων κατηγοριών, ο αλγόριθμος είναι ικανός να αναγνωρίσει κυρίως την ύπαρξη ή μη κυκλοφοριακής ουράς.

ΠΕΡΙΕΧΟΜΕΝΑ

1. ΕΙΣΑΓΩΓΗ	1
1.1 ΓΕΝΙΚΗ ΑΝΑΣΚΟΠΗΣΗ	1
1.2 ΣΚΟΠΟΣ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑΣ	4
1.3 ΔΙΑΡΘΡΩΣΗ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ	6
2. ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ	7
2.1 ΕΙΣΑΓΩΓΗ	7
2.2 ΧΡΗΣΗ ΣΜΗΜΑΤΑ ΣΕ ΚΥΚΛΟΦΟΡΙΑΚΕΣ ΑΝΑΛΥΣΕΙΣ	7
2.3 ΕΦΑΡΜΟΓΕΣ ΒΑΘΙΑΣ ΜΑΘΗΣΗΣ ΣΤΗΝ ΑΝΑΛΥΣΗ ΚΥΚΛΟΦΟΡΙΑΣ	10
2.4 ΣΥΝΟΨΗ ΒΙΒΛΙΟΓΡΑΦΙΑΣ	14
3. ΜΕΘΟΔΟΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ	15
3.1 ΕΙΣΑΓΩΓΗ	15
3.2 ΠΡΟΤΕΙΝΟΜΕΝΗ ΠΡΟΣΕΓΓΙΣΗ	15
3.3 ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ	17
3.3.1 Βαθιά Μάθηση	17
3.3.2 Εποπτευόμενη Μάθηση	18
3.3.3 Συνελκτικά Νευρωνικά Δίκτυα	21
3.3.4 Μεταφορά Μάθησης	23
4. ΑΝΑΛΥΣΗ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ	25
4.1 ΕΙΣΑΓΩΓΗ	25
4.2 ΣΥΛΛΟΓΗ ΔΕΔΟΜΕΝΩΝ	25
4.3 ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ	27
4.4 ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ	31
4.5 ΑΝΑΛΥΣΗ	34
4.6 ΑΠΟΤΕΛΕΣΜΑΤΑ	41
4.6.1 Οδός Πατησίων - Λεωφόρος Αλεξάνδρας (Video 1)	41
4.6.2 Λεωφόρος Αλεξάνδρας - Οδός Πατησίων (Video 2)	45
4.6.3 Πλατεία Ομονοίας (Video 3)	50
4.6.4 Συνδυασμός Βάσεων Δεδομένων (Video ALL)	55
5. ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΡΟΤΑΣΕΙΣ	59
5.1 ΕΙΣΑΓΩΓΗ	59
5.2 ΒΑΣΙΚΑ ΣΥΜΠΕΡΑΣΜΑΤΑ	59
5.3 ΠΡΟΤΑΣΕΙΣ ΓΙΑ ΠΕΡΑΙΤΕΡΩ ΕΡΕΥΝΑ	60
ΒΙΒΛΙΟΓΡΑΦΙΑ	62

ΕΥΡΕΤΗΡΙΟ ΔΙΑΓΡΑΜΜΑΤΩΝ

Διάγραμμα 1.1: Ποσοστό παραγόντων δημιουργίας κυκλοφοριακών συμφορήσεων ανάλογα με το είδος (Πηγή: U.S. Department of Transportation)	2
Διάγραμμα 1.2: Αναπαράσταση συστήματος ουράς (Πηγή: Rajashekara, et al., 2014).....	3
Διάγραμμα 1.3: Θεωρία ουρών στις μεταφορές - Καμπύλη Newell (Πηγή: Wang 2013).....	4
Διάγραμμα 1.4: Διάγραμμα ροής των σταδίων εκπόνησης της Διπλωματικής Εργασίας	5
Διάγραμμα 2.1: Διάγραμμα ροής συστήματος ανίχνευσης και παρακολούθησης οχημάτων με ΣμηΕΑ (Πηγή: Wang, et al., 2016).....	9
Διάγραμμα 2.2: Διάγραμμα ροής διαδικασιών προτεινόμενου μεθοδολογικού πλαισίου εντοπισμού κυκλοφοριακών συνθηκών (Πηγή: Vlahogianni, et al., 2021).....	13
Διάγραμμα 3.1: Διάγραμμα ροής ανάπτυξης ενός μοντέλου νευρωνικού δικτύου βαθιάς μάθησης με χρήση καταγεγραμμένων δεδομένων από ΣμηΕΑ	15
Διάγραμμα 3.2: Αλγόριθμοι παλινδρόμησης και ταξινόμησης	19
Διάγραμμα 3.3: Παραδοσιακή μηχανική μάθηση και μεταφορά μάθησης	23
Διάγραμμα 4.1: Ακρίβεια μοντέλου για δεδομένα εκπαίδευσης και επικύρωσης V1	42
Διάγραμμα 4.2: Απώλεια μοντέλου για δεδομένα εκπαίδευσης και επικύρωσης V1	42
Διάγραμμα 4.3: Καμπύλες ROC και τιμές AUC για τις κλάσεις του μοντέλου V1	44
Διάγραμμα 4.4: Ακρίβεια μοντέλου για δεδομένα εκπαίδευσης και επικύρωσης V2	46
Διάγραμμα 4.5: Απώλεια μοντέλου για δεδομένα εκπαίδευσης και επικύρωσης V2	46
Διάγραμμα 4.6: Καμπύλες ROC και τιμές AUC για τις κλάσεις του μοντέλου V2	48
Διάγραμμα 4.7: Καμπύλες ROC και τιμές AUC για τις κλάσεις του μοντέλου V2 (μεγέθυνση διαγράμματος).....	48
Διάγραμμα 4.8: Ακρίβεια μοντέλου για δεδομένα εκπαίδευσης και επικύρωσης V3.1	51
Διάγραμμα 4.9: Απώλεια μοντέλου για δεδομένα εκπαίδευσης και επικύρωσης V3.1	51
Διάγραμμα 4.10: Καμπύλες ROC και τιμές AUC για τις κλάσεις του μοντέλου V3.1	53
Διάγραμμα 4.11: Καμπύλες ROC και τιμές AUC για τις κλάσεις του μοντέλου V3.1 (μεγέθυνση διαγράμματος).....	53
Διάγραμμα 4.12: Ακρίβεια μοντέλου για δεδομένα εκπαίδευσης και επικύρωσης VALL	56
Διάγραμμα 4.13: Απώλεια μοντέλου για δεδομένα εκπαίδευσης και επικύρωσης VALL.....	57
Διάγραμμα 4.14: Καμπύλες ROC και τιμές AUC για τις κλάσεις του μοντέλου VALL.....	58

ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ

Πίνακας 1.1: Εξέλιξη χαμένων ωρών σε κυκλοφοριακή συμφόρηση στα κράτη της Ε.Ε. κατά τα έτη 2015 - 2017 (Πηγή: EU Transport Scoreboard)	2
Πίνακας 1.2: Κατάταξη ευρωπαϊκών πόλεων με βάση τα επίπεδα κυκλοφοριακής συμφόρησης (Πηγή: TomTom Traffic Index).....	3
Πίνακας 4.1: Κλάσεις κυκλοφοριακών ουρών και αντίστοιχο πλήθος	29
Πίνακας 4.2: Κλάσεις κυκλοφοριακών ουρών μοντέλων.....	31
Πίνακας 4.3: Μητρώο καρέ εικόνων βάσης δεδομένων (μεταβλητή x).....	33
Πίνακας 4.4: Μητρώο επισημάνσεων βάσης δεδομένων (μεταβλητή y).....	33
Πίνακας 4.5: Κωδικοποιημένο μητρώο επισημάνσεων για δοκιμές (y_test_one_hot)	35
Πίνακας 4.6: Ομαλοποιημένο μητρώο εικόνων για εκπαίδευση (x_train)	36
Πίνακας 4.7: Μητρώο αριθμών για αληθινές και ψευδείς ταξινομήσεις	39

ΕΥΡΕΤΗΡΙΟ ΕΙΚΟΝΩΝ

Εικόνα 3.1: Δομή επιστήμης υπολογιστών και τεχνητής νοημοσύνης και κατηγορίες αλγόριθμων βαθιάς μάθησης (Πηγή: Mysla, 2020)	18
Εικόνα 3.2: Αρχιτεκτονική συνελκτικού νευρωνικού δικτυού (Πηγή: Thomas, 2017)	23
Εικόνα 4.1: Σημεία τοποθέτησης ΣμηΕΑ (Πηγή: Barmprounakis, et al., 2019).....	26
Εικόνα 4.2: Οδός Πατησίων - Λεωφόρος Αλεξάνδρας (σημείο 7) - Video 1	27
Εικόνα 4.3: Λεωφόρος Αλεξάνδρας - Οδός Πατησίων (σημείο 8) - Video 2	27
Εικόνα 4.4: Πλατεία Ομονοίας (σημείο 10) - Video 3	28
Εικόνα 4.5: Περιβάλλον πλατφόρμας Supervisely	29
Εικόνα 4.6: Εργαλείο επισήμανσης ετικετών, λίστα καρέ εικόνων βίντεο, διάρκεια βίντεο και αριθμός εικόνων	30
Εικόνα 4.7: (αριστερά) Λίστα κλάσεων και αντίστοιχος αριθμός εικόνων, (δεξιά) Στοιχεία βίντεο και συνολικός αριθμός κλάσεων και επισημασμένων εικόνων.....	30
Εικόνα 4.8: (πάνω) Είδος μητρώων δειγμάτων, (κάτω) Μέγεθος μητρώων δειγμάτων	35
Εικόνα 4.9: Ονομασίες κλάσης καρέ εικόνας 2025 δείγματος εκπαίδευσης	36
Εικόνα 4.10: Πληροφορίες επαναλήψεων και ακρίβεια 94.79% μοντέλου V1	41
Εικόνα 4.11: (αριστερά) Μητρώα σύγχυσης μοντέλου V1, (δεξιά) Αναφορά ταξινόμησης μοντέλου V1	43
Εικόνα 4.12: (αριστερά) Μητρώα σύγχυσης δεδομένων Video 2 - μοντέλου V1, (δεξιά) Αναφορά ταξινόμησης δεδομένων Video 2 - μοντέλου V1	44
Εικόνα 4.13: (αριστερά) Μητρώα σύγχυσης δεδομένων Video 3.1 - μοντέλου V1, (δεξιά) Αναφορά ταξινόμησης δεδομένων Video 3.1 - μοντέλου V1	45
Εικόνα 4.14: Πληροφορίες επαναλήψεων και ακρίβεια 97.9% μοντέλου V2	45
Εικόνα 4.15: (αριστερά) Μητρώα σύγχυσης μοντέλου V2, (δεξιά) Αναφορά ταξινόμησης μοντέλου V2	47
Εικόνα 4.16: (αριστερά) Μητρώα σύγχυσης δεδομένων Video 1 - μοντέλου V2, (δεξιά) Αναφορά ταξινόμησης δεδομένων Video 1 - μοντέλου V2	49
Εικόνα 4.17: (αριστερά) Μητρώα σύγχυσης δεδομένων Video 3.1 - μοντέλου V2, (δεξιά) Αναφορά ταξινόμησης δεδομένων Video 3.1 - μοντέλου V2	49
Εικόνα 4.18: Πληροφορίες επαναλήψεων και ακρίβεια 96.69% μοντέλου V3.1	50
Εικόνα 4.19: Ακρίβεια 83.41% μοντέλου V3.2 και 85.25% μοντέλου V3.3	50
Εικόνα 4.20: (αριστερά) Μητρώα σύγχυσης μοντέλου V3.1, (δεξιά) Αναφορά ταξινόμησης μοντέλου V3.1	52
Εικόνα 4.21: (αριστερά) Μητρώα σύγχυσης δεδομένων Video 1 - μοντέλου V3.1, (δεξιά) Αναφορά ταξινόμησης δεδομένων Video 1 - μοντέλου V3.1	54
Εικόνα 4.22: (αριστερά) Μητρώα σύγχυσης δεδομένων Video 2 - μοντέλου V3.1, (δεξιά) Αναφορά ταξινόμησης δεδομένων Video 3.1 - μοντέλου V3.1	54
Εικόνα 4.23: (αριστερά) Μητρώα σύγχυσης δεδομένων Video 3.2 - μοντέλου V3.1, (δεξιά) Αναφορά ταξινόμησης δεδομένων Video 3.1 - μοντέλου V3.1	55
Εικόνα 4.24: (αριστερά) Μητρώα σύγχυσης δεδομένων Video 3.3 - μοντέλου V3.1, (δεξιά) Αναφορά ταξινόμησης δεδομένων Video 3.1 - μοντέλου V3.1	55
Εικόνα 4.25: Πληροφορίες επαναλήψεων και ακρίβεια 48,04% μοντέλου VALL	56
Εικόνα 4.26: (αριστερά) Μητρώα σύγχυσης μοντέλου VALL, (δεξιά) Αναφορά ταξινόμησης μοντέλου VALL.....	57

1. ΕΙΣΑΓΩΓΗ

1.1 ΓΕΝΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

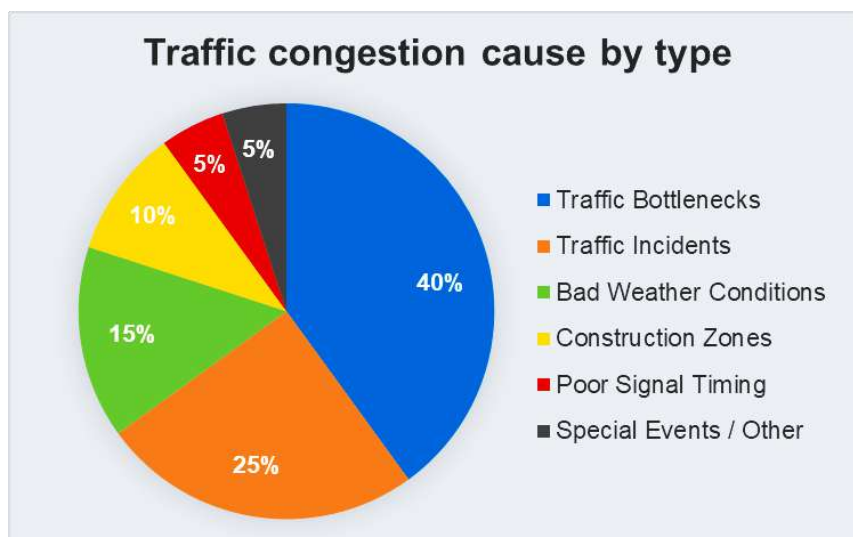
Η **αντιμετώπιση κυκλοφοριακών προβλημάτων** (συστηματικών ή περιστασιακών) αποτελεί ένα από τα σημαντικότερα αντικείμενα έρευνας ενός συγκοινωνιολόγου μηχανικού στη σημερινή εποχή. Η αποτελεσματική **αναγνώριση και ανάλυση των κυκλοφοριακών συνθηκών** για την βελτίωσή τους, αποτελεί τον κύριο στόχο των μηχανικών. Τα τελευταία χρόνια η εξέλιξη της τεχνολογίας έχει οδηγήσει στην εισαγωγή νέων μεθόδων με τη χρήση ηλεκτρονικών υπολογιστών, οι οποίες πραγματοποιούν ταχύτερα και ακριβέστερα αποτελέσματα, μειώνοντας ταυτόχρονα τη συμβολή του ανθρώπινου παράγοντα.

Σημαντική διάσταση των προβλημάτων αποτελούν οι **κυκλοφοριακές συμφορήσεις** που δημιουργούνται. Εμφανίζονται κυρίως σε πυκνοκατοικημένες περιοχές και σε οδικούς άξονες με υψηλό κυκλοφοριακό φόρτο, με αποτέλεσμα τις χρονικές καθυστερήσεις των χρηστών του συστήματος. Τα αίτια αυτών είναι ποικίλα και διαφέρουν ανάλογα την περιοχή όπως:

- ♦ Ανεπάρκεια χωρητικότητας συστήματος λόγω κακών οδικών υποδομών, ασταθών οδηγικών συμπεριφορών, μη ελκυστικών υπηρεσιών δημόσιων συγκοινωνιών και προβληματικής στάθμευσης
- ♦ Ατυχήματα
- ♦ Δυσμενείς καιρικές συνθήκες
- ♦ Οδικά έργα
- ♦ Μη συγχρονισμένοι χρόνοι σηματοδότησεων
- ♦ Άλλοι τυχαίοι παράγοντες

Σύμφωνα με πρόσφατα στοιχεία ερευνών της εταιρείας λογισμικού τεχνολογίας τοποθεσίας TomTom, υπολογίστηκαν μέσω συσκευών δορυφορικής πλοήγησης (GPS) για τα έτη 2015 - 2017, οι συνολικές ετήσιες οδικές ώρες που χάθηκαν λόγω κυκλοφοριακών συμφορήσεων από τον μέσο οδηγό σε ευρωπαϊκές χώρες (EU Transport Scoreboard, 2017). Σημειώνεται πως η Ελλάδα βρίσκεται στις πρώτες πέντε θέσεις για τα παραπάνω έτη με μέσο όρο περίπου 37 χαμένων ωρών, όμως με σταδιακή μείωση ανά έτος. Αντιθέτως, οι χώρες της Σκανδιναβίας και της Βαλτικής εμφανίζουν σταθερά το χαμηλότερο μέσο όρο για τις αντίστοιχες χρονιές (περίπου 20 ώρες). Επιπλέον, καταγράφηκαν οι ποσοστιαίες κατανομές των επιπέδων συμφόρησης 239 ευρωπαϊκών πόλεων για το έτος 2019, καθώς και η ποσοστιαία μεταβολή τους συγκριτικά με το προηγούμενο έτος (TomTom Traffic Index, 2019). Η Αθήνα και η Θεσσαλονίκη κατατάσσονται στην 14^η και 74^η θέση, με ποσοστά 43% και 30% αντίστοιχα, ενώ και οι δύο πόλεις παρουσίασαν αύξηση στις τιμές τους. Οι παραπάνω πόλεις συγκρίνονται και σε παγκόσμιο επίπεδο. Επομένως, παρατηρείται και στατιστικά η ανάγκη για αντιμετώπιση των συμφορήσεων.

Στο Διάγραμμα 1.1 παρατίθεται η ποσοστιαία κατανομή της κυκλοφοριακής συμφόρησης ανάλογα με τα αίτια δημιουργίας της. Στους Πίνακες 1.1 και 1.2 παρατίθενται οι κατανομές χαμένων ετήσιων οδικών ωρών ανά χώρα της Ευρωπαϊκής Ένωσης για την τριετία 2015 - 2017 και επιπέδων συμφόρησης ευρωπαϊκών πόλεων.



Διάγραμμα 1.1: Ποσοστό παραγόντων δημιουργίας κυκλοφοριακών συμφορήσεων ανάλογα με το είδος (Πηγή: U.S. Department of Transportation)

Πίνακας 1.1: Εξέλιξη χαμένων ωρών σε κυκλοφοριακή συμφόρηση στα κράτη της Ε.Ε. κατά τα έτη 2015 - 2017 (Πηγή: EU Transport Scoreboard)

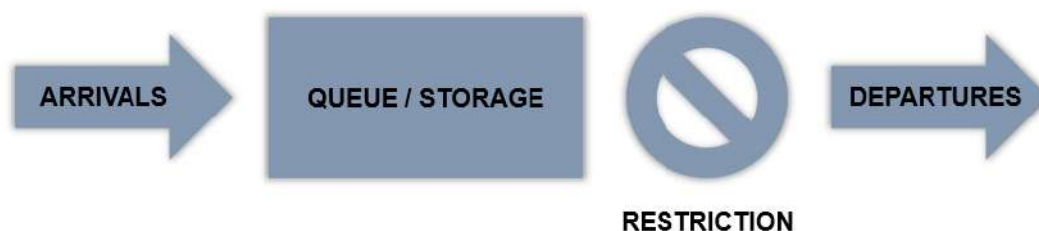
2015			2016		
#	Country	Hours	#	Country	Hours
1	United Kingdom	41.45	1	United Kingdom	45.62
2	Greece	38.85	2	Belgium	39.04
3	Belgium	36.11	3	Italy	37.50
4	Italy	35.43	4	Greece	36.46
5	Luxembourg	32.21	5	Luxembourg	35.68
:	:	:	:	:	:
22	Latvia	21.61	22	Sweden	21.26
23	Sweden	21.41	23	Latvia	20.56
24	Lithuania	21.37	24	Lithuania	20.21
25	Estonia	20.53	25	Estonia	19.21
26	Finland	19.94	26	Finland	18.19
2017					
#	Country	Hours			
1	United Kingdom	45.73			
2	Belgium	39.37			
3	Italy	37.36			
4	Luxembourg	36.88			
5	Greece	36.28			
:	:	:			
23	Latvia	21.81			
24	Sweden	21.56			
25	Lithuania	20.99			
26	Estonia	19.39			
27	Finland	18.13			



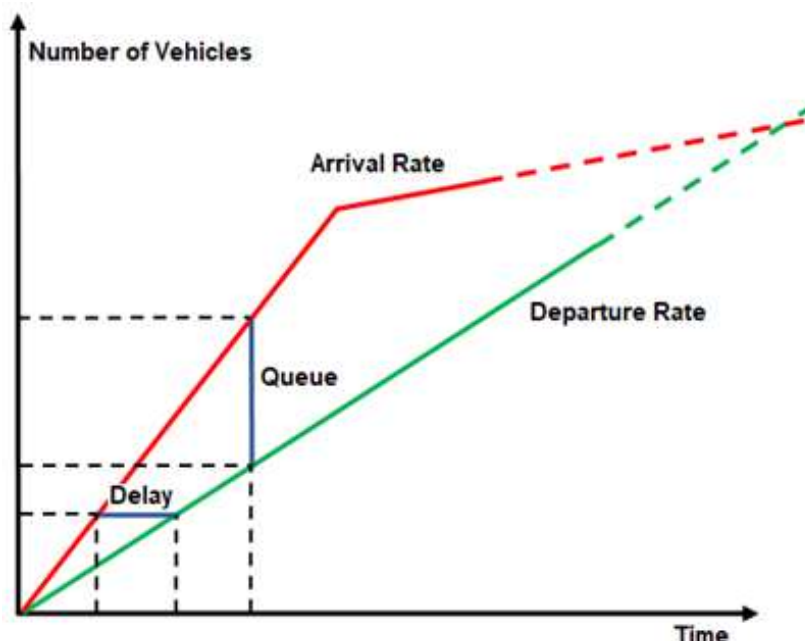
Πίνακας 1.2: Κατάταξη ευρωπαϊκών πόλεων με βάση τα επίπεδα κυκλοφοριακής συμφόρησης (Πηγή: TomTom Traffic Index)

Europe	World	City	Country	Congestion
1	6	Moscow	Russia	59% ↑3%
2	9	Instabul	Turkey	55% ↑2%
3	12	Kyiv	Ukraine	53% ↑7%
4	14	Bucharest	Romania	52% ↑4%
5	16	Saint Petersburg	Russia	49% ↑2%
6	17	Dublin	Ireland	48% ↑3%
7	18	Odessa	Ukraine	47% 0%
8	19	Lodz	Poland	47% ↑3%
9	22	Krakow	Poland	45% ↑5%
10	23	Novosibirsk	Russia	45% ↑1%
11	25	Samara	Russia	44% 0%
12	27	Poznan	Poland	44% ↑5%
13	29	Kharkiv	Ukraine	43% 0%
14	30	Athens	Greece	43% ↑3%
15	33	Edinburgh	United Kingdom	41% ↑1%
16	35	Yekaterinburg	Russia	41% ↑4%
17	37	Warsaw	Poland	40% ↑1%
18	41	Wroclaw	Poland	39% ↑4%
19	42	Paris	France	39% ↑3%
20	43	Rome	Italy	38% ↓1%
:	:	:	:	:
74	130	Thessaloniki	Greece	30% ↑2%
:	:	:	:	:
237	401	Bilbao	Spain	13% 0%
238	411	Almere	Netherlands	10% ↓2%
239	415	Cadiz	Spain	10% ↑1%

Ένας συγκεκριμένος παράγοντας που απαιτεί ανάλυση είναι ο **σχηματισμός ουρών** λόγω συμφόρησης (Διάγραμμα 1.2), ένα καθημερινό κυκλοφοριακό φαινόμενο που εμφανίζεται σε οδικούς άξονες παγκοσμίως. Όταν σε ένα οδικό τμήμα η ζήτηση ξεπερνά την ικανότητα για κάποια χρονική περίοδο ή ο ρυθμός εξυπηρέτησης του τμήματος υστερεί του ρυθμού άφιξης των οχημάτων, παρατηρείται σχηματισμός ουράς (Φραντζεσκάκης, et al., 2009). Η αναποτελεσματική αντιμετώπιση μιας ουράς μπορεί να οδηγήσει σε επέκταση της συμφόρησης του δικτύου ακόμη και σε μεγαλύτερη έκταση από το τμήμα στο οποίο εμφανίζεται, με αποτέλεσμα τη μείωση της κυκλοφοριακής ικανότητας και την αύξηση των χρονικών καθυστερήσεων (Διάγραμμα 1.3).



Διάγραμμα 1.2: Αναπαράσταση συστήματος ουράς (Πηγή: Rajashekhara, et al., 2014)



Διάγραμμα 1.3: Θεωρία ουρών στις μεταφορές - Καμπύλη Newell (Πηγή: Wang 2013)

Η ανάλυση τέτοιων φαινομένων δεν είναι αποτελεσματική με τις κλασσικές μεθόδους (μετρήσεις πεδίου και σταθερές κάμερες) και για το λόγο αυτό χρησιμοποιούμε σύγχρονες τεχνολογίες. Συγκεκριμένα, αξιοποιούνται **Συστήματα μη Επανδρωμένων Αεροσκαφών (ΣμηΕΑ)** για τη συλλογή δεδομένων μέσω **ψηφιακών εικόνων** και **βίντεο**, καθώς ξεπερνούν περιορισμούς τυφλών και κρυμμένων σημείων, καθώς και μικρών εκτάσεων καταγεγραμμένων περιοχών (Barmounakis, et al., 2016). Επιπλέον, χρησιμοποιούνται μέθοδοι προγραμματισμού και αλγόριθμων υπολογιστών όπως η **υπολογιστική όραση (computer vision)** και διάφοροι κλάδοι της, για τη δημιουργία μοντέλων, στατιστικών πληροφοριών και αποτελεσμάτων σε επίπεδα αυτοματισμού (Janga, et al., 2020).

1.2 ΣΚΟΠΟΣ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑΣ

Σκοπός της παρούσας Διπλωματικής Εργασίας είναι η **δημιουργία ενός αλγόριθμου υπολογιστικής όρασης με στόχο την ανίχνευση και κατηγοριοποίηση κυκλοφοριακών ουρών**. Συγκεκριμένα, εξετάζεται η δυνατότητα του αλγόριθμου να αντιλαμβάνεται αυτόματα κυκλοφοριακές ουρές που σχηματίζονται και να αντλεί πληροφορίες για τα χαρακτηριστικά τους. Ο αλγόριθμος αυτός θα πρέπει να έχει τη δυνατότητα να εντοπίζει τη σχηματιζόμενη ουρά, καθώς και να εκτιμά ποσοστιαία το είδος της κάθε ουράς για κάποια τυχαία καταγεγραμμένη περίπτωση.

Ο αλγόριθμος προορίζεται για τέτοια χρήση ώστε να μπορεί να αξιοποιηθεί από ΣμηΕΑ (drones) και να είναι δυνατή η ανίχνευση των ουρών ακόμη και σε πραγματικό χρόνο καταγραφής της κυκλοφορίας. Για το σκοπό αυτό, τα δεδομένα για την εκπαίδευση του μοντέλου λήφθηκαν από πραγματικές κυκλοφοριακές συνθήκες από πείραμα βιντεοσκόπησης. Τα ΣμηΕΑ του πειράματος πετούσαν σε σταθερό σημείο πάνω από

κεντρικούς οδικούς άξονες τριών περιοχών της Αθήνας, όπου και παρατηρείται η πλειονότητα των κυκλοφοριακών ουρών με τις σοβαρότερες επιπτώσεις.

Για τους αλγόριθμους υπολογιστικής όρασης εφαρμόστηκε η μέθοδος **βαθιάς μάθησης (deep learning)** και συγκεκριμένα η χρήση εποπτευόμενης μάθησης (supervised learning) για την αντιμετώπιση προβλημάτων ταξινόμησης, καθώς και για την πραγματοποίηση δομημένων προβλέψεων με την κατασκευή νευρωνικών δικτύων (neural networks). Επιπλέον, εξετάστηκε η εφαρμογή μεταφοράς μάθησης (transfer learning), για την κατηγοριοποίηση εικόνων διαφορετικής βάσης δεδομένων με ήδη εκπαιδευμένο μοντέλο.

Το νευρωνικό δίκτυο που χρησιμοποιήθηκε είναι το συνελκτικό **CNN (Convolutional Neural Network)**, το οποίο χρησιμοποιεί μια πραγματική βάση δεδομένων εικόνων. Η βάση αυτή περιέχει όλους τους συνδυασμούς ψηφιακών εικόνων και κατηγοριοποιήσεων των κυκλοφοριακών ουρών και βάσει αυτών το μοντέλο επιλέγει την κατάλληλη ταξινόμηση νέων δεδομένων εικόνων. Η τελική ταξινόμηση γίνεται βάσει των ποσοστών του κάθε είδους ουράς στο οποίο θεωρεί ότι ανταποκρίνεται περισσότερο μια εικόνα.

Στο Διάγραμμα 1.4 παρουσιάζονται υπό μορφή διαγράμματος ροής, τα διαδοχικά στάδια που ακολουθήθηκαν κατά την εκπόνηση της Διπλωματικής Εργασίας.



Διάγραμμα 1.4: Διάγραμμα ροής των σταδίων εκπόνησης της Διπλωματικής Εργασίας

1.3 ΔΙΑΡΘΡΩΣΗ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

Στο παρόν υποκεφάλαιο, παρουσιάζεται η διαδικασία υλοποίησης της Διπλωματικής Εργασίας μέσω συνοπτικής αναφοράς του περιεχομένου του εκάστοτε κεφαλαίου της.

Το **κεφάλαιο 1**, αποτελεί την **εισαγωγή** της Διπλωματικής Εργασίας και παρουσιάζει την ανάγκη για αντιμετώπιση των κυκλοφοριακών συμφορήσεων και συγκεκριμένα του φαινομένου των κυκλοφοριακών ουρών. Επιπλέον, παρουσιάζει την εξέλιξη της τεχνολογίας και ειδικά των ΣμηΕΑ, καθώς και τη συνεχώς αυξανόμενη συμβολή τους στην ανάλυση διαφόρων κυκλοφοριακών χαρακτηριστικών και παραγόντων. Έπειτα, περιγράφεται το αντικείμενο και ο στόχος της Διπλωματικής Εργασίας και τίθενται τα ερωτήματα προς διερεύνηση.

Στο **κεφάλαιο 2**, παρουσιάζονται τα αποτελέσματα της **βιβλιογραφικής ανασκόπησης** όπως αυτά προέκυψαν από την αναζήτηση ερευνών με παρεμφερές αντικείμενο. Σημειώνεται ότι οι παρατιθέμενες εργασίες αφορούν έρευνες σε ελληνικό αλλά και διεθνές επίπεδο, ενώ όλες έχουν δημοσιευθεί σε επιστημονικά περιοδικά, άρθρα ή συγγράμματα. Κατόπιν, πραγματοποιείται σύνοψη των μεθοδολογιών που χρησιμοποιήθηκαν και παρουσιάζονται συγκεντρωτικά όλα τα αποτελέσματα, ώστε να διαπιστωθεί η ποιότητα και ικανότητά τους να συμβάλλουν στην παρούσα Διπλωματική Εργασία.

Στο **κεφάλαιο 3**, το οποίο αποτελεί τη **μεθοδολογική προσέγγιση**, παρουσιάζεται και αναλύεται η επιλεχθείσα μεθοδολογία, καθώς και το θεωρητικό υπόβαθρο της Διπλωματικής Εργασίας. Δίνεται ιδιαίτερη έμφαση στις μαθηματικές σχέσεις και τις θεωρίες πάνω στις οποίες βασίζεται η βαθιά μάθηση, ενώ ταυτόχρονα περιγράφονται οι αρχές και οι τρόποι με τους οποίους εφαρμόζεται η μεθοδολογία.

Το **κεφάλαιο 4**, αφορά την **ανάλυση** και τα **αποτελέσματα** της Διπλωματικής Εργασίας. Αρχικά περιγράφεται η διαδικασία συλλογής και επεξεργασίας των δεδομένων που χρησιμοποιήθηκαν. Συγκεκριμένα, γίνεται εκτενής περιγραφή του πειράματος που εφαρμόστηκε και του τρόπου συλλογής των στοιχείων, ενώ ακολούθως αναλύεται ο τρόπος επεξεργασίας των στοιχείων για την διαμόρφωση της τελικής τους μορφής για την εισαγωγή τους στα μοντέλα. Έπειτα παρατίθενται η βάση δεδομένων που προέκυψε από το πείραμα, καθώς και τα βασικά χαρακτηριστικά αυτών. Τέλος, γίνεται αναφορά στον τρόπο με τον οποίο εφαρμόστηκε η μεθοδολογία που ακολουθήθηκε για την παραγωγή των μοντέλων και την εξαγωγή των τελικών αποτελεσμάτων.

Το **κεφάλαιο 5**, περιλαμβάνει τα **συμπεράσματα** που προέκυψαν από την ερμηνεία των αποτελεσμάτων και γίνεται αναφορά στον τρόπο με τον οποίο αυτά μπορούν να χρησιμοποιηθούν. Μέσω αυτών επιχειρείται να δοθεί απάντηση στο αντικείμενο της Διπλωματικής Εργασίας. Τέλος παρουσιάζονται **προτάσεις** για περαιτέρω έρευνα και εξέλιξη των συγκεκριμένων μοντέλων.

2. ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

2.1 ΕΙΣΑΓΩΓΗ

Το συγκεκριμένο κεφάλαιο αφορά στη **βιβλιογραφική ανασκόπηση**, η οποία πραγματοποιήθηκε στο πλαίσιο της εκπόνησης της παρούσας Διπλωματικής Εργασίας. Παρουσιάζονται τα αποτελέσματα ελληνικών και διεθνών επιστημονικών ερευνών, των οποίων το αντικείμενο παρουσιάζει συνάφεια με τα υπό μελέτη ζητήματα, όπως αυτά παρουσιάστηκαν στο προηγούμενο κεφάλαιο.

Αρχικά, παρατίθενται έρευνες οι οποίες αναφέρονται στη **χρήση ΣμηΕΑ** για αναλύσεις κυκλοφοριακών παραγόντων συμφόρησης και συγκεκριμένα την ανίχνευση οχημάτων και την παρακολούθηση της κυκλοφορίας. Στη συνέχεια, παρατίθενται έρευνες εφαρμογής της βαθιάς μάθησης για την κατασκευή **νευρωνικών δικτύων CNN** με διαφορετικές μεθοδολογίες, συγκεκριμένα για την αναγνώριση και την ανάλυση ψηφιακών εικόνων. Στις παρατιθέμενες έρευνες παρουσιάζονται συσχετίσεις με τις **κυκλοφοριακές ουρές** για τον εντοπισμό και την εκτίμηση του μεγέθους τους, είτε μέσω χρήσης ΣμηΕΑ είτε μέσω εφαρμογής βαθιάς μάθησης, κυρίως σε σηματοδοτούμενες διασταυρώσεις. Οι μεθοδολογίες και τα αποτελέσματά τους καταγράφονται αναλυτικά, καθώς απώτερος σκοπός της Διπλωματικής Εργασίας είναι η αναγνώριση κυκλοφοριακών ουρών, οι οποίες εξαρτώνται από τα παραπάνω χαρακτηριστικά.

Για κάθε επιστημονική εργασία πραγματοποιείται συνοπτική αναφορά στη μεθοδολογία που ακολουθήθηκε και στα συμπεράσματα που προέκυψαν. Μέσω της βιβλιογραφικής ανασκόπησης προσδιορίστηκε το ακριβές αντικείμενο της Διπλωματικής Εργασίας και η καταλληλότερη μεθοδολογία για την αντιμετώπισή του.

2.2 ΧΡΗΣΗ ΣΜΗΕΑ ΣΕ ΚΥΚΛΟΦΟΡΙΑΚΕΣ ΑΝΑΛΥΣΕΙΣ

Η χρήση των Συστημάτων μη Επανδρωμένων Αεροσκαφών έχει αυξηθεί ραγδαία τα τελευταία χρόνια. Τα ΣμηΕΑ αξιοποιούνται σε ένα ευρύ φάσμα εφαρμογών όπως ασφάλεια, παρακολούθηση, και περιβαλλοντική χαρτογράφηση, λόγω του δυναμικού σχεδιασμού τους υψηλής κινητικότητας και ταχείας ανταπόκρισης, καθώς και της αντιμετώπισης ανθρώπινων περιορισμών (Khan, et al., 2020). Τα χαρακτηριστικά τους τα καθιστούν κατάλληλα για χρήση σε κυκλοφοριακές αναλύσεις και ειδικότερα στην ανάπτυξη ευφυών συστημάτων μεταφορών και τη δημιουργία ενός αυτόνομου συστήματος παρακολούθησης της κυκλοφορίας.

Στην έρευνα των Puri, et al. (2007) πραγματοποιήθηκε μελέτη βασισμένη στους κυκλοφοριακούς παράγοντες που καταγράφηκαν από πείραμα βιντεοσκόπησης της κυκλοφορίας σηματοδοτούμενης διασταύρωσης, στην περιοχή του Τάμπα, στη Φλόριντα. **Σκοπός** της έρευνας, ήταν η μετατροπή των οπτικών δεδομένων σε «στατιστικά προφίλ» για την προσομοίωση της συμπεριφοράς της κυκλοφορίας. Σημειώνεται ότι για τη συλλογή των δεδομένων χρησιμοποιήθηκε ένα μη επανδρωμένο

ελικόπτερο με ένα πλήρως αυτόνομο σύστημα όρασης και μία τοποθετημένη κάμερα. Η χρήση των «στατιστικών προφίλ» ως μεταβλητές σε μοντέλα προσομοιώσεων έχει ως στόχο τη βελτίωση της βαθμονόμησης και της ακρίβειας των μοντέλων και ταυτόχρονα τη μείωση των διαφορών αποκλίσεων των προσομοιωμένων δεδομένων κίνησης.

Το πείραμα πραγματοποιήθηκε κατά την πρωινή ώρα αιχμής όπου και παρατηρείται ο μεγαλύτερος κυκλοφοριακός φόρτος. Λόγω περιορισμού του χρόνου πτήσης και κάλυψης μίας μόνο διασταύρωσης, τα δεδομένα αφορούν μία μόνο ώρα κυκλοφορίας και εμφανίζουν μεγάλες αποκλίσεις στην κίνηση των οχημάτων. Οι κυκλοφοριακές παράμετροι που εξήχθησαν ήταν: χωρητικότητα και πληρότητα λωρίδων, μέση ταχύτητα δικτύου, ροή και πυκνότητα οχημάτων, (ψευδο-) χωρική-χρονική πυκνότητα, καθυστερήσεις, κινήσεις στροφής και προέλευση - προορισμός οχημάτων.

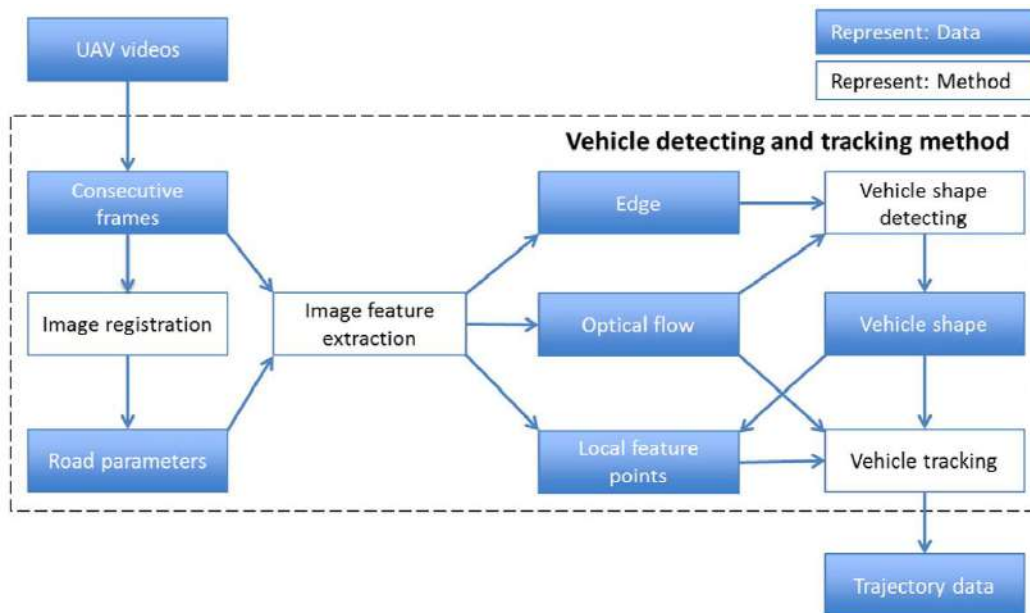
Για την επεξεργασία των δεδομένων χρησιμοποιήθηκε το πρόγραμμα Synchro 6. Με τη χρήση εικονικών ανιχνευτών, για τον εντοπισμό κάθε οχήματος σε κάθε λωρίδα και μέσω ενός αλγόριθμου επεξεργασίας εικόνων, παράχθηκαν εξισώσεις από τις οποίες προσομοιώθηκε το οδικό περιβάλλον και δημιουργήθηκαν τα «στατιστικά προφίλ». Επομένως, προέκυψε το **συμπέρασμα** ότι η χρήση διαθέσιμων δεδομένων από βίντεο σε πραγματικό χρόνο σε σχέση με απαρχαιωμένα δεδομένα παλαιότερων μελετών, προσφέρουν καλύτερα αποτελέσματα στην προσομοίωση της κυκλοφορίας.

Τέλος, **προτείνεται** σε μελλοντικές έρευνες η αύξηση του χρόνου πτήσης και της έκτασης της παρατηρούμενης περιοχής με τη χρήση πολλαπλών μη επανδρωμένων αεροσκαφών σε διαφορετικές διασταυρώσεις και ταυτόχρονη καταγραφή δεδομένων. Προτείνεται επίσης ο συνυπολογισμός της παραμέτρου κίνησης των ίδιων των αεροσκαφών, η οποία προκαλεί μείωση της διαύγειας των εικόνων.

Οι Wang et al. (2016) ασχολήθηκαν με τη συλλογή κυκλοφοριακών δεδομένων μέσω εικόνων και βίντεο από ΣμηΕΑ. **Σκοπός** της εργασίας, ήταν η παρουσίαση τεσσάρων συνεργατικών μεθόδων για τον εντοπισμό των θέσεων των οχημάτων και την εκτίμηση της ταχύτητά τους, με τη χρήση μη επανδρωμένων αεροσκαφών (Διάγραμμα 2.1). Συγκεκριμένα, πραγματοποιήθηκε καταγραφή εικόνων και εξαγωγή των χαρακτηριστικών τους, ενώ ακολούθησε ανίχνευση μεγέθους των οχημάτων και παρακολούθησή τους. Επιπλέον, για την αύξηση της ακρίβειας του παραπάνω συστήματος εισήχθησαν ορισμένα μοναδικά χαρακτηριστικά για να διαφοροποιηθεί η ροή της κυκλοφορίας, ώστε να χρησιμοποιηθούν πολλαπλές διαδοχικές καταγεγραμμένες εικόνες.

Για τις ανάγκες της εργασίας, πραγματοποιήθηκαν 66 πειράματα βιντεοσκόπησης πεδίου συνολικής διάρκειας 850 λεπτών, σε κεντρική αρτηρία του Πεκίνου, στην Κίνα, με τη χρήση ενσωματωμένης κάμερας σε drone, τοποθετημένο σε διαφορετικά υψόμετρα και διαφορετικά πεδία προβολής. Κατά τη διάρκεια της καταγραφής, το drone μετέφερε τη ζωντανή εικόνα στην οθόνη του χειριστή σε χαμηλή ευκρίνεια, καθώς και διάφορα σημαντικά δεδομένα πτήσης. Ταυτόχρονα, πραγματοποιούνταν αποθήκευση των εικόνων υψηλής ευκρίνειας στη μνήμη της κάμερας για ασφάλεια, αλλά και μετέπειτα χρήση.

Η επεξεργασία και η ανάλυση των δεδομένων πραγματοποιήθηκε με το πρόγραμμα MATLAB και την ψηφιακή βιβλιοθήκη VLFeat. Αρχικά, πραγματοποιήθηκε η σύνδεση των καρέ εικόνων του κάθε βίντεο σε ένα ενιαίο σύστημα, ενώ ακολούθησε η εισαγωγή των χαρακτηριστικών της αρτηρίας του πειράματος και η εξαγωγή των σημαντικών πληροφοριών. Η εξαγωγή των άκρων των εικόνων, της οπτικής ροής των οχημάτων και των χαρακτηριστικών τοπικών σημείων, οδήγησε στην ανίχνευση των μεγεθών που ορίζει το κάθε όχημα και έπειτα από επεξεργασία, στον τελικό εντοπισμό των οχημάτων και των κινήσεών τους.



Διάγραμμα 2.1: Διάγραμμα ροής συστήματος ανίχνευσης και παρακολούθησης οχημάτων με ΣμηΕΑ (Πηγή: Wang, et al., 2016)

Εξήχθη, λοιπόν, το **συμπέρασμα** ότι η μέθοδος που παρουσιάζεται παρέχει έναν εύχρηστο και αξιόπιστο τρόπο απόκτησης πληροφοριών οχημάτων. Τα δεδομένα τροχιάς που παράγονται μπορούν να αξιοποιηθούν στις περισσότερες κυκλοφοριακές μελέτες και αναλύσεις δεδομένων κίνησης, ενώ μπορούν να χρησιμοποιηθούν και σε ζωντανή παρακολούθηση της κυκλοφορίας με τον κατάλληλο φορητό εξοπλισμό. Τέλος, **προτείνεται** η επέκταση της μεθόδου σε περισσότερες εφαρμογές για διαφορετικά υψόμετρα και μεικτές καιρικές συνθήκες, καθώς και για διαφορετικό κυκλοφοριακό περιβάλλον.

Η έρευνα των Kaufmann et al. (2018) αφορά εξεταζόμενη μεθοδολογία αναγνώρισης μοτίβων κίνησης συγχρονισμένης ροής για υπερκορεσμένη κυκλοφορία, σε σηματοδοτούμενη διασταύρωση στο Ντίσελντορφ, στη Γερμανία. **Σκοπός** της έρευνας, ήταν η χρήση μη επανδρωμένων αεροσκαφών για την καταγραφή εναέριων παρατηρήσεων, για την εξαγωγή εμπειρικών δεδομένων αναφορικά με τα μοτίβα κίνησης των οχημάτων. Συγκεκριμένα, πραγματοποιήθηκε εποπτευόμενη παρακολούθηση της κυκλοφορίας από κάμερες σε drones, ενώ τα αποτελέσματα βασίστηκαν στα χαρακτηριστικά του οδικού δικτύου και στην εμφάνιση των μοτίβων συγχρονισμένης ροής στους σηματοδότες της διασταύρωσης. Η παραπάνω θεώρηση

αναπαριστά τους σχηματισμούς ουρών κατά την άφιξη οχημάτων στη διασταύρωση, οι οποίες καθιστούν το σύστημα υπερκορεσμένο καθώς αυτό δεν αποσυμφορείται πλήρως σε μία ολοκληρωμένη φάση σηματοδότησης.

Το οδικό τμήμα στο οποίο εφαρμόστηκε η μεθοδολογία έχει μήκος 600 μέτρων, δύο λωρίδες κυκλοφορίας στην αρχή του, ενώ καταλήγει σε τέσσερις στο σημείο της σηματοδότησης (δύο για ευθεία κίνηση και δύο για αριστερή στροφή). Επομένως, ελέγχθηκε και η δημιουργία μοτίβων για ουρές που δημιουργούνται λόγω παρεμπόδισης οχημάτων που θέλουν να πραγματοποιήσουν αριστερή στροφή. Για την επεξεργασία των δεδομένων χρησιμοποιήθηκαν εμπειρικές εξισώσεις, καθώς και η ψηφιακή βιβλιοθήκη OpenCV, η οποία προβλέπεται για εφαρμογές με έμφαση στον πραγματικό χρόνο.

Από τη παραπάνω μεθοδολογία παρατηρήθηκαν: σχηματισμοί μοτίβων κινούμενων ουρών, εισροή ήδη σχηματισμένων συγχρονισμένων μοτίβων στο οδικό τμήμα, σύμπτυξη συγχρονισμένων μοτίβων και σημεία διάλυσης αυτών. Από τα εμπειρικά αποτελέσματα, εξήχθη το **συμπέρασμα** ότι η συγκεκριμένη μεθοδολογία αποτελεί έναν σημαντικό τρόπο για τη διεξαγωγή χωροχρονικών μετρήσεων της κυκλοφορίας, αλλά και μικροσκοπικών δεδομένων για κάθε όχημα ξεχωριστά, με καταγεγραμμένα δεδομένα από ΣμηΕΑ σε πραγματικό χρόνο.

2.3 ΕΦΑΡΜΟΓΕΣ ΒΑΘΙΑΣ ΜΑΘΗΣΗΣ ΣΤΗΝ ΑΝΑΛΥΣΗ ΚΥΚΛΟΦΟΡΙΑΣ

Η εξέλιξη της βαθιάς μάθησης και συγκεκριμένα των νευρωνικών δικτύων σε εφαρμογές υπολογιστικής όρασης έχουν βελτιώσει την ακρίβεια των αλγόριθμων αναγνώρισης αντικειμένων και επεξεργασίας προβλημάτων σε πραγματικό χρόνο (Benjdira, et al., 2018). Οι πιο διαδεδομένες και αποτελεσματικές μορφές δικτύων στην ανάλυση εικόνας είναι τα συνελκτικά νευρωνικά δίκτυα καθώς έχουν τη δυνατότητα εκπαίδευσης ιδιαίτερων περιγραφικών χαρακτηριστικών υψηλής υπολογιστικής ισχύος λόγω της χρήσης ισχυρών γραφικών μονάδων επεξεργασίας (GPUs).

Οι Chakraborty et al. (2018) ασχολήθηκαν με τη δημιουργία νευρωνικών δικτύων βαθιάς μάθησης, με **σκοπό** την ανίχνευση κυκλοφοριακών συμφορήσεων που εμφανίζονται σε ένα οδικό σύστημα, με τη χρήση εικόνων από κάμερες κλειστών κυκλωμάτων παρακολούθησης (CCTV). Τα δεδομένα λήφθηκαν από καταγεγραμμένα βίντεο σε διάφορους αυτοκινητοδρόμους των μεγαλύτερων πόλεων της Αϊοβα. Σημειώνεται, ότι για την επεξεργασία των δεδομένων χρησιμοποιήθηκαν δύο διαφορετικά νευρωνικά δίκτυα βαθιάς μάθησης, YOLO και CNN, καθώς και ο βοηθητικός αλγόριθμος εποπτευόμενης μάθησης SVM (Support Vector Machine). Για την τελική επιλογή του δείγματος χρησιμοποιήθηκαν αισθητήρες ραντάρ στην περιοχή προέλευσης των εικόνων, για το διαχωρισμό τους ανάλογα με την ευκρίνειά τους λόγω διαφορετικών συνθηκών.

Η επεξεργασία των εικόνων πραγματοποιήθηκε με δύο προσεγγίσεις. Η πρώτη που ακολουθήθηκε ήταν η εισαγωγή της κάθε εικόνας σε μοντέλο αναγνώρισης αντικειμένου, το οποίο προσδιόριζε τον αριθμό των οχημάτων που παρατηρούνται και όταν αυτός

υπερέβαινε ένα όριο, τότε η εικόνα χαρακτηριζόταν ως «συμφορημένη». Η δεύτερη προσέγγιση ταξινομούσε ολόκληρη την εικόνα ως «συμφορημένη» ή μη, χωρίς τον έλεγχο ορίου των οχημάτων. Επιπλέον, για την πολυπλοκότητα των μοντέλων χρησιμοποιήθηκαν ξεχωριστά δείγματα εικόνων, καταγεγραμμένες κατά τη διάρκεια της ημέρας και της νύχτας, ικανοποιητικής και μη ικανοποιητικής ανάλυσης αντίστοιχα. Συνολικά χρησιμοποιήθηκαν 1400 εικόνες για εκπαίδευση και 1018 για δοκιμή.

Για την αξιολόγηση των αλγόριθμων χρησιμοποιήθηκαν οι τυπικές μετρικές ακρίβειας, ανάκλησης και συνοχής ακρίβειας (precision, recall, accuracy) και οι κατηγορίες TP, FP, TN, FN που αντιστοιχούνταν σε κάθε εικόνα όταν αυτή είχε χαρακτηριστεί σωστά ή λανθασμένα. Από τα αποτελέσματα κατασκευάστηκαν τα διαγράμματα χαρακτηριστικού λειτουργίας δέκτη (ROC) για τη σύγκριση των αλγόριθμων. Από αυτά, προέκυψε το **συμπέρασμα** ότι και οι τρεις αλγόριθμοι αποτελούν αξιόπιστες τεχνικές για την ανίχνευση κυκλοφοριακών συμφορήσεων με ικανοποιητική ακρίβεια, με τα δίκτυα YOLO και CCN βαθιάς μάθησης να παρουσιάζουν καλύτερα αποτελέσματα ακόμη και στις διαφορετικές μεικτές συνθήκες, από ότι ο αλγόριθμος εποπτευόμενης μάθησης SVM.

Στην έρευνα των Lee et al. (2019) αναπτύχθηκε ένα νέο μοντέλο εκτίμησης, σε πραγματικό χρόνο, του μήκους κυκλοφοριακής ουράς που βασίζεται σε λωρίδες κυκλοφορίας, με το συνδυασμό δύο διαδικασιών βαθιάς μάθησης, ενός νευρωνικού δικτύου CNN και ενός τεχνητού επαναλαμβανόμενου δικτύου μακροπρόθεσμης μνήμης LSTM. Εκτός της εκτίμησης του μήκους ουράς, **σκοπός** της έρευνας ήταν να ξεπεράσει τα αντίστοιχα πρόσφατα ανεπτυγμένα μοντέλα εκτίμησης, πραγματοποιώντας ταυτοχρόνως καταγραφή χωροχρονικών χαρακτηριστικών της κυκλοφορίας σε πραγματικό χρόνο, όπως χρόνοι αφίξεων οχημάτων και πλήθος οχημάτων σε κάθε λωρίδα. Τα δεδομένα λήφθηκαν από παλιά βάση δεδομένων εικόνων σηματοδοτούμενης διασταύρωσης, σε περιοχή του Λος Άντζελες, στην Καλιφόρνια.

Η διασταύρωση προσομοιώθηκε στο πρόγραμμα AIMSUN, στο οποίο τοποθετήθηκαν τα γεωμετρικά χαρακτηριστικά της, περίοδος σηματοδότησης 60sec και εναλλασσόμενος κυκλοφοριακός φόρτος 300, 600, 900 και 1200 οχημάτων ανά ώρα. Για την εκπαίδευση χρησιμοποιήθηκαν δεδομένα 863 περιόδων σηματοδότησης και 576 για δοκιμή. Η κατηγοριοποίηση των εικόνων πραγματοποιήθηκε με δυαδικό κώδικα, όπου στις περιπτώσεις εντοπισμού τουλάχιστον ενός τμήματος ενός οχήματος, οι εικόνες κατηγοριοποιούνταν ως 1 και στις υπόλοιπες περιπτώσεις ως 0. Έπειτα από εισαγωγή διαφόρων παραγόντων οδηγικής συμπεριφοράς, όπως αλλαγές λωρίδων, ακολούθησε η επεξεργασία των δεδομένων με το LSTM και εκπαιδεύτηκε το δίκτυο CNN με 7 διαφορετικά επίπεδα προσαρμογής.

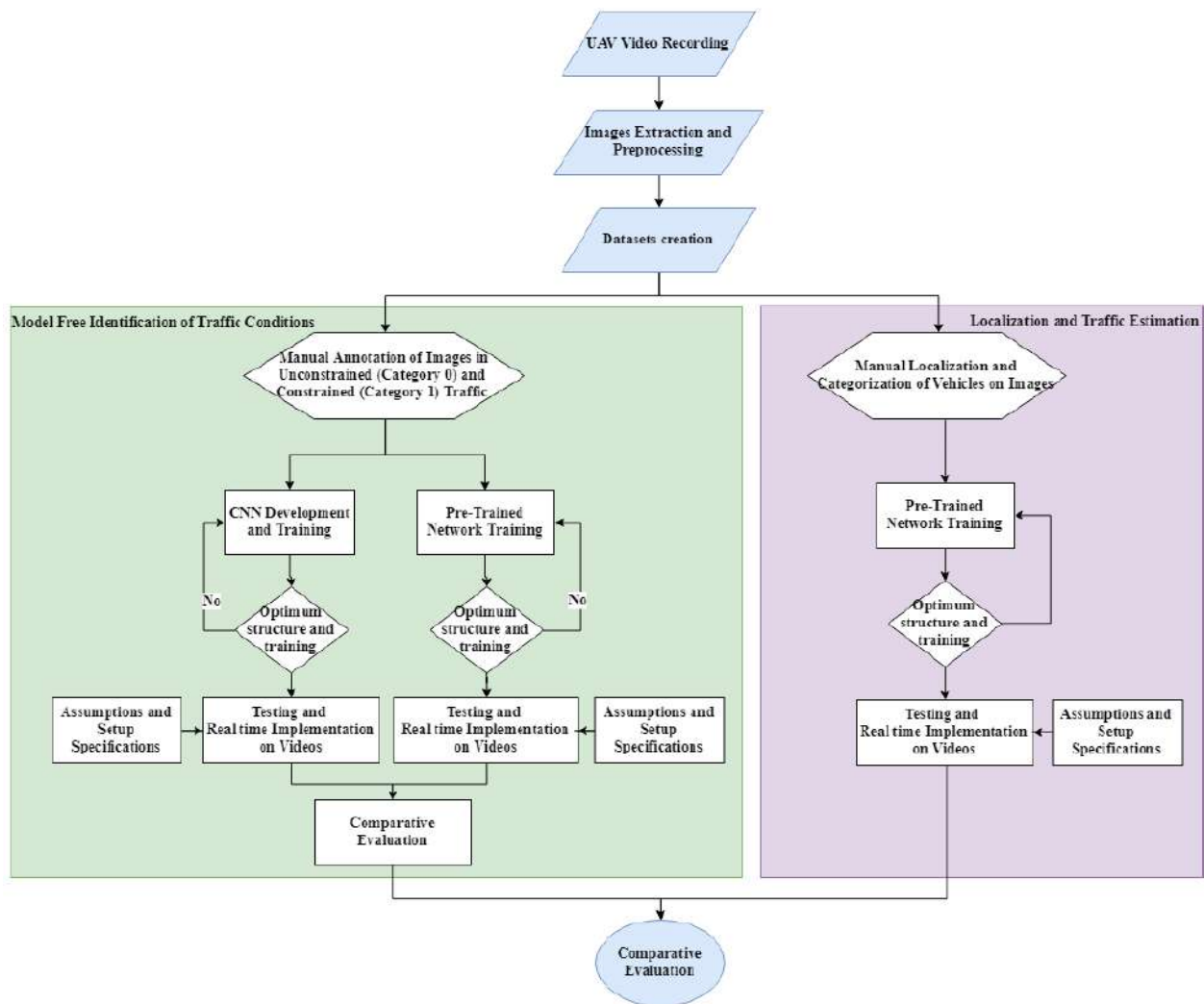
Για μεγαλύτερη ακρίβεια, εφαρμόστηκαν τρία διαφορετικά μοντέλα κυκλοφοριακών ουρών βασισμένων σε λωρίδες, από τα οποία τα δύο χρησιμοποιήθηκαν σε παλαιότερες έρευνες, ενώ το τρίτο ήταν το προτεινόμενο της παρούσας έρευνας. Από τα αποτελέσματα των μοντέλων, προέκυψαν οι τιμές για τις τυπικές μετρικές ακρίβειας και κατασκευάστηκαν οι καμπύλες ROC για την αξιολόγησή τους. Το **συμπέρασμα** που εξήχθη, ήταν ότι η αρχιτεκτονική του προτεινόμενου μοντέλου με τη χρήση πραγματικών και συνθετικών δεδομένων παρήγαγε βελτιωμένα αποτελέσματα συγκριτικά με τα

παλαιότερα μοντέλα, αξιόπιστα για χρήση ακόμη και για ζωντανή παρακολούθηση της κυκλοφορίας. Σημειώνεται ότι για το συνδυασμένο μοντέλο απαιτήθηκε χαμηλότερο επίπεδο υπολογιστικής επιβάρυνσης για τη βαθμονόμηση και τον έλεγχο του.

Η έρευνα των Vlahogianni et al. (2021) συγκροτήθηκε με **σκοπό** τη δημιουργία ενός μεθοδολογικού πλαισίου εντοπισμού κυκλοφοριακών συνθηκών, βασισμένο σε βιντεοσκοπημένα δεδομένα καταγεγραμμένα από μη επανδρωμένα αεροσκάφη και με χρήση βαθιάς μάθησης. Συγκεκριμένα, προτάθηκαν δύο συμπληρωματικές προσεγγίσεις (Διάγραμμα 2.2). Η πρώτη είχε ως στόχο τον προσδιορισμό, σε πραγματικό χρόνο και με ελάχιστο υπολογιστικό κόστος, των κυκλοφοριακών συνθηκών με περιορισμούς αλλά και χωρίς. Στόχος της δεύτερης ήταν ο εντοπισμός και η ταξινόμηση των οχημάτων, καθώς και η προσέγγιση των κυκλοφοριακών μεταβλητών (όγκος, ταχύτητα, πυκνότητα) των οχημάτων του οδικού τμήματος στο οποίο πραγματοποιήθηκε το πείραμα βιντεοσκόπησης. Για την αντιμετώπιση των παραπάνω εκπαιδεύτηκαν νευρωνικά δίκτυα CNN και SSD MobiNet v2, ενώ για λόγους σύγκρισης χρησιμοποιήθηκε και ένα προ-εκπαιδευμένο μοντέλο ResNet50 και πραγματοποιήθηκε έλεγχος με βάση την ακρίβεια και τη σκοπιμότητα εφαρμογής τους.

Για τη συλλογή των δεδομένων χρησιμοποιήθηκε ένα drone μέγιστης πτήσης 18 λεπτών με τοποθετημένη κάμερα υψηλής ανάλυσης (4K). Η πτήση πραγματοποιήθηκε στην περιοχή της Πολυτεχνειούπολης της Αθήνας, στην Ελλάδα, σε συγκεκριμένο τμήμα αρτηρίας με τρεις λωρίδες ανά κατεύθυνση, για την αποφυγή πιθανών περίπλοκων μοτίβων κυκλοφορίας από άλλους κυκλοφοριακούς παράγοντες. Η κατηγοριοποίηση των εικόνων πραγματοποιήθηκε με δυαδικό κώδικα με οπτική επιθεώρηση από τους μελετητές. Η «κατηγορία 1» αφορούσε τις περιπτώσεις στις οποίες εντοπίστηκαν τουλάχιστον τέσσερα οχήματα στην εικόνα, τουλάχιστον τρία στην ίδια ή και περισσότερες λωρίδες κίνησης, ή ακόμη και τουλάχιστον ένα σε κάθε λωρίδα. Όλες οι υπόλοιπες περιπτώσεις αφορούσαν την «κατηγορία 2».

Σημειώνεται ότι για την ανάπτυξη και την αξιολόγηση των μοντέλων της πρώτης προσέγγισης, το σύνολο των δεδομένων χωρίστηκε σε τρία δείγματα ίδιου πλήθους. Το πρώτο αξιοποιήθηκε για την προκαταρκτική αξιολόγηση διαφορετικών νευρωνικών δικτύων με 2000 εικόνες για εκπαίδευση και 200 για δοκιμή. Το δεύτερο χρησιμοποιήθηκε για την ανάπτυξη του μοντέλου με 50000 και 12000 εικόνες για εκπαίδευση και δοκιμές αντίστοιχα, ενώ το τρίτο δείγμα 2000 εικόνων για την αξιολόγηση των ανεπτυγμένων μοντέλων. Ο προγραμματισμός πραγματοποιήθηκε στο πρόγραμμα Python και με την ψηφιακή βιβλιοθήκη OpenCV. Αντιθέτως, για τα μοντέλα της δεύτερης προσέγγισης χρησιμοποιήθηκε ένα δείγμα 2000 εικόνων μεικτών συνθηκών, από το οποίο εξήχθησαν 4000 οχήματα (υποσύνολο για εκπαίδευση και δοκιμή 3600 και 400 αντίστοιχα) για τα οποία πραγματοποιήθηκε κατηγοριοποίηση και προσδιορισμός κυκλοφοριακών μεταβλητών.



Διάγραμμα 2.2: Διάγραμμα ροής διαδικασιών προτεινόμενου μεθοδολογικού πλαισίου εντοπισμού κυκλοφοριακών συνθηκών (Πηγή: Vlahogianni, et al., 2021)

Τα αποτελέσματα όλων των μοντέλων παρουσίασαν ποσοστά ακρίβειας τουλάχιστον 89% με τη κατασκευή διαγραμμάτων ROC. Συγκεκριμένα, το εκπαιδευόμενο CNN συγκριτικά με το προ-εκπαιδευόμενο ResNet50 εμφάνισε 91% μεγαλύτερη ακρίβεια, ενώ το SSD MobNet v2 παρουσίασε αξιόπιστες προβλέψεις της τάξης του 91% επίσης. Επιπλέον, εξήχθη το **συμπέρασμα**, ότι η πρώτη προσέγγιση ήταν αρκετά πιο αποδοτική όσον αφορά τους υπολογιστικούς χρόνους, έως και πέντε φορές ταχύτερη από τη δεύτερη προσέγγιση, η οποία όμως ήταν πολύ πιο ισχυρή αναφορικά με τις μεικτές συνθήκες καταγραφής των δεδομένων, όπως αστάθεια βίντεο και εμφάνιση σκίων και άλλων αντικειμένων στις εικόνες. Τέλος, παρατηρήθηκε η χρησιμότητα της μεταφοράς μάθησης (transfer learning) που δοκιμάστηκε στην παρούσα έρευνα, η οποία μπορεί να εφαρμοστεί σε προβλήματα γρήγορης και οικονομικής αναγνώρισης κυκλοφοριακών συνθηκών, με την δυνατότητα αύξησης της χωρικής κάλυψης των συστημάτων παρακολούθησης μίας εξαζόμενης περιοχής. Για μελλοντική έρευνα, **προτείνεται** η αναγνώριση σχηματισμού και διάλυσης κυκλοφοριακών ουρών και ο εντοπισμός θέσεων στάθμευσης σε πραγματικό χρόνο, καθώς και η ενσωμάτωση των ΣμηΕΑ σε συνεργατικά ευφυή συστήματα μεταφορών με παρόμοιες μεθοδολογίες.

2.4 ΣΥΝΟΨΗ ΒΙΒΛΙΟΓΡΑΦΙΑΣ

Στο κεφάλαιο αυτό παρουσιάστηκαν οι σημαντικότερες ελληνικές και διεθνείς έρευνες με αντικείμενο συναφές με αυτό της παρούσας Διπλωματικής Εργασίας. Ειδικότερα, παρατέθηκαν έρευνες οι οποίες αφορούν την εφαρμογή της τεχνολογίας των ΣμηΕΑ σε κυκλοφοριακές αναλύσεις, καθώς και έρευνες στις οποίες εφαρμόζονται κατασκευασμένα μοντέλα με χρήση βαθιάς μάθησης και νευρωνικών δικτύων. Επιπλέον, από τις παρατιθέμενες έρευνες παρατηρείται η συσχέτισή τους με το φαινόμενο των κυκλοφοριακών ουρών. Από τα ευρήματα της βιβλιογραφικής ανασκόπησης και τη σύνθεση των κύριων σημείων τους, προέκυψαν οι παρακάτω παρατηρήσεις:

- ♦ Τα **πειράματα βιντεοσκόπησης** με συστήματα μη επανδρωμένων αεροσκαφών, αποτελούν έναν εύκολο και γρήγορο τρόπο συλλογής δεδομένων κυκλοφοριακών μελετών σε σχέση με τις κλασσικές μεθόδους.
- ♦ Σημαντικό **πλεονέκτημα** της χρήσης των ΣμηΕΑ η συλλογή μεγάλων όγκων δεδομένων, όμως με **μειονεκτήματα** τους μικρούς χρόνους πτήσεων και τη μειωμένη διαύγεια των εικόνων λόγω καιρικών συνθηκών και μη σταθερότητας των συστημάτων.
- ♦ Τα μοντέλα που βασίζονται σε **πραγματικά δεδομένα** μπορούν να οδηγήσουν σε καλύτερα αποτελέσματα συγκριτικά με τα παραδοσιακά μοντέλα που χρησιμοποιούν τυπικές μαθηματικές σχέσεις.
- ♦ Οι μέθοδοι βαθιάς μάθησης παρέχουν τη δυνατότητα διαχείρισης **μεγάλων βάσεων δεδομένων** σε πραγματικό χρόνο και τη κατασκευή μοντέλων για την αντιμετώπιση πολύπλοκων προβλημάτων.
- ♦ Η ανάλυση προβλημάτων κυκλοφοριακής συμφόρησης απαιτεί **αναγνώριση και παρακολούθηση της κυκλοφορίας** της εξεταζόμενης περιοχής, καθώς και τον προσδιορισμό των κυκλοφοριακών μεταβλητών των οχημάτων.
- ♦ Η **αρχιτεκτονική** των νευρωνικών δικτύων επηρεάζει σημαντικά την ακρίβεια των εκπαιδευμένων μοντέλων χωρίς σημαντική υπολογιστική επιβάρυνση και μπορεί να αξιοποιηθεί για προβλήματα ταξινόμησης δυαδικού χαρακτήρα, αλλά και πολλαπλών κλάσεων.
- ♦ Η ανίχνευση των κυκλοφοριακών ουρών εξαρτάται κυρίως από τον εντοπισμό των οχημάτων τα οποία σχηματίζουν **μοτίβα κίνησης** στην κυκλοφορία, και επιβαρύνουν το σύστημα, ενώ επηρεάζονται και από μεικτούς μη σταθερούς παράγοντας.

3. ΜΕΘΟΔΟΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ

3.1 ΕΙΣΑΓΩΓΗ

Στο παρόν κεφάλαιο παρουσιάζεται το **θεωρητικό υπόβαθρο**, στο οποίο στηρίχτηκε η ανάλυση των στοιχείων και η εφαρμογή της μεθόδου της παρούσας Διπλωματικής Εργασίας. Αρχικά, παρουσιάζεται η μεθοδολογική προσέγγιση. Στη συνέχεια, γίνεται περιγραφή της βαθιάς μάθησης και της εποπτευόμενης μάθησης για την αντιμετώπιση προβλημάτων ταξινόμησης κλάσεων δυαδικού χαρακτήρα (0 και 1) και πολλαπλών κλάσεων (τουλάχιστον τριών διαφορετικών). Επιπλέον, περιγράφεται η αρχιτεκτονική των νευρωνικών δικτύων και οι μαθηματικές σχέσεις καθορισμού πολυπλοκότητας των μοντέλων εκπαίδευσης και συγκεκριμένα οι παράγοντες και οι τεχνικές που χρησιμοποιούνται στη διαδικασία ανάπτυξης δικτύων αντιμετώπισης προβλημάτων ταξινόμησης. Τέλος, πραγματοποιείται αναφορά στη μέθοδο μεταφοράς μάθησης.

3.2 ΠΡΟΤΕΙΝΟΜΕΝΗ ΠΡΟΣΕΓΓΙΣΗ

Η προτεινόμενη προσέγγιση για την ανάπτυξη ενός μοντέλου νευρωνικού δικτύου βαθιάς μάθησης με χρήση καταγεγραμμένων δεδομένων από ΣμηΕΑ, περιγράφεται σε βήματα στο Διάγραμμα 3.1.



Διάγραμμα 3.1: Διάγραμμα ροής ανάπτυξης ενός μοντέλου νευρωνικού δικτύου βαθιάς μάθησης με χρήση καταγεγραμμένων δεδομένων από ΣμηΕΑ

Βήμα 1: Τροποποίηση δεδομένων

Για τα δεδομένα που λήφθηκαν από τα πειράματα βιντεοσκόπησης απαιτήθηκε **αποκοπή των βίντεο σε εικόνες** για την εξαγωγή 25 καρέ ανά δευτερόλεπτο και τη δημιουργία μίας ψηφιακή βάση εικόνων με την **κατασκευή κώδικα γλώσσας προγραμματισμού Python**. Επιπλέον, οι κλάσεις με τις οποίες κατηγοριοποιήθηκαν οι ουρές κατά την επεξεργασία των δεδομένων, χωρίστηκαν με τον ίδιο τρόπο στο **πρόγραμμα Microsoft Excel**, ώστε να αντιστοιχούν στον ίδιο αριθμό καρέ εικόνων.

Βήμα 2: Μετατροπή ψηφιακής βάσης δεδομένων

Για τη μετατροπή της ψηφιακής βάσης εικόνων κάθε βίντεο, κατασκευάστηκε **κώδικας μετασχηματισμού**, ο οποίος μετέτρεψε τις εικόνες σε **λίστες δεδομένων εκπαίδευσης**, σε μορφή πινάκων όμοια με εκείνη των κλάσεων κατηγοριοποίησης. Κατά τη μετατροπή αυτή, πραγματοποιήθηκε επιπλέον αλλαγή των διαστάσεων των εικόνων για τη μείωση του μεγέθους τους και τη διευκόλυνση της εισαγωγής τους στο μοντέλο.

Βήμα 3: Προετοιμασία δεδομένων

Με την εισαγωγή των τροποποιημένων δεδομένων στον **αλγόριθμο**, αρχικά πραγματοποιήθηκε **διαχωρισμός δειγμάτων** για την εκπαίδευση και τη δοκιμή του μοντέλου (25% του συνόλου για δοκιμές). Στη συνέχεια, προσδιορίστηκαν οι **ονομασίες των κλάσεων** για την κατηγοριοποίηση, τις οποίες θα χρησιμοποιήσει το μοντέλο, ενώ έπειτα πραγματοποιήθηκε **ομαλοποίηση των τιμών των εικονοκυττάρων (pixels)** κάθε δείγματος εικόνων για εκπαίδευση και δοκιμή αντίστοιχα, για την υπολογιστική διευκόλυνση του μοντέλου.

Βήμα 4: Κατασκευή μοντέλου νευρωνικού δικτύου

Για την κατασκευή του μοντέλου επιλέχθηκε **αρχιτεκτονική επακόλουθης διαδικασίας εκπαίδευσης**, κατά την οποία τα δεδομένα περνούν από συνεχόμενα στάδια για την εκπαίδευση. Αρχικά, δημιουργήθηκαν τα διαφορετικά **στρώματα (layers)** με τα οποία πραγματοποιείται **εξαγωγή των πληροφοριών** των εικόνων και σχηματισμός νέων περίπλοκων χαρακτηριστικών. Η σύνταξη του μοντέλου ολοκληρώθηκε με τη δημιουργία των επιπέδων νευρώνων, τα οποία σχηματίζουν το νευρωνικό δίκτυο, το οποίο με τη σειρά του συνδέει τα παραπάνω χαρακτηριστικά και δημιουργεί τον **εγκέφαλο του μοντέλου**.

Βήμα 5: Έλεγχος μοντέλου

Ο έλεγχος πραγματοποιήθηκε με τον υπολογισμό των μετρικών του μοντέλου, καθώς και με την πρόβλεψη κατηγοριοποίησης εικόνων του δείγματος δοκιμής. Συγκεκριμένα, υπολογίστηκαν τα **μητρώα σύγχυσης (confusion matrices)** για κάθε κλάση με τη χρήση των **τυπικών μετρικών** ακρίβειας, ανάκλησης και συνοχής ακρίβειας (precision, recall, accuracy) και των χαρακτηρισμών που αντιστοιχούνταν σε κάθε εικόνα όταν αυτή είχε χαρακτηριστεί σωστά ή λανθασμένα κατά την εκπαίδευση του μοντέλου. Επιπλέον,

με τυχαία επιλογή κάποιας εικόνας του δείγματος δοκιμής και εφαρμογή του εκπαιδευμένου ταξινομητή, παρουσιάστηκαν οι ποσοστιαίες προβλέψεις της κάθε κλάσης που μπορεί να αντιστοιχεί στην επιλεγμένη εικόνα, ενώ ύστερα ελέγχθηκε οπτικά η ορθότητα της πρόβλεψης.

Βήμα 6: Αξιολόγηση μοντέλου

Η μεθοδολογία ολοκληρώθηκε με την χρήση των υπολογισμένων τυπικών μετρικών κατά τη διαδικασία ελέγχου, για την κατασκευή των **καμπυλών ROC** και της **τιμής AUC** για όλες τις κατηγορίες κλάσεων, η οποία δηλώνει την περιοχή κάτω από την καμπύλη, για την αξιολόγηση της ποιότητας του μοντέλου κατηγοριοποίησης.

3.3 ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

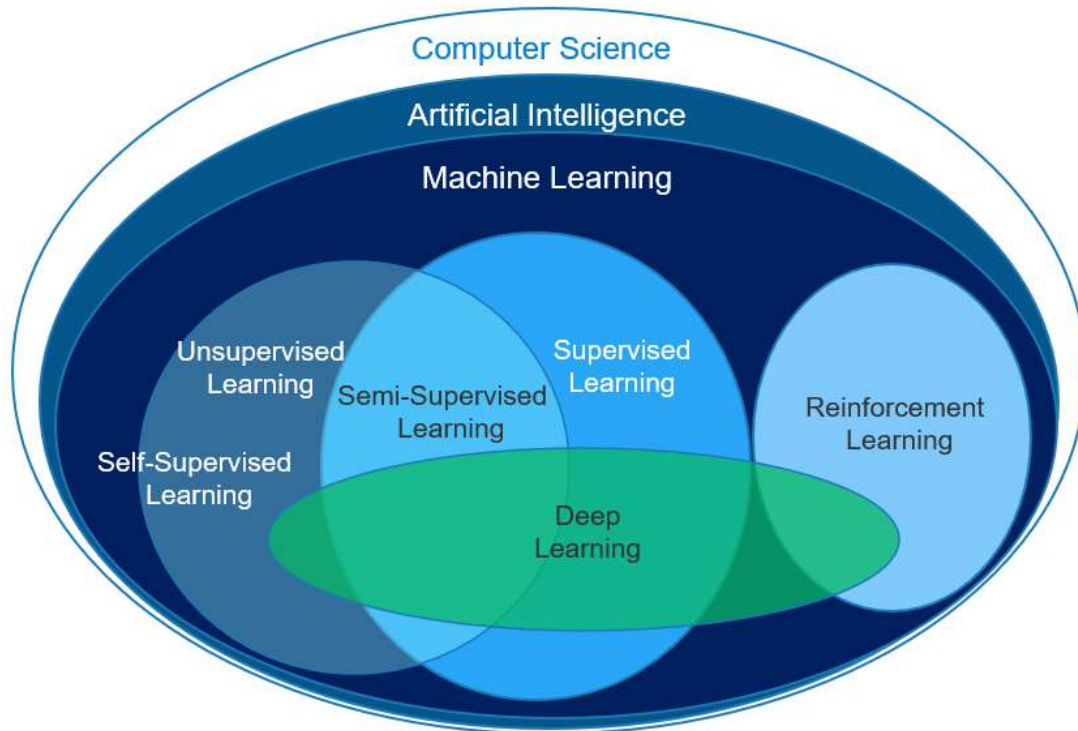
3.3.1 Βαθιά Μάθηση

Η βαθιά μάθηση είναι μία εξειδικευμένη μορφή μηχανικής μάθησης, η οποία σχετίζεται με αλγόριθμους οι οποίοι προσομοιάζουν τη δομή και τη λειτουργία ενός ανθρώπινου εγκεφάλου και επιτρέπουν συστήματα να εκτελούν πολύπλοκες εργασίες με αυξανόμενη ακρίβεια. Επικεντρώνεται στην ανάπτυξη προγραμμάτων ηλεκτρονικών υπολογιστών και απαιτεί **μεγάλες ποσότητες επισημασμένων δεδομένων (labeled data)**, καθώς και σημαντική **υπολογιστική ισχύ (computing power)** για τη μείωση του χρόνου εκπαίδευσης (Mysla, 2020). Το βασικότερο χαρακτηριστικό της βαθιάς μάθησης είναι η χρήση πολλαπλών επιπέδων στο δίκτυο, ώστε να πραγματοποιείται μια βελτιστοποιημένη πρακτική εφαρμογή σταδιακής εξαγωγής πληροφοριών.

Οι αλγόριθμοι βαθιάς μάθησης έχουν τέσσερις βασικές κατηγορίες (Εικόνα 3.1):

- ♦ Οι **αλγόριθμοι εποπτευόμενης μάθησης** εφαρμόζουν τη γνώση εκπαιδευμένων μοντέλων στο παρελθόν, για την πρόβλεψη μελλοντικών γεγονότων σε καινούργια δεδομένα. Η ανθρώπινη επίβλεψη πραγματοποιείται κατά το στάδιο επισημάνσης των δεδομένων και προσδιορισμού επιθυμητών τιμών εξόδου. Οι αλγόριθμοι παράγουν μία συνάρτηση πρόβλεψης σχετικά με τις τιμές εξόδου και συγκρίνουν τις τιμές αυτές με τις σωστές προβλεπόμενες, ώστε να εντοπιστούν λάθη και να τροποποιηθούν τα μοντέλα.
- ♦ Οι **αλγόριθμοι μη εποπτευόμενης μάθησης** εξετάζουν τον τρόπο με τον οποίο τα συστήματα καταλήγουν σε μία συνάρτηση πρόβλεψης χωρίς την ανθρώπινη επίβλεψη και τη χρήση μη επισημασμένων δεδομένων. Οι αλγόριθμοι δεν υπολογίζουν τις σωστές τιμές εξόδου, αλλά διερευνούν τα δεδομένα και εξαγουν συμπεράσματα από αυτά για τις κρυφές δομές τους.
- ♦ Οι **αλγόριθμοι ημι-εποπτευόμενης μάθησης** χρησιμοποιούν έναν συνδυασμό εποπτευόμενης και μη μάθησης, καθώς εκμεταλλεύονται επισημασμένα και μη δεδομένα. Οι αλγόριθμοι αυτοί μπορούν να βελτιώσουν σημαντικά την ακρίβεια της μάθησης και συνήθως επιλέγονται όταν τα συλλεγμένα δεδομένα απαιτούν συγκεκριμένη προσέγγιση και λογισμικό υπολογιστικού συστήματος για την εκπαίδευση.

- ♦ Οι **αλγόριθμοι ενισχυτικής μάθησης** αλληλοεπιδρούν με ένα διαδραστικό περιβάλλον, παράγουν ενέργειες και ανακαλύπτουν σφάλματα. Προσδιορίζουν αυτόματα την ιδανική συμπεριφορά μέσα σε ένα συγκεκριμένο πλαίσιο για την μεγιστοποίηση της απόδοσης ενός συστήματος, ενώ με τη χρήση ενός συνεχούς βρόγχου ανατροφοδότησης, τα μοντέλα μπορούν με την πάροδο του χρόνου να αυτοβελτιωθούν.



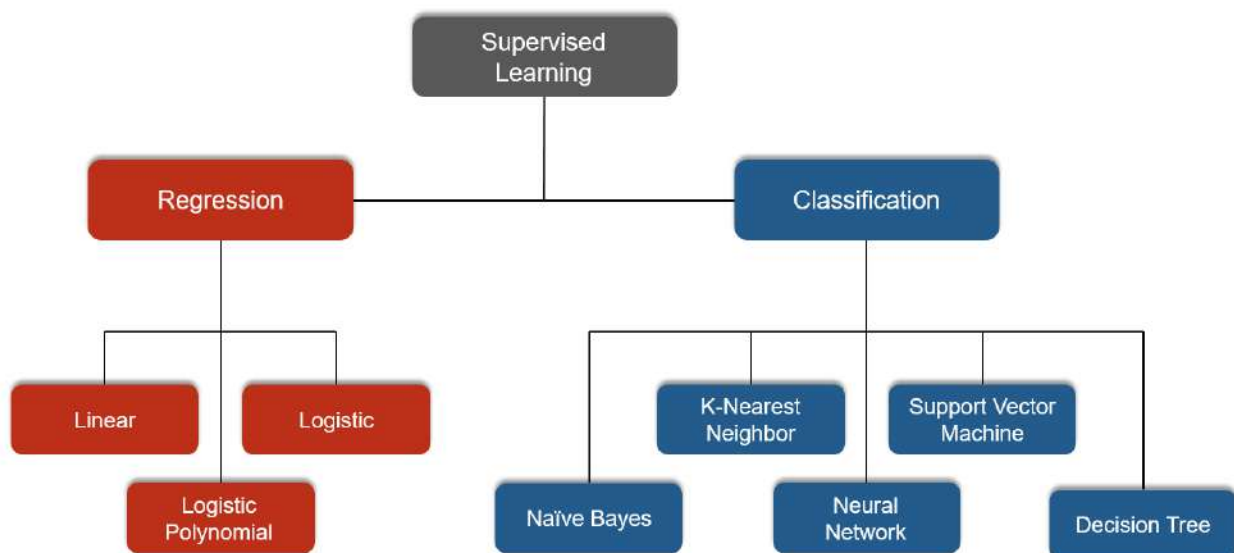
Εικόνα 3.1: Δομή επιστήμης υπολογιστών και τεχνητής νοημοσύνης και κατηγορίες αλγόριθμων βαθιάς μάθησης (Πηγή: Mysla, 2020)

3.3.2 Εποπτευόμενη Μάθηση

Τα μοντέλα εποπτευόμενης μάθησης είναι σχεδιασμένα ώστε να μπορούν να μαθαίνουν χρησιμοποιώντας άλλες πληροφορίες ως παράδειγμα (Wilson, 2019). Κατά την εκπαίδευσή τους, χρησιμοποιούνται **επισημασμένα δεδομένα** τα οποία συνδυάζονται με σωστές τιμές εξόδου με **ανθρώπινη επίβλεψη**. Αναζητούνται μοτίβα τα οποία σχετίζονται με τα επιθυμητά αποτελέσματα με στόχο την πρόβλεψη των σωστών τιμών εξόδου. Η εποπτευόμενη μάθηση χωρίζεται σε δύο τύπους προβλημάτων κατά την εξόρυξη δεδομένων: Παλινδρόμηση και Ταξινόμηση (Διάγραμμα 3.2)

Η **παλινδρόμηση (regression)** αποτελεί μία προγνωστική στατιστική διαδικασία, όπου το μοντέλο προσπαθεί να αναγνωρίσει τη σημαντική σχέση μεταξύ εξαρτημένων και ανεξάρτητων μεταβλητών. Ο στόχος ενός αλγόριθμου παλινδρόμησης είναι η πρόβλεψη μίας συνεχούς μεταβλητής. Οι κατηγορίες αλγόριθμων παλινδρόμησης είναι η γραμμική, η λογιστική και η πολυωνυμική λογιστική παλινδρόμηση.

Η **ταξινόμηση (classification)** αναφέρεται σε ένα προγνωστικό πρόβλημα μοντελοποίησης για το οποίο χρησιμοποιείται ένας αλγόριθμος για την ακριβή κατηγοριοποίηση δεδομένων δοκιμής σε συγκεκριμένες κατηγορίες. Ειδικότερα, κατά τη διάρκεια εκπαίδευσης ενός μοντέλου εισάγονται δεδομένα, από τα οποία το μοντέλο αναγνωρίζει συγκεκριμένες οντότητες και προσπαθεί να εξάγει συμπεράσματα σχετικά με τον τρόπο ορισμού αυτών σε δείγμα δεδομένων δοκιμής. Τα προβλήματα ταξινόμησης μπορούν να αντιμετωπιστούν με πολλούς αλγόριθμους, ανάλογα τα δεδομένα και την κατάσταση του προβλήματος. Οι σημαντικότεροι αλγόριθμοι ταξινόμησης είναι οι Naïve Bayes, οι K-Nearest Neighbors, τα νευρωνικά δίκτυα, τα μηχανήματα υποστήριξης φορέα (SVM) και τα δέντρα αποφάσεων.



Διάγραμμα 3.2: Αλγόριθμοι παλινδρόμησης και ταξινόμησης

Η αντιμετώπιση ενός προβλήματος ταξινόμησης εποπτευόμενης μάθησης περιγράφεται συνοπτικά στα παρακάτω βήματα:

- ♦ Προσδιορισμός του είδους των δεδομένων
- ♦ Συλλογή δεδομένων εκπαίδευσης και δοκιμής
- ♦ Επιλογή επισήμανσης δεδομένων εισαγωγής
- ♦ Προσδιορισμός δομής εξίσωσης και αντίστοιχου αλγόριθμου μάθησης
- ♦ Κατασκευή αρχιτεκτονικής μοντέλου και επιλογή παραμέτρων
- ♦ Εκπαίδευση μοντέλου με δεδομένα εκπαίδευσης
- ♦ Αξιολόγηση ακρίβειας παραγόμενης συνάρτησης πρόβλεψης
- ♦ Υπολογισμός απόδοσης προκύπτουσας συνάρτησης σε δεδομένα δοκιμής

Στη βαθιά μάθηση η ταξινόμηση αναφέρεται κυρίως σε ένα προγνωστικό πρόβλημα, όπου μία ετικέτα τάξης προβλέπεται για ένα δείγμα δεδομένων εισαγωγής. Τα σημαντικότερα αυτά προβλήματα είναι η **κατηγοριοποίηση εικόνων** και η **ανίχνευση αντικειμένων**, ενώ ο καλύτερος αλγόριθμος για την αντιμετώπισή τους είναι τα νευρωνικά δίκτυα. Οι δύο βασικές κατηγορίες είναι η δυαδική ταξινόμηση και η ταξινόμηση πολλαπλών κλάσεων (Brownlee, 2020).

Διαδική ταξινόμηση (Binary classification): περιλαμβάνει δύο κλάσεις και συνήθως η μία κλάση παρουσιάζει την κατηγορία κανονικής κατάστασης και η άλλη παρουσιάζει μία κατηγορία μη φυσιολογικής κατάστασης. Η πρόβλεψη ενός μοντέλου δυαδικής ταξινόμησης πραγματοποιείται με την διωνυμική κατανομή. Για μία μοναδική πρόβλεψη εφαρμόζεται η κατανομή Bernoulli. Η διωνυμική κατανομή και η Bernoulli είναι διακριτές συναρτήσεις κατανομής τυχαίας μεταβλητής, οι οποίες αφορούν περιπτώσεις όπου ένα πείραμα θα έχει δυαδικό αποτέλεσμα είτε 0 είτε 1. Για την ταξινόμηση, αυτό σημαίνει ότι το μοντέλο προβλέπει την πιθανότητα ενός παραδείγματος να ανήκει στην κανονική κατάσταση (κλάση 1) ή στην μη φυσιολογική κατάσταση (κλάση 0).

Η πιθανότητα κατανομής Bernoulli δίνεται από τη σχέση:

$$P(x) = \begin{cases} 1 - p & \text{για } x = 0 \\ p & \text{για } x = 1 \end{cases} \quad (1)$$

Η πιθανότητα διωνυμικής κατανομής δίνεται από τη σχέση:

$$P(x) = \binom{n}{x} p^x (1 - p)^{n-x} \quad \text{για } x = 0, 1, 2, \dots, n \quad (2)$$

όπου:

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

Ταξινόμηση πολλαπλών κλάσεων (Multi-Class classification): περιλαμβάνει τρεις τουλάχιστον κλάσεις διαφορετικών κατηγοριών, οι οποίες δεν αναφέρονται σε φυσιολογική ή μη κατάσταση. Η πρόβλεψη ενός μοντέλου ταξινόμησης πολλαπλών κλάσεων πραγματοποιείται με την πολυωνυμική κατανομή. Η πολυωνυμική κατανομή είναι η γενίκευση της διωνυμικής και χρησιμοποιείται για την εύρεση της πιθανότητας σωστής πρόβλεψης σειράς επαναλήψεων ανεξάρτητων τυχαίων ενδεχομένων, το καθένα εκ των οποίων έχει τη δική του γνωστή πιθανότητα. Για την ταξινόμηση, αυτό σημαίνει ότι το μοντέλο προβλέπει την πιθανότητα ενός παραδείγματος το οποίο ανήκει σε κάθε κατηγορία κλάσης.

Η πιθανότητα πολυωνυμικής κατανομής δίνεται από τη σχέση:

$$P(x_1 \dots x_k) = \begin{cases} \frac{n!}{x_1! \dots x_k!} (p_1^{x_1} \times \dots \times p_k^{x_k}), & \text{όταν } \sum_{i=1}^k x_i = n \\ 0, & \text{για κάθε άλλη περίπτωση} \end{cases} \quad (3)$$

3.3.3 Συνελικτικά Νευρωνικά Δίκτυα

Τα νευρωνικά δίκτυα αντικατοπτρίζουν τη συμπεριφορά του ανθρώπινου εγκεφάλου, επιτρέποντας σε προγράμματα υπολογιστών να αναγνωρίζουν μοτίβα και να επιλύουν προβλήματα τεχνητής νοημοσύνης, μηχανικής και βαθιάς μάθησης (Nielsen, 2019). Συγκεκριμένα, αποτελούν ένα υποσύνολο της μηχανικής μάθησης και έναν από τους σημαντικότερους αλγόριθμους εποπτευόμενης μάθησης για προβλήματα ταξινόμησης.

Στη γενική τους μορφή, τα νευρωνικά δίκτυα αποτελούνται από στρώματα κόμβων, τα οποία περιέχουν ένα **στρώμα εισόδου**, ένα ή περισσότερα **κρυμμένα στρώματα** και ένα **στρώμα εξόδου**. Μόλις προσδιοριστεί ένα στρώμα εισόδου, αντιστοιχίζονται τα βάρη τα οποία προσδιορίζουν τη σημασία οποιασδήποτε δεδομένης μεταβλητής, από τις οποίες οι μεγαλύτερες συμβάλλουν στις τιμές εξόδου. Εάν η τιμή εξόδου οποιουδήποτε μεμονωμένου κόμβου είναι μεγαλύτερη από κάποια καθορισμένη τιμή, ο κόμβος αυτός ενεργοποιείται και μεταδίδονται δεδομένα στο επόμενο στρώμα του δικτύου, διαφορετικά δεν διαβιβάζονται δεδομένα. Τα νευρωνικά δίκτυα βασίζονται σε δεδομένα εκπαίδευσης για την εκμάθηση μοντέλων και την αυτοβελτίωση της ακρίβειάς τους με την πάροδο του χρόνου. Ανάλογα με το μέγεθος και την πολυπλοκότητα της βάσης δεδομένων, η εκπαίδευση των μοντέλων μπορεί να διαρκέσει από μερικά λεπτά έως και αρκετές ώρες. Κάθε μεμονωμένος κόμβος μπορεί να εκφραστεί ως ένα μοντέλο γραμμικής παλινδρόμησης με την παρακάτω μορφή:

$$output = f(x) = \begin{cases} 1, & \text{όταν } \sum w_i x_i + b \geq 0 \\ 0, & \text{όταν } \sum w_i x_i + b < 0 \end{cases} \quad (4)$$

όπου:

- $\sum_{i=1}^m w_i x_i + bias = w_1 x_1 + w_2 x_2 + w_3 x_3 + bias$
- x_i , τα δεδομένα εισόδου
- w_i , τα βάρη
- $b = threshold$, το σημείο μετάβασης καθορισμένης τιμής

Τα συνελικτικά νευρωνικά δίκτυα (CNN) είναι ένα είδος αλγόριθμου βαθιάς μάθησης με κύρια εφαρμογή στην **ανάλυση οπτικών εικόνων**. Μπορούν να λάβουν μία εικόνα εισόδου και να εκχωρήσουν πληροφορίες σε διάφορες πτυχές της, ώστε να μπορεί να διαφοροποιηθεί η κάθε εικόνα. Μεγάλο πλεονέκτημα των συνελικτικών δικτύων είναι η χαμηλή απαιτούμενη προεπεξεργασία συγκριτικά με άλλους αλγόριθμους ταξινόμησης. Η αρχιτεκτονική των CNN είναι παρόμοια με εκείνη των απλών νευρωνικών δικτύων (Εικόνα 3.2). Λόγω της ιδιότητας τους να συλλέγουν επιτυχώς χωρικές και χρονικές εξαρτήσεις μιας εικόνας με την εφαρμογή σχετικών φίλτρων, τα καθιστά κατάλληλα για την αντιμετώπιση προβλημάτων ταξινόμησης βαθιάς μάθησης. Τα στρώματα στην αρχιτεκτονική ενός CNN επιτρέπουν στο δίκτυο να εκπαιδευτεί καλύτερα και να κατανοήσει την πολυπλοκότητα μίας εικόνας, λόγω της μείωσης του αριθμού των παραμέτρων και της επαναχρησιμοποίησης των βαρών στη βάση δεδομένων των εικόνων (Saha, 2018).

Το πρώτο στρώμα που εισάγεται στην αρχιτεκτονική του δικτύου είναι το **στρώμα συνέλιξης (convolution layer)**. Το στρώμα αυτό εξάγει λειτουργίες από την εικόνα εισόδου και δημιουργεί μητρώα στα οποία αποθηκεύονται περίπλοκα χαρακτηριστικά με μειωμένο μέγεθος εικονοκυττάρων, τα οποία ονομάζονται χάρτες χαρακτηριστικών. Μετά το στρώμα συνέλιξης ακολουθεί ένα **στρώμα ομαδοποίησης (pooling layer)**, το οποίο χρησιμοποιεί ένα φίλτρο εικονοκυττάρων για τη μείωση των συνδέσεων μεταξύ των στρωμάτων. Με τον τρόπο αυτό πραγματοποιείται μείωση του χωρικού μεγέθους των χαρτών χαρακτηριστικών, η οποία ορίζεται ως υπό δειγματοληψία. Υπάρχουν τρεις διαφορετικές κατηγορίες στρωμάτων ομαδοποίησης. Η μείωση που πραγματοποιείται και από τα δύο στρώματα έχει ως στόχο τη μείωση της υπολογιστικής ισχύος που απαιτείται για την επεξεργασία των δεδομένων.

Επιπλέον, εισάγονται δύο διαφορετικά βοηθητικά στρώματα. Το **στρώμα ισοπέδωσης (flattening layer)** μετατρέπει τις μειωμένες από τα προηγούμενα στρώματα εικόνες, σε μία γραμμική συστοιχία, η οποία απαιτείται για την τροφοδότηση και τη σύνδεση με τα νευρωνικά δίκτυα διαφορετικών στρωμάτων νευρώνων που ακολουθούν. Τέλος, κάθε στρώμα νευρώνων ακολουθεί ένα **στρώμα απόσυρσης (dropout layer)** για την εγκατάλειψη της διαδικασίας εκπαίδευσης κάποιου ποσοστού νευρώνων, ώστε να αποφευχθεί η υπερφόρτωση του μοντέλου και τα σφάλματα μοντελοποίησης.

Όλα τα παραπάνω στρώματα εκτός του τελευταίου, χρησιμοποιούν τη **συνάρτηση ενεργοποίησης ReLU** (Rectified Linear Unit). Η συνάρτηση διορθωμένης γραμμικής μονάδας ReLU είναι μία γραμμική συνάρτηση, η οποία εξάγει τις τιμές του μητρώου που εισάγεται απευθείας όταν αυτές είναι θετικές, διαφορετικά τις εξάγει ως μηδέν (Yang, 2017). Χρησιμοποιείται σε πολλούς τύπους νευρωνικών δικτύων, καθώς βελτιώνει την εκπαίδευση του μοντέλου και επιτυγχάνει καλύτερη απόδοση. Η συνάρτηση ReLU έχει την παρακάτω μορφή:

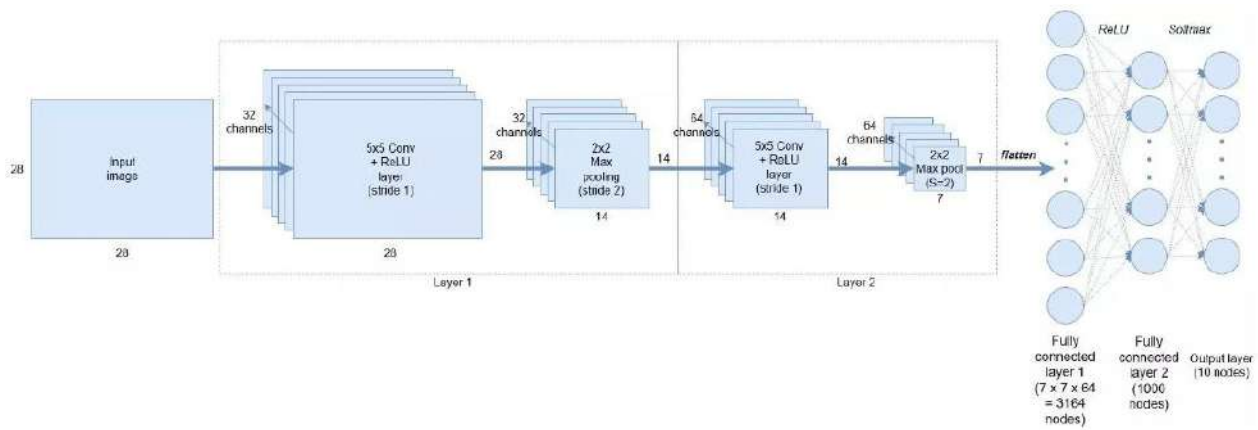
$$output = f(x) = \max(0, x) \quad (5)$$

Το τελευταίο στρώμα στο οποίο χρησιμοποιείται ο ίδιος αριθμός νευρώνων με αυτόν των κλάσεων ταξινόμησης του προβλήματος, χρησιμοποιεί τη **συνάρτηση softmax**. Η συνάρτηση αυτή είναι μία μορφή λογιστικής παλινδρόμησης, η οποία μετατρέπει τις τιμές εισόδου, σε τιμές μεταξύ 0 και 1 ώστε να ερμηνευτούν ως πιθανότητες (Yang, 2017) και είναι της μορφής:

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad (6)$$

όπου:

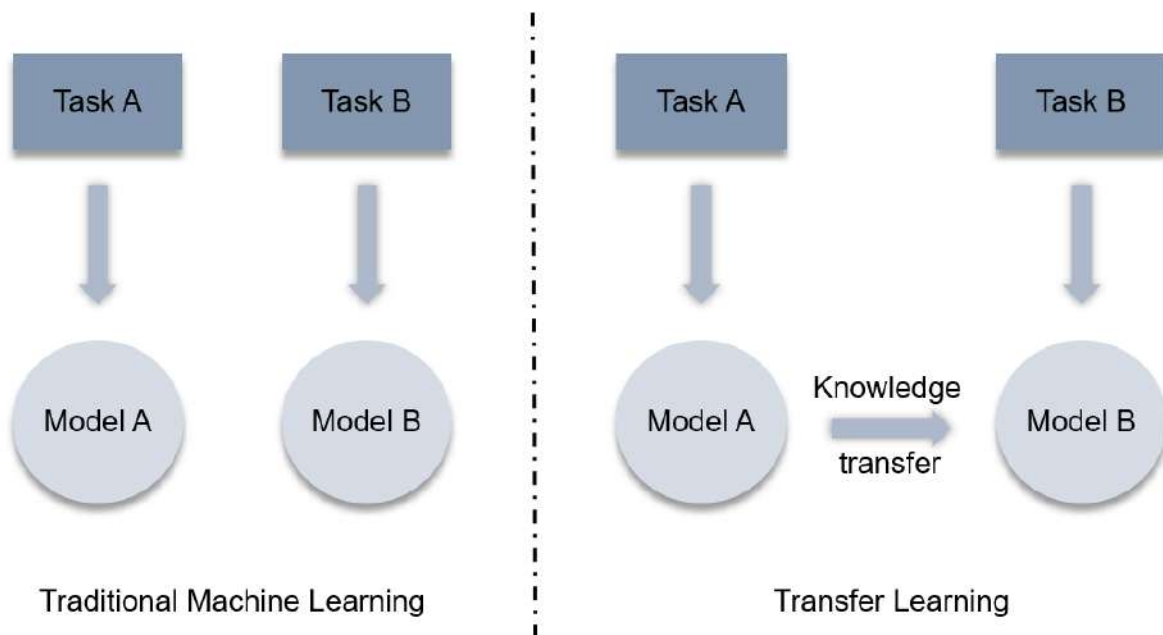
- $\vec{z}$, το διάνυσμα εισαγωγής
- K , ο αριθμός κλάσεων ταξινομητή πολλαπλών κλάσεων
- e^{z_i} , η τυπική εκθετική συνάρτηση για το διάνυσμα εισόδου
- e^{z_j} , η τυπική εκθετική συνάρτηση για το διάνυσμα εξόδου



Εικόνα 3.2: Αρχιτεκτονική συνελικτικού νευρωνικού δικτύου (Πηγή: Thomas, 2017)

3.3.4 Μεταφορά Μάθησης

Η μεταφορά μάθησης αξιοποιεί τις γνώσεις και τις δυνατότητες αναγνώρισης προτύπων που αναπτύχθηκαν με βάση ένα άλλο πρόβλημα, για τη διευκόλυνση της μαθησιακής διαδικασίας ενός διαφορετικού αλλά σχετικού προβλήματος (Διάγραμμα 3.3). Συγκεκριμένα, πρόκειται για μία δημοφιλή προσέγγιση στη βαθιά μάθηση, όπου **ήδη εκπαιδευμένα μοντέλα χρησιμοποιούνται ως σημείο εκκίνησης** για τις εργασίες επεξεργασίας άλλων μοντέλων, με δεδομένα τα οποία απαιτούν τεράστιους πόρους υπολογισμού και χρόνου για την ανάπτυξη μοντέλων νευρωνικών δικτύων (Brownlee, 2017). Η εφαρμογή της μεταφοράς μάθησης έχει ως κύριο στόχο την **επιτάχυνση εκπαίδευσης και τη βελτίωση απόδοσης** ενός μοντέλου βαθιάς μάθησης. Εμφανίζει καλύτερη λειτουργία μόνο στη βαθιά μάθηση εάν οι λειτουργίες του μοντέλου οι οποίες μαθαίνονται και αποθηκεύονται από την πρώτη εργασία είναι γενικές καθώς μεταφέρεται η αρχιτεκτονική και οι παράμετροί του.



Διάγραμμα 3.3: Παραδοσιακή μηχανική μάθηση και μεταφορά μάθησης

Η μεταφορά μάθησης μπορεί να αξιοποιηθεί για την αντιμετώπιση προβλημάτων ταξινόμησης με δύο κοινές προσεγγίσεις: Προ-εκπαιδευμένο μοντέλο προσέγγισης και Ανάπτυξη νέου μοντέλου προσέγγισης.

Προ-εκπαιδευμένα μοντέλα (pre-trained models)

Επιλογή μοντέλου πηγής: επιλέγεται ένα προ-εκπαιδευμένο μοντέλο προέλευσης από μία ποικιλία ήδη διαθέσιμων μοντέλων εταιρειών λογισμικών και ερευνητικών ιδρυμάτων, για χρήση σε μεγάλα και απαιτητικά σύνολα δεδομένων.

Επαναχρησιμοποίηση μοντέλου: το μοντέλο χρησιμοποιείται ως σημείο εκκίνησης για ένα νέο μοντέλο εκπαίδευσης του προβλήματος προς αντιμετώπιση, από το οποίο μπορεί να χρησιμοποιηθεί ένα τμήμα του ή και ολόκληρο το μοντέλο, ανάλογα με την τεχνική μοντελοποίησης που πραγματοποιείται.

Συντονισμός μοντέλου: μπορεί να απαιτηθεί προαιρετική προσαρμογή του μοντέλου ή βελτίωση των δεδομένων εισόδου και εξόδου τα οποία είναι διαθέσιμα για το πρόβλημα προς αντιμετώπιση.

Ανάπτυξη μοντέλων (model development)

Επιλογή εργασίας πηγής: επιλέγεται ένα σχετικό πρόβλημα πρόβλεψης μοντελοποίησης με αφθονία δεδομένων, το οποίο εμφανίζει σχέση μεταξύ των δεδομένων εισόδου, εξόδου και των προβλέψεων.

Ανάπτυξη πηγής μοντέλου: αναπτύσσεται ένα επιδέξιο μοντέλο για μία πρώτη εργασία, από το οποίο μπορεί να πραγματοποιηθεί κάποια εκμάθηση χαρακτηριστικών.

Επαναχρησιμοποίηση μοντέλου: το μοντέλο το οποίο ταιριάζει στην εργασία προέλευσης χρησιμοποιείται ως σημείο εκκίνησης για ένα νέο μοντέλο εκπαίδευσης του προβλήματος προς αντιμετώπιση, από το οποίο μπορεί να χρησιμοποιηθεί ένα τμήμα του ή και ολόκληρο το μοντέλο, αναλόγως της τεχνικής μοντελοποίησης που πραγματοποιείται.

Συντονισμός μοντέλου: μπορεί να απαιτηθεί προαιρετική προσαρμογή του μοντέλου ή βελτίωση των δεδομένων εισόδου και εξόδου τα οποία είναι διαθέσιμα για το πρόβλημα προς αντιμετώπιση.

4. ΑΝΑΛΥΣΗ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ

4.1 ΕΙΣΑΓΩΓΗ

Στο συγκεκριμένο κεφάλαιο παρουσιάζεται η **συλλογή και επεξεργασία** των πραγματικών δεδομένων πειράματος βιντεοσκόπησης, η κατασκευή των βάσεων δεδομένων που χρησιμοποιήθηκαν για τους αλγόριθμους ταξινόμησης, καθώς και τα **αποτελέσματα** που προέκυψαν για κάθε εξεταζόμενη περιοχή.

Η χρήση **Συστημάτων μη Επανδρωμένων Αεροσκαφών** έχει προταθεί ως ένα πρωτοποριακό και αποτελεσματικό εργαλείο των ευφυών συστημάτων μεταφορών, για την οπτικοποίηση και τη συλλογή πληροφοριών κίνησης μεγάλων δικτύων τα οποία δε μπορούν να καλυφθούν από σταθερές κάμερες, καθώς και για την ανάλυση διαφόρων κυκλοφοριακών φαινομένων (Barmounakis, et al., 2016). Τα ΣμηΕΑ αποτελούν μία νέα και οικονομικά αποδοτική λύση συλλογής μαζικών δεδομένων τροχιών οχημάτων συγκριτικά με την παλιά προσέγγιση χρήσης ήδη εγκατεστημένων καμερών, καθώς προσφέρουν δεδομένα καλύτερης ανάλυσης λόγω της σταθερής πτήσης τους. Η χρήση ΣμηΕΑ σε αστικές περιοχές αποτελεί έναν ιδανικό τρόπο παρακολούθησης και ανάλυσης της κυκλοφορίας, με την εφαρμογή τους να παρουσιάζει όλο και μεγαλύτερο ενδιαφέρον στις κυκλοφοριακές έρευνες. Ωστόσο, τα περισσότερα πειράματα είναι μικρής κλίμακας (χρήση ενός drone και κάλυψη μικρής περιοχής), με αποτέλεσμα την έλλειψη ανάλυσης συγκεκριμένων κυκλοφοριακών παραγόντων σε βάθος.

4.2 ΣΥΛΛΟΓΗ ΔΕΔΟΜΕΝΩΝ

Τα δεδομένα για την παρούσα Διπλωματική Εργασία λήφθηκαν από πείραμα βιντεοσκόπησης, το οποίο πραγματοποιήθηκε στο κέντρο της Αθήνας με χρήση Συστημάτων μη Επανδρωμένων Αεροσκαφών (Barmounakis, et al., 2019). Σκοπός του πειράματος ήταν η **καταγραφή κυκλοφοριακών ροών σε αστικό περιβάλλον** και η παροχή μίας ολοκληρωμένης εικόνας για τον τρόπο με τον οποίο τα ιδιαίτερα χαρακτηριστικά των ΣμηΕΑ μπορούν να ξεπεράσουν υφιστάμενους περιορισμούς στην καταγραφή της κυκλοφορίας.

Το πείραμα αποτέλεσε το πρώτο μεγάλης κλίμακας του είδους του, με στόχο την καταγραφή ροών κυκλοφορίας σε αστικό περιβάλλον. Συγκεκριμένα, χρησιμοποιήθηκε στόλος 10 drone σε μία ευρύτερη περιοχή του κέντρου της Αθήνας καθώς αποτελεί ένα πολυάσχολο κυκλοφοριακό περιβάλλον στο οποίο πραγματοποιούνται μετακινήσεις διαφόρων μέσων και μπορεί να επιτρέψει την παρακολούθηση διαφόρων κυκλοφοριακών φαινομένων. Το πείραμα πραγματοποιήθηκε κατά τη διάρκεια της πρωινής αιχμής, καθώς βάσει της ισχύουσας νομοθεσίας απαγορεύονται οι πτήσεις ΣμηΕΑ κατά τις νυχτερινές ώρες. Η συλλογή των δεδομένων, λόγω περιορισμού αυτονομίας των αεροσκαφών και χρόνου πτήσης, διεξήχθη σε φάσεις διάρκειας 30 λεπτών. Η τοποθέτηση των 10 drone, καθώς και τα δύο σημεία ελέγχου απογείωσης και προσγείωσης εμφανίζονται στην Εικόνα 4.1.



Εικόνα 4.1: Σημεία τοποθέτησης ΣμηΕΑ (Πηγή: Barmounakis, et al., 2019)

Τα **δεδομένα** που χρησιμοποιήθηκαν για την εκπαίδευση των μοντέλων της παρούσας εργασίας λήφθηκαν από τις καταγραφές της κυκλοφορίας στη **διασταύρωση Πατησίων και Λεωφόρου Αλεξάνδρας** στα σημεία 7 και 8 (Εικόνες 4.2 και 4.3) και στη **Πλατεία Ομονοίας** στο σημείο 10 (Εικόνα 4.4). Επιλέχθηκαν τα συγκεκριμένα σημεία καταγραφής καθώς αποτελούν περιοχές εμφάνισης του υψηλότερου κυκλοφοριακού φόρτου και διαφορετικών κινήσεων οχημάτων και συνεπώς διαφορετικών ειδών κυκλοφοριακών ουρών, συγκριτικά με τα υπόλοιπα σημεία καταγραφής διασταυρώσεων και οδικών τμημάτων μικρότερου φόρτου. Ειδικότερα, τα drone στα σημεία 7 και 8 καταγράφουν τις ίδιες τροχιές οχημάτων στη διασταύρωση, όμως από διαφορετική οπτική γωνία καθώς στο σημείο 7 καταγράφεται κυρίως μεγάλο τμήμα της οδού Πατησίων, ενώ αντίστοιχα στο σημείο 8 της Λεωφόρου Αλεξάνδρας. Με τον τρόπο αυτό επιδιώκεται η αναγνώριση και ταξινόμηση των ίδιων κυκλοφοριακών ουρών που σχηματίζονται από ανεστραμμένους άξονες κατά 90 μοίρες.

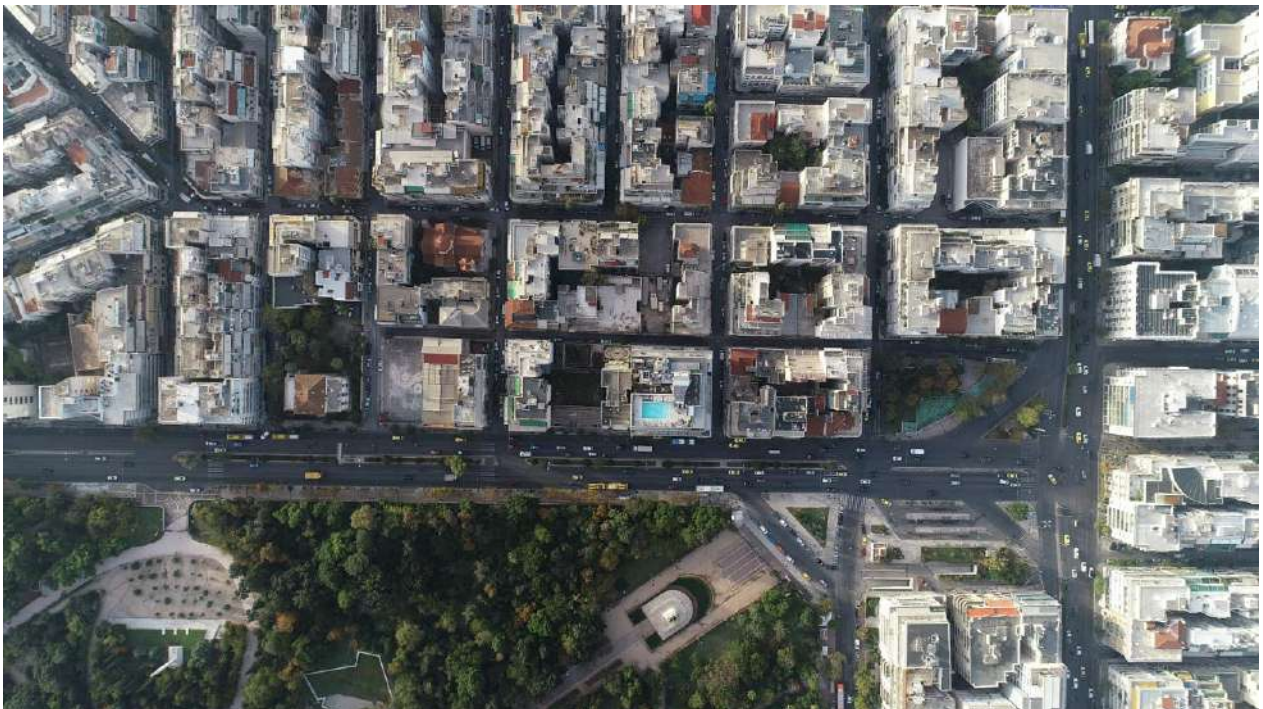
Τα ΣμηΕΑ αποτελούν έναν αποτελεσματικό και οικονομικό τρόπο **ανάλυσης μικροσκοπικών χαρακτηριστικών της κυκλοφορίας**, όπου με σωστή βαθμονόμηση και σταθεροποίηση του βίντεο μπορούν να εξαχθούν αποτελέσματα μεγάλης ακρίβειας με μέγιστο σφάλμα πρόβλεψης ταχύτητας μικρότερο του 1.0 km/h (Barmounakis, et al., 2016). Επιπλέον, η ταχύτητα των οχημάτων και το υψόμετρο των αεροσκαφών τα οποία μεταβάλλονται κατά τη διάρκεια πτήσης δεν επηρεάζουν την ακρίβεια των δεδομένων καθώς τα βίντεο καταγράφηκαν σε υψηλή ανάλυση 4K (4096x2160) στα **25 καρέ (frames) ανά δευτερόλεπτο** από DJI Drones.

4.3 ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ

Τα δεδομένα ήταν αρχικά διαχωρισμένα σε δύο αρχεία m4v και ένα mp4, ένα για κάθε σημείο καταγραφής. Τα βίντεο των σημείων 7 και 8 έχουν διάρκεια 5:00 και 5:27 λεπτών αντίστοιχα, ενώ του σημείου 10 έχει διάρκεια 15:18 λεπτών. Παρακάτω παρατίθενται οι περιοχές κάλυψης των αεροσκαφών τοποθετημένα στα επιλεγμένα σημεία.



Εικόνα 4.2: Οδός Πατησίων - Λεωφόρος Αλεξάνδρας (σημείο 7) - Video 1



Εικόνα 4.3: Λεωφόρος Αλεξάνδρας - Οδός Πατησίων (σημείο 8) - Video 2



Εικόνα 4.4: Πλατεία Ομονοίας (σημείο 10) - Video 3

Προκαταρκτική επεξεργασία

Για την σύγκριση συλλεγμένων δεδομένων ίδιας διάρκειας, το βίντεο 3 χωρίστηκε σε δύο διαφορετικά αρχεία διάρκειας 5:00 λεπτών (Video3.1, Video3.2) και ένα διάρκειας 5:18 λεπτών (Video3.3), ώστε τα μοντέλα να εκπαιδευτούν με βάσεις δεδομένων σχεδόν ίδιου πλήθους καρέ εικόνων. Για την επεξεργασία των δεδομένων, τα βίντεο εισήχθησαν στην πλατφόρμα **Supervisely**, η οποία επιτρέπει σε μεμονωμένους ερευνητές ή και μεγάλες ομάδες ερευνητών, να επισημάνουν εικόνες και βίντεο, να πειραματιστούν με σύνολα δεδομένων και να εργαστούν με νευρωνικά δίκτυα. Συγκεκριμένα, χρησιμοποιήθηκε το εργαλείο επισήμανσης για την **τοποθέτηση ετικετών σε κάθε καρέ εικόνας**, τα οποία εξήχθησαν αυτόματα από την πλατφόρμα.

Αρχικά, ορίστηκαν οι κατηγορίες ουρών προς ανίχνευση καθώς και οι παράγοντες με τους οποίους μία συμφόρηση οχημάτων ορίζεται ως κυκλοφοριακή ουρά. Συγκεκριμένα, οι παράγοντες για την επιτυχή ανίχνευση ουράς ήταν ο εντοπισμός συμφόρησης στην οποία συμμετείχαν **τουλάχιστον 3 οχήματα** και είχε **διάρκεια τουλάχιστον 5 δευτερολέπτων**. Αναλυτικά, επιλέχθηκε να εντοπιστούν οι παρακάτω κατηγορίες ουρών που σχηματίζονται:

- ♦ Στάση οχήματος (επιβατικό Ι.Χ., ταξί, λεωφορείο)
- ♦ Δεξιά στροφή προπορευόμενων οχημάτων
- ♦ Αριστερή στροφή χωρίς ξεχωριστή λωρίδα κίνησης
- ♦ Σύμπτυξη λωρίδων
- ♦ Υποχρεωτική αλλαγή λωρίδων
- ♦ Κατειλημμένη διασταύρωση (σταματημένα οχήματα στη διασταύρωση)
- ♦ Επίδραση διαρροής (κορεσμός τμήματος και εισροή οχημάτων στο προηγούμενο τμήμα)

Συνολικά καταγράφηκαν 88 κυκλοφοριακές ουρές (Πίνακας 4.1), όπου για την κάθε ουρά σημειώθηκε το είδος της, η προέλευση και ο προορισμός των εμπλεκόμενων οχημάτων, η χρονική στιγμή σχηματισμού και διάλυσης και η συνολική διάρκειά της.

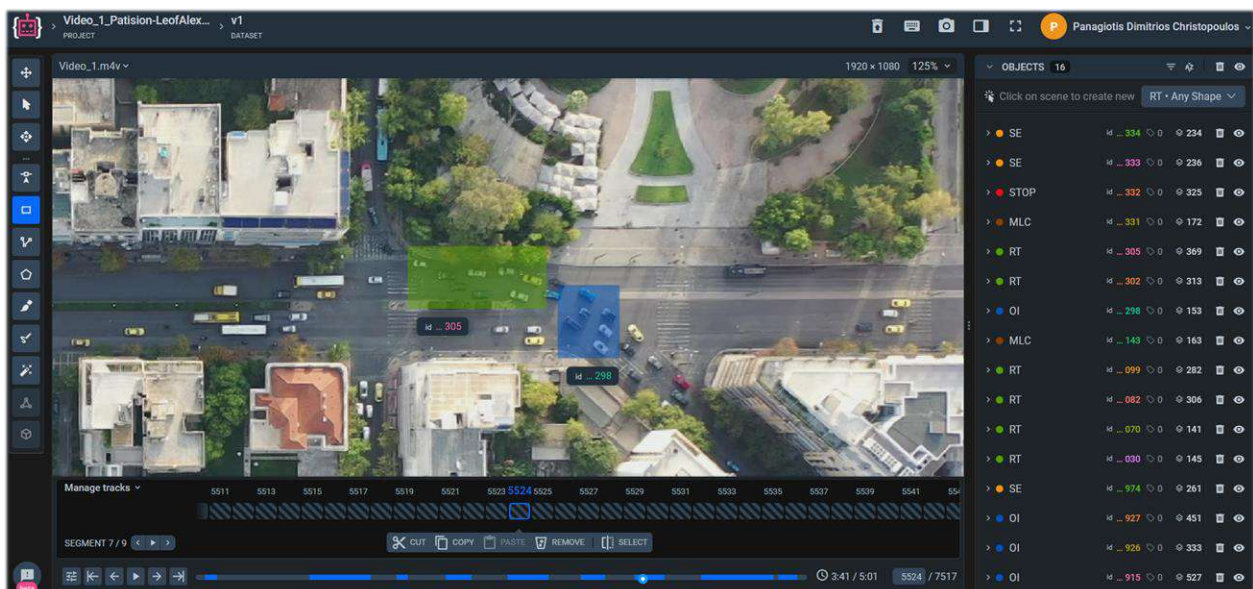
Πίνακας 4.1: Κλάσεις κυκλοφοριακών ουρών και αντίστοιχο πλήθος

Stop	11
Right Turn	21
Left Turn	9
Lane Contraction	5
Mandatory Lane Change	11
Occupied Intersection	20
Spillover Effect	11
TOTAL	88

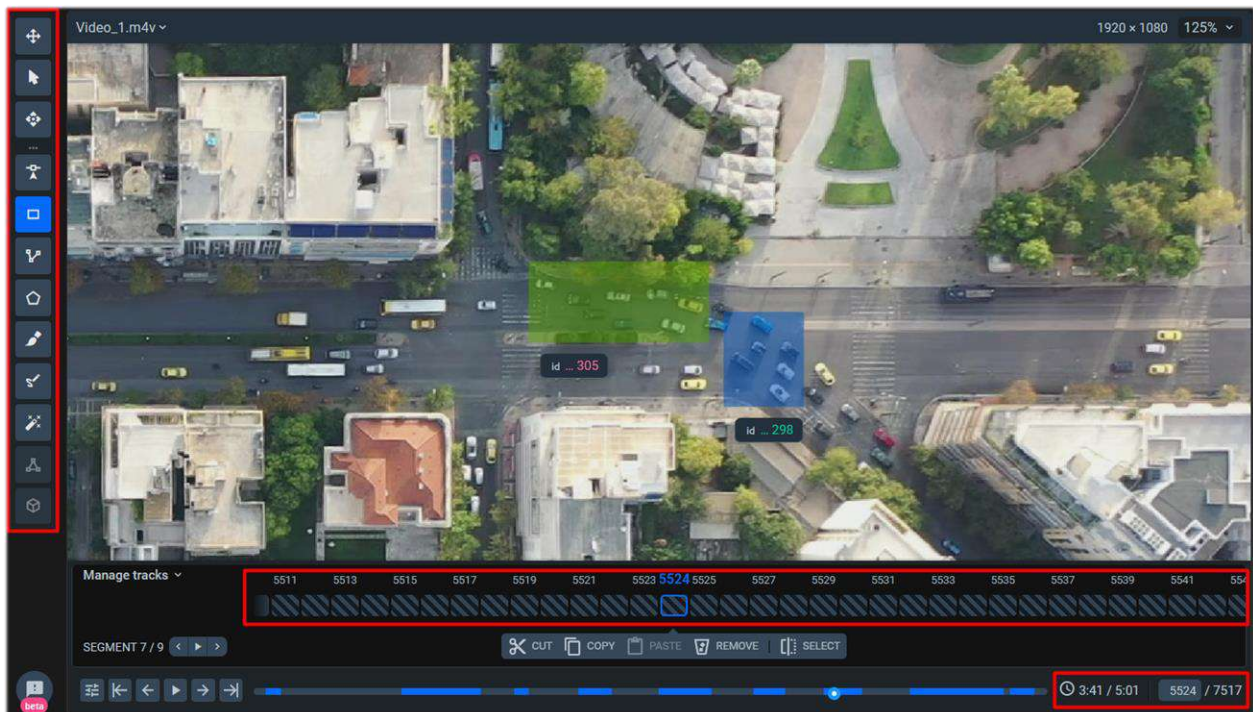
Επισημάνση δεδομένων

Στη συνέχεια, πραγματοποιήθηκε η επισημάνση των δεδομένων με το εργαλείο της πλατφόρμας. Επιλέχθηκε η χρήση ορθογωνίου παραλληλόγραμμου (rectangle) για την αναπαράσταση της ουράς και του μεγέθους της καθώς όλες οι πληροφορίες ενσωματώνονται στο απλό σχήμα του καθιστώντας το εύκολο να εκπαιδευτεί και απλό να ενσωματωθεί σε συστήματα που απαιτούν ένα στοιχείο ανίχνευσης.

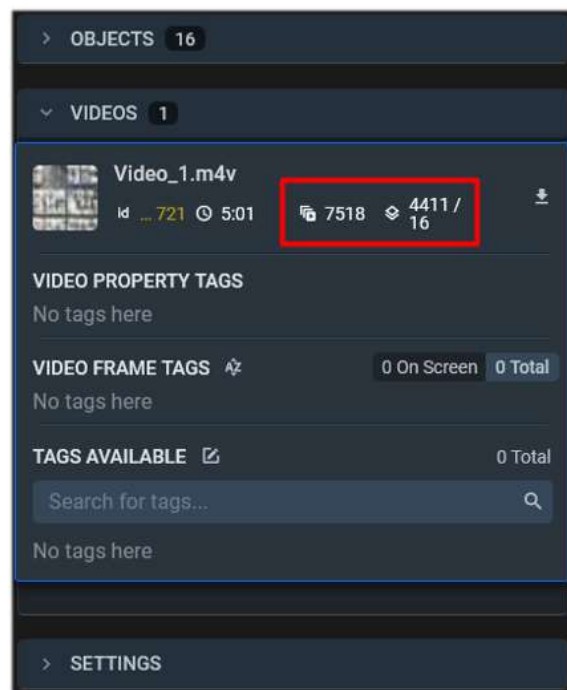
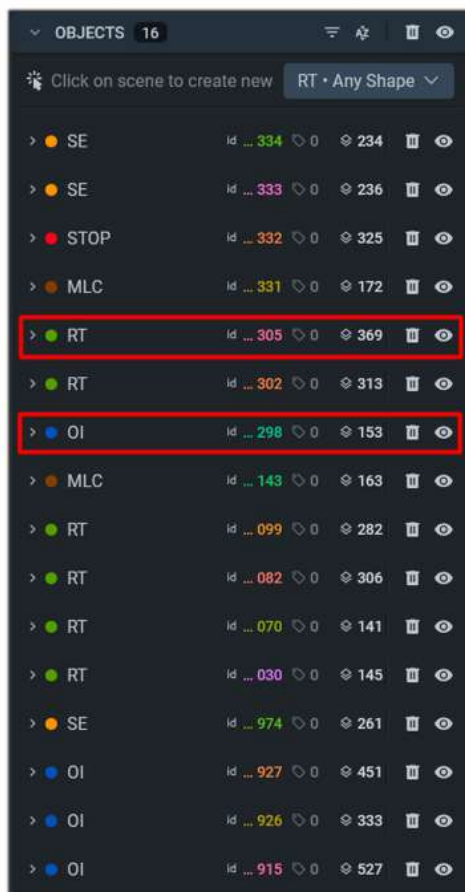
Από τις καταγεγραμμένες χρονικές διάρκειες, εντοπίστηκαν στο Supervisely τα καρέ εικόνων τα οποία αντιστοιχούσαν στα δευτερόλεπτα εμφάνισης κυκλοφοριακής ουράς, καθένα από τα οποία επισημάνθηκε με ετικέτα, αντίστοιχη του είδους ουράς. Η διαδικασία πραγματοποιήθηκε ξεχωριστά για τα πέντε βίντεο και προέκυψαν τα επισημασμένα δεδομένα για χρήση στα μοντέλα εκπαίδευσης, τα οποία εξήχθησαν από την πλατφόρμα για περαιτέρω επεξεργασία και τη δημιουργία βάσεων δεδομένων. Στις Εικόνες 4.5 έως 4.7 παρουσιάζεται το περιβάλλον και τα εργαλεία της πλατφόρμας.



Εικόνα 4.5: Περιβάλλον πλατφόρμας Supervisely



Εικόνα 4.6: Εργαλείο επισήμανσης ετικετών, λίστα καρέ εικόνων βίντεο, διάρκεια βίντεο και αριθμός εικόνων

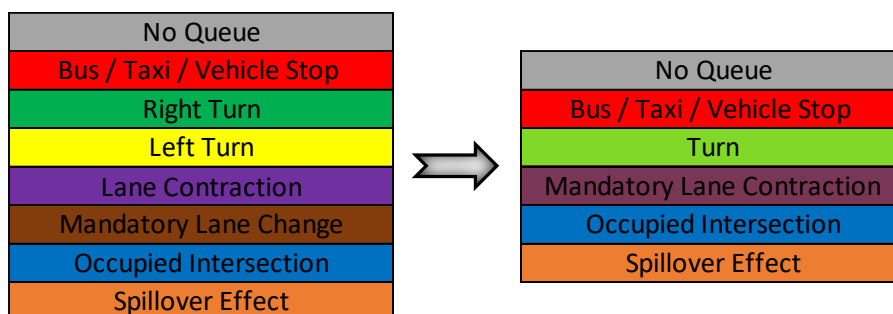


Εικόνα 4.7: (αριστερά) Λίστα κλάσεων και αντίστοιχος αριθμός εικόνων, (δεξιά) Στοιχεία βίντεο και συνολικός αριθμός κλάσεων και επισημασμένων εικόνων

Τελική επεξεργασία

Τα επισημασμένα δεδομένα εισήχθησαν σε πίνακες στο πρόγραμμα Microsoft Excel, όπου για κάθε κυκλοφοριακή ουρά αναγράφονταν τα καρέ εικόνων στα οποία εμφανίζεται. Για την συμπλήρωση των κενών εικόνων στις οποίες δεν παρατηρήθηκε σχηματισμός ουράς, κατασκευάστηκε **πλασματική κλάση** με ονομασία «No Queue». Καθώς συγκεκριμένες ουρές δεν εμφανίστηκαν σε όλα τα βίντεο, για την χρήση ισοβαρών βάσεων δεδομένων πραγματοποιήθηκε **σύμπτυξη κλάσεων** οι οποίες εμφάνιζαν συσχέτιση ή παρόμοια χαρακτηριστικά. Η σύμπτυξη πραγματοποιήθηκε για τη χρήση δεδομένων μεγαλύτερου πλήθους και συνοχής, καθώς και την αύξηση της ακρίβειας των μοντέλων και τη μείωση του υπολογιστικού χρόνου. Συγκεκριμένα, οι ουρές που σχηματίστηκαν λόγω δεξιάς και αριστερής στροφής συνδυάστηκαν σε μία κλάση που αφορά όλες τις στροφές, ενώ οι ουρές λόγω σύμπτυξης και υποχρεωτικής αλλαγής λωρίδων συνδυάστηκαν σε μία κλάση υποχρεωτικής σύμπτυξης λωρίδων. Οι έξι τελικές κλάσεις που προέκυψαν από τη σύμπτυξη παρουσιάζονται στον Πίνακα 4.2.

Πίνακας 4.2: Κλάσεις κυκλοφοριακών ουρών μοντέλων



Τέλος, για κάθε βίντεο κατασκευάστηκε πίνακας στον οποίο τοποθετήθηκαν οι παραπάνω κλάσεις με **σειρά εμφάνισης κάθε ουράς στα καρέ εικόνων**. Σημειώνεται ότι στις περιπτώσεις όπου σε κάποιες εικόνες εμφανιζόταν δύο ή και παραπάνω ουρές ταυτόχρονα, στην λίστα τοποθετήθηκε πρώτη εκείνη η οποία είχε εμφανιστεί ήδη σε προηγούμενα καρέ. Με τον τρόπο αυτό η διαχείριση των δεδομένων για τη μετατροπή τους σε μία βάση επισημάνσεων, ήταν αρκετά πιο εύκολη καθώς τα δεδομένα ήταν καλά διαχωρισμένα.

4.4 ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ

Από τους τελικούς πίνακες κατασκευάστηκαν οι βάσεις επισημασμένων δεδομένων σε αρχεία csv, ώστε να είναι εύκολα αναγνωρίσιμα από τον αλγόριθμο. Οι βάσεις αποτελούν πίνακα μίας στήλης, η οποία περιέχει τις **κωδικές ονομασίες** των κλάσεων με τη σειρά των ουρών όπως αυτή καταγράφηκε στους τελικούς πίνακες. Επομένως, τα μητρώα επισημάνσεων αποτελούνταν από 1 στήλη και σειρές όσες και ο αριθμός των καρέ εικόνων του κάθε βίντεο. Στη γλώσσα προγραμματισμού της Python η αρίθμηση ξεκινάει από το μηδέν. Επομένως, οι κατηγορίες αντιστοιχήθηκαν στους αριθμούς μηδέν έως πέντε με την παρακάτω σειρά:

- ♦ **(0):** *No Queue*
- ♦ **(1):** *Stop*
- ♦ **(2):** *Turn*
- ♦ **(3):** *Mandatory Lane Change*
- ♦ **(4):** *Occupied Intersection*
- ♦ **(5):** *Spillover Effect*

Για την ολοκλήρωση της συνολικής βάσης δεδομένων (καρέ εικόνων και αντίστοιχες επισημάνσεις), απαιτήθηκε **αποκοπή των βίντεο** που λήφθηκαν από το πείραμα στα οποία παρατηρήθηκε ο μέγιστος κυκλοφοριακός φόρτος στις επιλεγμένες περιοχές. Η διαδικασία πραγματοποιήθηκε με κώδικα γλώσσας προγραμματισμού **Python στο περιβάλλον PyCharm** με εξαγωγή 25 καρέ ανά δευτερόλεπτο, ώστε η ψηφιακή βάση εικόνων να αντιστοιχεί στον ίδιο αριθμό καρέ με αυτόν που πραγματοποίησε αυτόματα η πλατφόρμα Supervisely κατά την εισαγωγή κάθε βίντεο. Σημειώνεται ότι η αρίθμηση των εικόνων ξεκίνησε από το μηδέν όπως και των δεδομένων επισημάνσης λόγω χρήσης της γλώσσας Python. Παρακάτω παρατίθενται τα καρέ εικόνων τα οποία εξήχθησαν για κάθε βίντεο και ο συνολικό αριθμός αυτών στα οποία καταγράφηκε εμφάνιση ουράς:

- ♦ **Video 1:** 7518 (4411)
- ♦ **Video 2:** 8180 (4444)
- ♦ **Video 3.1:** 7500 (8545)
- ♦ **Video 3.2:** 7500 (4985)
- ♦ **Video 3.3:** 7970 (5057)

Έπειτα, πραγματοποιήθηκε μετατροπή των εικόνων σε **λίστες δεδομένων εκπαίδευσης** μορφής πινάκων με κατασκευή νέου κώδικα, ώστε να αναγνωρίζονται τα χαρακτηριστικά κάθε εικόνας από τον αλγόριθμο. Για τη διατήρηση της πολυπλοκότητας των δεδομένων, επιλέχθηκε η διατήρηση των έγχρωμων χαρακτηριστικών με τη χρήση της **κλίμακας χρωμάτων RGB (Red - Green - Blue)**, αντί της κλίμακας του γκρι με την οποία τα δεδομένα θα είχαν μειωμένο μέγεθος, αλλά και μικρότερη δυσκολία στην αναγνώριση συμφορήσεων χωρίς την εμφάνιση πολλαπλών αποχρώσεων.

Επιπλέον, καθώς οι εικόνες ήταν υψηλής ανάλυσης, πραγματοποιήθηκε **αλλαγή των διαστάσεων** για τη διευκόλυνση της εισαγωγής τους στον αλγόριθμο, σε μέγεθος ώστε να μην επηρεάζεται σημαντικά η ευκρίνειά τους. Έπειτα από αρκετές δοκιμές, επιλέχθηκε η ανάλυση των **256 x 256 pixels** με την οποία επιτεύχθηκε ταυτοχρόνως μείωση μεγέθους του αρχείου και υπολογιστικού χρόνου εισαγωγής. Η αρίθμηση των εικονοκυττάρων ξεκινάει από πάνω αριστερά και καταλήγει κάτω δεξιά. Τα μητρώα καρέ εικόνων αποτελούνταν από 3 στήλες και 256 σειρές, με τις τιμές της κλίμακας χρωμάτων κόκκινου - πράσινου - μπλε και του αριθμού των pixel των εικόνων αντίστοιχα. Το κάθε κελί μίας σειράς αντιστοιχεί στους συνδυασμούς χρωμάτων που περιέχει κάθε pixel.

Από τις παραπάνω τροποποιήσεις προέκυψαν τα μητρώα καρέ εικόνων και επισημάνσεων, ο συνδυασμός των οποίων αποτελεί τη βάση δεδομένων του αλγόριθμου κάθε βίντεο (Πίνακες 4.3 και 4.4).

Πίνακας 4.3: Μητρώο καρέ εικόνων βάσης δεδομένων (μεταβλητή x)

	0	1	2
20	192	211	218
21	207	222	225
22	168	170	178
23	161	166	172
24	163	169	174
25	158	166	173
26	151	155	160
27	115	140	135
28	104	127	134
29	115	119	146
30	233	242	253
31	160	170	185
32	199	208	217
33	201	208	211
34	128	142	139
35	118	150	150
36	193	213	224
37	200	212	218
38	181	189	196
39	198	196	195
40	188	182	183

Πίνακας 4.4: Μητρώο επισημάνσεων βάσης δεδομένων (μεταβλητή y)

	Class
2460	0
2461	0
2462	0
2463	0
2464	0
2465	0
2466	0
2467	0
2468	0
2469	0
2470	2
2471	2
2472	2
2473	2
2474	2
2475	2
2476	2
2477	2
2478	2
2479	2
2480	2

4.5 ΑΝΑΛΥΣΗ

Έπειτα από τη συλλογή και την επεξεργασία των δεδομένων και τη δημιουργία των βάσεων δεδομένων όπως περιεγράφηκε στα προηγούμενα υποκεφάλαια, ακολούθησε η εφαρμογή της μεθόδου του συνελικτικού νευρωνικού δικτύου βαθιάς μάθησης και ο έλεγχος χρήσης της μεταφοράς μάθησης στα αποτελέσματα των αλγόριθμων της Διπλωματικής Εργασίας. Για τη διαμόρφωση των αλγόριθμων του προβλήματος ταξινόμησης πολλαπλών κλάσεων, ακολουθείται μία κλασσική διαδικασία προετοιμασίας, εκπαίδευσης και αξιολόγησης των δεδομένων και των αποτελεσμάτων τους, η οποία χρησιμοποιείται για χρήση νευρωνικών δικτύων με εποπτευόμενη μάθηση.

Εγκατάσταση πακέτων

Αρχικά, πραγματοποιήθηκε η **εγκατάσταση των πακέτων** στο περιβάλλον της PyCharm, στα οποία περιέχονται οι βιβλιοθήκες με τις εντολές τις οποίες χρησιμοποιεί ο αλγόριθμος και η **εισαγωγή των ψηφιακών βιβλιοθηκών** στον κώδικά του. Συγκεκριμένα, χρησιμοποιήθηκαν βιβλιοθήκες για την ανάγνωση των δεδομένων και την τροποποίησή τους, τη σύνταξη του μοντέλου νευρωνικού δικτύου και την εκπαίδευση του, την αξιολόγηση του αλγόριθμου και των μετρικών τιμών του, καθώς και για την κατασκευή διαγραμμάτων και άλλων σημαντικών διαδικασιών.

Προετοιμασία δεδομένων

Για να χρησιμοποιηθούν τα δεδομένα για την εκπαίδευση του μοντέλου νευρωνικού δικτύου πραγματοποιήθηκε μία διαδικασία προετοιμασίας. Τα μητρώα της βάσης εισήχθησαν με συγκεκριμένες εντολές για τον τύπο του κάθε αρχείου, ως μεταβλητές x και y αντίστοιχα στον αλγόριθμο. Ακολούθησε **τυχαίος διαχωρισμός δειγμάτων** με ποσοστό 75% για την εκπαίδευση του μοντέλου (train) και 25% για τις δοκιμές και την αξιολόγησή του (test). Ο τυχαίος διαχωρισμός μπορεί κάθε φορά που τρέχει ο αλγόριθμος, να εξάγει δείγματα διαφορετικών ποσοστών κατηγοριών ουρών, χωρίς όμως με σημαντική διαφοροποίηση καθώς τα δεδομένα τροποποιήθηκαν ώστε να εμφανίζονται όλες οι κλάσεις σε ικανοποιητικό πλήθος εικόνων. Από τον διαχωρισμό προέκυψαν τα μητρώα x_train , x_test , y_train , y_test για τα δεδομένα εικόνων και επισημάνσεων αντίστοιχα.

Έπειτα, τυπώθηκαν στην κονσόλα του περιβάλλοντος το είδος και το μέγεθος των δειγμάτων για τον έλεγχο της συνάφειας των μητρώων (Εικόνα 4.8). Η μορφή ndarray αφορά ένα πολυδιάστατο μητρώο αντικειμένων σταθερού μεγέθους και ίδιου τύπου, ενώ η μορφή DataFrame αφορά ένα μητρώο αποθήκευσης δεδομένων για ανάλυση. Τα μεγέθη των δειγμάτων για το πρώτο βίντεο αναλύονται ως εξής: τα x_train και x_test παρουσιάζουν 4-διάστατους πίνακες που αποτελούνται από 5638 και 1880 σειρές στοιχείων (εικόνες) αντίστοιχα, διαστάσεων 256 επί 256 (pixels) και κλίμακας χρωμάτων με βάθος 3 (RGB), ενώ τα y_train και y_test παρουσιάζουν 2-διάστατους πίνακες που αποτελούνται από τον ίδιο αριθμό σειρών στοιχείων (επισημάνσεις) όπως τα αντίστοιχα x για εκπαίδευση και δοκιμές και μία στήλη.

```
x_train data type: <class 'numpy.ndarray'>
y_train data type: <class 'pandas.core.frame.DataFrame'>
x_test data type: <class 'numpy.ndarray'>
y_test data type: <class 'pandas.core.frame.DataFrame'>
```

```
x_train shape: (5638, 256, 256, 3)
y_train shape: (5638, 1)
x_test shape: (1880, 256, 256, 3)
y_test shape: (1880, 1)
```

Εικόνα 4.8: (πάνω) Είδος μητρώων δειγμάτων, (κάτω) Μέγεθος μητρώων δειγμάτων

Στη συνέχεια, ορίστηκαν οι **ονομασίες των κλάσεων** στις οποίες αντιστοιχούν οι κωδικές ονομασίες των αριθμών μηδέν έως πέντε όπως αυτές ορίστηκαν κατά την επεξεργασία των επισημάνσεων, τις οποίες θα τυπώνει ο αλγόριθμος. Επιπλέον, χρησιμοποιήθηκε η **κωδικοποίηση δυαδικού χαρακτήρα One-Hot** για τον μετασχηματισμό των κλάσεων σε σύνολο 6 αριθμών για την εισαγωγή τους στο νευρωνικό δίκτυο. Οι αριθμοί αυτοί αντιστοιχούν στις κωδικές ονομασίες των κλάσεων για την ταξινόμηση των εικόνων. Με την κωδικοποίηση μετασχηματίστηκαν τα μητρώα `y_train` και `y_test` μίας στήλης, σε μητρώα ίδιου αριθμού σειρών και 6 στηλών όπου κάθε στήλη αντιστοιχεί στις κλάσεις ταξινόμησης, όπως φαίνεται στον Πίνακα 4.5.

Πίνακας 4.5: Κωδικοποιημένο μητρώο επισημάνσεων για δοκιμές (`y_test_one_hot`)

	0	1	2	3	4	5
0	0	0	0	1	0	0
1	0	0	1	0	0	0
2	0	0	0	0	0	1
3	0	0	1	0	0	0
4	1	0	0	0	0	0
5	0	0	1	0	0	0
6	1	0	0	0	0	0
7	1	0	0	0	0	0
8	1	0	0	0	0	0
9	0	0	1	0	0	0
10	1	0	0	0	0	0
11	1	0	0	0	0	0
12	1	0	0	0	0	0
13	1	0	0	0	0	0
14	0	0	0	0	1	0
15	0	1	0	0	0	0
16	0	0	1	0	0	0
17	1	0	0	0	0	0
18	1	0	0	0	0	0
19	0	0	1	0	0	0
20	1	0	0	0	0	0

Για μία τυχαία εικόνα του πρώτου βίντεο παρατίθενται οι τρεις διαφορετικές ονομασίες με τις οποίες μπορεί να τυπωθεί η κατηγορία ουράς η οποία αντιστοιχεί σε αυτήν (Εικόνα 4.9). Στο καρέ εικόνας 2025 παρατηρείται κυκλοφοριακή ουρά λόγω κατειλημμένης διασταύρωσης (4^η κλάση ή διαφορετικά το 5^ο στοιχείο στο σύνολο 6 αριθμών κλάσεων).

```
The image label is: Class 4
Name: 2025, dtype: int64
The image class is: Occupied_Intersection
The one hot label is: [0. 0. 0. 0. 1. 0.]
```

Εικόνα 4.9: Ονομασίες κλάσης καρέ εικόνας 2025 δείγματος εκπαίδευσης

Η προετοιμασία των δεδομένων ολοκληρώθηκε με την **ομαλοποίηση των εικονοκυτάρων** των δειγμάτων εικόνων x_{train} και x_{test} , ώστε οι τιμές τους να λάβουν τιμές μεταξύ των αριθμών 0 και 1, αντί των τιμών μεταξύ 0 και 255 τις οποίες έλαβαν κατά την αλλαγή των διαστάσεών τους. Με τον τρόπο αυτό διευκολύνεται η ανάγνωση των δεδομένων από το νευρωνικό δίκτυο και μειώνεται ο υπολογιστικός χρόνος του αλγόριθμου κατά την εκπαίδευση και την αξιολόγηση. Ο μετασχηματισμός των μητρώων δειγμάτων εικόνων λόγω ομαλοποίησης φαίνεται στον Πίνακα 4.6.

Πίνακας 4.6: Ομαλοποιημένο μητρώο εικόνων για εκπαίδευση (x_{train})

	0	1	2
20	0.91373	0.99216	0.99608
21	0.70196	0.72157	0.73333
22	0.84314	0.87059	0.88235
23	0.65882	0.68235	0.70588
24	0.68235	0.71373	0.74118
25	0.57255	0.59608	0.61569
26	0.63529	0.69804	0.75686
27	0.54902	0.62353	0.71765
28	0.80392	0.85490	0.90588
29	0.70588	0.77647	0.81569
30	0.79608	0.86667	0.91765
31	0.71765	0.71373	0.73725
32	0.78431	0.81569	0.81569
33	0.47451	0.55686	0.53725
34	0.43922	0.54510	0.56863
35	0.74118	0.81569	0.85098
36	0.81961	0.83922	0.84314
37	0.76078	0.78431	0.78039
38	0.77647	0.76078	0.75686
39	0.72941	0.70980	0.69804
40	0.69020	0.67451	0.66667

Αρχιτεκτονική μοντέλου συνελικτικού νευρωνικού δικτύου

Για την κατασκευή του μοντέλου επιλέχθηκε **αρχιτεκτονική επακόλουθης διαδικασίας εκπαίδευσης** την οποία χρησιμοποιούν τα συνελικτικά νευρωνικά δίκτυα, κατά την οποία τα δεδομένα περνούν από συνεχόμενα πολλαπλά στρώματα εκπαίδευσης. Ειδικότερα, η αρχιτεκτονική των συνελικτικών δικτύων χωρίζεται σε δύο κύρια μέρη, ένα εργαλείο συνέλιξης το οποίο χωρίζει τα διάφορα χαρακτηριστικά της εικόνας για ανάλυση και ένα πλήρως συνδεδεμένο στρώμα το οποίο χρησιμοποιεί τις τιμές εξόδου για την πρόβλεψη της καλύτερης περιγραφής της εικόνας.

Οι νευρώνες του δικτύου CNN χωρίζονται σε μία 3-διάστατη δομή, όπου κάθε ομάδα νευρώνων ειδικεύεται στην αναγνώριση μίας μικρής περιοχής ή ενός χαρακτηριστικού της εικόνας. Τα συνελικτικά δίκτυα χρησιμοποιούν τις προβλέψεις που δημιουργούν τα στρώματα και παράγουν τις τελικές τιμές εξόδου, οι οποίες παρουσιάζουν ένα διάλυμα ποσοστών πιθανότητας ότι ένα συγκεκριμένο χαρακτηριστικό ανήκει σε μία συγκεκριμένη κλάση.

Για το μοντέλο του συνελικτικού νευρωνικού δικτύου χρησιμοποιήθηκαν οι τέσσερις τύποι στρωμάτων που απαρτίζουν τα κλασσικά δίκτυα CNN, τα συνελικτικά στρώματα (**convolution layers**), τα ομαδοποιημένα στρώματα (**pooling layers**), τα πλήρως συνδεδεμένα στρώματα ή διαφορετικά στρώματα ισοπέδωσης (**flattening layers**) και τα στρώματα εγκατάλειψης (**dropout layers**). Για τα στρώματα επιλέχθηκε η συνάρτηση ενεργοποίησης διορθωμένης γραμμικής μονάδας (**ReLU**), ενώ για το τελευταίο η συνάρτηση **softmax**.

Το **στρώμα συνέλιξης** είναι το πρώτο στρώμα το οποίο τοποθετήθηκε στο μοντέλο, το οποίο χρησιμοποιήθηκε για την **εξαγωγή των διάφορων χαρακτηριστικών των εικόνων** προς εκπαίδευση. Καθώς πρόκειται για το πρώτο στρώμα, απαιτήθηκε η εισαγωγή των διαστάσεων των εικόνων της βάσης δεδομένων για την αναγνώριση από το μοντέλο. Επομένως, η μαθηματική λειτουργία της συνέλιξης πραγματοποιήθηκε μεταξύ φίλτρου μεγέθους 5 x 5 και εικόνων μεγέθους 256 x 256 και βάθους 3 (RGB). Επιλέχθηκε η **κατασκευή 32 συνελικτικών χαρακτηριστικών για την μόρφωση του χάρτη χαρακτηριστικών** του πρώτου στρώματος. Η συνάρτηση ενεργοποίησης ReLU χρησιμοποιήθηκε για την βελτίωσή της εκπαίδευσης και της απόδοσης του μοντέλου, για την επιλογή και εξαγωγή των μέγιστων τιμών και συγκεκριμένα των θετικών ή μηδενικών τιμών μόνο από τα χαρακτηριστικά των εικόνων.

Στη συνέχεια τοποθετήθηκε ένα **στρώμα ομαδοποίησης** το οποίο στις περισσότερες περιπτώσεις ακολουθεί μετά το στρώμα συνέλιξης, για τη **μείωση του μεγέθους του μορφοποιημένου χάρτη χαρακτηριστικών** και επομένως του υπολογιστικού χρόνου του μοντέλου. Η ομαδοποίηση πραγματοποιήθηκε με φίλτρο εικονοκυττάρων 2 x 2 και κατηγορία Max Pooling, με την οποία εξάγεται το μεγαλύτερο στοιχείο από τον χάρτη για τη μείωση του μεγέθους του. Επιπλέον, το στρώμα ομαδοποίησης χρησιμοποιήθηκε για την σύνδεση του συνελικτικού στρώματος και του στρώματος ισοπέδωσης που ακολουθεί.

Τοποθετήθηκε ακόμη ένας συνδυασμός στρωμάτων συνέλιξης και ομαδοποίησης χωρίς την επισήμανση των διαστάσεων των εικόνων και επιλέχθηκε η **κατασκευή 64 συνελικτικών χαρακτηριστικών** για τη μόρφωση ενός ακόμη χάρτη χαρακτηριστικών.

Ύστερα χρησιμοποιήθηκε ένα **στρώμα ισοπέδωσης** με το οποίο οι εικόνες εισόδου από τα προηγούμενα στρώματα μετατράπηκαν σε μία γραμμική συστοιχία η οποία τροφοδότησε τα στρώματα διαφορετικών νευρώνων που ακολουθούν για τη σύνδεσή τους.

Τέλος, κατασκευάστηκαν τέσσερα διαφορετικά νευρωνικά δίκτυα των οποίων τα πρώτα στρώματα αποτελούνταν από 1000, 500, 250 και 6 νευρώνες αντίστοιχα. Για την **αποφυγή υπερφόρτωσης του μοντέλου** χρησιμοποιήθηκαν στα δύο πρώτα δίκτυα τα βοηθητικά **στρώματα εγκατάλειψης** με ποσοστό απόσυρσης 50% νευρώνων από το δίκτυο για τη μείωση του μεγέθους του μοντέλου και την αντιμετώπιση της αρνητικής απόδοσης. Η συνάρτηση ενεργοποίησης ReLU χρησιμοποιήθηκε για τα τρία πρώτα δίκτυα, ενώ για το τελευταίο με **αριθμό νευρώνων ίδιο με αυτόν των κλάσεων** για ταξινόμηση των εικόνων, χρησιμοποιήθηκε η συνάρτηση softmax για την μετατροπή των τιμών εισόδου σε τιμές πιθανοτήτων μεταξύ 0 και 1.

Για την εκπαίδευση του μοντέλου, χρησιμοποιήθηκαν:

- ♦ το ομαλοποιημένο δείγμα εικόνων **x_train** και το κωδικοποιημένο δείγμα επισημάνσεων **y_train_one_hot** για εκπαίδευση
- ♦ η **κατηγορική συνάρτηση απώλειας crossentropy** η οποία χρησιμοποιείται για προβλήματα ταξινόμησης πολλαπλών κλάσεων (περισσότερες από 2),
- ♦ ο **αλγόριθμος βελτιστοποίησης Adam** ο οποίος είναι μία στοχαστική μέθοδος που βασίζεται στην προσαρμοστική εκτίμηση με ποσοστό μάθησης ίσο με 0.001, κατάλληλος για εκπαίδευση μοντέλων βαθιάς μάθησης.
- ♦ η **μετρική ακρίβειας**

Η εκπαίδευση πραγματοποιήθηκε με τη μέθοδο προσαρμογής, με παραμέτρους:

- ♦ μέγεθος παρτίδας (**batch size**) = 225 παραδείγματα εκπαίδευσης σε μία παρτίδα
- ♦ αριθμός επαναλήψεων (**epochs**) = 10 επαναλήψεις για τη μεταφορά ολόκληρου του συνόλου των δεδομένων μέσω του νευρωνικού δικτύου μόνο μία φορά
- ♦ ποσοστό επικύρωσης (**validation**) = 20% των δεδομένων εκπαίδευσης

Με την ολοκλήρωση κάθε επανάληψης (epoch) τυπώθηκαν οι τιμές ακρίβειας και απώλειας των δεδομένων εκπαίδευσης και επικύρωσης αντίστοιχα, καθώς και ο υπολογιστικός χρόνος ολοκληρωμένης επανάληψης και κάθε βήματος ξεχωριστά.

Έλεγχος μετρικών μοντέλου

Για τον υπολογισμό και τον έλεγχο των μετρικών του μοντέλου χρησιμοποιήθηκαν το ομαλοποιημένο δείγμα εικόνων **x_test**, το κωδικοποιημένο δείγμα επισημάνσεων **y_test_one_hot** και το κανονικό δείγμα επισημάνσεων **y_test**.

Ακρίβεια (accuracy): με την εντολή `model.evaluate()` και τα δείγματα `x_test` και `y_test_one_hot`, υπολογίστηκε η ακρίβεια και η απώλεια του μοντέλου ταξινόμησης, καθώς και ο υπολογιστικός χρόνος εντολής.

Προβλέψεις (predictions): με την εντολή `model.predict()` και το δείγμα `x_test` υπολογίστηκαν οι προβλέψεις των κλάσεων του μοντέλου (`y_pred`), για τις εικόνες του δείγματος δοκιμής.

Μητρώο σύγχυσης (confusion matrix): με τις εντολές `confusion_matrix()` και `multilabel_confusion_matrix()` και τη χρήση του δείγματος `y_test` και των προβλέψεων του μοντέλου `y_pred`, υπολογίστηκαν το συνολικό μητρώο σύγχυσης και το μητρώο σύγχυσης πολλαπλών κλάσεων του μοντέλου αντίστοιχα.

Αναφορά ταξινόμησης (classification report): με την εντολή `classification_report()` και τη χρήση του δείγματος `y_test` και των προβλέψεων του μοντέλου `y_pred`, υπολογίστηκε η αναφορά ταξινόμησης με την οποία τυπώθηκαν οι τυπικές μετρικές ενός ταξινομητή. Οι τυπικές μετρικές υπολογίστηκαν από τους αριθμούς των σωστών και λανθασμένων ταξινομήσεων **True Positive (TP)**, **True Negative (TN)**, **False Positive (FP)**, **False Negative (FN)** βάσει του Πίνακα 4.7 και των παρακάτω σχέσεων:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Specificity = \frac{TN}{TN+FP} \quad (4)$$

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

$$False\ Positive\ Rate = \frac{FP}{FP+TN} \quad (5)$$

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

$$F1\ score = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (6)$$

Πίνακας 4.7: Μητρώο αριθμών για αληθινές και ψευδείς ταξινομήσεις

		Actual Class	
		Positive	Negative
Predicted Class	Positive	TP	FP
	Negative	FN	TN

Προβλέψεις τυχαίας εικόνας δείγματος δοκιμών: με τον προσδιορισμό μίας λίστας κλάσεων και τη χρήση των προβλέψεων `y_pred`, για μία τυχαία εικόνα του δείγματος πραγματοποιήθηκε διαχωρισμός των προβλέψεων από τη λιγότερο πιθανή στην πιο πιθανή κλάση που της αντιστοιχεί και εκτυπώθηκαν τα ποσοστά πιθανοτήτων τους.

Αξιολόγηση μοντέλου

Για την αξιολόγηση του μοντέλου κατασκευάστηκαν τα **διαγράμματα ακρίβειας και απώλειας** για τα δεδομένα εκπαίδευσης και επικύρωσης με τη χρήση της εντολής `.history()` και των τιμών του μοντέλου.

Επιπλέον, η απόδοση του μοντέλου αξιολογήθηκε με την κατασκευή της **καμπύλης λειτουργίας δέκτη (ROC)** για ταξινομητή πολλαπλών κλάσεων, και τον υπολογισμό των τιμών **περιοχής κάτω από την καμπύλη AUC** για κάθε κλάση, από τις οποίες συμπεραίνεται η ποιότητα του μοντέλου. Χρησιμοποιήθηκαν οι τιμές True Positive Rate, False Positive Rate οι οποίες αφορούν τις πραγματικές κλάσεις και το δείγμα `y_test`, το οποίο μετασχηματίστηκε σε δυαδική μορφή κλάσεων καθώς οι ROC κατασκευάζονται με δεδομένα δυαδικού χαρακτήρα, με την εντολή `label_binarize(classes=())`

Μετά τον υπολογισμό των διαγραμμάτων ROC, υπολογίστηκε ο μέσος όρος μακροεντολών (**macro average**), καθώς και ο μέσος όρος σταθμισμένης συχνότητας (**prevalence-weighted average**) με την εντολή `roc_auc_score()` και τη χρήση των δειγμάτων εικόνων `x_test` και επισημάνσεων `y_test`. Οι παραπάνω τιμές υπολογίστηκαν για τις κατηγορίες πολλαπλών κλάσεων **One-vs-One (OvO)** και **One-vs-Rest (OvR)**, οι οποίες συγκρίνουν κάθε μοναδικό συνδυασμό κατηγοριών κατά ζεύγη.

Δοκιμή μεταφοράς μάθησης

Ο αλγόριθμος ολοκληρώνεται με τη δοκιμή μεταφοράς μάθησης των μοντέλων κάθε βίντεο, για τα δεδομένα των άλλων βάσεων. Συγκεκριμένα πραγματοποιήθηκε μεταφορά μάθησης μεταξύ των βίντεο 1, 2, 3.1 και μεταξύ των βάσεων δεδομένων των αποσπασμάτων του τρίτου βίντεο 3.1, 3.2, 3.3 με την **ανάπτυξη μοντέλου μεταφοράς**. Η μεταφορά δοκιμάστηκε μεταξύ των διαφορετικών βίντεο λόγω των μικτών δεδομένων των τριών διαφορετικών βάσεων δεδομένων, ενώ μεταξύ των αποσπασμάτων του ίδιου βίντεο λόγω της ομοιότητας των δεδομένων καθώς πρόκειται για την ίδια περιοχή, αλλά με διαφορετική καταγραφή κυκλοφορίας.

Ειδικότερα, εισήχθησαν μόνο τα δεδομένα εικόνων των άλλων βίντεο και πραγματοποιήθηκε ο διαχωρισμός των δειγμάτων εκπαίδευσης και δοκιμής όπως στην αρχή του αλγόριθμου. Σημειώνεται ότι χρησιμοποιήθηκε **συγκεκριμένο ποσοστό διαχωρισμού** για κάθε δοκιμή, ώστε τα δείγματα δοκιμών να έχουν το ίδιο πλήθος δεδομένων με αυτό του δείγματος δοκιμών του επαναχρησιμοποιούμενου μοντέλου.

Η αξιολόγηση της μεταφοράς μάθησης πραγματοποιήθηκε με τον υπολογισμό των προβλέψεων, μητρώων σύγχυσης και της αναφοράς ταξινόμησης του ανεπτυγμένου μοντέλου με τη χρήση των δειγμάτων δοκιμών επισημάνσεων των βάσεων δεδομένων των άλλων βίντεο.

4.6 ΑΠΟΤΕΛΕΣΜΑΤΑ

Οι αλγόριθμοι βαθιάς μάθησης βασίζονται στην ψηφιακή βιβλιοθήκη και τη μεθοδολογία της TensorFlow. Η προετοιμασία των δεδομένων, η εκπαίδευση του μοντέλου και η αξιολόγηση των δοκιμών πραγματοποιήθηκαν με τα παρακάτω χαρακτηριστικά υπολογιστή:

- ♦ **Επεξεργαστής:** Intel® Core™ i7-7700HQ Processor, 4/8 (Cores/Threads) with 6MB Cache, Base Frequency 2.80 GHz up to 3.80 GHz
- ♦ **Κάρτα γραφικών:** NVIDIA GeForce GTX 1060 with Max-Q Design, Pipelines: 1280, VRAM: 6GB GDDR5
- ♦ **Μνήμη RAM:** 16 GB 2400 MHz

4.6.1 Οδός Πατησίων - Λεωφόρος Αλεξάνδρας (Video 1)

Για το μοντέλο V1 χρησιμοποιήθηκαν 7518 καρέ εικόνων, από τα οποία 5638 για εκπαίδευση και 1880 για δοκιμές. Παρουσιάστηκε σχηματισμός ουρών των κατηγοριών: Stop (1), Turn (6), Mandatory Lane Contraction (2), Occupied Intersection (4), Spillover Effect (3).

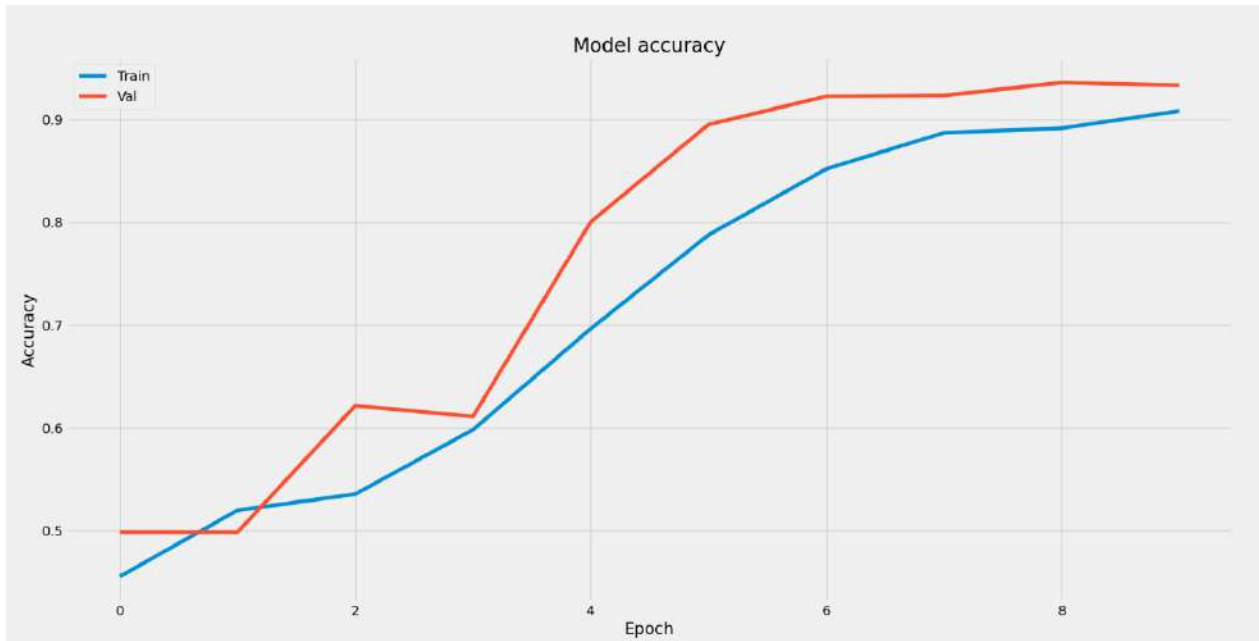
Ακρίβεια μοντέλου

Το μοντέλο παρουσίασε **πολύ ικανοποιητική ακρίβεια**, σχεδόν 95%, για ταξινομητή πολλαπλών κλάσεων και **αρκετά μικρή απώλεια** (Εικόνα 4.10). Παρατηρήθηκαν μικρές διακυμάνσεις κατά την εκπαίδευση και την επικύρωση. Ο συνολικός υπολογιστικός χρόνος του μοντέλου ήταν περίπου 81 λεπτά (8:06 min/epoch), ικανοποιητική διάρκεια για τα δεδομένα και τις παραμέτρους που χρησιμοποιήθηκαν.

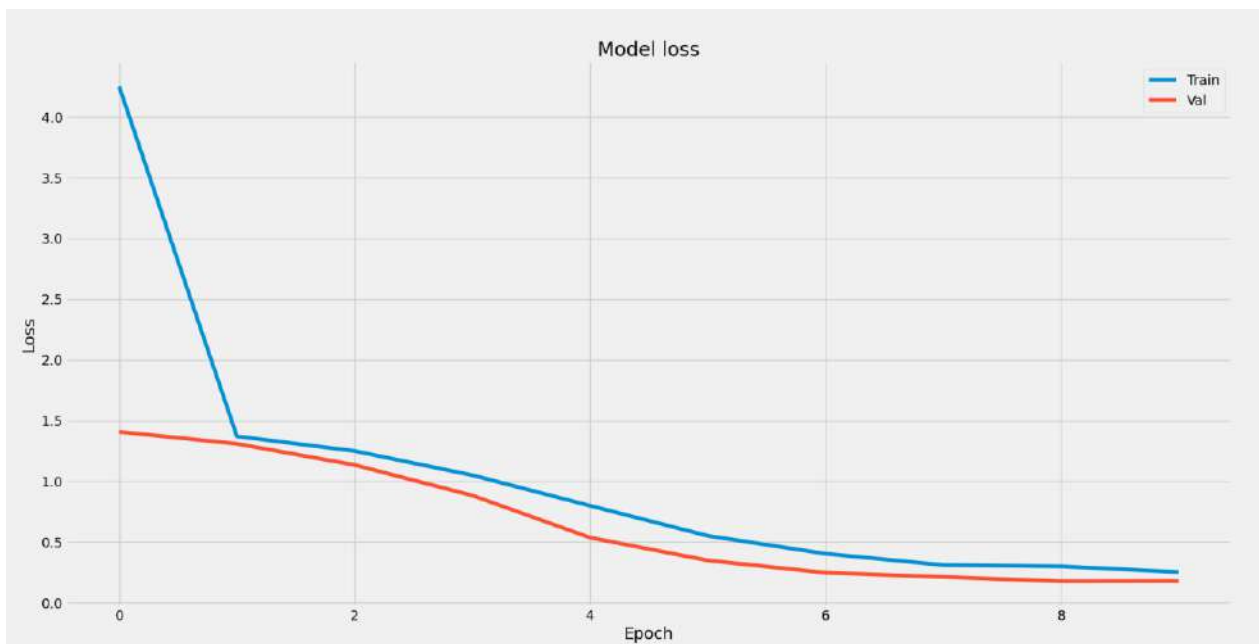
```
Epoch 1/10
21/21 [=====] - 381s 18s/step - loss: 4.2503 - accuracy: 0.4550 - val_loss: 1.4076 - val_accuracy: 0.4982
Epoch 2/10
21/21 [=====] - 369s 18s/step - loss: 1.3708 - accuracy: 0.5195 - val_loss: 1.3093 - val_accuracy: 0.4982
Epoch 3/10
21/21 [=====] - 467s 22s/step - loss: 1.2500 - accuracy: 0.5353 - val_loss: 1.1353 - val_accuracy: 0.6215
Epoch 4/10
21/21 [=====] - 500s 24s/step - loss: 1.0501 - accuracy: 0.5980 - val_loss: 0.8831 - val_accuracy: 0.6108
Epoch 5/10
21/21 [=====] - 498s 24s/step - loss: 0.7990 - accuracy: 0.6965 - val_loss: 0.5365 - val_accuracy: 0.8005
Epoch 6/10
21/21 [=====] - 495s 24s/step - loss: 0.5516 - accuracy: 0.7878 - val_loss: 0.3489 - val_accuracy: 0.8954
Epoch 7/10
21/21 [=====] - 518s 25s/step - loss: 0.4041 - accuracy: 0.8521 - val_loss: 0.2476 - val_accuracy: 0.9229
Epoch 8/10
21/21 [=====] - 501s 24s/step - loss: 0.3105 - accuracy: 0.8874 - val_loss: 0.2123 - val_accuracy: 0.9238
Epoch 9/10
21/21 [=====] - 505s 24s/step - loss: 0.3006 - accuracy: 0.8918 - val_loss: 0.1770 - val_accuracy: 0.9362
Epoch 10/10
21/21 [=====] - 611s 29s/step - loss: 0.2532 - accuracy: 0.9086 - val_loss: 0.1780 - val_accuracy: 0.9335
59/59 [=====] - 104s 2s/step - loss: 0.1571 - accuracy: 0.9479
```

Εικόνα 4.10: Πληροφορίες επαναλήψεων και ακρίβεια 94.79% μοντέλου V1

Στο Διάγραμμα 4.1 παρατηρήθηκαν απότομες μεταβολές των τιμών στις πρώτες επαναλήψεις λόγω μη ομαλής εισαγωγής των διαφορετικών επισημασμένων δεδομένων. Αντίθετα, στο Διάγραμμα 4.2 η απότομη πτώση της τιμής κατά την εκκίνηση της εκπαίδευσης λόγω εισαγωγής πολλών πληροφοριών στο μοντέλο ήταν αναμενόμενη. Σημειώνεται ότι τα αποτελέσματα εκπαίδευσης και επικύρωσης κατά την ολοκλήρωση του μοντέλου στην τελευταία επανάληψη παρουσίασαν μικρές αποκλίσεις στις τιμές τους.



Διάγραμμα 4.1: Ακρίβεια μοντέλου για δεδομένα εκπαίδευσης και επικύρωσης V1



Διάγραμμα 4.2: Απώλεια μοντέλου για δεδομένα εκπαίδευσης και επικύρωσης V1

Μητρώα σύγχυσης και αναφορά ταξινόμησης

Η ακρίβεια του μοντέλου επαληθεύτηκε με τα μητρώα σύγχυσης και την αναφορά ταξινόμησης, στα οποία παρατηρήθηκαν **ελάχιστες λανθασμένες προβλέψεις** (Εικόνα 4.11). Όλες οι κατηγορίες ουρών ανιχνεύτηκαν από το μοντέλο, με μικρότερο ποσοστό συνοχής ακρίβειας 92%, της κλάσης Stop (1), η οποία είχε πολύ μικρό δείγμα 39 εικόνων. Άξιο αναφοράς ότι οι προβλέψεις του δείγματος μέτριου πλήθους (392 εικόνες) της κλάσης Turn (2), παρουσίασαν ποσοστό 100%.

<pre> [[933 0 0 0 6 6] [0 38 0 0 1 0] [56 0 330 0 0 6] [3 2 0 81 0 1] [2 0 0 0 236 1] [6 0 0 3 5 164]] [[[868 67] [[1790 3] [12 933]] [6 81]] [[1839 2] [[1629 12] [1 38]] [3 236]] [[1488 0] [[1688 14] [62 330]] [14 164]] </pre>					
	precision	recall	f1-score	support	
0	0.93	0.99	0.96	945	
1	0.95	0.97	0.96	39	
2	1.00	0.84	0.91	392	
3	0.96	0.93	0.95	87	
4	0.95	0.99	0.97	239	
5	0.92	0.92	0.92	178	
accuracy			0.95	1880	
macro avg	0.95	0.94	0.95	1880	
weighted avg	0.95	0.95	0.95	1880	

Εικόνα 4.11: (αριστερά) Μητρώα σύγχυσης μοντέλου V1, (δεξιά) Αναφορά ταξινόμησης μοντέλου V1

Πρόβλεψη ταξινόμησης τυχαίων εικόνων δείγματος

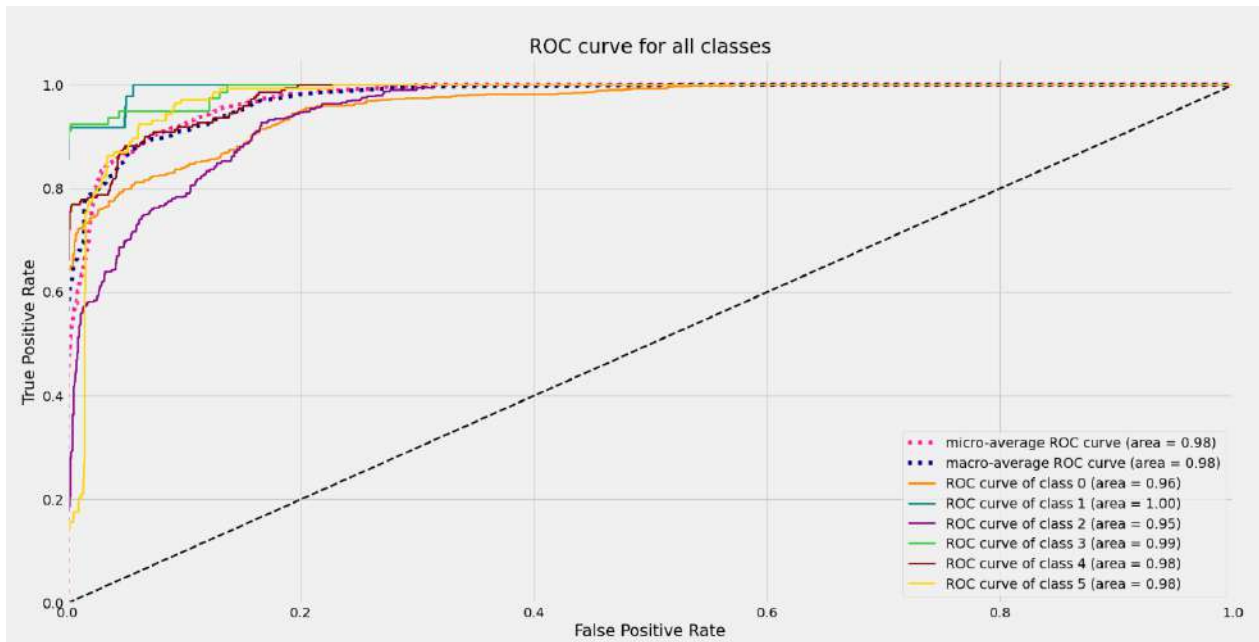
Με τη εφαρμογή του μοντέλου στις παρακάτω τυχαίες εικόνες του δείγματος δοκιμής, προέκυψαν **ακριβείς προβλέψεις υψηλού ποσοστού** ανίχνευσης ουράς:

Εικόνα 1682 - Κλάση (4): [Occupied_Intersection] 86.34%, [Stop] 6.2%, [No_Queue] 5.68%, [Manatory_Lane_Contraction] 1.22%, [Spillover_Effect] 0.47%, [Turn] 0.1%

Εικόνα 4712 - Κλάση (0): [No_Queue] 99.98%, [Turn] 0.01%, [Stop] 0.0%, [Manatory_Lane_Contraction] 0.0%, [Occupied_Intersection] 0.0%, [Spillover_Effect] 0.0%

Καμπύλες ROC και περιοχές AUC

Η **υψηλή ποιότητα του ταξινομητή** αξιολογήθηκε από τις περιοχές AUC και τις καμπύλες ROC (Διάγραμμα 4.3). Τα ποσοστά των περιοχών AUC όλων των κλάσεων παρουσίασαν τιμή τουλάχιστον ίση με 0.95, εφόσον ταυτίζονται στο μεγαλύτερο τμήμα τους με τις τιμές των καμπυλών ROC και ένα πολύ μικρό ποσοστό τους εμφανίστηκε κάτω από αυτές. Επομένως, ο ταξινομητής χαρακτηρίζεται εξαιρετικός για τιμές μεταξύ 0.9 Και 1.0 και οι προβλέψεις του είναι αξιόπιστες και με πολύ μικρό ποσοστό λανθασμένων προβλέψεων



Διάγραμμα 4.3: Καμπύλες ROC και τιμές AUC για τις κλάσεις του μοντέλου V1

Μεταφορά μάθησης

Η μεταφορά μάθησης του ταξινομητή του Video 1 στις εικόνες του Video 2 παρουσίασε **μέτρια αποτελέσματα**, με αναγνώριση μόνο τεσσάρων κλάσεων (Εικόνα 4.12). Πραγματοποιήθηκε πρόβλεψη κυρίως των κλάσεων παρόμοιου ποσοστού εμφάνισης στα ίδια σημεία καταγραφής των δύο πειραμάτων, ενώ οι κλάσεις που είχαν καταγραφεί σε κάθε βίντεο ξεχωριστά δεν ανιχνεύθηκαν. Η μεταφορά μάθησης του ίδιου ταξινομητή στις εικόνες του Video 3.1 παρουσίασε **ανεπαρκή αποτελέσματα** λόγω μεγάλης διαφοροποίησης των περιοχών καταγραφής και των ποσοστών εμφάνισης των κυκλοφοριακών ουρών, εκτός της κλάσης No Queue (0) όπου παρατηρήθηκε συνοχή ακρίβειας προβλέψεων περίπου 50% (Εικόνα 4.13).

[[736	0	142	72	0	5]
[42	0	6	3	0	0]
[312	0	65	36	0	4]
[64	0	17	4	0	0]
[162	0	47	18	0	1]
[106	0	24	10	0	4]]
[[[239	686]	[[1656	139]		
[219	736]]	[81	4]]		
[[1829	0]	[[1652	0]		
[51	0]]	[228	0]]		
[[1227	236]	[[1726	10]		
[352	65]]	[140	4]]		

	precision	recall	f1-score	support
0	0.52	0.77	0.62	955
1	0.00	0.00	0.00	51
2	0.22	0.16	0.18	417
3	0.03	0.05	0.04	85
4	0.00	0.00	0.00	228
5	0.29	0.03	0.05	144
accuracy			0.43	1880
macro avg	0.17	0.17	0.15	1880
weighted avg	0.33	0.43	0.36	1880

Εικόνα 4.12: (αριστερά) Μητρώα σύγχυσης δεδομένων Video 2 - μοντέλου V1, (δεξιά) Αναφορά ταξινόμησης δεδομένων Video 2 - μοντέλου V1

```

[[345  0  0  81 490 39]
 [ 17  0  0  2 25  7]
 [148  0  3 30 220 16]
 [ 22  1  1 10  47  4]
 [ 82  0  1 22 108 15]
 [ 40  0  1 12  81 10]]

[[[ 616 309]  [[1648 147]
 [ 610 345]]  [  75  10]]

[[1828  1]  [[ 789 863]
 [  51  0]]  [ 120 108]]

[[1460  3]  [[1655 81]
 [ 414  3]]  [ 134 10]]]

```

	precision	recall	f1-score	support
0	0.53	0.36	0.43	955
1	0.00	0.00	0.00	51
2	0.50	0.01	0.01	417
3	0.06	0.12	0.08	85
4	0.11	0.47	0.18	228
5	0.11	0.07	0.09	144
accuracy			0.25	1880
macro avg	0.22	0.17	0.13	1880
weighted avg	0.40	0.25	0.25	1880

Εικόνα 4.13: (αριστερά) Μητρώα σύγχυσης δεδομένων Video 3.1 - μοντέλου V1, (δεξιά) Αναφορά ταξινόμησης δεδομένων Video 3.1 - μοντέλου V1

4.6.2 Λεωφόρος Αλεξάνδρας - Οδός Πατησίων (Video 2)

Για το μοντέλο V2 χρησιμοποιήθηκαν 8180 καρέ εικόνων, από τα οποία 6135 για εκπαίδευση και 2045 για δοκιμές. Παρουσιάστηκε σχηματισμός ουρών των κατηγοριών: Stop (3), Turn (5), Mandatory Lane Contraction (2), Occupied Intersection (3), Spillover Effect (2).

Ακρίβεια μοντέλου

Το μοντέλο παρουσίασε **πολύ ικανοποιητική ακρίβεια**, σχεδόν 98%, για ταξινομητή πολλαπλών κλάσεων και **σχεδόν μηδενική απώλεια** (Εικόνα 4.14). Δεν παρατηρήθηκαν διακυμάνσεις κατά την εκπαίδευση και την επικύρωση. Ο συνολικός υπολογιστικός χρόνος του μοντέλου ήταν περίπου 65 λεπτά (6:30 min/epoch), αρκετά ικανοποιητική διάρκεια για τα δεδομένα και τις παραμέτρους που χρησιμοποιήθηκαν.

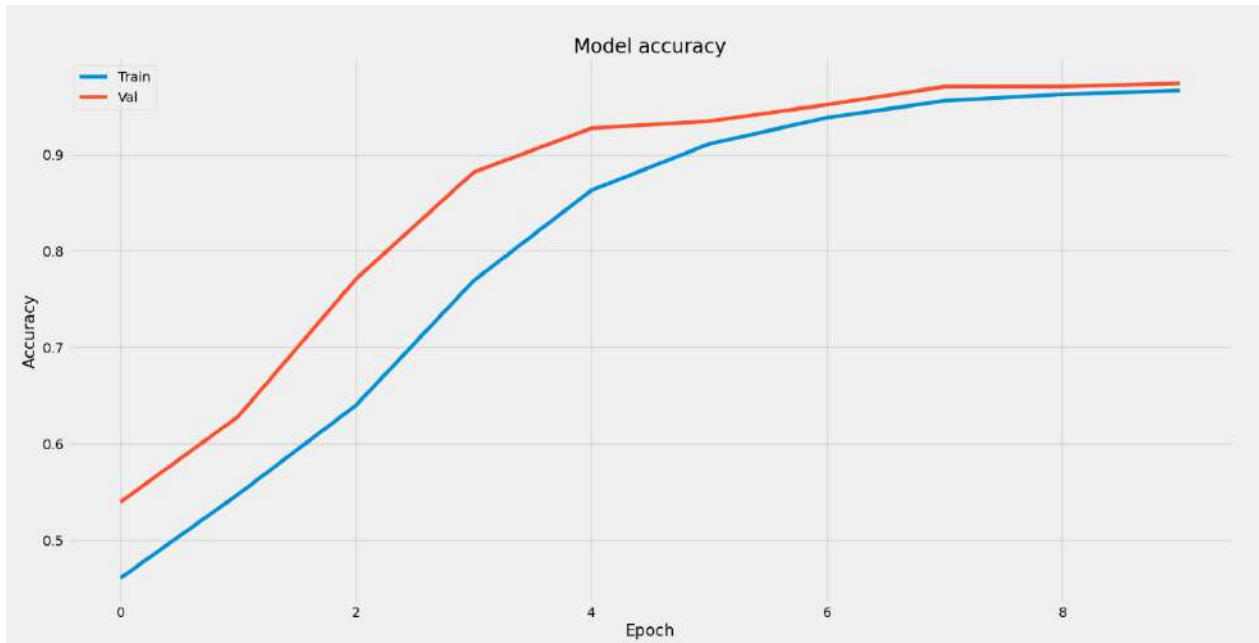
```

Epoch 1/10
22/22 [=====] - 367s 17s/step - loss: 4.8030 - accuracy: 0.4607 - val_loss: 1.3482 - val_accuracy: 0.5395
Epoch 2/10
22/22 [=====] - 363s 17s/step - loss: 1.2720 - accuracy: 0.5477 - val_loss: 0.9825 - val_accuracy: 0.6284
Epoch 3/10
22/22 [=====] - 360s 16s/step - loss: 0.9654 - accuracy: 0.6400 - val_loss: 0.6455 - val_accuracy: 0.7710
Epoch 4/10
22/22 [=====] - 373s 17s/step - loss: 0.6336 - accuracy: 0.7694 - val_loss: 0.3592 - val_accuracy: 0.8818
Epoch 5/10
22/22 [=====] - 426s 19s/step - loss: 0.3783 - accuracy: 0.8631 - val_loss: 0.1989 - val_accuracy: 0.9275
Epoch 6/10
22/22 [=====] - 400s 18s/step - loss: 0.2552 - accuracy: 0.9112 - val_loss: 0.1591 - val_accuracy: 0.9348
Epoch 7/10
22/22 [=====] - 392s 18s/step - loss: 0.1816 - accuracy: 0.9385 - val_loss: 0.1157 - val_accuracy: 0.9519
Epoch 8/10
22/22 [=====] - 367s 17s/step - loss: 0.1309 - accuracy: 0.9560 - val_loss: 0.0794 - val_accuracy: 0.9707
Epoch 9/10
22/22 [=====] - 418s 19s/step - loss: 0.1009 - accuracy: 0.9627 - val_loss: 0.0709 - val_accuracy: 0.9707
Epoch 10/10
22/22 [=====] - 384s 17s/step - loss: 0.0937 - accuracy: 0.9668 - val_loss: 0.0596 - val_accuracy: 0.9739
64/64 [=====] - 35s 541ms/step - loss: 0.0563 - accuracy: 0.9790

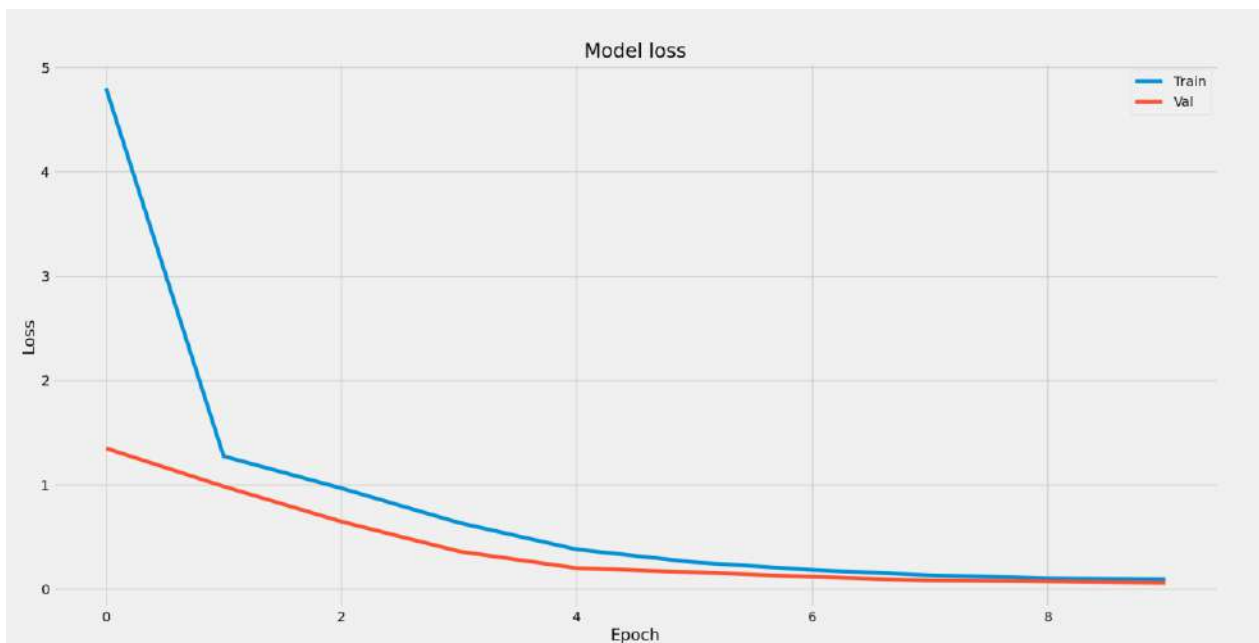
```

Εικόνα 4.14: Πληροφορίες επαναλήψεων και ακρίβεια 97.9% μοντέλου V2

Στο Διάγραμμα 4.4 παρατηρήθηκαν ομοιόμορφες αυξήσεις και μειώσεις των τιμών σε όλες τις επαναλήψεις λόγω ομαλής εισαγωγής των διαφορετικών επισημασμένων δεδομένων. Αντίθετα, στο Διάγραμμα 4.5 η απότομη πτώση της τιμής κατά την εκκίνηση της εκπαίδευσης λόγω εισαγωγής πολλών πληροφοριών στο μοντέλο ήταν αναμενόμενη. Σημειώνεται ότι τα αποτελέσματα εκπαίδευσης και επικύρωσης κατά την ολοκλήρωση του μοντέλου στην τελευταία επανάληψη παρουσίασαν σχεδόν ίσες τιμές.



Διάγραμμα 4.4: Ακρίβεια μοντέλου για δεδομένα εκπαίδευσης και επικύρωσης V2



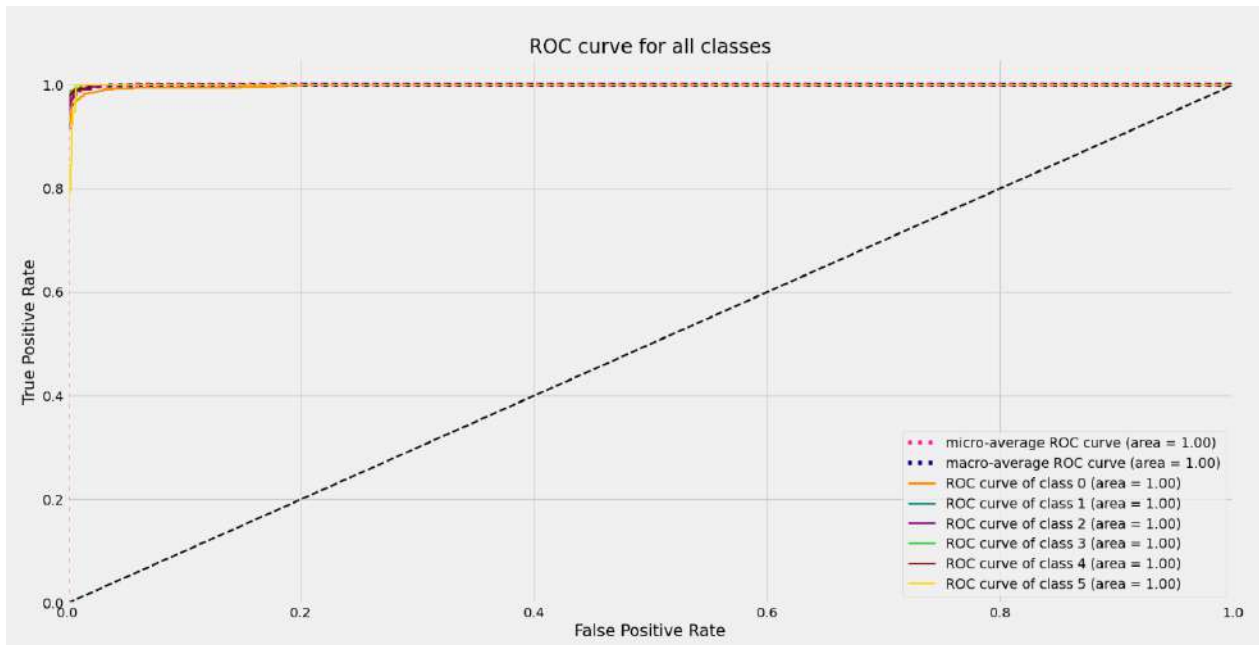
Διάγραμμα 4.5: Απώλεια μοντέλου για δεδομένα εκπαίδευσης και επικύρωσης V2

Μητρώα σύγχυσης και αναφορά ταξινόμησης

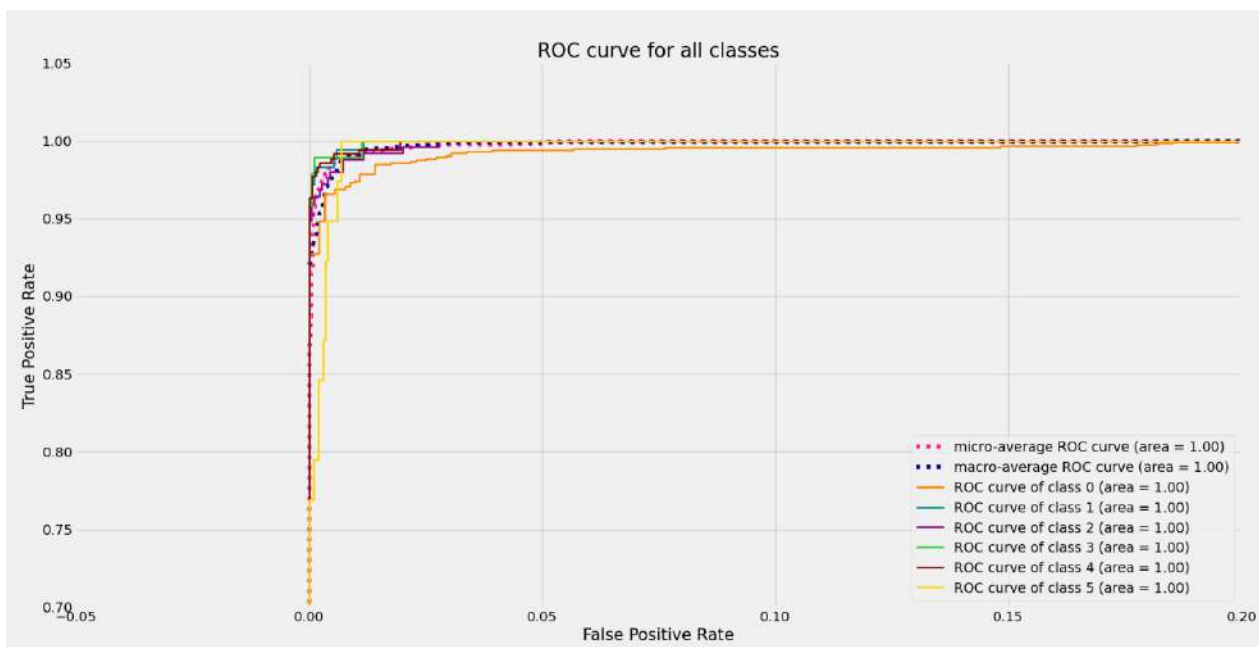
Η ακρίβεια του μοντέλου επαληθεύτηκε με τα μητρώα σύγχυσης και την αναφορά ταξινόμησης, στα οποία παρατηρήθηκαν **ελάχιστες λανθασμένες προβλέψεις** (Εικόνα 4.15). Όλες οι κατηγορίες ουρών ανιχνεύτηκαν από το μοντέλο, με μικρότερο ποσοστό συνοχής ακρίβειας 92% της κλάσης Spillover Effect (5), η οποία είχε πολύ μικρό δείγμα 35 εικόνων. Άξιο αναφοράς ότι οι προβλέψεις του δείγματος μέτριου πλήθους (237 εικόνες) της κλάσης Turn (2), παρουσίασαν ποσοστό 100%.

[[1144 13 1 5 2 0]					
[5 177 0 0 0 0]					
[5 0 229 0 0 3]					
[1 0 0 107 0 0]					
[7 0 0 0 311 0]					
[1 0 0 0 0 34]]					
[[[861 19] [[1932 5]					
[21 1144]] [1 107]]					
[[1850 13] [[1725 2]					
[5 177]] [7 311]]					
[[1807 1] [[2007 3]					
[8 229]] [1 34]]]					

	precision	recall	f1-score	support
0	0.98	0.98	0.98	1165
1	0.93	0.97	0.95	182
2	1.00	0.97	0.98	237
3	0.96	0.99	0.97	108
4	0.99	0.98	0.99	318
5	0.92	0.97	0.94	35
accuracy			0.98	2045
macro avg	0.96	0.98	0.97	2045
weighted avg	0.98	0.98	0.98	2045



Διάγραμμα 4.6: Καμπύλες ROC και τιμές AUC για τις κλάσεις του μοντέλου V2



Διάγραμμα 4.7: Καμπύλες ROC και τιμές AUC για τις κλάσεις του μοντέλου V2 (μεγέθυνση διαγράμματος)

Μεταφορά μάθησης

Η μεταφορά μάθησης του ταξινομητή του Video 1 στις εικόνες του Video 2 παρουσίασε **μέτρια αποτελέσματα**, με αναγνώριση μόνο δύο κλάσεων (Εικόνα 4.16). Πραγματοποιήθηκε πρόβλεψη κυρίως των κλάσεων παρόμοιου ποσοστού εμφάνισης στα ίδια σημεία καταγραφής των δύο πειραμάτων, ενώ οι κλάσεις που είχαν καταγραφεί σε κάθε βίντεο ξεχωριστά δεν ανιχνεύτηκαν. Η μεταφορά μάθησης του ίδιου ταξινομητή στις εικόνες του Video 3.1 παρουσίασε **ανεπαρκή αποτελέσματα** λόγω μεγάλης διαφοροποίησης των περιοχών καταγραφής και των ποσοστών εμφάνισης των κυκλοφοριακών ουρών, εκτός της κλάσης No Queue (0) όπου παρατηρήθηκε συνοχή ακρίβειας προβλέψεων περίπου 50% (Εικόνα 4.17).

[[983 182 0 0 0 0]					
[156 26 0 0 0 0]					
[203 34 0 0 0 0]					
[88 20 0 0 0 0]					
[264 54 0 0 0 0]					
[27 8 0 0 0 0]]					
[[[142 738] [[1937 0]					
[182 983]] [108 0]]					
[[1565 298] [[1727 0]					
[156 26]] [318 0]]					
[[1808 0] [[2010 0]					
[237 0]] [35 0]]]					

	precision	recall	f1-score	support
0	0.57	0.84	0.68	1165
1	0.08	0.14	0.10	182
2	0.00	0.00	0.00	237
3	0.00	0.00	0.00	108
4	0.00	0.00	0.00	318
5	0.00	0.00	0.00	35
accuracy			0.49	2045
macro avg	0.11	0.16	0.13	2045
weighted avg	0.33	0.49	0.40	2045

Εικόνα 4.16: (αριστερά) Μητρώα σύγχυσης δεδομένων Video 1 - μοντέλου V2, (δεξιά) Αναφορά ταξινόμησης δεδομένων Video 1 - μοντέλου V2

[[270 857 6 0 0 0]					
[42 136 0 0 0 0]					
[50 202 1 0 0 0]					
[19 77 1 0 0 0]					
[91 262 1 0 0 0]					
[9 30 0 0 0 0]]					
[[[707 211] [[1949 0]					
[857 270]] [96 0]]					
[[439 1428] [[1692 0]					
[42 136]] [353 0]]					
[[1785 8] [[2006 0]					
[251 1]] [39 0]]]					

	precision	recall	f1-score	support
0	0.56	0.24	0.34	1127
1	0.09	0.76	0.16	178
2	0.11	0.00	0.01	252
3	0.00	0.00	0.00	96
4	0.00	0.00	0.00	353
5	0.00	0.00	0.00	39
accuracy			0.20	2045
macro avg	0.11	0.17	0.08	2045
weighted avg	0.32	0.20	0.20	2045

Εικόνα 4.17: (αριστερά) Μητρώα σύγχυσης δεδομένων Video 3.1 - μοντέλου V2, (δεξιά) Αναφορά ταξινόμησης δεδομένων Video 3.1 - μοντέλου V2

4.6.3 Πλατεία Ομονοίας (Video 3)

Για τα μοντέλα V3.1 και V3.2 του Video 3 χρησιμοποιήθηκαν 7500 καρέ εικόνων, από τα οποία 5625 για εκπαίδευση και 1875 για δοκιμές. Στο V3.1 παρουσιάστηκε σχηματισμός ουρών των κατηγοριών: Stop (2), Turn (7), Mandatory Lane Contraction (5), Occupied Intersection (7), Spillover Effect (4). Στο V3.2 παρουσιάστηκε σχηματισμός ουρών των κατηγοριών: Stop (3), Turn (7), Mandatory Lane Contraction (5), Occupied Intersection (4), Spillover Effect (1). Για το μοντέλο V3.3 του Video 3 χρησιμοποιήθηκαν 7970 καρέ εικόνων, από τα οποία 5977 για εκπαίδευση και 1993 για δοκιμές. Παρουσιάστηκε σχηματισμός ουρών των κατηγοριών: Stop (2), Turn (5), Mandatory Lane Contraction (2), Occupied Intersection (2), Spillover Effect (1). Αναλύεται το μοντέλο V3.1 καθώς παρουσίασε τα καλύτερα αποτελέσματα.

Ακρίβεια μοντέλου

Το μοντέλο παρουσίασε **πολύ ικανοποιητική ακρίβεια**, σχεδόν 97%, για ταξινομητή πολλαπλών κλάσεων και **σχεδόν μηδενική απώλεια** (Εικόνα 4.18). Παρατηρήθηκαν μικρές διακυμάνσεις κατά την εκπαίδευση και την επικύρωση. Ο συνολικός υπολογιστικός χρόνος του μοντέλου ήταν περίπου 56 λεπτά (5:36 min/epoch), αρκετά ικανοποιητική διάρκεια για τα δεδομένα και τις παραμέτρους που χρησιμοποιήθηκαν.

```
Epoch 1/10
20/20 [=====] - 347s 17s/step - loss: 3.9766 - accuracy: 0.2647 - val_loss: 1.5854 - val_accuracy: 0.4542
Epoch 2/10
20/20 [=====] - 328s 16s/step - loss: 1.5368 - accuracy: 0.3820 - val_loss: 1.2988 - val_accuracy: 0.4560
Epoch 3/10
20/20 [=====] - 328s 16s/step - loss: 1.2082 - accuracy: 0.5133 - val_loss: 0.7975 - val_accuracy: 0.7556
Epoch 4/10
20/20 [=====] - 320s 16s/step - loss: 0.7632 - accuracy: 0.7169 - val_loss: 0.4164 - val_accuracy: 0.8596
Epoch 5/10
20/20 [=====] - 369s 18s/step - loss: 0.4474 - accuracy: 0.8353 - val_loss: 0.2072 - val_accuracy: 0.9333
Epoch 6/10
20/20 [=====] - 353s 18s/step - loss: 0.2799 - accuracy: 0.9058 - val_loss: 0.1510 - val_accuracy: 0.9422
Epoch 7/10
20/20 [=====] - 334s 17s/step - loss: 0.2022 - accuracy: 0.9300 - val_loss: 0.1470 - val_accuracy: 0.9493
Epoch 8/10
20/20 [=====] - 324s 16s/step - loss: 0.1586 - accuracy: 0.9462 - val_loss: 0.0886 - val_accuracy: 0.9689
Epoch 9/10
20/20 [=====] - 323s 16s/step - loss: 0.1576 - accuracy: 0.9456 - val_loss: 0.0918 - val_accuracy: 0.9618
Epoch 10/10
20/20 [=====] - 330s 16s/step - loss: 0.1476 - accuracy: 0.9498 - val_loss: 0.0693 - val_accuracy: 0.9716
59/59 [=====] - 37s 622ms/step - loss: 0.0865 - accuracy: 0.9669
```

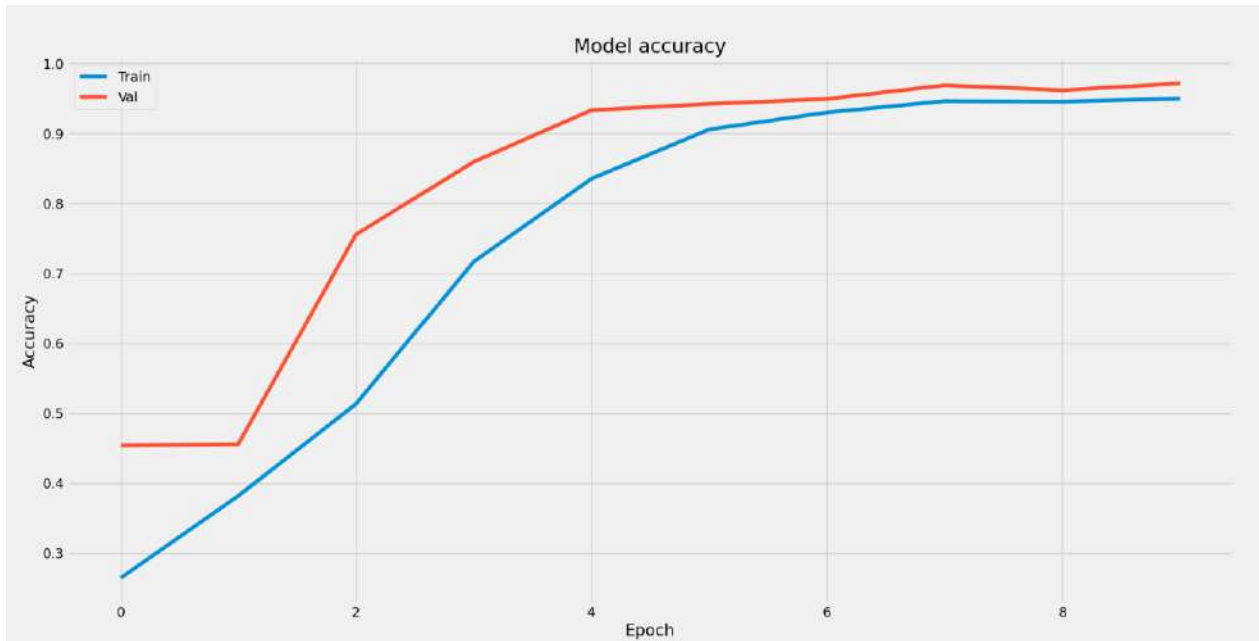
Εικόνα 4.18: Πληροφορίες επαναλήψεων και ακρίβεια 96.69% μοντέλου V3.1

Τα μοντέλα V3.2 και V3.3 παρουσίασαν ικανοποιητική ακρίβεια, περίπου 83% και 84% αντίστοιχα για ταξινομητή πολλαπλών κλάσεων και αρκετά μικρές απώλειες (Εικόνα 4.19). Ο συνολικός υπολογιστικός χρόνος του μοντέλου V3.2 ήταν περίπου 56 λεπτά (5:46 min/epoch) και του μοντέλου V3.3 περίπου 62 λεπτά (6:12 min/epoch).

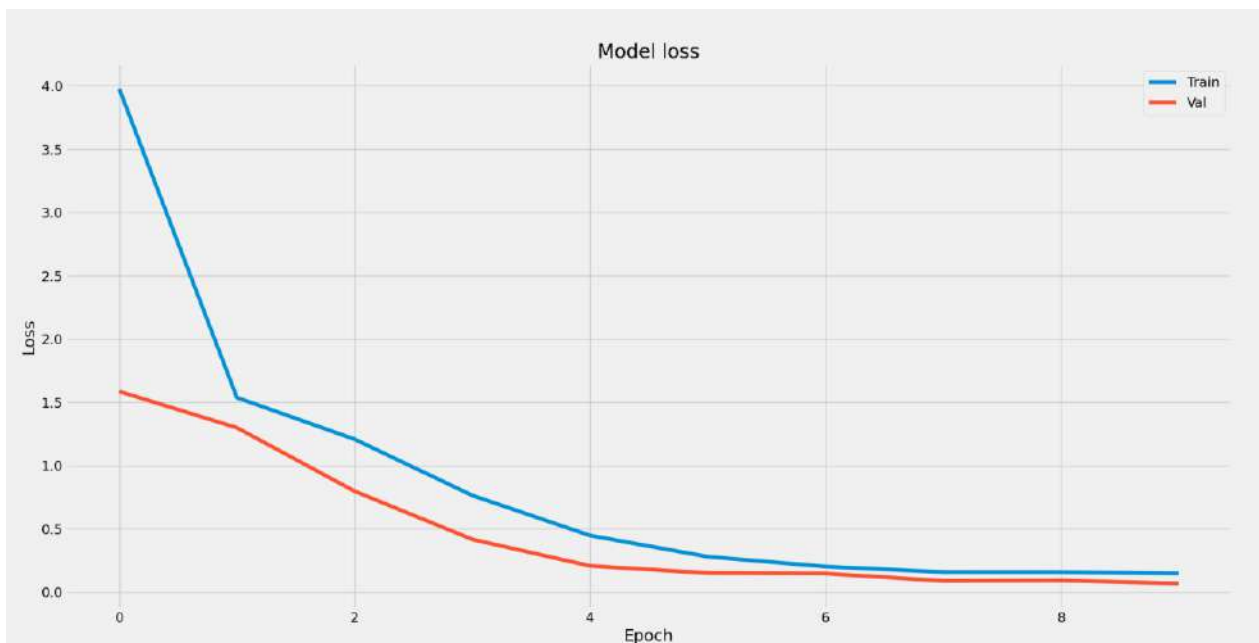
```
59/59 [=====] - 38s 647ms/step - loss: 0.4694 - accuracy: 0.8341
63/63 [=====] - 41s 644ms/step - loss: 0.3864 - accuracy: 0.8525
```

Εικόνα 4.19: Ακρίβεια 83.41% μοντέλου V3.2 και 85.25% μοντέλου V3.3

Στο Διάγραμμα 4.8 παρατηρήθηκαν απότομες μεταβολές των τιμών στις πρώτες επαναλήψεις λόγω μη ομαλής εισαγωγής των διαφορετικών επισημασμένων δεδομένων. Αντίθετα, στο Διάγραμμα 4.9 η απότομη πτώση της τιμής κατά την εκκίνηση της εκπαίδευσης λόγω εισαγωγής πολλών πληροφοριών στο μοντέλο ήταν αναμενόμενη. Σημειώνεται ότι τα αποτελέσματα εκπαίδευσης και επικύρωσης κατά την ολοκλήρωση του μοντέλου στην τελευταία επανάληψη παρουσίασαν σχεδόν ίσες τιμές.



Διάγραμμα 4.8: Ακρίβεια μοντέλου για δεδομένα εκπαίδευσης και επικύρωσης V3.1



Διάγραμμα 4.9: Απώλεια μοντέλου για δεδομένα εκπαίδευσης και επικύρωσης V3.1

Μητρώα σύγχυσης και αναφορά ταξινόμησης

Η ακρίβεια του μοντέλου επαληθεύτηκε με τα μητρώα σύγχυσης και την αναφορά ταξινόμησης, στα οποία παρατηρήθηκαν **ελάχιστες λανθασμένες προβλέψεις** (Εικόνα 4.20). Όλες οι κατηγορίες ουρών ανιχνεύτηκαν από το μοντέλο, με μικρότερο ποσοστό συνοχής ακρίβειας 84% της κλάσης Mandatory Lane Contraction (3), η οποία είχε το μικρότερο δείγμα 94 εικόνων. Άξιο αναφοράς ότι οι προβλέψεις του δείγματος μέτριου πλήθους (138 εικόνες) της κλάσης Turn (2), παρουσίασαν ποσοστό 99%.

[[549 5 0 6 3 3]	
[3 303 0 1 1 0]	
[1 0 123 10 4 0]	
[2 3 0 89 0 0]	
[7 3 1 0 502 0]	
[3 4 0 0 2 247]]	
[[[1293 16] [[1764 17]	
[17 549]] [5 89]]	
[[1552 15] [[1352 10]	
[5 303]] [11 502]]	
[[1736 1] [[1616 3]	
[15 123]] [9 247]]]	

	precision	recall	f1-score	support
0	0.97	0.97	0.97	566
1	0.95	0.98	0.97	308
2	0.99	0.89	0.94	138
3	0.84	0.95	0.89	94
4	0.98	0.98	0.98	513
5	0.99	0.96	0.98	256
accuracy			0.97	1875
macro avg	0.95	0.96	0.95	1875
weighted avg	0.97	0.97	0.97	1875

Εικόνα 4.20: (αριστερά) Μητρώα σύγχυσης μοντέλου V3.1, (δεξιά) Αναφορά ταξινόμησης μοντέλου V3.1

Πρόβλεψη ταξινόμησης τυχαίων εικόνων δείγματος

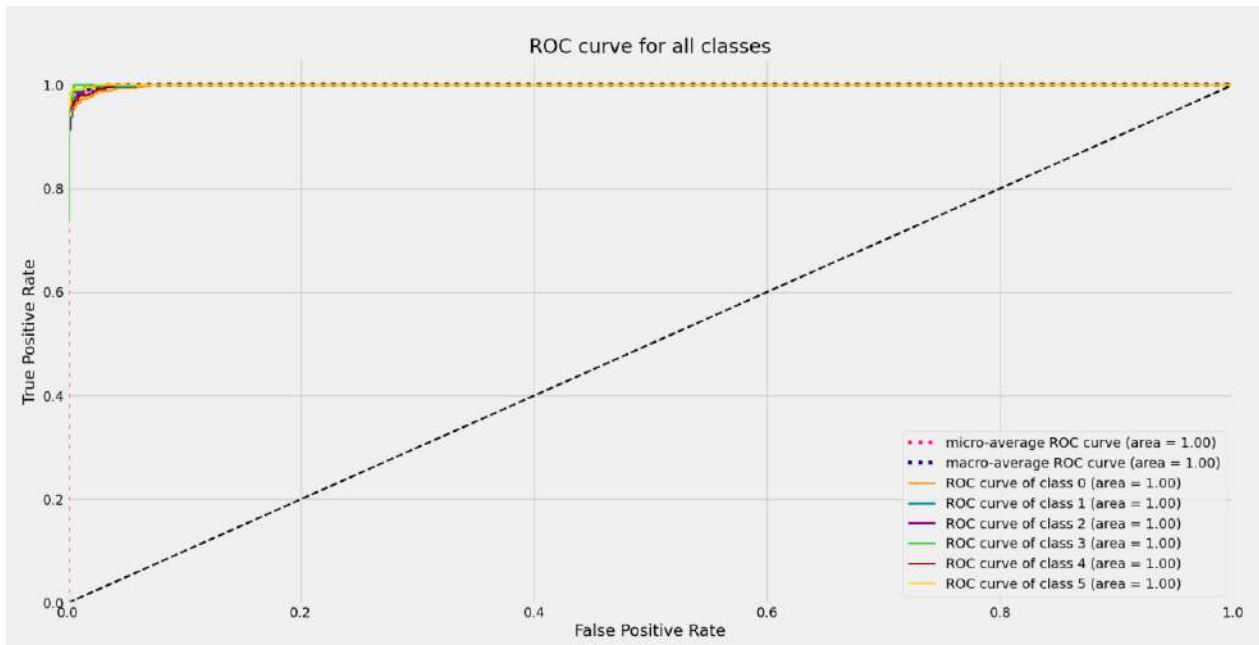
Με τη εφαρμογή του μοντέλου στις παρακάτω τυχαίες εικόνες του δείγματος δοκιμής, προέκυψαν **ακριβείς προβλέψεις υψηλού ποσοστού** ανίχνευσης ουράς:

Εικόνα 6501 - Κλάση (4): [Occupied_Intersection] 98.65%, [Turn] 0.8%, [No_Queue] 0.27%, [Stop] 0.19%, [Spillover_Effect] 0.09%, [Mandatory_Lane_Contraction] 0.01%

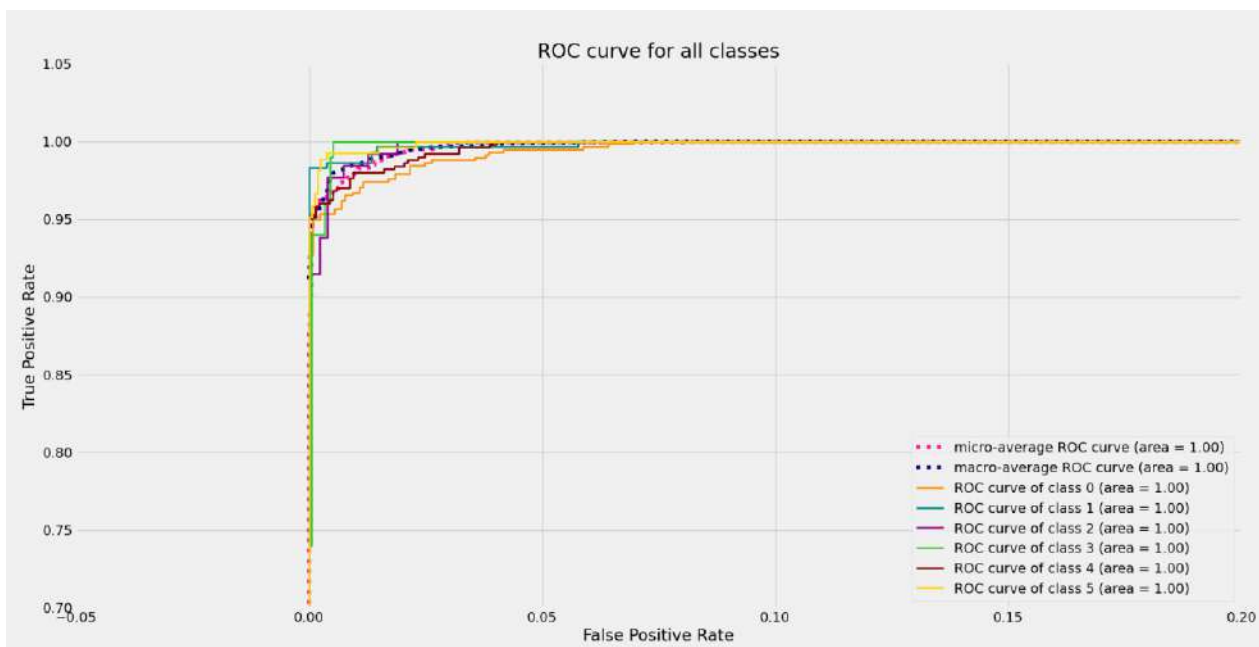
Εικόνα 4364 - Κλάση (0): [No_Queue] 99.51%, [Spillover_Effect] 0.48%, [Stop] 0.0%, [Occupied_Intersection] 0.0%, [Mandatory_Lane_Contraction] 0.0%, [Turn] 0.0%

Καμπύλες ROC και περιοχές AUC

Η **υψηλή ποιότητα του ταξινομητή** αξιολογήθηκε από τις περιοχές AUC και τις καμπύλες ROC (Διαγράμματα 4.10 και 4.11). Τα ποσοστά των περιοχών AUC όλων των κλάσεων παρουσίασαν τιμή ίση με 1, εφόσον σχεδόν ολόκληρο το τμήμα τους ταυτίζεται με τις τιμές των καμπυλών ROC και ένα μηδαμινό ποσοστό τους εμφανίστηκε κάτω από αυτές, όπως παρατηρήθηκε και από το διάγραμμα μεγέθυνσης. Επομένως, ο ταξινομητής χαρακτηρίζεται εξαιρετικός για τιμές μεταξύ 0.9 και 1.0 και οι προβλέψεις του είναι αξιόπιστες και με σχεδόν μηδενικό ποσοστό λανθασμένων προβλέψεων.



Διάγραμμα 4.10: Καμπύλες ROC και τιμές AUC για τις κλάσεις του μοντέλου V3.1



Διάγραμμα 4.11: Καμπύλες ROC και τιμές AUC για τις κλάσεις του μοντέλου V3.1 (μεγέθυνση διαγράμματος)

Μεταφορά μάθησης

Η μεταφορά μάθησης του ταξινομητή του Video 3.1 στις εικόνες των Video 1 και Video 2 παρουσίασε **ανεπαρκή αποτελέσματα** λόγω μεγάλης διαφοροποίησης των περιοχών καταγραφής και των ποσοστών εμφάνισης των κυκλοφοριακών ουρών (Εικόνες 4.21 και 4.22). Η μεταφορά μάθησης του ίδιου ταξινομητή στις εικόνες των Video 3.2 Video 3.3 παρουσίασε **μέτρια αποτελέσματα**, με αναγνώριση σχεδόν όλων των κλάσεων (Εικόνες 4.23 και 4.24). Πραγματοποιήθηκε πρόβλεψη κυρίως των κλάσεων παρόμοιου ποσοστού εμφάνισης στα ίδια σημεία καταγραφής του πειράματος, ενώ οι κλάσεις που είχαν καταγραφεί σε διαφορετικά σημεία καταγραφής δεν ανιχνεύτηκαν. Τα ποσοστά συνοχής ακρίβειας ήταν αρκετά μικρότερα από τα αναμενόμενα.

[[18 494 67 0 0 0]	
[6 253 37 0 0 0]	
[3 113 13 0 0 0]	
[4 87 9 0 0 0]	
[10 427 67 0 0 0]	
[10 223 34 0 0 0]]	
[[[1263 33] [[1775 0]	
[561 18]] [100 0]]	
[[235 1344] [[1371 0]	
[43 253]] [504 0]]	
[[[1532 214] [[1608 0]	
[116 13]] [267 0]]]	

	precision	recall	f1-score	support
0	0.35	0.03	0.06	579
1	0.16	0.85	0.27	296
2	0.06	0.10	0.07	129
3	0.00	0.00	0.00	100
4	0.00	0.00	0.00	504
5	0.00	0.00	0.00	267
accuracy			0.15	1875
macro avg	0.09	0.16	0.07	1875
weighted avg	0.14	0.15	0.06	1875

Εικόνα 4.21: (αριστερά) Μητρώα σύγχυσης δεδομένων Video 1 - μοντέλου V3.1, (δεξιά) Αναφορά ταξινόμησης δεδομένων Video 1 - μοντέλου V3.1

[[2 577 0 0 0 0]	
[2 294 0 0 0 0]	
[0 129 0 0 0 0]	
[0 100 0 0 0 0]	
[4 500 0 0 0 0]	
[1 266 0 0 0 0]]	
[[[1289 7] [[1775 0]	
[577 2]] [100 0]]	
[[7 1572] [[1371 0]	
[2 294]] [504 0]]	
[[[1746 0] [[1608 0]	
[129 0]] [267 0]]]	

	precision	recall	f1-score	support
0	0.22	0.00	0.01	579
1	0.16	0.99	0.27	296
2	0.00	0.00	0.00	129
3	0.00	0.00	0.00	100
4	0.00	0.00	0.00	504
5	0.00	0.00	0.00	267
accuracy			0.16	1875
macro avg	0.06	0.17	0.05	1875
weighted avg	0.09	0.16	0.05	1875

Εικόνα 4.22: (αριστερά) Μητρώα σύγχυσης δεδομένων Video 2 - μοντέλου V3.1, (δεξιά) Αναφορά ταξινόμησης δεδομένων Video 3.1 - μοντέλου V3.1

[[330 142 2 0 73 32]	
[158 72 0 0 48 18]	
[59 38 0 0 25 7]	
[46 24 0 0 19 11]	
[292 123 2 0 58 29]	
[148 73 0 0 34 12]]	
[[[593 703] [[1775 0]	
[249 330]] [100 0]]	
[[1179 400] [[1172 199]	
[224 72]] [446 58]]	
[[1742 4] [[1511 97]	
[129 0]] [255 12]]]	

	precision	recall	f1-score	support
0	0.32	0.57	0.41	579
1	0.15	0.24	0.19	296
2	0.00	0.00	0.00	129
3	0.00	0.00	0.00	100
4	0.23	0.12	0.15	504
5	0.11	0.04	0.06	267
accuracy			0.25	1875
macro avg	0.13	0.16	0.14	1875
weighted avg	0.20	0.25	0.21	1875

Εικόνα 4.23: (αριστερά) Μητρώα σύγχυσης δεδομένων Video 3.2 - μοντέλου V3.1, (δεξιά) Αναφορά ταξινόμησης δεδομένων Video 3.1 - μοντέλου V3.1

[[532 22 0 0 25 0]	
[263 16 1 0 16 0]	
[122 3 1 0 3 0]	
[87 4 0 0 9 0]	
[450 24 3 0 27 0]	
[249 11 1 0 6 0]]	
[[[125 1171] [[1775 0]	
[47 532]] [100 0]]	
[[1515 64] [[1312 59]	
[280 16]] [477 27]]	
[[1741 5] [[1608 0]	
[128 1]] [267 0]]]	

	precision	recall	f1-score	support
0	0.31	0.92	0.47	579
1	0.20	0.05	0.09	296
2	0.17	0.01	0.01	129
3	0.00	0.00	0.00	100
4	0.31	0.05	0.09	504
5	0.00	0.00	0.00	267
accuracy			0.31	1875
macro avg	0.17	0.17	0.11	1875
weighted avg	0.22	0.31	0.18	1875

Εικόνα 4.24: (αριστερά) Μητρώα σύγχυσης δεδομένων Video 3.3 - μοντέλου V3.1, (δεξιά) Αναφορά ταξινόμησης δεδομένων Video 3.1 - μοντέλου V3.1

4.6.4 Συνδυασμός Βάσεων Δεδομένων (Video ALL)

Για το μοντέλο VALL, με συνδυασμό όλων των βάσεων δεδομένων, χρησιμοποιήθηκαν συνολικά 38668 καρέ εικόνων, από τα οποία 29001 για εκπαίδευση και 9667 για δοκιμές. Από το συνδυασμό των επισημάνσεων παρουσιάστηκε σχηματισμός ουρών των κατηγοριών: Stop (11), Turn (30), Mandatory Lane Contraction (16), Occupied Intersection (20), Spillover Effect (11).

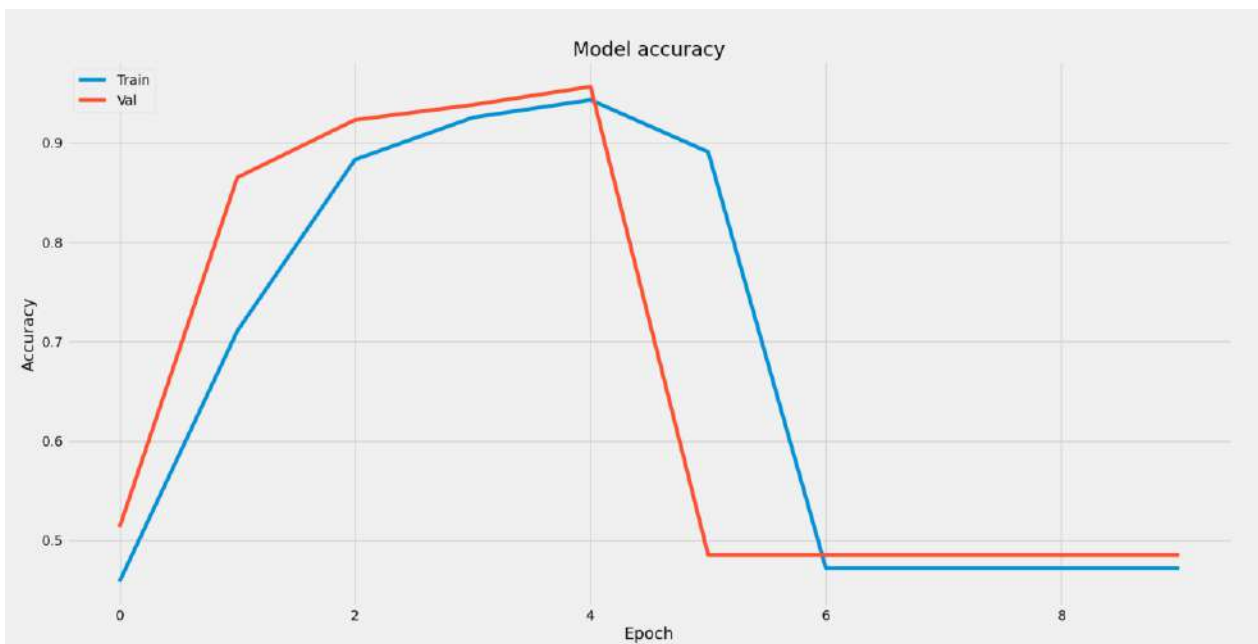
Ακρίβεια μοντέλου

Το μοντέλο παρουσίασε **καθόλου ικανοποιητική ακρίβεια**, περίπου 48%, για ταξινομητή πολλαπλών κλάσεων και **σημαντική απώλεια** (Εικόνα 4.25). Παρατηρήθηκε τεράστια διακύμανση κατά την εκπαίδευση και την επικύρωση. Ο συνολικός υπολογιστικός χρόνος του μοντέλου ήταν περίπου 594min (59:20 min/epoch), υπερβολικά μεγάλη διάρκεια για τα δεδομένα και τις παραμέτρους που χρησιμοποιήθηκαν.

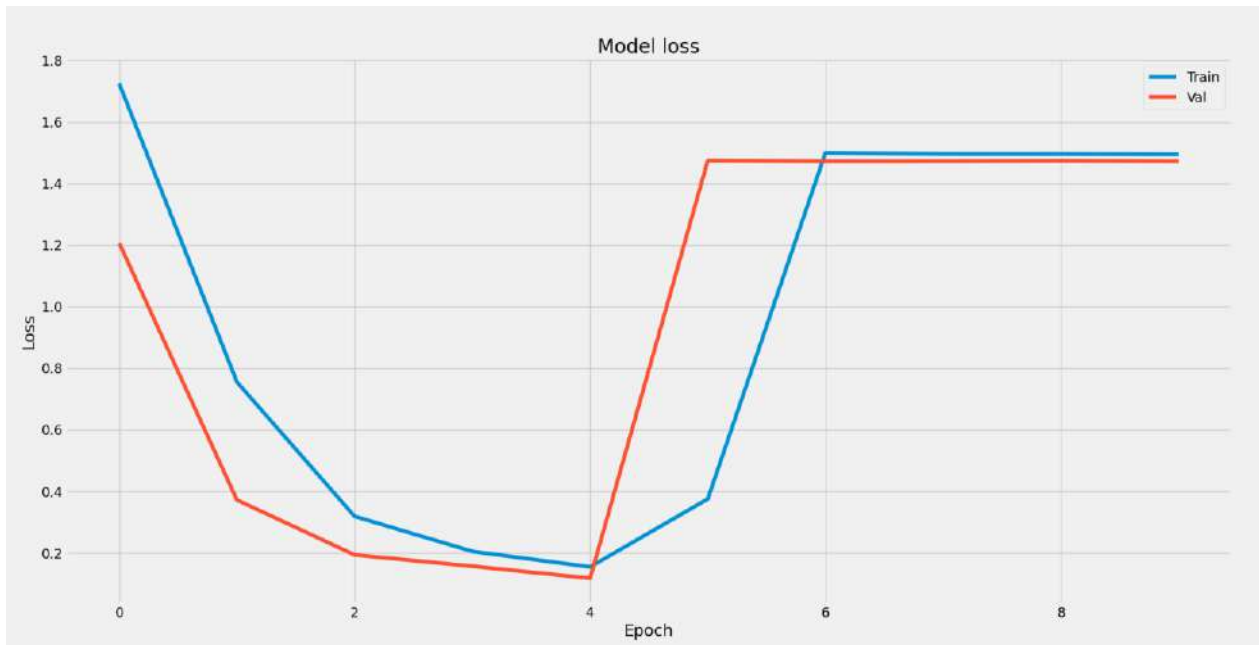
```
Epoch 1/10
104/104 [=====] - 4406s 42s/step - loss: 1.7245 - accuracy: 0.4588 - val_loss: 1.2066 - val_accuracy: 0.5135
Epoch 2/10
104/104 [=====] - 4827s 46s/step - loss: 0.7554 - accuracy: 0.7106 - val_loss: 0.3723 - val_accuracy: 0.8650
Epoch 3/10
104/104 [=====] - 3726s 36s/step - loss: 0.3191 - accuracy: 0.8830 - val_loss: 0.1939 - val_accuracy: 0.9229
Epoch 4/10
104/104 [=====] - 3393s 33s/step - loss: 0.2050 - accuracy: 0.9254 - val_loss: 0.1570 - val_accuracy: 0.9381
Epoch 5/10
104/104 [=====] - 3301s 32s/step - loss: 0.1547 - accuracy: 0.9431 - val_loss: 0.1179 - val_accuracy: 0.9564
Epoch 6/10
104/104 [=====] - 3184s 31s/step - loss: 0.3756 - accuracy: 0.8907 - val_loss: 1.4751 - val_accuracy: 0.4853
Epoch 7/10
104/104 [=====] - 3237s 31s/step - loss: 1.4999 - accuracy: 0.4722 - val_loss: 1.4730 - val_accuracy: 0.4853
Epoch 8/10
104/104 [=====] - 3221s 31s/step - loss: 1.4963 - accuracy: 0.4722 - val_loss: 1.4731 - val_accuracy: 0.4853
Epoch 9/10
104/104 [=====] - 3219s 31s/step - loss: 1.4961 - accuracy: 0.4722 - val_loss: 1.4744 - val_accuracy: 0.4853
Epoch 10/10
104/104 [=====] - 3083s 30s/step - loss: 1.4950 - accuracy: 0.4722 - val_loss: 1.4729 - val_accuracy: 0.4853
303/303 [=====] - 342s 1s/step - loss: 1.4799 - accuracy: 0.4804
```

Εικόνα 4.25: Πληροφορίες επαναλήψεων και ακρίβεια 48,04% μοντέλου VALL

Στο Διάγραμμα 4.12 παρατηρήθηκε απότομη πτώση και στο Διάγραμμα 4.13 απότομη αύξηση των τιμών στην πέμπτη επανάληψη. Οι ακραίες μεταβολές οφείλονται στην εισαγωγή των διαφορετικών δεδομένων των πειραμάτων, τα οποία προκάλεσαν **υπερφόρτωση του μοντέλου** με πληροφορίες και **εγκατάλειψη της διαδικασίας**.



Διάγραμμα 4.12: Ακρίβεια μοντέλου για δεδομένα εκπαίδευσης και επικύρωσης VALL



Διάγραμμα 4.13: Απώλεια μοντέλου για δεδομένα εκπαίδευσης και επικύρωσης VALL

Μητρώα σύγχυσης και αναφορά ταξινόμησης

Η εγκατάλειψη και μη ολοκλήρωση εκπαίδευσης του μοντέλου επαληθεύτηκαν με τα μητρώα σύγχυσης και την αναφορά ταξινόμησης, στα οποία παρατηρήθηκαν **μηδενικές προβλέψεις** για τις κλάσεις (Εικόνα 4.26). Συγκεκριμένα πραγματοποιήθηκε ταξινόμηση μόνο της κλάσης No Queue (0) με ποσοστό συνοχής ακρίβειας 48%.

[[4644 0 0 0 0 0]					
[849 0 0 0 0 0]					
[1620 0 0 0 0 0]					
[651 0 0 0 0 0]					
[1412 0 0 0 0 0]					
[491 0 0 0 0 0]]					
[[0 5023]			[[9016 0]		
[0 4644]]			[651 0]]		
[[8818 0]			[[8255 0]		
[849 0]]			[1412 0]]		
[[8047 0]			[[9176 0]		
[1620 0]]			[491 0]]]		

	precision	recall	f1-score	support
0	0.48	1.00	0.65	4644
1	0.00	0.00	0.00	849
2	0.00	0.00	0.00	1620
3	0.00	0.00	0.00	651
4	0.00	0.00	0.00	1412
5	0.00	0.00	0.00	491
accuracy			0.48	9667
macro avg	0.08	0.17	0.11	9667
weighted avg	0.23	0.48	0.31	9667

Εικόνα 4.26: (αριστερά) Μητρώα σύγχυσης μοντέλου VALL, (δεξιά) Αναφορά ταξινόμησης μοντέλου VALL

Πρόβλεψη ταξινόμησης τυχαίων εικόνων δείγματος

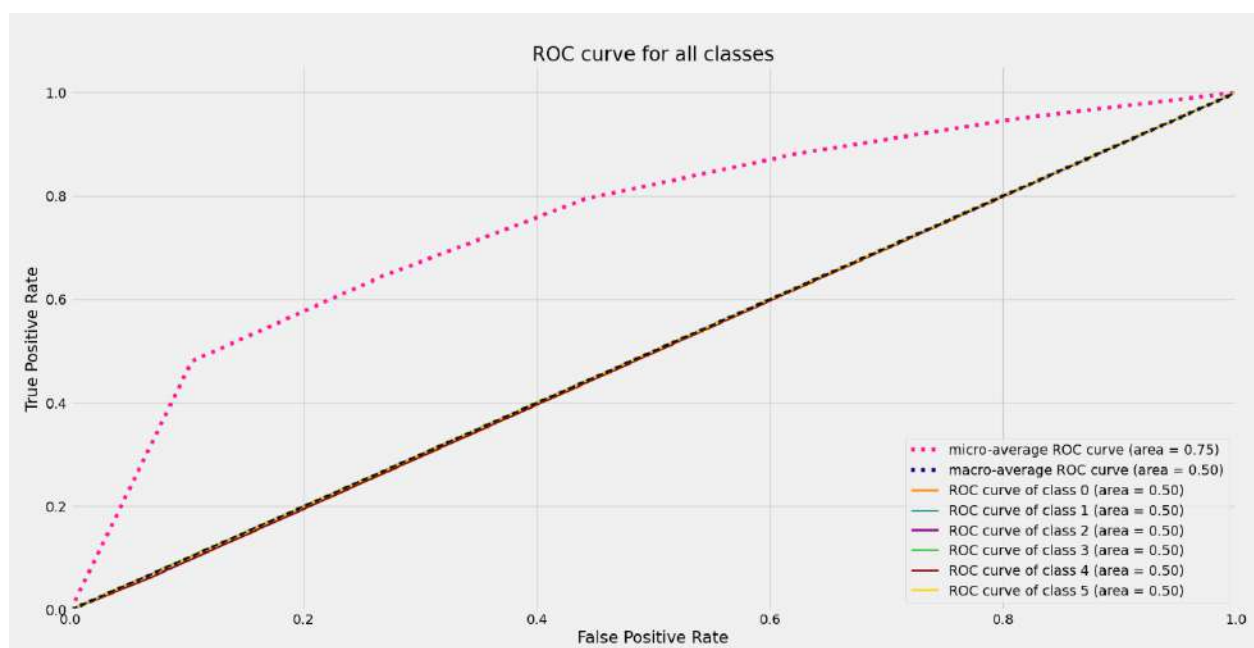
Με τη εφαρμογή του μοντέλου στις παρακάτω τυχαίες εικόνες του δείγματος δοκιμής, προέκυψαν **αδύναμες προβλέψεις χαμηλού ποσοστού** ανίχνευσης ουράς:

Εικόνα 14184 - Κλάση (0): [No_Queue] 43.17%, [Occupied_Intersection] 18.92%, [Stop] 16.1%, [Turn] 12.2%, [Spillover_Effect] 7.01%, [Manatory_Lane_Contraction] 2.59%

Εικόνα 25186 - Κλάση (0): [No_Queue] 46.96%, [Turn] 17.06%, [Occupied_Intersection] 15.8%, [Stop] 8.37%, [Manatory_Lane_Contraction] 6.83%, [Spillover_Effect] 4.99%

Καμπύλες ROC και περιοχές AUC

Η **χαμηλή ποιότητα του ταξινομητή** αξιολογήθηκε από τις περιοχές AUC και τις καμπύλες ROC (Διάγραμμα 4.14). Τα ποσοστά των περιοχών AUC όλων των κλάσεων παρουσίασαν τιμή ίση με 0.5, εφόσον δεν πραγματοποιήθηκε ανίχνευση κλάσεων. Επομένως, ο ταξινομητής χαρακτηρίζεται ακατάλληλος για τιμές μεταξύ 0.5 και 0.6 και οι προβλέψεις αναξιόπιστες και με τεράστιο ποσοστό λανθασμένων προβλέψεων.



Διάγραμμα 4.14: Καμπύλες ROC και τιμές AUC για τις κλάσεις του μοντέλου VALL

5. ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΡΟΤΑΣΕΙΣ

5.1 ΕΙΣΑΓΩΓΗ

Η αντιμετώπιση κυκλοφοριακών συμφορήσεων αποτελεί αντικείμενο μείζονος σημασίας και απαιτεί λεπτομερή ανάλυση για τη βελτίωση των κυκλοφοριακών ικανοτήτων κορεσμένων οδικών συστημάτων. Συγκεκριμένα, θα πρέπει να αναλυθούν οι κυκλοφοριακοί παράγοντες οι οποίοι προκαλούν συμφορήσεις μεγάλου επιπέδου και να αντιμετωπιστούν μεμονωμένα. Ένας από τους σημαντικότερους παράγοντες προς ανάλυση είναι οι κυκλοφοριακές ουρές που παρατηρούνται σε σηματοδοτούμενους κόμβους. Το αντικείμενο της παρούσας Διπλωματικής Εργασίας είναι η **δημιουργία ενός αλγόριθμου υπολογιστικής όρασης με στόχο την αναγνώριση και κατηγοριοποίηση κυκλοφοριακών ουρών**. Ο αλγόριθμος θα πρέπει να ανιχνεύει τυχόν ουρές που σχηματίζονται σε μία περιοχή και να πραγματοποιεί ποσοστιαίες προβλέψεις για το είδος τους.

Τα στοιχεία τα οποία χρησιμοποιήθηκαν για την εκπαίδευση του αλγόριθμου λήφθηκαν από ένα πείραμα βιντεοσκόπησης μεγάλης έκτασης, το οποίο πραγματοποιήθηκε σε μια ευρύτερη περιοχή στο κέντρο της Αθήνας, όπου καταγράφηκε η κυκλοφορία με χρήση ΣμηΕΑ. Στη συνέχεια, ακολούθησε η επεξεργασία των δεδομένων στη πλατφόρμα Supervisely και σε περιβάλλον γλώσσας προγραμματισμού Python, στο οποίο πραγματοποιήθηκε η εκπαίδευση και η τελική διαμόρφωση του αλγόριθμου υπολογιστικής όρασης, καθώς και οι δοκιμές για την πρόβλεψη κυκλοφοριακών ουρών.

Έπειτα από μια σειρά δοκιμών και τροποποιήσεων του μοντέλου, αναπτύχθηκε ο αλγόριθμος βαθιάς μάθησης ο οποίος προβλέπει την εμφάνιση κυκλοφοριακής ουράς από ψηφιακές εικόνες και την κατηγορία που τους αντιστοιχεί.

5.2 ΒΑΣΙΚΑ ΣΥΜΠΕΡΑΣΜΑΤΑ

Στο παρόν υποκεφάλαιο συνοψίζονται τα συμπεράσματα που προέκυψαν από τα διάφορα στάδια εκπόνησης της παρούσας Διπλωματικής Εργασίας, μέσω της σύνθεσης των αποτελεσμάτων των προηγούμενων κεφαλαίων.

Η συγκεκριμένη Διπλωματική Εργασία αποτελεί μία από τις πρώτες έρευνες που **πραγματοποιεί προβλέψεις κατηγοριοποίησης κυκλοφοριακών ουρών** για την αντιμετώπιση και αποφυγή κυκλοφοριακών συμφορήσεων, με χρήση υπολογιστικής όρασης και καταγεγραμμένων δεδομένων από ΣμηΕΑ. Οι περισσότερες έρευνες πάνω στο αντικείμενο αντιμετώπισης συμφορήσεων εστιάζουν περισσότερο στην εξαγωγή και ανάλυση των κυκλοφοριακών μεταβλητών των οχημάτων με χρήση παρόμοιας μεθοδολογίας και βιντεοσκοπημένων δεδομένων. Όμως, για την αποτελεσματική αντιμετώπιση του προβλήματος και τη δημιουργία ενός αυτόνομου συστήματος παρακολούθησης της κυκλοφορίας, απαιτείται η μελέτη ποικίλων κυκλοφοριακών συνθηκών σε βάθος.

Τα συνελκτικα νευρωνικα δίκτυα βαθιάς μάθησης αποτελούν **κατάλληλη μέθοδο για τη διαμόρφωση των αλγόριθμων ταξινόμησης ουρών** καθώς είναι σε θέση να διαχειριστούν πολύπλοκα επισημασμένα δεδομένα μεγάλου πλήθους. Η αναγνώριση τέτοιων κυκλοφοριακών συνθηκών σε πραγματικούς χρόνους αποτελεί τεράστιο πλεονέκτημα για κυκλοφοριακές έρευνες, η οποία μπορεί να επιτευχθεί με το συνδυασμό νευρωνικών δικτύων CNN και ισχυρής υπολογιστική ισχύος.

Ο αλγόριθμος είναι εξαιρετικής ακρίβειας για δεδομένα ίδιων χαρακτηριστικών, όπως προκύπτει από τις επαναλήψεις των μοντέλων και τις τελικές τιμές των δεδομένων εκπαίδευσης και επικύρωσης των διαγραμμάτων ακρίβειας και απώλειας, οι οποίες συγκλίνουν. Αντιθέτως, για πολλαπλά δεδομένα διαφορετικών χαρακτηριστικών, ο αλγόριθμος δεν είναι καθόλου ικανοποιητικός καθώς το μοντέλο υπερφορτώνεται με πληροφορίες και διακόπτει τη διαδικασία εκπαίδευσης.

Ο πολύ ικανοποιητικός υπολογιστικός χρόνος των μοντέλων είναι ένα από τα βασικά πλεονεκτήματα του αλγόριθμου καθώς επιτυγχάνεται μεγάλη ακρίβεια προβλέψεων με **χρήση βάσεων δεδομένων μεγάλου πλήθους**. Με τη χρήση μεγάλων βάσεων, ο αλγόριθμος δεν περιορίζεται αποκλειστικά στην αναγνώριση της ύπαρξης ή μη κυκλοφοριακής ουράς, αλλά στην ανίχνευση του είδους κάθε ουράς. Βάσει αυτής της πολυπλοκότητας ο αλγόριθμος απαιτεί περισσότερο υπολογιστικό χρόνο, χωρίς αυτός να υπερβαίνει την αναμενόμενη διάρκεια εκπαίδευσης του μοντέλου.

Η υψηλή ποιότητα των προβλέψεων του αλγόριθμου παραμένει ακόμη και στις περιπτώσεις χρήσης μικρού δείγματος δοκιμών. Ο αλγόριθμος είναι ικανός να αναγνωρίσει όλες τις κατηγορίες ουρών που έχουν οριστεί και να προσφέρει έως και **μηδενικές λανθασμένες προβλέψεις**, για κατηγορίες ουρών μεγάλου ποσοστού εμφάνισης στα δεδομένα δοκιμών. Ειδικότερα, οι ποσοστιαίες εκτιμήσεις που πραγματοποιεί ο αλγόριθμος είναι αξιόπιστες καθώς ακόμη και στις λανθασμένες ταξινομήσεις ουρών, παρουσιάζονται ελάχιστες αποκλίσεις από την πραγματικότητα.

Η χρήση των γνώσεων του μοντέλου του αλγόριθμου σε έναν αλγόριθμο διαφορετικής βάσης δεδομένων με **μεταφορά μάθησης παρουσιάζει μη ικανοποιητικά, αλλά ενθαρρυντικά αποτελέσματα**. Η ποιότητα των προβλέψεων είναι αναμενόμενη καθώς το επαναχρησιμοποιούμενο μοντέλο έχει εκπαιδευτεί με διαφορετικά καρέ εικόνων, όμως ο αλγόριθμος μπορεί και ανιχνεύει την πλειοψηφία των κοινών κατηγοριών ουρών που εμφανίζονται και στα δύο βίντεο. Αντιθέτως, για διαφορετικά καρέ εικόνων και ποσοστών εμφάνισης συγκεκριμένων κατηγοριών, ο αλγόριθμος είναι ικανός να αναγνωρίσει κυρίως την ύπαρξη ή μη κυκλοφοριακής ουράς.

5.3 ΠΡΟΤΑΣΕΙΣ ΓΙΑ ΠΕΡΑΙΤΕΡΩ ΕΡΕΥΝΑ

Για την περαιτέρω διερεύνηση του αντικειμένου της παρούσας Διπλωματικής Εργασίας, ενδιαφέρον θα παρουσίαζε να διερευνηθούν τα παρακάτω:

Ανίχνευση περισσότερων χαρακτηριστικών κυκλοφοριακής ουράς

Ο αλγόριθμος μπορεί να εκπαιδευτεί με τέτοιο τρόπο ώστε να χρησιμοποιεί όλες τις πληροφορίες των επισημασμένων δεδομένων. Συγκεκριμένα, μπορεί εκτός της κατηγοριοποίησης να πραγματοποιεί αναγνώριση της θέσης της κυκλοφοριακής ουράς την οποία ανιχνεύει στην σταθερή, κατά τη διάρκεια βιντεοσκόπησης, περιοχή κάλυψης του αεροσκάφους. Ακόμη, μπορεί να πραγματοποιεί εκτίμηση του μεγέθους κάθε ουράς βάσει του σχήματος με το οποίο έχει επισημανθεί, η οποία όμως προϋποθέτει την λεπτομερή επεξεργασία των στοιχείων κατά το στάδιο επισήμανσης, ώστε να ορίζεται κατάλληλα το μέγεθος της ουράς ανάλογα τον αριθμό των εμπλεκόμενων οχημάτων.

Χρήση δεδομένων μικτών συνθηκών

Οι καιρικές συνθήκες μπορούν να προκαλέσουν μείωση της ορατότητας κατά τη διάρκεια συλλογής των δεδομένων μέσω βιντεοσκόπησης. Με τον τρόπο αυτό μπορούν να δημιουργηθούν βάσεις δεδομένων εικόνων με μικτές συνθήκες, αυξημένης και μειωμένης ορατότητας για την ίδια περιοχή κάλυψης. Ο αλγόριθμος θα εκτίθεται σε πιο πολύπλοκες καταστάσεις ανίχνευσης ουράς καθώς η μειωμένη ευκρίνεια των εικόνων λόγω δυσμενών καιρικών συνθηκών και θα αντιμετωπίζει το πρόβλημα ταξινόμησης με μικρότερη ακρίβεια. Ενδιαφέρον παρουσιάζει και η χρήση μίας βάσης εικόνων μόνο μειωμένης ευκρίνειας, για τη διαπίστωση της ποιότητας του αλγόριθμου στη χρήση του σε αντίστοιχες περιπτώσεις καταγραφής της κυκλοφορίας και ανίχνευσης ουρών σε πραγματικό χρόνο.

Εκπαίδευση με μεγάλη βάση δεδομένων διαφορετικών σημείων

Στην παρούσα Διπλωματική Εργασία το μοντέλο του αλγόριθμου της ενιαίας βάσης δεδομένων από τον συνδυασμό όλων των δεδομένων, για τη χρήση βάσης αρκετά μεγάλου μεγέθους και αυξημένης πολυπλοκότητας, υπερφορτώθηκε και δεν ολοκλήρωσε τη διαδικασία εκπαίδευσης. Ο έλεγχος της ακρίβειας των προβλέψεων του αλγόριθμου για κυκλοφοριακές ουρές σε διαφορετικές περιοχές καταγραφής ταυτόχρονα, μπορεί να πραγματοποιηθεί με τη χρήση ενός μικρότερου μεγέθους με επιλογή συγκεκριμένων εικόνων από τις μεμονωμένες βάσεις, οι οποίες να προσφέρουν μία αυξημένη επιθυμητή πολυπλοκότητα. Ένας τέτοιος αλγόριθμος ικανός να ανιχνεύει σύνθετα δεδομένα εικόνων και επισημάνσεων, μπορεί να αξιοποιηθεί για πειράματα σε πραγματικό χρόνο από σμήνος ΣμηΕΑ σε μεγαλύτερες εξεταζόμενες περιοχές.

Μεταφορά μάθησης για διαφορετικές γωνίες καταγραφής και χρονικές στιγμές

Η αύξηση της συνοχής ακρίβειας των προβλέψεων αλγόριθμων με μεταφορά μάθησης και δεδομένα ίδιου σημείου καταγραφής, αλλά διαφορετικής γωνίας λήψης, μπορεί να πραγματοποιηθεί με τροποποίηση των εικόνων με κώδικα στην Python και συγκεκριμένα με την περιστροφή των αξόνων, ώστε να ταυτίζονται τα εξεταζόμενα οδικά τμήματα. Επιπλέον, μπορεί να χρησιμοποιηθεί βάση καταγεγραμμένων δεδομένων ίδιου σημείου και γωνίας λήψης διαφορετικής χρονικής στιγμής, για τον ποιοτικό έλεγχο του αλγόριθμου και τη χρήση του σε πειράματα πραγματικού χρόνου ίδιου αντικειμένου και εξεταζόμενης περιοχής.

BIBΛΙΟΓΡΑΦΙΑ

1. Barmounakis, E. N., Vlahogianni, E. I., & Golias, J. C. (2016). Unmanned Aerial Aircraft Systems for transportation engineering: Current practice and future challenges. *International Journal of Transportation Science and Technology*, 5(3), 111-122.
2. Barmounakis, E., & Geroliminis, N. (2019). Utilizing a swarm of drones for large-scale traffic measurements.
3. Benjdira, B., Khursheed, T., Koubaa, A., Ammar, A., & Ouni, K. (2018). Car Detection using Unmanned Aerial Vehicles: Comparison between Faster R-CNN and YOLOv3.
4. Brownlee, J. (2017). A Gentle Introduction to Transfer Learning for Deep Learning. Retrieved February 6, 2021, from <https://machinelearningmastery.com/transfer-learning-for-deep-learning/>
5. Brownlee, J. (2020). 4Types of Classification Tasks in Machine Learning. Retrieved February 6, 2021, from <https://machinelearningmastery.com/types-of-classification-in-machine-learning/>
6. Chakraborty, P., Adu-Gyamfi, Y. O., Poddar, S., Ahsani, V., Sharma, A., & Sarkar, S. (2018). Traffic congestion detection from camera images using deep convolution neural networks. *Transportation Research Record*, 2672(45), 222-231.
7. EU Transport Scoreboard. (2017). https://ec.europa.eu/transport/facts-fundings/scoreboard/compare/energy-union-innovation/road-congestion_en#2017
8. Janga, V., Kumar V., S., Enugala, V. K. (2020). Advanced machine learning-based implementation patterns for computer vision and real-time multimedia applications. *Materials Today: Proceedings*.
9. Kaufmann, S., Kerner, B. s., Hubert Rehborn, M. K., & Klenov, S. L. (2018). Aerial observations of moving synchronized flow patterns in over-saturated city traffic.
10. Khan, N. A., Jhanjhi, N. Z., Brohi, S. N., Usmani, R. S. A., Nayyar, A. (2020). Smart traffic monitoring system using Unmanned Aerial Vehicles (UAVs). *Computer Communications*, 157, 434-443.
11. Lee, S., Xie, K., Ngoduy, D., Keyvan-Ekbatani, M. (2019). An advanced deep learning approach to real-time estimation of lane-based queue lengths at a signalized junction. *Transportation Research Part C: Emerging Technologies*, 109, 117-136.
12. Mysla, V. (2020). Machine Learning in 10 minutes. Retrieved February 6, 2021, from <https://medium.datadriveninvestor.com/machine-learning-in-10-minutes-354d83e5922e>

13. Nielsen, M. A. (2015). Neural networks and deep learning (Vol. 25). San Francisco, CA: Determination press. Accessed at: <http://neuralnetworksanddeeplearning.com/chap1.html>
14. Puri, A., Valavanis, K., Kontitsis, M., 2007. Statistical profile generation for traffic monitoring using realtime UAV based video data. 2007 Mediterr. Conf. Control Autom. 1–6.
15. Rajashekhara, M. R. (2014). Characteristics of Development and Dissipation of Queue Length Under Mixed Traffic Conditions. International Journal of Engineering Research & Technology, 3(7).
16. Saha, S. (2018). A Comprehensive Guide to Convolutional Neural Networks - the ELI5 way. Retrieved February 6, 2021, from <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>
17. Thomas, A. (2017). Keras tutorial - build a convolutional neural network in 11 lines. Retrieved February 6, 2021, from <https://adventuresinmachinelearning.com/keras-tutorial-cnn-11-lines/>
18. TomTom Traffic Index. (2019). https://www.tomtom.com/en_gb/traffic-index/ranking/?country=AT,BE,BG,CZ,DK,EE,FI,FR,DE,GR,HU,IS,IE,IT,LV,LT,LU,NL,NO,PL,PT,RO,RU,SK,SI,ES,SE,CH,TR,UA,UK
19. U.S. Department of Transportation. (2020). Traffic Congestion and Reliability: Trends and Advanced Strategies for Congestion Mitigation. Retrieved January 18, 2021, from https://ops.fhwa.dot.gov/congestion_report/chapter3.htm?fbclid=IwAR0SXvYwTxxukWmGRqrpmwP_HSe4rNLDZ_ng6dkGNXyPqiq_zoBDnN1a5zo
20. Vlahogianni, E. I., Del Ser, J., Kepaptsoglou, K., & Laña, I. (2021). Model Free Identification of Traffic Conditions Using Unmanned Aerial Vehicles and Deep Learning, Journal of Big Data Analytics in Transportation, in press.
21. Wang, Z. (2013). Queue Storage Design for Metered On-Ramps. International Journal of Transportation Science and Technology, 2(1), 47-64.
22. Wang, L., Chen, F., & Yin, H. (2016). Detecting and tracking vehicles in traffic by unmanned aerial vehicles. Automation in construction, 72, 294-308.
23. Wilson, A. (2019). A Brief Introduction to Supervised Learning. Retrieved February 6, 2021, from <https://towardsdatascience.com/a-brief-introduction-to-supervised-learning-54a3e3932590>

24. Yang, J. (2017). ReLU and Softmax Activation Functions. Retrieved February 6, 2021, from <https://github.com/Kulbear/deep-learning-nano-foundation/wiki/ReLU-and-Softmax-Activation-Functions>
25. Κοκκολάκης Γ., Σπηλιώτης Ι. (2008) “Θεωρία Πιθανοτήτων και Στατιστική με Εφαρμογές”, Εκδόσεις Συμεών, Αθήνα 2010 Emerging Technologies 16.3: 320-331.
26. Φραντζεσκάκης, Ι. Μ., Ι. Κ. Γκόλιας & Μ. Χ. Πιτσιάβα–Λατινοπούλου (2009) "Κυκλοφοριακή Τεχνική." Εκδόσεις Παπασωτηρίου, Αθήνα.