ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΠΟΛΙΤΙΚΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΜΕΤΑΦΟΡΩΝ ΚΑΙ ΣΥΓΚΟΙΝΩΝΙΑΚΗΣ ΥΠΟΔΟΜΗΣ

Πρότυπα Υπολογιστικής Όρασης για την Αναγνώριση Κυκλοφοριακών Συνθηκών από Μη Επανδρωμένο Αεροσκάφος

Διπλωματική Εργασία

Γεώργιος Δάβος

Επιβλέπουσα Καθηγήτρια: Ελένη Βλαχογιάννη
Επίκουρη Καθηγήτρια Σχολής Πολιτικών Μηχανικών

Αθήνα, Μάρτιος 2019

ΕΥΧΑΡΙΣΤΙΕΣ

Με την ολοκλήρωση της διπλωματικής μου εργασίας θα ήθελα να ευχαριστήσω όσους με υποστήριξαν και με βοήθησαν να κλείσω αυτόν τον σημαντικό κύκλο της ζωής μου.

Αρχικά θα ήθελα να ευχαριστήσω την επιβλέπουσα καθηγήτρια μου, κυρία Ελένη Βλαχογιάννη για την δυνατότητα που μου προσέφερε να εκπονήσω ένα θέμα διπλωματικής εργασίας κοντά στα ενδιαφέροντα μου. Εκτιμάω την βοήθεια και καθοδήγηση που μου προσέφερε σε κάθε στάδιο της διπλωματικής εργασίας. Θα ήθελα επίσης να ευχαριστήσω και τον Διδάκτορα Εμμανουήλ Μπαρμπουνάκη ο οποίος ήταν διαθέσιμος να μου προσφέρει οδηγίες και καθοδήγηση όποτε τον χρειάστηκα.

Τέλος, επιθυμώ να ευχαριστήσω όλους τους κοντινούς μου ανθρώπους που ήταν δίπλα μου σε κάθε βήμα και κυρίως την οικογένεια μου που με στήριξε όλα αυτά τα χρόνια στην προσπάθεια μου.

Πρότυπα Υπολογιστικής Όρασης για την Αναγνώριση Κυκλοφοριακών Συνθηκών από Μη Επανδρωμένο Αεροσκάφος

Γεώργιος Δάβος

Επιβλέπουσα Καθηγήτρια: Ελένη Βλαχογιάννη

ΣΥΝΟΨΗ

Η ραγδαία ανάπτυξη της τεχνολογίας των μη επανδρωμένων αεροσκαφών (Unmanned Aerial Vehicles -UAV ή Drones) τα τελευταία χρόνια έχει οδηγήσει σε σημαντικές προόδους στη συλλογή, κατανόηση και προτυποποίηση της κυκλοφορίας με βάση την μεγάλη ροή πληροφορίας μέσω βίντεο. Επιπλέον η υπολογιστική όραση βρίσκεται πλέον στο προσκήνιο της μηχανικής μάθησης. Υπάρχουν αλγόριθμοι και προ-εκπαιδευμένα μοντέλα βαθιάς μάθησης υπεύθυνα για την ανάλυση τεράστιων όγκων εικόνας και βίντεο με τεράστια δυνατότητα εντοπισμού πληροφορίας. Αυτά τα δίκτυα πλέον εμφανίζονται σε προβλήματα μεταφορών. Σκοπός της παρούσας διπλωματικής είναι να δημιουργήσει ένα αξιόπιστο μεθοδολογικό πλαίσιο για τον εντοπισμό κυκλοφοριακών συνθήκων που να βασίζονται σε μοντέλα βαθιάς μάθησης και σε βίντεο καταγεγραμμένα από μη επανδρωμένο αεροσκάφος. Η πρώτη προσέγγιση που ακολουθείται αποσκοπεί στην αξιολόγηση, σε πραγματικό χρόνο, κυκλοφοριακών συνθηκών ενός τμήματος οδού σε κορεσμένο ή μη κορεσμένο, με ελάχιστο υπολογιστικό κόστος. Για να επιτευχθεί αυτό αναπτύσσονται και αξιολογούνται πολλά συνελικτικά νευρωνικά δίκτυα με διαφορετικές δομές. Η αναγνώριση της βέλτιστης δομής βασίζεται στην ελαχιστοποίηση του σφάλματος ανίχνευσης στο δείγμα αξιολόγησης. Ο σκοπός της δεύτερης προσέγγισης είναι να εντοπίσει, να κατηγοριοποιήσει τα οχήματα και να προσεγγίσει κυκλοφοριακά μεγέθη σε ένα βίντεο ενός οδικού τμήματος. Αυτή η μέθοδος εκμεταλλεύεται προ-εκπαιδευμένα συνελικτικά νευρωνικά δίκτυα τροποποιημένα για χρήση με τα δεδομένα βίντεο από το μη επανδρωμένο αεροσκάφος. Οι προτεινόμενες προσεγγίσεις βασίζονται σε δεδομένα από βίντεο που καταγράφηκαν από μη επανδρωμένο αεροσκάφος εξοπλισμένο με κάμερα, πάνω από την οδό Κατεχάκη, Αθήνα. Οι δύο προσεγγίσεις στη συνέχεια συγκρίνονται με βάση την ακρίβεια και τη δυνατότητα τους να εφαρμοστούν και συζητούνται οι πιθανές εφαρμογές στον τομέα των μεταφορών και της παρακολούθησης της κυκλοφορίας.

Λέξεις Κλειδιά: Συνελικτικά Νευρωνικά Δίκτυα, Υπολογιστική Όραση, Μηχανική Μάθηση, Βαθιά Μάθηση, Αξιολόγηση Κυκλοφορίας, Εντοπισμός και Κατηγοριοποίηση Οχημάτων, Υπολογισμός Μακροσκοπικών Μεγεθών.

Computer Vision Models for Identifying Traffic Conditions from Unmanned Aerial Vehicles

Georgios Davos

Supervising Professor: Eleni Vlachogianni

ABSTRACT

The rapid development of unmanned aerial vehicles (UAV/Drones) technology in recent years has resulted in significant advancements in collecting, understanding and modeling traffic based on large video streams. In addition, Computer Vision (CV) is now at the forefront of Machine Learning; algorithms and pre-trained deep learning networks exist for analyzing huge volumes of images with enormous detection capabilities. These networks are emerging in transportation problems. The purpose of this diploma thesis is to provide a consistent methodological framework to detect traffic conditions based on deep learning models and video data captured from an unmanned aerial vehicle. The first approach followed aims to evaluate in real time, with minimal computation cost, the general traffic conditions of a road segment in two categories. To achieve that, many convolutional neural networks with different structures are developed and evaluated. The identification of the optimal structure is based on the minimization of the detection error on the evaluation sample. The purpose of the second approach is to localize, classify vehicles and approximate traffic metrics and conditions in a video of a road segment. This method takes advantage of pre-trained convolutional neural networks modified to be used with the video data from the unmanned aerial vehicle. The two approaches are implemented on video data captured from an unmanned aerial vehicle (UAV/Drone) equipped with a camera, over Katehaki Street, Athens. Both approaches are then compared based on their accuracy and feasibility in implementation and possible applications in the field of transportation and traffic monitoring are discussed.

Key words: Convolutional Neural Networks, Computer Vision, Machine Learning, Deep Learning, Unmanned Aerial Vehicle, Traffic Metrics, Vehicle Classification and Localization.

ΠΕΡΙΛΗΨΗ

Τα τελευταία χρόνια η κοινωνία έχει δει απότομη ανάπτυξη της τεχνολογίας. Αυτό έχει ως αποτέλεσμα να έχει γίνει αναπόσπαστο κομμάτι της καθημερινότητας σχεδόν σε κάθε τομέα του σύγχρονου ανθρώπου. Επόμενο είναι αυτή η ανάπτυξη να έχει επηρεάσει σημαντικά και τον συγκοινωνιακό τομέα προσφέροντας νέες τεχνολογίες και εναλλακτικές σε πληθώρα ζητημάτων. Πιο συγκεκριμένα έχουν εμφανιστεί νέες εφαρμογές συλλογής και επεξεργασίας κυκλοφοριακών δεδομένων και φαινομένων. Η ανάλυση εικόνας που προέρχεται από μη επανδρωμένο αεροσκάφος (Unmanned Aerial Vehicles – UAV ή Drone) εμφανίζει μεγάλο ενδιαφέρον καθώς τα τελευταία χρόνια έχει οδηγήσει σε σημαντικές προόδους στην συλλογή και κατανόηση της κυκλοφορίας από οπτικά μέσα. Επιπλέον η υπολογιστική όραση έχει θεμελιώσει τη θέση της στον τομέα της μηχανικής μάθησης. Με χρήση αλγόριθμων και προ-εκπαιδευμένων μοντέλων βαθιάς μάθησης αναλύεται τεράστιος εικόνας και βίντεο. Αυτά τα μοντέλα πλέον εμφανίζονται σε πολλές εφαρμογές της κυκλοφορίας και των μεταφορών λόγω της εξαιρετικής δυνατότητας ανάλυσης μεγάλου όγκου δεδομένων.

Από την ανασκόπηση της βιβλιογραφίας, αναδεικνύονται τα πλεονεκτήματα και οι περιορισμοί των εφαρμογών που αξιοποιούν μη επανδρωμένα αεροσκάφη και τεχνικές υπολογιστικής όρασης, ενώ παρατηρείται πως οι υπάρχουσες εφαρμογές που αξιοποιούν μη επανδρωμένα αεροσκάφη σε συνδυασμό με εφαρμογές βαθιάς μάθησης για την επεξεργασία των δεδομένων, είναι περιορισμένες και δεν ασχολούνται με εφαρμογές στην κυκλοφορία.

Η παρούσα εργασία αποσκοπεί στην ανάπτυξη και αξιολόγηση αξιόπιστων προτύπων βαθιάς μάθησης και υπολογιστικής όρασης για τον εντοπισμό κυκλοφοριακών συνθήκων από βίντεο καταγεγραμμένα με μη επανδρωμένο αεροσκάφος.

Στο πλαίσιο της εργασίας θα εξεταστούν δύο προσεγγίσεις που ακολουθούν διαφορετικές μεθοδολογίες οι οποίες διαφέρουν στην χρήση τους αλλά και στα χαρακτηριστικά τους. Η πρώτη προσέγγιση αποσκοπεί στην αξιολόγηση, σε πραγματικό χρόνο, κυκλοφοριακών συνθηκών ενός τμήματος οδού σε δύο κατηγορίες (κορεσμένο ή μη κορεσμένο), με ελάχιστο υπολογιστικό κόστος. Ο σκοπός της δεύτερης προσέγγισης είναι να εντοπίσει, να κατηγοριοποιήσει τα οχήματα και να εκτιμήσει κυκλοφοριακά μεγέθη από ένα βίντεο ενός οδικού τμήματος.

Και στις δύο προσεγγίσεις αξιοποιήθηκαν δεδομένα μορφής βίντεο που έχουν καταγράψει την κυκλοφορία επι την οδό Κατεχάκη, Αθήνα από μη επανδρωμένο αεροσκάφος. Κάθε προσέγγιση υποστηρίζεται από ξεχωριστή προεπεξεργασία των δεδομένων. Πρώτο βήμα είναι η κατάλληλη μετατροπή του βίντεο σε εικόνες για δημιουργία βάσεων δεδομένων. Για το πρώτο πρότυπο απομονώνεται ένα τμήμα από την οδό και εξάγονται εικόνες που χωρίζονται σε δύο κατηγορίες, Χαμηλής/Καθόλου Κυκλοφορίας και Υψηλής Κυκλοφορίας με κριτήρια τον αριθμό και θέση των οχημάτων στις εικόνες. Για την δεύτερη μεθοδολογική προσέγγιση απομονώνονται επίσης επιθυμητά τμήματα της οδού από τα βίντεο, όπου ακολουθώντας μετατρέπονται σε εικόνες στις οποίες επισημαίνεται με κατάλληλη μορφή η θέση και κατηγορία των οχημάτων που βρίσκονται εντός αυτών.

Στο δεύτερο στάδιο της εργασίας, εξετάζονται αρχικά διαφορετικές μεθοδολογίες για κάθε μια από τις δυο προσεγγίσεις. Για την επίτευξη του στόχου της πρώτης προσέγγισης

αναπτύσσονται, εκπαιδεύονται και αξιολογούνται συνελικτικά νευρωνικά δίκτυα με διαφορετικές δομές. Επίσης εξετάζεται η εφαρμογή διαβιβαστικής μάθησης από προ-εκπαιδευμένα δίκτυα. Η επιλογή της βέλτιστης δομής βασίζεται στην ελαχιστοποίηση του σφάλματος στο δείγμα αξιολόγησης και στην μεγιστοποίηση ακρίβειας. Το βέλτιστο δίκτυο χαρακτηρίζεται από υψηλή ακρίβεια (91%), ενώ η εφαρμογή διαβιβαστικής μάθησης, παρά την υψηλή ακρίβεια (89%) της, κρίνεται μη ιδανική για την συγκεκριμένη προσέγγιση, για λόγους όπως οι υπολογιστικές απαιτήσεις. Τελικά το μοντέλο εφαρμόζεται σε βίντεο με δεδομένα που δεν έχουν χρησιμοποιηθεί στην εκπαίδευση, για την αξιολόγηση της δυνατότητας γενίκευσης σε δεδομένα που θα λαμβάνει σε πραγματικό χρόνο.

Για τις απαιτήσεις της δεύτερης προσέγγισης αξιοποιούνται προ-εκπαιδευμένα συνελικτικά νευρωνικά δίκτυα για εντοπισμό αντικειμένων, ρυθμισμένα καταλλήλως για τα βίντεο από το μη επανδρωμένο αεροσκάφος. Το τελικό μοντέλο εντοπίζει και κατηγοριοποιεί τα οχήματα σε ένα τμήμα οδού με ικανοποιητική ακρίβεια (89% F1 Score), στη συνέχεια εφαρμόζεται σε βίντεο για την προσέγγιση κυκλοφοριακών μεγεθών και αξιολογείτε με βάση τις υπολογιστικές και χρονικές απαιτήσεις του.

Κατά την διεξαγωγή της διπλωματικής εργασίας εντοπίστηκαν κάποιοι περιορισμοί όπως, η αξιόπιστη αξιολόγηση των μοντέλων λόγω υψηλής ομοιότητας των δεδομένων εκπαίδευσης και αξιολόγησης. Για την αντιμετώπιση τους είναι προφανής η ανάγκη για περισσότερα και καλύτερης ποιότητας δεδομένα, καθώς είναι η κυριότερη προϋπόθεση για την διαμόρφωση αξιόπιστων προτύπων βαθιάς μάθησης. Επίσης σημαντικός περιοριστικός παράγοντας είναι οι υψηλές υπολογιστικές απαιτήσεις των μεθοδολογιών υπολογιστικής όρασης και βαθιάς μάθησης σε σχέση με άλλες μεθόδους. Επομένως πρέπει να γίνουν κατάλληλες επιλογές και ρυθμίσεις για την βέλτιστη αξιοποίηση των διαθέσιμων δεδομένων και εξοπλισμού.

Συνοψίζοντας, συνιστάται περαιτέρω ερευνά, με περισσότερα δεδομένα και επανεξέταση των προτύπων για την κατάλληλη, πιο γενικευμένη, εφαρμογή τους στις περιπτώσεις που εξετάστηκαν με μη επανδρωμένα αεροσκάφη και στα ακόλουθα βήματα τους. Επίσης αναφέρεται η δυνατότητα διαφορετικών εφαρμογών όπως, για παράδειγμα, ο εντοπισμός θέσεων στάθμευσης, μήκος ουρών αναμονής σε διασταυρώσεις και παρακολούθηση κυκλοφορίας πεζών από εναέρια λήψη. Τελικά οι εφαρμογές βαθιάς μάθησης σε συνδυασμό την αξιοποίηση μη επανδρωμένων αεροσκαφών παρουσιάζουν πληθώρα δυνατών εφαρμογών με τη συνεχή ανάπτυξη της τεχνολογίας και περαιτέρω έρευνα.

ΠΕΡΙΕΧΟΜΕΝΑ

1 - ΕΙΣΑΓΩΓΗ ..... - 1 -
1.1 - Εξέλιξη και Εφαρμογές Μη Επανδρωμένων Αεροσκαφών ..... - 1 -
1.2 - Σύγχρονες Εφαρμογές Επεξεργασίας Οπτικού Υλικού ..... - 3 -
1.3 - Σκοπός Διπλωματικής Εργασίας ..... - 4 -
1.4 - Διάρθρωση Διπλωματικής Εργασίας ..... - 5 -
2 - ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ ..... - 7 -
2.1 - Εισαγωγή ..... - 7 -
2.2 - Παρακολούθηση Κυκλοφορίας ..... - 7 -
2.2.1 - Μέθοδοι Χωρίς Αξιοποίηση Οπτικού Υλικού ..... - 7 -
2.2.2 - Μέθοδοι με Αξιοποίηση Οπτικού Υλικού ..... - 8 -
2.2.3 - Εφαρμογές Μη Επανδρωμένων Αεροσκαφών στην Κυκλοφορία ..... - 10 -
2.3 - Σύγχρονες Εφαρμογές Βαθιάς Μάθησης ..... - 20 -
2.4 - Γενικά Συμπεράσματα Βιβλιογραφικής Ανασκόπησης ..... - 23 -
3 - ΜΕΘΟΔΟΛΟΓΙΑ ..... - 25 -
3.1 - Γενικά ..... - 25 -
3.2 - Παραδοχές ..... - 25 -
3.3 - Περιγραφή Προτεινόμενων Προσεγγίσεων ..... - 26 -
3.3.1 - Αναγνώριση κυκλοφορικών συνθηκών σε πραγματικό χρόνο ..... - 27 -
3.3.2 - Μακροσκοπική αναγνώριση κυκλοφορικών συνθηκών ..... - 28 -
3.4 - Θεωρητικό Υπόβαθρο ..... - 29 -
3.4.1 - Κυκλοφοριακά Μεγέθη ..... - 29 -
3.4.2 - Υπολογιστική Όραση ..... - 31 -
3.4.3 - Μέθοδος Αξιολόγησης ..... - 43 -
3.5 - Επιλογή Λογισμικού - Υπολογιστικού Εξοπλισμού και Περιορισμοί τους ..... - 45 -
4 - ΣΥΛΛΟΓΗ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ ..... - 47 -
4.1 - Συλλογή Δεδομένων ..... - 47 -
4.2 - Προετοιμασία Δεδομένων ..... - 49 -
4.2.1 - Αναγνώριση Κυκλοφοριακών Συνθηκών σε Πραγματικό Χρόνο ..... - 49 -
4.2.2 - Μακροσκοπική Αναγνώριση Κυκλοφοριακών Συνθηκών ..... - 52 -
5 - ΑΝΑΛΥΣΗ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ ..... - 55 -
5.1 - Εισαγωγή ..... - 55 -
5.2 - Αναγνώριση Κυκλοφορικών Συνθηκών σε Πραγματικό Χρόνο ..... - 55 -
5.2.1 - Συνοπτική Θεώρηση ..... - 55 -

5.2.2 - Εκτενής Θεώρηση ..... - 57 -
5.2.3 - Βέλτιστο Μοντέλο ..... - 63 -
5.2.4 - Εφαρμογή Συνελικτικού Νευρωνικού Δικτύου σε Πραγματικό Χρόνο ..... - 71 -
5.2.5 - Εφαρμογή Διαβιβαστικής Μάθησης ..... - 74 -
5.3 - Μακροσκοπική Αναγνώριση Κυκλοφορικών Συνθηκών ..... - 77 -
5.3.1 - Περιγραφή ..... - 77 -
5.3.2 - Εκπαίδευση Προ-Εκπαιδευμένου Δικτύου ..... - 77 -
5.3.3 - Αποτελέσματα Εντοπισμού και Κατηγοριοποίησης Οχημάτων ..... - 79 -
5.4 - Σύγκριση Προτεινόμενων Προσεγγίσεων ..... - 85 -
6 - ΣΥΜΠΕΡΑΣΜΑΤΑ ..... - 87 -
6.1 - Σύνοψη Μεθοδολογίας και Αποτελεσμάτων ..... - 87 -
6.2 - Βασικά Συμπεράσματα ..... - 88 -
6.3 - Προτάσεις για Περαιτέρω Έρευνα ..... - 90 -
7 - ΒΙΒΛΙΟΓΡΑΦΙΑ ..... - 93 -

ΕΥΡΕΤΗΡΙΟ ΔΙΑΓΡΑΜΜΑΤΩΝ

Διάγραμμα 3-1: Διάγραμμα Ροής με τις δύο Διαφορετικές Προσεγγίσεις..... - 26 -
Διάγραμμα 3-2: Φαινόμενο Υπερπροσαρμογής ..... - 35 -
Διάγραμμα 3-3: Διάγραμμα Καμπύλης ROC (Powers, 2007)..... - 45 -
Διάγραμμα 5-1: Γραφήματα Ακρίβειας (πάνω) – Συνάρτησης Κόστους-Απωλειών (κάτω) της Εκπαίδευσης Αρχικών Μοντέλων (ανά εποχή)..... - 56 -
Διάγραμμα 5-2: Γράφημα Ακρίβειας της Αξιολόγησης των Μοντέλων (2ο Γράφημα: Τα Αποτελέσματα Έχουν Εξομαλυνθεί για Καλύτερη Ορατότητα του Γραφήματος ..... - 56 -
Διάγραμμα 5-3: Γράφημα Τιμών Συνάρτησης Κόστους της Αξιολόγησης των Μοντέλων (2ο Γράφημα: Τα Αποτελέσματα Έχουν Εξομαλυνθεί για Καλύτερη Ορατότητα του Γραφήματος)..... - 57 -
Διάγραμμα 5-4: Γραφήματα Ακρίβειας (πάνω) – Συνάρτησης Κόστους (κάτω) της Εκπαίδευσης των Μοντέλων ..... - 59 -
Διάγραμμα 5-5: Γράφημα Ακρίβειας της Αξιολόγησης των Μοντέλων (2ο Γράφημα: Τα Αποτελέσματα Έχουν Εξομαλυνθεί για Καλύτερη Ορατότητα του Γραφήματος)..... - 60 -
Διάγραμμα 5-6: Γράφημα Τιμών Συνάρτησης Κόστους της Αξιολόγησης των Μοντέλων (2ο Γράφημα: Τα Αποτελέσματα Έχουν Εξομαλυνθεί για Καλύτερη Ορατότητα του Γραφήματος)..... - 61 -
Διάγραμμα 5-7: Απεικόνιση Δομής- Χαρακτηριστικών Εξεταζόμενου Μοντέλου και Ροής Λειτουργειών. (Διάγραμμα Ροής από TensorBoard)..... - 64 -
Διάγραμμα 5-8: Γραφήματα Ακρίβειας(πάνω) – Συνάρτησης Κόστους (κάτω) της Εκπαίδευσης Εξεταζόμενου Μοντέλου ..... - 65 -
Διάγραμμα 5-9: Γραφήματα Ακρίβειας(πάνω) – Συνάρτησης Κόστους (κάτω) της Αξιολόγησης Εξεταζόμενου Μοντέλου..... - 65 -
Διάγραμμα 5-10: Διάγραμμα Απωλειών Εντοπισμού Οχημάτων κατά την Αξιολόγηση (Localization Loss) ..... - 78 -
Διάγραμμα 5-11: Διάγραμμα Απωλειών Κατηγοριοποίησης Οχημάτων κατά την Αξιολόγηση (Classification Loss) ..... - 78 -
Διάγραμμα 5-12: Διάγραμμα Συνολικών Απωλειών κατά την Αξιολόγηση (Total Loss) . - 79 -
Διάγραμμα 5-13: Ενδεικτικό Διάγραμμα Φόρτου-Ταχύτητας από Βίντεο ..... - 83 -
Διάγραμμα 5-14: Ενδεικτικό Διάγραμμα Πυκνότητάς-Ταχύτητας από Βίντεο ..... - 83 -
Διάγραμμα 5-15: Ενδεικτικό Διάγραμμα Πυκνότητας-Φόρτου από Βίντεο ..... - 84 -

ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ

Πίνακας 2-1: Συγκριτική Αξιολόγηση των Τριών Μέσων Καταγραφής ως προς τις Απαιτήσεις ανά Κατηγορία (Barmpounakis, 2017) - 10 -
Πίνακας 2-2: Συγκεντρωτικός Πίνακας Βιβλιογραφίας στη Χρήση Εναέριων Μέσων Καταγραφής της Κυκλοφορίας (Barmpounakis, 2017) - 11 -
Πίνακας 3-1: Μερικές Συναρτήσεις Ενεργοποίησης (Wikipedia, 2019) - 33 -
Πίνακας 3-2: Χαρακτηριστικά Προ-Εκπαιδευμένων Μοντέλων στη Βάση Δεδομένων COCO (Google-Tensorflow, 2019) - 41 -
Πίνακας 3-3: Πίνακας Σύγχυσης (confusion matrix) - 44 -
Πίνακας 4-1: Ενδεικτικός Διαχωρισμός του Βίντεο με Όνομα 1797.mp4 (Με Βάση τα Στιγμιότυπα - Frames) - 50 -
Πίνακας 5-1: Μερικά Εκτενή Εξεταζόμενα Μοντέλα και Χαρακτηριστικά τους - 58 -
Πίνακας 5-2: Χαρακτηριστικά Εκπαίδευσης Δικτύου 1ης Μεθόδου - 63 -
Πίνακας 5-3: Αποτελέσματα Αξιολόγησης Συνελικτικού Νευρωνικού Δικτύου - 66 -
Πίνακας 5-4: Αποτελέσματα ανά Χρονική Περίοδο στο Ενδεικτικό Βίντεο–1η Μέθοδος - 73 -
Πίνακας 5-5: Χαρακτηριστικά Υπολογιστικές Επιδώσεις Εφαρμογής Δικτύου 1ης Μεθόδου στο Ενδεικτικό Βίντεο - 74 -
Πίνακας 5-6: Προ-Εκπαιδευμένα Μοντέλα που Υποστηρίζονται από το TensorFlow με τις Καταγεγραμμένες Ακρίβειες τους (Google-Tensorflow, 2019) - 75 -
Πίνακας 5-7: Στατιστικά Αποτελέσματα Μοντέλου από Διαβιβαστική Μάθηση με Βάση το ResNet50 - 76 -
Πίνακας 5-8: Χαρακτηριστικά Εκπαίδευσης Δικτύου 2ης Μεθόδου - 79 -
Πίνακας 5-9: Αποτελέσματα Εντοπισμού/Κατηγοριοποίησης - 80 -
Πίνακας 5-10: Αποτελέσματα-Κυκλοφοριακά Μεγέθη ανά Χρονική Περίοδο στο Ενδεικτικό Βίντεο - 82 -
Πίνακας 5-11: Χαρακτηριστικά Υπολογιστικές Επιδώσεις Εφαρμογής Δικτύου 2ης Μεθόδου στο Ενδεικτικό Βίντεο - 84 -
Πίνακας 5-12: Συγκριτικά Αποτελέσματα 1ης Μεθοδολογίας - 85 -
Πίνακας 5-13: Αποτελέσματα των δύο Μεθόδων ανά Χρονική Περίοδο στο Ενδεικτικό Βίντεο - 85 -
Πίνακας 5-14: Σύγκριση Απαιτήσεων Υπολογιστικής Ισχύς των δύο Μεθόδων - 86 -

ΕΥΡΕΤΗΡΙΟ ΕΙΚΟΝΩΝ

Εικόνα 1-1: Σταθερή Κάμερα για Παρακολούθηση Κυκλοφορίας ..... - 1 -
Εικόνα 1-2: Δύο Είδη Σύγχρονων Μη Επανδρωμένων Αεροσκαφών (Αριστερά: Εξακόπτερο, Δεξιά: Τετρακόπτερο) (DJI, 2019) ..... - 2 -
Εικόνα 1-3: Προσομοίωση Αυτόνομου Αυτοκινήτου (NVIDIA, 2018) ..... - 4 -
Εικόνα 2-1: Προβλεπόμενο Πλήθος Συσκευών Έξυπνων Τηλεφώνων (Cisco, 2019)..... - 8 -
Εικόνα 2-2: Πρόβλεψη Ατυχήματος Από Σταθερή Κάμερα (Shah et al. 2018)..... - 9 -
Εικόνα 2-3: Μετρητής Οχημάτων από Σταθερή Κάμερα (Grant et al. 2000) ..... - 9 -
Εικόνα 2-4: Πίνακας Προέλευσης/Προορισμού σε Κυκλικό Κόμβο (Khan, et al., 2018). - 13 -
Εικόνα 2-5: Απεικόνιση Τροχιών Οχημάτων και Διάφορων Συνθηκών σε Τετρασκελή Διασταύρωση με Φωτεινό Σηματοδότη (Khan et al. 2018)..... - 14 -
Εικόνα 2-6: Εξαγωγή Κυκλοφοριακών Χαρακτηριστικών σε Αστική Κυκλοφορία υπό Συμφόρηση από Εναέρια Λήψη (Kaufmann et al. 2018). ..... - 15 -
Εικόνα 2-7: Συνεργατικό Σύστημα UAV και Σταθερών Καμερών (Liu, et al. 2019)..... - 16 -
Εικόνα 2-8: Εξεταζόμενη Διασταύρωση Ερευνάς (Barmpounakis et al. 2015)..... - 17 -
Εικόνα 2-9: Τροχιές Οχημάτων της Διασταύρωσης και πίνακας Προέλευσης – Προορισμού (Barmpounakis et al. 2015)..... - 17 -
Εικόνα 2-10: Καταγραφή Τροχιάς Πεζών (Barmpounakis et al. 2015). ..... - 18 -
Εικόνα 2-11: Παράμετροι Απόφασης για Τεχνική Καταγραφής Βίντεο (Barmpounakis et al. 2017) ..... - 19 -
Εικόνα 2-12: Εντοπισμός Πινακίδων Κυκλοφορίας σε Εικόνα με Χρήση Συνελικτικών Νευρωνικών Δικτύων (Rezaei et al. 2018) ..... - 21 -
Εικόνα 2-13: Μετατροπή Εικόνας Πινακίδας Κυκλοφορίας σε Ψηφιακή Πληροφορία με Χρήση Συνελικτικών Νευρωνικών Δικτύων (Zherzdez et al. 2018)..... - 21 -
Εικόνα 2-14: Βέλτιστα Πρότυπα Συνελικτικών Νευρωνικών Δικτύων για Σωστή Αναγνώριση Οδικής Σήμανσης ..... - 21 -
Εικόνα 2-15: Αποτελέσματα Μοντέλων Faster R-CNN και YOLOv3 Για Εντοπισμό Οχημάτων (Benjdira et al., 2018) ..... - 22 -
Εικόνα 3-1: Σύγκριση Τυπικής Επίλυσής και Επίλυσης με Νευρωνικά Δίκτυα..... - 31 -
Εικόνα 3-2: Τυπική Μορφή Νευρωνικού Δικτύου (Wikipedia, 2019) ..... - 32 -
Εικόνα 3-3: Τυπική Μορφή Νευρώνα (Nielsen, 2018) ..... - 32 -
Εικόνα 3-4: Ενέργειες Τυπικού Νευρώνα ενός Νευρωνικού Δικτύου (Ahirwar, 2017).... - 33 -
Εικόνα 3-5: Εφαρμογή Μεθόδου Απόρριψης Νευρώνων (Srivastava et al., 2014) ..... - 36 -
Εικόνα 3-6: Λειτουργεία Νευρώνων κατά την Εκπαίδευση και Δοκιμή με Εφαρμογή Απόρριψης Νευρώνων (Srivastava et al., 2014)..... - 36 -
Εικόνα 3-7: Μετατροπή με την Μέθοδο Κανονικοποίησης Τμημάτων Δεδομένων σε Ένα Τμήμα (Ioffe et al., 2015) ..... - 37 -
Εικόνα 3-8: Παράδειγμα Συνέλιξης Φίλτρων σε Εικόνα (XRDS, 2016) ..... - 38 -
Εικόνα 3-9: Παράδειγμα Υπό-Δειγματοληψίας (Computer Science Wiki, 2018) ..... - 38 -
Εικόνα 3-10: Ροή Συνελικτικού Νευρωνικού Δικτύου με Πολλά Επίπεδα Συνέλιξης (Medium, 2018) ..... - 39 -
Εικόνα 3-11: Παράδειγμα Μοτίβων Ενεργοποίησης Φίλτρων σε Συνελικτικό Νευρωνικό Δίκτυο (Honglak et al.,, 2011) ..... - 39 -
Εικόνα 3-12: Παραδείγματα Παραμόρφωσης Εικόνας (Taylor et al., 2017) ..... - 40 -

Εικόνα 3-13: Δίκτυο Εντοπισμού Πολλαπλών Πλαισίων με Μία Λήψη (Single Shot Multibox Detection Network – SSD Network) (Liu, et al., 2016) ..... - 42 -
Εικόνα 3-14: Παράδειγμα Εικόνας σε Δίκτυο Εντοπισμού Πολλαπλών Πλαισίων με Μία Λήψη (Single Shot Multibox Detection Network – SSD Network) (Liu, et al., 2016) ..... - 42 -
Εικόνα 4-1: Χαρακτηριστικά της κάμερας Go Pro Hero 3 ..... - 48 -
Εικόνα 4-2: Ενδεικτικά Στιγμιότυπα από τα Βίντεο..... - 48 -
Εικόνα 4-3: Ενδεικτικές Εικόνες Κατηγορίας 0 (Χαμηλή/Καθόλου Κυκλοφορία)..... - 51 -
Εικόνα 4-4: Ενδεικτικές Εικόνες Κατηγορίας 1 (Αυξημένη Κυκλοφορία) ..... - 51 -
Εικόνα 4-5: Παραδείγματα Χρήσης Προγράμματος LabelImage ..... - 54 -
Εικόνα 5-1: Αρχική Εικόνα Κατηγορίας 1 (Διαστάσεις 150:60) ..... - 67 -
Εικόνα 5-2: 1ο Επίπεδο Συνέλιξης - 32 Φίλτρα – Κατηγορία 1..... - 67 -
Εικόνα 5-3: 2ο Επίπεδο Συνέλιξης - 64 Φίλτρα - Κατηγορία 1 ..... - 68 -
Εικόνα 5-4: 3ο Επίπεδο Συνέλιξης - 128 Φίλτρα - Κατηγορία 1 ..... - 68 -
Εικόνα 5-5: Ενδεικτικά - 31ο Φίλτρο 1ου Επιπέδου Συνέλιξης - Κατηγορία 1..... - 69 -
Εικόνα 5-6: Αρχική εικόνα κατηγορίας 0 (Διαστάσεις 150:60)..... - 69 -
Εικόνα 5-7: Επίπεδο Συνέλιξης - 32 Φίλτρα – Κατηγορία 0..... - 70 -
Εικόνα 5-8: 2ο Επίπεδο Συνέλιξης - 64 Φίλτρα - Κατηγορία 0 ..... - 70 -
Εικόνα 5-9: 3ο Επίπεδο Συνέλιξης - 128 Φίλτρα - Κατηγορία 0 ..... - 70 -
Εικόνα 5-10: Ενδεικτικά - 31ο Φίλτρο 1ου Επιπέδου Συνέλιξης - Κατηγορία 0..... - 71 -
Εικόνα 5-11: Ενδεικτικό Στιγμιότυπο Πρόβλεψης - Κατηγορία 1..... - 72 -
Εικόνα 5-12: Ενδεικτικό Στιγμιότυπο Πρόβλεψης - Κατηγορία 0..... - 72 -
Εικόνα 5-13: Ενδεικτικά Στιγμιότυπα Εντοπισμού και Κατηγοριοποίησης Οχημάτων σε Βίντεο με Χρήση του Προ-Εκπαιδευμένου Δικτύου..... - 80 -
Εικόνα 5-14: Ενδεικτικά Στιγμιότυπα Εντοπισμού και Κατηγοριοποίησης Οχημάτων σε Βίντεο με Χρήση του Προ-Εκπαιδευμένου Δικτύου και Υπολογισμού Κυκλοφοριακών Μεγεθών. ..... - 81 -

1 - ΕΙΣΑΓΩΓΗ

1.1 - Εξέλιξη και Εφαρμογές Μη Επανδρωμένων Αεροσκαφών

Η συλλογή κυκλοφοριακών δεδομένων και μοτίβων αποτελεί ένα θεμελιώδες τμήμα του τομέα των μεταφορών και παρακολούθησης κυκλοφορίας. Η ανάλυση των κυκλοφοριακών δεδομένων ενισχύει την δυνατότητα αντίληψης και ερμηνείας της κυκλοφορίας και της κίνησης των οχημάτων.

Πιο συγκεκριμένα η ραγδαία ανάπτυξης της τεχνολογίας έχει ως αποτέλεσμα την εξέλιξη των οπτικών μέσων καταγραφής, ενισχύοντας την βαθύτερη κατανόηση κυκλοφοριακών φαινομένων σε μακροσκοπικό και μικροσκοπικό επίπεδο.

Η συνήθης πρακτική για καταγραφή κυκλοφορίας εδώ και πολλά έτη, είναι η εγκατάσταση σταθερών καμερών (Εικόνα 1-1) στο μήκος μίας οδού ή σε σημεία υψηλού ενδιαφέροντος όπως κόμβοι ή σημεία προς μελέτη. Η χρήση σταθερών καμερών παρόλο που είναι μία καθιερωμένη και αποδοτική μέθοδος καταγραφής οπτικού υλικού παρουσιάζει κάποια προβλήματα. Ένα από αυτά είναι η εξ ορισμού χωρική δέσμευση καθώς είναι σταθερές κάμερες, γεγονός που περιορίζει το εύρος της καταγραφόμενης περιοχής, ενώ δεν δύναται να ακολουθεί πορείες οχημάτων εκτός της περιοχής αυτής. Επίσης η εγκατάσταση των καμερών για ασφαλή φύλαξη και ικανοποιητική ορατότητα του σημείου ενδιαφέροντος απαιτεί τοποθέτηση σε κατάλληλό ύψος και απαραίτητη υποδομή, γεγονός που πιθανώς να εμφανίσει περιορισμούς ενώ συνοδεύεται από σημαντικό κόστος για τοποθέτηση κάθε κάμερας. Μερικά προβλήματα που μπορούν να εμφανιστούν επίσης είναι η πιθανή παραμόρφωσης της εικόνας λόγω χρήσης φακών και η ύπαρξη κρυφών σημείων επί της οδού λόγω προσωρινών ή μόνιμων εμποδίων.

Εικόνα 1-1: Σταθερή Κάμερα για Παρακολούθηση Κυκλοφορίας

Τα τελευταία χρόνια έχει παρατηρηθεί απότομη αύξηση στην αξιοποίηση **Μη Επανδρωμένων Αεροσκαφών (Unmanned Aerial Vehicles-UAV ή Drones)** (Εικόνα 1-2) σε πολλούς τομείς. Αυτό οφείλεται στις τεχνολογικές εξελίξεις που έχουν υποστεί τα UAV που τους επιτρέπουν περισσότερους αυτοματισμούς, αυτονομία και αξιοπιστία, σε κομμάτι λειτουργικού και υλικού εξοπλισμού καθώς και στις δυνατότητες τηλεμετρίας (Mairaj et al., 2019). Τα UAV πλέον έχουν θεμελιώσει την θέση τους σε πολλές εφαρμογές στο γενικό επάγγελμα του πολιτικού μηχανικού, όπως έλεγχος δομικών στοιχείων πού δεν είναι εύκολα προσβάσιμα, χαρτογράφηση περιοχών ή έλεγχος της υποδομής οδικού δικτύου (Rednoa Inc., 2017).

Πιο συγκεκριμένα η χρήση UAV στο επάγγελμα στου Συγκοινωνιολόγου Μηχανικού έχει αποδειχτεί πολύ χρήσιμη καθώς η πληροφορία της εικόνας που προσφέρουν μπορεί να αξιοποιηθεί με πολλούς τρόπους για αξιολόγηση κυκλοφοριακών συνθηκών.

Εικόνα 1-2: Δύο Είδη Σύγχρονων Μη Επανδρωμένων Αεροσκαφών (Αριστερά: Εξακόπτερο, Δεξιά: Τετρακόπτερο) (DJI, 2019)

Μία από της πρώτες εφαρμογές που αξιοποιήθηκαν UAV στον συγκοινωνιακό τομέα είναι το Witas UAV Project (Doherty et al., 2000) το οποίο αποσκοπούσε στον σχεδιασμό ενός UAV που θα κινούταν πανω από οδικές αρτηρίες για να εντοπίζει δίαφορα φαινόμενα όπως προσπεράσεις και η παρακολόυθηση ορισμένων τύπων οχημάτων. Έως σήμερα έχουν γίνει πολλές ακόμα προσπάθειες για εξαγωγή κυκλοφοριακής πληροφορίας και χαρακτηριστικών με χρήση μη επανδρομένων αεροσκάφων, χωρίς να έχει κορεστεί η δυνατότητα αξιοπόιησης UAV σε μεθοδολογίες για μελλοντική ερευνα.

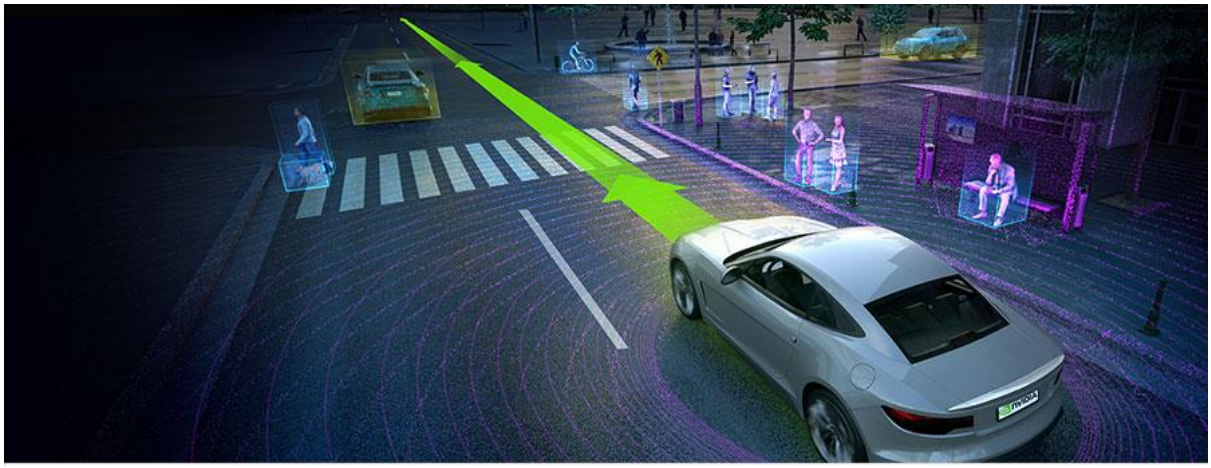
Παρά το πληθος των εφαρμογών με χρήση UAV, οι έρευνες και οι εφαρμογές βρίσκοντα ακόμα σε αρχικό στάδιο για τον κύριο λογω πως η ανάλυση τέτοιας μορφής πληροφορίας παρουσιάζει αρκετούς περιορισμούς, όπως θα φανεί και από την βιβλιογραφία.

1.2 - Σύγχρονες Εφαρμογές Επεξεργασίας Οπτικού Υλικού

Στη σύγχρονη κοινωνία υπάρχει σε καθημερινή βάση τεράστιος όγκος πληροφοριών και δεδομένων που κινούνται με μανιακούς ρυθμούς μεταξύ όλων των βασικών τμημάτων της. Αυτή η ραγδαία αύξηση πληροφορίας είχε ως άμεσο αποτέλεσμα την ανάπτυξη μεγάλου πλήθους νέων βελτιστοποιημένων μεθόδων αξιολόγησης και επεξεργασίας πληροφορίας κάθε μορφής. Πολλές από αυτές τις μεθοδολογίες βρίσκονται υπό την ομπρέλα της έννοιας **Μηχανική Μάθηση (Machine Learning),** που έχουν ως γενικό σκοπό την εύκολη και γρήγορη επεξεργασία δεδομένων και τον αυτοματισμό αυτών, εκμεταλλεύοντας τα άλματα της τεχνολογίας κυρίως στα υπολογιστικά συστήματα. Βασικό κομμάτι της μηχανικής μάθησης είναι η έννοια της **Βαθιάς Μάθησης (Deep Learning)**, που περιέχει κατά κύριο λόγο τις διαφορετικές κατηγορίες **Νευρωνικών Δικτύων (Neural Networks)**, τα οποία εφαρμόζονται σε πληθώρα τομέων λόγω της ικανότητας να κάνουν καλές γενικεύσεις.

Η **Υπολογιστική Όραση (Computer Vision)** είναι ένα επιστημονικό πεδίο που ασχολείται με την δυνατότητα των υπολογιστών να αποκτήσουν μια κατανόηση υψηλού επιπέδου από ψηφιακές εικόνες ή βίντεο. Απώτερος σκοπός της αξιοποίησης μεθόδων υπολογιστικής όρασης είναι να αυτοματοποιηθούν εργασίες που μπορούν να γίνουν με το ανθρώπινο οπτικό σύστημα. Βασικό κομμάτι της υπολογιστικής όρασης είναι οι εφαρμογές βαθιάς μάθησης που αξιοποιούν δομές **Συνελικτικών Νευρωνικών Δικτυών (Convolutional Neural Networks)**, που ενδείκνυται για εξαγωγή πληροφοριών από οπτικό υλικό.

Στον τομέα των μεταφορών τα τελευταία χρόνια οι εφαρμογές υπολογιστικής όρασης και της βαθιάς μάθησης έχουν έρθει στο προσκήνιο, καθώς χρησιμοποιούνται σε πολλούς τομείς. Μία πολύ γνωστή εφαρμογή είναι στην έρευνα και παραγωγή ημιαυτόνομων και αυτόνομων οχημάτων, που κυρίως σε χώρες του εξωτερικού όπως οι Ηνωμένες Πολιτείες της Αμερικής, έχουν γίνει τεράστια άλματα προς την καθιέρωση, αρχικά των ημιαυτόνομων και εν συνεχεία των αυτόνομων οχημάτων. Μερικά παραδείγματα είναι η εταιρία Tesla που ήδη παρέχει με τα οχήματα της ημιαυτόνομές δυνατότητες, με την πλήρη αυτονομία να είναι σε τελικά στάδια. Επίσης η εταιρία nVidia βρίσκεται σε πολύ καλό στάδιο με την πλατφόρμα της για αυτόνομα οχήματα nVidia Drive Platform όπου έχει συνδυάσει, τον απίστευτα υψηλών αποδόσεων, εξοπλισμό της με δικές της αρχιτεκτονικές υπολογιστικής όρασης και τεχνητής νοημοσύνης, υποστηρίζοντας πως τα αυτόνομα οχήματα θα είναι εφικτά στο άμεσο μέλλον (Εικόνα 1-3). Μία σημαντική προσπάθεια επίσης σε παρόμοια θέματα κάνει το πανεπιστήμιο του Berkeley με την ερευνητική του πλατφόρμα DeepDrive που εξετάζει διάφορες εφαρμογές σχετικά με τα αυτόνομα οχήματα και την κυκλοφορία.



Εικόνα 1-3: Προσομοίωση Αυτόνομου Αυτοκινήτου (NVIDIA, 2018)

1.3 - Σκοπός Διπλωματικής Εργασίας

Η ανάλυση εικόνας που προέρχεται από μη επανδρωμένο αεροσκάφος (Unmanned Aerial Vehicles – UAV ή Drone) για αναγνώριση χαρακτηριστικών και μοτίβων της κυκλοφορίας είναι κρίσιμη και παρουσιάζει πολλές προκλήσεις.

Σκοπός της διπλωματικής εργασίας είναι η ανάπτυξη και αξιολόγηση ενός αξιόπιστου μεθοδολογικού πλαισίου για τον εντοπισμό κυκλοφοριακών συνθήκων που να βασίζεται σε μοντέλα βαθιάς μάθησης και υπολογιστικής όρασης και δεδομένα μορφής βίντεο καταγεγραμμένα από μη επανδρωμένο αεροσκάφος.

Θα εξεταστούν δύο προσεγγίσεις που ακολουθούν διαφορετικές μεθοδολογίες που διαφέρουν στην χρήση τους αλλά και στα χαρακτηριστικά τους.

Η πρώτη προσέγγιση αποσκοπεί στην αξιολόγηση, σε πραγματικό χρόνο, κυκλοφοριακών συνθηκών ενός τμήματος οδού σε δύο κατηγορίες, με ελάχιστο υπολογιστικό κόστος. Για να επιτευχθεί αυτό, αναπτύσσονται, εκπαιδεύονται και αξιολογούνται πολλαπλά συνελικτικά νευρωνικά δίκτυα με διαφορετικές δομές Εξετάζεται επίσης η εφαρμογή διαβιβαστικής μάθησης με χρήση προπαιδευμένων συνελικτικών νευρωνικών δικτύων. Η αναγνώριση της βέλτιστης δομής βασίζεται στην ελαχιστοποίηση του σφάλματος ανίχνευσης στο δείγμα αξιολόγησης και στην ακρίβεια των προβλέψεων.

Ο σκοπός της δεύτερης προσέγγισης είναι να εντοπίσει, να κατηγοριοποιήσει οχήματα και να προσεγγίσει κυκλοφοριακά μεγέθη σε ένα βίντεο ενός οδικού τμήματος. Αυτή η μέθοδος εκμεταλλεύεται προ-εκπαιδευμένα συνελικτικά νευρωνικά δίκτυα τροποποιημένα για χρήση με τα δεδομένα βίντεο από το μη επανδρωμένο αεροσκάφος

Στο τέλος θα αξιολογηθούν οι προσεγγίσεις και θα συγκριθούν, με σκοπό να εξεταστεί η δυνατότητα εφαρμογής τους, μαζί με προτάσεις για μελλοντική έρευνα σε εφαρμογές στον τομέα μεταφορών και παρακολούθησης κυκλοφορίας.

1.4 - Διάρθρωση Διπλωματικής Εργασίας

Παρουσιάζεται η διάταξη των κεφαλαίων της διπλωματική εργασίας και το περιεχόμενο τους.

Το πρώτο κεφάλαιο της εργασίας αναφέρεται εισαγωγικά στην εξέλιξη εφαρμογών των μη επανδρωμένων αεροσκαφών, και των νεών δυνατοτήτων αξιοποίησης και ανάλυσης πληροφοριών με μεθοδολογίες βαθιάς μάθησης και υπολογιστικής όρασης. Αναφέρεται ο σκοπός της εργασίας να εξετάσει τον συνδυασμό των παραπάνω εφαρμογών. Τέλος παρουσιάζεται η διάρθρωση της διπλωματικής εργασίας.

Το κεφάλαιο 2 περιλαμβάνει την βιβλιογραφική ανασκόπηση ερευνών που έχουν γίνει στις εφαρμογές μη επανδρωμένων αεροσκαφών, με αναφορές στις σύγχρονες εφαρμογές μεθόδων υπολογιστικής όρασης και βαθιάς μάθησης στο τομέα των μεταφορών για εξαγωγή κυκλοφοριακών δεδομένων.

Στο τρίτο κεφάλαιο γίνεται αναφορά στο σκοπό της εργασίας και στην συνέχεια αναλύονται οι μεθοδολογίες που ακολουθήθηκαν κατά την εκπόνηση της διπλωματικής εργασίας μαζί με αντίστοιχο διάγραμμα ροής. Παρουσιάζεται το απαραίτητο θεωρητικό υπόβαθρο και αναφέρεται η επιλογή λογισμικού και υπολογιστικού εξοπλισμού.

Στο κεφάλαιο 4 αναφέρεται ο τρόπος συλλογής των δεδομένων βίντεο με τα χαρακτηριστικά τους και στην συνέχεια αναλύονται οι διαδικασίες επεξεργασίας αυτών των δεδομένων για αξιοποίηση τους στις μεθοδολογίες.

Το κεφάλαιο 5 περιλαμβάνει την αναλυτική πορεία επίλυσης του προβλήματος των δύο μεθοδολογιών με ταυτόχρονη παράθεση των αποτελεσμάτων. Το κεφάλαιο περιέχει τις διαδικασίες εκπαίδευσης αξιολόγησης και δοκιμής των μοντέλων, και στην συνέχεια την εφαρμογή τους σε βίντεο για αξιολόγηση της δυνατότητας εφαρμογής τους για την αξιολόγηση της κυκλοφορίας.

Στο κεφάλαιο 6 γίνεται μία σύνοψη της εργασίας και των αποτελεσμάτων με τον σχολιασμό αυτών. Αναφέρονται οι περιοριστικοί παράγοντες και τα πλεονεκτήματα των προσεγγίσεων. Τέλος γίνονται προτάσεις για μελλοντική έρευνα η πιθανές εφαρμογές.

Στο κεφάλαιο 7 παρατίθεται η πλήρης βιβλιογραφία που χρησιμοποιήθηκε κατά την εκπόνηση της διπλωματικής εργασίας.

.

2 - ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

2.1 - Εισαγωγή

Στο κεφάλαιο αυτό γίνεται ανάλυση προγενέστερων ερευνών ώστε να κατανοηθούν οι υπάρχουσες εφαρμογές στον συγκοινωνιακό τομέα μαζί με τα πλεονεκτήματα και τα μειονεκτήματα τους. Η βιβλιογραφική ανασκόπηση θα καθοδηγήσει τα κριτήρια επιλογής προτύπων κατά την διάρκεια της διπλωματικής εργασίας. Το κεφάλαιο θα αναφερθεί σε μεθόδους καταγραφής και παρακολούθησης την κυκλοφορίας με έμφαση σε εφαρμογές με μη επανδρωμένα αεροσκάφη (Unmanned Aerial Vehicles-UAV ή Drones) και στις σύγχρονες εφαρμογές προτύπων υπολογιστικής όρασης στον τομέα των μεταφορών.

2.2 - Παρακολούθηση Κυκλοφορίας

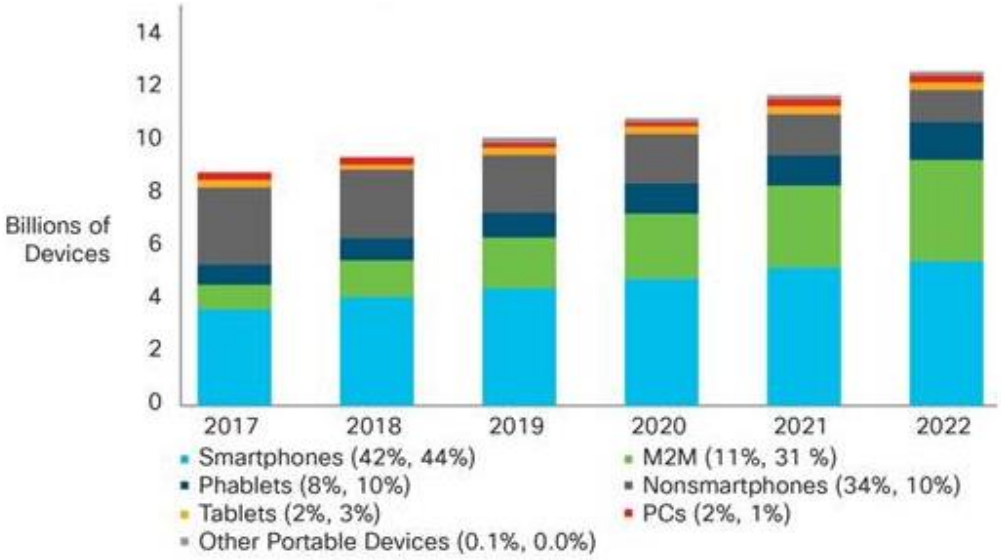
2.2.1 - Μέθοδοι Χωρίς Αξιοποίηση Οπτικού Υλικού

Οι πλειοψηφία των μεθόδων καταγραφής δεδομένων βασίζονται σε τεχνολογίες που εφαρμόζονται πολλά έτη, έχοντας ωριμάσει προσφέρουν ακριβή και αξιόπιστα κυκλοφοριακά δεδομένα. Οι βασικότερες μεθοδολογίες είναι οι φωρατές, αισθητήρες ανίχνευσης εντός και εκτός του οδοστρώματος, με πιο σύγχρονες τεχνολογίες να αξιοποιούν ακουστικές μεθόδους, ραντάρ, λέιζερ και μαγνητικούς ανιχνευτές. Επίσης τα τελευταία χρόνια αξιοποιούνται σε μεγάλο βαθμό αισθητήρες GPS είτε ενσωματωμένοι στα οχήματα είτε από τα έξυπνα τηλέφωνα και συσκευές, για καταγραφή κυκλοφοριακών δεδομένων και συνθηκών.

Η αξιοποίηση αισθητήρων είναι πολύ αξιόπιστη (Leduc, 2008) και για αυτόν τον λόγο χρησιμοποιούνται κατά κύριο λόγο από τον τομέα των μεταφορών για καταγραφή κυκλοφοριακών δεδομένων. Παρόλα αυτά παρουσιάζουν μερικά περιοριστικά προβλήματα όπως το υψηλό κόστος εγκατάστασης και συντήρησης των σταθερών μετρητών και ο περιορισμός στις περιοχές τοποθέτησης για να υπάρχουν αξιόπιστες μετρήσεις. Επίσης αξίζει να αναφερθεί πως μέθοδοι όπως η χρήση πνευματικών σωλήνων σε μία διατομή οδού θεωρούνται ευκολά εφαρμόσιμες για καταγραφή οχημάτων και μεγεθών, παρόλα αυτά υπάρχουν έρευνες που αμφισβητούν την ακρίβεια και την αξιοπιστία τους για καταγραφή δεδομένων (Sanderson et al., 2011).

Μία σύγχρονη μέθοδος επίσης είναι η συλλογή πληροφοριών από τις ηλεκτρονικές συσκευές έξυπνων τηλεφώνων. Η απότομη ανάπτυξη της τεχνολογίας και των δυνατοτήτων τους, τα έχει εξοπλίσει με προηγμένες τεχνολογίες όπως αισθητήρες κίνησης, προσανατολισμού, και GPS, που αξιοποιούνται σε πολλές έρευνες για συλλογή δεδομένων. Αυτά τα δεδομένα μπορεί να

είναι κυκλοφοριακά δεδομένα, δεδομένα τροχιών και διαδρομών, δεδομένα για συμπεριφορά του οδηγού ή και δεδομένα για καταγραφή συμβάντων (Demissie et al., 2013) (Steenbruggen et al., 2011). Η αξιοποίηση των έξυπνων τηλεφώνων ενδείκνυται για συλλογή δεδομένων τέτοιας μορφής γιατί παρέχουν υψηλή ποιότητα δεδομένων με ευρεία περιοχή μελέτης χωρίς ιδιαίτερη δυσκολία εφαρμογής. Επίσης βασικός παράγοντας της χρησιμότητας τους είναι η σημαντική αύξηση του πλήθους των έξυπνων συσκευών και της πληροφορίας που διακινούν τα τελευταία χρόνια όπως παρουσιάζουν και έρευνες της Cisco (Cisco, 2019) (Εικόνα 2-1).



Εικόνα 2-1: Προβλεπόμενο Πλήθος Συσκευών Έξυπνων Τηλεφώνων (Cisco, 2019)

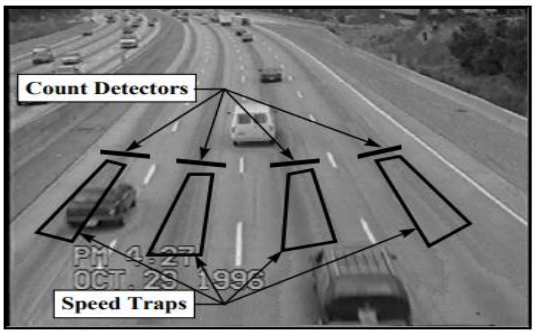
2.2.2 - Μέθοδοι με Αξιοποίηση Οπτικού Υλικού

Στα πλαίσιο της διπλωματικής εργασίας περισσότερο ενδιαφέρον έχουν οι εφαρμογές καταγραφής και αξιοποίησης οπτικού υλικού. Όπως αναφέρθηκε και στο πρώτο κεφάλαιο οι καθιερωμένες μέθοδοι είναι οι σταθερές κάμερες και σε ειδικές περιπτώσεις τα επανδρωμένα αεροσκάφη (Manned Aerial Vehicles - MAV), όπως για παράδειγμα τα ελικόπτερα της αστυνομίας.

Η χρήση σταθερών καμερών ενδείκνυται για πολλές εφαρμογές επεξεργασίας οπτικού υλικού. Το χαρακτηριστικό των καμερών που προκαλεί τους περισσότερους περιορισμούς είναι και αυτό που κάνει την χρήση τους πολύ καλή επιλογή σε πολλές περιπτώσεις. Πέρα από την ζωντανή αναμετάδοση πληροφορίας για οπτικό έλεγχο, το βίντεο που καταγράφεται από μία κάμερα επι μίας οδού μπορεί να επεξεργαστεί για απόσπαση πληροφοριών με εφαρμογές επεξεργασίας βίντεο. Σε αυτές τις εφαρμογές η σταθερή λήψη του βίντεο χωρίς απότομες αλλαγές λήψης η συνθηκών ενδείκνυται για την αποτελεσματική επεξεργασία όπως φαίνεται και στην έρευνα των Shah et al. (2018) που επιδιώκουν να προβλέψουν και να καταγράψουν περιστατικά ατυχημάτων από σταθερές κάμερες (Εικόνα 2-2)

Εικόνα 2-2: Πρόβλεψη Ατυχήματος Από Σταθερή Κάμερα (Shah et al. 2018)

Μία ακόμα εφαρμογή από τους Grant et al. (2000) επιδιόκει την καταγραφή οχημάτων για μέτρηση αυτών από σταθερή κάμερα επι εθνικής οδού τεσσάρων λωρίδων (Εικόνα 2-3). Πολύ σημαντικός παράγοντας είναι πάλι η σταθερότητα της κάμερας για την σωστή λειτουργεία της εφαρμογής καθώς εχουν οριστεί σταθερές περιοχές αναγνωρισης σε κάθε λωρίδα για την μέτρηση οχημάτων.



Εικόνα 2-3: Μετρητής Οχημάτων από Σταθερή Κάμερα (Grant et al. 2000)

Η εικόνα αποτελεί πλέον βασικό είδος δεδομένων στον τομέα των μεταφορών, με πολλές εφαρμογές να έχουν θεμελιωθεί για την παρακολούθηση κυκλοφορίας και την συλλογή κυκλοφοριακών δεδομένων με σκοπό δημιουργία λεπτομερών βάσεων δεδομένων. Όπως αναφέρθηκε και στο πρώτο κεφάλαιο υπάρχουν βασικοί περιορισμοί στις εφαρμογές με σταθερές κάμερες. Για την καταγραφή πολύπλοκων κυκλοφοριακών φαινομένων σε ένα οδικό τμήμα απαιτείται μεγάλος αριθμός καμερών σε κατάλληλες αποστάσεις γιατί η λήψη κάθε κάμερας είναι περιορισμένη λόγω της σταθερής θέσης τους. Επομένως για την καταγραφή μεγάλων οδικών δικτύων απαιτείται μεγάλο πλήθος καμερών κάτι που ισοδυναμεί με πολύ μεγάλο κόστος εγκατάστασης υποδομών και συντήρησης για μία περίπτωση οδού. Άλλα προβλήματα καταγραφής βίντεο παρουσιάζονται στην εργασία του Lee (2007), με τα βασικότερα να είναι η ύπαρξη κρυφών σημείων, σκιών από το περιβάλλον, σκιές των οχημάτων και η παραμόρφωση από τον φακό της κάμερας τα οποία επηρεάζουν σημαντικά τις συνήθης εφαρμογές επεξεργασίας βίντεο δημιουργώντας σφάλματα και αποκλίσεις που μπερδεύουν τα μοντέλα επεξεργασίας με αποτέλεσμα την μείωση της αξιοπιστίας τους.

Για την ακριβέστερη καταγραφή δεδομένων απαιτείται να ξεπεραστούν τα προηγούμενα προβλήματα κάτι που ενδέχεται να πραγματοποιηθεί σε ένα κομμάτι από τον εναλλακτικό

τρόπο συλλογής οπτικού υλικού μέσω μη επανδρωμένων αεροσκαφών, εφαρμογές των οποίων θα αναφερθούν στην συνέχεια. Στον πίνακα που ακολουθεί (Πίνακας 2-1) παρουσιάζεται μία συγκριτική αξιολόγηση των τριών μέσων καταγραφής εικόνας ως προς τις απαιτήσεις και τα βασικά χαρακτηριστικά τους (Barmpounakis, 2017).

Πίνακας 2-1: Συγκριτική Αξιολόγηση των Τριών Μέσων Καταγραφής ως προς τις Απαιτήσεις ανά Κατηγορία (Barmpounakis, 2017)

	Σταθερές Κάμερες	***MAV***	***UAV***
Μήκος Καταγραφής	Χαμηλή	Υψηλή	Υψηλή
Ασφάλεια/Ιδιωτικότητα	Μέτρια	Μέτρια	Χαμηλή
Κόστος Απόκτησης και Συντήρησης	Χαμηλή	Υψηλή	Χαμηλή
Πολλαπλές χρήσεις	Χαμηλή	Υψηλή	Υψηλή
Κατανάλωση Ενέργειας	Χαμηλή	Χαμηλή	Υψηλή
Ανάπτυξη/Τοποθέτηση	Χαμηλή	Υψηλή	Χαμηλή
Επιχειρησιακός χρόνος	Υψηλή	Υψηλή	Χαμηλή
Χρήση υπό δυσμενείς συνθήκες	Μέτρια	Χαμηλή	Χαμηλή
Ρίσκο	Χαμηλή	Υψηλή	Μέτρια
Αντοχή	Υψηλή	Υψηλή	Χαμηλή
Μεταεπεξεργασία Βίντεο	Μέτρια	Υψηλή	Υψηλή
Μετάδοση δεδομένων, επικοινωνία και αποθήκευση	Χαμηλή	Υψηλή	Υψηλή
Δεξιότητες χρήσης	Χαμηλή	Υψηλή	Μέτρια
Ανάγκη εκπαίδευσης	Χαμηλή	Υψηλή	Μέτρια
Περιπλοκότητα	Μέτρια	Υψηλή	Μέτρια

2.2.3 - Εφαρμογές Μη Επανδρωμένων Αεροσκαφών στην Κυκλοφορία

Οι εφαρμογές των μη επανδρωμένων αεροσκαφών στον συγκοινωνιακό τομέα έχουν εμφανιστεί εδώ και πολλά χρόνια, με διάφορες απόπειρες επίλυσης κυκλοφοριακών προβλημάτων και καταγραφής κυκλοφοριακών στοιχείων. Ο λόγος που τα τελευταία χρόνια έχει αυξηθεί ο αριθμός σχετικών ερευνών είναι πλήρως εξαρτημένος με την τεχνολογική ανάπτυξη των μηχανικών μερών, των λειτουργικών πλοήγησης και τηλεμετρίας αυτών. Τα παραπάνω έχουν κάνει την απόκτηση, χρήση και συντήρηση των UAV πολύ πιο εύκολή και προσιτή για τον μελετητή. Επίσης πολλές σχετικές έρευνες συνιστούν πλέον την αξιοποίηση μη επανδρωμένων αεροσκαφών για καταγραφή βίντεο γιατί όπως φαίνεται και στον προηγούμενο Πίνακα (Πίνακας 2-1) έχουν πολλά πλεονεκτήματα έναντι των άλλων μεθόδων. Τα βασικότερα πλεονεκτήματα είναι η καταγραφή δεδομένων χωρίς να επηρεαστεί συμπεριφορά των οδηγών της οδού, η δυνατότητα της καταγραφής δεδομένων σε μεγάλες περιοχές χάρη στην ευελιξία των σύγχρονων UAV και η ευκολία αποθήκευσης και μετάδοσης πληροφορίας.

Να αναφερθεί πως παρά την ανάπτυξη τεχνολογιών UAV, υπάρχουν ακόμα μερικοί πολύ σημαντικοί περιοριστικοί παράγοντες στην αξιοποίηση τους. Ο κυριότερος περιορισμός είναι η περιορισμένη διάρκεια πτήσης λόγω μπαταριών. Τα πιο σύγχρονα UAV φτάνουν αυτονομίες κοντά στα 30 λεπτά πτήσης με ένα σετ μπαταριών. Επίσης η σταθερότητα λήψης βίντεο είναι πρόβλημα πολλές φορές καθώς η επιρροή ανέμων μπορεί να είναι σημαντικά αισθητή στο βίντεο. Αυτό το πρόβλημα βελτιώνετε με τον καιρό με εφαρμογή καλύτερων αυτοματισμών μηχανισμών σταθεροποίησης του UAV αλλά και της κάμερας. Με εξαίρεση μερικά UAV που το κόστος τους είναι εξαιρετικά μεγάλο, η πλειοψηφία τους περιορίζεται πολύ από τις καιρικές συνθήκες, καθώς τα φαινόμενα βροχής και υψηλών ανέμων καθιστούν αδύνατη τη πτήση τους. Τέλος ένας βασικός περιορισμός είναι ο νομικός περιορισμός πτήσης αεροσκαφών. Αναλυτικοί νόμοι και κανονισμοί για τα Ελληνικά πλαίσιο (ΦΕΚ Β-3152/30-9-2016) υπάρχουν στην αντίστοιχη ιστοσελίδα της Υπηρεσίας Πολιτικής Αεροπορίας του Υπουργείου Υποδομών και Μεταφορών (https://uas.hcaa.gr/).

Στην συνέχεια (Πίνακας 2-2) εμφανίζεται ένας συγκεντρωτικός πίνακας βιβλιογραφίας στη χρήση εναέριων μέσων καταγραφής της κυκλοφορίας για συγκοινωνιακές εφαρμογές μαζί με τα κύριο σημεία τους (Barmpounakis, 2017)

Πίνακας 2-2: Συγκεντρωτικός Πίνακας Βιβλιογραφίας στη Χρήση Εναέριων Μέσων Καταγραφής της Κυκλοφορίας (Barmpounakis, 2017)

Συγγραφέας	***Έτος***	***Τύπος Α/Φ***	***Σκοπός της εργασίας***	***Κεντρικά Σημεία***
Doherty et al.	2000	UAV	Αναγνώριση, εντοπισμός και παρακολούθηση οχημάτων (WITAS Project)	Ένα πλήρες σχέδιο για χρήση UAVs σε κυκλοφοριακές εφαρμογές
Granlund et al.	2000	UAV	Αναγνώριση, εντοπισμός και παρακολούθηση οχημάτων (WITAS Project)	Θέματα Όρασης Υπολογιστών
Mirchandani et al.	2002	MAV	Εξαγωγή κινηματικών χαρακτηριστικών όπως πυκνότητα, χρόνοι ταξιδιού, αριθμός στρεφουσών κινήσεων κτλ.	Αναγνώριση δυνατοτήτων των UAVs για παρακολούθηση και διαχείριση κυκλοφορίας
Shastry and Schowengerdt	2002	MAV	Σταθεροποίηση βίντεο και εντοπισμός οχημάτων	Συνδυασμών GIS και βίντεο από αέρα
Frew et al.	2004	UAV	Παρακολούθηση της οδού (UC Berkeley)	Θέματα ακρίβειας λόγω υψομέτρου
Srinivasan et al.	2004	-	Έρευνα σκοπιμότητας χρήσης των UAVs για παρακολούθηση κυκλοφορίας (Florida DOT)	Έρευνα σκοπιμότητας χρήσης των UAVs
Farradyne	2005	-	Έρευνα σκοπιμότητας χρήσης των UAVs σαν μέρος της υποδομής των ITS	Έρευνα σκοπιμότητας χρήσης των UAVs
Kaaniche et al.	2005	UAV	Παρακολούθηση της οδού και εντοπισμός κινούμενων αντικειμένων	Αποτελεσματικός, γρήγορος και πραγματικού χρόνου αλγόριθμος
Coifman et al.	2006	UAV	Εξαγωγή παραμέτρων όπως Επίπεδο Εξυπηρέτησης, ΕΜΗΚ, μητρώα προέλευσης προορισμού κ.α.	Δυνατότητες των βίντεο από αέρα για εφαρμογές κυκλοφορίας
Lin et al.	2006	-	Γενική ιδέα και πρώτη προσέγγιση για μετατροπή δεδομένων πραγματικού χρόνου σε κυκλοφοριακή πληροφορία	Δυνατότητες των βίντεο από αέρα για εφαρμογές κυκλοφορίας
Reinartz et al.	2006	MAV	Αυτόματος εντοπισμός οχημάτων και υπολογισμός ταχυτήτων	Δυνατότητες των βίντεο από αέρα για εφαρμογές κυκλοφορίας
Heintz et al.	2007	UAV	Αναγνώριση, εντοπισμός και παρακολούθηση οχημάτων (WITAS Project)	Παρακολούθηση σε πραγματικό χρόνο
Puri et al.	2007	UAV	Χρήση δεδομένων πραγματικού χρόνου σαν είσοδος σε Πρότυπα προσομοίωσης	Δυνατότητες σμήνους UAVs για την κυκλοφορία

Ro et al.	2007	UAV	Έρευνα σκοπιμότητας χρήσης των UAVs σαν μέρος της υποδομής των ITS.	Επόμενα ερευνητικά βήματα
Yu Min Chen and Oh	2007	UAV	Έλεγχος ποιότητας σύνδεσης με ένα DOT (Michigan DOT)	Ασφάλεια και θέματα συνδεσιμότητας
Lenhart et al.	2008	-	Αυτοματοποιημένος τρόπος παρακολούθησης της κυκλοφορίας με εντοπισμό οχημάτων	Χρήση σειριακών εικόνων
Li	2008	UAV	Εντοπισμός οχημάτων για υπολογισμός κυκλοφοριακού φόρτου και κατηγοριοποίηση οχημάτων	Αποτελεσματικός ασαφής (fuzzy) αλγόριθμος
McCormack and Trepanier	2008	UAV	Έρευνα σκοπιμότητας χρήσης των UAVs για παρακολούθηση της κυκλοφορίας	Έρευνα σκοπιμότητας χρήσης των UAVs
Cheng et al.	2009	UAV	Εντοπισμός και παρακολούθηση οχημάτων σε βίντεο από πραγματικό χρόνο	Αποτελεσματικός αλγόριθμος εντοπισμού
Braut et al.	2012	UAV	Υπολογισμός μητρώων Προέλευσης Προορισμού σε σύνθετες διασταυρώσεις	Αναγνώριση θεμάτων βίντεο για εντοπισμό οχημάτων
Liu et al.	2013	UAV	Πρότυπο βελτιστοποίησης επιλογής διαδρομής για UAV	Εντοπισμός θεμάτων για χρήση των UAV ως μέσο συλλογής κυκλοφοριακής πληροφορίας
Salvo et al.	2013	UAV	Μέτρηση κενού αποδοχής	Απόδειξη της αξιοπιστίας των UAVs
Zhang et al.	2013	UAV	Αναγνώριση συμβάντων, όπως για παράδειγμα ακινητοποιημένα οχήματα, ή οχήματα που κινούνται σε λάθος κατεύθυνση	Υψηλή ακρίβεια στον εντοπισμό συμβάντων
Gao et al.	2014	UAV	Αυτόματος εντοπισμός τροχιών οχημάτων και υπολογισμός ταχυτήτων και επιταχύνσεων	Υψηλή ακρίβεια στον αυτόματο εντοπισμό οχημάτων
Kanistras et al.	2014	-	Ανασκόπηση σε εφαρμογές UAVs για παρακολούθηση κυκλοφορίας	Δυνατότητες των UAV και πλεονεκτήματα για την παρακολούθηση της κυκλοφορίας
Lima et al.	2014	MAV	Αυτόματος εντοπισμός οχημάτων	Υψηλή ακρίβεια στην εξαγωγή τροχιών
Salvo et al.	2014a	UAV	Σύγκριση μακροσκοπικών προτύπων κάνοντας χρήση βίντεο από UAVs	Πλεονεκτήματα των UAVs στην παρακολούθηση της κυκλοφορίας
Salvo et al.	2014b	UAV	Σύγκριση ταχυτήτων μεταξύ GPS και UAVs	Υψηλή ακρίβεια και ελάχιστος αριθμός σημείων για βαθμονόμηση
Wang et al.	2014	UAV	Συλλογή δεδομένων από τροχιές οχημάτων	Επιρροή του υψομέτρου στην ακρίβεια
Xin et al.	2014	UAV	Σύγκριση ταχυτήτων μεταξύ επίγειων ραντάρ και UAVs	Λεπτομερής περιγραφή της διαδικασίας εντοπισμού και υπολογισμού ταχύτητας
Zhang et al.	2014	UAV	Αναγνώριση ορίων δρόμου	Υψηλή ακρίβεια στον εντοπισμό ορίων της οδού
Gheisari et al.	2015	-	Προτάσεις για χρήση UAVs σαν υποδομή των ITS	Συνεντεύξεις με προσωπικό από DOT
Lee et al.	2015	UAV	Παρακολούθηση κυκλοφοριακών συμβάντων	Επιρροή της ποιότητας των βίντεο στην αποτελεσματικότητα των UAV
Liu et al.	2015	UAV	Εντοπισμός σταματημένων οχημάτων	Πειράματα και προσομοιώσεις πτήσεων
Wang et al.	2015	UAV	Εντοπισμός και παρακολούθηση οχημάτων	Υψηλή ακρίβεια στον αυτόματο εντοπισμό οχημάτων και στην παρακολούθησή τους
Barmpounakis et al.	2016	UAV	Εξαγωγή κινηματικών παραμέτρων όπως ταχύτητα, επιταχύνσεις και τροχιές	Υψηλή ακρίβεια και ανάλυση βήμα-προς-βήμα
Erdelj and Natalizio	2016	-	Ανασκόπηση βασικών εφαρμογών διαχείρισης καταστροφών	Λεπτομερής ανασκόπηση των εφαρμογών που χρησιμοποιούν UAV και ανοιχτά θέματα προς επίλυση

Όπως αναφέρθηκε και στην εισαγωγή έχουν γίνει απόπειρες αξιοποίησης UAV από το έτος 2000, με το Witas UAV Project (Doherty et al., 2000) να είναι από τα πρώτα. Μεγάλο πλήθος από τις πρώτες έρευνες αποσκοπούν να αντιμετωπίσουν πολλά βασικά στοιχεία στις μεταφορές και την κυκλοφορία. Πιο συγκεκριμένα η έρευνα των Liu et al. (2002) εξετάζει την δυνατότητα συλλογής, από βίντεο, κυκλοφοριακών στοιχείων. Σε άλλες έρευνες εξετάζεται η σταθεροποίηση του βίντεο κατά την καταγραφή και συλλογή δεδομένων καθώς αποτελεί έναν βασικό περιορισμό αυτής (Shastry et al., 2002). Οι Coifman et al. (2006) κατάφεραν να καθορίσουν σε αστικό περιβάλλον στάθμη εξυπηρέτησης, ετήσια μέση κυκλοφορία και την μέτρηση κυκλοφοριακών συνθηκών σε κόμβους.

Μία ενδιαφέρουσα έρευνα χρησιμοποίησε UAV για τον εντοπισμό συμβάντων όπως σταματημένα οχήματα, οχήματα κινούμενα σε λάθος κατεύθυνση και τα όρια οδών με εξαιρετικά ποσοστά επιτυχίας (Zhang et al. 2013).

Επίσης ενδιαφέρον έχουν πιο σύγχρονες έρευνες που παρατηρείτε η εξέλιξη της τεχνολογίας των UAV και η αυξημένη ποιότητα καταγραφής οπτικού υλικού. Οι Khan, et al. (2018) χρησιμοποίησαν μη επανδρωμένο αεροσκάφος για καταγραφή και ανάλυση κυκλοφοριακών ροών σε αστικούς κυκλικούς κόμβους, εξάγοντας πίνακα προέλευσης προορισμού (Εικόνα 2-4).

		Destination				
	Counters	2	4	6	8	Total
Origin	1	25	27	20	0	72
	3	0	7	26	12	45
	5	1	4	33	36	74
	7	19	61	0	19	99
	Total	45	99	79	67	

Εικόνα 2-4: Πίνακας Προέλευσης/Προορισμού σε Κυκλικό Κόμβο (Khan, et al., 2018)

Σε διαφορετική έρευνα των Khan et al. (2018) εξετάζεται κόμβος τετρασκελής διασταύρωσης με φωτεινό σηματοδότη για καταγραφή της κυκλοφοριακής ροής, τροχιών οχημάτων, και υπολογισμό κυκλοφοριακών μεγεθών όπως πυκνότητα (Εικόνα 2-5). Έγινε επίσης μία προσπάθεια αυτοματοποίησης αυτών των υπολογισμών.

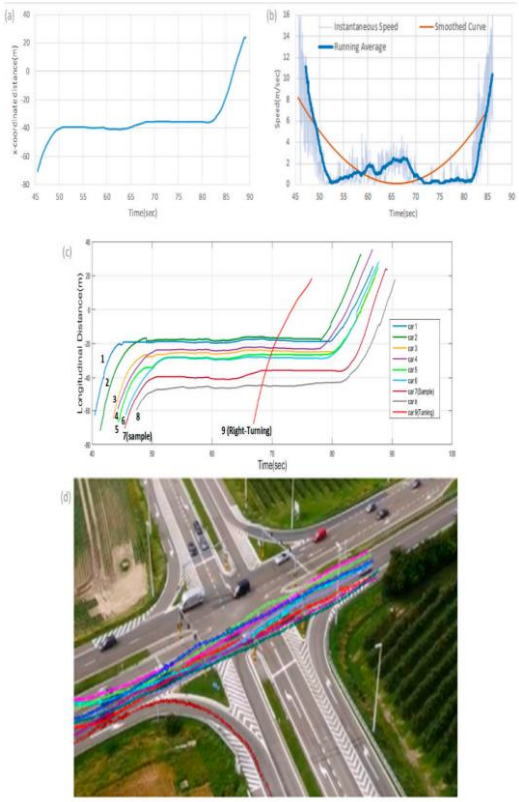


Figure 6. (a) The space-time diagram of the trajectory of a sample vehicle; (b) the speed profile of the sample vehicle; (c) the space-time diagram of extracted trajectories (labelled with assigned number) of a group of vehicles; and (d) an illustration of extracted trajectories.

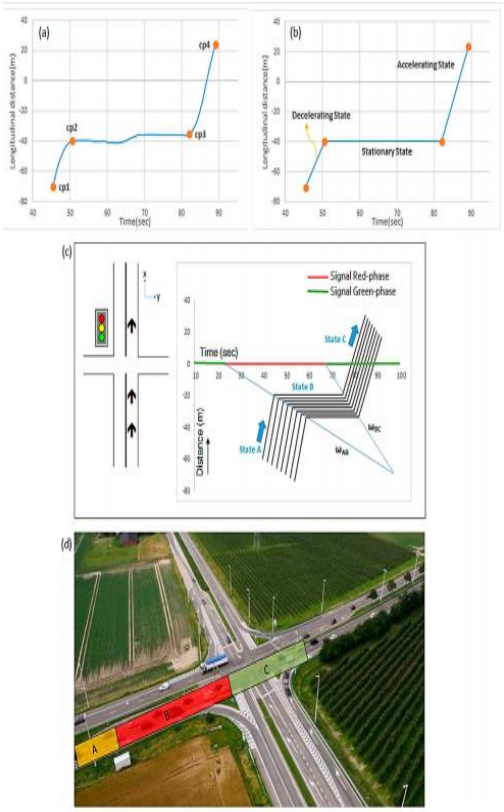


Figure 7. (a) Extracted critical points (CP) on vehicle trajectory; (b) simplified trajectories with identified traffic states; (c) the generated shockwaves and signal cycle length; and (d) an illustration of various traffic states on the UAV-acquired image.

Εικόνα 2-5: Απεικόνιση Τροχιών Οχημάτων και Διάφορων Συνθηκών σε Τετρασκελή Διασταύρωση με Φωτεινό Σηματοδότη (Khan et al. 2018)

Αναφέρεται επίσης στην έρευνα πως παρά το γεγονός ότι τα αποτελέσματα προήλθαν αποκλειστικά από UAV, εξακολουθούν να υπάρχουν πολλοί περιορισμοί, επομένως η μέθοδος απέχει πολύ από την ιδανική λύση και η ερευνά σκοπεύει κυρίως στο να παρουσιάσει τις μελλοντικές δυνατότητες αξιοποίησης των UAV. Περιορισμοί που αναφέρονται είναι ο εσφαλμένος εντοπισμός οχημάτων λόγω κρυφών σημείων, σκιών, αντικειμένων που περιόριζαν το οπτικό πεδίο προσωρινά, και λάθος προβλέψεις της διαδικασίας που προκάλεσαν σημαντικές παρεμβολές στα αποτελέσματα.

Η έρευνα των Kaufmann et al. (2018) αποσκοπεί στην παρατήρηση συγχρονισμένων ροών σε αστική κυκλοφορία υπό συμφόρηση. Εφαρμόζονται διάφορες τεχνικές για την επιτυχία απόσπασης μικροσκοπικών δεδομένων όπως δεδομένα για κάθε όχημα, για εμπειρική ανάλυση κυκλοφορίας και τροχιών (Εικόνα 2-6).

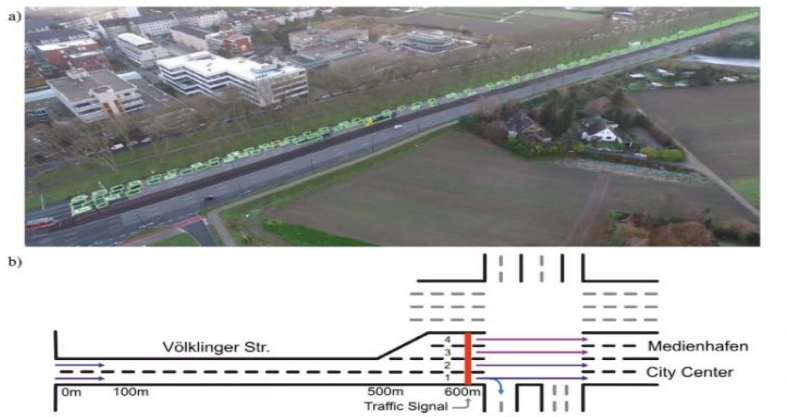


Fig. 1. Measurement setup: (a) video frame of the measured road section together with the tracked vehicles on the relevant lanes (unnecessary bottom part of the image has been omitted), leftmost vehicles are waiting on the stop line of the traffic signal, (b) Schematic of the measured road with lane numbers and traffic directions at the intersection.

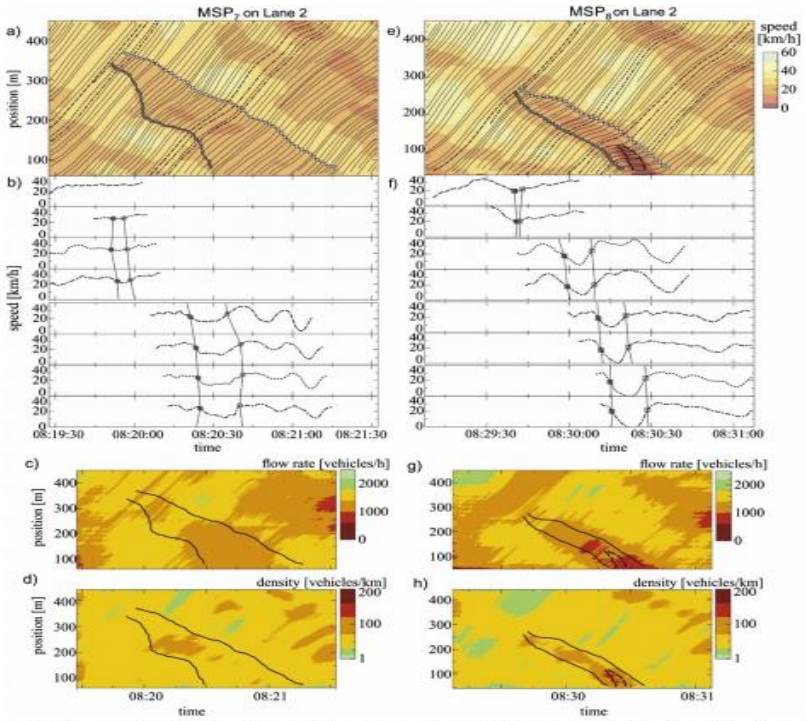


Fig. 11. Emerging of MSP_7 and MSP_8: (a, e) vehicle trajectories and mean speed over time and space, (b, f) speed according to the highlighted trajectories in (a, e) and in the same order, (c, g) traffic flow and (d, h) density over time and space. For the determination of the upstream and downstream fronts we have used the threshold $v_\theta = 25$ km/h for MSP_7 and $v_\theta = 20$ km/h for MSP_8.

Εικόνα 2-6: Εξαγωγή Κυκλοφοριακών Χαρακτηριστικών σε Αστική Κυκλοφορία υπό Συμφόρηση από Εναέρια Λήψη (Kaufmann et al. 2018).

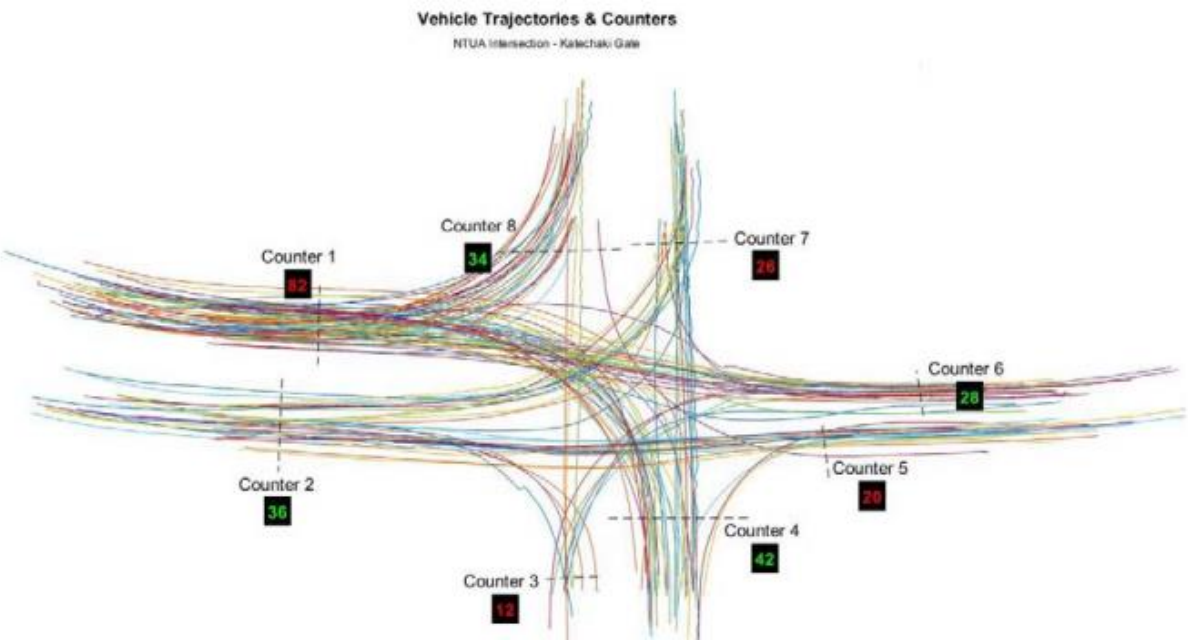
Η εφαρμογή της παρακολούθησης και εντοπισμού στόχων σε αστικό περιβάλλον είναι ένα ζήτημα που εξετάζουν οι Liu, et al. (2019). Η έρευνα εξετάζει την δυνατότητα ένταξης των UAV σε υπάρχοντα συστήματα παρακολούθησης εδάφους με σταθερές κάμερες που θα λειτουργούν συνεργατικά και η επεξεργασία των δεδομένων θα πραγματοποιείτε σε ένα απομακρυσμένο σύμπλεγμα υπολογιστών, με σκοπό την συνεχή παρακολούθηση ενός επιθυμητού στόχου σε αστικό περιβάλλον που η οπτική επαφή διακόπτεται συχνά (Εικόνα 2-7). Η εφαρμογή έχει πολύ ενδιαφέρον γιατί εξετάζει την δυνατότητα ένταξης των UAV σε υπάρχοντα συστήματα καταγραφής και παρακολούθησης κυκλοφορίας.

.

Εικόνα 2-7: Συνεργατικό Σύστημα UAV και Σταθερών Καμερών (Liu, et al. 2019).

Μία επίσης αντιπροσωπευτική έρευνα που παρουσιάζει τις δυνατότητες χρήσης UAV και μεθόδων υπολογιστικής όρασης, αλλά και τις δυσκολίες τους, παρουσιάζουν οι Barmpounakis et al. (2015). Η εξεταζόμενη περιοχή είναι μια διασταύρωση (Εικόνα 2-8), στην οποία καταγράφονται οι τροχιές οχημάτων και σχηματίζεται αντίστοιχος πίνακας προέλευσης – προορισμού (Εικόνα 2-9).

Εικόνα 2-8: Εξεταζόμενη Διασταύρωση Ερευνάς (Barmpounakis et al. 2015).

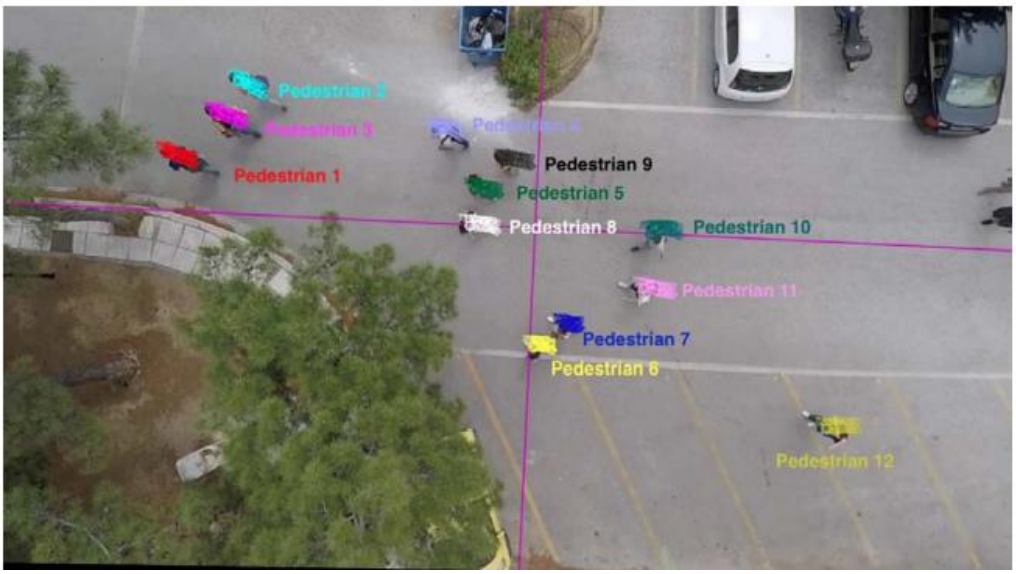


Destination

Origin: Counters	1	3	5	7	SUM
2		4	16	16	36
4	31		3	8	42
6	21	5		2	28
8	30	3	1		34
SUM	82	12	20	26	**140**

Εικόνα 2-9: Τροχιές Οχημάτων της Διασταύρωσης και πίνακας Προέλευσης – Προορισμού (Barmpounakis et al. 2015).

Επιπλέον παρουσιάζεται και η εφαρμογή καταγραφής τροχιάς και πεζών σε ένα χώρο (Εικόνα 2-10). Σε αυτήν την έρευνα αναφέρονται πάλι οι περιοριστικοί παράγοντες των UAV όπως η σταθερότητα λήψης, διάρκεια πτήσης και προβλήματα που αφορούν την ιδιωτικότητα και την ασφάλεια των εφαρμογών με χρήση UAV.



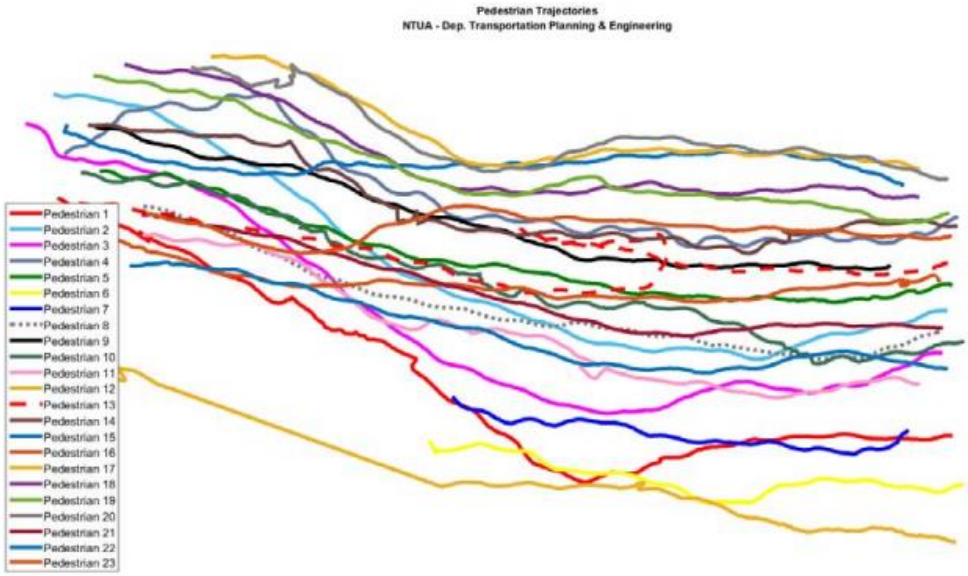


FIGURE 7: Graphical representation of pedestrians' trajectories

Εικόνα 2-10: Καταγραφή Τροχιάς Πεζών (Barmpounakis et al. 2015).

Τέλος μία άλλη έρευνα των Barmpounakis et al. (2017) παρουσιάζει τις υπάρχουσες εφαρμογές στον συγκοινωνιακό τομέα, μελλοντικές προκλησεις και υποσχόμενες εφαρμογές. Παρουσιάζεται ένα χρήσιμο διάγραμμα για την διαδικασία επιλογής μέσου και τεχνικής συλλογής και επεξεργασίας δεδοένων

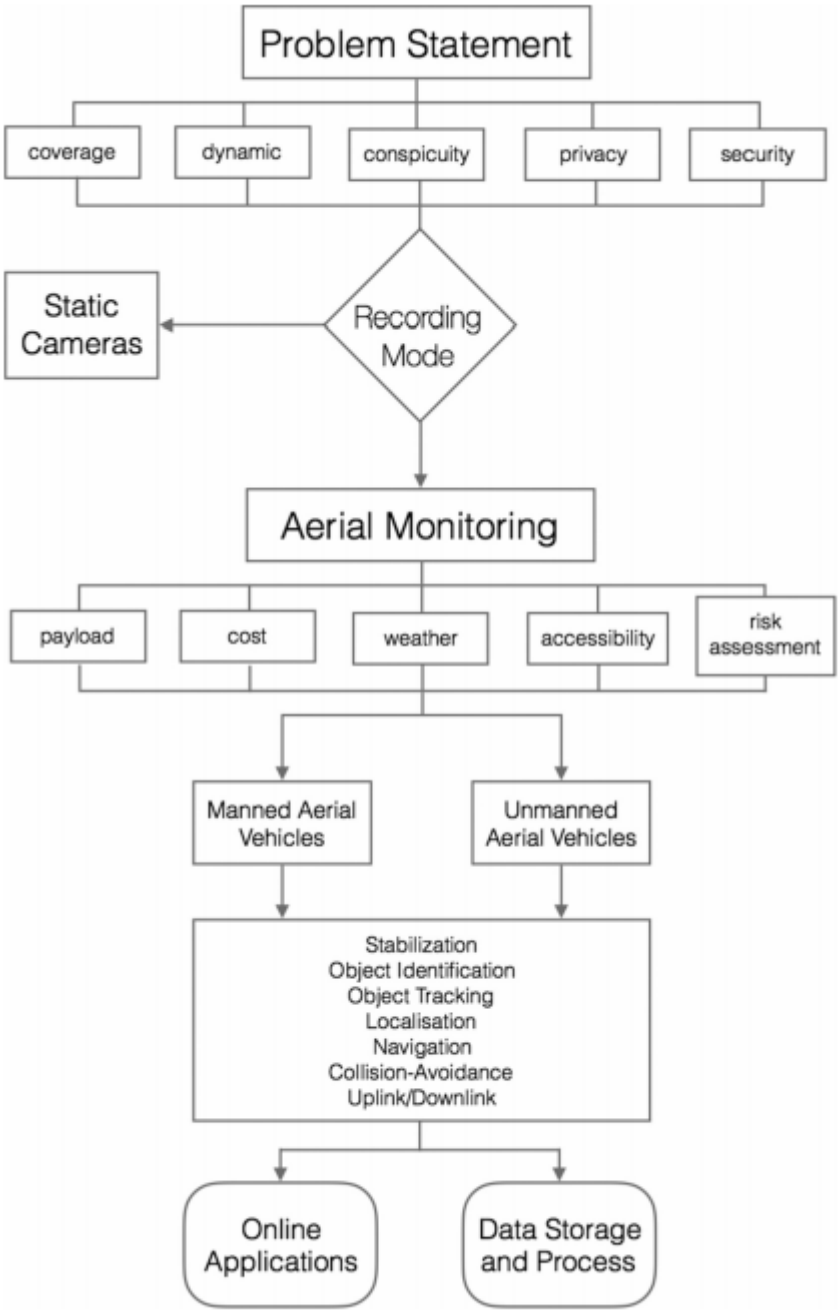


Fig. 2. Decision parameters for video recording mode choice.

Εικόνα 2-11: Παράμετροι Απόφασης για Τεχνική Καταγραφής Βίντεο (Barmpounakis et al. 2017)

2.3 - Σύγχρονες Εφαρμογές Βαθιάς Μάθησης

Από τις πρώτες έρευνες (LeCun et al., 1995) σχετικά με τα συνελικτικά νευρωνικά δίκτυα έχουν αναφερθεί τα σημαντικά πλεονεκτήματα των δικτύων για ανάλυση οπτικού υλικού και αναγνώριση μοτίβων, τονίζοντας το λαμπρό μέλλον που έχουν σε μελλοντικές εφαρμογές. Στην σύγχρονη κοινωνία, αρκετά χρόνια μετά τις πρώτες έρευνες, παρατηρείτε πως υπάρχει πληθώρα εφαρμογών συνελικτικών δικτύων, από καθημερινές εφαρμογές του μέσου ανθρώπου έως επιστημονικές έρευνες και εξειδικευμένες εφαρμογές. Τα τελευταία χρόνια η ανάπτυξη και έρευνα νέων προτύπων βαθιάς μάθησης σε συνδυασμό με τα άλματα που έχουν πραγματοποιηθεί στα υπολογιστικά μέσα και στον εξειδικευμένο εξοπλισμό που αξιοποιούν αυτές οι μέθοδοι, έχουν κάνει την χρήση μεθοδολογιών βαθιάς μάθησης πολύ προσιτές για έναν ερευνητή.

Στην έρευνα των Huang, et al. (2017), γίνεται μια εκτενής παρουσίαση διαφορετικών εφαρμογών, των πιο γνωστών προ-εκπαιδευμένων συνελικτικών νευρωνικών δικτύων και εξετάζονται παράγοντες που επηρεάζουν την απόδοση, ταχύτητα και ακρίβεια των μοντέλων. Γίνονται πολύ χρήσιμες αναφορές σε περιοριστικούς παράγοντές και στις απαιτήσεις υπολογιστικής ισχύς. Αυτή η μελέτη προσφέρει πολύ χρήσιμές πληροφορίες για την εξέταση εφαρμογών, και επίσης υποστηρίζεται από την πλατφόρμα του TensorFlow της Google, που προσφέρει γενικευμένες εφαρμογές σε μορφή ανοιχτού κώδικα.

Επόμενο είναι λοιπόν να έχουν εμφανιστεί και σχετικές εφαρμογές στον συγκοινωνιακό τομέα. Η πιο γνωστή εφαρμογή υπολογιστικής όρασης και βαθιάς μάθησης στον τομέα των μεταφορών είναι τα αυτόνομα αυτοκίνητα. Πλέον όλοι οι μεγάλοι κατασκευαστές οχημάτων έχουν μπει στον αγώνα των αυτόνομων και ημιαυτόνομων οχημάτων. Η πλατφόρμα DeepDrive του πανεπιστημίου Berkeley αναπτύσσει πολλά πρότζεκτ σχετικά με τα αυτόνομα οχήματα και την αναγνώριση και εντοπισμό αντικειμένων (Berkeley, 2018). Οι εφαρμογές αξιοποιούν μια εξαιρετική βάση δεδομένων (Yu, et al., 2018) για χρήση σε περιπλοκά συνελικτικά νευρωνικά δίκτυα όπως αυτό που έχουν δημιουργήσει, το SkipNet (Wang et al., 2018).

Όσον αφορά στον συγκοινωνιακό τομέα οι εφαρμογές βαθιάς μάθησης στην υπολογιστική όραση δεν είναι ακόμα πολύ διαδεδομένες, καθώς όπως αναφέρθηκε είναι ένας τομέας που έχει γίνει προσιτός τα τελευταία χρόνια και βρίσκεται σε συνεχή ανάπτυξη και βελτίωση. Παρόλα αυτά τεχνικές βαθιάς μάθησης και η αξιοποίησή συνελικτικών νευρωνικών δικτύων μπορούν να προσφέρουν λύσεις σε υπάρχοντα προβλήματα της επεξεργασίας εικόνας.

Η αναγνώριση χαρακτήρων στις πινακίδες αυτοκινήτων από βίντεο είναι από τις πρώτες εφαρμογές που εξετάστηκαν με χρήση συνελικτικών νευρωνικών δικτύων, και επομένως υπάρχουν πολλές ενναλακτικές έρευνες όπως των Rezaei et al. (2018) και των Zherzdez et al. (2018). Υπάρχει πληθώρα ερευνών καθώς οι προσεγγίσεις από τον κάθε ερευνητή είναι διαφορετικές ενώ συνεχώς εμφανίζονται νέες τεχνικές που βελτιώνουν τις υφιστάμενες μεθοδολογίες.

Εικόνα 2-12: Εντοπισμός Πινακίδων Κυκλοφορίας σε Εικόνα με Χρήση Συνελικτικών Νευρωνικών Δικτύων (Rezaei et al. 2018)

Εικόνα 2-13: Μετατροπή Εικόνας Πινακίδας Κυκλοφορίας σε Ψηφιακή Πληροφορία με Χρήση Συνελικτικών Νευρωνικών Δικτύων (Zherzdez et al. 2018)

Μία ακόμα εφαρμογή είναι η αναγνώριση οδικής σήμανσης με χρήση συνελικτικών νευρωνικών δικτύων. Στην έρευνα των Sermanet et al. (2011) παρουσιάζεται μία μέθοδος που το 2011 είχε την υψηλότερη ακρίβεια πρόβλεψης 99.14% (Εικόνα 2-14). Σημειώνεται επίσης πως τα πέντε καλύτερα μοντέλα ξεπερνούν στην ακρίβεια ενός ανθρώπου παρατηρητή που έχει υπολογιστεί στο 98.81%, γεγονός που δείχνει τις δυνατότητες αυτών των εφαρμογών.

#	Team	Method	Accuracy
	sermanet	**EBLearn 2LConvNet ms 108-108 + 100-feats CF classifier + No color**	**99.17%**
197	IDSIA	cnn_hog3	98.98%
196	IDSIA	cnn_cnn_hog3	98.98%
178	**sermanet**	**EBLearn 2LConvNet ms 108-108**	**98.97%**
	sermanet	**EBLearn 2LConvNet ms 108-200 + 100-feats CF classifier + No color**	**98.85%**
	sermanet	**EBLearn 2LConvNet ms 108-200 + 100-feats CF classifier + No color + Ramdom features + No jitter**	**97.33%**

Εικόνα 2-14: Βέλτιστα Πρότυπα Συνελικτικών Νευρωνικών Δικτύων για Σωστή Αναγνώριση Οδικής Σήμανσης

Οι έρευνες που συνδυάζουν πρότυπα βαθιάς μάθησης και οπτικό υλικό από μη επανδρωμένο αεροσκάφος δεν είναι πολλές και έχουν εμφανιστεί κατά κύριο λόγο το τελευταίο έτος (2018-2019).

Μία πρώτη έρευνα (Benjdira et al., 2018) συγκρίνει την απόδοση δύο διαφορετικών δομών συνελικτικών δικτύων για εντοπισμό αυτοκινήτων σε φωτογραφίες καταγεγραμμένες UAV. Τα εξεταζόμενα μοντέλα είναι τα Faster R_CNN και το Yolo V3. Τα αποτελέσματα των δύο μοντέλων είναι πολύ καλά, με το YOLOv3 να υπερτερεί και στην ακρίβεια προβλέψεων αλλά και στον μέσο χρόνο επεξεργασίας κάθε εικόνας(Εικόνα 2-15). Αξίζει να σημειωθεί πως τα μοντέλα είναι αυξημένης περιπλοκότητας και μεγέθους επομένως είναι πολύ ακριβή με αντάλλαγμα την ταχύτητα ανάλυσης της εικόνας. Η εφαρμογή τους σε βίντεο υψηλού ρυθμού καταγραφής δεν θα είναι ιδανική καθώς η ανάλυση θα γίνεται με χαμηλότερο ρυθμό από αυτόν που παράγονται οι εικόνες. Η έρευνα περιορίζεται στην αξιολόγηση των μοντέλων.

Measure	Faster R-CNN (test dataset)	YOLOv3 (test dataset)
TP (True positives)	578	751
FP (False positives)	2	2
FN (False negatives)	150	7
Precision (TPR)	99.66%	99.73%
Sensitivity (recall)	79.40%	99.07%
F1 Score	88.38%	99.94%
Quality	79.17%	98.81%
Processing time (Av. in ms)	1.39 s	0.057 ms

TABLE I

EVALUATION METRICS OF FASTER R-CNN AND YOLOv3

Εικόνα 2-15: Αποτελέσματα Μοντέλων Faster R-CNN και YOLOv3 Για Εντοπισμό Οχημάτων (Benjdira et al., 2018)

Η έρευνα των Wang et al. (2018) είναι παρόμοια, εξετάζοντας διαφορετικά είδη μοντέλων για εντοπισμό αντικειμένων από UAV. Η ανάλυση περιορίζεται στην απόδοση των μοντέλων προσφέροντας χρήσιμες πληροφορίες για τις ιδιότητες και την συμπεριφορά τους σε αυτό το σενάριο.

Τελική αναφορά θα γίνει σε μία πολύ ενδιαφέρουσα έρευνα και εφαρμογή (Kyrkou et al., 2018) που αποσκοπεί στον εξοπλισμό ενός UAV με συσκευή Raspberry Pi 3 για επιτόπου εντοπισμό οχημάτων σε πραγματικό χρόνο από το σύστημα του UAV- Raspberry Pi 3. Εξετάζεται η απόδοση υπάρχων μοντέλων μαζί με το DroNet που αναπτύχθηκε στα πλαίσια της έρευνας. Τα αποτελέσματα της έρευνας περιορίζονται από τον εξοπλισμό ανάλυσης Raspberry Pi 3 καθώς ο ρυθμός επεξεργασίας εικόνων κυμαινόταν στις 5-6 εικόνες το δευτερόλεπτο που είναι χαμηλός για αξιόπιστη χρήση σε πραγματικό χρόνο.

2.4 - Γενικά Συμπεράσματα Βιβλιογραφικής Ανασκόπησης

Η επεξεργασία εικόνας αποτελεί πλέον βασικό κομμάτι της καταγραφής και ανάλυσης κυκλοφοριακών δεδομένων με μεγάλο πλήθος εφαρμογών να έχουν καθιερωθεί στον συγκοινωνιακό τομέα.

Η ταυτόχρονη εξέλιξη των μη επανδρωμένων αεροσκαφών τα τελευταία χρόνια προσφέρει νέες δυνατότητες στους ερευνητές για νέες η βελτιωμένες εφαρμογές συλλογής πληροφορίας. Οι σχετικές έρευνες αναζητούν νέες τεχνικές με σκοπό να ξεπεραστούν οι περιορισμοί που υπάρχουν από τα καθιερωμένα μέσα καταγραφής εικόνας, όπως οι σταθερές κάμερες. Η πλειοψηφία των ερευνητών είναι αισθητά αισιόδοξη ως προς την σημαντικότητα και την καθιέρωση των μη επανδρωμένων αεροσκαφών ως βασικό μέσο καταγραφής οπτικού υλικού στον τομέα των μεταφορών, για επιστημονικούς σκοπούς και για υπολογισμό κυκλοφοριακών συνθηκών και μεγεθών.

Παρά τις πρωτοτυπίες που παρουσιάζουν οι έρευνες η, πλειοψηφία τους αναφέρει τους ίδιους περιοριστικούς παράγοντες που εμφανίστηκαν λόγω των μη επανδρωμένων αεροσκαφών. Τονίζονται μερικά σημεία όπως η ανάγκη σταθερής λήψης από τα μη επανδρωμένα αεροσκάφη, ο περιορισμός διάρκειας πτήσης λόγω μπαταριών, ο περιορισμός λόγω καιρικών συνθηκών για πτήση των αεροσκαφών και οι ανησυχίες που αφορούν στα θέματα ιδιωτικότητα και ασφαλείας. Βέβαια κρίνοντας από τον ρυθμό των τεχνολογικών εξελίξεων οι ερευνητές αναφέρουν πως είναι θέμα χρόνου η επίλυση των κύριων περιορισμών.

Οι τεχνικές που αξιοποιούν την εικόνα ως δεδομένο ανήκουν στην έννοια υπολογιστική όραση. Αντίστοιχα οι έρευνες που αναφέρθηκαν αξιοποιούν τεχνικές υπολογιστικής όρασης για την επεξεργασία της εικόνας από μη επανδρωμένο αεροσκάφος. Ένα χαρακτηριστικό σε αυτές τις μεθόδους είναι πως το οπτικό υλικό που καταγράφεται συνήθως απαιτεί σημαντική επεξεργασία για να εφαρμοστούν τεχνικές υπολογιστικής όρασης. Η επεξεργασία αυτή μπορεί να είναι σταθεροποίηση του υλικού με κατάλληλες τεχνικές, να αντιμετωπιστεί το πρόβλημα εμφάνισης σκιών που επηρεάζουν πολύ έντονα τα αποτελέσματα και σε μερικές περιπτώσεις απαιτείται η εφαρμογή φίλτρων για κατάλληλη παραμόρφωση της εικόνας. Αυτή η διαδικασία ονομάζεται ρύθμιση-καλιμπράρισμα (calibration) του οπτικού υλικού, και σε πολλές περιπτώσεις απαιτεί σημαντικό χρόνο επεξεργασίας από τον ερευνητή. Επίσης για διαφορετικά δεδομένα και άλλες συνθήκες στο οπτικό υλικό είναι πολύ πιθανόν η διαδικασία αυτή να απαιτείτε ξανά για εφαρμογή μίας μεθοδολογίας. Σημαντικό είναι να διευκρινιστεί πως αυτές οι μεθοδολογίες δεν εφαρμόζονται σε πραγματικό χρόνο αλλά η επεξεργασία των δεδομένων είναι μία διαδικασία που ακολουθεί την συλλογή τους.

Σε αυτό το σημείο θα αναφερθούν οι τεχνικές βαθιάς μάθησης στην υπολογιστική όραση, οι οποίες ίσως μπορούν να αντιμετωπίσουν μερικά από τα παραπάνω προβλήματα. Η βιβλιογραφία παρουσιάζει την εξαιρετική ικανότητα των συνελικτικών νευρωνικών δικτύων να αναλύουν οπτικό υλικό. Αυτό επιβεβαιώνεται από την ευρεία χρήση τους σε εφαρμογές όπως τα αυτόνομα οχήματα, και σε άλλες εφαρμογές όπως αυτές που αναφέρθηκαν στην βιβλιογραφική ανασκόπηση.

Η βασική ιδέα της βαθιάς μάθησής είναι ότι ο ερευνητής θα επιχειρήσει να αντιμετωπίσει ένα πρόβλημα, ιδανικά, με πολύ μεγάλο πλήθος διακριτών δεδομένων, δημιουργώντας μέσα από μία διαδικασία δοκιμών και σφαλμάτων ένα βέλτιστο μοντέλο που θα μπορεί να εφαρμοστεί

σε γενικευμένες και ξένες περιπτώσεις του αρχικού προβλήματος χωρίς να χρειαστεί κάποια εκ νέου επεξεργασία από τον ερευνητή.

Μερικές έρευνες αξιοποιούν συνελικτικά νευρωνικά δίκτυα σε δεδομένα από εναέρια λήψη για να εξετάσουν την απόδοση διαφορετικών μοντέλων και την δυνατότητα εφαρμογής αυτών σε πραγματικό χρόνο. Αυτές οι έρευνές θα καθοδηγήσουν την επιλογή χαρακτηριστικών και μοντέλων στο πλάσιμο της διπλωματικής εργασίας.

Τέλος, συμπεραίνεται από την βιβλιογραφία πως οι εφαρμογές βαθιάς μάθησης σε δεδομένα από μη επανδρωμένο αεροσκάφος δεν έχουν επεκταθεί έντονα ακόμα. Αυτές οι εφαρμογές είναι που θα εξεταστούν και στο πλαίσιο της παρούσας διπλωματικής εργασίας.

3 - ΜΕΘΟΔΟΛΟΓΙΑ

3.1 - Γενικά

Στο κεφάλαιο αυτό περιγράφεται αναλυτικά η προτεινόμενη μεθοδολογική προσέγγιση για την ανάπτυξη και αξιολόγηση προτύπων, ως προς την αποτελεσματικότητα τους και την εφαρμοσιμότητα τους, με χρήση εφαρμογών μηχανικής μάθησης (machine learning), βαθιάς μάθησης(deep leaning), και υπολογιστικής όρασης (computer vision), για την αναγνώριση κυκλοφοριακών δεδομένων σε βίντεο καταγεγραμμένα από μη επανδρωμένο αεροσκάφος (UAV), μια σύγχρονη μέθοδος συλλογής πληροφορίας σε μορφή εικόνας και βίντεο

Συγκεκριμένα θα παρουσιαστούν τα χαρακτηριστικά των προσεγγίσεων για την συλλογή και ανάλυση κυκλοφοριακών δεδομένων από κοινή βάση δεδομένων μορφής βίντεο. Οι δύο προσεγγίσεις θα εξετάσουν δυο διαφορετικές εφαρμογές υπολογιστικής όρασης με χρήση συνελικτικών νευρωνικών δικτύων και θα αξιολογήσουν πιθανές εφαρμογές, πλεονεκτήματα, μειονεκτήματα και περιοριστικούς παράγοντες.

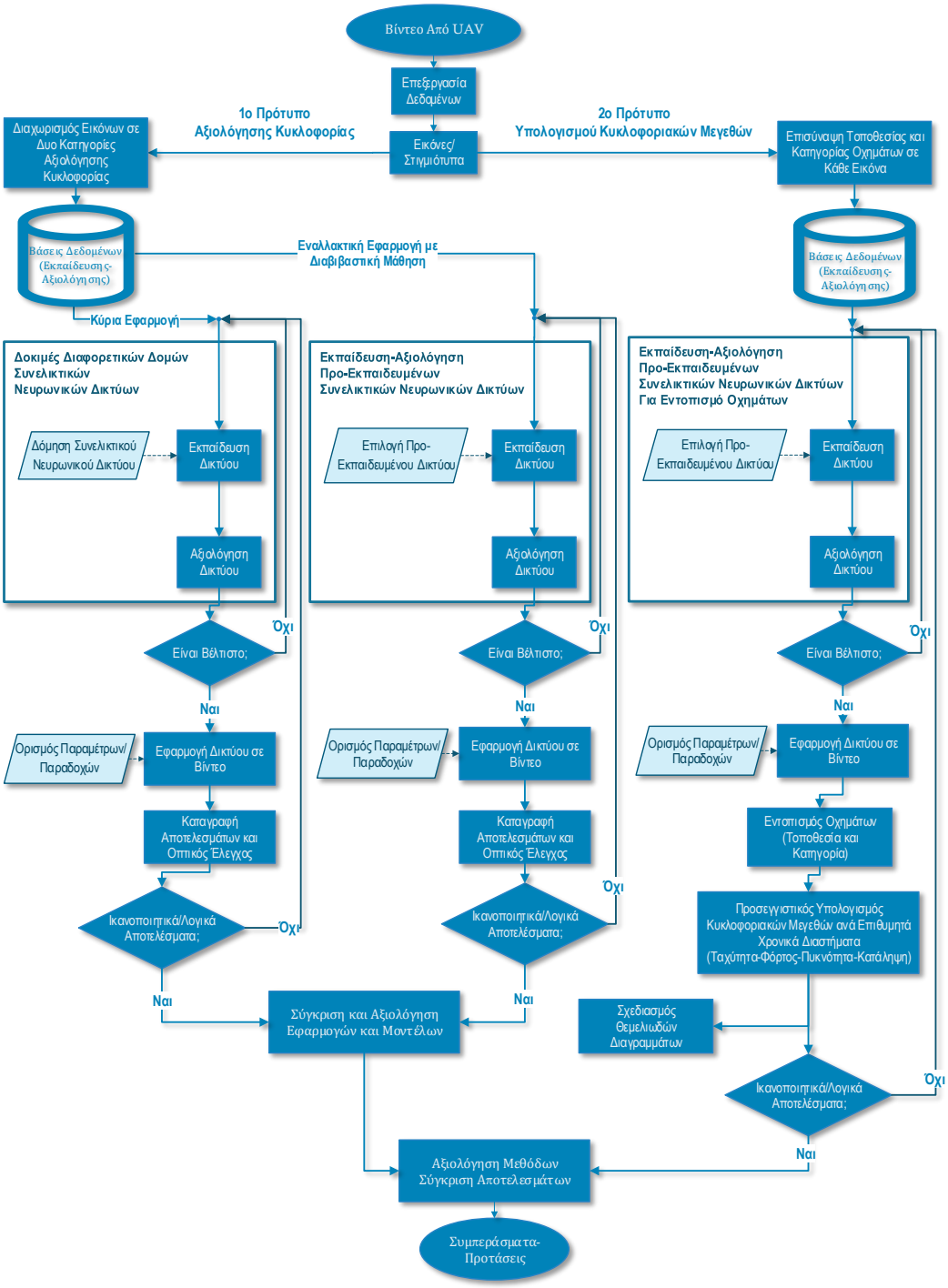
3.2 - Παραδοχές

Η φύση της διπλωματικής εργασίας και των διαθέσιμων δεδομένων με μορφή βίντεο απαιτούν τη διατύπωση ορισμένων παραδοχών, για την κατασκευή και εφαρμογή προτύπων, καθώς στα αρχικά στάδια υπάρχει σημαντικό πλήθος περιορισμών. Οι βασικότερες παραδοχές είναι οι εξής:

- Δεδομένα από την οδό Κατεχάκη στο ύψος της Πολυτεχνειούπολης (Αστικό περιβάλλον).
- Εξεταζόμενη κυκλοφορία μόνο στην άνοδο της παραπάνω οδού (3 λωρίδες).
- Καταγραφή δεδομένων κατά την διάρκεια της ημέρας (κυρίως μεσημβρινές ώρες με επαρκή φωτεινότητα επι της οδού).
- Αίθριες καιρικές συνθήκες (Ηλιοφάνεια ή ήπια συννεφιά).
- Μέσο μήκος αυτοκινήτου ίσο με 4,5 μέτρα.
- Τήρηση του κώδικα οδικής ασφαλείας από τους οδηγούς.

3.3 - Περιγραφή Προτεινόμενων Προσεγγίσεων

Η επεξεργασία βίντεο από μη επανδρωμένο αεροσκάφος (UAV) παρουσιάζει πολλές προκλήσεις, επομένως στο πλαίσιο της διπλωματικής εργασίας εξετάστηκε μεγάλο πλήθος μεθοδολογιών, για την ανάπτυξη προτύπων, ως προς την εφαρμοσιμότητα τους. Τελικώς επιλέχτηκαν και αναπτύχθηκαν δύο πρότυπα, των οποίων η μεθοδολογία θα παρουσιαστεί στη συνέχεια. Η συνολική διαδικασία περιγράφεται σχηματικά στο Διάγραμμα 3-1 και αναλύεται στις παραγράφους που ακολουθούν.



Διάγραμμα 3-1: Διάγραμμα Ροής με τις δύο Διαφορετικές Προσεγγίσεις

3.3.1 - Αναγνώριση κυκλοφορικών συνθηκών σε πραγματικό χρόνο

Σκοπός της πρώτης μεθοδολογίας είναι η δημιουργία προτύπου ανάλυσης βίντεο από UAV σε πραγματικό χρόνο, και στην τοποθεσία λήψης. Η εφαρμογή αυτή παρουσιάζει μία σημαντική πρόκληση, καθώς η επεξεργασία βίντεο, συγκεκριμένα υψηλής ανάλυσης, λόγω του τεράστιου πλήθους πληροφορίας που περιέχει, απαιτεί σημαντική υπολογιστική ισχύι, κάτι που δεν είναι πάντα εφικτό με φορητές συσκευές (κινητά τηλέφωνα, φορητοί υπολογιστές).

Για να αντιμετωπιστεί αυτός ο περιορισμός σε έναν βαθμό, θα δομηθούν, εκπαιδευτούν συνελικτικά νευρωνικά δίκτυα εκ νέου, και θα αξιολογηθούν με τα δεδομένα από τα βίντεο, που θα έχουν επεξεργαστεί κατάλληλα για αποκοπή περιττής πληροφορίας.

Αρχικά, τα βίντεο θα περικοπούν και θα απομονωθούν μόνο τα επιθυμητά τμήματα όσον αφορά χρονικά αλλά και ως προς το τμήμα της εικόνας που περιέχει το τμήμα της οδού που θα εξεταστεί. Στην συνέχεια θα γίνει διαχωρισμός των τμημάτων βίντεο από τις επιθυμητές περιοχές και τμήματα οδού σε δύο κατηγορίες κυκλοφορίας, Υψηλή Κυκλοφορία (κατηγορία 1) και Χαμηλή / Καθόλου Κυκλοφορία (κατηγορία 0). Έπειτα θα εξαχθούν εικόνες από τα παραπάνω για δημιουργία βάσεων δεδομένων για την εκπαίδευση και αξιολόγηση των συνελικτικών νευρωνικών δικτύων στις δύο κατηγορίες. Ο διαχωρισμός πραγματοποιήθηκε αρχικά με αυθαίρετα κριτήρια για την αρχική αξιολόγηση της μεθόδου, και στην συνέχεια οριστικοποιήθηκαν κατάλληλα κριτήρια διαχωρισμού των εικόνων ανάλογα το πλήθος και την θέση των οχημάτων σε κάθε εικόνα.

Με τα δεδομένα μοιρασμένα σε κατάλληλες αναλογίες, θα ακολουθήσει μία διαδικασία στην οποία θα δομούνται διαφορετικά συνελικτικά νευρωνικά δίκτυα, με διαφορετικές παραμέτρους, θα εκπαιδεύονται και θα αξιολογούνται και τελικώς θα συγκρίνονται με βάση την ακρίβειά και στατιστικά αποτελέσματα τους για την επιλογή των βέλτιστων. Έχοντας απομονώσει τα βέλτιστα δίκτυα, θα εφαρμοστούν σε βίντεο που δεν έδωσε πληροφορία για την εκπαίδευση τους, και θα οριστεί ένα χρονικό διάστημα (π.χ. 10 δευτερόλεπτα). Θα πραγματοποιηθεί πρόβλεψη κατηγορίας για κάθε εικόνα-στιγμιότυπο του βίντεο και για κάθε χρονικό διάστημα θα παράγεται μια κατηγορία κυκλοφορίας που θα το χαρακτηρίζει, με χρήση αλγόριθμου επιλογής πλειοψηφίας. Έπειτα θα εξεταστούν τα αποτελέσματα όσον αφορά στο αν είναι λογικά και αντιπροσωπεύουν το βίντεο που εξετάστηκε και όσον αφορά στην χρονική διάρκεια και υπολογιστικές απαιτήσεις της ανάλυσης, ώστε να καθοριστεί αν μπορεί να γίνει η ανάλυση σε πραγματικό χρόνο, δηλαδή σε βίντεο ζωντανής μετάδοσης.

Εξετάστηκε επίσης η ίδια μεθοδολογία με χρήση προ-εκπαιδευμένων συνελικτικών δικτύων έναντι δόμησης νέων, για σύγκριση των δύο ως προς τα αποτελέσματα τους σε κοινό δείγμα αξιολόγησης αλλά και ως προς τις απαιτήσεις υπολογιστικής ισχύς. Η μεθοδολογία κατά κύριο λόγο ακολουθεί την ροή που αναφέρθηκε προηγουμένως με μερικές διαφοροποιήσεις. Μία είναι οι δοκιμές και η προετοιμασία των προ-εκπαιδευμένων δικτύων καθώς και η προ-επεξεργασία των δεδομένων ανάλογα τις απαιτήσεις του κάθε δικτύου, ώστε να είναι εφαρμόσιμα στο πλαίσιο της μεθοδολογίας. Το μοντέλο που εξετάστηκε τελικά και εφαρμόζεται καλύτερα στα δεδομένα της εργασίας είναι το ResNet 50.

Τέλος θα σχολιαστούν τα αποτελέσματα, η απαιτούμενη διάρκεια ανάλυσης και υπολογιστικής ισχύς των εναλλακτικών δικτύων.

3.3.2 - Μακροσκοπική αναγνώριση κυκλοφορικών συνθηκών

Η δεύτερη μεθοδολογία αποσκοπεί στην δημιουργία προτύπου που θα μπορεί να εξάγει από βίντεο, καταγεγραμμένο από UAV, κυκλοφοριακά μεγέθη. Ο στόχος θα πραγματοποιηθεί εντοπίζοντας οχήματα ως προς την κατηγορία τους και την θέση τους στο εξεταζόμενο τμήμα οδού και σε συνέχεια θα υπολογίζονται, με βάση τα προηγούμενα, κυκλοφοριακά μεγέθη όπως φόρτος, ταχύτητα, πυκνότητα και κατάληψη στο ίδιο τμήμα οδού. Να σημειωθεί πως το βίντεο είναι ακαλιμπράριστο και χωρίς πραγματικά κυκλοφοριακά δεδομένα επί της οδού.

Οι περιορισμοί που υπάρχουν στην επεξεργασία βίντεο, όπως αναφέρθηκαν στην πρώτη μεθοδολογία, ισχύουν και στο πλαίσιο της δεύτερης μεθοδολογίας, επομένως θα αποτελέσουν πάλι σημαντικό κριτήριο επιλογής προτύπου.

Αρχικά θα απομονωθεί ένα πλήθος εικόνων από διαφορά τμήματα των δεδομένων βίντεο, οι οποίες θα περιέχουν πολλά οχήματα διαφορετικών κατηγοριών, όπως αυτοκίνητα, φορτηγά, βανάκια, μοτοσυκλέτες. Έπειτα με χρήση κατάλληλου προγράμματος θα γίνει επισύναψη, με κατάλληλη μορφή, σε κάθε εικόνα των πληροφοριών θέσης και κατηγορίας κάθε οχήματος, και θα χωριστούν τυχαία σε δύο βάσεις δεδομένων για εκπαίδευση και αξιολόγηση.

Σε αυτήν την μεθοδολογία θα εξεταστεί η εφαρμογή μιας σύγχρονής πιο περιπλοκής μορφής συνελικτικών δικτύων. Αυτά τα δίκτυα ονομάζονται Συνελικτικά Νευρωνικά Δίκτυα Πρότασης Περιοχής (Region Proposal Convolutional Neural Networks ή R-CNN) και Δίκτυα Εντοπισμού Πολλαπλών Πλαισίων με Μία Λήψη (Single Shot Multibox Detection Network – SSD Network). Με βάση την βιβλιογραφία και την φύση την διπλωματικής εργασίας, έπειτα από δοκιμές ενναλακτικών, έγινε επιλογή του δικτύου SSD MobilNet V2. Το δίκτυο αυτό είναι αρκετά γρήγορο και ελαφρύ όσον αφορά στην υπολογιστική ισχύ που απαιτεί, ενώ επίσης είναι δομημένο για βέλτιστη εφαρμογή του σε συσκευές χαμηλής υπολογιστικής ισχύς, όπως συσκευές Android, γεγονός που ενδείκνυται στο πλαίσιο της παρούσας διπλωματικής εργασίας.

Στην συνέχεια θα ρυθμιστούν οι παράμετροι του δικτύου, θα εκπαιδευτεί και θα αξιολογηθεί με τα δεδομένα του πειράματος. Όταν τα αποτελέσματα του μοντέλου είναι τα βέλτιστα δυνατά θα εφαρμοστεί στο βίντεο που χρησιμοποιήθηκε και στην πρώτη μεθοδολογία, και θα εξεταστεί οπτικά για έλεγχο σωστής και αναμενόμενης λειτουργείας όσον αφορά στον σωστό εντοπισμό και καθορισμό των οχημάτων.

Εφόσον το πρόγραμμα λειτουργεί σωστά, θα οριστεί ένα χρονικά διάστημα (π.χ. 10 δευτερόλεπτα), οπού στο τέλος κάθε χρονικού διαστήματος θα επιστρέφονται από το πρόγραμμα οι τιμές των κυκλοφοριακών μεγεθών αυτού, ή όποια άλλη πληροφορία απαιτηθεί από την εφαρμογή. Στο τέλος του βίντεο θα σχεδιαστούν και τα θεμελιώδη διαγράμματα των κυκλοφοριακών μεγεθών κάθε χρονικής περιόδου.

Τέλος θα σχολιαστούν τα αποτελέσματα, η απαιτούμενη διάρκεια ανάλυσης και υπολογιστικής ισχύς.

3.4 - Θεωρητικό Υπόβαθρο

Οι μεθοδολογίες βασίζονται στο μεγαλύτερο τμήμα τους στην λειτουργεία συνελικτικών νευρωνικών δικτύων διαφορετικών μορφών. Η λειτουργεία τους είναι σύνθετη και είναι ένα αντικείμενο που ερευνάτε και εξελίσσεται συνεχώς. Θα ακολουθήσει μία σύντομη αναφορά στις γενικές έννοιες της βαθιάς μάθησης και των νευρωνικών δικτύων και στην συνέχεια μία εκτενέστερη επεξήγηση στα συνελικτικά νευρωνικά δίκτυα και τις διάφορες μορφές τους, καθώς αυτά αξιοποιήθηκαν στην διπλωματική εργασία.

Επίσης θα γίνει αναφορά στα κυκλοφοριακά μεγέθη που υπολογίζονται και στα στατιστικά μεγέθη αξιολόγησης που αξιοποιούνται κατά την διάρκεια της εργασίας για την αξιολόγηση των αποτελεσμάτων.

3.4.1 - Κυκλοφοριακά Μεγέθη

Τα κυκλοφοριακά μεγέθη που υπολογίζονται στην δεύτερη μεθοδολογία είναι ο κυκλοφοριακός φόρτος, η πυκνότητα, η κατάληψη και η μέση ταχύτητα.

Οι μαθηματικοί τύποι προέκυψαν από το βιβλίο Κυκλοφοριακή Τεχνική (Φραντζεσκάκης et al., 2009).

Ο Κυκλοφοριακός Φόρτος – *q*, δίνεται από τη Σχέση 1:

$$q(x) = \frac{n(x)}{T} \tag{1}$$

Τη περίοδο μέτρησης, *n(x)* είναι ο αριθμός οχημάτων που περνούν από μία διατομή στην θέση x κατά το χρονικά διάστημα *T*.

Η Πυκνότητα – *k*, δίνεται για συγκεκριμένη χρονική στιγμη *t* από τη Σχέση 2 και για συγκεκριμένο χρονικό διάστημα *T* από τη Σχέση 3:

$$k(t) = \frac{n(t)}{L} \tag{2}$$

Όπου t η εξεταζόμενη χρονική στιγμή, στην οποία σε μήκος οδού *L* βρίσκεται *n(t)* πλήθος οχημάτων.

$$k(T) = \frac{\textit{μέσος αριθμός οχημάτων στο εξεταζόμενο οδικό τμήμα (για το δίαστημα T)}}{L} \tag{3}$$

Όπου η τιμή προκύπτει από τον μέσο αρθμό των οχημάτων που βρίσκονται στο εξεταζόμενο τμήμα οδού μήκου *L*, κατά το χρονικό διάστημα *T*.

Η τιμή της Κατάληψης – *O* δίνεται από τη Σχέση 4:

$$O(x) = \frac{\sum_{i}^{n(x)} \frac{L_i}{u_i(x)}}{T} \tag{4}$$

Η κατάληψη για το χρονικό διάστημα *T* ισούται με το αθροισμα του χρόνου κατά τον οποίο τα οχήματα καλύπτουν μια διατομή, προς το χρονικό διαάστημα *T*. Για κάθε όχημα *i* , ο χρόνος που καταλαμβάνει μια διατομή στην θέση *x* εξαρτάται από την ταχύτητα του οχήματος $u_i(x)$ στην θέση *x* και το μήκος του οχήματος L_i.

Η Μέση Ταχύτητα Χρόνου - u_t δίνεται από τη Σχέση 5:

$$u_t = \frac{1}{n(x)} \sum_{i}^{n(x)} u_i(x) \tag{5}$$

Η μέση ταχύτητα χρόνου προκύπτει από το άθροισμα της ταχύτητας $u_i(x)$ κάθε οχήματος *i* στην θέση *x* προς το πλήθος των οχημάτων που πέρασαν από την διατομή στην θέση *x* το χρονικό διάστημα *T*.

Η μέση ταχύτητα χώρου – u_s δίνεται από τη Σχέση 6:

$$u_s = \frac{N(T) * L}{\sum_{j=1}^{N(T)} t_j} \tag{6}$$

Όπου t_j είναι ο χρόνος που χρειάζεται το όχημα *j*, από το πλήθος *N(T)* οχημάτων ποου διένυσαν ένα οδικό τμήμα μήκους *L* σε χρονικό διάστημα *T*, για να διανύσει ολόκληρο το εν λόγω οδικό τμήμα.

Τέλος η Θεμελιώδης Σχέση της Κυκλοφοριακής Ροής δίνεται απο τη Σχέση 7:

$$q = k * u_s \tag{7}$$

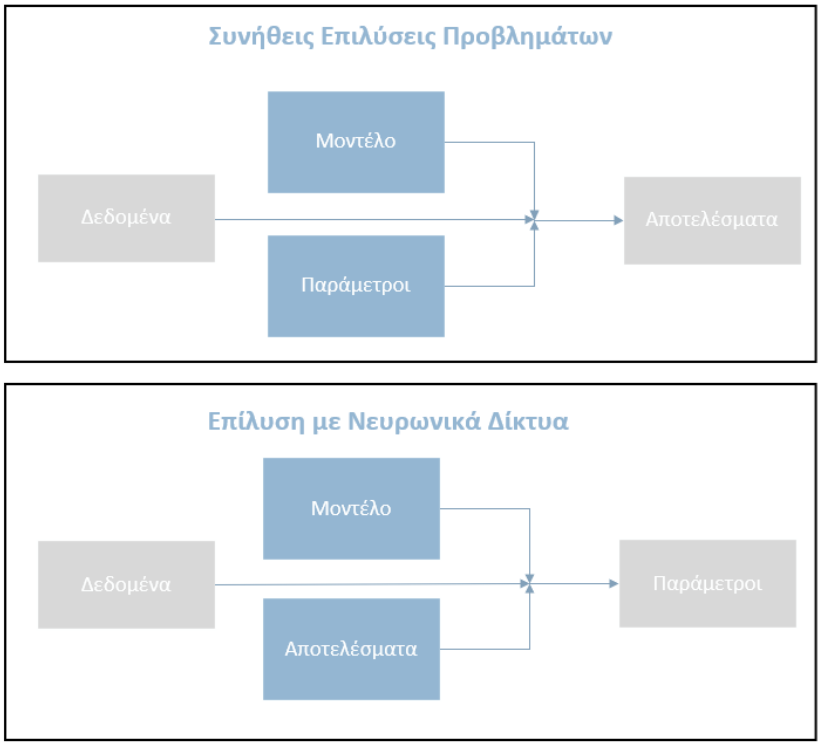
Όπους q – κυκλοφοριακός φόρτος (Σχέση 1), k – πυκνότητα (Σχέση 3) και u_s – μέση ταχύτητα χώρου (Σχέση 6)

3.4.2 - Υπολογιστική Όραση

Εισαγωγικά Βαθιάς Μάθησης και Νευρωνικών Δικτύων

Η έννοια της βαθιάς μάθησης (Deep Learning) ανήκει στην ευρεία οικογένεια των εφαρμογών μηχανικής μάθησης (Machine learning), και περιέχει μεγάλη ποικιλία αρχιτεκτονικών νευρωνικών δικτύων (Neural Networks), ένα από τα οποία είναι τα συνελικτικά νευρωνικά δίκτυα που χρησιμοποιούνται σε εφαρμογές υπολογιστικής όρασης (Computer Vision).

Οι εφαρμογές βαθιάς μάθησης έχουν μια βασική διαφορά με τις κλασσικές εφαρμογές μοντελοποίησης. Σε ένα τυπικό πρόβλημα δεδομένα είναι το αξιοποιήσιμο μοντέλο και οι παράμετροι του, οι οποίοι τροποποιούνται έως τα αποτελέσματα να είναι επιθυμητά. Στις εφαρμογές νευρωνικών δικτύων, ως δεδομένα είναι το εξεταζόμενο μοντέλο και το ζεύγος δεδομένων και σωστών αποτελεσμάτων τους, ενώ και το αποτέλεσμα της μεθόδου είναι οι κατάλληλοι παράμετροι για την επίτευξη της σωστής αντιστοιχίας δεδομένων-αποτελεσμάτων (Εικόνα 3-1).

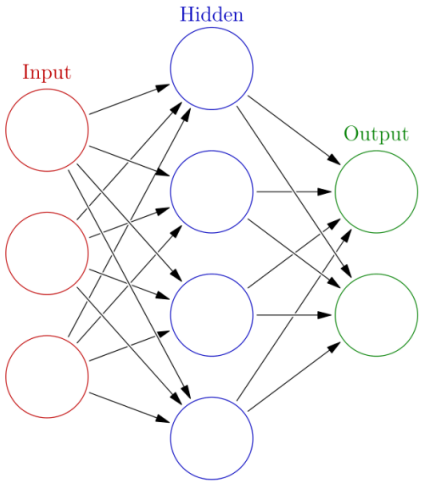


Εικόνα 3-1: Σύγκριση Τυπικής Επίλυσής και Επίλυσης με Νευρωνικά Δίκτυα

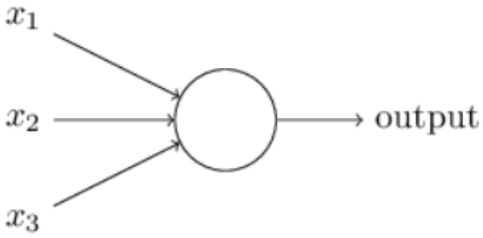
Η εφαρμογή νευρωνικών δικτύων δύναται να προσομοιώσει την λειτουργεία ενός εγκεφάλου, χρησιμοποιώντας την ιδέα των νευρώνων και της αλληλεξάρτησης τους καθώς βασίζεται στον τρόπο επικοινωνίας των νευρώνων σε ένα νευρικό σύστημα.

Η δομή των **νευρωνικών δικτύων** αποτελείται από τρία βασικά επίπεδα, το επίπεδο εισόδου, το επίπεδο εξόδου και τα ενδιάμεσα η διαφορετικά τα κρυφά επίπεδα, τα οποία αποτελούνται από νευρώνες. Το επίπεδο εισόδου είναι το επίπεδο που θα δεχθεί τα δεδομένα για την εκπαίδευση και έπειτα τα δεδομένα για την πρόβλεψη. Αυτό το επίπεδο αποτελείται από ένα σύνολο νευρώνων, το οποίο εξαρτάται συνήθως από τα δεδομένα, όπου το σύνολο αυτό

επικοινωνεί με το κρυφό επίπεδο. Το ενδιάμεσο κρυφό επίπεδο μπορεί να αποτελείται από οποιοδήποτε πλήθος υπό-επιπέδων και νευρώνων, τα οποία είναι υπεύθυνα για την επεξεργασία των δεδομένων. Όσο πιο πολλά και περίπλοκα επίπεδα υπάρχουν το μοντέλο χαρακτηρίζεται ως βαθύ νευρωνικό δίκτυο. Εφόσον γίνει η επεξεργασία στο κρυφό επίπεδο θα ενεργοποιηθούν οι νευρώνες του επιπέδου εξόδου, όπου θα δώσουν το αποτέλεσμα που αντιστοιχεί στα δεδομένα εισόδου. Να σημειωθεί πως οι νευρώνες κάθε επιπέδου δεν επικοινωνούν μεταξύ τους, αλλά ο κάθε νευρώνας επικοινωνεί με όλους τους νευρώνες των γειτονικών επιπέδων, όπως εμφανίζεται ενδεικτικά και στις παρακάτω εικόνες (Εικόνα 3-2, Εικόνα 3-3).



Εικόνα 3-2: Τυπική Μορφή Νευρωνικού Δικτύου (Wikipedia, 2019)



Εικόνα 3-3: Τυπική Μορφή Νευρώνα (Nielsen, 2018)

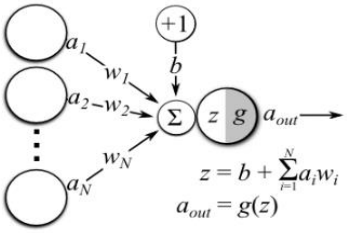
Η κατανόηση της λειτουργείας του νευρώνα είναι το βασικότερο στοιχείο για την κατανόηση λειτουργείας ενός νευρωνικού δικτύου.

Ένας **νευρώνας** δεν είναι κάτι παραπάνω από μία συνάρτηση με δεδομένα εισόδου τιμές από τους προηγουμένως συνδεδεμένους νευρώνες σε συνδυασμό με την βαρύτητα κάθε δεδομένου εισόδου για τον συγκεκριμένο νευρώνα και μία σταθερά τιμή. Οι τιμές που επηρεάζουν την βαρύτητα κάθε δεδομένου εισόδου ονομάζονται **βάρη (weights)** και η σταθερά τιμή ονομάζεται **προκατάληψη (bias)** (Εικόνα 3-4).

Το αποτέλεσμα αυτής της συνάρτησης είναι η τιμή του νευρώνα. Η συνάρτηση αυτή ονομάζεται **Συνάρτηση Ενεργοποίησης (Activation Function)** και υπάρχουν διάφορες

συναρτήσεις (Πίνακας 3-1) με διαφορετικό εύρος τιμών αποτελεσμάτων είτε με συνεχείς τιμές (0-1) είτε με δυαδικές τιμές (0/1), των οποίων η κατάλληλη επιλογή γίνεται ανάλογα την εφαρμογή. Η πιο γνωστή συνάρτηση ενεργοποίησης είναι η σιγμοειδής συνάρτηση και τα τελευταία χρόνια η συνάρτηση ReLU (Rectified Linear Unit).

$$z = x_1 * w_1 + x_2 * w_2 + \ldots\ldots + x_n * w_n + b * 1$$
$$\hat{y} = a_{out} = sigmoid(z)$$
$$sigmoid(z) = \frac{1}{1 + e^{-z}}$$



Εικόνα 3-4: Ενέργειες Τυπικού Νευρώνα ενός Νευρωνικού Δικτύου (Ahirwar, 2017)

Πίνακας 3-1: Μερικές Συναρτήσεις Ενεργοποίησης (Wikipedia, 2019)

Name	Plot	Equation	Derivative
Identity		$f(x) = x$	$f'(x) = 1$
Binary step		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x \neq 0 \\ ? & \text{for } x = 0 \end{cases}$
Logistic (a.k.a Soft step)		$f(x) = \frac{1}{1+e^{-x}}$	$f'(x) = f(x)(1 - f(x))$
TanH		$f(x) = \tanh(x) = \frac{2}{1+e^{-2x}} - 1$	$f'(x) = 1 - f(x)^2$
ArcTan		$f(x) = \tan^{-1}(x)$	$f'(x) = \frac{1}{x^2+1}$
Rectified Linear Unit (ReLU)		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Parameteric Rectified Linear Unit (PReLU) [2]		$f(x) = \begin{cases} \alpha x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Exponential Linear Unit (ELU) [3]		$f(x) = \begin{cases} \alpha(e^x - 1) & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} f(x) + \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
SoftPlus		$f(x) = \log_e(1 + e^x)$	$f'(x) = \frac{1}{1+e^{-x}}$

Επομένως ένα νευρωνικό δίκτυο εισάγοντας πληροφορία στο επίπεδο εισόδου, ο κάθε νευρώνας θα ενεργοποιηθεί και θα επηρεάσει τους επόμενους του, έως το τελικό επίπεδο εξόδου όπου η ενεργοποίηση των νευρώνων θα δώσει και το αποτέλεσμα. Στο τελευταίο επίπεδο επιλέγεται συνάρτηση κατάλληλη ώστε να εκφράζει την μορφή των αποτελεσμάτων.

Η γενική διαδικασία δόμησης νευρωνικών δικτύων είναι αρχικά η επιλογή δομής, στην συνέχεια η εκπαίδευση με γνωστά δεδομένα με τα αποτελέσματα τους, η αξιολόγηση των αποτελεσμάτων, και αν δεν είναι σωστά ή ικανοποιητικά, τότε αλλάζουν οι τιμές των προκαταλήψεων και βαρών και ξανά εξετάζεται το δίκτυο. Αυτή η διαδικασία επαναλαμβάνεται έως ότου τα αποτελέσματα του δικτύου να ταιριάζουν με τα πραγματικά αποτελέσματα.

Για την αξιολόγηση των αποτελεσμάτων γίνεται χρήση **Συνάρτησης Κόστους/Συνάρτηση Απωλειών**, η οποία υπολογίζει το σφάλμα του αποτελέσματος ενός δεδομένου κατά την εκπαίδευση και εν συνεχεία οι συνολικές απώλειες του δικτύου είναι ο μέσος όρος των σφαλμάτων όλων των δεδομένων. Υπάρχουν διάφορες συναρτήσεις υπολογισμού απωλειών, με τις πιο ευρέων χρησιμοποιούμενες την MSE-Mean Square Error, την Binary Crossentropy για δυαδικές λογαριθμικές απώλειες και την Categorical Crossentropy για πολλαπλών κατηγοριών λογαριθμικές απώλειες.

Μέθοδος ελαχίστων τετραγώνων (Σχέση 8):

$$C(w,b) = \frac{1}{2n}\sum || \mathrm{yi} - \mathrm{f(w,b)} ||^2 \tag{8}$$

Cross Entropy (Σχέση 9):

$$C(w,b) = \frac{1}{n}\sum[\mathrm{yi} * \ln[\mathrm{f(w,b)}] + (1 - \mathrm{yi}) * \ln[1 - \mathrm{f(w,b)}] \tag{9}$$

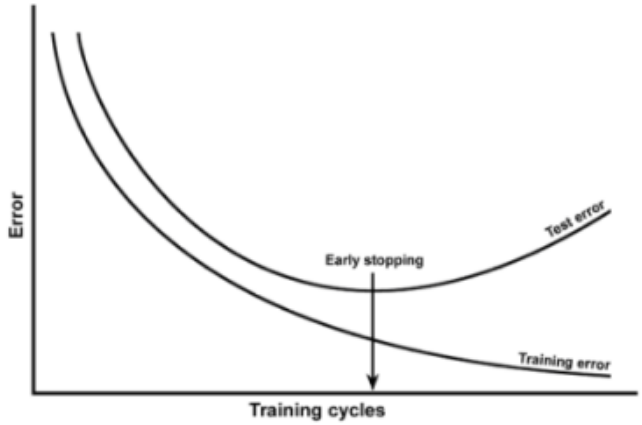
Η εκπαίδευση των νευρωνικών δικτύων είναι μια πολύ περίπλοκη και δύσκολη διαδικασία λόγω της αλληλεξάρτησης των νευρώνων που δημιουργεί τεράστιο αριθμό παραμέτρων που πρέπει να ρυθμιστούν για την επίτευξη κάλων αποτελεσμάτων, αν αυτά είναι και εφικτά. Για να ελαχιστοποιηθεί η συνάρτηση κόστους του μοντέλου πρέπει να ρυθμιστούν οι τιμές των βαρών και των προκαταλήψεων. Αυτό γίνεται μέσω μια διαδικασίας που ονομάζεται **Οπισθόδρομη Διόρθωση Σφάλματος (Backpropagation)** όπου χρησιμοποιούνται τεχνικές βελτιστοποίησης των βαρών και των προκαταλήψεων που ονομάζονται **Βελτιστοποιητές Μοντέλου (Model Optimizers)**. Σε αυτές τις τεχνικές βελτιστοποίησης οι πιο γνωστές είναι ο Αλγόριθμος Στοχαστικής Απότομης Καθόδου (SGD-Stochastic Gradient Descent), ο αλγόριθμός RMSprop και ο πλέον ευρέως χρησιμοποιούμενος αλγόριθμος Adam. Ο αλγόριθμος **Adam** είναι πλέον διαδεδομένος γιατί ρυθμίζει αυτόματα κατά την εκπαίδευση υπερπαραμέτρους όπως ο ρυθμός εκμάθησης και η ροπή (Diederik P. Kingma et al. 2015).

Για την εκπαίδευση και αξιολόγηση δικτύων πρέπει να διαχωριστούν το δείγμα δεδομένων σε τρία μέρη. Το πρώτο μέρος είναι το **Σύνολο Εκπαίδευσης** το οποίο θα αξιοποιηθεί για την εκπαίδευση του δικτύου και ενδείκνυται να αποτελεί το μεγαλύτερο ποσοστό του δείγματος (π.χ. 70%). Το δεύτερο μέρος είναι στο **Σύνολο Αξιολόγησης** που θα αξιοποιηθεί για να ελέγχεται το δίκτυο κατά την εκπαίδευση σε ξένα δεδομένα, κυρίως για παρακολούθηση των φαινομένων της **Υπερπροσαρμογής (Overfitting)**. Το σύνολο αυτό ενδείκνυται να αποτελεί ένα μικρό ποσοστό από το αρχικό δείγμα (π.χ. 15%) ανεξάρτητο από το σύνολο εκπαίδευσης. Τελευταίο μέρος είναι το **Σύνολο Δοκιμής** που θα αποτελείται από τα υπόλοιπα δεδομένα του

αρχικού δείγματος (π.χ. 15%) στο οποίο θα δοκιμαστεί το μοντέλο και η τελική ακρίβεια των προβλέψεων.

Η διάρκεια της εκπαίδευσης ενός νευρωνικού δικτύου ορίζεται σε **Εποχές (Epochs)** από τον ερευνητή. Ως μία εποχή ορίζεται όταν το δίκτυο κατά την εκπαίδευση έχει δει όλο το πλήθος των δεδομένων εκπαίδευσης. Ανάλογα την εφαρμογή προτείνεται από ερευνητές ένα ενδεικτικό αρχικό πλήθος εποχών για καλά αποτελέσματα. Επίσης σε περίπτωση μεγάλου πλήθους δεδομένων, λόγω υπολογιστικών περιορισμών, για κάθε εποχή τα δεδομένα εκπαίδευσης χωρίζονται σε **τμήματα σταθερού μεγέθους** ορισμένα από τον ερευνητή **(Batch Size)**, και η εκπαίδευση για μία εποχή γίνεται σε βήματα το πλήθος των οποίων ορίζεται από το μέγεθος των τμημάτων.

Σκοπός μιας εφαρμογής νευρωνικού δικτύου είναι να εκπαιδευτεί σε ικανοποιητικά μεγάλο δείγμα με απώτερο σκοπό την δυνατότητα γενίκευσης του μοντέλου και σε ξένα δεδομένα. Το μεγαλύτερο πρόβλημα που μπορεί να εμφανιστεί είναι το φαινόμενο της υπερπροσαρμογής. Δηλαδή το μοντέλο κατά την εκπαίδευση έπειτα από ένα σημείο αρχίζει να μαθαίνει 'απέξω' τα χαρακτηριστικά και τις ιδιομορφίες του δείγματος εκπαίδευσης με αποτέλεσμα να επιστρέφει εξαιρετικά αποτελέσματα ακρίβειας και απωλειών, αλλά αν έρθει αντιμέτωπο με ξένα δεδομένα οι προβλέψεις είναι οριακά τυχαίες. Για να αποφευχθεί αυτό το φαινόμενο το μοντέλο κατά την αξιολόγηση εξετάζεται συνεχώς και στο σύνολο αξιολόγησης, το οποίο περιέχει δεδομένα που είναι ξένα, συνεπώς τα αποτελέσματα σε αυτό το σύνολο είναι πολύ πιο αντιπροσωπεύτηκα στην πραγματική απόδοση του μοντέλου. Στο Διάγραμμα 3-2 φαίνεται το φαινόμενο της υπερπροσαρμογής.

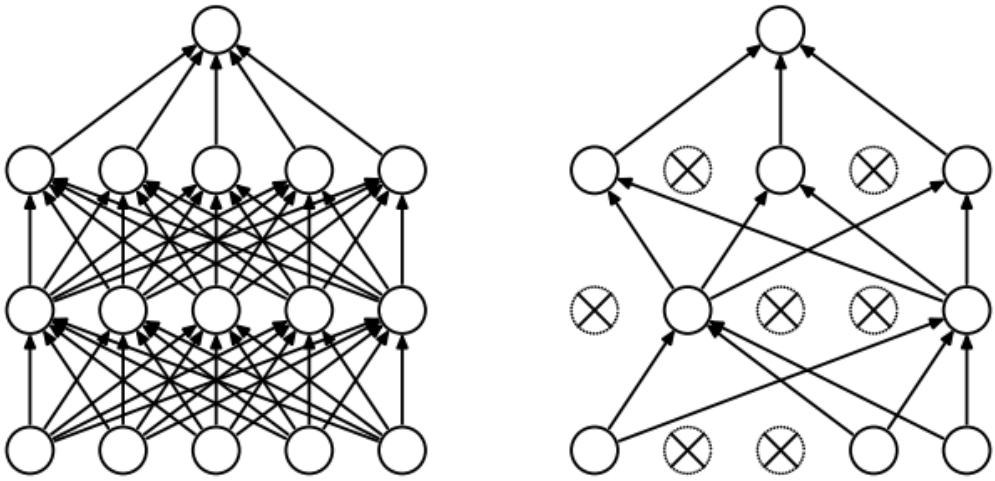


Διάγραμμα 3-2: Φαινόμενο Υπερπροσαρμογής

Για να περιοριστεί το πρόβλημα που εμφανίζεται με το φαινόμενο της **υπερπροσαρμογής**, έχουν ερευνηθεί και εξεταστεί αρκετές εφαρμογές και προτεινόμενες λύσεις. Αρχικά πρέπει να αναφερθεί πως είναι καλή μέθοδος και προτιμάτε τα δεδομένα της εκπαίδευσης να είναι μεγάλα σε πλήθος και διακριτά μεταξύ τους. Η ποικιλία χαρακτηριστικών έναντι των παρόμοιων δεδομένων ενισχύει στην αποφυγή υπερπροσαρμογής και ενισχύει την δυνατότητα γενίκευσης του μοντέλου. Επίσης είναι πολύ σημαντικό να καθοριστεί σωστό πλήθος εποχών ώστε να απομονωθεί το μοντέλο που δίνει την ελάχιστη τιμή απωλειών στο δείγμα αξιολόγησης. Αυτό δεν είναι πάντα εύκολο να καθοριστεί, για αυτό μπορούν να εφαρμοστούν αυτοματισμοί που παρακολουθούν τα αποτελέσματα και αποθηκεύουν το μοντέλο με τα βάρη και τις προκαταλήψεις που δίνουν τα βέλτιστα (checkpoint saver), και εν συνεχεία σταματάνε την εκπαίδευση όταν αρχίσει το φαινόμενο της υπερπροσαρμογής (early stopping).

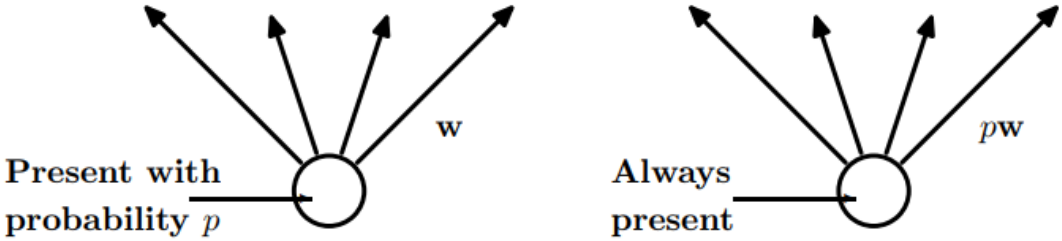
Βέβαια πολλές φορές το παραπάνω δεν είναι εφικτό, η δεν έχει μεγάλη επιθυμητή επίδραση, επομένως εφαρμόζονται επιπλέον οι επόμενες τεχνικές οι ανεξάρτητα οι σε συνδυασμό μεταξύ τους. Η **Μέθοδος Απόρριψής Νευρώνων (Dropout)** η και η **Κανονικοποίηση Τμημάτων Δεδομένων (Batch Normalization),** όταν αυτά υπάρχουν, είναι οι βασικότερες και πιο αποδοτικές τεχνικές για μείωση του φαινομένου της υπερπροσαρμογής.

Η **μέθοδος απόρριψής νευρώνων** (Srivastava et al., 2014) είναι μία τεχνική όπου ορίζοντας την πιθανότητα απόρριψης για κάθε νευρώνα κατά την εκπαίδευση σε κάθε εποχή, άρα δεν θα συμβάλουν στην συγκεκριμένη περίπτωση εκπαίδευσης (Εικόνα 3-5). Αυτό έχει ως αποτέλεσμα να περιορίζεται η ελευθερία του δικτύου να απομνημονεύει τα χαρακτηριστικά των δεδομένων, επομένως μειώνεται σημαντικά το φαινόμενο της υπερπροσαρμογής. Να σημειωθεί πως το ποσοστό απόρριψής είναι μια παράμετρος που θα πρέπει να ρυθμιστεί αναλόγως το πείραμα για βέλτιστο αποτέλεσμα. Επίσης σημαντικό είναι να αναφερθεί πως η απόρριψη πραγματοποιείται μόνο κατά την εκπαίδευση και όχι κατά την δοκιμή του δικτύου (Εικόνα 3-6).



(a) Standard Neural Net (b) After applying dropout.

Εικόνα 3-5: Εφαρμογή Μεθόδου Απόρριψης Νευρώνων (Srivastava et al., 2014)



(a) At training time (b) At test time

Εικόνα 3-6: Λειτουργεία Νευρώνων κατά την Εκπαίδευση και Δοκιμή με Εφαρμογή Απόρριψης Νευρώνων (Srivastava et al., 2014)

Η **Κανονικοποίηση Τμημάτων** (Ioffe et al., 2015) αποσκοπεί στην σταθεροποίηση του νευρωνικού δικτύου εφαρμόζοντας κανονικοποίηση στα δεδομένα εξόδου του προηγούμενου επιπέδου ενεργοποίησης νευρώνων αφαιρώντας τον μέσο όρο του τμήματος δεδομένων και διαιρώντας με την τυπική απόκλιση. Η μέθοδος επίσης **Συστηματοποιεί (Regularization)** το μοντέλο και έχει επίσης παρατηρηθεί πως όταν εφαρμόζεται η μέθοδος σε συνδυασμό με την μέθοδο απόρριψης νευρώνων τότε μπορεί να μειωθεί η πιθανότητα απόρριψης για καλύτερα αποτελέσματα.

Input: Values of x over a mini-batch: $\mathcal{B} = \{x_{1\ldots m}\}$;
Parameters to be learned: γ, β

Output: $\{y_i = \mathrm{BN}_{\gamma,\beta}(x_i)\}$

$$\mu_\mathcal{B} \leftarrow \frac{1}{m}\sum_{i=1}^{m} x_i \qquad \text{// mini-batch mean}$$

$$\sigma_\mathcal{B}^2 \leftarrow \frac{1}{m}\sum_{i=1}^{m} (x_i - \mu_\mathcal{B})^2 \qquad \text{// mini-batch variance}$$

$$\widehat{x}_i \leftarrow \frac{x_i - \mu_\mathcal{B}}{\sqrt{\sigma_\mathcal{B}^2 + \epsilon}} \qquad \text{// normalize}$$

$$y_i \leftarrow \gamma \widehat{x}_i + \beta \equiv \mathrm{BN}_{\gamma,\beta}(x_i) \qquad \text{// scale and shift}$$

Algorithm 1: Batch Normalizing Transform, applied to activation x over a mini-batch.

Εικόνα 3-7: Μετατροπή με την Μέθοδο Κανονικοποίησης Τμημάτων Δεδομένων σε Ένα Τμήμα (Ioffe et al., 2015)

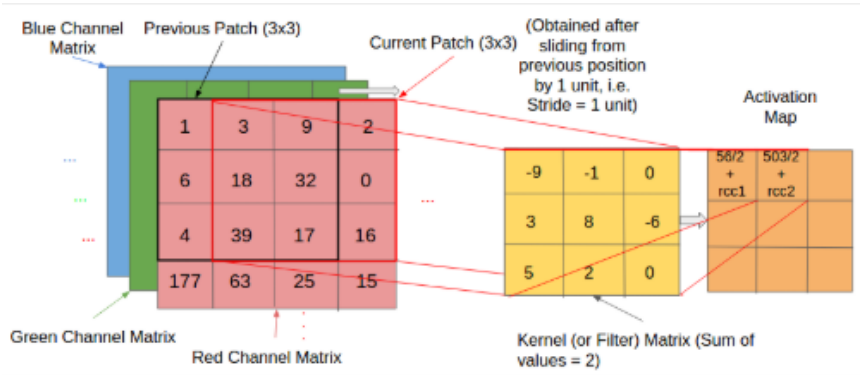
Συνελικτικά Νευρωνικά Δίκτυα

Τα συνελικτικά νευρωνικά δίκτυα είναι μία μορφή νευρωνικών δικτύων, επομένως ότι αναφέρθηκε προηγουμένως για τα νευρωνικά δίκτυα ισχύουν κατ' επέκταση και για τα συνελικτικά νευρωνικά δίκτυα. Η κύριες διαφορές είναι στην δομή τους, καθώς η αρχική εισαγωγή και επεξεργασία πληροφορίας γίνεται με διαφορετικό τρόπο από τα κλασσικά νευρωνικά δίκτυα. Αυτές οι αλλαγές στην δομή είναι ο λόγος που τα συνελικτικά νευρωνικά δίκτυα ενδείκνυται για ανάλυση εικόνας και βίντεο, και πάνω σε αυτά έχουν βασιστεί εφαρμογές στον τομέα της υπολογιστικής όρασης.

Οι έγχρωμες εικόνες έχουν τρείς διαστάσεις, πλάτος, ύψος, και βάθος. Το βάθος ορίζεται από την χρωματική κλίμακα της εικόνας, για την χρωματική κλίμακα κόκκινο-πράσινο-μπλε (RGB) που είναι η πιο συνήθης, η εικόνα έχει βάθος τριών επιπέδων, ένα για κάθε χρώμα της κλίμακας επομένως κάθε πίξελ της εικόνας έχει τρεις τιμές μία για κάθε βασικό χρώμα τις κλίμακας που σε συνδυασμό καθορίζουν το τελικό χρώμα του πίξελ. Αντίστοιχα οι

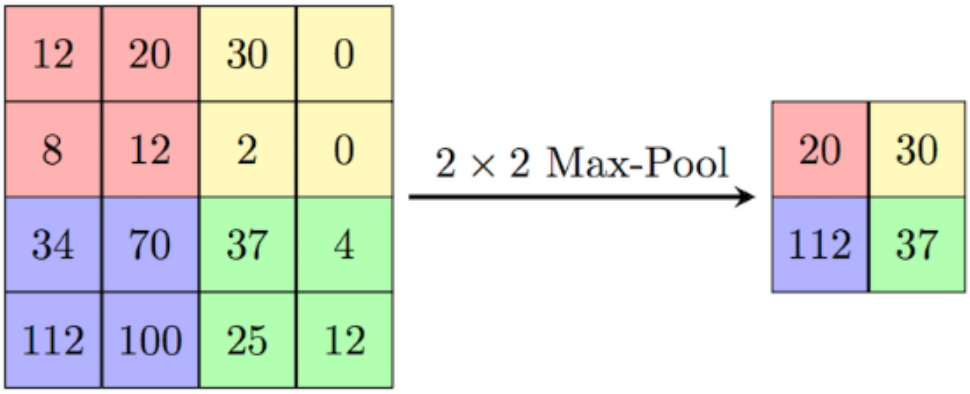
ασπρόμαυρες φωτογραφίες έχουν βάθος ενός επιπέδου, όπου κάθε πίξελ έχει μία τιμή που καθορίζει το χρώμα του μεταξύ του λευκού και του μαύρου. Να σημειωθεί πως οι ασπρόμαυρες φωτογραφίες έχουν σχεδόν το ένα τρίτο της πληροφορίας από τις έγχρωμες, κάτι που μπορεί να είναι θετικό και αρνητικό αναλόγως την εφαρμογή.

Ένα συνελικτικό νευρωνικό δίκτυο αποτελείται από ένα ή περισσότερα **Επίπεδα Συνέλιξης (Convolutional Layers)** και συνήθως να επακολουθεί ένα επίπεδο **Υπό-Δειγματοληψίας (Pooling/Max-Pooling)** ακολουθούμενο από ένα ή περισσότερα πλήρως συνδεδεμένα επίπεδα όπως συμβαίνει και σε ένα κλασικό νευρωνικό δίκτυο. Η είσοδος δεδομένων σε ένα επίπεδο συνέλιξης είναι μία εικόνα διαστάσεων $\pi*\upsilon*\beta$ οπού π, υ και β είναι τα χαρακτηριστικά που αναφέρθηκαν προηγουμένως, πλάτος, ύψος, βάθος αντίστοιχα. Το επίπεδο έχει **φίλτρα (kernels)** διαστάσεων $\nu*\nu*\kappa$ οπού ν είναι μία τιμή μικρότερη από τις διαστάσεις την εικόνας και κ είναι το βάθος του φίλτρου που μπορεί να είναι ίσο η μικρότερο από τα κανάλια χρώματος της εικόνας, οι διαστάσεις αυτές ορίζονται από τον μελετητή αναλόγως την εφαρμογή. Επίσης κάθε επίπεδο συνέλιξης μπορεί να έχει φίλτρα διαφορετικών μεγεθών. Το μέγεθος των φίλτρων προκαλεί τοπικά συνδεδεμένη δομή όπου καθένα συνελίσσεται σε κάθε εικόνα με αποτέλεσμα ένα πλήθος από **χάρτες χαρακτηριστικών (feature maps)** μεγέθους $\pi-\kappa+1*\theta-\kappa+1$ (Εικόνα 3-8).



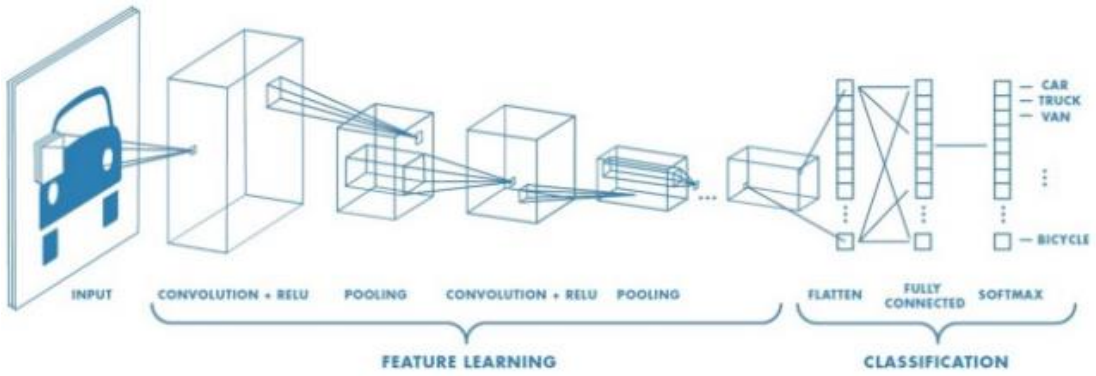
Εικόνα 3-8: Παράδειγμα Συνέλιξης Φίλτρων σε Εικόνα (XRDS, 2016)

Κάθε χαρακτηριστικό από τα προηγουμένως παραγμένα υπόκειται σε υπό-δειγματοληψία στο επόμενο επίπεδο συνήθως με την μέθοδο max-pooling σε περιοχές διαστάσεων συνήθως 2*2 έως 5*5 για εικόνες υψηλής διάστασης (Εικόνα 3-9).



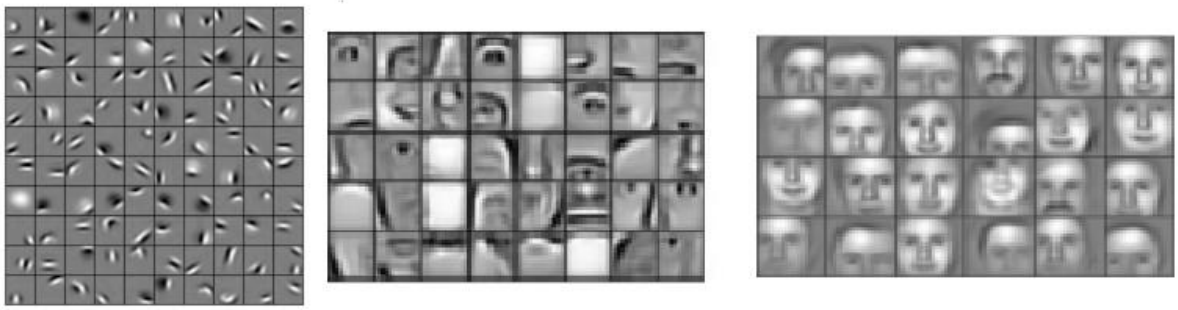
Εικόνα 3-9: Παράδειγμα Υπό-Δειγματοληψίας (Computer Science Wiki, 2018)

Πριν ή μετά το επίπεδο της υπό-δειγματοληψίας ορίζονται προκαταλήψεις και μια συνάρτηση ενεργοποίησης, συνήθως η σιγμοειδής ή ReLU. Στην συνέχεια επαναλαμβάνεται η διαδικασία για κάθε επίπεδο συνέλιξης έως το τελευταίο όπου σε εκείνη την φάση του δικτύου τα δεδομένα απλώνονται σε ένα επίπεδο που ονομάζεται **Flatten Layer**. Το επίπεδο αυτό λειτουργεί σαν επίπεδο εισόδου για ένα κλασσικό νευρωνικό δίκτυο που ακολουθεί, εισάγοντας τουλάχιστον ένα πλήρως συνδεδεμένο επίπεδο (fully connected). Στην εικόνα που ακολουθεί φαίνεται η ροή λειτουργείας ενός συνελικτικού δικτύου (Εικόνα 3-10).



Εικόνα 3-10: Ροή Συνελικτικού Νευρωνικού Δικτύου με Πολλά Επίπεδα Συνέλιξης (Medium, 2018)

Μέσω των παραπάνω διαδικασιών ένα συνελικτικό νευρωνικό δίκτυο δημιουργεί χάρτες χαρακτηριστικών για αναγνώριση μοτίβων. Σε κάθε επίπεδο συνέλιξης αυξάνεται η περιπλοκότητα των μοτίβων, όπως στο γνωστό παράδειγμα των Honglak et al. (2011) όπως φαίνεται στην Εικόνα 3-11 τα μοτίβα στο πρώτο επίπεδο είναι γραμμές και απλά σχήματα, στο δεύτερο επίπεδο είναι χαρακτηριστικά προσώπου, και στο τρίτο πλέον τα μοτίβα είναι ολόκληρα πρόσωπα.

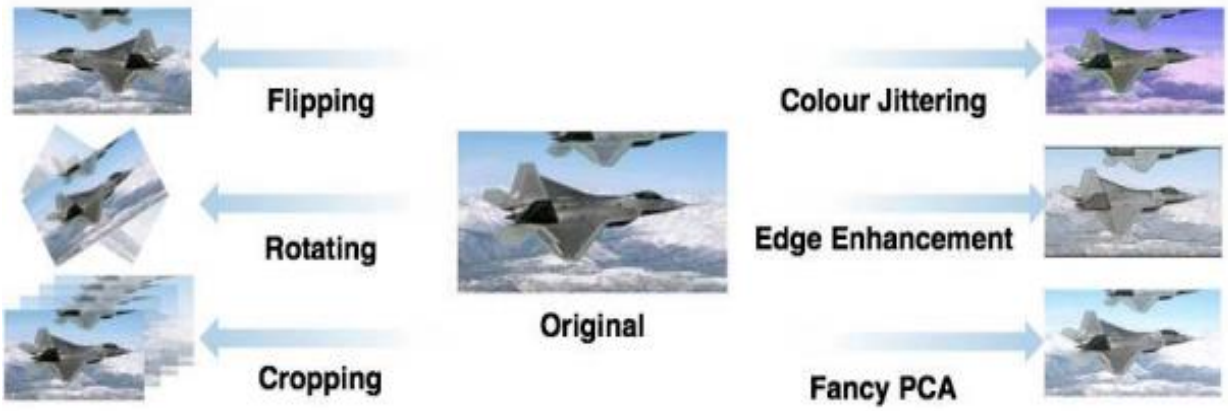


Εικόνα 3-11: Παράδειγμα Μοτίβων Ενεργοποίησης Φίλτρων σε Συνελικτικό Νευρωνικό Δίκτυο (Honglak et al.,, 2011)

Το πρόβλημα του φαινομένου της υπερπροσαρμογής υπάρχει και σε αυτήν την μορφή δικτύων, και η αντιμετώπιση του πραγματοποιείται με τις ίδιες τεχνικές που έχουν αναφερθεί και προηγουμένως.

Επίσης να αναφερθεί πως η διαδικασία εκπαίδευσης αξιολόγησής και δοκιμής του συνελικτικού νευρωνικού δικτύου είναι ίδια με προηγουμένως, καθώς οι προηγούμενες τεχνικές και μεθοδολογίες εφαρμόζονται και σε αυτήν την μορφή δικτύου.

Μία τεχνική στα συνελικτικά νευρωνικά δίκτυα κατά την εκπαίδευση, για καλύτερη γενίκευση του μοντέλου και διαφοροποίηση των δεδομένων σε κάθε εποχή είναι η τυχαία παραμόρφωση την εικόνων πριν την εκπαίδευση (Taylor et al., 2017). Τέτοιες παραμορφώσεις μπορεί να είναι η αντιστροφή την εικόνας, τυχαία περιστροφή, τυχαία περικοπή, τυχαία αλλαγή χρωμάτων και φωτεινότητας και άλλα (Εικόνα 3-12). Τα προηγούμενα εφαρμόζονται αυτόνομα η σε συνδυασμό μεταξύ τους με τυχαία πιθανότητα εφαρμογής και ένταση παραμόρφωσης ορισμένη από την ερευνητή. Η μέθοδος ονομάζεται **Παραμόρφωση Δεδομένων (Data Augmentation)**



Εικόνα 3-12: Παραδείγματα Παραμόρφωσης Εικόνας (Taylor et al., 2017)

Προ-Εκπαιδευμένα Συνελικτικά Νευρωνικά Δίκτυα

Τα προ-εκπαιδευμένα συνελικτικά νευρωνικά δίκτυα, είναι δίκτυα βασισμένα στα βασικά στοιχεία των συνελικτικών δικτύων, με περίπλοκες δομές και μεγάλο πλήθος επιπέδων. Αυτά έχουν δημιουργηθεί από ερευνητές, εταιρίες ή πανεπιστήμια, για συγκεκριμένους σκοπούς και έχουν εκπαιδευτεί σε τεράστιες βάσεις δεδομένων εικόνων.

Τα δίκτυα αυτά μπορούν να χρησιμοποιηθούν στο ίδιο πλαίσιο με τα συνελικτικά νευρωνικά δίκτυα που αναλύθηκαν προηγουμένως με νέα δεδομένα. Δηλαδή να αξιοποιηθεί η δομή τους και οι παράμετροι τους για εκπαίδευση και δοκιμή σε δεδομένα ανεξάρτητα με αυτά που έχουν αρχικά εκπαιδευτεί. Αυτή η διαδικασία ονομάζεται **Διαβιβαστική Μάθηση (Transfer Learning)** (Tan, et al., 2018).

Επίσης μία βασική εφαρμογή αυτών των δικτύων είναι ο εντοπισμός αντικειμένων εντός μιας εικόνας πέρα από την κατηγοριοποίηση της εικόνας σε μία κατηγορία. Αυτό πραγματοποιείται από δύο μορφές δικτύων, τα **Συνελικτικά Νευρωνικά Δίκτυα Πρότασης Περιοχής (Region Proposal Convolutional Neural Networks ή R-CNN)** (Girshick et al., 2014) και **Δίκτυα**

Εντοπισμού Πολλαπλών Πλαισίων με Μία Λήψη (Single Shot Multibox Detection Network – SSD Network) (Liu, et al., 2016).

Τα περισσότερα από αυτά τα δίκτυα είναι ανοιχτού κώδικα και μπορούν να αξιοποιηθούν από οποιονδήποτε, γεγονός πολύ σημαντικό για κάποιον ερευνητή καθώς η δόμηση δικτύων και εκπαίδευση τους σε τέτοια κλίμακα είναι πρακτικά αδύνατη με τον τυπικό υπολογιστικό εξοπλισμό και τα διαθέσιμα δεδομένα.

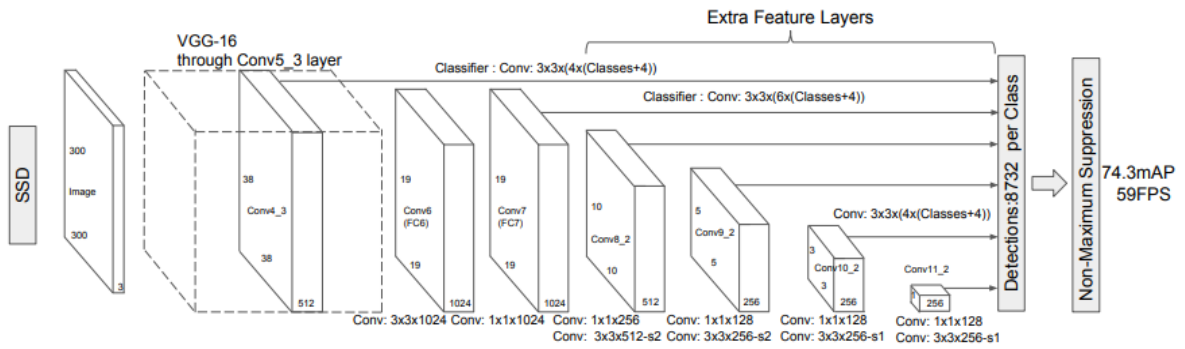
Τα δίκτυα αυτά εξετάζονται σε γνωστές βάσεις δεδομένων, όπως η βάση δεδομένων COCO (Common Objects in Context) της Microsoft (Lin, et al., 2015), με πάνω από 300.000 εικόνες από συνήθη αντικείμενα και αξιολογούνται ως προς την ταχύτητα τους και την ακρίβεια όπως παρουσιάζεται στον επόμενο πίνακα (Πίνακας 3-2).

Αξίζει να αναφερθεί και το μοντέλο YOLOv3 (You Only Look Once Version 3) και οι διάφορες εκδοχές του, που έχει διαφορετική αρχιτεκτονική και έχει εξίσου καλά αποτελέσματα και ταχύτητα επεξεργασίας (Redmon & Farhadi, 2018).

Η λειτουργεία των Δικτύων Εντοπισμού Πολλαπλών Πλαισίων με Μία Λήψη βασίζεται στον εντοπισμό περιοχών πιθανών αντικειμένων μέσω της δημιουργίας επιπέδων για την δημιουργία χαρτών χαρακτηριστικών και στην συνέχεια με χρήση επιπέδων συνέλιξης να γίνεται η πρόβλεψη της περιοχής ως μία κατηγορία. Η βασική δομή ενός τέτοιου δικτύου παρουσιάζεται στην Εικόνα 3-13.

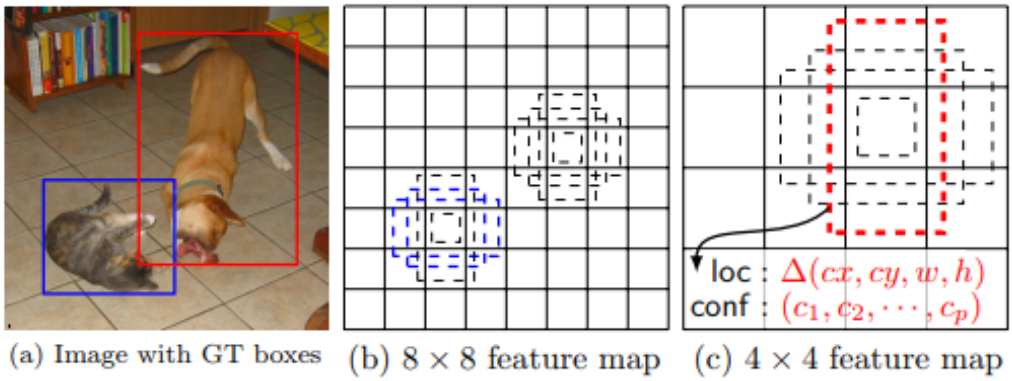
Πίνακας 3-2: Χαρακτηριστικά Προ-Εκπαιδευμένων Μοντέλων στη Βάση Δεδομένων COCO (Google-Tensorflow, 2019)

Model name	Speed (ms)	COCO mAP[^1]	Outputs
ssd_mobilenet_v1_coco	30	21	Boxes
ssd_mobilenet_v1_0.75_depth_coco ☆	26	18	Boxes
ssd_mobilenet_v1_quantized_coco ☆	29	18	Boxes
ssd_mobilenet_v1_0.75_depth_quantized_coco ☆	29	16	Boxes
ssd_mobilenet_v1_ppn_coco ☆	26	20	Boxes
ssd_mobilenet_v1_fpn_coco ☆	56	32	Boxes
ssd_resnet_50_fpn_coco ☆	76	35	Boxes
ssd_mobilenet_v2_coco	31	22	Boxes
ssd_mobilenet_v2_quantized_coco	29	22	Boxes
ssdlite_mobilenet_v2_coco	27	22	Boxes
ssd_inception_v2_coco	42	24	Boxes
faster_rcnn_inception_v2_coco	58	28	Boxes
faster_rcnn_resnet50_coco	89	30	Boxes
faster_rcnn_resnet50_lowproposals_coco	64		Boxes
rfcn_resnet101_coco	92	30	Boxes
faster_rcnn_resnet101_coco	106	32	Boxes
faster_rcnn_resnet101_lowproposals_coco	82		Boxes
faster_rcnn_inception_resnet_v2_atrous_coco	620	37	Boxes
faster_rcnn_inception_resnet_v2_atrous_lowproposals_coco	241		Boxes
faster_rcnn_nas	1833	43	Boxes
faster_rcnn_nas_lowproposals_coco	540		Boxes
mask_rcnn_inception_resnet_v2_atrous_coco	771	36	Masks
mask_rcnn_inception_v2_coco	79	25	Masks
mask_rcnn_resnet101_atrous_coco	470	33	Masks
mask_rcnn_resnet50_atrous_coco	343	29	Masks



Εικόνα 3-13: Δίκτυο Εντοπισμού Πολλαπλών Πλαισίων με Μία Λήψη (Single Shot Multibox Detection Network – SSD Network) (Liu, et al., 2016)

Λόγω της φύσης του μοντέλου, η διαδικασία της εκπαίδευσης διαφέρει καθώς θα πρέπει να οριστούν στις εικόνες των δεδομένων για κάθε αντικείμενο, η κατηγορία του αλλά και η θέση του στην εικόνα ως ορθογώνιο παραλληλεπίπεδο πλαίσιο ορίζοντας τις συντεταγμένες των γωνιών του. Τα δεδομένα εκπαίδευσης και τα αποτελέσματα θα είναι της μορφής που φαίνονται στην Εικόνα 3-14.



Εικόνα 3-14: Παράδειγμα Εικόνας σε Δίκτυο Εντοπισμού Πολλαπλών Πλαισίων με Μία Λήψη (Single Shot Multibox Detection Network – SSD Network) (Liu, et al., 2016)

Επιπλέον διαφοροποιείται ο τρόπος αξιολόγησης του μοντέλου, καθώς δεν εξετάζεται από την συνάρτηση απωλειών μόνο η σωστή πρόβλεψη, αλλά ο συνδυασμός σωστής πρόβλεψης σε σωστή θέση εντός της εικόνας. Οι τελικές τιμές των απωλειών ονομάζονται Συνολικές Απώλειες (Total Loss) η οποίες προκύπτουν από συνδυασμό των συναρτήσεων Απωλειών Κατηγοριοποίησης/Αυτοπεποίθησης (Classification/Confidence Loss) και των Απωλειών Καθορισμού Θέσης (Localization Loss) των οποίων οι εξισώσεις παρατίθενται στην συνέχεια.

Συνολικές Απώλειες (Total Loss) (Σχέση 10):

$$L(x, c, l, g) = \frac{1}{N}(L_{conf}(x, c) + \alpha L_{loc}(x, l, g)) \tag{10}$$

Απώλειες Καθορισμού Θέσης (Localization Loss) (Σχέση 11):

$$L_{loc}(x,l,g) = \sum_{i \in Pos}^{N} \sum_{m \in \{cx,cy,w,h\}} x_{ij}^{k} \text{smooth}_{\text{L1}}(l_i^m - \hat{g}_j^m)$$
$$\hat{g}_j^{cx} = (g_j^{cx} - d_i^{cx})/d_i^w \qquad \hat{g}_j^{cy} = (g_j^{cy} - d_i^{cy})/d_i^h$$
$$\hat{g}_j^{w} = \log\left(\frac{g_j^w}{d_i^w}\right) \qquad \hat{g}_j^{h} = \log\left(\frac{g_j^h}{d_i^h}\right) \tag{11}$$

Απώλειες Κατηγοριοποίησης/Αυτοπεποίθησης (Classification/Confidence Loss) (Σχέση 12):

$$L_{conf}(x,c) = -\sum_{i \in Pos}^{N} x_{ij}^{p} log(\hat{c}_i^p) - \sum_{i \in Neg} log(\hat{c}_i^0) \quad \text{where} \quad \hat{c}_i^p = \frac{\exp(c_i^p)}{\sum_p \exp(c_i^p)} \tag{12}$$

Μία βελτιωμένη εκδοχή αυτού του μοντέλου αποτελεί το δίκτυο SSD MobileNetV2 (Sandler et al., 2018) η οποία εχει βελτιστοποιήσει την δομή και τον τρόπο λειτουργίας του δικτύου ώστε να έιναι πιο αποδοτικό, για εφαρμοφή σε συσκευές με χαμηλή υπολογσιτική ισχύ και χωρτητικότητα μνήμης, όπως είναι οι συσκευές Andoid.

Σύνοψη Βασικών Στοιχείων Συνελικτικών Δικτύων

Τα νευρωνικά δίκτυα και κατ' επέκταση τα συνελικτικά νευρωνικά δίκτυα αποτελούν μία σύγχρονη μέθοδο επίλυσης προβλημάτων με πολλές δυνατότητες και μεθοδολογίες διαθέσιμες στον ερευνητή.

Τα συνελικτικά νευρωνικά δίκτυα δύναται να επεξεργαστούν πληροφορίες σε μορφή εικόνας, κάτι που μέχρι πρόσφατα θεωρούταν πολύ δύσκολο καθώς δεν είχαν αναπτυχθεί πολλές τεχνικές και κυρίως λόγω της έλλειψης επαρκούς υπολογιστικής ισχύς. Αυτό γιατί η εικόνα περιέχει τεράστιο πλήθος πληροφορίας, ειδικά με τα σύγχρονα μέσα καταγραφής οπτικού υλικού που έχουν δυνατότητα καταγραφής σε μεγάλες αναλύσεις.

Για την βέλτιστη λειτουργεία των συνελικτικών νευρωνικών δικτύων πρέπει να δοθεί μεγάλη βάση στα δεδομένα εκπαίδευσης καθώς είναι ο βασικότερος παράγοντας στην δημιουργία ενός αξιόπιστου δικτύου. Τα δεδομένα πρέπει να έχουν μεγάλο πλήθος και είναι όσο το δυνατόν ποικίλα και διακριτά μεταξύ τους.

Επομένως είναι κρίσιμο ο μελετητής να γνωρίζει τις απαιτήσεις του προβλήματος που θα εξετάσει, και να έχει αποκτήσει τον κατάλληλο εξοπλισμό για την γρήγορη και αποδοτική επεξεργασία και ανάλυση των δεδομένων του.

3.4.3 - Μέθοδος Αξιολόγησης

Η οριστική μέθοδος αξιολόγησης στην πρώτη μεθοδολογία πέρα από τις τιμές απωλειών και ακρίβειας κατά την εκπαίδευση-αξιολόγηση των συνελικτικών νευρωνικών δικτύων είναι τα

στατιστικά μεγέθη που προκύπτουν από μητρώο Προβλέψεων/Πραγματικότητας (ή διαφορετικά πίνακας σύγχυσης) (Πίνακας 3-3).

Πίνακας 3-3: Πίνακας Σύγχυσης (confusion matrix)

Μητρώο Πραγματικότητας/Προβλέψεων		Πραγματικότητα	
		0	1
Πρόβλεψη	0	TP	FP
	1	FN	TN

Όπου:

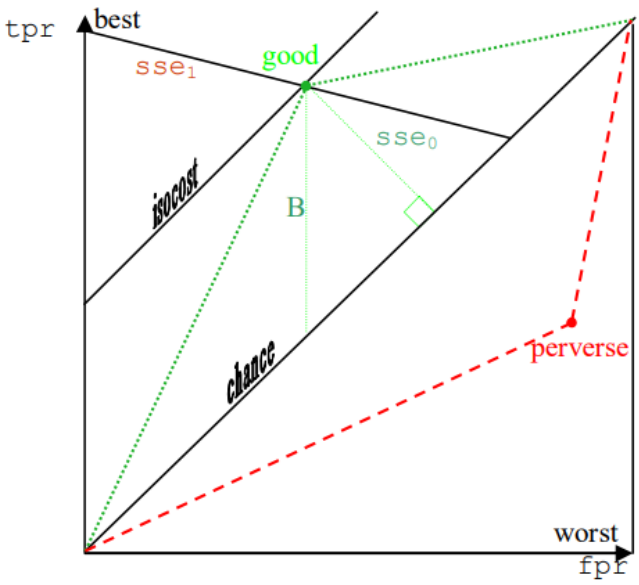
- TP (True Positive): Αριθμός των στοιχείων που προβλέπονται θετικά και είναι πραγματικά θετικά.
- TN (True Negative): Αριθμός των στοιχείων που προβλέπονται αρνητικά και είναι πραγματικά αρνητικά.
- FP (False Positive): Αριθμός των στοιχείων που προβλέπονται θετικά ενώ είναι αρνητικά.
- FN (False Negative): Αριθμός των στοιχείων που προβλέπονται αρνητικά ενώ είναι θετικά.

Από τον παραπάνω πίνακα προκύπτουν οι βασικές μετρικές κατηγοριοποίησης που φαίνονται παρακάτω:

- Ορθότητα (Accuracy) $ACC = \frac{TP+TN}{TP+TN+FP+FN}$
- Ακρίβεια (Precision) $PPV = \frac{TP}{TP+FP}$
- Ευαισθησία (Sensitivity) $TPR = \frac{TP}{TP+FN}$
- Εξειδίκευση (Specificity) $SPC = \frac{TN}{TN+FP}$ (= TNR)
- Ποσοστό εσφαλμένων προβλέψεων (False Positive Rate) $FPR = \frac{FP}{FP+TN} = 1 - SPC$
- Αρμονικό μέσο $F1 = \frac{2*TP}{2*TP+FP+FN}$

Η απόδοση ενός μοντέλου μπορεί να αναπαραχθεί από ένα σημείο στην καμπύλη ROC (Receiver Operating Curve) η οποία στον κάθετο άξονα έχει τις τιμές της ευαισθησίας του μοντέλου (TPR) και στον οριζόντιο άξονα τις τιμές 1 – SPC (FPR) (Διάγραμμα 3-3).

Η επιφάνεια κάτω από την καμπύλη ROC είναι μετρική της ακρίβειας του μοντέλου και μπορεί να υπολογιστεί ως εξής: Area Under Curve $AUC = \frac{TPR-FPR+1}{2}$



Διάγραμμα 3-3: Διάγραμμα Καμπύλης ROC (Powers, 2007)

Το ιδανικό μοντέλο θα είχε στον πίνακα σύγχυσης όλα τα στοιχεία που δεν ανήκουν στην κύρια διαγώνιο μηδενικά. Δηλαδή θα προέβλεπε το κάθε στοιχείο στην πραγματική του κλάση. Ακόμα ένα ικανοποιητικό μοντέλο κατηγοριοποίηση θα πρέπει να έχει AUC μεγαλύτερο από 0.7.

3.5 - Επιλογή Λογισμικού - Υπολογιστικού Εξοπλισμού και Περιορισμοί τους

Στο πλαίσιο της διπλωματικής εργασίας επιλέχτηκε να χρησιμοποιηθεί η γλώσσα προγραμματισμού **Python**, μία γλώσσα ανοιχτού κώδικα με τεράστιο εύρος εφαρμογών, και μεγάλη κοινότητα υποστήριξης. Επίσης η Python θεωρείται από πολλούς η ιδανική γλώσσα σε θέματα επιστήμης των δεδομένων (data science), καθώς την υποστηρίζουν οι μεγαλύτερες βιβλιοθήκες (libraries) και τεχνολογικά πλαίσια που έχουν να κάνουν με μηχανική μάθηση, βαθιά μάθηση, υπολογιστική όραση κλπ.

Το μεγαλύτερο μέρος της επεξεργασίας δεδομένων στην εργασία πραγματοποιήθηκε στην Python με την βοήθεια εξειδικευμένων βιβλιοθηκών για επεξεργασία βίντεο και εικόνας, όπως το **OpenCV3.**

Η δημιουργία, ανάλυση αξιολόγηση και ότι έχει να κάνει με τα συνελικτικά νευρωνικά δίκτυα πραγματοποιήθηκε στην Python, με την χρήση της βιβλιοθήκης **TensorFlow** της Google, που θεωρείται πλέον κορυφαία για εφαρμογές μηχανικής μάθησης και βαθιάς μάθησης. Μερικές από εταιρίες που χρησιμοποιούν το TensorFlow είναι η Airbnb, AMD, nVIDIA, UBER, SAP, DEEPMIND, Dropbox, ebay, intel, CocaCola και πολλές άλλες. Σε συνδυασμό με την βιβλιοθήκη του TensorFlow σε μερικές εφαρμογές αξιοποιήθηκε και η βιβλιοθήκη **Keras** που είναι μια βιβλιοθήκη γραμμένη στην Python, για εύκολη και ταχύτερη ανάπτυξη εφαρμογών βαθιάς μάθησης, με υπόβαθρο την πλατφόρμα του TensorFlow.

Η εκπαίδευση νευρωνικών δικτύων είναι μια πολύ χρονοβόρα διαδικασία καθώς το πλήθος των δεδομένων και των παραμέτρων προς εκπαίδευση είναι μεγάλο. Πιο συγκεκριμένα η εκπαίδευση συνελικτικών νευρωνικών δικτύων είναι ακόμα πιο απαιτητική καθώς ένα μεγάλο πλήθος εικόνων μπορεί να απαιτεί πολύ μεγαλύτερη χωρητικότητα συγκριτικά με δεδομένα μορφής κειμένου και αριθμών. Αυτό μπορεί να παρουσιάσει πρόβλημα καθώς κατά την εκπαίδευση τα δεδομένα αποθηκεύονται προσωρινά στην μνήμη RAM του υπολογιστή για να είναι άμεσα διαθέσιμα σε υψηλές ταχύτητες μετάδοσης από τον επεξεργαστή. Αυτό μας περιορίζει άμεσα στο πλήθος των δεδομένων που μπορούν να αξιοποιηθούν. Επίσης πρέπει να αναφερθεί πως η κεντρική μονάδα επεξεργασίας ενός υπολογιστή δεν είναι το βέλτιστο μέσο για να πραγματοποιεί τις ενεργείς που απαιτούνται κατά την εκπαίδευση ενός δικτύου γιατί οι εντολές που εκτελεί είναι λιγότερες και πολύ πιο περίπλοκες σε μικρό αριθμό πυρήνων ενώ οι εντολές που χρειάζονται στην εκπαίδευση ενός νευρωνικού δικτύου είναι πολύ απλές αλλά τεράστιες σε πλήθος, ουσιαστικά οι ενέργειες αυτές είναι απλές πράξεις μεταξύ πολύ μεγάλων μητρώων. Επομένως ενδείκνυται η χρήση κάρτας γραφικών όπου αποτελείτε από πολύ μεγαλύτερο πλήθος πυρήνων συγκριτικά λιγότερο ισχυρών αλλά βελτιστοποιημένους για να πραγματοποιούν τις επιθυμητές ενέργειες ενώ χρησιμοποιείται και η μνήμη VRAM της κάρτας γραφικών σε αυτήν την περίπτωση για προσωρινή αποθήκευση των δεδομένων, η οποία είναι σημαντικά ταχύτερη από την μνήμη RAM του υπολογιστή. Για αυτόν τον λόγο οι εργασίες σε μία κάρτα γραφικών μπορούν να γίνουν έως και 45 φορές ταχύτερα, σε ορισμένες περιπτώσεις, σε σχέση με τον επεξεργαστή. Σημαντική παρατήρηση επίσης είναι πως στις εφαρμογές που αξιοποιείται εικόνα ως δεδομένο ενδείκνυται η κάρτα γραφικών να έχει τουλάχιστον 8 GB (Giga Byte) μνήμη VRAM και το σύστημα να έχει τουλάχιστον την διπλάσιά μνήμη RAM. Η επιλογή του TensorFlow έγινε επίσης γιατί υποστηρίζει την επιτάχυνση των εργασιών με κάρτες γραφικών από την εταιρία nVidia, που αξιοποιούν την γλωσσα προγραμματισμού CUDA, και τα πακέτα cuDNN της nVidia, που είναι δομημένα για τέτοιου τύπου εργασίες και εφαρμογές.

Το σύστημα που αξιοποιήθηκε έχει τα εξής χαρακτηριστικά:

Πλατφόρμα:	Z370
Επεξεργαστής:	Intel® Core™ i7-8700K Processor 6/12 (Cores/Threads) with 12M Cache, up to 4.70 GHz
Μνήμη RAM:	16 GB 3200 MHz
Κάρτα γραφικών:	nVidia GeForce GTX 1070 Ti , Cores: 2432, VRAM: 8 GB DDR5

Σε αυτό το σύστημα κύριος περιοριστικός παράγοντας αλλά και κύριος ενισχυτικός παράγοντας είναι η κάρτα γραφικών. Η χωρητικότητα της VRAM επέτρεψε την εκπαίδευση μοντέλων με μεγάλες βάσεις δεδομένων με τον διαχωρισμό τους σε όσον το δυνατόν λιγότερα υποσύνολα, κάτι που ενίσχυσε χρονικά σε μεγάλο βαθμό τις εφαρμογές τις διπλωματικής εργασίας αλλά και επέτρεψε την εφαρμογή μίας μεθοδολογίας που δεν θα ήταν εφικτή χωρίς την χρήση του συγκεκριμένου εξοπλισμού. Από την άλλη ο περιοριστικός παράγοντας ήταν την χωρητικότητα της μνήμης VRAM καθώς αξιοποιούταν στο μέγιστο σε κάθε περίπτωση εφαρμογής.

4 - ΣΥΛΛΟΓΗ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ

4.1 - Συλλογή Δεδομένων

Για την πραγματοποίηση της διπλωματικής εργασίας αξιοποιήθηκαν βίντεο καταγεγραμμένα από μη επανδρωμένο αεροσκάφος, με πανοραμική λήψη επί της οδού Κατεχάκη. Από τα βίντεο προέκυψε μεγάλο πλήθος εικόνων, οι οποίες επεξεργάστηκαν, ανάλογα με τις απαιτήσεις της κάθε μεθοδολογίας, για την δόμηση, εκπαίδευση και αξιολόγηση συνελικτικών νευρωνικών δικτύων.

Τα βίντεο είναι 25 σε αριθμό με διάρκεια που κυμαίνεται από δύο έως δώδεκα λεπτά.

Τα δεδομένα σε μορφή βίντεο προσφέρουν μεγάλη ευελιξία στον ερευνητή για επιλογή και απομόνωση μεγάλου πλήθους δεδομένων ή εικόνων που απαιτούνται για το πείραμα. Επίσης η καταγραφή τους είναι πιο εύκολη και ελεγχόμενη συγκριτικά με την καταγραφή απομονωμένων εικόνων καθώς είναι εφικτή η καταγραφή οποιαδήποτε χρονική στιγμή και για οποιοδήποτε χρονικά διάστημα, παράγοντας μεγάλο πλήθος εικόνων προσφέροντας στον ερευνητή και την πλήρη εικόνα και αντίληψη των δεδομένων που θα επιλεχτούν για το πείραμα, με δυνατότητα ανά πάσα στιγμή να γίνουν αλλαγές στα δεδομένα χωρίς την απαίτηση συλλογής νέων.

Τα βίντεο της παρούσας διπλωματικής είναι καταγεγραμμένα από κάμερα τοποθετημένη σε μη επανδρωμένο αεροσκάφος, επομένως υπόκεινται περιορισμών που πηγάζουν από την χρήση του μη επανδρωμένου αεροσκάφους. Οι βασικότεροι περιορισμοί είναι η περιορισμένη διάρκεια των βίντεο λόγω περιορισμού από την δυνατή χρονική διάρκεια πτήσης του αεροσκάφους λόγω μπαταριών, ή έλλειψη σταθερότητας στην λήψη συγκριτικά με μια σταθερή στάσιμη κάμερα καθώς και ο νομικός περιορισμός της πτήσης μη επανδρωμένων αεροσκαφών στον εναέριο χώρο.

Η κάμερα που χρησιμοποιήθηκε είναι Go Pro Hero 3 που προσφέρει επαρκή ανάλυση και ευρύ οπτικό πεδίο, καθώς η λήψη είναι σε σημαντική απόσταση από την οδό και καθ΄ ύψος και κατά πλάτος (Εικόνα 4-1).

Τα 25 βίντεο έχουν ανάλυση 1920*1080 πίξελ, πανοραμική λήψη, και ρυθμό καταγραφής 24, 25, 30 η 60 εικόνες το δευτερόλεπτο.

Video Resolution	NTSC/ PAL	FOV	Screen Resolution
4K	30/25 24/24	Ultra Wide	3840x2160, 16:9
4K SuperView	24/24	Ultra Wide	3840x2160, 16:9
2.7K	48/50 30/25 24/24	Ultra Wide, Medium	2704x1520, 16:9
2.7K SuperView	30/25	Ultra Wide	2704x1520, 16:9
2.7K 4:3	30/25	Ultra Wide	2704x2028, 4:3
1440p	80/80 60/50 48/48 30/25 24/24	Ultra Wide	1920x1440, 4:3
1080p	120/120 90/90	Ultra Wide, Narrow	1920x1080, 16:9
1080p	60/50 48/48 30/25 24/24	Ultra Wide, Medium, Narrow	1920x1080, 16:9
1080p SuperView	80/80 60/50 48/48 30/25 24/24	Ultra Wide	1920x1080, 16:9
960p	120/120 60/50	Ultra Wide	1280x960, 4:3
720p	120/120 60/50 30/25	Ultra Wide, Medium, Narrow	1280x720, 16:9
720p SuperView	120/120 60/50	Ultra Wide	1280x720, 16:9
WVGA	240/240	Ultra Wide	848x480, 16:9

Εικόνα 4-1: Χαρακτηριστικά της κάμερας Go Pro Hero 3

Η λήψη των βίντεο πραγματοποιήθηκε επί της οδού Κατεχάκη, στο ύψος της Πολυτεχνειούπολης, μεσημεριανές ώρες για επαρκή φωτισμό της οδού σε ώρες αιχμής ώστε να υπάρχει ποικιλία δεδομένων. Λόγω θέσης του μη επανδρωμένου αεροσκάφους υπάρχει περιορισμένη ορατότητα σε τμήματα της οδού, επομένως θα εξεταστεί η οδός στην άνοδο , η οποία έχει τρείς λωρίδες κυκλοφορίας καθώς και έναν φωτεινό σηματοδότη (Εικόνα 4-2).

Εικόνα 4-2: Ενδεικτικά Στιγμιότυπα από τα Βίντεο

Για την αναλυτική παρουσίαση της διαδικασίας συλλογής δεδομένων του πειράματος ο αναγνώστης παρατίθεται στον Μπαρμπουνάκη (Barmpounakis, Investigating the decision-making process of drivers during overtaking by Powered Two Wheelers, 2017).

4.2 - Προετοιμασία Δεδομένων

Ακολουθείται διαφορετική διαδικασία προετοιμασίας δεδομένων ανάλογα με το στάδιο προτυποποίησης.

4.2.1 - Αναγνώριση Κυκλοφοριακών Συνθηκών σε Πραγματικό Χρόνο

Για την εφαρμογή της πρώτης προσέγγισης θα εξεταστούν τα βίντεο, θα απομονωθούν τα χρονικά διαστήματα που χρειαστούν από τα βίντεο και στην συνέχεια θα απομονωθούν τα τμήματα της οδού που θα αξιοποιηθούν. Σκοπός είναι η δημιουργία βάσεων δεδομένων **δυο κατηγοριών**, **Υψηλή Κυκλοφορία** (κατηγορία 1) και **Χαμηλή / Καθόλου Κυκλοφορία** (κατηγορία 0), για την εκπαίδευση και αξιολόγηση των συνελικτικών νευρωνικών δικτύων.

Ο διαχωρισμός στις δύο κατηγορίες, Υψηλή Κυκλοφορία (κατηγορία 1) και Χαμηλή/Καθόλου Κυκλοφορία (κατηγορία 0), γίνεται ανά εικόνα. Έπειτα από οπτική αξιολόγηση του κάθε βίντεο και της κυκλοφορίας, τέθηκαν κάποια κριτήρια με βάση τα οποία θα γίνει ο διαχωρισμός κάθε εικόνας σε κατηγορίες.

Τα κριτήρια για να θεωρηθεί μια εικόνα κατηγορία 1 είναι τα εξής (αρκεί να ισχύει ένα από τα επόμενα):

- Να υπάρχουν, έστω και μερικώς, στην εικόνα σε οποιαδήποτε θέση 4 (τέσσερα) ή παραπάνω οχήματα.
- Να υπάρχουν στην ίδια λωρίδα κυκλοφορίας 3 (τρία) ή παραπάνω οχήματα.
- Να υπάρχουν σε γειτονικές λωρίδες 3 (τρία) ή παραπάνω οχήματα.
- Να υπάρχει ένα όχημα, ή παραπάνω, σε κάθε λωρίδα.

Στις υπόλοιπες περιπτώσεις θα θεωρείται κατηγορία 0. Να σημειωθεί πως οι μοτοσυκλέτες δεν λαμβάνονται υπόψιν στον διαχωρισμό

Η επεξεργασία των βίντεο πραγματοποιήθηκε κυρίως με χρήση της γλώσσας προγραμματισμού Python και μερικώς με το πρόγραμμα επεξεργασίας βίντεο Adobe Premier Pro.

Αρχικά αποκόπηκαν τα τμήματα απογείωσης και προσγείωσης του UAV, διατηρώντας τα τμήματα των βίντεο όπου το αεροσκάφος έχει σταθεροποιηθεί πάνω από την οδό. Επίσης αφαιρέθηκαν τα χρονικά διαστήματα όπου ο φωτεινός σηματοδότης έχει κόκκινη σήμανση, και υπάρχουν σταματημένα οχήματα στην οδό.

Στην συνέχεια εξετάστηκε κάθε βίντεο για την επιλογή κατάλληλου τμήματος της οδού καθώς και για τον διαχωρισμό της κυκλοφορίας κάθε χρονική στιγμή, στο επιλεγμένο τμήμα της οδού, στις κατηγορίες: Χαμηλή/Καθόλου Κυκλοφορία και Αυξημένη Κυκλοφορία.

Η επιλογή τμήματος οδού έγινε με κριτήρια την βέλτιστη ορατότητα των τριών λωρίδων κυκλοφορίας και των οχημάτων. Λόγω μεταβαλλόμενης θέσης του αεροσκάφους σε κάθε λήψη, το επιλεγμένο τμήμα ενδέχεται να διαφέρει ανά βίντεο. Επίσης για ομοιομορφία των

δεδομένων τα τμήματα θα διατηρούν σταθερό λόγο διαστάσεων (aspect ratio) 1 : 2,5 καθώς βρέθηκε ως ιδανικός για βέλτιστη ορατότητα και ελάχιστη περιττή πληροφορία (εμπόδια, δέντρα κλπ.)

Έπειτα για κάθε βίντεο εξετάστηκε η κυκλοφορία στα επιλεγμένα τμήματα και έγινε διαχωρισμός στις δύο κατηγορίες με βάση τα στιγμιότυπα (Frames), και στην συνέχεια καταγράφηκαν σε αρχείο λογιστικών φύλλων (Excel) (Πίνακας 4-1). Ο αρχικός διαχωρισμός πραγματοποιήθηκε με τα κριτήρια που αναφέρθηκαν προηγουμένως.

Για λόγους διευκόλυνσης η κατηγορία « Χαμηλή/Καθόλου Κυκλοφορία» θα αναφέρεται ως κατηγορία 0 και η «Αυξημένη Κυκλοφορία» ως κατηγορία 1.

Πίνακας 4-1: Ενδεικτικός Διαχωρισμός του Βίντεο με Όνομα 1797.mp4 (Με Βάση τα Στιγμιότυπα - Frames)

Video Name	1797		
File No	11	crop No	1
Total Frames	13,679	fps	25

FRAMES			
CATEGORIES			
0		1	
from	to	from	to
-	-	1	342
343	667	668	800
801	872	873	968
969	1,007	1,008	1,207
1,208	1,240	1,241	1,644
1,645	1,832	1,833	1,959
1,960	2,382	2,383	2,477
2,478	3,367	3,368	3,478
3,479	3,690	3,691	4,084
4,085	4,169	4,170	4,220
4,221	4,368	4,369	4,453
4,454	4,503	4,504	5,110
5,111	6,116	6,117	6,236
6,237	6,300	6,301	6,808
6,809	7,040	7,041	7,078
7,079	7,189	7,190	8,105
8,106	8,851	8,852	8,948
8,949	8,980	8,981	9,068
9,069	9,303	9,304	9,431
9,432	9,440	9,441	10,846
10,847	11,727	11,728	11,757
11,758	11,808	11,809	12,348
12,349	12,418	12,419	12,747
12,748	13,246	13,247	13,341
13,342	13,679		

TIME (Sec:Frames)			
CATEGORIES			
0		1	
from	to	from	to
0:0	0:0	0:1	13:17
13:18	26:17	26:18	32:0
32:1	34:22	34:23	38:18
38:19	40:7	40:8	48:7
48:8	49:15	49:16	65:19
65:20	73:7	73:8	78:9
78:10	95:7	95:8	99:2
99:3	134:17	134:18	139:3
139:4	147:15	147:16	163:9
163:10	166:19	166:20	168:20
168:21	174:18	174:19	178:3
178:4	180:3	180:4	204:10
204:11	244:16	244:17	249:11
249:12	252:0	252:1	272:8
272:9	281:15	281:16	283:3
283:4	287:14	287:15	324:5
324:6	354:1	354:2	357:23
357:24	359:5	359:6	362:18
362:19	372:3	372:4	377:6
377:7	377:15	377:16	433:21
433:22	469:2	469:3	470:7
470:8	472:8	472:9	493:23
493:24	496:18	496:19	509:22
509:23	529:21	529:22	533:16
533:17	547:4	0:0	0:0

TIME (Sec)			
CATEGORIES			
0		1	
from	to	from	to
0	0	0.04	13.68
13.72	26.68	26.72	32
32.04	34.88	34.92	38.72
38.76	40.28	40.32	48.28
48.32	49.6	49.64	65.76
65.8	73.28	73.32	78.36
78.4	95.28	95.32	99.08
99.12	134.68	134.72	139.12
139.16	147.6	147.64	163.36
163.4	166.76	166.8	168.8
168.84	174.72	174.76	178.12
178.16	180.12	180.16	204.4
204.44	244.64	244.68	249.44
249.48	252	252.04	272.32
272.36	281.6	281.64	283.12
283.16	287.56	287.6	324.2
324.24	354.04	354.08	357.92
357.96	359.2	359.24	362.72
362.76	372.12	372.16	377.24
377.28	377.6	377.64	433.84
433.88	469.08	469.12	470.28
470.32	472.32	472.36	493.92
493.96	496.72	496.76	509.88
509.92	529.84	529.88	533.64
533.68	547.16	0	0

Στην συνέχεια για την εξαγωγή των στιγμιότυπων από τα βίντεο, το διαχωρισμό τους ανά κατηγορία και ανά τμήμα οδού και τέλος την αποθήκευση τους ως εικόνες, αναπτύχθηκε πρόγραμμα (script) στη γλώσσα προγραμματισμού Python.

Το πρόγραμμα λειτουργεί με δεδομένα εισόδου τα εξής:

- Το αρχείο του βίντεο.
- Τοποθεσία αποθήκευσης εικόνων.
- Τα στοιχεία διαχωρισμού που έχουν καταγραφεί προηγουμένως.
- Το χρονικό διάστημα ανά το οποίο θα εξάγεται και θα αποθηκεύονται τα στιγμιότυπά. (Για παράδειγμα ανά μισό δευτερόλεπτο)
- Το αρχικό σημείο περικοπής του βίντεο καθώς και τις επιθυμητές διαστάσεις των τελικών εικόνων. (Με σταθερό λόγο διαστάσεων 1:2.5)

- Λοιπά στοιχεία επεξεργασίας του βίντεο, όπως επιλογή αποκοπής τμημάτων σε οποιοδήποτε σημείο του βίντεο, παραμόρφωση ή οπτική επεξεργασία αυτού.

Τα παραπάνω δεδομένα εισόδου μπορούν ευκολά να διαμορφωθούν και να προσαρμοστούν ανάλογα το κάθε βίντεο, και τις απαιτήσεις του πειράματος, προσφέροντας μεγάλη ευελιξία. Τα δεδομένα εξόδου του προγράμματος είναι του οι εικόνες των επιλεγμένων τμημάτων οδού διαχωρισμένες στις κατηγορίες. Το σύνολο των εικόνων από τα βίντεο της κάθε κατηγορίας θα είναι τα δεδομένα που θα αξιοποιηθούν στο πείραμα.

Εφόσον έχει πραγματοποιηθεί ο διαχωρισμός με το παραπάνω πρόγραμμα, θα γίνει και οπτικός έλεγχος των δεδομένων με βάση τα κριτήρια, καθώς λόγω των υψηλών ρυθμών καταγραφής των βίντεο σε εικόνες ανά δευτερόλεπτο είναι αναμενόμενο να υπάρχουν και εικόνες σε λάθος κατηγορία. Οι εικόνες σε κάθε κατηγορία έχουν την παρακάτω μορφή (Εικόνα 4-3 και Εικόνα 4-4).

Εικόνα 4-3: Ενδεικτικές Εικόνες Κατηγορίας 0 (Χαμηλή/Καθόλου Κυκλοφορία)

Εικόνα 4-4: Ενδεικτικές Εικόνες Κατηγορίας 1 (Αυξημένη Κυκλοφορία)

Εκπαίδευση και Αξιολόγηση Προτύπων

Η εκπαίδευση των συνελικτικών νευρωνικών δικτύων είναι μια χρονοβόρα διαδικασία, λόγω του μεγάλου όγκου πληροφορίας που περιέχουν οι εικόνες, επομένως θα δημιουργηθούν τυχαία οι εξής βάσεις δεδομένων:

- Βάση με μικρό αριθμό εικόνων για την αρχική **συνοπτική αξιολόγηση** δομών και χαρακτηριστικών των νευρωνικών δικτύων, που αποτελείται από:
    - **2000** εικόνες για **εκπαίδευση** (1000 ανά κατηγορία)
    - **200** εικόνες για **αξιολόγηση** (100 ανά κατηγορία)
- Βάση με μεγάλο αριθμό εικόνων για **εκτενή και πιο αξιόπιστη εκπαίδευση-αξιολόγηση** των νευρωνικών δικτύων, που αποτελείται από:
    - **50000** εικόνες για **εκπαίδευση** (25000 ανά κατηγορία)
    - **12000** εικόνες για **αξιολόγηση** (6000 ανά κατηγορία)
- Βάση με **2000** εικόνες (1000 ανά κατηγορία) για **τελική αξιολόγηση/δοκιμή** κάθε μοντέλου και υπολογισμός στατιστικών μεγεθών (κοινό δείγμα για κάθε μοντέλο).

Σημειώνεται ότι το δείγμα αξιολόγησης πρέπει να μην περιέχει εικόνες από το δείγμα που χρησιμοποιήθηκε για την εκπαίδευση. Όσο διαφορετικό/ξένο είναι το δείγμα αξιολόγησης τόσο καλύτερα μπορεί να αξιολογηθεί η δυνατότητα γενίκευσης του μοντέλου(με κόστος την ακρίβεια).

Στην συνέχεια απαιτείται η διαμόρφωση των εικόνων ώστε να έχουν κατάλληλη μορφή (κοινές διαστάσεις κλπ.) για το νευρωνικό δίκτυο καθώς και η επιλογή των διαστάσεων και της χρωματικής κλίμακας (RGB, Grayscale), που επηρεάζουν άμεσα την χρονική διάρκεια εκπαίδευσης και την απαιτούμενη υπολογιστική ισχύι, καθώς και την ποιότητα των αποτελεσμάτων.

Γενικά, θα πρέπει να σημειωθεί ότι μεγαλύτερη ανάλυση εικόνας και πληροφορίες για την χρωματική κλίμακα RGB, αντιστοιχούν σε υψηλότερες υπολογιστικές απαιτήσεις, αλλά όμως δεν ισοδυναμούν απαραίτητα σε χρήσιμη πληροφορία για το πρότυπο κατηγοριοποίησης, δηλαδή δεν οδηγεί πάντα σε πιο αξιόπιστο μοντέλο.

Για το πείραμα της διπλωματικής θα εξεταστούν οι περιπτώσεις με εικόνες στην **κλίμακα RGB και Grayscale**, με ανάλυση **300:120 πίξελ** και **150:60 πίξελ**. Τα δεδομένα προετοιμάζονται με πρόγραμμα (script).

4.2.2 - Μακροσκοπική Αναγνώριση Κυκλοφοριακών Συνθηκών

Για την εφαρμογή της δεύτερης προσέγγισης πρέπει να εξεταστούν τα βίντεο, να γίνει επιλογή και απομόνωση του επιθυμητού τμήματος της οδού, και στην συνέχεια να επιλεγούν χρονικά διαστήματα με υψηλή κυκλοφορία ώστε να υπάρχουν πολλά οχήματα σε κάθε στιγμιότυπο. Από το σύνολο των εικόνων θα γίνει επιλογή ενός πλήθους με διαφορετικά χαρακτηριστικά και θα γίνει επισύναψη σε κάθε εικόνα της τοποθεσίας του κάθε οχήματος και της κατηγορίας του.

Για την εξαγωγή εικόνων από τα δεδομένα βίντεο, δημιουργείται πρόγραμμα στο οποίο ορίζεται το τμήμα της οδού που είναι επιθυμητό, και τα χρονικά διαστήματα από τα οποία θα εξαχθούν οι εικόνες.

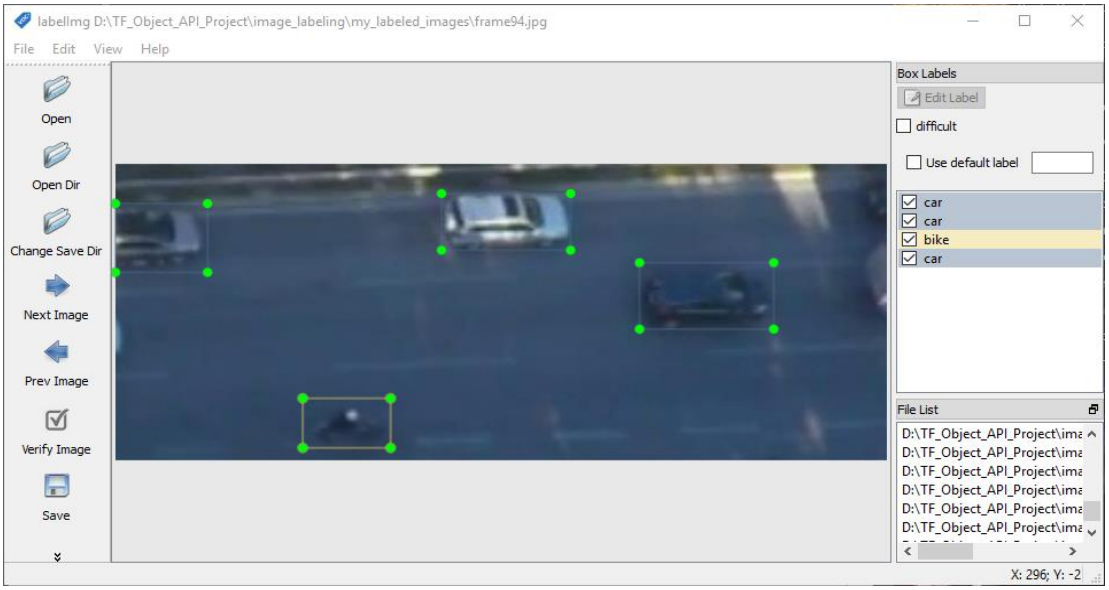
Στην συνέχεια ακολούθησε η επιλογή **1200 εικόνων**, από διαφορετικές λήψεις, διαφορετικές συνθήκες φωτισμού και σκίασης και διαφορετικής ανάλυσης. Σε αυτές τις εικόνες εμφανίζονται συνολικά **4000 οχήματα**, όπου με χρήση του προγράμματος ανοιχτού κώδικα (*LabelImg* (https://github.com/tzutalin/labelImg)) θα οριστούν οι θέσεις τους και οι κατηγορίες τους με μορφή κατάλληλη για την εκπαίδευση του προ-εκπαιδευμένου συνελικτικού νευρωνικού δικτύου (Εικόνα 4-5).

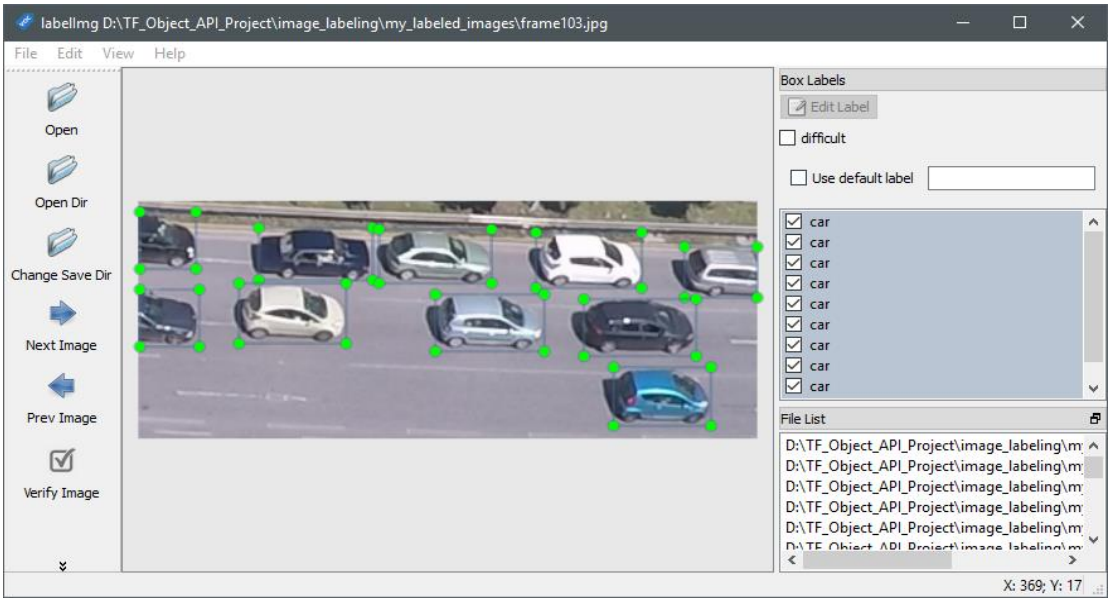
Τα οχήματα είναι τεσσάρων κατηγοριών:

- Αυτοκίνητα (Car)
- Φορτηγάκια-Βαν (Van)
- Φορτηγά-Λεωφορεία (Truck-Bus)
- Μοτοσυκλέτες

Στην συνέχεια πραγματοποιείται τυχαίος διαχωρισμός των εικόνων σε δύο δείγματα, ένα για **εκπαίδευση των δικτύων με 3600 οχήματα**, και ένα δείγμα για **αξιολόγηση του δικτύου κατά την εκπαίδευση με 400 οχήματα**. Τέλος, έπειτα από την παραπάνω επεξεργασία, για να αξιοποιηθούν οι εικόνες μαζί με τις αντίστοιχες πληροφορίες που περιέχουν, όσον αφορά στο πλήθος, τη θέση και την κατηγορία οχήματος θα μετατραπούν σε κατάλληλη μορφή που αναγνωρίζει το TensorFlow, για να ξεκινήσει η διαδικασία της εκπαίδευσης στο επιλεγμένο προ-εκπαιδευμένο συνελικτικό νευρωνικό δίκτυο με τα δεδομένα του πειράματος.

Για την οριστική αξιολόγηση και υπολογισμό μετρικών την ικανότητας πρόβλεψης του μοντέλου πραγματοποιείται η επιλογή ενός νέου συνόλου **80** ξένων εικόνων που περιέχουν συνολικά **400** οχήματα, εξαιρουμένων των μοτοσυκλετών.





Εικόνα 4-5: Παραδείγματα Χρήσης Προγράμματος LabelImage

5 - ΑΝΑΛΥΣΗ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ

5.1 - Εισαγωγή

Σε αυτό το κεφάλαιο, θα παρουσιαστούν δυο προτεινόμενες μεθοδολογικές προσεγγίσεις. Η πρώτη μεθοδολογία αποσκοπεί στην γρήγορη, υπολογιστικά μη απαιτητική, αξιολόγηση της κυκλοφορίας επι μιας οδού εντός δύο κατηγοριών, από βίντεο για χρήση σε πραγματικά χρόνο. Η δεύτερη μεθοδολογία στοχεύει στο να αναγνωρίζει και να εντοπίζει οχήματα σε βίντεο από μη επανδρωμένο αεροσκάφος χωρίς κάποια προσαρμογή ή ρύθμιση αυτού και στην συνέχεια να υπολογίζει προσεγγιστικά κυκλοφοριακά μεγέθη σε ένα τμήμα οδού. Τέλος θα συγκριθούν τα αποτελέσματα των δυο μεθόδων. Για τις δύο παραπάνω μεθοδολογίες θα εξεταστούν τα βήματα τους, οι απαιτήσεις του και τελικά θα αξιολογηθούν.

5.2 - Αναγνώριση Κυκλοφορικών Συνθηκών σε Πραγματικό Χρόνο

Περιγράφεται η διαδικασία ανάπτυξης, παραμετροποίησης, εκπαίδευσης και αξιολόγησης των προτύπων μέσα από μια συνοπτική και μια εκτενή θεώρηση.

5.2.1 - Συνοπτική Θεώρηση

Για τον καθορισμό αποδοτικών μοντέλων θα δοκιμαστούν βασικές δομές και συνδυασμός αυτών με την πρώτη βάση δεδομένων, για σύντομη εκπαίδευση και αξιολόγηση των δομών.

Αρχικά εξετάζονται πρότυπα με τις ακόλουθες δομές:

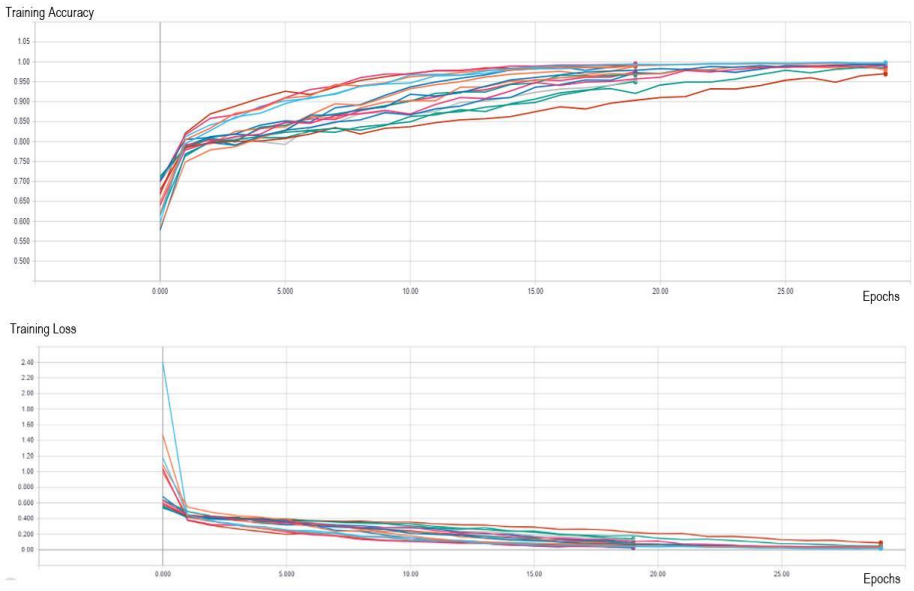
- 1-3 επίπεδα συνέλιξης (Convolutional layers) με διαφορετικά χαρακτηριστικά-μεγέθη.
- Συμπαγής Στρώση (Dense layer) με διαφορετικά μεγέθη.
- Χρήση η μη χρήση Απόρριψης Νευρώνων (Dropout).
- Συνδυασμό των παραπάνω δομών.

Η εκπαίδευση θα γίνει με μικρό αριθμό εποχών, αρκετό ώστε να αναγνωριστούν σύντομα φαινόμενα και προβλήματα του μοντέλου όπως Υπερπροσαρμογή (Overfitting) στα αποτελέσματα της αξιολόγησης. Στα νευρωνικά δίκτυα το πιο αξιόπιστο μέτρο για την γρήγορη αξιολόγηση ενός μοντέλου είναι κατά κανόνα η ελαχιστοποίηση της τιμής της

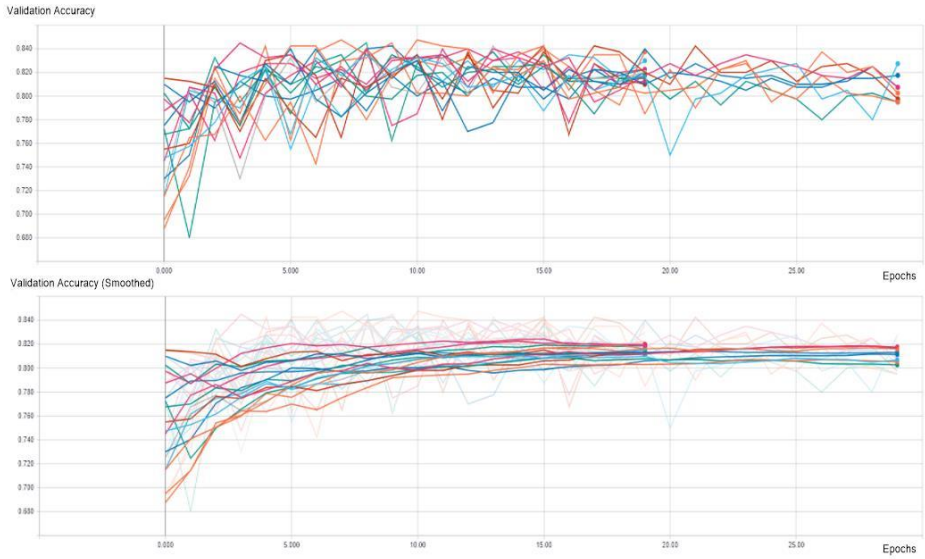
Συνάρτησης Κόστους (Loss Function) και στην συνέχεια οι τιμές ακρίβειας στο δείγμα αξιολόγησης (Validation Loss και Validation Accuracy αντίστοιχα).

Στην συνέχεια παρουσιάζονται μερικά από τα μοντέλα και τα αποτελέσματα τους σε γραφήματα, όπου ο άξονας των x αντιστοιχεί στις εποχές (Epochs) και ο άξονας των y στην τιμή είτε της ακρίβειας ανά εποχή (accuracy) η στην τιμή απωλειών ανά εποχή (loss). Η οπτικοποίηση των αποτελεσμάτων έγινε σε περιβάλλον TensorBoard.

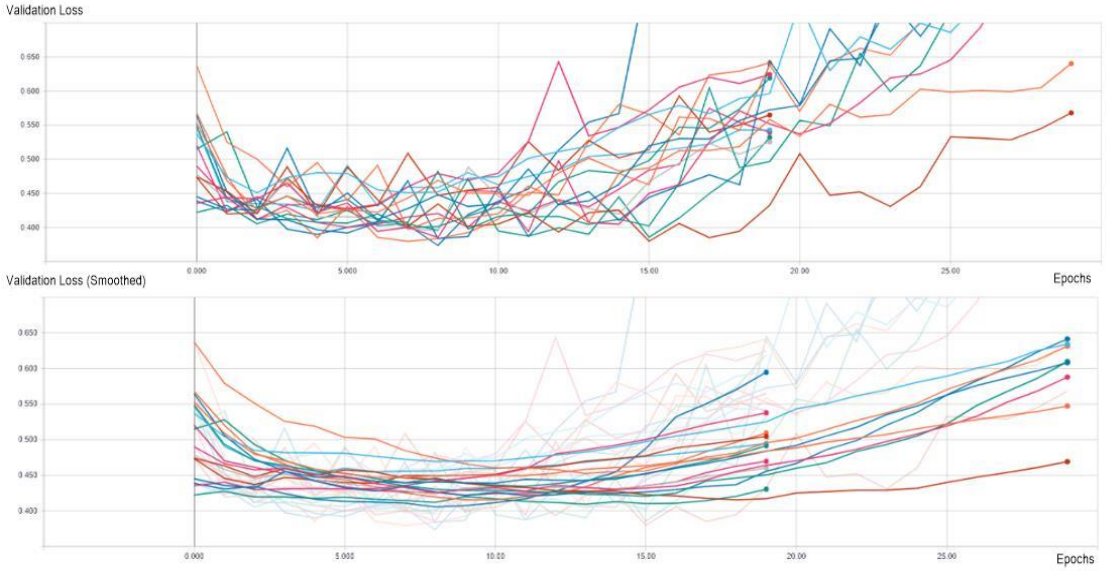
Στο Διάγραμμα 5-1 παρουσιάζεται η ακρίβεια και οι απώλειες των μοντέλων στο δείγμα εκπαίδευσης. Τα αποτελέσματα ακρίβειας και απωλειών στο δείγμα αξιολόγησης εμφανίζονται στο Διάγραμμα 5-2 και Διάγραμμα 5-3 αντόιστοιχα.



Διάγραμμα 5-1: Γραφήματα Ακρίβειας (πάνω) – Συνάρτησης Κόστους-Απωλειών (κάτω) της Εκπαίδευσης Αρχικών Μοντέλων (ανά εποχή)



Διάγραμμα 5-2: Γράφημα Ακρίβειας της Αξιολόγησης των Μοντέλων (2ο Γράφημα: Τα Αποτελέσματα Έχουν Εξομαλυνθεί για Καλύτερη Ορατότητα του Γραφήματος



Διάγραμμα 5-3: Γράφημα Τιμών Συνάρτησης Κόστους της Αξιολόγησης των Μοντέλων (2ο Γράφημα: Τα Αποτελέσματα Έχουν Εξομαλυνθεί για Καλύτερη Ορατότητα του Γραφήματος)

Παρατείνοντας το αποτελέσματα των απωλειών στην αξιολόγηση των μοντέλων(Διάγραμμα 5-3), βγαίνουν τα εξής αρχικά συμπεράσματα, που θα καθοδηγήσουν τα επόμενα βήματα:

- Τα μοντέλα με ένα επίπεδο συνέλιξης (convolutional layer) , έχουν αυξημένες απώλειες σε σχέση με τα υπόλοιπα.
- Τα χαρακτηριστικά των επίπεδων συνέλιξης επηρεάζουν τα αποτελέσματα.
- Τα μοντέλα χωρίς συμπαγές επίπεδο (Dense Layer) αποδίδουν χειρότερα από τα υπόλοιπα.
- Η χρήση της μεθόδου απόρριψης νευρώνων (Dropout) μειώνει το φαινόμενο της υπερπροσαρμογής.
- Σε αυτή την βάση δεδομένων τα βέλτιστα αποτελέσματα εμφανίζονται στις εποχές 5 – 10, για όλα τα μοντέλα.

5.2.2 - Εκτενής Θεώρηση

Με χρήση των παραπάνω πληροφοριών θα δομηθούν και θα δοκιμαστούν μοντέλα για εκπαίδευση στην δεύτερη βάση δεδομένων με διαστάσεις εικόνων 300:150 και 150:60, στις δυο χρωματικές κλίμακες RGB-Grayscale, σε 30-40 εποχές, περισσότερες από όσες χρειάζονται, ώστε να αποθηκευτεί το κάθε μοντέλο με τα βάρη (weights) και προκαταλήψεις (bias) που αντιστοιχούν στην εποχή που το δείγμα αξιολόγησης έδωσε το ελάχιστο σφάλμα. Επίσης να σημειωθεί πως σε αυτό το στάδιο, επιπλέον των παραπάνω συνδυασμών, εξετάζεται και η επιρροή της χρήσης κανονικοποίησης τμημάτων δεδομένων (Batch Normalization) σε διάφορα επίπεδα καθώς και η χρήση απόρριψης νευρώνων (Dropout) με ποσοστό από 0,2-0,75.

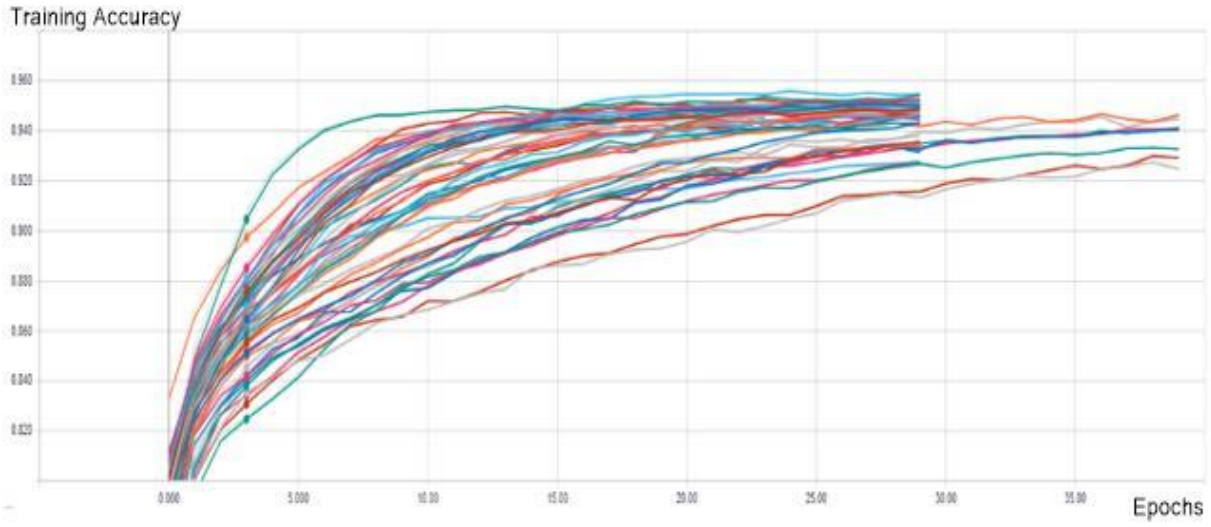
Στην συνέχεια παρουσιάζονται μερικά από τα μοντέλα και τα αποτελέσματα τους σε γραφήματα, όπου ο άξονας των x αντιστοιχεί στις εποχές και ο άξονας των y στην τιμή είτε της ακρίβειας ανά εποχή (accuracy) η στην τιμή της συνάρτησης κόστους ανά εποχή (loss).

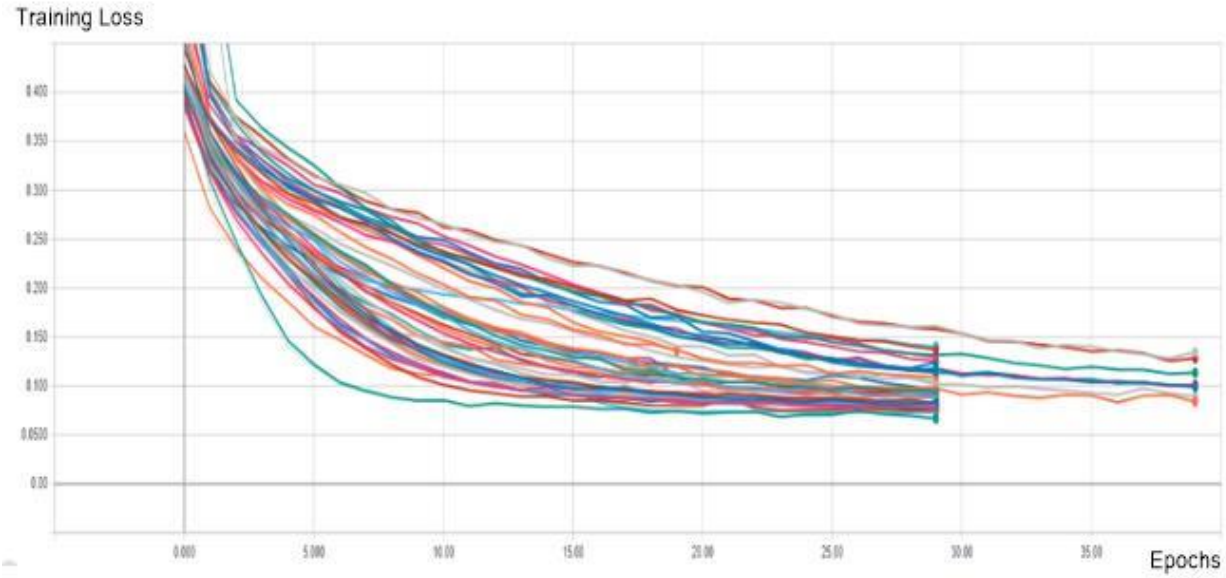
Στο Διάγραμμα 5-4 παρουσιάζεται η ακρίβεια και οι απώλειες των μοντέλων στο δείγμα εκπαίδευσης. Τα αποτελέσματα ακρίβειας και απωλειών στο δείγμα αξιολόγησης εμφανίζονται στο Διάγραμμα 5-5 και Διάγραμμα 5-6.

Όπως και στην παραπάνω περίπτωση χρησιμοποιήθηκε ονοματολογία μοναδική για κάθε πρότυπο που αναπτύσσεται (Πίνακας 5-1).

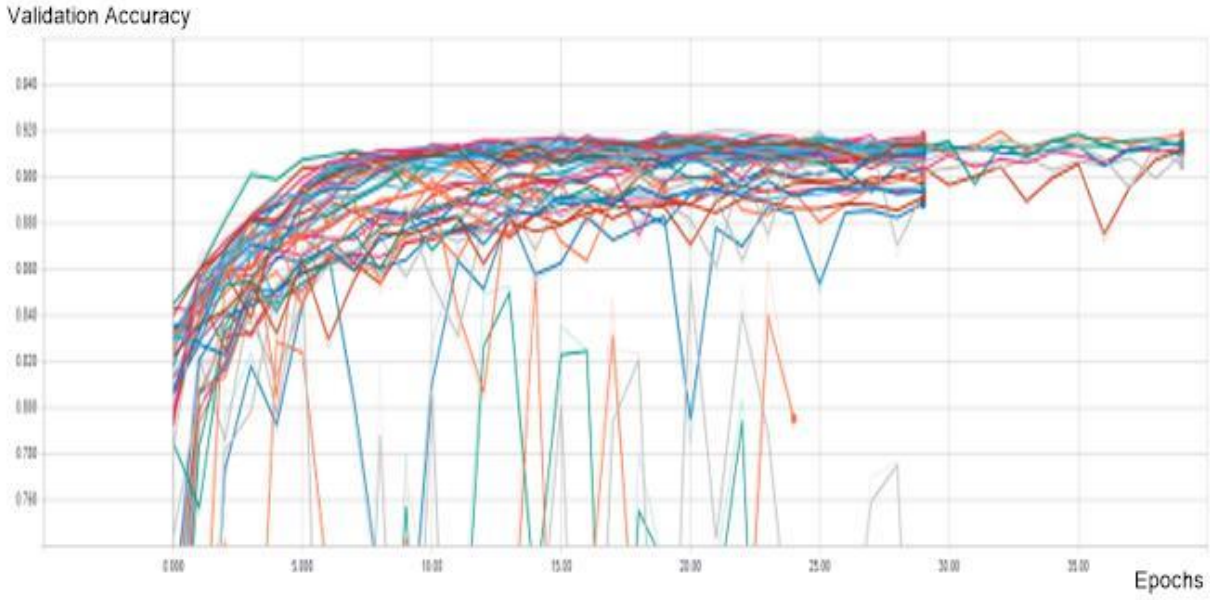
Πίνακας 5-1: Μερικά Εκτενή Εξεταζόμενα Μοντέλα και Χαρακτηριστικά τους

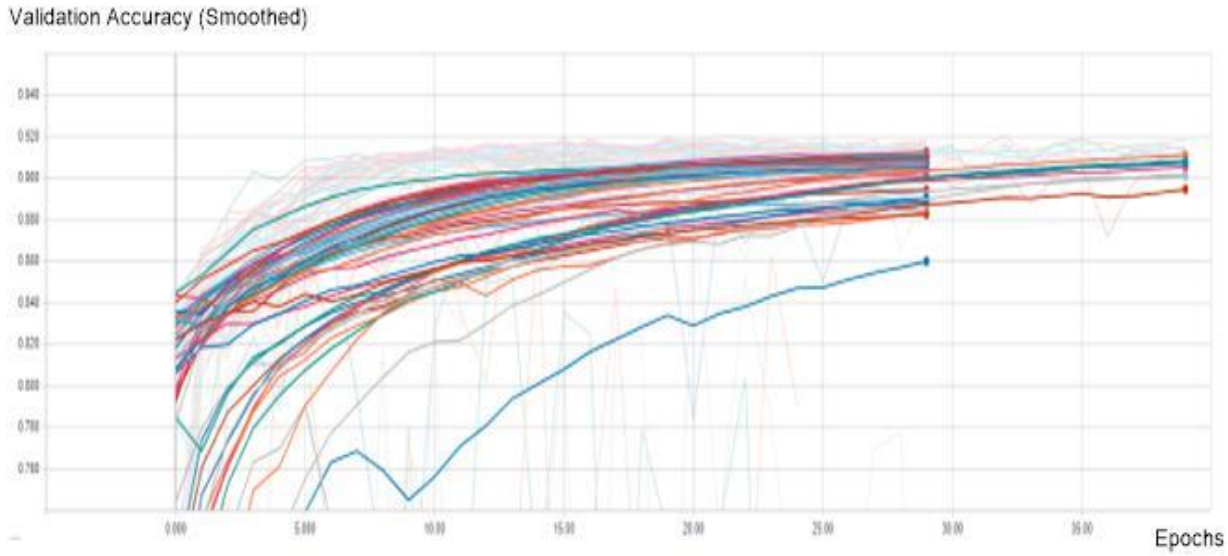
Name
AD_DS\2_CON-128_256_0_NODES-1_DEN-512_DSiz-0.5_DOPr-150_BS-30_E-0.3_VS-1539108964
AD_DS\2_CON-32_32_0_NODES-1_DEN-256_DSiz-0.5_DOPr-150_BS-30_E-0.3_VS-1539110700
AD_DS\2_CON-32_32_0_NODES-1_DEN-512_DSiz-0.5_DOPr-150_BS-30_E-0.3_VS-1539091923
AD_DS\2_CON-32_64_0_NODES-0_DEN-0_DSiz-0_DOPr-150_BS-30_E-0.3_VS-1539091046
AD_DS\2_CON-32_64_0_NODES-1_DEN-256_DSiz-0.5_DOPr-150_BS-30_E-0.3_VS-1539111056
AD_DS\2_CON-32_64_0_NODES-1_DEN-512_DSiz-0.5_DOPr-150_BS-30_E-0.3_VS-1539091416
AD_DS\2_CON-64_128_0_NODES-1_DEN-512_DSiz-0.5_DOPr-150_BS-30_E-0.3_VS-1539107313
AD_DS\2_CON-64_128_0_NODES-1_DEN-512_DSiz-0.5_DOPr-150_BS-30_E-0.3_VS-1539108141
AD_DS\2_CON-64_64_0_NODES-1_DEN-256_DSiz-0.5_DOPr-150_BS-30_E-0.3_VS-1539112400
AD_DS\2_CON-64_64_0_NODES-1_DEN-512_DSiz-0.5_DOPr-150_BS-30_E-0.3_VS-1539106681
AD_DS\2_CON-64_64_0_NODES-2_DEN-512_DSiz-0.5_DOPr-150_BS-30_E-0.3_VS-1539112957
AD_DS\3_CON-32_64_128_NODES-0_DEN-0_DSiz-0_DOPr-150_BS-30_E-0.3_VS-1539090535
AD_DS\3_CON-32_64_128_NODES-1_DEN-128_DSiz-0.2_DOPr-150_BS-30_E-0.3_VS-1539089151
AD_DS\3_CON-32_64_128_NODES-1_DEN-128_DSiz-0.5_DOPr-150_BS-30_E-0.3_VS-1539087837
AD_DS\3_CON-32_64_128_NODES-1_DEN-256_DSiz-0.2_DOPr-150_BS-30_E-0.3_VS-1539087191
AD_DS\3_CON-32_64_128_NODES-1_DEN-256_DSiz-0.5_DOPr-150_BS-30_E-0.3_VS-1539086749
AD_DS\3_CON-32_64_128_NODES-1_DEN-512_DSiz-0.2_DOPr-150_BS-30_E-0.3_VS-1539086205
AD_DS\3_CON-32_64_128_NODES-1_DEN-512_DSiz-0.5_DOPr-150_BS-30_E-0.3_VS-1539085448
AD_DS\3_CON-32_64_128_NODES-1_DEN-64_DSiz-0.5_DOPr-150_BS-30_E-0.3_VS-1539089673
AD_DS\3_CON-32_64_64_NODES-1_DEN-512_DSiz-0.5_DOPr-150_BS-30_E-0.3_VS-1539084124
AD_DS\3_CON-32_64_64_NODES-1_DEN-512_DSiz-0.5_DOPr-150_BS-30_E-0.3_VS_BN-1539163874
AD_DS\3_CON-32_64_64_NODES-1_DEN-512_DSiz-0.5_DOPr-150_BS-30_E-0.3_VS_BNallLayers-1539179720
AD_FS\3_CON-32_64_128_NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-30_E-0.3_VS-1539340693
AD_FS\3_CON-32_64_64_NODES-1_DEN-512_DSiz-0.5_DOPr-150_BS-30_E-0.3_VS-1539251087
AD_FS\3_CON-32_64_64_NODES-1_DEN-512_DSiz-0.5_DOPr-150_BS-30_E-0.3_VS-a1539252622
AD_FS\3_CON-32_64_64_NODES-1_DEN-512_DSiz-0.5_DOPr-150_BS-30_E-0.3_VS-b1539254664
TEST_LOGS\3_CON-32_64_64_NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-30_E-0.3_VS-1539344218
TEST_LOGS\BN_3_CON-32_64_64_NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-40_E-0.3_VS-1539357492
TEST_LOGS\BN_3_CON-64_64_64_NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-40_E-0.3_VS-1539429500
TEST_LOGS\BN_4_CON-32_64_128_256NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-40_E-0.3_VS-1539437060
TEST_LOGS\SM3_CON-32_64_64_NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-30_E-0.3_VS-1539354350
TEST_LOGS\SM3_CON-32_64_64_NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-30_E-0.3_VS-1539356129
cor_logs\BN_3_CON-32_64_128_0NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-40_E-0.3_VS-1539688405
cor_logs\BN_3_CON-32_64_128_0NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-40_E-0.3_VS-1539690557
cor_logs\DS_COR_BN_4_CON-32_64_128_256NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-40_E-0.3_VS-1539697623
cor_logs\FS_BN_3_CON-32_64_128_0NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-40_E-0.3_VS-1539691721
test1\DS_COR_4_CON-32_64_128_256NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-30_E-0.3_VS-1539787184
test1\DS_COR_BN_4_CON-32_64_128_256NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-20_E-0.3_VS-1539785672
test1\DS_COR__3_CON-32_64_64_0NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-40_E-0.3_VS-1539791926
test2\DS_COR_3_CON-64_64_124_0NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-30_E-0.3_VS-1540210156
test2\DS_COR_BN_3_CON-64_64_128_0NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-30_E-0.3_VS-1540200943
test2\DS_COR_BN_3_CON-64_64_128_0NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-30_E-0.3_VS-1540207477
test2\DS_COR_BN_3_CON-64_64_128_0NODES-1_DEN-512_DSiz-0.75_DOPr-256_BS-30_E-0.3_VS-1540200124
test2\DS_COR__2_CON-128_128_0_0NODES-1_DEN-512_DSiz-0.75_DOPr-256_BS-30_E-0.3_VS-1540198312
test2\DS_COR__3_CON-64_64_128_0NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-30_E-0.3_VS-1540208394
test2\DS_COR__3_CON-64_64_128_0NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-30_E-0.3_VS-1540209150
test2\SM_DS_COR_3_CON-32_64_64_0NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-30_E-0.2_VS-1540282598



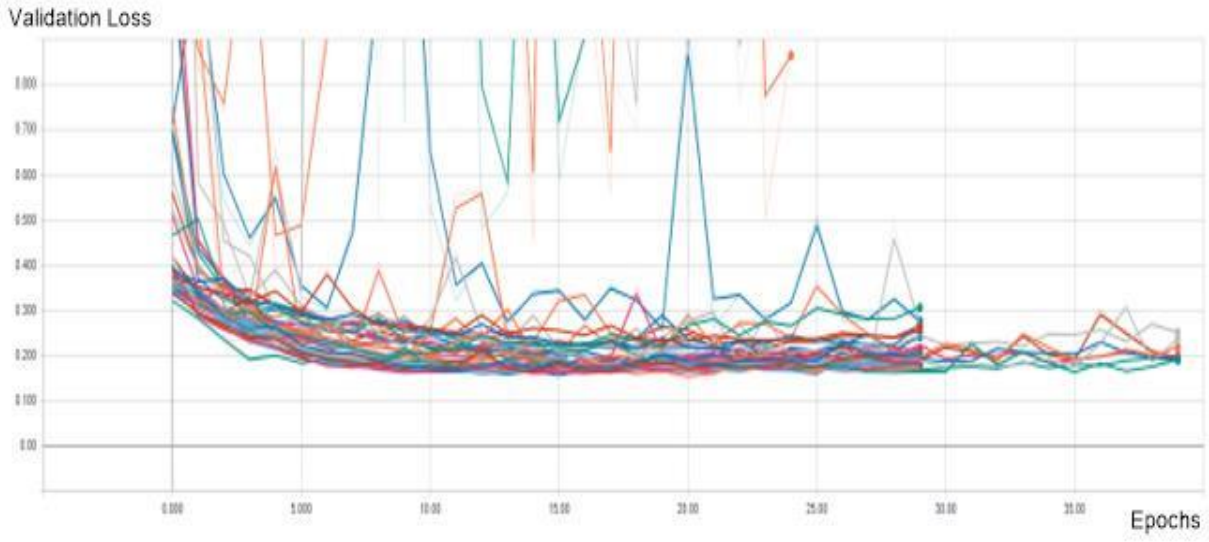


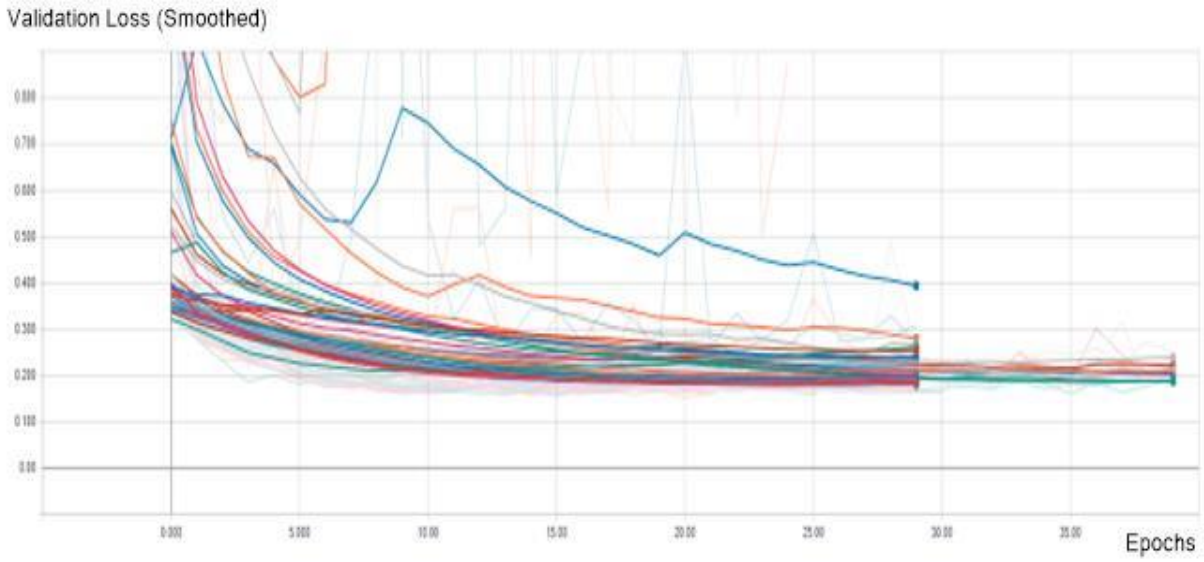
Διάγραμμα 5-4: Γραφήματα Ακρίβειας (πάνω) – Συνάρτησης Κόστους (κάτω) της Εκπαίδευσης των Μοντέλων





Διάγραμμα 5-5: Γράφημα Ακρίβειας της Αξιολόγησης των Μοντέλων (2ο Γράφημα: Τα Αποτελέσματα Έχουν Εξομαλυνθεί για Καλύτερη Ορατότητα του Γραφήματος)





Διάγραμμα 5-6: Γράφημα Τιμών Συνάρτησης Κόστους της Αξιολόγησης των Μοντέλων (2ο Γράφημα: Τα Αποτελέσματα Έχουν Εξομαλυνθεί για Καλύτερη Ορατότητα του Γραφήματος)

Από τα αποτελέσματα της αξιολόγησης που παρουσιάστηκαν προηγουμένως (Διάγραμμα 5-5, Διάγραμμα 5-6), θα απομονωθούν εκείνα τα πρότυπα που είχαν τις ελάχιστες απώλειες.

Στην συνέχεια παρατηρώντας το αποτελέσματα των απωλειών (Διάγραμμα 5-6) και της ακρίβειας (Διάγραμμα 5-4) κατά την αξιολόγηση, με την δεύτερη βάση δεδομένων, βγαίνουν τα εξής συμπεράσματα:

- Τα μοντέλα με δύο η παραπάνω επίπεδα συνέλιξης (convolutional layers), αποδίδουν καλύτερα.
- Τα δεδομένα στην χρωματική κλίμακα Grayscale εκπαιδεύονται σημαντικά πιο γρήγορα από αυτά που είναι στην χρωματική κλίμακα RGB, δίνοντας παρόμοια αποτελέσματα. Αυτό μπορεί να συμβαίνει γιατί στο συγκεκριμένο πείραμα η πληροφορία του χρώματος δεν αποτελεί κριτήριο που επηρεάζει το αποτέλεσμα. Πιθανόν αν προστεθούν και άλλες απαιτήσεις στο πείραμα να είναι χρήσιμη, υπερδιπλασιάζοντας τον χρόνο εκπαίδευσης.
- Τα δεδομένα με μειωμένες διαστάσεις 150:60, δίνουν σε γενικές καλύτερα αποτελέσματα. Γεγονός που βοηθάει στην εκπαίδευση καθώς η χρονική διάρκεια της εκπαίδευσης του ιδίου μοντέλου στις εικόνες με διπλάσια ανάλυση 300:150 είναι παραπάνω από διπλάσια.
- Τα χαρακτηριστικά των επίπεδων συνέλιξης όπως ο αριθμός των convolutions/filters/kernel size ανά επίπεδο επηρεάζουν τα αποτελέσματα.
- Η βέλτιστη δομή αποτελείται από ένα συμπαγές επίπεδο (Dense layer), και μεγέθους 512.
- Η χρήση της μεθόδου απόρριψης νευρώνων (Dropout) μειώνει σημαντικά το φαινόμενο Υπερπροσαρμογής, με βέλτιστα αποτελέσματα για τιμές 0.5 έως 0.75.
- Η εφαρμογή κανονικοποίησης τμημάτων (Batch Normalization) σε συγκεκριμένα επίπεδα μειώνει σημαντικά το φαινόμενο υπερπροσαρμογής, αλλιώς επηρεάζει πολύ αρνητικά την απόδοση του μοντέλου.
- Σε αυτή την βάση δεδομένων τα βέλτιστα αποτελέσματα εμφανίζονται στις εποχές 10-15, για όλα τα μοντέλα. Από την εποχή 15 εμφανίζεται πλατό με λίγα μοντέλα να έχουν σημαντική βελτίωση.

Σημειώνεται ότι με περαιτέρω βελτιστοποίηση των παραμέτρων εκπαίδευσης και δομής των προτύπων, τα αποτελέσματα ενδέχεται να είναι όμοια (εντός των περιθωρίων σφάλματος) η και καλύτερα

Στην συνέχεια τα βέλτιστα μοντέλα θα εξεταστούν σε κοινή βάση δεδομένων με 1000 εικόνες για κάθε κατηγορία, και θα υπολογιστούν στατιστικά μεγέθη για την τελική αξιολόγηση του κάθε μοντέλου.

5.2.3 - Βέλτιστο Μοντέλο

Θα παρουσιαστούν τα στοιχεία και τα αποτελέσματα για το μοντέλο με τα βέλτιστα αποτελέσματα με όνομα:

«BN_3_CON-32_64_128_0NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-40_E-0.3_VS-1539690557»

Να σημειωθεί ότι το μοντέλο αυτό λειτουργεί για δεδομένα εισόδου μόνο εικόνες διαστάσεων 150:60 pixels και στην χρωματική κλίμακα Grayscale.

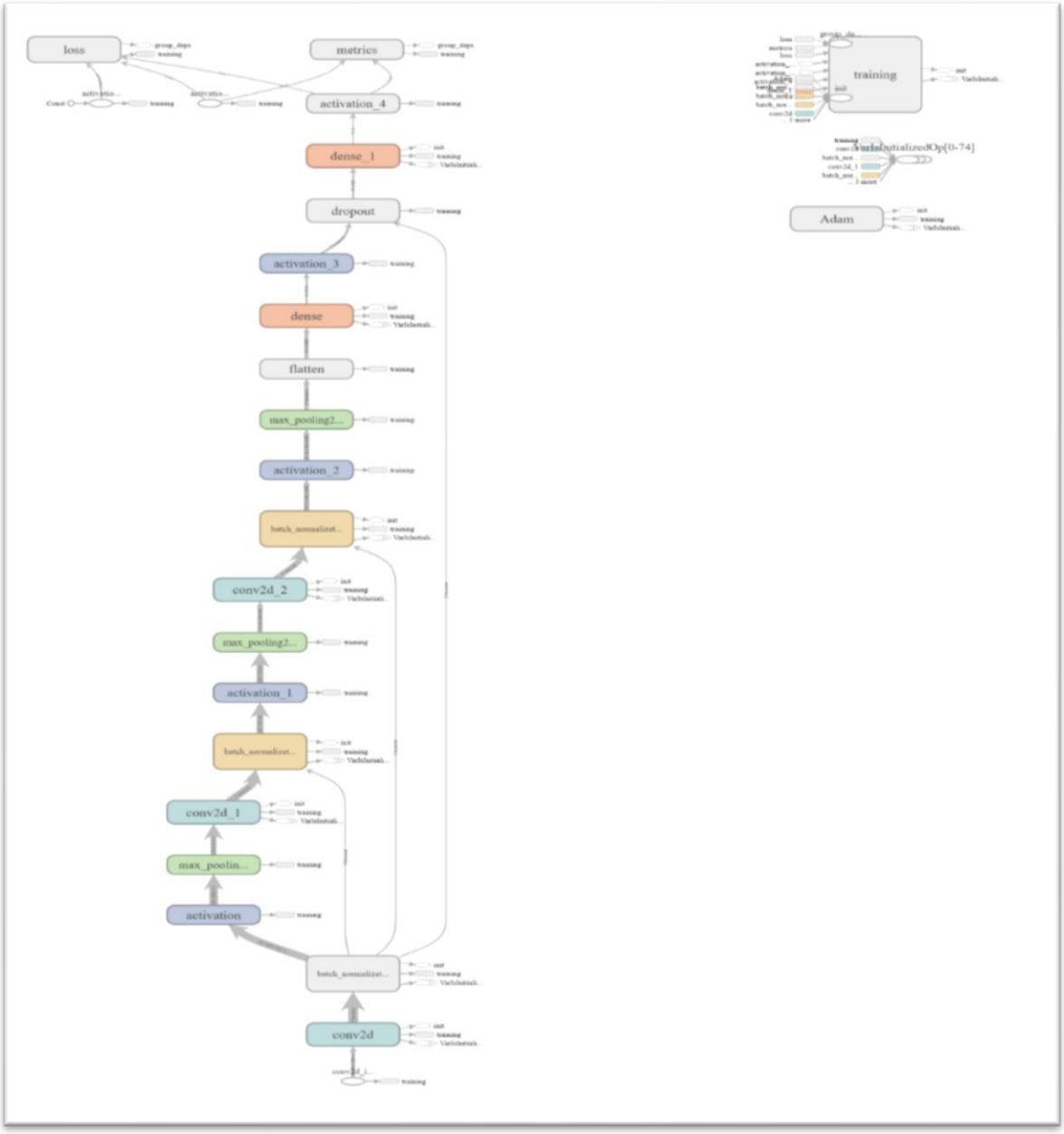
Η εκπαίδευση έγινε για 40 εποχές, με συνολική διάρκεια 15.16 λεπτά σε κάρτα γραφικών. Σε αυτό το σημείο είναι εμφανής η ανάγκη χρήσης κατάλληλου εξοπλισμού καθώς το ίδιο δίκτυο με ίδιες παραμέτρους αν εκπαιδευτεί σε επεξεργαστή, η εκπαίδευση έχει διάρκεια πάνω από 40 φορές (περισσότερες από 10 ώρες) σε σύγκριση με την εκπαίδευση σε κάρτα γραφικών με επαρκή μνήμη (Πίνακας 5-2).

Πίνακας 5-2: Χαρακτηριστικά Εκπαίδευσης Δικτύου 1ης Μεθόδου

Εκπαίδευση Συνελικτικού Νευρωνικού Δικτύου		
Μοντέλο	BN_3_CON-32_64_128_0NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-40_E-0.3_VS-1539690557	
Μέσο Εικπαίδευσης και Υπολογισμών	GPU	
Πλήθος Δεδομένων Εκπαίδευσης	50000	(εικόνες)
Εποχές Εκπαίδευσης	40	
Μέγεθος Τμημάτων Εκπαίδευσης (Batch Size)	150	(εικόνες ανα τμήμα)
Διάρκεια Εικπαίδευσης Δικτύου	15.16	(λεπτα)

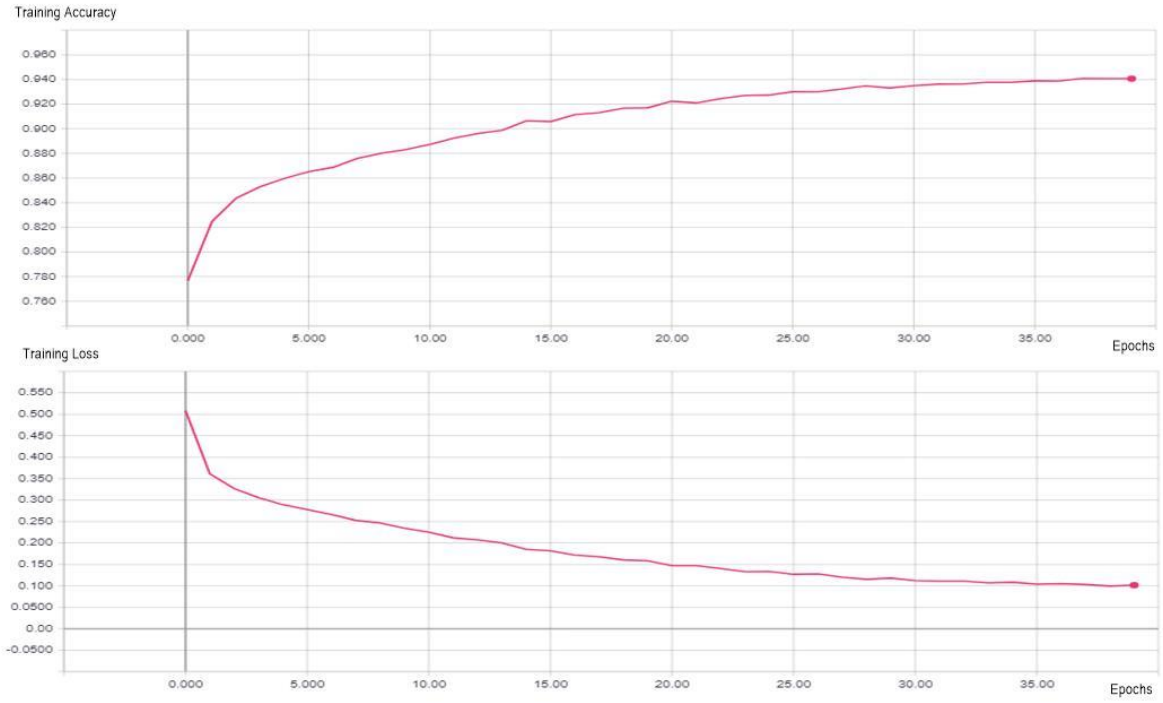
Στο Διάγραμμα 5-7 παρουσιάζονται η δομή και τα χαρακτηριστικά διαγράμματα του συνελικτικού μοντέλου και της εκπαίδευσης/αξιολόγησης του.

Στο Διάγραμμα 5-8 και Διάγραμμα 5-9 απομονώνονται τα γραφήματα εκπαίδευσης και αξιολόγησης του συγκεκριμένου μοντέλου αντίστοιχα.

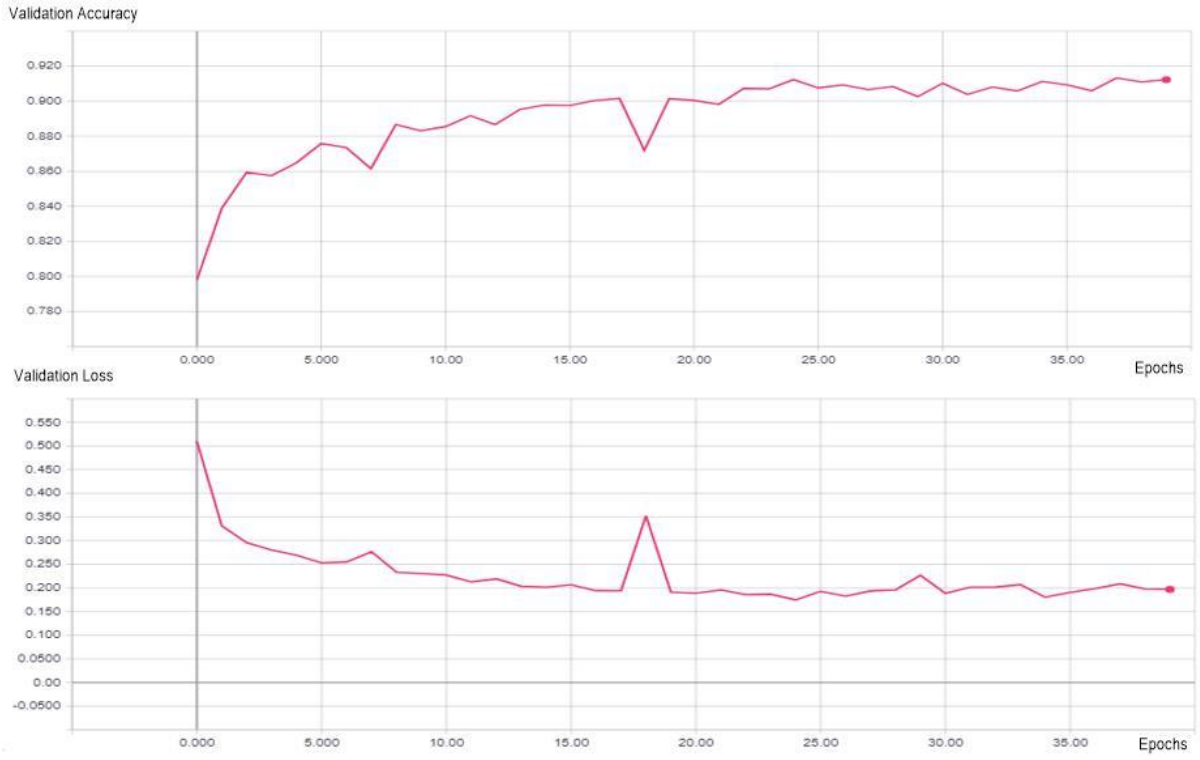


Διάγραμμα 5-7: Απεικόνιση Δομής- Χαρακτηριστικών Εξεταζόμενου Μοντέλου και Ροής Λειτουργειών. (Διάγραμμα Ροής από TensorBoard)

.



Διάγραμμα 5-8: Γραφήματα Ακρίβειας(πάνω) – Συνάρτησης Κόστους (κάτω) της Εκπαίδευσης Εξεταζόμενου Μοντέλου



Διάγραμμα 5-9: Γραφήματα Ακρίβειας(πάνω) – Συνάρτησης Κόστους (κάτω) της Αξιολόγησης Εξεταζόμενου Μοντέλου

Όπως παρατηρείται (Διάγραμμα 5-8, Διάγραμμα 5-9), το μοντέλο εμφανίζει μέγιστη ακρίβεια και ελάχιστες απώλειες κατά τις εποχές 24-25, επομένως το μοντέλο θα εξεταστεί με τα βάρη (Weights) και τις προκαταλήψεις (Biases) που είχε σε εκείνη την εποχή, το οποίο είναι εφικτό καθώς κατά την εκπαίδευση έχει αποθηκευτεί ξεχωριστά το μοντέλο με τις ελάχιστες απώλειες.

Στην συνέχεια (Πίνακας 5-3) παρουσιάζονται τα **αποτελέσματα και τα στατιστικά μεγέθη** του μοντέλου. Τα αποτελέσματα είναι πολύ ικανοποιητικά καθώς επιτυγχάνεται μεγάλη ακρίβεια προβλέψεων στο δείγμα δοκιμής. Σημαντικό είναι να σημειωθεί πως με τελείως ξένο δείγμα δοκιμής (από άλλο δρόμο) το μοντέλο αναμένεται να είναι σημαντικά πιο ευαίσθητο και τα αποτελέσματα αναμένεται να είναι σημαντικά χαμηλότερα Πρόβλημα που ενδεχομένως να αντιμετωπίζεται με μεγαλύτερο πλήθος δεδομένων υπό διαφορετικές συνθήκες και περιβάλλον.

Πίνακας 5-3: Αποτελέσματα Αξιολόγησης Συνελικτικού Νευρωνικού Δικτύου

Μητρώο Προβλέψεων/Πραγματικότητας

Συνολικά Δεδομένα Αξιολόγησης - 2000		Πραγματικές Συνθήκες	
		0	1
Προβλεπόμενες Συνθήκες	0	947	53
	1	123	877

Κατηγορία 0	Χαμήλη/Καθόλου Κυκλοφορία
Κατηγορία 1	Υψηλή Κυκλοφορία

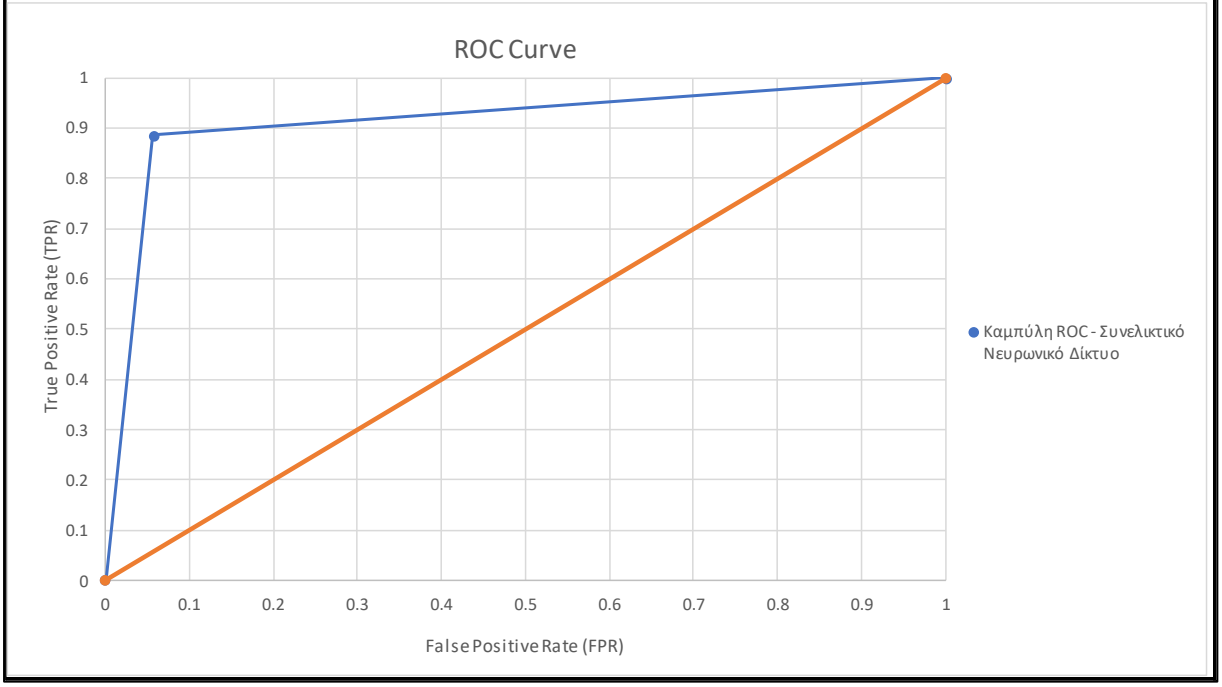
True Positive	947	(TP)
False Positive	53	(FP)
True Negative	877	(TN)
False Negative	123	(FN)

Αποτελέσματα

Ευαισθησία-Sensitivity	**0.885**	(TPR)
Εξειδίκευση-Specificity	**0.943**	(TNR)
Ακρίβεια-Precision	**0.947**	(PPV)
False Negative Rate - Error Type II	**0.115**	(FNR)
False Positive Rate - Error Type I	**0.057**	(FPR)

Ορθότητα-Accuracy	**0.912**	(ACC)
Αρμονικό Μέσο-F1 Score	**0.915**	(F1)

Area Under Curve	**0.914**	(AUC)

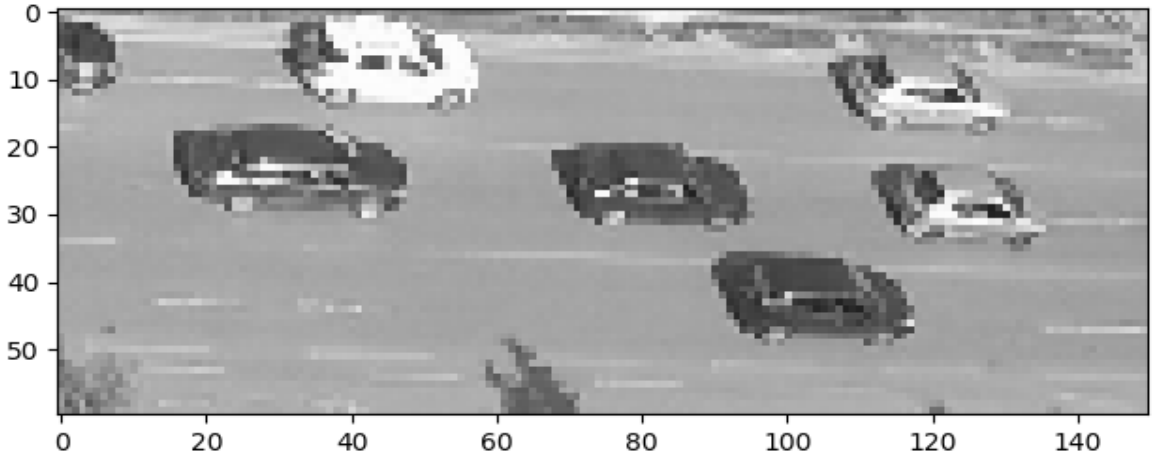


Οπτικοποίηση των Συνελικτικών Επιπέδων

Τα συνελικτικά επίπεδα κατά την εκπαίδευση μαθαίνουν μοτίβα και εφαρμόζουν φίλτρα για την αναγνώριση αυτών. Ξεκινώντας από τα πρώτα επίπεδα αναγνωρίζουν απλά χαρακτηριστικά όπως γραμμές και καταλήγουν σε επόμενα επίπεδα να αναγνωρίζουν πιο περίπλοκα μοτίβα και χαρακτηριστικά. Αυτά τα μοτίβα, υπάρχει η δυνατότητα, να οπτικοποιηθούν και να βοηθήσουν στην αναγνώριση των χαρακτηριστικών της εικόνας που ενεργοποιούν τα φίλτρα για κάθε περίπτωση.

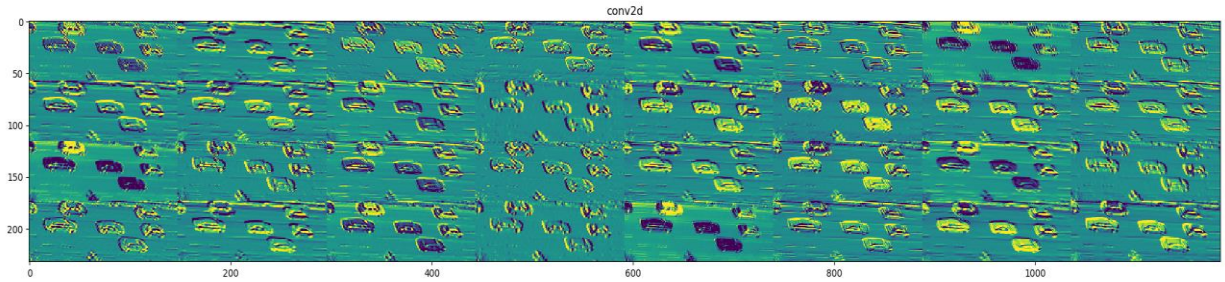
Στο συγκεκριμένο πείραμα θα εξεταστεί μία τυχαία εικόνα από κάθε κατηγορία, με το μοντέλο που εξετάστηκε προηγουμένως, επομένως η εικόνα πρέπει πρώτα να μετασχηματιστεί σε διαστάσεις 150:60 και Grayscale (Εικόνα 5-1, Εικόνα 5-6).

Κατηγορία 1 – Υψηλή Κυκλοφορία

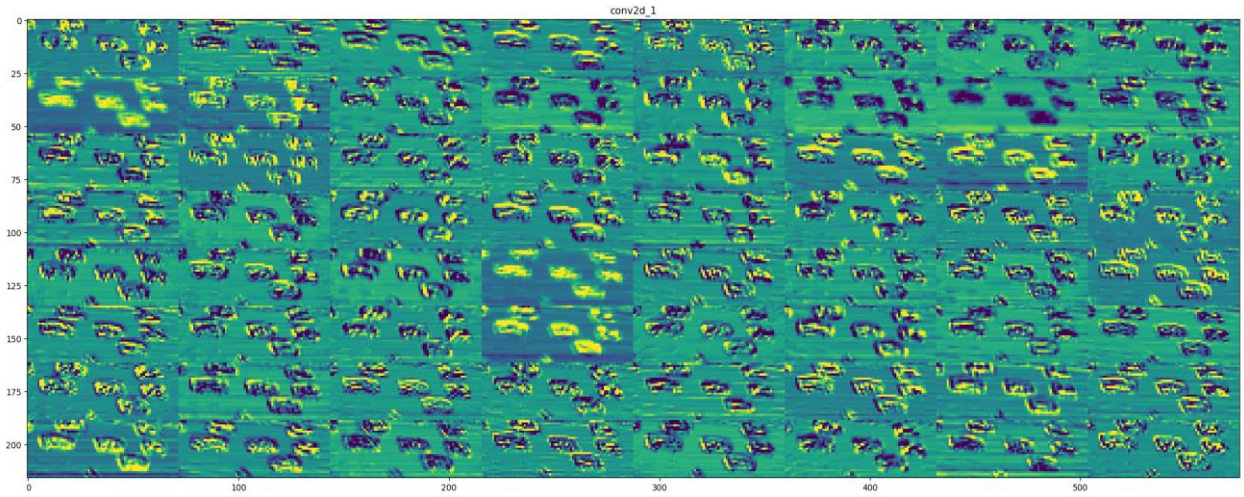


Εικόνα 5-1: Αρχική Εικόνα Κατηγορίας 1 (Διαστάσεις 150:60)

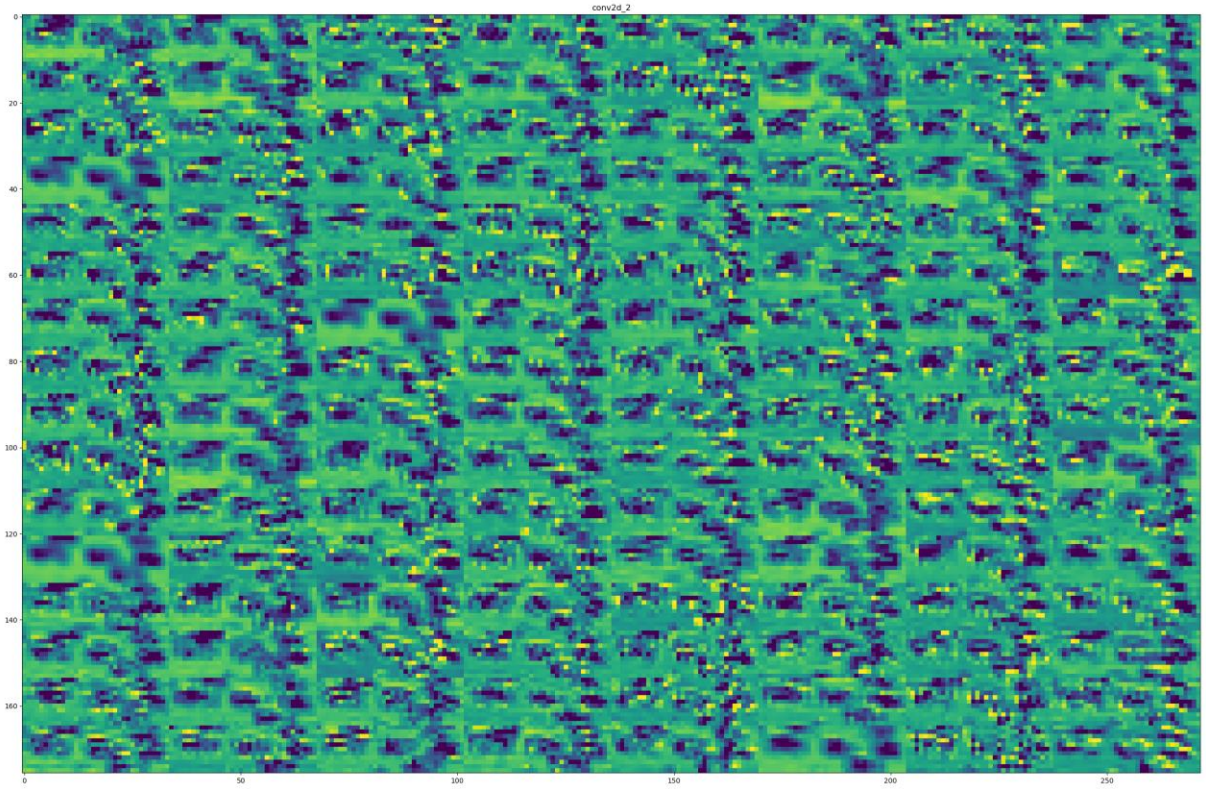
Κάθε επίπεδο συνέλιξης έχει διαφορετικό πλήθος φίλτρων, που κάθε ένα ενεργοποιείται από διαφορετικά χαρακτηριστικά όπως φαίνεται και στην συνέχεια (Εικόνα 5-2, Εικόνα 5-3, Εικόνα 5-4,Εικόνα 5-5, Εικόνα 5-5, Εικόνα 5-7, Εικόνα 5-8, Εικόνα 5-9, Εικόνα 5-10).



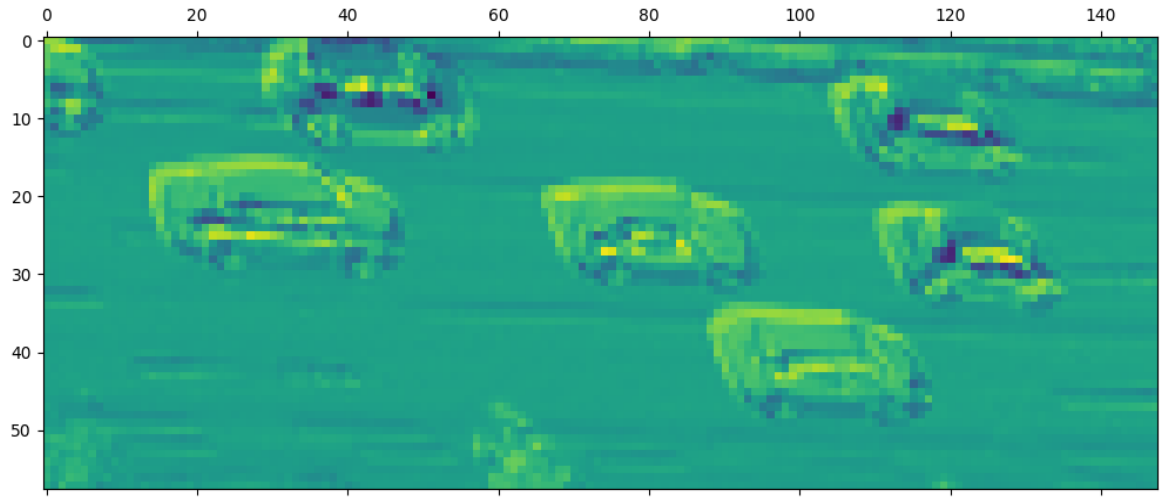
Εικόνα 5-2: 1ο Επίπεδο Συνέλιξης - 32 Φίλτρα – Κατηγορία 1



Εικόνα 5-3: 2ο Επίπεδο Συνέλιξης - 64 Φίλτρα - Κατηγορία 1



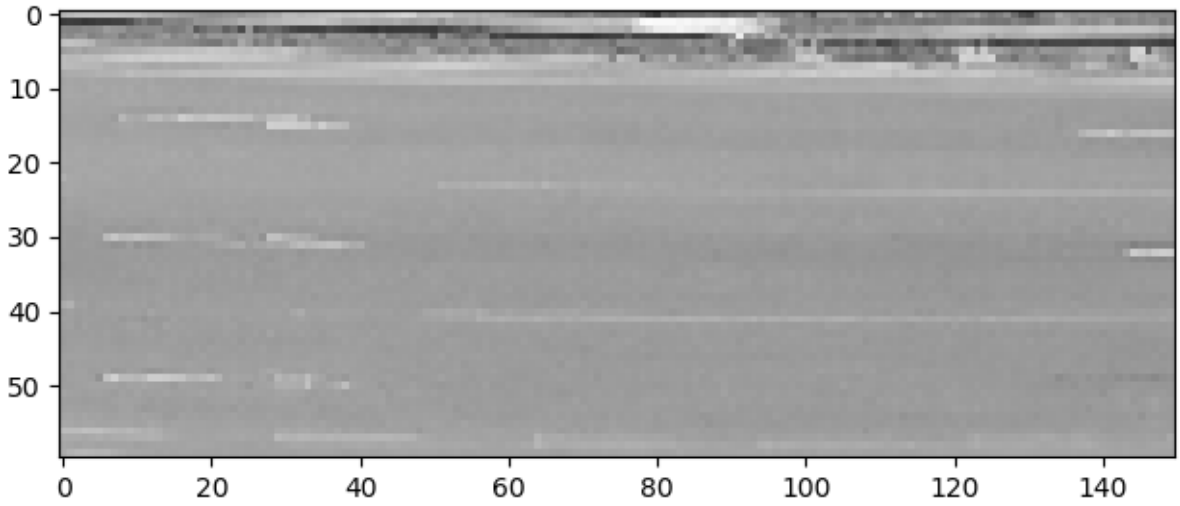
Εικόνα 5-4: 3ο Επίπεδο Συνέλιξης - 128 Φίλτρα - Κατηγορία 1



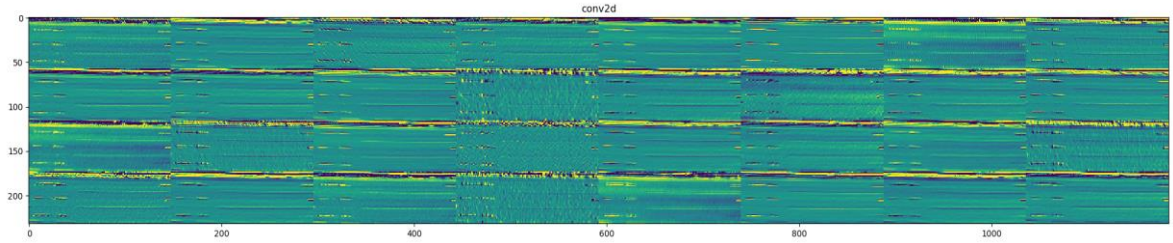
Εικόνα 5-5: Ενδεικτικά - 31ο Φίλτρο 1ου Επιπέδου Συνέλιξης - Κατηγορία 1

Από τα παραπάνω παρατηρείται πως το μοτίβο που ενεργοποιεί πολλά φίλτρα είναι το σύνολο των αυτοκινήτων στην οδό, η χαρακτηριστικά αυτών.

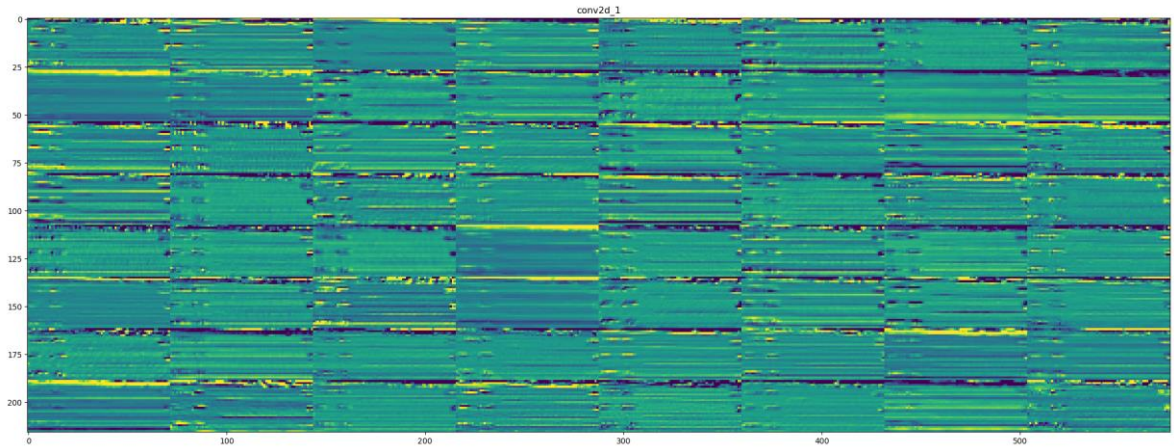
Κατηγορία 0 – Χαμηλή/Καθόλου Κυκλοφορία



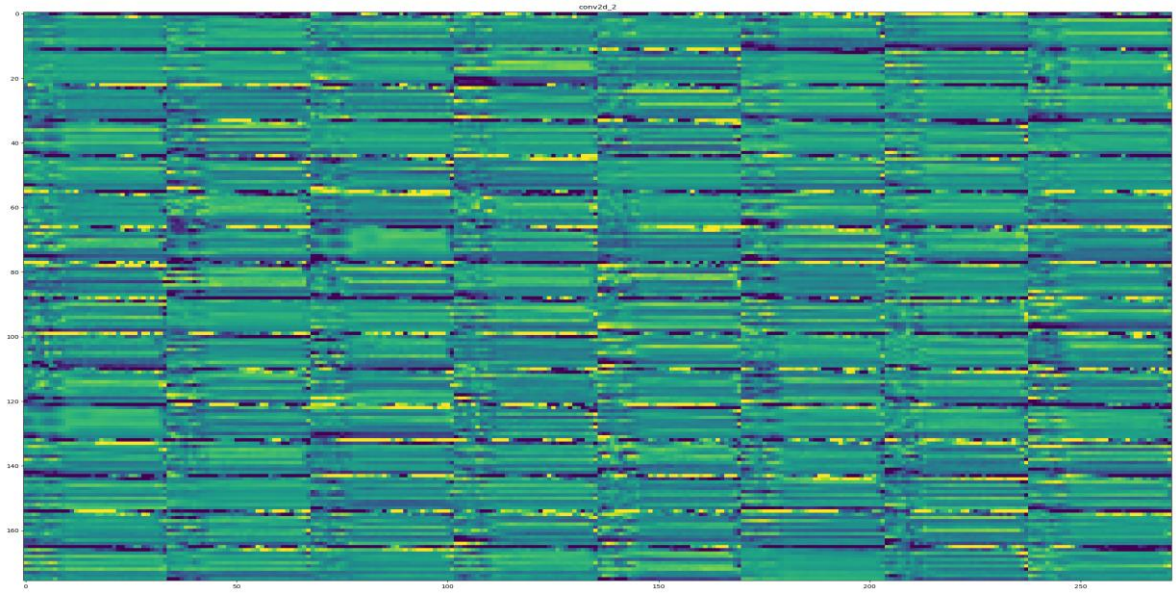
Εικόνα 5-6: Αρχική εικόνα κατηγορίας 0 (Διαστάσεις 150:60)



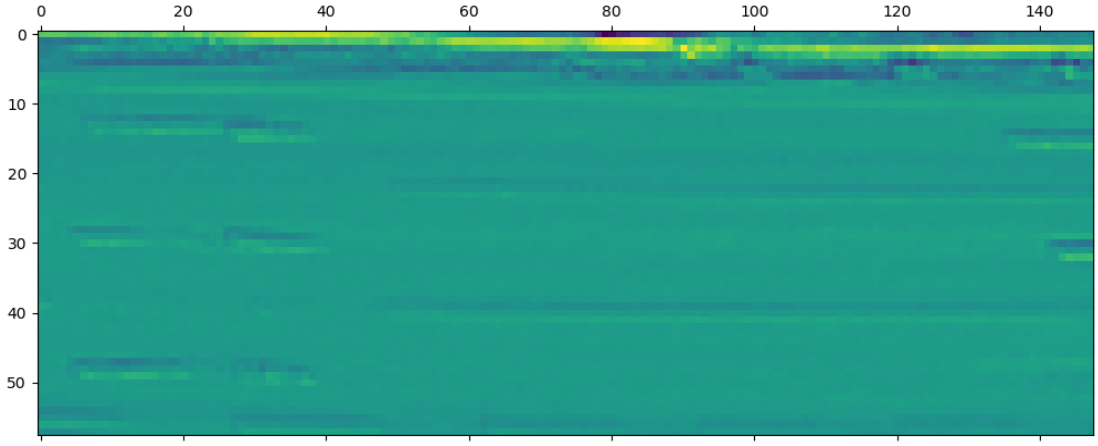
Εικόνα 5-7: Επίπεδο Συνέλιξης - 32 Φίλτρα – Κατηγορία 0



Εικόνα 5-8: 2ο Επίπεδο Συνέλιξης - 64 Φίλτρα - Κατηγορία 0



Εικόνα 5-9: 3ο Επίπεδο Συνέλιξης - 128 Φίλτρα - Κατηγορία 0



Εικόνα 5-10: Ενδεικτικά - 31ο Φίλτρο 1ου Επιπέδου Συνέλιξης - Κατηγορία 0

Από τα παραπάνω παρατηρείται πως στην περίπτωση που δεν υπάρχουν οχήματα δεν ενεργοποιούνται τα φίλτρα που εντόπιζαν οχήματα και χαρακτηριστικά τους αλλά φίλτρα που ενεργοποιούνται από τις λωρίδες της οδού και τα όρια της.

5.2.4 - Εφαρμογή Συνελικτικού Νευρωνικού Δικτύου σε Πραγματικό Χρόνο

Με την επιλογή κατάλληλου συνελικτικού νευρωνικού δικτύου, θα δομηθεί πρόγραμμα που θα δέχεται ως δεδομένο εισόδου ένα υπάρχον βίντεο ή βίντεο από κάποια πηγή σε πραγματικό χρόνο (κάμερα) και το επιθυμητό νευρωνικό δίκτυο ώστε να γίνεται πρόβλεψη για κάθε στιγμιότυπο του βίντεο.

Επίσης θα δοθεί ένα χρονικό διάστημα σε δευτερόλεπτα (π.χ. 10 δευτερότερα) και με χρήση ενός αλγόριθμου επιλογής πλειοψηφίας (Majority Voting Algorithm), θα καθορίζεται η κυκλοφορία σε κάθε χρονικό διάστημα ως κατηγορία 1 η κατηγορία 0

Το πρόγραμμα λειτουργεί με δεδομένα εισόδου τα εξής:

- Το αρχείο του βίντεο.
- Το μοντέλο με το οποίο θα γίνονται οι προβλέψεις.
- Το χρονικό διάστημα ανά το οποίο θα γίνεται πρόβλεψη του στιγμιότυπου ή επιλογή πρόβλεψης όλων των στιγμιότυπων.
- Το χρονικό διάστημα σε δευτερόλεπτά που θα γίνεται η πρόβλεψη.
- Η επιλογή τμήματος που θα γίνεται η ανάλυση επιλέγοντας το αρχικό σημείο περικοπής του βίντεο καθώς και τις επιθυμητές διαστάσεις του τμήματος. (Με σταθερό λόγο διαστάσεων 1:2.5).
- Ανάλογα το μοντέλο, κατάλληλος μετασχηματισμός και επεξεργασία των εικόνων.

Το πρόγραμμα επιστρέφει τις προβλέψεις κάθε στιγμιότυπου και προβάλει το εξεταζόμενο τμήμα του βίντεο αναγράφοντας την πρόβλεψη κάθε στιγμή. Στην συνέχεια παρουσιάζονται

δυο ενδεικτικά στιγμιότυπα του προβαλλόμενου βίντεο για τις δύο κατηγορίες πρόβλεψής. (Εικόνα 5-11, Εικόνα 5-12)

Το βίντεο που θα χρησιμοποιηθεί, δεν έχει συμπεριληφθεί στα δεδομένα εκπαίδευσης των δικτύων, ώστε θα αντιμετωπιστεί ως ξένο δεδομένο. Να σημειωθεί πως λόγω δεδομένων από μία οδό τα δεδομένα από το βίντεο δεν θα είναι πλήρως ανεξάρτητα από τα δεδομένα εκπαίδευσης καθώς θα υπάρχουν σημαντικές ομοιότητες. Για βέλτιστα αποτελέσματα μπορεί να πραγματοποιηθεί σταθεροποίηση του εξεταζόμενου βίντεο με χρήση κατάλληλου προγράμματος.

Εικόνα 5-11: Ενδεικτικό Στιγμιότυπο Πρόβλεψης - Κατηγορία 1

Εικόνα 5-12: Ενδεικτικό Στιγμιότυπο Πρόβλεψης - Κατηγορία 0

Στο ενδεικτικό βίντεο τα αποτελέσματα των περιόδων διάρκειας 10 δευτερολέπτων παρουσιάζονται στην συνέχεια (Πίνακας 5-4).

Πίνακας 5-4: Αποτελέσματα ανά Χρονική Περίοδο στο Ενδεικτικό Βίντεο–1η Μέθοδος

Συνολική Διάρκεια Βίντεο	5,38	(λεπτά - 338 δευτερόλεπτα)

t -Διάρκεια Περιόδου	10	(δευτερόλεπτα)

Περίοδοι	**Μέθοδος 1 - Αναγνώριση Κυκλοφορικών Συνθηκών σε Πραγματικό Χρόνο**	
(10 δευτερολεπτων)	Κατηγορία	Κατηγορία
1	1	Υψηλή Κυκλοφορία
2	0	Χαμηλή/Καθόλου Κυκλοφορία
3	0	Χαμηλή/Καθόλου Κυκλοφορία
4	1	Υψηλή Κυκλοφορία
5	1	Υψηλή Κυκλοφορία
6	1	Υψηλή Κυκλοφορία
7	1	Υψηλή Κυκλοφορία
8	1	Υψηλή Κυκλοφορία
9	0	Χαμηλή/Καθόλου Κυκλοφορία
10	0	Χαμηλή/Καθόλου Κυκλοφορία
11	0	Χαμηλή/Καθόλου Κυκλοφορία
12	0	Χαμηλή/Καθόλου Κυκλοφορία
13	1	Υψηλή Κυκλοφορία
14	1	Υψηλή Κυκλοφορία
15	1	Υψηλή Κυκλοφορία
16	1	Υψηλή Κυκλοφορία
17	1	Υψηλή Κυκλοφορία
18	1	Υψηλή Κυκλοφορία
19	1	Υψηλή Κυκλοφορία
20	0	Χαμηλή/Καθόλου Κυκλοφορία
21	0	Χαμηλή/Καθόλου Κυκλοφορία
22	1	Υψηλή Κυκλοφορία
23	1	Υψηλή Κυκλοφορία
24	1	Υψηλή Κυκλοφορία
25	1	Υψηλή Κυκλοφορία
26	1	Υψηλή Κυκλοφορία
27	1	Υψηλή Κυκλοφορία
28	1	Υψηλή Κυκλοφορία
29	0	Χαμηλή/Καθόλου Κυκλοφορία
30	0	Χαμηλή/Καθόλου Κυκλοφορία
31	0	Χαμηλή/Καθόλου Κυκλοφορία
32	1	Υψηλή Κυκλοφορία
33	1	Υψηλή Κυκλοφορία
34	0	Χαμηλή/Καθόλου Κυκλοφορία

Στην συνέχεια παρουσιάζεται η χρονική απαίτηση για την ανάλυση του ενδεικτικού βίντεο στον επεξεργαστή και στην κάρτα γραφικών και σύγκριση των δύο (Πίνακας 5-5).

Πίνακας 5-5: Χαρακτηριστικά Υπολογιστικές Επιδώσεις Εφαρμογής Δικτύου 1ης Μεθόδου στο Ενδεικτικό Βίντεο

Εφαρμογή Συνελικτικού Νευρωνικού Δικτύου			
Μοντέλο	BN_3_CON-32_64_128_0NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-40_E-0.3_VS-1539690557		
Μέσο Επεξεργασίας και Υπολογισμών	GPU	CPU	
Διάρκεια Εξεταζόμενου Βίντεο	5.64		(λεπτα)
Διάρκεια Ανάλυσης/Πρόβλεψης Βίντεο	0.51	0.92	(λεπτα)
Μέση Διάρκεια Ανάλυσης Εικόνας	1.4	4.4	(ms)
Διάρκεια Ανάλυσης σε Σχέση με το Εξεταζόμενο Βίντεο	-91%	-84%	(% Μείωση)
Σχέση Διάρκειας Ανάλυσης GPU με CPU	-45%		(% Μείωση)

Παρατηρείται πως και με τους δύο τρόπους ανάλυσης του βίντεο, η ανάλυση γίνεται πολύ πιο γρήγορα από την διάρκεια του βίντεο, επομένως με αυτόν τον εξοπλισμό είναι δυνατή η ανάλυση από βίντεο σε πραγματικό χρόνο. Επίσης με τη κάρτα γραφικών είναι δύο φορές πιο γρήγορη η ανάλυση σε σχέση με την ανάλυση στον επεξεργαστή.

Επίσης σημαντικό είναι να σημειωθεί για την παραπάνω μεθοδολογία, πως ο χρόνος ανάλυσης είναι πλήρως εξαρτημένος από τον ρυθμό καταγραφής στιγμιότυπων της κάμερας ή του βίντεο (στιγμιότυπα ανά δευτερόλεπτο - frames/second-fps), καθώς όσο υψηλότερος ο ρυθμός καταγραφής τόσο μεγαλύτερο το πλήθος των εικόνων που θα πρέπει να αναλυθούν από το πρόγραμμα άρα μεγαλύτερη διάρκεια ανάλυσης. Στην περίπτωση που ένα μοντέλο από αυτήν την μεθοδολογία δεν είναι ικανό να αναλύσει το βίντεο σε ικανοποιητικό χρόνο λόγω πολύ υψηλού ρυθμού καταγραφής, αξίζει να αναφερθεί η ικανότητα επιλογής ανάλυσης βίντεο ανά έναν αριθμό στιγμιότυπων, για παράδειγμα να γίνεται πρόβλεψη ανά 1,2,3,… στιγμιότυπα, κάτι που θα μειώσει ενδεχομένως ικανοποιητικά την απαιτούμενη διάρκεια ανάλυσης.

5.2.5 - Εφαρμογή Διαβιβαστικής Μάθησης

Μια εναλλακτική μέθοδος εκπαίδευσης Συνελικτικών Νευρωνικών Δικτύων είναι η μέθοδος Διαβιβαστικής Μάθησης (Transfer Learning), με την χρήση προ-εκπαιδευμένων μοντέλων (Pre-Trained Models).

Τα προ-εκπαιδευμένα μοντέλα για κατηγοριοποίηση εικόνων (Image Classification), έχουν δημιουργηθεί και εκπαιδευτεί σε μεγάλες βάσεις δεδομένων από εταιρίες όπως η Google και πανεπιστήμια. Υπάρχουν μοντέλα για διάφορες χρήσεις, όπως αργά, μεγάλα μοντέλα που σκοπός τους είναι υψηλή ακρίβεια κατηγοριοποίησης, και υπολογιστικά ελαφριά, γρήγορα μοντέλα που θυσιάζουν την ακρίβεια για γρήγορη και εύκολη εφαρμογή τους, ακόμα και σε συστήματα με μικρή υπολογιστική ισχύι (Πίνακας 5-6).

Πίνακας 5-6: Προ-Εκπαιδευμένα Μοντέλα που Υποστηρίζονται από το TensorFlow με τις Καταγεγραμμένες Ακρίβειες τους (Google-Tensorflow, 2019)

Model	TF-Slim File	Checkpoint	Top-1 Accuracy	Top-5 Accuracy
Inception V1	Code	inception_v1_2016_08_28.tar.gz	69.8	89.6
Inception V2	Code	inception_v2_2016_08_28.tar.gz	73.9	91.8
Inception V3	Code	inception_v3_2016_08_28.tar.gz	78.0	93.9
Inception V4	Code	inception_v4_2016_09_09.tar.gz	80.2	95.2
Inception-ResNet-v2	Code	inception_resnet_v2_2016_08_30.tar.gz	80.4	95.3
ResNet V1 50	Code	resnet_v1_50_2016_08_28.tar.gz	75.2	92.2
ResNet V1 101	Code	resnet_v1_101_2016_08_28.tar.gz	76.4	92.9
ResNet V1 152	Code	resnet_v1_152_2016_08_28.tar.gz	76.8	93.2
ResNet V2 50^	Code	resnet_v2_50_2017_04_14.tar.gz	75.6	92.8
ResNet V2 101^	Code	resnet_v2_101_2017_04_14.tar.gz	77.0	93.7
ResNet V2 152^	Code	resnet_v2_152_2017_04_14.tar.gz	77.8	94.1
ResNet V2 200	Code	TBA	79.9*	95.2*
VGG 16	Code	vgg_16_2016_08_28.tar.gz	71.5	89.8
VGG 19	Code	vgg_19_2016_08_28.tar.gz	71.1	89.8
MobileNet_v1_1.0_224	Code	mobilenet_v1_1.0_224.tgz	70.9	89.9
MobileNet_v1_0.50_160	Code	mobilenet_v1_0.50_160.tgz	59.1	81.9
MobileNet_v1_0.25_128	Code	mobilenet_v1_0.25_128.tgz	41.5	66.3
MobileNet_v2_1.4_224^*	Code	mobilenet_v2_1.4_224.tgz	74.9	92.5
MobileNet_v2_1.0_224^*	Code	mobilenet_v2_1.0_224.tgz	71.9	91.0
NASNet-A_Mobile_224#	Code	nasnet-a_mobile_04_10_2017.tar.gz	74.0	91.6
NASNet-A_Large_331#	Code	nasnet-a_large_04_10_2017.tar.gz	82.7	96.2
PNASNet-5_Large_331	Code	pnasnet-5_large_2017_12_13.tar.gz	82.9	96.2
PNASNet-5_Mobile_224	Code	pnasnet-5_mobile_2017_12_13.tar.gz	74.2	91.9

Τα παραπάνω μοντέλα μπορούν να αξιοποιηθούν με πολλούς τρόπους, εκ των οποίων ένας είναι η εφαρμογή Transfer Learning όπως αναφέρθηκε προηγουμένως. Με την εφαρμογή αυτής της μεθόδου αξιοποιούνται τα μεγάλα νευρωνικά δίκτυα που έχουν ήδη εκπαιδευτεί σε πολύ μεγάλες βάσεις δεδομένων και χρησιμοποιούνται για να πρόβλεψή-κατηγοριοποίηση εικόνων σε νέες κατηγορίες που δεν εμπεριέχονται στο αρχικό μοντέλο, ανάλογα με τις απαιτήσεις του νέου πειράματος.

Αυτό πραγματοποιείται, αρχικά, παγώνοντας τον κύριο σκελετό του μοντέλου, μαζί με τα βάρη και τις προκαταλήψεις του, εκτός από το τελευταίο επίπεδο που μετατρέπεται όπως απαιτεί το νέο πείραμα και στην συνέχεια εκπαιδεύοντας το νέο δίκτυο στην βάση δεδομένων του πειράματος. Κατά την εκπαίδευση αλλάζουν μόνο τα βάρη και οι προκαταλήψεις που συνδέονται με το τελευταίο επίπεδο που δίνει το αποτέλεσμα. Αυτή η μέθοδος μπορεί να λειτουργήσει γιατί το δίκτυο έχει πολλά συνελικτικά επίπεδα που έχουν ήδη εκπαιδευτεί σε πολύ μεγάλη βάση δεδομένων, και ξέρουν ήδη να αναγνωρίζουν μοτίβα από πολύ απλά έως πολύ περίπλοκα, και η αλλαγή του τελευταίου επιπέδου ουσιαστικά αλλάζει τον τρόπο ερμήνευσης αυτών για να εφαρμόζονται στην νέα βάση δεδομένων.

Να σημειωθεί πως η μέθοδος αυτή έχει το σημαντικό πλεονέκτημα πως μπορεί σε πολλές περιπτώσεις να αποφέρει καλά αποτελέσματα και με μικρές βάσεις δεδομένων και σε λίγες εποχές. Επίσης κάθε μοντέλο είναι απόλυτο ως προς την μορφή που θα έχουν τα δεδομένα εκπαίδευσης και αξιολόγησης όσον αφορά τις διαστάσεις και την χρωματική κλίμακα.

Εφαρμογή Διαβιβαστικής Μάθησης με το Μοντέλο ResNet50

Στο πείραμα της πρώτης μεθοδολογίας θα γίνει δοκιμή με Transfer Learning με χρήση του μοντέλου ResNet50 για σύγκριση με το προηγουμένως επιλεγμένο βέλτιστο μοντέλο.

Για να προσαρμοστεί το μοντέλο στο πείραμα, θα παγώσει το μοντέλο με τα βάρη και τις προκαταλήψεις του, εκτός από το τελευταίο επίπεδο που θα μετατραπεί σε ένα συμπαγές επίπεδο με δύο κόμβους, έναν για κάθε κατηγορία. Στην συνέχεια θα εκπαιδευτεί στην βάση δεδομένων του πειράματος για 5 εποχές και θα αξιολογηθεί στην τρίτη βάση δεδομένων με υπολογισμό των στατιστικών μεγεθών των αποτελεσμάτων, τα οποία παρουσιάζονται στην συνέχεια (Πίνακας 5-7).

Τα αποτελέσματα και τα στατιστικά μεγέθη, είναι ικανοποιητικά, και δείχνουν τα θετικά της χρήσης της μεθόδου, καθώς με περιορισμένη βάση δεδομένων και λίγες εποχές τα αποτελέσματα είναι καλύτερα από αρκετά μοντέλα που εξετάστηκαν προηγουμένως. Επίσης όμως παρουσιάζονται και τα αρνητικά της μεθόδου, καθώς αν εφαρμόσουμε το μοντέλο στο βίντεο της προηγούμενης ενότητας τα αποτελέσματα δεν είναι ικανοποιητικά και λόγω πιο περίπλοκης δομής του προ-εκπαιδευμένου μοντέλου η ανάλυση του βίντεο γίνεται σημαντικά πιο αργά από το βέλτιστο μοντέλο που επιλέχτηκε στην προηγουμένη ενότητα, κάτι που δεν ενδείκνυται για την εξεταζόμενη προσέγγιση.

Πίνακας 5-7: Στατιστικά Αποτελέσματα Μοντέλου από Διαβιβαστική Μάθηση με Βάση το ResNet50

Μητρώο Προβλέψεων/Πραγματικότητας

Συνολικά Δεδομένα Αξιολόγησης - 2000		Πραγματικές Συνθήκες	
		0	1
Προβλεπόμενες Συνθήκες	0	896	104
	1	181	819

Κατηγορία 0	Χαμήλη/Καθόλου Κυκλοφορία
Κατηγορία 1	Υψηλή Κυκλοφορία

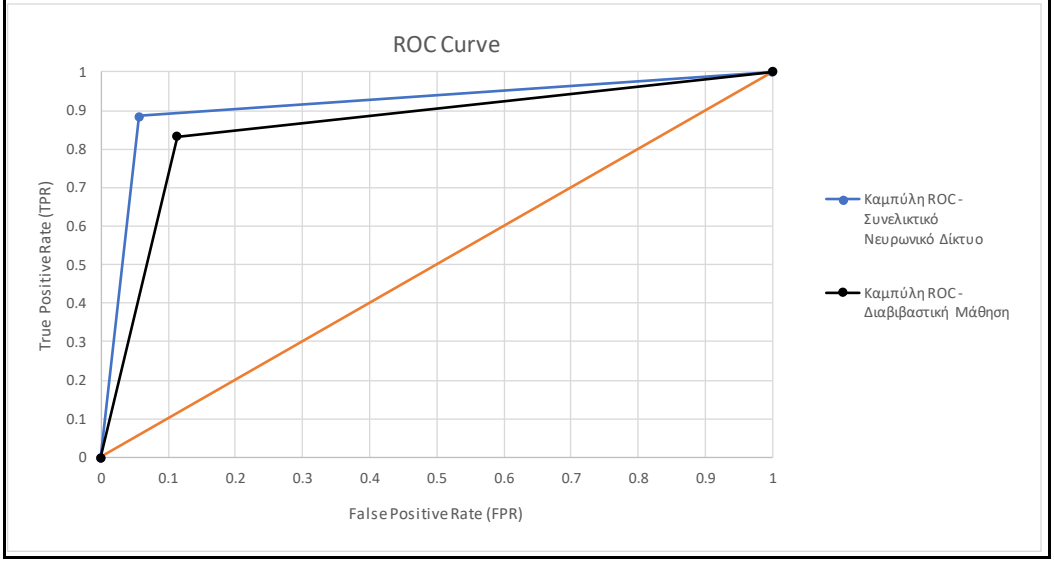
True Positive	896	(TP)
False Positive	104	(FP)
True Negative	819	(TN)
False Negative	181	(FN)

Αποτελέσματα

Ευαισθησία-Sensitivity	**0.832**	(TPR)
Εξειδίκευση-Specificity	**0.887**	(TNR)
Ακρίβεια-Precision	**0.896**	(PPV)
False Negative Rate - Error Type II	**0.168**	(FNR)
False Positive Rate - Error Type I	**0.113**	(FPR)

Ορθότητα-Accuracy	**0.858**	(ACC)
Αρμονικό Μέσο-F1 Score	**0.863**	(F1)

Area Under Curve	**0.860**	(AUC)



5.3 - Μακροσκοπική Αναγνώριση Κυκλοφορικών Συνθηκών

5.3.1 - Περιγραφή

Στην δεύτερη μεθοδολογία που εξετάστηκε, θα αξιοποιηθούν οι επεξεργασμένες εικόνες στις οποίες έχουν εντοπιστεί και καταγραφεί θέσεις και τύποι οχημάτων. Στην συνέχεια θα εκπαιδευτεί με αυτά τα δεδομένα το προ-εκπαιδευμένο, στην βάση δεδομένων COCO της Microsoft, SSD MobilNet Version 2 συνελικτικό νευρωνικό δίκτυο.

Σύμφωνα με τη βιβλιογραφία, το **SSD MobilNet Version 2** είναι ικανοποιητικά ακριβές, γρήγορο, καθώς και υπολογιστικά μη απαιτητικό σε σχέση με πιο ακριβή δίκτυα πχ RCNN_ResNet, Faster_RCNN.

Αξίζει να αναφερθεί πως το μοντέλο MobilNet είναι δομημένο για χρήση σε συσκευές με μικρή υπολογιστική ισχύι όπως συσκευές Android.

Θα εξεταστεί η ακρίβεια πρόβλεψης και κατηγοριοποίησης του μοντέλου και τελικά θα εφαρμοστεί σε βίντεο για να αξιολογηθεί η απόδοση του στα δεδομένα της μορφής του πειράματος. Τέλος θα δημιουργηθεί πρόγραμμα που να υπολογίζει κυκλοφοριακά μεγέθη με βάση τις προβλέψεις του μοντέλου σε εξεταζόμενο οδικό τμήμα.

5.3.2 - Εκπαίδευση Προ-Εκπαιδευμένου Δικτύου

Αρχικά θα μετατραπούν τα δεδομένα σε μορφή που μπορούν να αναγνωριστούν από το TensorFlow. Στην συνέχεια θα οριστούν παράμετροι στο δίκτυο SSD MobilNet Version 2 για να προσαρμοστεί στο πείραμα. Αρχικά θα εντοπιστούν οι λεζάντες-τίτλοι πιθανών προβλέψεων και εντοπισμού αντικειμένων από την βάση δεδομένων COCO, στην οποία έχει προ-εκπαιδευτεί το δίκτυο, και θα αντικατασταθούν με της λεζάντες των αντικειμένων που υπάρχουν στα δεδομένα του πειράματος που είναι τα εξής:

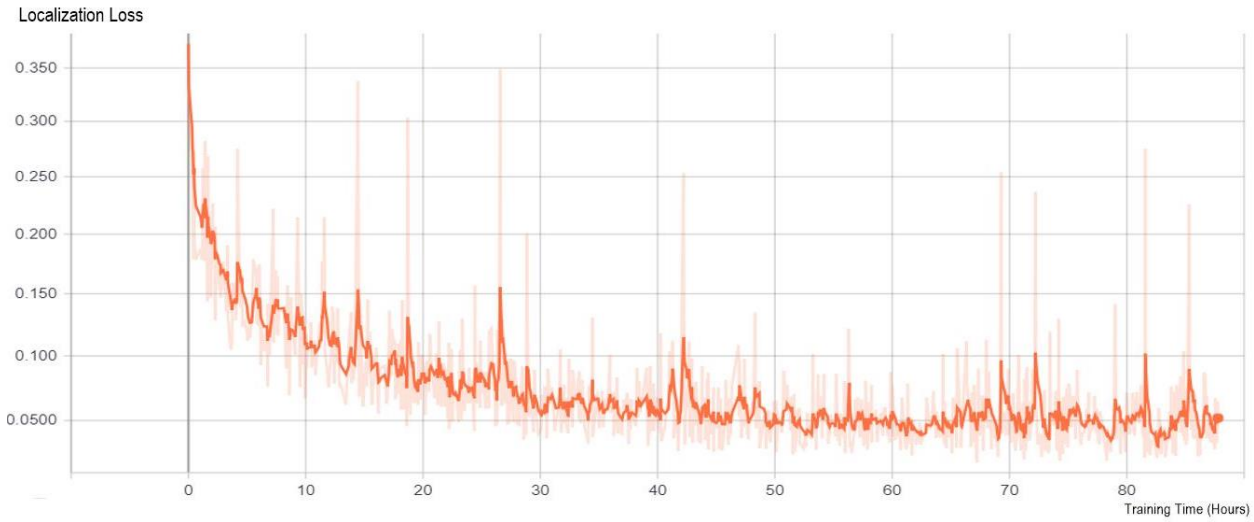
- Αμάξι (Car)
- Φορτηγάκι-Βανάκι (Van)
- Φορτηγό-Λεωφορείο (Truck)
- Μοτοσυκλέτα (Bike)

Τέλος θα οριστούν παράμετροι σχετικά με τον ρυθμό εκπαίδευσής μέγεθος συνόλων εκπαίδευσης, και την παραμόρφωση των δεδομένων ώστε να αξιοποιηθεί σε βέλτιστο βαθμό ο υπολογιστικός εξοπλισμός (Κάρτα Γραφικών – GPU) για ικανοποιητικά αποτελέσματα.

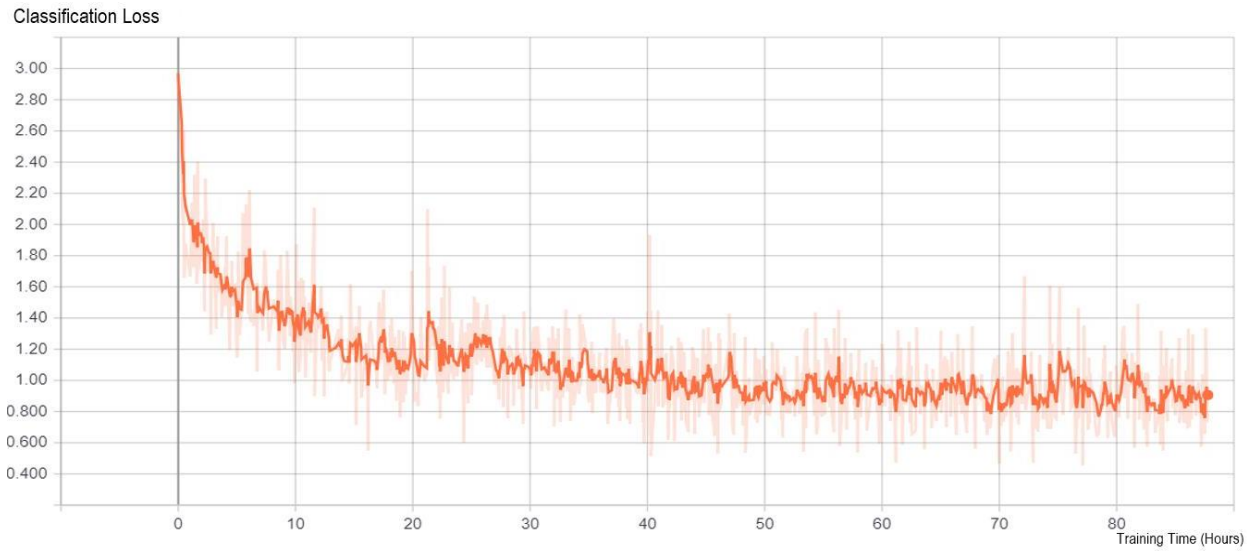
Εφόσον οριστούν οι παράμετροι, θα ξεκινήσει η εκπαίδευση του δικτύου, η οποία διαρκεί πολλές ώρες, καθώς τα πλήθος των δεδομένων είναι μεγάλο και το δίκτυο είναι αυξημένης περιπλοκότητας. Επιπλέον για την κατηγοριοποίηση των αντικειμένων πρέπει να εντοπιστεί και η θέση τους στην εικόνα επομένως και η διαδικασία εκπαίδευσης είναι επίσης αυξημένης πολυπλοκότητας.

Η εκπαίδευση θα διακοπεί όταν οι συνολικές απώλειες στο δείγμα αξιολόγησης προσεγγίσουν και σταθεροποιηθούν προς την τιμή ένα (1). Οι **συνολικές απώλειες** είναι ο συνδυασμός των **απωλειών εντοπισμού αντικειμένων** (localization loss) και των **απωλειών κατηγοριοποίησης τους** (classification loss).

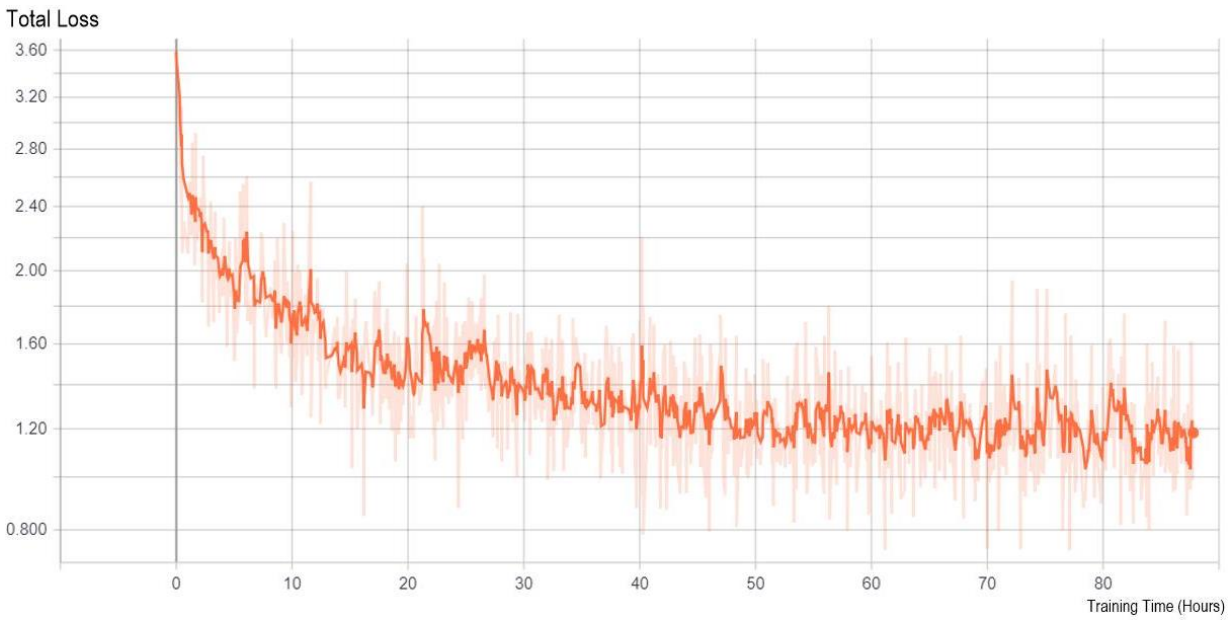
Στα διαγράμματα που ακολουθούν (Διάγραμμα 5-10, Διάγραμμα 5-11, Διάγραμμα 5-12) στον **άξονα των x ορίζεται η διάρκεια**, σε ώρες, της εκπαίδευσης και στον **άξονα y οι τιμές απωλειών** του κάθε διαγράμματος. Η εκπαίδευση διακόπηκε μετά από 87 ώρες (3 ημέρες και 15 ώρες), όπου η τιμή των απωλειών έχει σταθεροποιηθεί κοντά στην τιμή του ένα (Πίνακας 5-8).



Διάγραμμα 5-10: Διάγραμμα Απωλειών Εντοπισμού Οχημάτων κατά την Αξιολόγηση (Localization Loss)



Διάγραμμα 5-11: Διάγραμμα Απωλειών Κατηγοριοποίησης Οχημάτων κατά την Αξιολόγηση (Classification Loss)



Διάγραμμα 5-12: Διάγραμμα Συνολικών Απωλειών κατά την Αξιολόγηση (Total Loss)

Πίνακας 5-8: Χαρακτηριστικά Εκπαίδευσης Δικτύου 2ης Μεθόδου

Εφαρμογή Προ-Εκπαιδευμένου Συνελικτικού Νευρωνικού Δικτύου		
Μοντέλο	ssd_mobilnet_v2_coco	
Μέσο Εικπαίδευσης και Υπολογισμών	GPU	
Πλήθος Δεδομένων Εκπαίδευσης	3600	(θέσεις και κατηγορίες οχημάτων σε εικόνες)
Βήματα εκπαίδευσης (steps)	400000	
Διάρκεια Εικπαίδευσης Δικτύου	87	(ώρες - 3 ημέρες και 15 ώρες)

5.3.3 - Αποτελέσματα Εντοπισμού και Κατηγοριοποίησης Οχημάτων

Πραγματοποιείτε συλλογή 80 εικόνων, που δεν έχουν αξιοποιηθεί κατά την εκπαίδευση του δικτύου με διαφορετικά χαρακτηριστικά η καθεμία, όπως πλήθος οχημάτων, γωνία λήψης και συνθήκες φωτισμού (σκιές). Αυτές οι εικόνες περιέχουν συνολικά 400 οχήματα που εμφανίζονται ή ολόκληρά στην εικόνα ή στο μεγαλύτερο τμήμα τους. Εξετάζεται η ακρίβεια εντοπισμού και σωστής κατηγοριοποίησης των οχημάτων. Σημειώνεται πως οι μοτοσυκλέτες δεν συμπεριλαμβάνονται την διαδικασία παρά την δυνατότητα εντοπισμού τους.

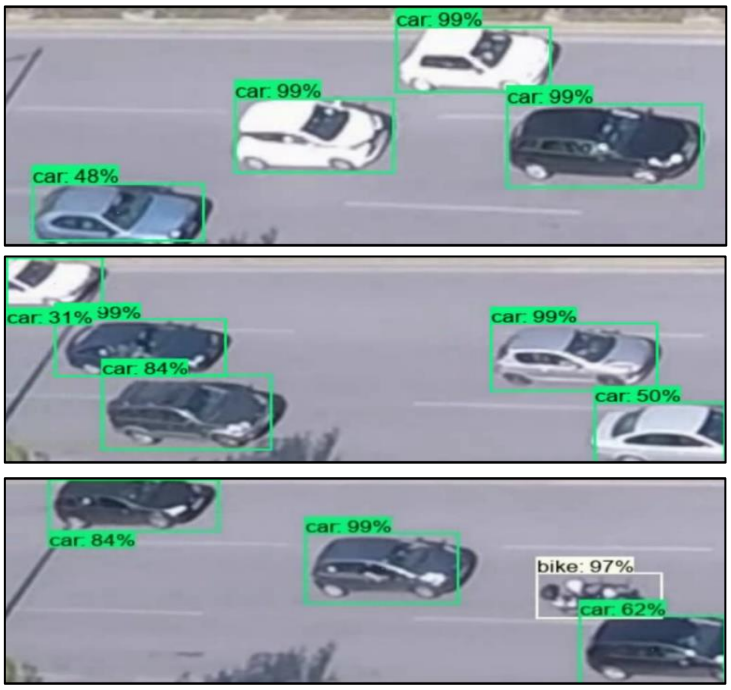
Στον επόμενο πίνακα (**Πίνακας 5-9**) παρουσιάζονται τα **αποτελέσματα του εντοπισμού/κατηγοριοποίησης** όπου TP είναι οι περιπτώσεις σωστού εντοπισμού και κατηγοριοποίησης των εξεταζόμενων οχημάτων, FP είναι οι περιπτώσεις σωστού εντοπισμού οχημάτων αλλά λάθος κατηγοριοποίηση τους ή λάθος εντοπισμοί και τέλος **FN** είναι το σύνολο των οχημάτων που δεν εντοπίστηκαν επομένων δεν κατηγοριοποιήθηκαν.

Πίνακας 5-9: Αποτελέσματα Εντοπισμού/Κατηγοριοποίησης

Αποτελέσματα Εντοπισμού/Κατηγοριοποίησης		
Συνολικά Οχήματα		**400**
TP		323
FP		3
FN		77
Ακρίβεια-Precision (PPV)		0.991
Ευαισθησία-Sensitivity (TPR)		0.808
Αρμονικό Μέσο-F1 Score		0.890
Μέση Χρονική Διάρκεια Ανάλυσης Εικόνας		
Μέσο Επεξεργασίας	Δευτερόλεπτα (sec)	(ms)
CPU	0.044	44.2
GPU	0.026	26.0

Εφαρμογή Νευρωνικού Δικτύου σε Βίντεο για Οπτικοποίηση Αποτελεσμάτων

Το επιλεγμένο μοντέλο εφαρμόζεται στην συνέχεια για οπτικοποίηση των αποτελεσμάτων. Τα εντοπισμένα οχήματα περικλείονται από ένα ορθογώνιο κουτί διαφορετικού χρώματος ανά κατηγορία, όπου το κουτί περιέχει την περιοχή στην εικόνα που το αντικείμενο έδωσε το υψηλότερο ποσοστό ακρίβειας πρόβλεψης κατά την ανάλυση κάθε εικόνας (Εικόνα 5-13).



Εικόνα 5-13: Ενδεικτικά Στιγμιότυπα Εντοπισμού και Κατηγοριοποίησης Οχημάτων σε Βίντεο με Χρήση του Προ-Εκπαιδευμένου Δικτύου

Υπολογισμός Μακροσκοπικών Μεγεθών Κυκλοφορίας σε Βίντεο

Για τον υπολογισμό κυκλοφοριακών μεγεθών, θα δημιουργηθεί πρόγραμμα με δεδομένα εισόδου:

- Το βίντεο προς εξέταση.
- Το προ-εκπαιδευμένο μοντέλο θα εφαρμόζεται για τις προβλέψεις.
- Το χρονικό διάστημα ανά το οποίο θα υπολογίζονται τα μεγέθη.

Το πρόγραμμα με την παραδοχή ότι το μέσο μήκος αυτοκινήτου είναι 4.5 μέτρα και αγνοώντας τις μοτοσυκλέτες στις προβλέψεις, αξιοποιεί τις τοποθεσίες των οχημάτων που παράγονται από το δίκτυο σε κάθε εικόνα και εντοπίζει τα οχήματα από εικόνα σε εικόνα και υπολογίζει το πλήθος τους, την ταχύτητα τους, το μήκος της οδού, τον φόρτο, την πυκνότητα, και την κατάληψη για την διάρκεια κάθε χρονικής περιόδου στο εξεταζόμενο οδικό τμήμα (Εικόνα 5-14).

Για αυτούς τους υπολογισμούς ορίζεται μία θεωρητική διατομή της οδού στο μέσο της εικόνας. Επίσης απαιτείται τουλαχιστον ένα μήκος αναφοράς επειδή οι διαστάσεις του οδικού τμήματος δεν θα είναι σταθερές σε κάθε λήψη. Για την καλύτερη γενίκευση του προγράμματος αυτό το μήκος είναι το μέσο μήκος αυτοκινήτου το οποίο αντιστοιχίζεται στην μέση οριζόντια διάσταση σε πίξελ των πλαισίων πρόβλεψης των οχημάτων κατηγορίας αυτοκίνητο με αυτοπεποίηθηση πρόβλεψης πάνω από 90%. Το ποσοστό επιλέγεται γιατί επιθυμούμε το μήκος των πλαισίων όταν το όχημα βρίσκεται ολόκληρο εντος της εικόνας. Επίσης γίνεται μία μικρή διόρθωση στην αντιστοίχηση των διαστάσεων λόγω γωνίας λήψης της κάμερας και την απόκλησης των πλαισίων από τα οχήματα στις εικόνες.

Εικόνα 5-14: Ενδεικτικά Στιγμιότυπα Εντοπισμού και Κατηγοριοποίησης Οχημάτων σε Βίντεο με Χρήση του Προ-Εκπαιδευμένου Δικτύου και Υπολογισμού Κυκλοφοριακών Μεγεθών.

Στο τέλος του βίντεο αποθηκεύονται τα αποτελέσματα κάθε χρονικής περιόδου και θα σχεδιαστούν τα θεμελιώδη διαγράμματα, φόρτου-ταχύτητας, πυκνότητας- ταχύτητας, πυκνότητας- φόρτου, αυτών.

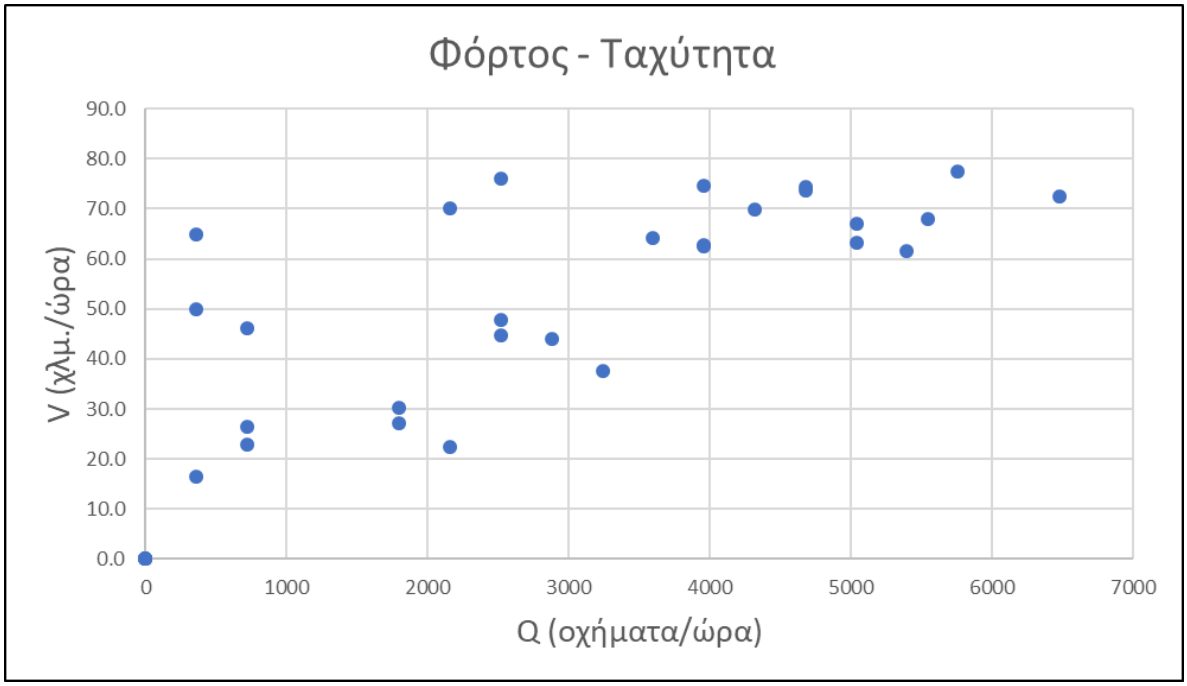
Για το ενδεικτικό βίντεο διάρκειας 5 λεπτών και 38 δευτερόλεπτων, γίνονται υπολογισμοί κάθε 10 δευτερόλεπτα. Στον **Πίνακας 5-10** φαίνονται τα **αποτελέσματα του υπολογισμού κυκλοφοριακών μεγεθών** και στο **Διάγραμμα 5-13, Διάγραμμα 5-14 και Διάγραμμα 5-15** οι σχέσεις φόρτου-ταχύτητας, ταχύτητας-πυκνότητας και πυκνότητας-φόρτου αντίστοιχα.

Πίνακας 5-10: Αποτελέσματα-Κυκλοφοριακά Μεγέθη ανά Χρονική Περίοδο στο Ενδεικτικό Βίντεο

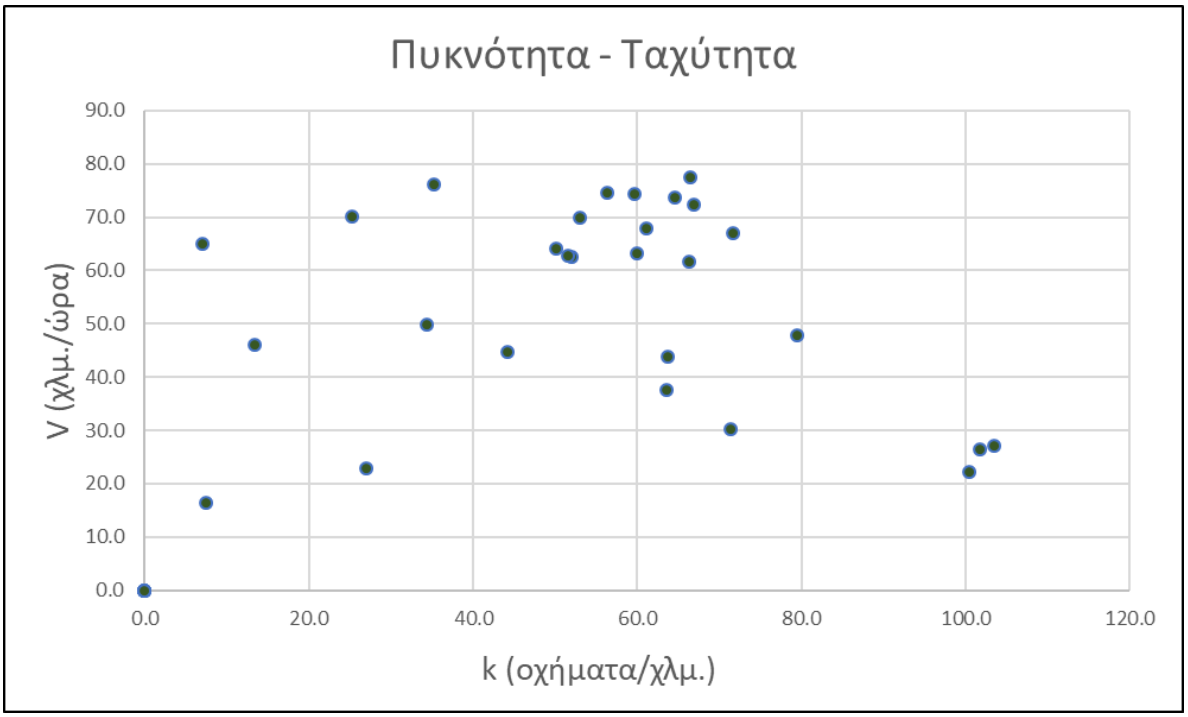
Συνολική Διάρκεια Βίντεο	5,38	(λεπτά - 338 δευτερόλεπτα)

t -Διάρκεια Περιόδου	10	(δευτερόλεπτα)

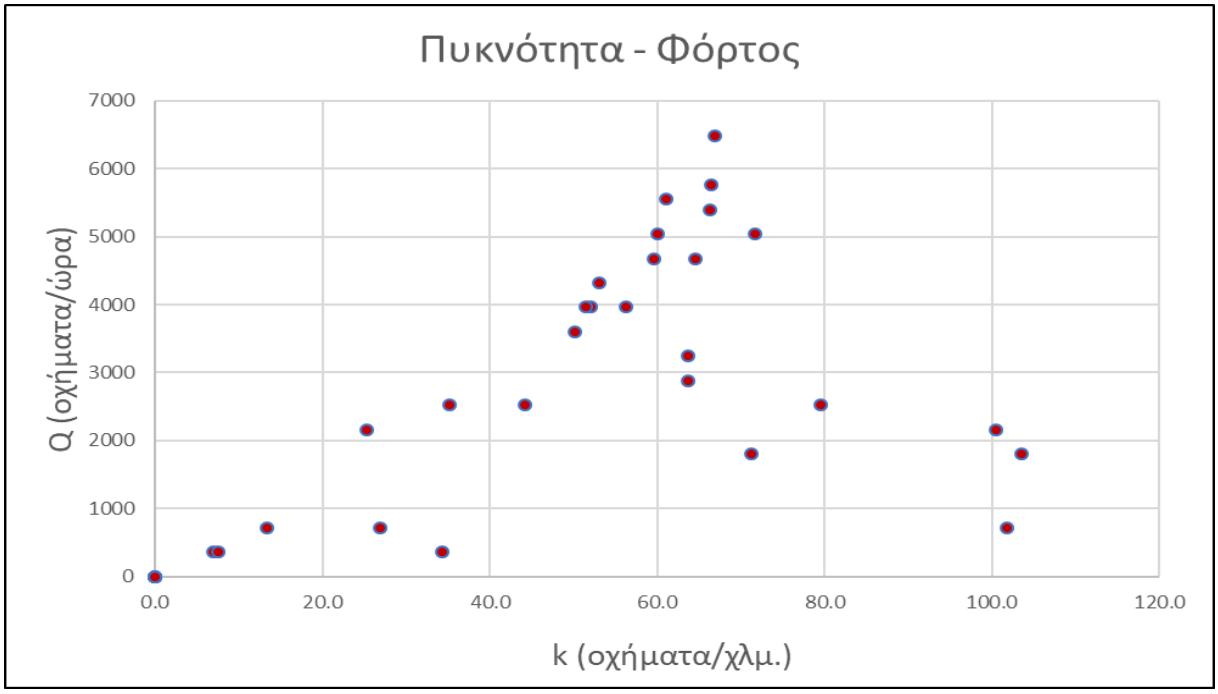
Περίοδοι	Μέθοδος 2 - Μακροσκοπική Αναγνώριση Κυκλοφορικών Συνθηκών			
(10 δευτερολεπτων)	Φόρτος	Μέση Ταχύτητα	Πυκνότητα	Κατάληψη
	Q (οχήματα/ώρα)	V (χλμ./ώρα)	k (οχήματα/χλμ.)	O
1	2520	44.7	44.2	0.26
2	0	0.0	0.0	0.00
3	360	64.9	7.0	0.02
4	4680	73.6	64.6	0.30
5	5760	77.5	66.4	0.36
6	5040	63.2	60.0	0.37
7	5400	61.6	66.3	0.42
8	2880	43.9	63.7	0.34
9	720	22.9	26.9	0.14
10	0	0.0	0.0	0.00
11	0	0.0	0.0	0.00
12	0	0.0	0.0	0.00
13	3960	62.5	52.0	0.22
14	3960	74.5	56.3	0.26
15	4680	74.4	59.6	0.30
16	3600	64.2	50.1	0.27
17	3240	37.6	63.6	0.44
18	2520	47.9	79.5	0.31
19	2160	22.3	100.5	0.50
20	720	46.1	13.4	0.15
21	0	0.0	0.0	0.00
22	3960	62.7	51.5	0.22
23	5550	67.9	61.1	0.28
24	4320	69.9	53.0	0.28
25	6480	72.4	66.9	0.42
26	1800	30.3	71.3	0.30
27	720	26.4	101.8	0.16
28	1800	27.2	103.5	0.55
29	360	16.5	7.5	0.10
30	0	0.0	0.0	0.00
31	2160	70.2	25.3	0.14
32	5040	67.1	71.7	0.35
33	2520	76.1	35.2	0.16
34	360	49.9	34.3	0.01



Διάγραμμα 5-13: Ενδεικτικό Διάγραμμα Φόρτου-Ταχύτητας από Βίντεο



Διάγραμμα 5-14: Ενδεικτικό Διάγραμμα Πυκνότητάς-Ταχύτητας από Βίντεο



Διάγραμμα 5-15: Ενδεικτικό Διάγραμμα Πυκνότητας-Φόρτου από Βίντεο

Στην συνέχεια παρουσιάζεται η χρονική απαίτηση για την ανάλυση του ενδεικτικού βίντεο στον επεξεργαστή και στην κάρτα γραφικών και σύγκριση των δύο (Πίνακας 5-11).

Πίνακας 5-11: Χαρακτηριστικά Υπολογιστικές Επιδώσεις Εφαρμογής Δικτύου 2ης Μεθόδου στο Ενδεικτικό Βίντεο

Εφαρμογή Προ-Εκπαιδευμένου Συνελικτικού Νευρωνικού Δικτύου			
Μοντέλο	ssd_mobilnet_v2_coco		
Μέσο Επεξεργασίας και Υπολογισμών	GPU	CPU	
Διάρκεια Εξεταζόμενου Βίντεο	5.64		(λεπτα)
Διάρκεια Ανάλυσης/Πρόβλεψης Βίντεο	3.88	6.11	(λεπτα)
Μέση Διάρκεια Ανάλυσης Εικόνας	26.0	44.2	(ms)
Διάρκεια Ανάλυσης σε Σχέση με το Εξεταζόμενο Βίντεο	-31%	8%	(% Μείωση)
Σχέση Διάρκειας Ανάλυσης GPU με CPU	-36%		(% Μείωση)

Παρατηρείται πως η ανάλυση με την κάρτα γραφικών γίνεται πιο γρήγορα από την διάρκεια του βίντεο, επομένως με αυτόν τον εξοπλισμό είναι δυνατή η ανάλυση από βίντεο και ο υπολογισμός μεγεθών σε πραγματικό χρόνο. Σε αντίθεση η ανάλυση με χρήση επεξεργαστή είναι πιο χρονοβόρα επομένως δεν είναι εφικτή η ανάλυση κάθε εικόνας του βίντεο σε πραγματικό χρόνο. Αξίζει να αναφερθεί ότι το MobilNet είναι φτιαγμένο για χρήση σε συσκευές με μικρή υπολογιστική ισχύι όπως συσκευές Android.

5.4 - Σύγκριση Προτεινόμενων Προσεγγίσεων

Στη πρώτη μεθοδολογία αξιολογήθηκαν δύο περιπτώσεις, του βέλτιστου απλού συνελικτικού νευρωνικού δικτύου και του προ-εκπαιδευμένου δικτύου ResNet50. Όπως παρουσιάζεται και στον επόμενο πίνακα (Πίνακας 5-12) τα αποτελέσματα του απλού συνελικτικού νευρωνικού δικτύου είναι καλύτερα, επομένως εξετάζεται περεταίρω μόνο το πρώτο.

Πίνακας 5-12: Συγκριτικά Αποτελέσματα 1ης Μεθοδολογίας

Συνελικτικό Νευρωνικό Δίκτυο			Διαβιβαστική Μάθηση - Προ-Εκπαιδευμένο Μοντέλο		
Ευαισθησία-Sensitivity	0.885	(TPR)	Ευαισθησία-Sensitivity	0.832	(TPR)
Εξειδίκευση-Specificity	0.943	(TNR)	Εξειδίκευση-Specificity	0.887	(TNR)
Ακρίβεια-Precision	0.947	(PPV)	Ακρίβεια-Precision	0.896	(PPV)
False Negative Rate - Error Type II	0.115	(FNR)	False Negative Rate - Error Type II	0.168	(FNR)
False Positive Rate - Error Type I	0.057	(FPR)	False Positive Rate - Error Type I	0.113	(FPR)
Ορθότητα-Accuracy	0.912	(ACC)	Ορθότητα-Accuracy	0.858	(ACC)
Αρμονικό Μέσο-F1 Score	0.915	(F1)	Αρμονικό Μέσο-F1 Score	0.863	(F1)
Area Under Curve	0.914	(AUC)	Area Under Curve	0.860	(AUC)

Οι δύο μεθοδολογίες εφαρμόστηκαν σε κοινό βίντεο, το οποίο δεν προσέφερε δεδομένα για την εκπαίδευση των νευρωνικών δικτύων της κάθε μεθοδολογίας. Έγινε επιλογή διάρκειας χρονικού διαστήματος 10 δευτερόλεπτα και στις δύο περιπτώσεις για γίνει σύγκριση των αποτελεσμάτων (Πίνακας 5-13).

Πίνακας 5-13: Αποτελέσματα των δύο Μεθόδων ανά Χρονική Περίοδο στο Ενδεικτικό Βίντεο

Συνολική Διάρκεια Βίντεο	5,38	(λεπτά - 338 δευτερόλεπτα)
t -Διάρκεια Περιόδου	10	(δευτερόλεπτα)
Κατηγορία 0	Χαμηλή/Καθόλου Κυκλοφορία	
Κατηγορία 1	Υψηλή Κυκλοφορία	

Περίοδοι	Μέθοδος 1 - Αναγνώριση Κυκλοφορικών Συνθηκών σε Πραγματικό Χρόνο	Μέθοδος 2 - Μακροσκοπική Αναγνώριση Κυκλοφορικών Συνθηκών				
(10 δευτερολεπτων)	Κατηγορία	Φόρτος Q (οχήματα/ώρα)	Μέση Ταχύτητα V (χλμ./ώρα)	Πυκνότητα k (οχήματα/χλμ.)	Κατάληψη O	
1	1	2520	44.7	44.2	0.26	
2	0	0	0.0	0.0	0.00	
3	0	360	64.9	7.0	0.02	
4	1	4680	73.6	64.6	0.30	
5	1	5760	77.5	66.4	0.36	
6	1	5040	63.2	60.0	0.37	
7	1	5400	61.6	66.3	0.42	
8	1	2880	43.9	63.7	0.34	
9	0	720	22.9	26.9	0.14	
10	0	0	0.0	0.0	0.00	
11	0	0	0.0	0.0	0.00	
12	0	0	0.0	0.0	0.00	
13	1	3960	62.5	52.0	0.22	
14	1	3960	74.5	56.3	0.26	
15	1	4680	74.4	59.6	0.30	
16	1	3600	64.2	50.1	0.27	
17	1	3240	37.6	63.6	0.44	
18	1	2520	47.9	79.5	0.31	
19	1	2160	22.3	100.5	0.50	
20	0	720	46.1	13.4	0.15	
21	0	0	0.0	0.0	0.00	
22	1	3960	62.7	51.5	0.22	
23	1	5550	67.9	61.1	0.28	
24	1	4320	69.9	53.0	0.28	
25	1	6480	72.4	66.9	0.42	
26	1	1800	30.3	71.3	0.30	
27	1	720	26.4	101.8	0.16	(Σταματημένα Οχήματα Λόγω Κόκκινου Σηματοδότη)
28	1	1800	27.2	103.5	0.55	
29	0	360	16.5	7.5	0.10	
30	0	0	0.0	0.0	0.00	
31	0	2160	70.2	25.3	0.14	
32	1	5040	67.1	71.7	0.35	
33	1	2520	76.1	35.2	0.16	
34	0	360	49.9	34.3	0.01	

Παρατηρείται πως για υπολογισμένες τιμές φόρτου, από την δεύτερη μεθοδολογία, άνω των 1000 οχημάτων ανά ώρα στο συγκεκριμένο βίντεο και στην συγκεκριμένη τοποθεσία, η πρώτη μεθοδολογία κάνει πρόβλεψη, στο ίδιο χρονικό διάστημα, Υψηλής Κυκλοφορίας.

Οι τιμές που υπολογίζονται είναι εντός λογικών πλαισίων και αντιστοιχούν στην πραγματική κατάσταση του βίντεο, με εξαίρεση την 27 και 28 περίοδο στην οποία είναι σταματημένα τα οχήματα στο βίντεο λόγω κόκκινου σηματοδότη όπου η πυκνότητα παρουσιάζει αυξημένες τιμές.

Τελικά βγαίνει το συμπέρασμα πως οι δύο μέθοδοι μπορούν να εφαρμοστούν σε βίντεο με ικανοποιητικά αποτελέσματα, λαμβάνοντας υπόψη την έλλειψη διαφορετικών συνθηκών και περιβάλλοντος στα δεδομένα.

Στον παρακάτω πίνακα (Πίνακας 5-14) παρουσιάζονται οι χρονικές απαιτήσεις για εκπαίδευση και εφαρμογή των συνελικτικών νευρωνικών δικτύων των δυο μεθοδολογιών.

Πίνακας 5-14: Σύγκριση Απαιτήσεων Υπολογιστικής Ισχύς των δύο Μεθόδων

Μέθοδος 1 - Αναγνώριση Κυκλοφορικών Συνθηκών σε Πραγματικό Χρόνο			
Εφαρμογή Συνελικτικού Νευρωνικού Δικτύου			
Μοντέλο	BN_3_CON-32_64_128_0NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-40_E-0.3_VS-1539690557		
Μέσο Επεξεργασίας και Υπολογισμών	GPU	CPU	
Διάρκεια Εξεταζόμενου Βίντεο	5.64		(λεπτα)
Διάρκεια Ανάλυσης/Πρόβλεψης Βίντεο	0.51	0.92	(λεπτα)
Μέση Διάρκεια Ανάλυσης Εικόνας	1.4	4.4	(ms)
Διάρκεια Ανάλυσης σε Σχέση με το Εξεταζόμενο Βίντεο	-91%	-84%	(% Μείωση)
Σχέση Διάρκειας Ανάλυσης GPU με CPU	-45%		(% Μείωση)

Εκπαίδευση Συνελικτικού Νευρωνικού Δικτύου		
Μοντέλο	BN_3_CON-32_64_128_0NODES-1_DEN-512_DSiz-0.75_DOPr-150_BS-40_E-0.3_VS-1539690557	
Μέσο Εικπαίδευσης και Υπολογισμών	GPU	
Πλήθος Δεδομένων Εκπαίδευσης	50000	(εικόνες)
Εποχές Εκπαίδευσης	40	
Μέγεθος Τμημάτων Εκπαίδευσης (Batch Size)	150	(εικόνες ανα τμήμα)
Διάρκεια Εικπαίδευσης Δικτύου	15.16	(λεπτα)

Μέθοδος 2 - Μακροσκοπική Αναγνώριση Κυκλοφορικών Συνθηκών			
Εφαρμογή Προ-Εκπαιδευμένου Συνελικτικού Νευρωνικού Δικτύου			
Μοντέλο	ssd_mobilnet_v2_coco		
Μέσο Επεξεργασίας και Υπολογισμών	GPU	CPU	
Διάρκεια Εξεταζόμενου Βίντεο	5.64		(λεπτα)
Διάρκεια Ανάλυσης/Πρόβλεψης Βίντεο	3.88	6.11	(λεπτα)
Μέση Διάρκεια Ανάλυσης Εικόνας	26.0	44.2	(ms)
Διάρκεια Ανάλυσης σε Σχέση με το Εξεταζόμενο Βίντεο	-31%	8%	(% Μείωση)
Σχέση Διάρκειας Ανάλυσης GPU με CPU	-36%		(% Μείωση)

Εφαρμογή Προ-Εκπαιδευμένου Συνελικτικού Νευρωνικού Δικτύου		
Μοντέλο	ssd_mobilnet_v2_coco	
Μέσο Εικπαίδευσης και Υπολογισμών	GPU	
Πλήθος Δεδομένων Εκπαίδευσης	3600	(θέσεις και κατηγορίες οχημάτων σε εικόνες)
Βήματα εκπαίδευσης (steps)	400000	
Διάρκεια Εικπαίδευσης Δικτύου	87	(ώρες - 3 ημέρες και 15 ώρες)

Το συμπέρασμα είναι πως σε εφαρμογές με χρήση του TensorFlow που μπορεί να αξιοποιήσει την υπολογιστική ισχύ κατάλληλων καρτών γραφικών, η αξιοποίηση κατάλληλου εξοπλισμού είναι χρονικά σημαντικά συμφέρων και σε πολλές περιπτώσεις απαραίτητη.

6 - ΣΥΜΠΕΡΑΣΜΑΤΑ

6.1 - Σύνοψη Μεθοδολογίας και Αποτελεσμάτων

Η ανάλυση εικόνας που προέρχεται από μη επανδρωμένο αεροσκάφος (Unmanned Aerial Vehicles – UAV) παρουσιάζει πολλές προκλήσεις όπως συμπεραίνεται στην βιβλιογραφία αλλά και κατά την διάρκεια της διπλωματικής εργασίας.

Οι υπάρχουσες εφαρμογές που αξιοποιούν UAV και εφαρμογές βαθιάς μάθησης, όπως παρουσιάζονται και στο κεφάλαιο της βιβλιογραφικής ανασκόπησης, είναι περιορισμένες και δεν επικεντρώνονται σε κυκλοφοριακά δεδομένα και φαινόμενα.

Η παρούσα εργασία αποσκοπεί στην ανάπτυξη και αξιολόγηση αξιόπιστων μεθοδολογικών πλαισίων για τον εντοπισμό κυκλοφοριακών συνθήκων από δεδομένα μορφής βίντεο καταγεγραμμένα από μη επανδρωμένο αεροσκάφος, που να βασίζονται σε μοντέλα βαθιάς μάθησης και υπολογιστικής όρασης. Στο πλαίσιο της εργασίας θα εξεταστούν δύο προσεγγίσεις που ακολουθούν διαφορετικές μεθοδολογίες οι οποίες διαφέρουν στην χρήση τους αλλά και στα χαρακτηριστικά τους. Η πρώτη προσέγγιση αποσκοπεί στην αξιολόγηση, σε πραγματικό χρόνο, κυκλοφοριακών συνθηκών ενός τμήματος οδού σε δύο κατηγορίες, με ελάχιστο υπολογιστικό κόστος. Ο σκοπός της δεύτερης προσέγγισης είναι να εντοπίσει, να κατηγοριοποιήσει τα οχήματα και να προσεγγίσει κυκλοφοριακά μεγέθη σε ένα βίντεο ενός οδικού τμήματος.

Το αρχικό στάδιο της εργασίας είναι η κατάλληλη επεξεργασία των δεδομένων για κάθε μία από τις προσεγγίσεις. Τα δεδομένα που αξιοποιήθηκαν είναι βίντεο της κυκλοφορίας επι την οδό Κατεχάκη, Αθήνα που έχουν καταγραφεί από μη επανδρωμένο αεροσκάφος.

Για την πρώτη προσέγγιση απομονώνεται ένα τμήμα από την οδό και εξάγονται εικόνες για δημιουργία βάσης δεδομένων. Έπειτα γίνεται διαχωρισμός των εικόνων σε δύο κατηγορίες, Χαμηλής/Καθόλου Κυκλοφορίας και Υψηλής Κυκλοφορίας με κριτήρια το πλήθος και θέση των οχημάτων στις εικόνες.

Για την δεύτερη μεθοδολογική προσέγγιση απομονώνονται επίσης επιθυμητά τμήματα της οδού από τα βίντεο. Στην συνέχεια, από τις παραγμένες εικόνες γίνεται επιλογή καταλλήλων και έπειτα με χρήση προγράμματος επισημαίνεται με κατάλληλη μορφή, σε κάθε εικόνα, η θέση και κατηγορία κάθε οχήματος εντός αυτής.

Στο δεύτερο στάδιο της εργασίας, εξετάζονται αρχικά διαφορετικές μεθοδολογίες για κάθε μια από τις προσεγγίσεις.

Για να επιτευχθεί ο στόχος της πρώτης προσέγγισης αναπτύσσονται, εκπαιδεύονται και αξιολογούνται πολλαπλά συνελικτικά νευρωνικά δίκτυα με διαφορετικές δομές. Επίσης εξετάζεται η χρήση προ-εκπαιδευμένων δικτύων με την τεχνική της διαβιβαστικής μάθησης. Η αναγνώριση της βέλτιστης δομής βασίζεται στην ελαχιστοποίηση του σφάλματος ανίχνευσης στο δείγμα αξιολόγησης. Το βέλτιστο συνελικτικό νευρωνικό δίκτυο επιτυγχάνει ακρίβεια 91% στο τελικό δείγμα αξιολόγησης. Επίσης το προ-εκπαιδευμένο μοντέλο επιτυγχάνει επίσης υψηλές ακρίβειες (89%) αλλά κρίνεται χειρότερο, λόγω υπολογιστικών απαιτήσεων, επομένως δεν εξετάζεται περαιτέρω. Τελικά το μοντέλο εφαρμόζεται σε βίντεο με ξένα δεδομένα για την αξιολόγηση της πραγματικής εφαρμογής του και των υπολογιστικών και χρονικών απαιτήσεων για την ανάλυση.

Για τις απαιτήσεις της δεύτερης προσέγγισης εκμεταλλεύονται προ-εκπαιδευμένα συνελικτικά νευρωνικά δίκτυα τροποποιημένα για χρήση με τα δεδομένα βίντεο από το μη επανδρωμένο αεροσκάφος. Το τελικό μοντέλο προβλέπει με ικανοποιητική ακρίβεια και περιορισμένες απώλειες, τις θέσεις και τις κατηγορίες των οχημάτων σε κάθε εικόνα (89% F1 Score). Τελικά εφαρμόζεται σε βίντεο στο οποίο εντοπίζονται τα οχήματα, και στην συνέχεια αξιοποιώντας τις προβλέψεις προσεγγίζονται κυκλοφοριακά μεγέθη για το εξεταζόμενο οδικό τμήμα.

Στην συνέχεια θα παρουσιαστούν τα βασικά συμπεράσματα για κάθε προσέγγιση μαζί με τους περιορισμούς και προβλήματα που εμφανίστηκαν, τα πλεονεκτήματα και προτεινόμενες εφαρμογές τους.

6.2 - Βασικά Συμπεράσματα

Στο πλαίσιο της διπλωματικής εργασίας αναλύθηκαν δύο διαφορετικές προσεγγίσεις με κοινά δεδομένα που προήλθαν από μη επανδρωμένο αεροσκάφος. Όλα τα βίντεο είναι από την οδό Κατεχάκη στο ίδιο ύψος της οδού και με παρόμοια λήψη στην άνοδο της οδού. Επομένως τα δεδομένα είναι παρόμοια μεταξύ τους, γεγονός που περιορίζει από την αρχή τον εκτενή έλεγχο και αξιολόγηση των μοντέλων. Πιο συγκεκριμένα το δείγμα αξιολόγησης σε κάθε προσέγγιση σχηματίστηκε από δεδομένα ξένα στα μοντέλα για πιο αντικειμενικά αποτελέσματα. Η αλήθεια είναι όμως πως τα δεδομένα δεν είναι πλήρως ξένα για το μοντέλο λόγω των αιτιών που αναφέρθηκαν προηγουμένως. Η ιδανική περίπτωση θα ήταν να υπήρχε μεγάλο πλήθος δεδομένων από διαφορετικούς τύπους οδών, σε διαφορετικό περιβάλλον με διαφορετικές λήψεις βίντεο και με διαφορετικές κυκλοφοριακές συνθήκες. Μια τέτοια ποικιλία δεδομένων θα παρήγαγε ένα μοντέλο με πολύ καλύτερη δυνατότητα γενίκευσης.

Το πρότυπο που αναπτύχθηκε στην πρώτη προσέγγιση χαρακτηρίζεται από πολύ υψηλές ακρίβειες όταν εξετάζεται στο δείγμα αξιολόγησης. Για τους λόγους που προαναφέρθηκαν τα αποτελέσματα δεν είναι πλήρως αντιπροσωπεύτηκα αλλά είναι μια καλή ένδειξη για την δυνατότητα εφαρμογής του προτύπου. Το συγκεκριμένο μοντέλο αναμένετε να είναι σημαντικά πιο ευαίσθητο σε ξένα δεδομένα. Επίσης ένας από τους κύριους στόχους είναι να εξεταστούν οι υπολογιστικές και χρονικές απαιτήσεις του μοντέλου, με σκοπό την βελτιστοποίηση αυτών καθώς απώτερος σκοπός είναι η ανάλυση της εικόνας σε πραγματικό χρόνο. Αυτό σημαίνει πως η ανάλυση πρέπει να πραγματοποιείτε με ρυθμό πιο γρήγορο από

τον ρυθμό καταγραφής στιγμιότυπων της κάμερας. Για να επιτευχθεί αυτό προτιμήθηκαν δομές μειωμένης περιπλοκότητας ώστε το μοντέλο να είναι λιγότερο υπολογιστικά απαιτητικό. Τα αποτελέσματα υποδεικνύουν πως με τον κατάλληλο εξοπλισμό η ανάλυση γίνεται έως και 90% πιο γρήγορα από τον ρυθμό ανανέωσης στιγμιότυπων του βίντεο, τιμές που ικανοποιούν τον επιθυμητό σκοπό. Η εφαρμογή της διαβιβαστικής μάθησης από προ-εκπαιδευμένο δίκτυο παρουσιάζει καλή ακρίβεια, αλλά οι υπολογιστικές και χρονικές απαιτήσεις του κατά την εκπαίδευση και αξιολόγηση είναι σημαντικά αυξημένες λόγω του μεγέθους και της πολυπλοκότητάς του δικτύου, επομένως δεν προτιμάται για τις συνθήκες της πρώτης προσέγγισης.

Οι περιορισμοί του πρώτου προτύπου είναι ότι τα δεδομένα εισόδου με μορφή βίντεο πρέπει να περιέχουν μόνο το εξεταζόμενο τμήμα της οδού με την καλύτερη δυνατή ορατότητα. Επίσης η μέθοδος μπορεί να επηρεαστεί σημαντικά από την ποιότητα του βίντεο. Η μη σταθερή λήψη προκαλεί σε πολλές περιπτώσεις εσφαλμένες ή τυχαίες προβλέψεις. Επίσης παρατηρήθηκε πως οι σκιασμένες περιοχές μπορούν να επηρεάσουν αρνητικά την ακρίβεια σε μερικές περιπτώσεις. Επιπλέον η μέθοδος εφαρμόζεται μόνο κατά την διάρκεια της ημέρας καθώς δεν εξετάστηκε περίπτωση νυχτερινών λήψεων λόγω συνθηκών φωτισμού. Τέλος να αναφερθεί πως σε περίπτωση σταματημένων οχημάτων στην οδό, για παράδειγμα σε κόκκινο σηματοδότη, το μοντέλο προβλέπει κατηγορία υψηλής κυκλοφορίας.

Τα πλεονεκτήματα του προτύπου είναι η χαμηλή υπολογιστική απαίτηση για εύκολη εφαρμογή του για τη γρήγορη αξιολόγηση της στιγμιαίας κατάστασης στο τμήμα της οδού. Επίσης το βίντεο δεν απαιτεί κάποια επεξεργασία για να αναλυθεί από το μοντέλο, παρόλο που η σταθεροποίηση του βίντεο αυξάνει την αξιοπιστία των αποτελεσμάτων. Επομένως το πρότυπο μπορεί να εφαρμοστεί και επι τόπου στην εξεταζόμενη τοποθεσία υπό τις κατάλληλες συνθήκες.

Το δεύτερο πρότυπο που εφαρμόζεται στο πλαίσιο της διπλωματικής εργασίας αξιοποιεί προ-εκπαιδευμένα συνελικτικά νευρωνικά δίκτυα για εντοπισμό οχημάτων σε ένα τμήμα οδού για προσεγγιστικό υπολογισμό κυκλοφοριακών μεγεθών. Τα δίκτυα αυτά είναι αυξημένης περιπλοκότητας που έχουν ως αποτέλεσμα την υψηλή ακρίβεια, αλλά με κόστος αυξημένες απαιτήσεις υπολογιστικής ισχύς. Η επιλογή κατάλληλου δικτύου είναι απαραίτητη για την επίτευξη ισορροπίας ταχύτητας και ακρίβειας ώστε να καλύπτει τις ανάγκες της εφαρμογής. Το δίκτυο που επιλέχτηκε πετυχαίνει σε κατάλληλο εξοπλισμό ταχύτητες έως και 30% πιο γρήγορες σε σχέση με τον ρυθμό ανανέωσης στιγμιότυπων του βίντεο, διατηρώντας εξαιρετικές ακρίβειες για τα δεδομένα της διπλωματικής εργασίας. Σημαντικό είναι να αναφερθεί πως στην εφαρμογή σε κοινό βίντεο, το μοντέλο της πρώτης προσέγγισης μπορεί να αναλύσει το ίδιο βίντεο έως πέντε φορές ταχύτερα από το μοντέλο της δεύτερης προσέγγισης, λόγω των μειωμένων, συγκριτικά υπολογιστικών απαιτήσεων.

Οι κύριοι περιορισμοί της δεύτερης προσέγγισης είναι οι αυξημένες υπολογιστικές απαιτήσεις σε σχέση με την πρώτη. Ο υπολογισμός των κυκλοφοριακών μεγεθών είναι προσεγγιστικός καθώς απαιτούνται ορισμένες παραδοχές. Αν τα βίντεο συνοδεύονταν με πραγματικές μετρήσεις των κυκλοφοριακών μεγεθών θα μπορούσε να γίνει ψιλότερη ρύθμιση ορισμένων παραμέτρων για ακριβέστερο υπολογισμό των μεγεθών. Επίσης το μοντέλο που επιλέχτηκε τελικά, έχει τον περιορισμό ενός μέγιστου πλήθους προβλέψεων σε κάθε εικόνα. Βέβαια αυτό το πλήθος είναι μεγαλύτερο από το μέγιστο πλήθος οχημάτων που μπορούν να βρίσκονται

ταυτόχρονα στο εξεταζόμενο τμήμα της οδού. Επίσης όπως και η πρώτη μέθοδος, υπάρχει περιορισμός εφαρμογής κατά την διάρκεια την ημέρας.

Τα πλεονεκτήματα της δεύτερης προσέγγισης είναι ο πιο αξιόπιστος εντοπισμός και πρόβλεψη των οχημάτων. Αυτό συμβαίνει γιατί η μέθοδος δεν επηρεάζεται σημαντικά από τα χαρακτηριστικά του βίντεο. Πιο συγκεκριμένα δεν είναι απαραίτητο τα βίντεο προς εξέταση να περιέχουν μόνο το τμήμα οδού με την βέλτιστη ορατότητα. Επίσης οι σκιές επηρεάζουν μερικές φορές τον σωστό εντοπισμό γκρίζων η μαύρων οχημάτων. Τέλος η δυνατότητα εξαγωγής ποσοτικών κυκλοφοριακών δεδομένων είναι ένα σημαντικό πλεονέκτημα έναντι τις πρώτης προσέγγισης.

Συνοψίζοντας, πρέπει να τονιστεί πως τα δυο πρότυπα δεν είναι άμεσα συγκρίσιμα καθώς οι προτεινόμενες εφαρμογές τους δεν είναι απαραίτητα οι ίδιες. Πιο συγκεκριμένα το πρώτο πρότυπο θα μπορούσε να αξιοποιηθεί για τον γρήγορο γενικό καθορισμό κυκλοφοριακών συνθηκών σε διαδοχικά τμήματα μιας κεντρικής οδικής αρτηρίας. Το δεύτερο πρότυπο μπορεί να αξιοποιηθεί για καταγραφή οχημάτων και υπολογισμό κυκλοφοριακών μεγεθών σε ένα ή πολλά τμήματα μίας οδού που δεν είναι εξοπλισμένη με σταθερές κάμερες ή που η πρόσβαση δεν είναι εύκολη με διαφορετικό τρόπο.

6.3 - Προτάσεις για Περαιτέρω Έρευνα

Οι εφαρμογές βαθιάς μάθησης η χρήση των μη επανδρωμένων αεροσκαφών είναι πολύ ενδιαφέροντα και σύγχρονα θέματα για τον συγκοινωνιολόγο μηχανικό, και οι πιθανές εφαρμογές φαντάζουν απεριόριστες με την βελτίωση της τεχνολογίας και την περαιτέρω έρευνα.

Η παρούσα διπλωματική εργασία αποτελεί μία πρώτη έρευνα με περιορισμένα δεδομένα, οπού ο κύριος σκοπός της είναι να εξεταστούν μερικές από τις πιθανές προσεγγίσεις και να αξιολογηθεί η δυνατότητα εφαρμογής τους στον τομέα των μεταφορών.

Οι προτάσεις για περαιτέρω έρευνα είναι σε πρώτο στάδιο η συλλογή μεγάλης βάσης δεδομένων με διαφορετικά, διακριτά δεδομένα. Αυτό θα συνεισφέρει σημαντικά στην δημιουργία πολύ πιο αξιόπιστων μοντέλων με καλύτερη δυνατότητα γενίκευσης. Επίσης ο περιορισμός δεδομένων χωρίς επαρκή φωτισμό ή νυχτερινής λήψης ίσως μπορεί να αντιμετωπιστεί με χρήση κάμερα υπέρυθρων. Τα επόμενα βήματα αυτής της εργασίας είναι η εφαρμογή των προτύπων με μη επανδρωμένο αεροσκάφος επί τόπου στο τμήμα μίας οδού. Επίσης μπορεί να εξεταστεί το ενδεχόμενο εξοπλισμού του μη επανδρωμένου αεροσκάφους με κάποια συσκευή (π.χ. Raspberry Pi 3), που θα αναλύει το βίντεο επιτόπου σε πραγματικό χρόνο. Ένα τέτοιο ενδεχόμενο θα ήταν ιδανικό αν υπάρχει η δυνατότητα να ξεπεραστεί ο περιορισμός της περιορισμένης υπολογιστικής ισχύς των μικρών συσκευών. Μία άλλη εναλλακτική είναι να εξεταστεί το ενδεχόμενο της ανάλυσης του βίντεο από έξυπνο κινητό (Smartphone π.χ. συσκευές Android) συνδεδεμένο ασύρματα με το μη επανδρωμένο αεροσκάφος. Να αναφερθεί πως το μοντέλο που επιλέχτηκε στην δεύτερη προσέγγιση (SSD MobilNet V2) έχει δομή κατάλληλη για βέλτιστη χρήση σε τέτοιες συσκευές με περιορισμένη υπολογιστική ισχύ.

Προτεινόμενες προτάσεις για ενναλακτικές εφαρμογές είναι η αξιοποίηση παρόμοιων προτύπων για γρήγορο υπολογισμό μήκους ουράς αναμονής σε διασταυρώσεις με η και χωρίς φωτεινό σηματοδότη με την ελάχιστη δυνατή επεξεργασία του βίντεο.

Εναλλακτικά, θα μπορούσε να πραγματοποιηθεί εντοπισμός και παρακολούθηση πεζών και υπολογισμός σχετικών μεγεθών.

Επίσης μια πολύ χρήσιμη εφαρμογή που ενδείκνυται η χρήση μη επανδρωμένου αεροσκάφους είναι ο καθορισμός θέσεων στάθμευσης σε χώρους με οριοθετημένες θέσεις αλλά και σε οδούς. Η εναέρια λήψη μπορεί να προσφέρει άμεση και ευρεία πληροφορία σχετικά με τις διαθέσιμες θέσεις στάθμευσης η για το ποσοστό κατάληψης αυτών σε μία μεγάλη περιοχή.

Συνοψίζοντας, πιο εξειδικευμένες μελλοντικές εφαρμογές μπορεί να αποτελέσουν η αυτοματοποίηση σμήνους μη επανδρωμένων αεροσκαφών, για αυτοματοποιημένη καταγραφή κυκλοφοριακών συνθηκών. Επίσης η ένταξη των αεροσκαφών σε υπάρχοντα συστήματα παρακολούθησης κυκλοφορίας για δημιουργία ευφυών συνεργατικών συστημάτων παρακολούθησης και καταγραφής πληροφορίας.

7 - ΒΙΒΛΙΟΓΡΑΦΙΑ

Φραντζεσκάκης, Ι. Μ., Γκόλιας, Ι. Κ., & Πιτσιάβα-Λατινοπούλου, Μ. Χ. (2009). *Κυκλοφοριακή Τεχνική*. Αθήνα: Παπασωτηρίου.

Ahirwar, K. (2017). *Everything you need to know about Neural Networks*. Retrieved 2019, from Medium: https://hackernoon.com/everything-you-need-to-know-about-neural-networks-8988c3ee4491

Barmpounakis, E. N. (2017). Investigating the decision-making process of drivers during overtaking by Powered Two Wheelers.

Barmpounakis, E. N., Vlahogianni, E. I., & Golias, J. C. (2015). *Extracting Kinematic Characteristics From Unmanned Aerial Vehicles.*

Barmpounakis, E. N., Vlahogianni, E. I., & Golias, J. C. (2017). *Unmanned Aerial Aircraft Systems for transportation engineering: Current practice and future challenges.*

Benjdira, B., Khursheed, T., Koubaa, A., Ammar, A., & Ouni, K. (2018). *Car Detection using Unmanned Aerial Vehicles: Comparison between Faster R-CNN and YOLOv3.*

Berkeley. (2018). *Berkeley DeepDrive*. Retrieved 2019, from https://deepdrive.berkeley.edu/

Buduma , N. (2017). *Fundamentals of Deep Learning.* O'Reilly.

Chollet, F. (2018). *Deep Learning with Python.* Manning.

Cisco. (2019). *Global Mobile Data Traffic Forecast Update, 2017–2022 White Paper.*

Claude Sammut, G. I. (2010). *Encyclopedia of Machine Learning.* Boston: Springer.

Coifman, B., Mishalani, R., Wang, C., & Krishnamurthy, S. (2006). *Airborne video registration for visualization and parameter estimation of traffic flows.*

Computer Science Wiki. (2018). *Max-pooling / Pooling*. Ανάκτηση 2019, από https://computersciencewiki.org/index.php/Max-pooling_/_Pooling

Demissie, M. G., Correia, G. d., & Bento, C. L. (2013). *TRAFFIC VOLUME ESTIMATION THROUGH CELLULAR NETWORKS HANDOVER INFORMATION.* Rio de Janeiro.

Diederik P. Kingma, J. L. (2015). *Adam: A Method For Stochastic Optimization.*

Doherty, P., Granlund, G., & Kuchinski, K. (2000). *The WITAS Unmanned Aerial Vehicle Project.* ECAI.

Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). *Rich feature hierarchies for accurate object detection and semantic segmentation.* UC Berkeley.

Google-Tensorflow. (2019). *github.com/tensorflow*. Retrieved 2019, from TensorFlow-Slim image classification model library: https://github.com/tensorflow/models/tree/master/research/slim

Grant, C., Gillis, B., & Guensler, R. (2000). *Collection of Vehicle Activity Data by Video Detection for Use in Transportation Planning.*

Honglak, L., Grosse, R., Ranganath, R., & Ng, A. Y. (2011). *Unsupervised Learning of Hierarchical Representations with Convolutional Deep Belief Networks.*

Huang, J., Rathod, V., Sun, C., Zhu, M., Korattikara, A., Fathi, A., . . . Murphy, K. (2017). *Speed/accuracy trade-offs for modern convolutional object detectors.* Google Research.

Ioffe, S., & Szegedy, C. (2015). *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.*

Kaiming, H., Xiangyu, Z., Shaoqing, R., & Jian , S. (2015). *Deep Residual Learning for Image Recognition.*

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.*

Khan, M. A., Ectors, W., Bellemans, T., Janssens, D., & Wets, G. (2018). *Unmanned Aerial Vehicle-Based Traffic Analysis:.*

Khan, M. A., Ectorsa, W., Bellemansa, T., Ruichekb, Y., Yasara, A.-H., Janssensa, D., & Wets, G. (2018). *Unmanned Aerial Vehicle-based Traffic Analysis: A Case Study to Analyze Traffic Streams at Urban Roundabouts.*

Kyrkou, C., Plastiras, G., Theocharides, T., Venieris, S. I., & Bouganis, C.-S. (2018). *DroNet: Efficient Convolutional Neural Network.*

LeCun, Y., & Bengio, Y. (1995). *Convolutional Networks for Images, Speech, and Time-Series.*

Leduc, G. (2008). Road Traffic Data: Collection Methods and Applications. *JRC.*

Lee, T.-C. (2007). *An agent-based model to simulate motorcycle behaviour in mixed traffic flow.*

Lin, T.-Y., Maire, M., Bolongie, S., Bourdev, L., Girshick, R., Hays, J., . . . Dollar, P. (2015). *Microsoft COCO: Common Objects in Context.*

Liu, H. X., Ban, X., Ran, B., & Mirchandani, P. (2002). *Analytical dynamic traffic assignment model with probabilistic travel times and perceptions.*

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016). *SSD: Single Shot MultiBox Detector.*

Liu, Y., Zhu, C., Deng, X., Guan, P., Wan, Z., JieLuo, . . . Zhang, H. (2019). *UAV-aided urban target tracking system based on edge computing.*

Mairaj, A., Baba, A. I., & Javaid, A. Y. (2019). *Application Specific Drone Simulators: Recent Advances and Challenges.* Toledo: The University of Toledo.

Medium. (2018). *Understanding of Convolutional Neural Network (CNN)—Deep Learning*. Retrieved 2019, from Medium: https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148

Nielsen, M. (2018). *Neural Networks and Deep Learning.*

NVIDIA. (2018). *Deep Learning - Autonomous Cars*. Ανάκτηση 2019, από https://www.nvidia.com/object/autonomous-cars.html

Powers, D. M. (2007). *Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation.* Adelaide: Flinders University.

Redmon, J., & Farhadi, A. (2018). *YOLOv3: An Incremental Improvement.* University of Washington.

Rednoa Inc. (2017). *A Report on The use of Unmanned Aerial Systems (UAS) to Remotely Collect Data for Road Infrastructure.*

Rezaei, M., & Isehaghi, M. (2018). *An Efficient Method for License Plate Localization Using Multiple Statistical Features in a Multilayer Perceptron Neural Network.*

Sanderson, P., & Michael, M. (2011). *Accuracy of Pneumatic Road Tube Counters.* Anchorage: 2011 Western District Annual Meeting.

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). *MobileNetV2: Inverted Residuals and Linear Bottlenecks.* Google Inc.

Sermanet, P., & LeCun, Y. (2011). *Traffic Sign Recognition with Multi-Scale Convolutional Networks.*

Shah, A., Lamare, J. B., Anh, T. N., & Hauptmann, A. (2018). *Accident Forecasting in CCTV Traffic Camera Videos.*

Shastry, A., & Schowengerdt, R. (2002). *Airborne video registration for visualization and parameter estimation of traffic flows.*

Sinno Jialin Pan, Q. Y. (2010). *A Survey on Transfer Learning.*

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from. *Journal of Machine Learning Research 15 (2014) 1929-1958.*

Steenbruggen, J., Borzacchiello, M. T., Nijkamp, P., & Scholten, H. (2011). *Mobile phone data from GSM networks for trafficparameter and urban spatial pattern assessment:a review of applications and opportunities.*

Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., & Liu, C. (2018). *A Survey on Deep Transfer Learning.*

Taylor, L., & Nitschke, G. (2017). *Improving Deep Learning using Generic Data Augmentation.* Cape Town: University of Cape Town.

Wang, X., Cheng, P., Liu, X., & Uzochukwu, B. (2018). *Fast and Accurate, Convolutional Neural Network Based Approach for Object Detection From UAV.*

Wang, X., Yu, F., Dou, Z.-Y., Darrel, T., & Gonzalez, J. E. (2018). *SkipNet: Learning Dynamic Routing in Convolutional Networks.* University of California, Berkeley.

Wikipedia. (2019). *Activation function*. Retrieved 2019, from https://en.wikipedia.org/wiki/Activation_function

Wikipedia. (2019). *Artificial neural network*. Retrieved 2019, from https://en.wikipedia.org/wiki/Artificial_neural_network

XRDS. (2016). *Convolutional Neural Networks (CNNs): An Illustrated Explanation*. Ανάκτηση 2018, από XRDS: https://blog.xrds.acm.org/2016/06/convolutional-neural-networks-cnns-illustrated-explanation/

Yu, F., Xian, W., Chen, Y., Liu, F., Liao , M., Madhavan, V., & Darrell, T. (2018). *BDD100K: A Diverse Driving Video Database with.* UC Berkeley.

Zherzdez, S., & Gruzdev, A. (2018). *LPRNet: License Plate Recognition via Deep Neural Networks.*