

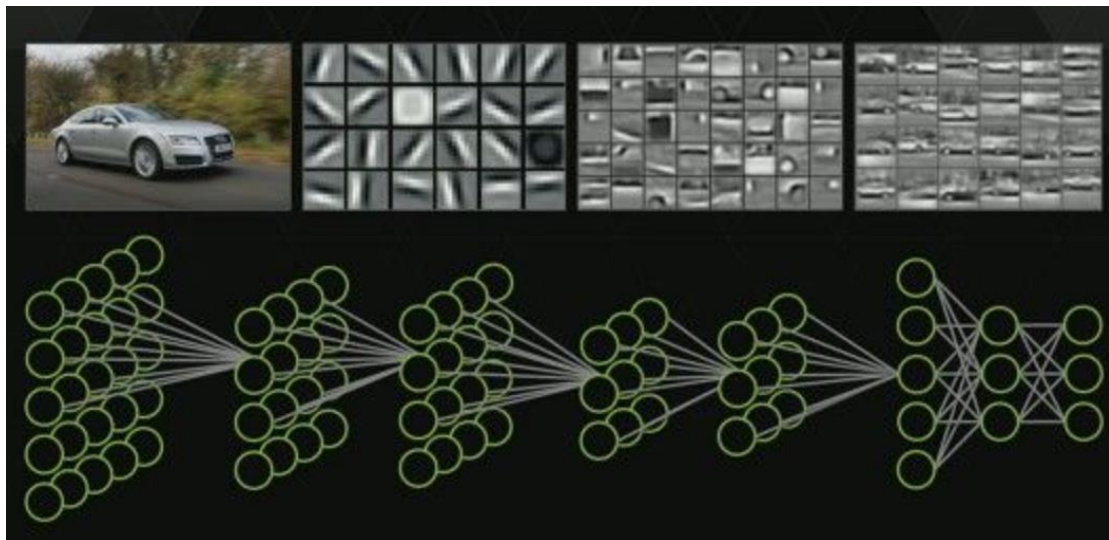


ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

Σχολή Πολιτικών Μηχανικών

Τομέας Μεταφορών και Συγκοινωνιακής Υποδομής

Ανάπτυξη Συνελκτικών Νευρωνικών Δικτύων για την Αναγνώριση Ελεύθερων Θέσεων Στάθμευσης επί της Οδού



Διπλωματική Εργασία

Κωνσταντίνος Γκόλιας

Επιβλέπουσα Καθηγήτρια : Ελένη Βλαχογιάννη,
Επίκουρη Καθηγήτρια Σχολής Πολιτικών Μηχανικών ΕΜΠ

Αθήνα, Ιούλιος 2016

Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Πολιτικών Μηχανικών
Τομέας Μεταφορών και Συγκοινωνιακής Υποδομής

**Ανάπτυξη Συνελικτικών Νευρωνικών Δικτύων για την Αναγνώριση Ελεύθερων
Θέσεων Στάθμευσης επί της Οδού**

Κωνσταντίνος Γκόλιας
Επιβλέπουσα Καθηγήτρια : Ελένη Βλαχογιάννη
Αθήνα, 2016.

National Technical University of Athens
School of Civil Engineering
Department of Transportation Planning and Engineering,

**Development of Convolutional Neural Networks for On-Street Free Parking
Space Detection**

Thesis Author: Konstantinos Gkolas
Supervising Professor: Eleni Vlachogianni
Athens, 2016

Copyright © Κωνσταντίνος Γκόλιας, 2016

Με επιφύλαξη παντός δικαιώματος

Απαγορεύεται η αντιγραφή, αποθήκευση σε αρχείο πληροφοριών, διανομή, αναπαραγωγή, μετάφραση ή μετάδοση της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό, υπό οποιαδήποτε μορφή και με οποιοδήποτε μέσο επικοινωνίας, ηλεκτρονικό ή μηχανικό, χωρίς την προηγούμενη έγγραφη άδεια της συγγραφέως. Επιτρέπεται η αναπαραγωγή, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν στη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Η έγκριση της διπλωματικής εργασίας από τη Σχολή Πολιτικών Μηχανικών του Εθνικού Μετσοβίου Πολυτεχνείου δεν υποδηλώνει αποδοχή των απόψεων του συγγραφέως (Ν. 5343/1932, Άρθρο 202).

Copyright © Konstantinos Gkolias, 2016

All Rights Reserved

Neither the whole nor any part of this diploma thesis may be copied, stored in a retrieval system, distributed, reproduced, translated, or transmitted for commercial purposes, in any form or by any means now or hereafter known, electronic or mechanical, without the written permission from the author. Reproducing, storing and distributing this thesis for non-profitable, educational or research purposes is allowed, without prejudice to reference to its source and to inclusion of the present text. Any queries in relation to the use of the present thesis for commercial purposes must be addressed to its author.

Approval of this diploma thesis by the School of Civil Engineering of the National Technical University of Athens (NTUA) does not constitute in any way an acceptance of the views of the author contained herein by the said academic organization (L. 5343/1932, art. 202).

ΕΥΧΑΡΙΣΤΙΕΣ

Κατ'αρχάς, θα ήθελα να ευχαριστήσω θερμά, την επιβλέπουσα καθηγήτριά μου, κυρία Ελένη Βλαχογιάννη, Επίκουρη Καθηγήτρια στον Τομέα Μεταφορών και Συγκοινωνιακής Υποδομής της Σχολής Πολιτικών Μηχανικών, για την καθολική της στήριξη και καθοδήγηση σε όλο το διάστημα της συνεργασίας μας, για την εκπόνηση της διπλωματικής μου εργασίας. Ακόμη την ευχαριστώ, επειδή αποτέλεσε τον άνθρωπο που με μύησε στην λογική των νευρωνικών δικτύων, δεδομένης της μεγάλης εμπειρίας της στο θέμα αυτό, συμβάλλοντας καθοριστικά στην επαφή μου με ένα τόσο σύγχρονο και εφαρμόσιμο εργαλείο.

Ακόμη, θα ήθελα να ευχαριστήσω τον Υποψήφιο Διδάκτορα, Εμμανουήλ Μπαρμπουνάκη για την διαθεσιμότητά του και την προθυμία του να βοηθήσει με κάθε τρόπο στην πραγματοποίηση αυτής της διπλωματικής εργασίας. Κάθε παρατήρηση ή ενέργειά του, συνέβαλε στην υπέρβαση εμποδίων, που υπό άλλες συνθήκες θα απαιτούσαν πολύ μεγάλη προσπάθεια για να επιλυθούν.

Τέλος, ευχαριστώ ολόψυχα την οικογένειά μου, τους φίλους μου και τα κοντινά μου πρόσωπα που ήταν δίπλα μου όλο αυτό το διάστημα. Ειδικότερα, ευχαριστώ θερμά τον πατέρα μου, που με την ιδιότητα του καθηγητή, μπορούσε να με συμβουλέψει και να με κατευθύνει σε μία σειρά θεμάτων, αλλά και τον αδερφό μου, ο οποίος σπουδάζει Computer Science στις Η.Π.Α. και οι συζητήσεις μου μαζί του συνέβαλαν πολύ, όχι μόνο στην καλύτερη κατανόηση του αντικειμένου, αλλά και στην επαφή μου με προσεγγίσεις και μεθόδους που χρησιμοποιούνται και διδάσκονται σε άλλες χώρες.

**Τίτλος: Ανάπτυξη Συνελκτικών Νευρωνικών Δικτύων για την Αναγνώριση
Ελεύθερων Θέσεων Στάθμευσης επί της Οδού**

Συγγραφέας Διπλωματικής Εργασίας: Κωνσταντίνος Γκόλιας

Επιβλέπουσα Καθηγήτρια: Ελένη Βλαχογιάννη

ΣΥΝΟΨΗ

Η επεξεργασία της εικόνας και τα οπτικά μέσα αποτελούν την βάση της υπολογιστικής όρασης και αναπόφευκτα αναπόσπαστο κομμάτι των μελλοντικών δομών των αυτοκινήτων. Σκοπός της παρούσας διπλωματικής εργασίας είναι η ανάπτυξη προτύπων μηχανικής μάθησης για την αναγνώριση των ελεύθερων θέσεων στάθμευσης επί της οδού, σε αστικά οδικά δίκτυα από βίντεο που προκύπτουν από τις κάμερες οι οποίες υπάρχουν ή θα υπάρχουν πάνω στα οχήματα. Αρχικά, διερευνήθηκε και αξιολογήθηκε η σχέση της θέσης της κάμερας πάνω στο όχημα (ύψος και γωνία λήψης) με το μήκος παρά το κράσπεδο που αποτυπώνεται στις παραγόμενες φωτογραφίες, και κατ' επέκταση, την ανίχνευση της ύπαρξης ή μη ελεύθερης θέσης παρά το κράσπεδο. Αναπτύχθηκαν, εκπαιδεύτηκαν και αξιολογήθηκαν συνελκτικά νευρωνικά δίκτυα με τη χρήση φωτογραφιών από κάμερα σε κινούμενο όχημα, αφού πρώτα αυτές μετατράπηκαν σε κατάλληλους πίνακες για επεξεργασία, ώστε να διατηρηθεί η χρήσιμη πληροφορία τους. Η βελτιστοποίηση της δομής και των παραμέτρων εκπαίδευσης των συνελκτικών νευρωνικών δικτύων, με στόχο την ελαχιστοποίηση του παραγόμενου σφάλματος γενίκευσης, οδήγησε σε σημαντικά βελτιωμένη ακρίβεια πρόβλεψης κενών θέσεων, της τάξης του 90%. Τέλος, διερευνήθηκε ο τρόπος εφαρμογής των προτύπων που αναπτύχθηκαν, ώστε να προκύπτουν χρήσιμες για τους μετακινούμενους πληροφορίες σε πραγματικό χρόνο.

Λέξεις Κλειδιά : συνελκτικά δίκτυα, νευρωνικά δίκτυα, μηχανική μάθηση, υπολογιστική όραση, στάθμευση επί της οδού, ευφυή συστήματα διαχείρισης στάθμευσης, εύρεση κενών θέσεων στάθμευσης, ανίχνευση χαρακτηριστικών

Title: Development of Convolutional Neural Networks for On-Street Free Parking Space Detection

Thesis Author: Konstantinos Gkolias

Supervising Professor: Eleni Vlachogianni

ABSTRACT

Image processing and video analytics are the cornerstone of computer vision and inevitably integral part of the vehicles of the future. The purpose of this diploma thesis is the development of machine learning models for the recognition of on-street, free parking spaces in urban road networks from video recordings that come from cameras that exist or will exist on the vehicles. First, the relation between the position of the camera on the vehicle (height and video angle) with the measured length of the road section, which is depicted in the resulted video frames and as a consequence, with the trace of the existence of an on-street free parking space, was studied and evaluated. Second, Convolutional Neural Networks were developed, trained and evaluated with the use of images from a camera on a moving vehicle, after their conversion to suitable matrices for processing, so that their useful information is preserved. The optimized Convolutional Networks, in terms of structural and learning parameters, provided predictions for the existence of on-street free parking spaces with approximately 90% average accuracy. Finally, the implementation of a framework, which integrates the developed models to produce meaningful parking information for commuters in real time was discussed.

Key Words: convolutional networks, neural networks, machine learning, computer vision, on-street parking, intelligent parking management systems, detection of free parking space, feature detection

ΠΕΡΙΛΗΨΗ

Η δυσκολία εύρεσης ελεύθερης θέσης στάθμευσης για τους μετακινούμενους στα αστικά κέντρα δημιουργεί πολλά προβλήματα τόσο στον χρήστη όσο και στο σύστημα και το περιβάλλον. Η αναζήτηση για θέση στάθμευσης στα σύγχρονα αστικά κέντρα, έχει οδηγήσει στην αύξηση του μέσου χρόνου εύρεσής της, με αποτέλεσμα να αυξάνεται σημαντικά ο μέσος χρόνος ταξιδιού του χρήστη. Ακόμη η διαδικασία αυτή, έχει αρνητικό αντίκτυπο στην ψυχοσύνθεσή του, καθώς δημιουργεί άγχος, κούραση και κακή οδηγική συμπεριφορά. Η αύξηση της κυκλοφοριακής συμφόρησης, δεδομένης της αύξησης του αριθμού των οχημάτων που αναζητούν θέση στάθμευσης αλλά και του χρόνου εύρεσής της, οδηγούν στην επιβάρυνση του οδικού δικτύου αλλά και την εκπομπή πολλών επιπλέον επιβλαβών καυσαερίων στο περιβάλλον. Η επίλυση των παραπάνω προβλημάτων βασιζόταν στην επίλυση του προβλήματος της εύρεσης κενής θέσης στάθμευσης επί της οδού και αποτέλεσε το κίνητρο για την πραγματοποίηση της παρούσας διπλωματικής εργασίας.

Η χρήση της εικόνας κρίθηκε αναγκαία, εφόσον αποτελεί τον πιο διαδεδομένο και αποτελεσματικό τρόπο για την βελτίωση της απόδοσης της υπολογιστικής όρασης. Εξετάστηκαν εναλλακτικές προσεγγίσεις της βιβλιογραφίας στις οποίες χρησιμοποιήθηκαν αισθητήρες, υπέρηχοι, ΜΔΥ ή άλλοι τρόποι ανίχνευσης της διαθεσιμότητας των θέσεων επί της οδού. Κρίθηκε ότι τα συνελκτικά νευρωνικά δίκτυα αποτελούν την βέλτιστη μεθοδολογία επίλυσης, εφόσον έχουν την δυνατότητα να αναγνωρίζουν με μεγάλη ακρίβεια χαρακτηριστικά εντός των εικόνων, άρα εν προκειμένω τα οχήματα ή άλλα εμπόδια που χαρακτηρίζουν μία θέση ως κενή ή κατειλημμένη.

Τα συνελκτικά δίκτυα ανήκουν στην ευρύτερη οικογένεια των συστημάτων μηχανικής μάθησης και αποτελούν μια ειδική κατηγορία νευρωνικών δικτύων με κύριο χαρακτηριστικό την ικανότητά τους να επεξεργάζονται επιτυχώς εικόνες. Η δυνατότητά τους αυτή, οφείλεται αφενός στην δομή τους, η οποία επιτρέπει την εισαγωγή δεδομένων τριών διαστάσεων (όπως οι εικόνες) και στην εκπαίδευσή τους, η οποία δεν ακολουθεί την τακτική σύνδεσης των νευρώνων «όλα με όλα» όπως στα απλά νευρωνικά δίκτυα, αλλά δημιουργεί τοπικούς τομείς ανάλυσης, στους οποίους οφείλεται και η ανίχνευση των ιδιαίτερων χαρακτηριστικών των εικόνων.

Σε πρώτη φάση, πραγματοποιήθηκε η συλλογή των απαραίτητων δεδομένων για την ανάλυση και εκπαίδευση των συνελκτικών νευρωνικών δικτύων. Διερευνήθηκε η σχέση της θέσης της κάμερας πάνω στο όχημα (ύψος και γωνία λήψης) με το μήκος

παρά το κρiάσπεδο που αποτυπώνεται στις παραγόμενες φωτογραφίες, και κάτ' επέκταση, την ύπαρξη ή μη ελεύθερης θέσης παρά το κρiάσπεδο.

Στην συνέχεια, και αφού συλλέχθηκαν τα δεδομένα, ακολούθησε μία εκτεταμένη ανάλυση, η οποία πραγματοποιήθηκε σε ορισμένα στάδια αυξανόμενης δυσκολίας. Σε πρώτο στάδιο, επιλύθηκε το πρόβλημα με ένα υποσύνολο δεδομένων στο οποίο οι εικόνες που χρησιμοποιήθηκαν ήταν πιο ξεκάθαρες ως προς την κατηγοριοποίησή τους ως 0/1. Ουσιαστικά, άλλου τύπου εμπόδια εντός των εικόνων, θεωρήθηκαν συνθετότερα ως προς την αναγνώρισή τους. Σε δεύτερο στάδιο, χρησιμοποιήθηκαν όλες οι φωτογραφίες και αφού έγινε αντιληπτό ένα συστηματικό λάθος που υπονόμει την απόδοση του δικτύου, συγκεκριμένα η ύπαρξη κόμβων, αφαιρέθηκαν τα αντίστοιχα δεδομένα και το δίκτυο εκπαιδεύτηκε με ακρίβεια 90%. Πραγματοποιήθηκε μια σειρά από δοκιμές, για πολλές διαφορετικές μορφές του συνελκτικού δικτύου ή παραμέτρους του, ώστε να επιβεβαιωθεί η βέλτιστη απόδοσή του.

Τέλος, λαμβάνοντας υπόψη τόσο την εξέλιξη της τεχνολογίας όσο και τις δυνατότητες που αυτή προσφέρει, διερευνήθηκε μία στρατηγική εφαρμογής του προτύπου, η οποία βασίστηκε στις σύγχρονες επικρατούσες τάσεις του πληθοπορισμού και των «Ευφυών Συστημάτων» στα αστικά κέντρα. Ειδικότερα, σχεδιάστηκε η αρχιτεκτονική ενός συστήματος, που καλείται να αντιμετωπίσει πολλά από τα προβλήματα της εφαρμογής του υπό πραγματικές συνθήκες, και διατυπώθηκαν επιμέρους προβλήματα που πρέπει να επιλυθούν με αντίστοιχους αλγόριθμους για την υλοποίησή του.

ΠΕΡΙΕΧΟΜΕΝΑ

1.	ΕΙΣΑΓΩΓΗ	1
1.1	Η ΣΤΑΘΜΕΥΣΗ ΣΤΑ ΑΣΤΙΚΑ ΚΕΝΤΡΑ	1
1.2	ΤΕΧΝΟΛΟΓΙΚΗ ΕΞΕΛΙΞΗ ΑΥΤΟΚΙΝΗΤΩΝ-ΕΥΡΕΙΑ ΧΡΗΣΗ ΤΗΣ ΚΑΜΕΡΑΣ	4
1.3	ΑΠΟ ΤΗΝ ΥΠΟΛΟΓΙΣΤΙΚΗ ΎΡΑΣΗ ΣΤΗ ΣΥΓΚΟΙΝΩΝΙΑΚΗ ΠΛΗΡΟΦΟΡΙΑ.....	8
1.4	ΣΚΟΠΟΣ ΤΗΣ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ	9
1.5	ΔΟΜΗ ΤΗΣ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ	10
2.	ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ	12
2.1	Ο ΠΛΗΘΟΠΟΡΙΣΜΟΣ ΣΤΑ ΣΥΓΚΟΙΝΩΝΙΑΚΑ ΣΥΣΤΗΜΑΤΑ	13
2.2	ΜΕΘΟΔΟΙ ΧΩΡΙΣ ΤΗΝ ΧΡΗΣΗ ΦΩΤΟΓΡΑΦΙΑΣ.....	14
2.2.1	<i>Αισθητήρες επί της οδού</i>	14
2.2.2	<i>Πομποί Υπερήχων επί των Οχημάτων</i>	15
2.2.3	<i>Ραντάρ επί της Οδικής Υποδομής</i>	16
2.2.4	<i>Συγκριτική Αξιολόγηση Μεθόδων χωρίς την Χρήση της Φωτογραφίας</i>	17
2.3	ΜΕΘΟΔΟΙ ΜΕ ΤΗΝ ΧΡΗΣΗ ΤΗΣ ΦΩΤΟΓΡΑΦΙΑΣ	19
2.3.1	<i>Ανίχνευση Σχημάτων</i>	19
2.3.2	<i>Χρήση Κινητών Τηλεφώνων και Διαδικτύου Με προαιρετική τη χρήση video</i>	20
2.3.3	<i>Μηχανές Διανυσμάτων Υποστήριξης</i>	21
2.3.4	<i>Ταυτοποίηση Οχημάτων</i>	22
2.3.5	<i>Τρισδιάστατη Προσέγγιση</i>	23
2.3.6	<i>Συνελικτικά Νευρωνικά Δίκτυα</i>	24
2.3.7	<i>Βασικά Συμπεράσματα</i>	26
3.	ΜΕΘΟΔΟΛΟΓΙΑ	28
3.1	ΔΙΑΤΥΠΩΣΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ.....	28
3.2	ΠΑΡΑΔΟΧΕΣ	29
3.3	ΕΠΙΛΟΓΗ ΜΕΘΟΔΟΥ ΕΠΙΛΥΣΗΣ.....	30
3.4	ΑΠΑΙΤΟΥΜΕΝΟ ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ	31
3.5	ΔΟΜΗ & ΕΚΠΑΙΔΕΥΣΗ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ.....	32
3.5.1	<i>Δομή Νευρωνικών Δικτύων</i>	32
3.5.2	<i>Εκπαίδευση Νευρωνικών Δικτύων</i>	35
3.6	ΣΥΝΕΛΙΚΤΙΚΑ ΔΙΚΤΥΑ.....	44
3.6.1	<i>Γενικά</i>	44
3.6.2	<i>Τρισδιάστατη Δομή</i>	44
3.6.3	<i>Τοπικοί Τομείς Ανάλυσης</i>	45
3.6.4	<i>Επίπεδα Συγκέντρωσης</i>	50

3.6.5	<i>Αλληλουχία Επιπέδων και ένα Παράδειγμα.....</i>	51
3.7	ΕΠΙΛΟΓΗ ΛΟΓΙΣΜΙΚΟΥ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΟΙ ΠΕΡΙΟΡΙΣΜΟΙ	54
3.8	ΜΕΘΟΔΟΣ ΑΞΙΟΛΟΓΗΣΗΣ.....	55
3.9	ΔΙΑΓΡΑΜΜΑ ΡΟΗΣ ΠΕΙΡΑΜΑΤΟΣ.....	57
4.	ΣΥΛΛΟΓΗ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ	58
4.1	ΣΥΛΛΟΓΗ ΒΙΝΤΕΟ.....	58
4.1.1	<i>Εξοπλισμός.....</i>	59
4.1.2	<i>Επιλογή Θέσης Λήψης.....</i>	60
4.1.3	<i>Επι Τόπου Έρευνα.....</i>	62
4.2	ΠΡΟΕΤΟΙΜΑΣΙΑ ΔΕΔΟΜΕΝΩΝ	63
4.2.1	<i>Μετατροπή Βίντεο σε Φωτογραφίες</i>	63
4.2.2	<i>Διαχωρισμός Φωτογραφιών και Υποπροβλήματα</i>	64
4.2.3	<i>Τελικό Στάδιο Επεξεργασίας</i>	65
5.	ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΑΞΙΟΛΟΓΗΣΗ ΔΙΚΤΥΩΝ	68
5.1	ΠΡΟΒΛΗΜΑ ΜΕΙΩΜΕΝΗΣ ΠΟΛΥΠΛΟΚΟΤΗΤΑΣ	69
5.1.1	<i>Δεδομένα.....</i>	69
5.1.2	<i>Παραμετροποίηση Δομής και Εκπαίδευσης.....</i>	69
5.1.3	<i>Αποτελέσματα.....</i>	72
5.2	ΠΡΟΒΛΗΜΑ ΑΥΞΗΜΕΝΗΣ ΠΟΛΥΠΛΟΚΟΤΗΤΑΣ.....	75
5.2.1	<i>Δεδομένα.....</i>	75
5.2.2	<i>Παραμετροποίηση Δομής και Εκπαίδευσης</i>	75
5.2.3	<i>Αποτελέσματα.....</i>	77
5.3	ΣΥΝΟΛΙΚΟ ΠΡΑΓΜΑΤΙΚΟ ΠΡΟΒΛΗΜΑ	80
5.3.1	<i>Δεδομένα.....</i>	80
5.3.2	<i>Παραμετροποίηση Δομής και Εκπαίδευσης.....</i>	80
5.3.3	<i>Αποτελέσματα.....</i>	86
5.4	ΕΝΑ ΕΥΦΥΕΣ ΣΥΣΤΗΜΑ ΣΤΑΘΜΕΥΣΗΣ	89
6.	ΣΥΜΠΕΡΑΣΜΑΤΑ	94
6.1	ΓΕΝΙΚΑ	94
6.2	ΜΕΘΟΔΟΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ.....	95
6.3	ΒΑΣΙΚΑ ΣΥΜΠΕΡΑΣΜΑΤΑ	96
6.4	ΠΡΟΤΑΣΕΙΣ ΓΙΑ ΠΕΡΑΙΤΕΡΩ ΈΡΕΥΝΑ.....	98
7.	ΒΙΒΛΙΟΓΡΑΦΙΑ	100

ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ

Πίνακας 1.1 Απόλυτος αριθμός οχημάτων στην Ευρώπη 1970-2013, Ευρωπαϊκή Επιτροπή Statistical Pocketbook 2015	2
Πίνακας 1.2 Ετήσιες επιπλέον εκπομπές των Ι.Χ που αναζητούν θέσεις Στάθμευσης	4
Πίνακας 2.1 Σύγκριση Εγχειρημάτων για στάθμευση επί της οδού χωρίς την χρήση της φωτογραφίας (LIAO et al., 2016)	18
Πίνακας 4.1 Χαρακτηριστικά Βίντεο Κάμερας	59
Πίνακας 4.2 Μετρήσεις Πλάτους Φωτογραφίας για διάφορες Αποστάσεις από τον Φακό (d), διάφορες Γωνίες ως προς την Κατακόρυφο (θ) και διάφορα Ύψη (h)	62
Πίνακας 5.1 Δομή Συνελικτικού Δικτύου Προβλήματος Μειωμένης Πολυπλοκότητας	69
Πίνακας 5.2 Δεδομένα και Παράμετροι Συνελικτικού Δικτύου Προβλήματος Μειωμένης Πολυπλοκότητας	70
Πίνακας 5.3 Πίνακας Προβλέψεων/Πραγματικότητας και Στατιστικά Μέτρα Κατηγοριοποίησης του Συνελικτικού Δικτύου για το Πρόβλημα Μειωμένης Πολυπλοκότητας	73
Πίνακας 5.4 Δομή Συνελικτικού Δικτύου Προβλήματος Αυξημένης Πολυπλοκότητας	75
Πίνακας 5.5 Δεδομένα και Παράμετροι Συνελικτικού Δικτύου Προβλήματος Αυξημένης Πολυπλοκότητας	76
Πίνακας 5.6 Πίνακας Προβλέψεων/Πραγματικότητας και Στατιστικά Μέτρα Κατηγοριοποίησης του Συνελικτικού Δικτύου για το Πρόβλημα Αυξημένης Πολυπλοκότητας	77
Πίνακας 5.7 Αύξων Αριθμός των Λαθών του Συνελικτικού Δικτύου Εύκολων και Δύσκολων Φωτογραφιών (με Κίτρινο Υπογραμμισμένες οι Δύσκολες Φωτογραφίες)	78
Πίνακας 5.8 Δομή Συνελικτικού Δικτύου Συνολικού Πραγματικού Προβλήματος με 2 Κρυφά Επίπεδα	81
Πίνακας 5.9 Δομή Συνελικτικού Δικτύου Συνολικού Πραγματικού Προβλήματος με 3 Κρυφά Επίπεδα	82
Πίνακας 5.10 Δεδομένα και Παράμετροι Συνελικτικού Δικτύου Συνολικού Πραγματικού Προβλήματος με 2 Κρυφά Επίπεδα για 2 Στάδια Εκπαίδευσης	83
Πίνακας 5.11 Πίνακας Προβλέψεων/Πραγματικότητας και Στατιστικά Μέτρα Κατηγοριοποίησης του Συνελικτικού Δικτύου για το Συνολικό Πραγματικό Πρόβλημα (2 κρυφά επίπεδα)	84
Πίνακας 5.12 Δεδομένα και Παράμετροι Συνελικτικού Δικτύου Συνολικού Πραγματικού Προβλήματος με 2 Κρυφά Επίπεδα για 2 Στάδια Εκπαίδευσης	85

Πίνακας 5.13 Πίνακας Προβλέψεων/Πραγματικότητας και Στατιστικά Μέτρα Κατηγοριοποίησης του Συνελικτικού Δικτύου για το Συνολικό Πραγματικό Πρόβλημα (3 κρυφά επίπεδα) _____ 86

Πίνακας 5.14 Θέσεις Λαθών των Συνελικτικών Δικτύων Με 2 και 3 Κρυφά Επίπεδα. Τα υπογραμμισμένα με Κίτρινο Αντιστοιχούν στα Κοινά Λάθη _____ 87

ΕΥΡΕΤΗΡΙΟ ΔΙΑΓΡΑΜΜΑΤΩΝ

Διάγραμμα 2.1 Χρήση υπερήχων για την ανίχνευση ελεύθερης θέσης στάθμευσης (LIAO, et al., 2016) _____ 16

Διάγραμμα 3.1 Συνάρτηση RELU _____ 49

Διάγραμμα 5.1 Ακρίβεια Συνόλου Εκπαίδευσης και Συνόλου Επικύρωσης ως προς τον αριθμό των Εποχών Εκπαίδευσης του Συνελικτικού Δικτύου _____ 71

Διάγραμμα 5.2 Ακρίβεια Συνόλου Εκπαίδευσης και Συνόλου Επικύρωσης ως προς τον αριθμό των Εποχών Εκπαίδευσης του Συνελικτικού Δικτύου _____ 76

Διάγραμμα 5.3 Ακρίβεια Συνόλου Εκπαίδευσης και Συνόλου Επικύρωσης ως προς τον αριθμό των Εποχών Εκπαίδευσης του Συνελικτικού Δικτύου _____ 85

ΕΥΡΕΤΗΡΙΟ ΕΙΚΟΝΩΝ

Εικόνα 1.1 Αναπαράσταση του περιβάλλοντος ενός οχήματος FORD με την χρήση LIDAR το βράδυ __ 6

Εικόνα 1.2 Οι 4 κάμερες της INFINITI στο σύστημά της AROUND VIEW MONITOR _____ 7

Εικόνα 2.1 Τριπλέτες των θέσεων πανοραμικής λήψης με την χρήση ΜΔΥ (WU et al., 2009) _____ 22

Εικόνα 2.2 Απεικόνιση μίας δοκιμής των (JONES & LI, 2012) _____ 24

Εικόνα 3.1 Δομή Νευρωνικού Δικτύου με δύο κρυμμένα επίπεδα _____ 32

Εικόνα 3.2 Αναπαράσταση Συνάρτησης Κόστους με δύο μεταβλητές _____ 35

Εικόνα 3.3 Αλλαγή Τρισδιάστατης Μορφής Συνελικτικών Δικτύων _____ 45

Εικόνα 3.4 Συνέλιξη εικόνας 28x28 για Πυρήνα 5x5 (Nielsen, 2015)	46
Εικόνα 3.5 Ανίχνευση Επιμέρους Στοιχείων του Προσώπου και Επιμέρους Στοιχείων αυτών. (Dettmers, 2015)	48
Εικόνα 3.6 $P=1$ και για τις δύο διαστάσεις σε Πυρήνα Συγκέντρωσης 4x4	49
Εικόνα 3.7 Μέγιστη Συγκέντρωση και Τρισδιάστατη Απεικόνισή της	51
Εικόνα 3.8 Δίκτυο LUNET-5	52
Εικόνα 3.9 Παραλλαγή Συνελικτικού δικτύου LUNET-5	53
Εικόνα 3.10 Καμπύλη ROC - Όσο αριστερότερα και ψηλότερα τόσο καλύτερο αποτέλεσμα	56
Εικόνα 4.1 Go Pro Hero 4	59
Εικόνα 4.2 Μηχανισμός Στήριξης και Προστασίας της Κάμερας	60
Εικόνα 4.3 Εύκολες Εικόνες και οι Περικοπές τους (128x72-128x48)	67
Εικόνα 4.4 Δύσκολες-Rest Εικόνες και οι Περικοπές τους (128x72-128x56)	67
Εικόνα 5.1 Διάγραμμα Ροής Εφαρμογής του Προτύπου	90

1. ΕΙΣΑΓΩΓΗ

1.1 Η ΣΤΑΘΜΕΥΣΗ ΣΤΑ ΑΣΤΙΚΑ ΚΕΝΤΡΑ

Η μικρή, αλλά συνεχής αύξηση του πληθυσμού στην πλειονότητα των χωρών παγκοσμίως, όσο και Ευρωπαϊκά, σε συνδυασμό με την αύξηση του δείκτη ιδιοκτησίας Ι.Χ. (C.I.A., 2015) ή την μη μείωση (E.E.A., 2012), αναδεικνύουν το πρόβλημα της αύξησης του απόλυτου αριθμού των οχημάτων που εισέρχονται στα αστικά κέντρα. Τα τελευταία χρόνια το πρόβλημα αυτό προκαλεί έντονα και εκθετικά μεταβαλλόμενα κυκλοφοριακά προβλήματα. Το ίδιο συμπέρασμα επαληθεύεται και από τα στοιχεία του Πίνακα 1.1 που δημοσίευσε η Ευρωπαϊκή Επιτροπή το 2015. Στην πλήρη του μορφή ο Πίνακας είναι αναλυτικός για κάθε χώρα της Ευρωζώνης (European Commission, 2015).

Αξιοσημείωτο είναι να αναφερθεί πως, ενώ η ζήτηση για μετακινήσεις αυξάνεται στα αστικά κέντρα, οι θέσεις στάθμευσης στο κράσπεδο παραμένουν υπό φυσιολογικές συνθήκες σταθερές. Επομένως αυξάνεται η δυσκολία εύρεσης κενής θέσης στάθμευσης. Ο ίδιος αριθμός κενών θέσεων καταλαμβάνεται πιο γρήγορα από τον ήδη μεγαλύτερο αριθμό των χρηστών, με αποτέλεσμα ένα μεγαλύτερο (σε σχέση με πριν) ποσοστό τους να υποχρεώνεται να αναζητά περισσότερη ώρα κενή θέση επιβαρύνοντας περαιτέρω την κυκλοφοριακή λειτουργία του τοπικού οδικού δικτύου.

Πίνακας 1.1 Απόλυτος αριθμός οχημάτων στην Ευρώπη 1970-2013, Ευρωπαϊκή Επιτροπή Statistical Pocketbook 2015

	1970	1980	1990	1991	1992	1993
EU-28			163 628	168 476	172 906	175 795
EU-15	62 477	104 284	148 041	151 510	155 037	156 868
EU-13			15 587	16 965	17 869	18 926
1994	1995	1996	1997	1998	1999	2000
178 967	182 462	186 759	186 415	191 291	197 037	201 993
159 119	161 539	164 596	163 009	166 939	171 790	175 513
19 848	20 923	22 163	23 406	24 352	25 247	26 480
2001	2002	2003	2004	2005	2006	
EU-28	206 758	210 891	213 781	217 238	221 608	225 891
EU-15	179 303	182 283	184 246	186 466	189 611	192 764
EU-13	27 455	28 609	29 535	30 772	31 996	33 127
2007	2008	2009	2010	2011	2012	2013
231 027	235 387	237 573	240 427	243 450	245 906	247 997
195 932	197 481	198 863	200 953	202 794	204 142	205 037
35 095	37 906	38 710	39 473	40 655	41 765	42 960

Η αλήθεια είναι πως είναι πολύ δύσκολα μετρήσιμο, πρακτικά, το ποσοστό της ροής που οφείλεται στην κυκλοφορία η οποία αναζητά θέση στάθμευσης, γεγονός που εξηγείται από την σχετικά μικρή βιβλιογραφία που υπάρχει παγκοσμίως σχετικά με αυτό το θέμα. Μία από τις σημαντικότερες αναφορές στο ποσοστό της κίνησης που αντιστοιχεί σε οχήματα που αναζητούν θέση στάθμευσης είναι η σύνοψη 16 ερευνών σε διάφορες πόλεις του κόσμου μεταξύ 1927 και 2001 που έκανε ο καθηγητής Donald Shoup του UCLA (Shoup, 2006). Οι έρευνες έδειξαν ότι το παραπάνω ποσοστό κυμαίνεται μεταξύ 8-75% με μέσο όρο το 30%, ενώ ο μέσος χρόνος αναζήτησης κενής θέσης κυμαίνεται από 3.5 έως 14 λεπτά. Βεβαίως αυτή η πληροφορία είναι δύσκολο να γενικευτεί διότι, αφενός οι μετρήσεις αυτές την εποχή

που πραγματοποιήθηκαν εγκυμονούσαν κινδύνους ως προς την ακρίβειά τους και αφετέρου επειδή οι μετρήσεις έχουν γίνει σε σημεία της πόλης όπου αναμένεται να παρουσιάζεται τέτοια συμπεριφορά. Σημειωτέον, πρέπει να λαμβάνεται σοβαρά υπόψη ότι οι θέσεις στάθμευσης επί της οδού έχουν χαμηλό κόστος χρήσης στις παραπάνω μετρήσεις.

Η διαφορετική κρατική πολιτική χρέωσης των θέσεων στάθμευσης επί της οδού δημιουργεί επιπλέον προβλήματα στην γενίκευση αυτών των στατιστικών. Για παράδειγμα, αντίστοιχες μετρήσεις σε Ευρωπαϊκές χώρες του Βορρά, όπως το Βέλγιο και η Ολλανδία, όπου η χρέωση των θέσεων στάθμευσης επί της οδού είναι η ίδια με εκείνες των χώρων στάθμευσης, ο χρόνος αναζήτησης κενής θέσης αφορά μόνο το 30% του συνολικού φόρτου και είναι ίσος με 36 δευτερόλεπτα κατά μέσο όρο (Ommeren et al., 2012).

Ενδιαφέρον παρουσιάζει και μία αντίστοιχη μελέτη σε διάφορες περιοχές της Γαλλίας στην οποία υπολογίστηκε ότι ο μέσος χρόνος αναζήτησης κενής θέσης στάθμευσης ήταν 3.3 λεπτά για την Grenoble, 11.8 λεπτά για την Lyon και για το Παρίσι 10 λεπτά στο Commerce district και 7.7 λεπτά στο Saint-Germain district. Ακόμη, υπολογίστηκε ότι το ποσοστό των οχημάτων που αναζητά θέση στάθμευσης είναι το 5-10% της κυκλοφοριακής συμφόρησης με το ποσοστό να αυξάνεται έως και το 60% σε μικρές οδούς (Gantelet & Lefauconnier, 2006).

Εντυπωσιακά επίσης ποσοστά είναι αυτά που εμφανίζονται σε κάποια σημεία της Νέας Υόρκης. Συγκεκριμένα στην 7th Avenue στο Brooklyn 45% των οχημάτων αναζητούν θέσεις στάθμευσης στην ώρα αιχμής (Transportation Alternatives, 2007). Το αντίστοιχο ποσοστό στο SoHo District του Manhattan είναι 28% (Schaller Consulting, 2006).

Είναι λοιπόν σαφές, ότι το προαναφερθέν πρόβλημα υπάρχει στα σύγχρονα αστικά κέντρα και αποτελεί την γενεσιουργό αιτία και άλλων προβλημάτων. Οι κυκλοφοριακές επιπτώσεις αφορούν την αύξηση του μέσου χρόνου αναμονής και διαδρομής του χρήστη και την επιβάρυνση του τοπικού δικτύου σε επίπεδο κυκλοφοριακού συστήματος. Ακόμη προκύπτουν περιβαλλοντικές επιπτώσεις οι οποίες είναι απόρροια των πολλών παραπάνω χιλιομέτρων που διανύουν τα οχήματα όσο αναζητούν θέση στάθμευσης. Σύμφωνα με έναν απλό υπολογισμό κάθε θέση αντιστοιχεί σε 1825 παραπάνω μίλια από τους χρήστες το χρόνο (Shoup, 2006). Υποθέτοντας ότι το μέσο όχημα διανύει 5 μίλια/λίτρο βενζίνης τότε κάθε θέση στάθμευσης αντιστοιχεί σε 365 χαμένα λίτρα βενζίνης το χρόνο με ότι αυτό συνεπάγεται για το κόστος στο περιβάλλον. Ενδεικτικά παρατίθεται ο Πίνακας 1.2 με το μέσο όρο των εκπομπών των επιβατικών Ι.Χ. στην ατμόσφαιρα όπως αυτή

δημοσιεύτηκε σε σχετική έρευνα (U.S.E.P.A., 2008). Στην τελευταία στήλη έγιναν και οι υπολογισμοί του βάρους των εκπομπών για τα επιπλέον 1825 μίλια ανά θέση, ανά έτος.

Πίνακας 1.2 Ετήσιες επιπλέον εκπομπές των Ι.Χ που αναζητούν θέσεις Στάθμευσης

Ρύπος	Σύμβολο	Κατανάλωση σε g/mi	Κατανάλωση (g) στα extra 1825 μίλια/χρόνο
Συνολικοί Υδρογονάνθρακες	THC	1,077	1965,525
Συνολικοί Υδρογονάνθρακες + αλδεΐδες,-μεθάνιο,-αιθάνιο	VOC	1,034	1887,05
Μονοξείδιο του άνθρακα	CO	9,4	17155
Οξειδία του αζώτου	NOx	0,693	1264,725
Σωματίδια < 10 microns	PM10	0,0044	8,03
Σωματίδια < 2.5 microns	PM2.5	0,0041	7,4825
Διοξείδιο του άνθρακα	CO2	368,4	672330

1.2 ΤΕΧΝΟΛΟΓΙΚΗ ΕΞΕΛΙΞΗ ΑΥΤΟΚΙΝΗΤΩΝ-ΕΥΡΕΙΑ ΧΡΗΣΗ ΤΗΣ ΚΑΜΕΡΑΣ

Η συνεχής βελτίωση των τεχνολογικών μέσων καθιστούν αναγκαία την ραγδαία αλλαγή της βιομηχανίας των αυτοκινήτων και προσαρμογή της στα νέα δεδομένα. Συγκεκριμένα, οι παραδοσιακές έως σήμερα δυνάμεις της αυτοκινητοβιομηχανίας καλούνται να αντιμετωπίσουν τεχνολογικούς γίγαντες όπως η GOOGLE και η

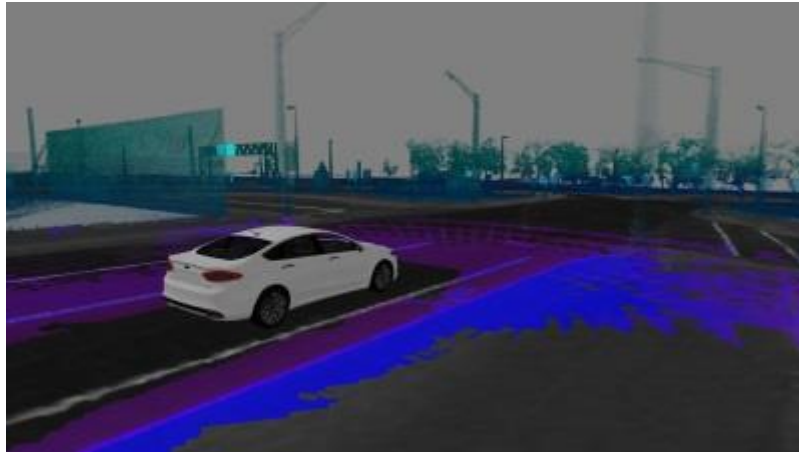
APPLE σύμφωνα με συνέντευξη του Mark Fields, διευθυντή της Ford, στο BBC (Ahmed, 2016).

Δεν είναι μυστικό ότι η αυτοκινητοβιομηχανία στρέφεται προς τα αυτόματα αυτοκίνητα, με τις τεχνολογικές εταιρείες να έχουν πρωταγωνιστικό ρόλο σε αυτές τις εξελίξεις. Ενδεικτικές είναι και οι συμμαχίες των εταιρειών είτε για την πραγματοποίηση αυτού του οράματος είτε για την πολιτική πίεση για να δημιουργηθούν οι συνθήκες πραγματοποίησής του (Newcomb, 2016). Βεβαίως η πλήρης αυτοματοποίηση των αυτοκινήτων απαιτεί μια σειρά αλλαγών στο παραδοσιακό αυτοκίνητο με μία από τις σημαντικότερες να είναι ο τρόπος με τον οποίο το αυτοκίνητο «βλέπει» το περιβάλλον του.

Οι επικρατούσες τάσεις σήμερα είναι δύο: η χρήση καμερών ή η χρήση LIDAR. Στην πρώτη περίπτωση χρησιμοποιείται ένας αριθμός καμερών στο όχημα σε συνδυασμό με ορισμένους σένσορες ή/και ραντάρ για την αποκωδικοποίηση του εξωτερικού χώρου ενώ στην δεύτερη περίπτωση οι κάμερες αντικαθίστανται από ένα προηγμένο σύστημα 64 laser το οποίο μετρά αποστάσεις με στόχο την δημιουργία ενός 3D χάρτη 360 μοιρών με ακρίβεια έως και 200 μέτρα μακριά (Google, 2016). Η δεύτερη μέθοδος είναι ιδιαίτερα ακριβή και αυτό είναι και το μεγαλύτερό της μειονέκτημα.

Έρευνες γίνονται και προς τις δύο κατευθύνσεις με τις αυτοκινητοβιομηχανίες και τις τεχνολογικές εταιρείες να παίρνουν ήδη θέση. Η Google είναι ο μεγαλύτερος υποστηρικτής της χρήσης LIDAR (Google, 2016) μαζί με αυτοκινητοβιομηχανίες όπως η Ford η οποία εξετάζει την χρήση του συστήματος αυτού ακόμα και το βράδυ που οι κάμερες υπολειτουργούν χωρίς φωτισμό (FORD(a), 2016). Μία αναπαράσταση της τρισδιάστατης απεικόνισης του περιβάλλοντος που αντλαμβάνεται το όχημα φαίνεται στην Εικόνα 1.1. Ο πάροχος του συστήματος LIDAR με τον οποίο συνεργάζονται οι παραπάνω εταιρείες είναι η Velodyne. Θερμοί υποστηρικτές της χρήσης καμερών είναι η TESLA, η Volvo, η GM και άλλες εταιρείες οι οποίες ήδη δοκιμάζουν τα αυτοκίνητά τους με ειδικές άδειες σε δρόμους πόλεων. Ο διευθύνων σύμβουλος της TESLA, Elon Musk, έχει δημοσίως τοποθετηθεί κατά του LIDAR χαρακτηρίζοντάς το σε συνέντευξή του περιττό για τα αυτοκίνητα εφόσον η TESLA, μπορεί να έχει αντίστοιχες επιδόσεις με έναν πολύ πιο οικονομικό συνδυασμό ραντάρ, υπερήχων και καμερών που έχουν την ίδια σχεδόν ανάλυση με αυτές των κινητών τηλεφώνων (Tillemann & McCormick, 2016). Στην ίδια κατεύθυνση κινείται και η Volvo η οποία στοχεύει σε μία 360 μοιρών οπτική με συνδυασμό ραντάρ, laser και καμερών, 4άρων για τα κοντινά αντικείμενα και μια τριπλής εστίασης για τα απρόβλεπτα εμπόδια (VOLVO, 2015). Και οι δύο παραπάνω εταιρείες συνεργάζονται με την Mobileye, μια Ισραηλινή εταιρεία computer-vision, η

οποία ισχυρίζεται ότι το λογισμικό της ανιχνεύει οχήματα μέχρι και σε απόσταση 230 ποδιών χρησιμοποιώντας μία τυπική κάμερα (Tillemann & McCormick, 2016).

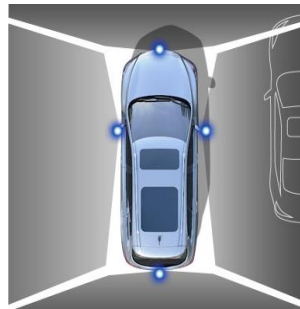


Εικόνα 1.1 Αναπαράσταση του περιβάλλοντος ενός οχήματος FORD με την χρήση LIDAR το βράδυ

Είναι λοιπόν σαφές ότι η αυτοκινητοβιομηχανίες κινούνται προς την κατεύθυνση της αυτοματοποίησης και είναι ήδη ορατές πολλές από τις προσπάθειές τους να τις εντάξουν και στα υπάρχοντα μοντέλα τους. Αναφέρονται παρακάτω ορισμένα από τα πιο ευρέως διαδεδομένα οχήματα σήμερα, των οποίων τα συστήματα εμπεριέχουν την χρήση κάμερας:

- 2016 BMW 750i xDrive - Μία stereo camera κοντά στον καθρέφτη δίνει την αίσθηση του βάθους της BMW για τον προσδιορισμό πεζών και σημάτων των λωρίδων κυκλοφορίας (Sherman, 2016)
- 2015 Infiniti Q50S - Τέσσερις φωτογραφικές μηχανές μία σε κάθε πλευρά του αυτοκινήτου όπως φαίνονται στην Εικόνα 1.2 για τον σχηματισμό μίας 360 μοιρών οπτική του περιβάλλοντος με πιθανότητα να αυξηθούν άμεσα (INFINITI, 2016)
- 2015 Mercedes-Benz S65 AMG - Μία stereo camera πίσω από το παρμπρίζ για τον έλεγχο της λωρίδας κυκλοφορίας (Sherman, 2016)
- 2015 Tesla Model S P85D - Μία φωτογραφική μηχανή στο παρμπρίζ που σε συνδυασμό με έναν σένσορα επιτυγχάνουν σχεδόν 20/20 ορατότητα (Sherman, 2016)
- 2017 Ford Explorer SUV, 2016 Ford Edge – Μία 180 μοιρών κάμερα για χαμηλές ταχύτητες επάνω στη μάσκα η οποία έχει τη δυνατότητα να διώχνει τη σκόνη και τη λάσπη από το φακό της (FORD(b), 2016) (FORD(c), 2016),

- 2016 Ford F150 - Τέσσερις φωτογραφικές μηχανές, μία σε κάθε πλευρά του αυτοκινήτου για τον σχηματισμό μίας 360 μοιρών οπτικής του περιβάλλοντος, με πιθανότητα να αυξηθούν άμεσα (Birch, 2016)
- 2017 Jaguar XJ R SPORT AWD – Πέντε κάμερες με βασικό στόχο την διευκόλυνση του παρκαρίσματος (JAGUAR, 2016)



Εικόνα 1.2 Οι 4 κάμερες της INFINITI στο σύστημά της AROUND VIEW MONITOR

Όπως γίνεται κατανοητό από τα παραπάνω, ακόμη και αυτοκινητοβιομηχανίες που πιστεύουν στη επικράτηση του LIDAR όπως η Ford, δεν παραλείπουν να εξοπλίσουν τα αυτοκίνητά τους με κάμερες. Ακόμη και στα πλήρως αυτοματοποιημένα αυτοκίνητα που σχεδιάζονται από τους υπέρμαχους του LIDAR, βασική προϋπόθεση αποτελεί η ύπαρξη τουλάχιστον μίας κάμερας για να ανιχνεύει κυκλοφοριακές πληροφορίες όπως η σήμανση κάθετη ή οριζόντια (Google, 2016).

Σύμφωνα με όλα τα παραπάνω η τεχνολογική εξέλιξη έχει ήδη αρχίσει να μεταβάλλει την εικόνα του παραδοσιακού αυτοκινήτου. Οι κάμερες πολύ σύντομα θα μετατραπούν από πολυτέλεια σε ανάγκη. Οι Η.Π.Α. ήδη θέσπισαν μέσω του αρμόδιου οργανισμού NHTSA, κανόνα που υποχρεώνει όλα τα νέα οχήματα με βάρος μικρότερο των 4.5 τόνων από το Μάιο του 2018 να έχουν στον εξοπλισμό τους μία κάμερα στο πίσω μέρος του οχήματος με οπτικό πεδίο 3x6 μέτρα (NHTSA, 2014). Με φόντο το πλήρως αυτοματοποιημένο αυτοκίνητο κάθε ενδιάμεσο βήμα θα απαιτεί την χρήση καμερών για ποικίλες διεργασίες. Ακόμη και στην περίπτωση της επικράτησης του LIDAR επί των καμερών, γεγονός αμφίβολο, οι κάμερες θα εξακολουθούν να απαιτούνται απλά σε μικρότερο βαθμό. Σε αντίθετη περίπτωση όλες οι διαδικασίες του αυτοκινήτου θα ξεκινούν από την αναπαράσταση του περιβάλλοντος από τις κάμερες. Επομένως, και στα δύο ενδεχόμενα η χρήση της κάμερας από τις αυτοκινητοβιομηχανίες στο μέλλον θεωρείται αυτονόητη.

1.3 ΑΠΟ ΤΗΝ ΥΠΟΛΟΓΙΣΤΙΚΗ ΟΡΑΣΗ ΣΤΗ ΣΥΓΚΟΙΝΩΝΙΑΚΗ ΠΛΗΡΟΦΟΡΙΑ

Η έννοια της εικόνας και των οπτικών μέσων αποτελούν την νέα πραγματικότητα και αναπόφευκτα αναπόσπαστο κομμάτι των μελλοντικών δομών των αυτοκινήτων. Η φωτογραφική κάμερα θα πρωταγωνιστεί επομένως στα νέα αυτοκίνητα και, συνεπώς, και όλες οι τεχνικές που σχετίζονται είτε με την ανάλυσή της, είτε με αναγνώριση αντικειμένων, είτε με τον συνδυασμό της με άλλα πολυμέσα.

Η χρήση των καμερών στα αυτοκίνητα μπορεί να είναι πολυδιάστατη. Ήδη οι κάμερες χρησιμοποιούνται ώστε να δίνεται μία καλή εικόνα στον οδηγό για τον περιβάλλοντα χώρο ή για νεκρά σημεία και μελετάται η οπτική των 360 μοιρών. Επίσης χρησιμοποιούνται ώστε να αναγνωρίζουν την σήμανση, όπως οι γραμμές των λωρίδων, όρια ταχυτήτων κ.α. Στην ίδια λογική μπορούν να επιλύσουν και άλλα προβλήματα τα οποία δεν αφορούν την λειτουργία του ίδιου του οχήματος αλλά την πρόθεση του οδηγού.

Για την ουσιαστική χρησιμοποίηση όλων όσων προσφέρει η τεχνολογική εξέλιξη και συγκεκριμένα η λήψη εικόνας σε πραγματικό χρόνο του υπό μελέτη συστήματος, προς όφελος του συγκοινωνιακού συστήματος σε θέματα ελέγχου και διαχείρισης της ζήτησης και των κυκλοφοριακών συνθηκών, απαιτούνται προηγμένοι αλγόριθμοι οι οποίοι να μετατρέπουν την πληροφορία της εικόνας σε πληροφορία χρήσιμη για το σύστημα και τον χρήστη. Όπως αναφέρουν οι Tian et al. (Tian et al., 2015), ενώ το απαραίτητο τεχνολογικό υπόβαθρο για την παρακολούθηση των συνθηκών σε μια πόλη είναι ήδη σε αρκετά υψηλό επίπεδο, οι προκλήσεις σε ότι αφορά στην παροχή πληροφόρησης στους χρήστες και στο σύστημα είναι σημαντικές (σύνθεση κυκλοφορίας, πολύπλοκα περιβάλλοντα, εμπόδια και προσωπικά δεδομένα κτλ) και προϋποθέτουν την ανάπτυξη προχωρημένων μαθηματικών πρότυπων για την αντιμετώπισή τους. Η σχετική βιβλιογραφία, η οποία θα αναλυθεί σε επόμενο κεφάλαιο, καταδεικνύει ότι η ανάπτυξη αλγορίθμων για την εξόρυξη πληροφορίας από δεδομένα συλλεγμένα από κινούμενες κάμερες, αποτελεί ένα από τα σημαντικότερα ερευνητικά πεδία της ανάπτυξης ευφυών συγκοινωνιακών εφαρμογών (Huang & Chen, 2013).

1.4 ΣΚΟΠΟΣ ΤΗΣ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

Η παρούσα διπλωματική εργασία έχει ως στόχο την ανάπτυξη προτύπων μηχανικής μάθησης για την αναγνώριση των ελεύθερων θέσεων στάθμευσης σε αστικά οδικά δίκτυα από βίντεο που προκύπτουν από τις κάμερες που υπάρχουν ή θα υπάρχουν πάνω στα οχήματα. Για τον σκοπό αυτό, αναπτύσσονται, εκπαιδεύονται και αξιολογούνται συνελκτικά νευρωνικά με χρήση φωτογραφιών από κάμερες σε κινούμενο όχημα. Η συλλογή δεδομένων θα βασιστεί σε εξοπλισμό που είναι προσιτός και διαθέσιμος σε ευρεία γκάμα οχημάτων.

Στο πλαίσιο της έρευνας, παρουσιάζεται και αξιολογείται μια ολοκληρωμένη μεθοδολογία συσχέτισης της θέσης της κάμερας πάνω στο όχημα με την απόσταση που αποτυπώνεται στις παραγόμενες φωτογραφίες και κάτ' επέκταση την ύπαρξη ή όχι ελεύθερης θέσης παρά το κράσπεδο. Σε δεύτερο στάδιο αναπτύσσεται και κωδικοποιείται στην υπολογιστική πλατφόρμα R μια μεθοδολογία μετατροπής των φωτογραφιών σε πίνακες δεδομένων και επεξεργασίας τους, ώστε να διατηρηθεί η χρήσιμη πληροφορία για την πρόβλεψη ελεύθερης θέσης στάθμευσης.

Στη συνέχεια, αναπτύσσονται συνελκτικά νευρωνικά, τα οποία βελτιστοποιούνται ως προς τη δομή και τις παραμέτρους εκπαίδευσης τους, με στόχο την ελαχιστοποίηση του παραγόμενου σφάλματος γενίκευσης.

Τέλος, διερευνάται ο τρόπος εφαρμογής των προτύπων που αναπτύχθηκαν, ώστε να προκύπτουν χρήσιμες για τους μετακινούμενους πληροφορίες σε πραγματικό χρόνο.

1.5 ΔΟΜΗ ΤΗΣ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

Παρουσιάζεται συνοπτικά παρακάτω η διάταξη των κεφαλαίων της παρούσας διπλωματικής εργασίας και το περιεχόμενό τους:

Κεφάλαιο 1

Εισαγωγικό κεφάλαιο της διπλωματικής εργασίας με σκοπό να αναδειχθεί το πρόβλημα της στάθμευσης στα μεγάλα αστικά κέντρα τον 21^ο αιώνα και οι συνέπειές του. Ακόμη παρουσιάζεται η επικράτηση της φωτογραφίας ως το μέσο αναγνώρισης του περιβάλλοντα χώρου των αυτοκινήτων. Οι δύο πραγματικότητες όπως αναλύονται συμβάλλουν στην τεκμηρίωση του σκοπού της παρούσας διπλωματικής εργασίας. Τέλος παρατίθεται η δομή της διπλωματικής εργασίας.

Κεφάλαιο 2

Περιλαμβάνει την βιβλιογραφική ανασκόπηση των ερευνών που έχουν γίνει σε αντίστοιχη θεματολογία. Συγκεκριμένα εξετάζονται διαφορετικές μέθοδοι προσέγγισης του προβλήματος της εύρεσης κενής θέσης στάθμευσης χωρίς την χρήση της φωτογραφίας. Στην συνέχεια ακολουθεί μία βιβλιογραφική ανασκόπηση για την προσέγγιση του προβλήματος με μεθόδους ανάλυσης των φωτογραφιών και αιτιολογείται η καταλληλότητα των συνελκτικών νευρωνικών δικτύων.

Κεφάλαιο 3

Αναλύεται διεξοδικά η μεθοδολογία που θα ακολουθηθεί για την ολοκλήρωση της διπλωματικής εργασίας. Ειδικότερα, διατυπώνεται αρχικά το πρόβλημα και επεξηγείται η επιλογή του λογισμικού, οι σχετικοί περιορισμοί και ο τρόπος επίλυσης του προβλήματος, ενώ δίνεται και το θεωρητικό υπόβαθρο που απαιτείται για την κατανόηση της μεθοδολογίας. Ακόμη αναλύεται ο τρόπος που θα εκπαιδευτούν και θα αξιολογηθούν τα συνελκτικά νευρωνικά δίκτυα που αναπτύσσονται στη συνέχεια.

Κεφάλαιο 4

Αναλύεται η συλλογή των δεδομένων και η προ-επεξεργασία τους για το συγκεκριμένο πρόβλημα. Εξετάζονται πιθανές θέσεις της κάμερας και διαφορετικά

οπτικά πεδία. Επίσης εξετάζονται οι μορφές των εικόνων, το μέγεθός τους αλλά και το βάρος τους και η επάρκειά τους. Τέλος γίνεται ο επιμερισμός σε απλούστερα προβλήματα σύμφωνα με την μεθοδολογία.

Κεφάλαιο 5

Παρουσιάζεται η πορεία επίλυσης του συγκεκριμένου προβλήματος όπου εξηγείται κάθε επιλογή του μελετητή. Σε αυτό το κεφάλαιο γίνεται η εκπαίδευση των συνελκτικών νευρωνικών δικτύων όπου επιλέγεται η μορφή τους, οι παράμετροί τους και παρουσιάζονται τα εκάστοτε αποτελέσματα. Ακολουθεί στο πέρας κάθε τρεξίματος μία στατιστική ανάλυση για καλύτερη επεξήγηση των αποτελεσμάτων. Τέλος, παρουσιάζεται μία ολοκληρωμένη μελέτη για τον τρόπο εφαρμογής του μοντέλου. Συγκεκριμένα δίνονται διευκρινίσεις σχετικά με τις θέσεις που απαγορεύεται η στάθμευση, την έγκαιρη διαδικτυακή πληροφόρηση και την συνεργασία του συστήματος με άλλα που συμβάλλουν στην επίλυση χωρικών και άλλων ποιοτικών ζητημάτων.

Κεφάλαιο 6

Περιλαμβάνεται μία σύνοψη των πιο σημαντικών αποτελεσμάτων του κεφαλαίου 5 και ακολουθούν ποιοτικά συμπεράσματα. Ακόμη γίνονται προτάσεις για περαιτέρω έρευνα που θα συμπλήρωνε την υπάρχουσα, κυρίως μεταβάλλοντας τις προϋποθέσεις που αναφέρονται στο κεφάλαιο 3.

Κεφάλαιο 7

Παρατίθεται η πλήρης βιβλιογραφία που χρησιμοποιήθηκε για την πραγματοποίηση της παρούσας διπλωματικής εργασίας.

2. ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

Αυτό το Κεφάλαιο αφορά την βιβλιογραφική ανασκόπηση του θέματος που κατ' αρχήν συζητήθηκε στο Κεφάλαιο 1 με σκοπό την σύγκριση παρόμοιων μεθόδων που χρησιμοποιήθηκαν για την επίλυση του βασικού αυτού θέματος στάθμευσης και την επιβεβαίωση ότι η παρούσα διπλωματική θα αποτελέσει μία σημαντική προσθήκη στις υπάρχουσες προσπάθειες.

Το παρόν Κεφάλαιο χωρίζεται σε δύο μέρη: Στο πρώτο, περιγράφονται εναλλακτικές προσεγγίσεις για την επίλυση του προβλήματος της εύρεσης κενής θέσης στάθμευσης από άλλους ερευνητές χωρίς την χρήση της φωτογραφίας. Στο δεύτερο περιγράφονται οι διαφορετικοί τρόποι προσέγγισης και ανάλυσης της φωτογραφίας για την επίλυση του ίδιου προβλήματος και αναδεικνύεται η καταλληλότητα των συνελκτικών νευρωνικών δικτύων.

Είναι αξιοσημείωτο να αναφερθεί ότι οι οπτικές αναλύσεις των διαφορετικών μεθόδων ποικίλουν ως προς την λήψη τους (πανοραμική, πλαϊνή, μπροστινή). Σε κάθε περίπτωση θα εξεταστεί όχι μόνο η ακρίβεια του προτεινόμενου μοντέλου αλλά και οι συνθήκες υπό τις οποίες αποδίδει στο συγκεκριμένο βαθμό. Ακόμη θα εξεταστεί η εφαρμοσιμότητά του.

2.1 Ο ΠΛΗΘΟΠΟΡΙΣΜΟΣ ΣΤΑ ΣΥΓΚΟΙΝΩΝΙΑΚΑ ΣΥΣΤΗΜΑΤΑ

Οι Estellés-Arolas και González Ladrón-de- Guevara (Estellés-Arolas & González Ladrón-de-Guevara, 2012) ορίζουν τον πληθοπορισμό ως εξής:

“Ο πληθοπορισμός είναι μία μορφή συλλογικής διαδικτυακής δραστηριότητας στην οποία ένα άτομο, ένα ίδρυμα, ένας μη κερδοσκοπικός οργανισμός ή μία εταιρεία προτείνει σε μία ομάδα ατόμων με ποικίλλες γνώσεις, ετερογένεια και αριθμό, μέσω μίας ανοιχτής πρόσκλησης, να αναλάβουν εθελοντικά μια εργασία. Οι χρήστες λαμβάνουν την ικανοποίηση κάποιας ανάγκης τους, είτε αυτή είναι οικονομική, είτε κοινωνική αναγνώριση, προσωπική ικανοποίηση, ανάπτυξη ατομικών ικανοτήτων σε κάποιο τομέα, ενώ ο εκκινητής της πρωτοβουλίας (πληθοποριστής) αποκτά και χρησιμοποιεί προς όφελός του, αυτά που έχει συνεισφέρει ο χρήστης στο εγχείρημα, τα οποία εξαρτώνται από τη δραστηριότητα που έχει αναλάβει ο χρήστης.”

Ο πληθοπορισμός είναι βασική έννοια των ευφών συγκοινωνιακών συστημάτων και αρκετά διαδεδομένη σε θέματα στάθμευσης. Μία ενδιαφέρουσα έρευνα (Chen et al., 2012), μελετά ακόμη και τον τρόπο άντλησης της πληροφορίας από πληθοπορισμό, ο οποίος είναι απόλυτα συσχετισμένος με το επίπεδο της παροχής της υπηρεσίας. Όσο φιλικότερο είναι το περιβάλλον της εφαρμογής στο χρήστη, τόσο καλύτερη αντιμετώπιση θα λάβει η εφαρμογή με αποτέλεσμα να βελτιώνεται η λειτουργία της αφού αυτή εξαρτάται από την επιλογή των χρηστών να δώσουν πληροφορίες.

Στη συνέχεια του κεφαλαίου θα αναλυθούν ορισμένες προσπάθειες που πραγματοποιήθηκαν από άλλους ερευνητές ή υπηρεσίες, πολλές από τις οποίες ανεξάρτητα της μεθόδου ανάλυσης που χρησιμοποιείται βασίζονται στην επιτυχή εφαρμογή του πληθοπορισμού.

2.2 ΜΕΘΟΔΟΙ ΧΩΡΙΣ ΤΗΝ ΧΡΗΣΗ ΦΩΤΟΓΡΑΦΙΑΣ

Είναι γεγονός ότι οι περισσότερες προσπάθειες για την αντιμετώπιση του προβλήματος της εύρεσης κενής θέσης λύνονται με ορισμένη διαχείριση και επεξεργασία φωτογραφιών. Βεβαίως έχουν γίνει αξιόλογες προσπάθειες αξιοποίησης και εναλλακτικών μέσων πέραν της φωτογραφίας. Η χρήση αισθητήρων (διαφόρων ειδών) ή/και παρκόμετρων, η χρήση ραντάρ και η χρήση υπερήχων αποφέρουν επίσης χρήσιμα και σχετικά ακριβή αποτελέσματα.

2.2.1 ΑΙΣΘΗΤΗΡΕΣ ΕΠΙ ΤΗΣ ΟΔΟΥ

Μία διαδεδομένη μεν ακριβή δε μέθοδος για την εύρεση κενών ή κατειλημμένων θέσεων στάθμευσης είναι η εγκατάσταση αισθητήρων στην άσφαλτο. Πολλές πόλεις ανά τον κόσμο έχουν δοκιμάσει κάποια παραλλαγή τέτοιου συστήματος με τα αποτελέσματα να είναι εμφανώς θετικά αλλά όχι πλήρως ικανοποιητικά. Η εγκατάστασή τους αλλά και η συντήρησή τους είναι επίσης ζητήματα που απαιτούν προσοχή και επιμέλεια. Με τους αισθητήρες γίνεται συνήθως και μία εγκατάσταση πυλών, ανά τακτά χωρικά διαστήματα, ώστε να μεταδίδεται η πληροφορία από τους αισθητήρες στην τοπική πύλη και από αυτήν σε ένα κεντρικό σύστημα επεξεργασίας ώστε να διατίθεται πίσω στον χρήστη. Επίσης συνηθίζεται η εγκατάσταση παρκόμετρων ανά κάποιο συγκεκριμένο αριθμό θέσεων, τα οποία επικοινωνώντας με τους αισθητήρες διασφαλίζουν ότι τα σταθμευμένα αυτοκίνητα έχουν πληρώσει το ποσό που τους αναλογεί. Ενδεικτικά παραδείγματα της παραπάνω μεθόδου είναι το FASTPRK, το Smart Santander, το Smart Parking και το SFpark.

1. Το FASTPRK (Worldsensing, n.d.) χρησιμοποιεί ένα μαγνητικό αισθητήρα για κάθε θέση στάθμευσης ο οποίος είναι ενσωματωμένος στην άσφαλτο και όλα τα προαναφερθέντα συστήματα. Το σύστημα έχει δοκιμαστεί στην Μόσχα, όπου ξεκίνησε με 1000 αισθητήρες το Νοέμβριο του 2012 οι οποίοι

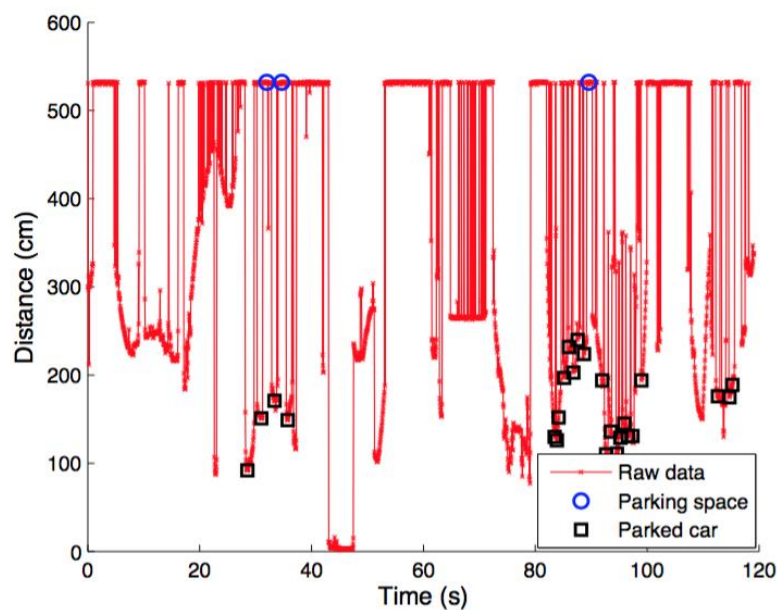
- έγιναν 50000 ως τα μέσα του 2014. Ο χρόνος αναζήτησης θέσης μειώθηκε κατά 35%.
2. Το Smart Santander (Smart Santander, 2016) χρησιμοποιεί ένα σιδηρομαγνητικό αισθητήρα για κάθε θέση στάθμευσης ο οποίος είναι ενσωματωμένος στην άσφαλο αλλά έχει ως προϋπόθεση την έννοια της έξυπνης πόλης. Αυτή τη στιγμή υπάρχουν 4 εγκαταστάσεις στο Βελιγράδι, στο Σανταντέρ, στο Γκίλντφορντ και στο Λούμπεκ.
 3. Το Smart Parking (Smart Parking, 2016) χρησιμοποιεί έναν αισθητήρα (Smart Eye) με υπέρυθρες για κάθε θέση στάθμευσης, ο οποίος μεταδίδει την πληροφορία σε μία ασύρματη πλατφόρμα η οποία με την σειρά της διεξάγει όλες τις προαναφερθείσες διαδικασίες. Σήμερα χρησιμοποιείται σε πάνω από 16 χώρες όπου έχουν εγκατασταθεί πάνω από 15000 αισθητήρες με το Westminster, το New Plymouth και το Ταυρο της Νέας Ζηλανδίας να αποτελούν τρία από τα πιο πετυχημένα τους σχέδια.
 4. Το SFpark (SFpark, 2012) είναι ένα σχέδιο του Αμερικανικού κράτους για την βελτίωση της διαχείρισης της στάθμευσης στο Σαν Φρανσίσκο και μέχρι σήμερα αριθμεί 8200 θέσεις με αισθητήρες. Χρησιμοποιούνται μαγνητικοί αισθητήρες και η μέθοδος που περιγράφηκε παραπάνω για την μετάδοση της πληροφορίας.

2.2.2 ΠΟΜΠΟΙ ΥΠΕΡΗΧΩΝ ΕΠΙ ΤΩΝ ΟΧΗΜΑΤΩΝ

Μία εναλλακτική μέθοδος εύρεσης θέσεων στάθμευσης χωρίς φωτογραφία είναι η χρήση υπερήχων. Σε αυτή την μέθοδο εγκαθίσταται ένας πομπός υπερήχων σε κάποιο υπολογισμένο σημείο επάνω στο όχημα και γίνεται καταμέτρηση της απόστασης των αντικειμένων γύρω από το όχημα με βάση το χρόνο επιστροφής των υπερηχητικών κυμάτων. Έτσι ανιχνεύονται παρκαρισμένα ή μη οχήματα. Με την βοήθεια ενός GPS δίνονται οι θέσεις που ανιχνεύονται τα εκάστοτε σήματα επομένως αντιστοιχείται πλήρως η γεωγραφική θέση με την ύπαρξη ή μη θέσης στάθμευσης.

Ένα τέτοιο σύστημα είναι το ParkNet που εμφανίστηκε το 2010 και χρησιμοποιεί ακριβώς την δομή που περιγράφηκε παραπάνω. Το σύστημα αυτό είναι σχεδιασμένο να λειτουργεί από οχήματα όπως ταξί ή κρατικά Μ.Μ.Μ. ώστε να μην χρειάζεται να εξοπλισθούν όλα τα οχήματα με ένα τέτοιο σύστημα (Marthur et al., 2010). Αντίστοιχη έρευνα έχει πραγματοποιηθεί και από άλλους (Liao et al., 2016), με την διαφορά ότι προτείνεται το σύστημα υπερήχων να είναι πάνω στο όχημα κάθε χρήστη

και να λειτουργεί υπό την μορφή πληθοπορισμού. Στο Διάγραμμα 2.1 παρουσιάζεται μία απεικόνιση αποστάσεων των εμποδίων από το όχημα σε σχέση με το χρόνο, όπου με τα κατάλληλα όρια γίνεται πρόβλεψη για το είδος τους. Τέλος, σε άλλη έρευνα έγινε χρήση των ήδη εγκατεστημένων αισθητήρων του οχήματος (Coric & Gruteser, 2013) με στόχο όχι μόνο την εύρεση της διαθεσιμότητας κενών θέσεων αλλά κυρίως την εύρεση των νόμιμων θέσεων στάθμευσης γιατί όπως αναφέρεται σε άλλες έρευνες η νομιμότητα θεωρείται δεδομένη ενώ στην πραγματικότητα δεν είναι.



Διάγραμμα 2.1 Χρήση υπερήχων για την ανίχνευση ελεύθερης θέσης στάθμευσης (LIAO, et al., 2016)

2.2.3 ΡΑΝΤΑΡ ΕΠΙ ΤΗΣ ΟΔΙΚΗΣ ΥΠΟΔΟΜΗΣ

Μία λιγότερο διαδεδομένη τεχνική αλλά εξίσου αποτελεσματική είναι τα ραντάρ από μεγάλο ύψος. Η Siemens (SIEMENS, 2015) πρωτοστατεί στην έρευνα πάνω στο Integrated Smart Parking Solution, ένα εγχείρημα πολλά υποσχόμενο. Συγκεκριμένα, ειδικά ραντάρ εγκαθίστανται στους στύλους φωτισμού του δρόμου και μπορούν να ανιχνεύουν και να αναλύουν ποικίλα χαρακτηριστικά της κυκλοφορίας, από την κίνηση των αυτοκινήτων ή των πεζών έως και όρια ταχυτήτων, θέσεις στάθμευσης

κ.λ.π. Ένα σημαντικό του πλεονέκτημα είναι ότι δεν επηρεάζεται από τις δυσμενείς καιρικές συνθήκες ή το φωτισμό και η ακρίβειά του είναι αρκετά υψηλή.

Μία ενδιαμέση λύση ερευνά η Ford στο νέο της σύστημα Parking Spotter με το οποίο αισθητήρες υπερήχων και ραντάρ ήδη εγκατεστημένα στα οχήματά της παρέχουν πληροφόρηση αξιοποιήσιμη σχετικά με την κατάσταση της στάθμευσης από τις περιοχές που διέρχονται (The Ford Motor Company Media Center, 2015).

2.2.4 ΣΥΓΚΡΗΤΙΚΗ ΑΞΙΟΛΟΓΗΣΗ ΜΕΘΟΔΩΝ ΧΩΡΙΣ ΤΗΝ ΧΡΗΣΗ ΤΗΣ ΦΩΤΟΓΡΑΦΙΑΣ

Αν και πολλές από τις παραπάνω μεθόδους βρίσκουν εφαρμογή σε κάποιες πόλεις του κόσμου ή έχουν πολλούς υποστηρικτές, όλες παρουσιάζουν κάποια μειονεκτήματα που τις καθιστούν μη επαρκείς.

Η χρήση των αισθητήρων είναι μία λύση της οποίας το ποσοστό ακρίβειας είναι πολύ μεγάλο αφού η κάλυψη τους από ένα όχημα δύσκολα δεν αναγνωρίζεται. Στα μειονεκτήματα αυτής της επιλογής ανήκουν η εγκατάσταση και η συνεχής επίβλεψη και συντήρηση του συστήματος. Ακόμη το συνολικό τους κόστος είναι δυσανάλογα μεγαλύτερο από την ακρίβεια που προσφέρουν σε σχέση με άλλες μεθόδους. Επίσης απαιτείται ένα πολύ καλό τοπικό και συνολικό σύστημα σύνδεσης για την άμεση μεταβίβαση της πληροφορίας, το οποίο δύσκολα διατηρείται σε αυτό το επίπεδο 24/7. Τέλος απαιτείται να είναι από πριν καθορισμένες οι θέσεις στάθμευσης διότι σε κάθε μία αντιστοιχεί ένας αισθητήρας, το οποίο δεν συμβαίνει σε κάθε περίπτωση. Για τους παραπάνω λόγους η εν λόγω μέθοδος δεν μπορεί να χρησιμοποιηθεί παρά τοπικά.

Η χρήση των πομπών υπερήχων αποτελεί μία υποσχόμενη λύση με ακρίβειες τις τάξεως του 90%. (Coric & Gruteser, 2013) σε αστικά κέντρα. Παρ'όλα αυτά, η ακρίβεια αυτή δεν ισχύει σε περιοχές στις οποίες είναι πιθανόν να εμφανισθούν και άλλα εμπόδια το οποίο την καθιστά αρκετά ευαίσθητη. Ακόμη, η πληροφορία που παρέχεται δύσκολα μπορεί να μετουσιωθεί και σε άλλα συμπεράσματα σε αντίθεση με την φωτογραφία η οποία δίνει πολλές αξιοποιήσιμες πληροφορίες με μία μέτρηση.

Η χρήση των ραντάρ είναι πανοραμική και πολυδιάστατη. Αυτό αποτελεί και το πλεονέκτημα αλλά και το μειονέκτημα της μεθόδου. Από την μία, πραγματοποιούνται πολλές δραστηριότητες ταυτόχρονα και υπάρχει συνολικότερη εικόνα με καλό

ποσοστό ακρίβειας. Από την άλλη, μία τέτοια λύση, όσον αφορά την στάθμευση, είναι χρήσιμη κυρίως για χώρους στάθμευσης όπου η πανοραμική λήψη ενδείκνυται και οι θέσεις είναι από πριν καθορισμένες. Επομένως, η εφαρμοσιμότητά του στις αρτηρίες ενός αστικού κέντρου δεν κρίνεται σκόπιμη.

Συνοψίζοντας, οι παραπάνω μέθοδοι αποτελούν πολύ θετικές προσπάθειες επίλυσης του προβλήματος αλλά δεν επαρκούν. Είναι προφανές, ότι απαιτείται η δύναμη της φωτογραφίας για να λύσει ορισμένα από τα προβλήματα των έως τώρα μεθόδων.

Στον πίνακα της επόμενης σελίδας συνοψίζονται οι περισσότερες από τις μεθόδους που επεξηγήθηκαν στο Κεφάλαιο 2.2, έτσι όπως καταγράφηκαν σε σχετική δημοσίευση (Liao et al., 2016)

Πίνακας 2.1 Σύγκριση Εγχειρημάτων για στάθμευση επί της οδού χωρίς την χρήση της φωτογραφίας (LIAO et al., 2016)

On-street Parking Projects	Sensing Technologies	Remarks
SFpark [6], 2014	magnetometer >1 sensor needed per spot	a complete on/off-street solution; digging roads required
FASTPRK [7]	magnetic 1 sensor per spot	a company portfolio including sensing, analysis, open data interface and APP; digging roads required
Street Parking System [8], 2013	magnetic 1 sensor per spot	82 sensors are deployed at Shenzhen Institute of Advanced Technology; digging roads required
Smart Parking [9]	infrared unspecified number	a company portfolio including sensing, guiding, payment and management; digging roads required
Smart Santander [10], 2014	ferromagnetic 1 sensor per spot	part of a smart city project; digging roads required
Integrated Smart Parking [11], ongoing	radar <1 sensor needed per spot	sensing not only parking but also traffic flows; mounting overhead
Parking Spotter [12], ongoing	sonar or radar <1 sensor needed per spot	crowdsourcing sensing; Ford proprietary application
ParkNet [13], 2010	sonar <1 sensor needed per spot	crowdsourcing sensing; environmental fingerprinting to reduce GPS errors
ParkSense [14], 2013	mobile phone and WiFi 1 mobile phone per spot	use beacons between mobile phone and WiFi to infer parking status; presence of both mobile phone and WiFi infrastructure required

2.3 ΜΕΘΟΔΟΙ ΜΕ ΤΗΝ ΧΡΗΣΗ ΤΗΣ ΦΩΤΟΓΡΑΦΙΑΣ

Η πλειονότητα των ερευνητών τα τελευταία χρόνια, έχει στραφεί στην επεξεργασία των εικόνων ή την ανάλυσή τους με προηγμένα μοντέλα με σκοπό την εξαγωγή πληροφοριών από αυτές στην μορφή που θα τις αντιλαμβάνονταν ο ανθρώπινος εγκέφαλος.

Οι ποικίλες λήψεις των φωτογραφιών, το μέγεθός τους και τα χαρακτηριστικά τους (φωτεινότητα, χρώματα κ.λ.π.) είναι παράγοντες καθοριστικοί στον τρόπο χειρισμού τους και στην συμβολή τους στην εξαγωγή χρήσιμων συμπερασμάτων. Παρακάτω γίνεται μία ανασκόπηση των κυριότερων από αυτές τις μεθόδους.

2.3.1 ΑΝΙΧΝΕΥΣΗ ΣΧΗΜΑΤΩΝ

Η πανοραμική λήψη των θέσεων στάθμευσης εξασφαλίζει μεγαλύτερες επιφάνειες κάλυψης και καλύτερη εποπτεία της συνολικής εικόνας. Παρά τα προβλήματα που αναφέρθηκαν ήδη στην σελίδα 18, η πανοραμική ανίχνευση θέσεων ήδη σχηματισμένων στο δρόμο είναι η προσέγγιση που προτιμήθηκε από τους ερευνητές Matteo D'Aloia et al. (D'Aloia et al., 2015). Υποθέτοντας ότι κάθε παραλληλόγραμμη θέση στάθμευσης, έχει και ένα σχήμα εντός αυτής, με την βοήθεια ενός drone έλαβαν φωτογραφίες πανοραμικές και στόχος τους ήταν η ανίχνευση αυτού του σχήματος, το οποίο ουσιαστικά υποδηλώνει κενή θέση. Το μοντέλο που ανέπτυξαν έπρεπε να αντιμετωπίζει τόσο διαφορετικούς φωτισμούς όσο και διαφορετικές οπτικές γωνίες. Σε πρώτο στάδιο εισήγαγαν το σχήμα ως προς το οποίο θα γίνεται η σύγκριση και δημιούργησαν ένα προφίλ για αυτό. Κάθε νέα εικόνα μετατρέπεται από έγχρωμη σε ασπρόμαυρη, ύστερα σε διωνυμική και με χρήση κατάλληλων φίλτρων γίνεται ανίχνευση των ακμών του σχήματος που απομονώθηκε από την εικόνα. Το σχήμα αυτό αφού διορθωθεί ως προς το μέγεθός του, την φωτεινότητα, την κάλυψη και άλλες παραμέτρους μετατρέπεται σε ένα προφίλ συγκρίσιμο με αυτό του αρχικού σχεδίου. Αν τα δύο σχέδια έχουν ομοιότητα πάνω από ένα συγκεκριμένο ποσοστό,

τότε η θέση αναγνωρίζεται ως κενή. Το ποσοστό επιτυχημένης αναγνώρισης των σχημάτων αγγίζει το 99%.

Στην πραγματικότητα πάντως, ένα τέτοιο σύστημα παρότι τόσο ακριβές είναι πολύ δύσκολο να εφαρμοστεί στην πράξη, ακόμη και με τα drones, επειδή δεν έχουν όλες οι θέσεις στάθμευσης οριοθετηθεί, ούτε έχουν κάποιο σχέδιο εντός αυτών. Ακόμα και να είχαν όμως, τα σχήματα μπορεί να καλύπτονται από άλλα αντικείμενα πχ. μία εφημερίδα καθιστώντας το παραπάνω αποτέλεσμα παραπλανητικό.

2.3.2 ΧΡΗΣΗ ΚΙΝΗΤΩΝ ΤΗΛΕΦΩΝΩΝ ΚΑΙ ΔΙΑΔΙΚΤΥΟΥ ΜΕ ΠΡΟΑΙΡΕΤΙΚΗ ΤΗ ΧΡΗΣΗ VIDEO

Η λογική του πληθοπορισμού που αναφέρθηκε και στο Κεφάλαιο 2.1 αποσκοπεί στην συλλογή δεδομένων από πολλούς χρήστες ταυτόχρονα ώστε να αξιοποιείται η πληροφορία και να προσφέρεται πίσω στους χρήστες. Μία τέτοια εφαρμογή για να γίνεται με άμεση ανταπόκριση απαιτεί καλή σύνδεση με το διαδίκτυο και κινητό τηλέφωνο προηγμένης τεχνολογίας.

Υπάρχουν ήδη στην αγορά πολλές εφαρμογές κινητών οι οποίες απαιτούν από τον χρήστη να δώσει χειροκίνητα μια πληροφορία απαντώντας σε μία ερώτηση για να του δοθεί κάποια άλλη. Υπάρχουν όμως και εφαρμογές οι οποίες είναι σχεδιασμένες έτσι ώστε να μπορεί ο χρήστης να μεταδώσει εικόνα ή/και ήχο είτε χειροκίνητα, είτε αυτόματα. Το City Watcher π.χ. είναι μία εφαρμογή στην οποία ο χρήστης όταν προσέξει μία ύποπτη ή κακή συμπεριφορά, ένα ατύχημα ή κακές συνθήκες οδήγησης, ή υπάρξει κάποια ανάγκη μπορεί να επικοινωνήσει με τις αρμόδιες αρχές είτε με ήχο είτε με βίντεο της κατάστασης που παρακολουθεί (Medvedev et al., 2015).

Από την άλλη υπάρχουν εφαρμογές όπως το Park Master (Grassi et al., 2015) το οποίο βασίζεται στην αυτόματη μετάδοση της πληροφορίας. Συγκεκριμένα, όσο ο οδηγός κινείται, το κινητό τηλέφωνο πρέπει να είναι εντός της εφαρμογής, συνδεδεμένο στο διαδίκτυο και στερεωμένο κάτω από τον μπροστά καθρέφτη των οχημάτων. Η εφαρμογή καταγράφει σε βίντεο την διαδρομή του χρήστη και η πληροφορία δίνεται σε online server ο οποίος έχει εκπαιδευτεί να αναγνωρίζει οχήματα. Γνωρίζοντας εξ αρχής τον μέγιστο αριθμό οχημάτων που μπορούν να σταθμεύσουν σε μία οδό και τον αριθμό που μετρήθηκε από οχήματα παρκαρισμένα σε αυτήν, είναι δυνατός ο υπολογισμός των κενών θέσεων ανά οδό. Η προσέγγιση αυτή παρουσιάζει ιδιαίτερο ενδιαφέρον αφού στον χρήστη παρέχεται η εικόνα της

διαθεσιμότητας των θέσεων για πολλές οδούς κοντά στον προορισμό του. Επομένως μπορεί πλέον να ζυγίσει την πιθανότητα εύρεσης μιας θέσης σε μία οδό χωρίς να την χάσει από κάποιον άλλο χρήστη. Βέβαια δεν του δίνεται η ακριβής τοποθεσία της θέσης με αποτέλεσμα το σύστημα να έχει ένα σοβαρό βαθμό χωρικής ανακρίβειας.

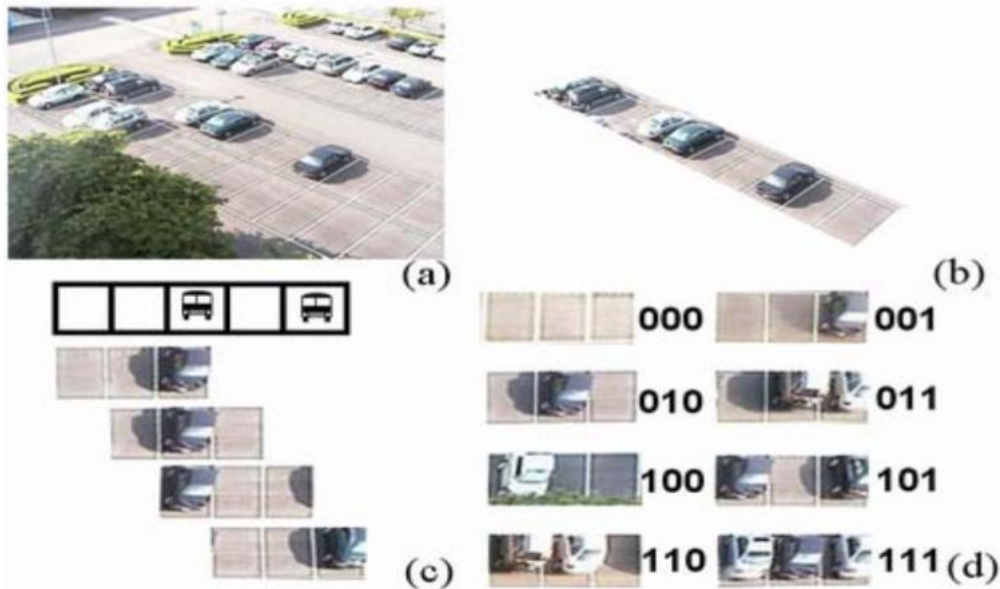
2.3.3 ΜΗΧΑΝΕΣ ΔΙΑΝΥΣΜΑΤΩΝ ΥΠΟΣΤΗΡΙΞΗΣ

Οι Μηχανές Διανυσμάτων Υποστήριξης (ΜΔΥ) επινοήθηκαν από τον Varnik και αποτελούν έναν αποδοτικό τρόπο επιβλεπόμενης μηχανικής μάθησης η οποία χρησιμοποιήθηκε και ευρέως σε προβλήματα κατηγοριοποίησης. Η μέθοδος χρησιμοποιεί μία συνάρτηση μετασχηματισμού για την προβολή των διανυσμάτων ιδιοτήτων σε ένα χώρο περισσότερων διαστάσεων επιδιώκοντας την εύρεση ενός υπερεπιπέδου, γραμμικού διαχωριστή, που να διαχωρίζει τις κατηγορίες αυτές με μέγιστο περιθώριο στο νέο διανυσματικό χώρο. Οι ΜΔΥ αποτελούν ένα τύπο μηχανής εκμάθησης, δηλαδή ένα σύστημα το οποίο κατηγοριοποιεί μία νέα είσοδο στη σωστή έξοδο αφού πρώτα έχει εκπαιδευτεί με μεγάλο αριθμό αντιστοιχιών εισόδου-εξόδου. Στην ίδια κατηγορία ανήκουν και τα νευρωνικά δίκτυα που θα αναλυθούν στην συνέχεια όπως και τα συνελκτικά δίκτυα τα οποία αποτελούν μία εξειδίκευσή τους. Τα ΜΔΥ έχουν χρησιμοποιηθεί σε πολλούς τομείς με μεγάλη επιτυχία. Επομένως εκτιμάται ότι μπορούν να χρησιμοποιηθούν και στο πρόβλημα της εύρεσης κενής θέσης στάθμευσης.

Συγκεκριμένα, σε έρευνα με πανοραμική λήψη χρησιμοποιήθηκε ένας ΜΔΥ ταξινομητής 8 κλάσεων οποίος σε συνδυασμό με την διόρθωση του Markov Random Field πλαισίου άγγιζε ακρίβεια 93,5% (Wu et al., 2009). Η αιτία που χρησιμοποιήθηκαν 8 κλάσεις αντί για 2 βρίσκεται στο γεγονός ότι η ανάλυση γίνεται με περιγραφικές συναρτήσεις των χρωμάτων που προκύπτουν από τα ιστογράμματα ταξινόμησης χρωμάτων. Οι σκιές των οχημάτων επιβάρυναν την ανάλυση με αποτέλεσμα να απαιτούνται τριάδες θέσεων στάθμευσης, επομένως $2^3=8$ συνδυασμοί, οι οποίοι είναι εμφανείς στην Εικόνα 2.1. Μετά την αποτύπωση τριάδων αποτελεσμάτων του μοντέλου, το σύστημα του Markov Random Field διόρθωνε όλα τα γειτονικά αντιφατικά αποτελέσματα που υπερκαλύπτονταν και διέφεραν.

Η μέθοδος αυτή, αν και είναι αποτελεσματική, βασίζεται σε παραδοχές οι οποίες την καθιστούν εφαρμόσιμη μόνο στα πλαίσια χώρων στάθμευσης διότι μόνο εκεί μπορούν να συνδυαστούν οι διαδοχικές θέσεις και να διορθωθεί το αποτέλεσμα τους

ώστε να επιτευχθεί τέτοια υψηλή ακρίβεια. Η πανοραμική λήψη αποτελεί επίσης θέμα προς εξέταση για την γενίκευση του μοντέλου.



Εικόνα 2.1 Τριπλέτες των θέσεων πανοραμικής λήψης με την χρήση ΜΔΥ (WU et al., 2009)

2.3.4 ΤΑΥΤΟΠΟΙΗΣΗ ΟΧΗΜΑΤΩΝ

Από την παραπάνω έρευνα και τα προβλήματα στάθμευσης στο USCD, εμπνεύστηκε ο True, (True, 2007), ο οποίος προσπάθησε να συνδυάσει τόσο την παραπάνω μεθοδολογία όσο και την ανίχνευση χαρακτηριστικών των σταθμευμένων οχημάτων για την αναγνώρισή τους, παραθέτοντας αποτελέσματα και των δύο μεθόδων. Η πρώτη αναλύθηκε και παραπάνω. Επομένως θα γίνει εστίαση στην δεύτερη προσέγγιση. Σε αυτήν εκτελείται ο αλγόριθμος της εύρεσης γωνιών του Harris και αφού πραγματοποιηθεί μία αναδιαστασιολόγηση των σημείων που καταγράφει, επειδή η προοπτική της λήψης μικραίνει τα αντικείμενα στο βάθος, δημιουργεί kernels 25x25 pixels, τα οποία αναζητά το μοντέλο για να ταυτοποιήσει τις νέες εικόνες ως οχήματα σταθμευμένα. Τα προβλήματα που αντιμετωπίζει αυτή η μέθοδος είναι ότι πολλές φορές καταγράφονται και χαρακτηριστικά που δεν αποτελούν στοιχεία των αυτοκινήτων αλλά σκουπίδια ή άλλα αντικείμενα που έτυχε να βρίσκονται στις θέσεις στάθμευσης την στιγμή της λήψης της φωτογραφίας. Επίσης σημαντικό πρόβλημα αποτελεί η ανίχνευση των γωνιών με το μοντέλο Harris, το

οποίο προκύπτει από την οπτική της κάμερας που φράζεται μερικές φορές από άλλα αντικείμενα ή ακόμα και οχήματα τα οποία βρίσκονται στις «μπροστά» θέσεις κρύβοντας αυτές που βρίσκονται στο βάθος.

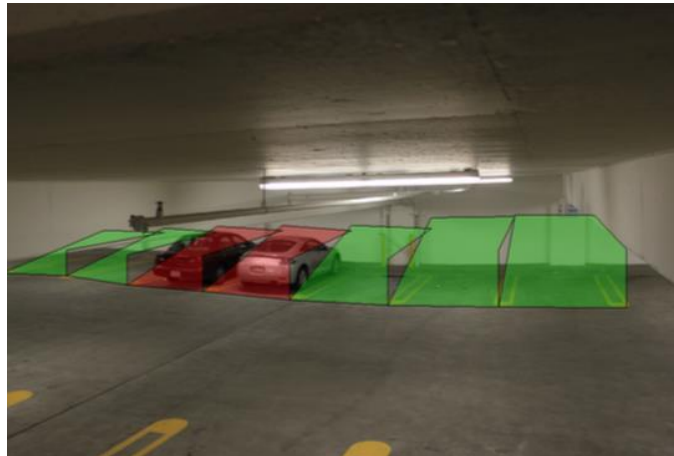
2.3.5 ΤΡΙΣΔΙΑΣΤΑΤΗ ΠΡΟΣΕΓΓΙΣΗ

Στις παραγράφους 2.3.1 και 2.3.3 αναλύθηκαν προσεγγίσεις ανάλυσης της εικόνας στις δύο διαστάσεις και από πανοραμική λήψη. Μία εναλλακτική προσέγγιση είναι η χρήση και των τριών διαστάσεων της φωτογραφίας οι οποίες μπορούν να γίνουν αντιληπτές με άλλους τρόπους προσέγγισης. Συγκεκριμένα έχουν δημιουργηθεί μεθοδολογίες οι οποίες αντιμετωπίζουν τις θέσεις στάθμευσης σαν τρισδιάστατα παραλληλεπίπεδα οι οποίες είτε ανεξάρτητες είτε σε συνδυασμό με άλλες μεθόδους αποφέρουν εντυπωσιακά αποτελέσματα.

Σχετική εργασία (Jones & Li, 2012), έχει πραγματοποιηθεί σε κλειστό χώρο στάθμευσης αλλά η λογική που ακολουθείται θα μπορούσε να εφαρμοστεί με τις κατάλληλες τροποποιήσεις και σε ανοιχτούς χώρους στάθμευσης με εξίσου μεγάλη επιτυχία. Αναλυτικότερα η μεθοδολογία τους χωρίζεται σε δύο βήματα. Αρχικά χρησιμοποιώντας απλά φίλτρα και τροποποιήσεις της εικόνας γίνεται ο εντοπισμός των γραμμών που διαχωρίζουν τις θέσεις στάθμευσης. Τα άκρα κάθε γραμμής δίνουν δύο σημεία αναφοράς. Ανά 2 γραμμές δίνουν 4 σημεία αναφοράς τα οποία αποτελούν μία θέση στάθμευσης όπως φαίνεται και στην Εικόνα 2.2. Δίνοντας το κατάλληλο ύψος δίνεται ο όγκος μέσα στον οποίο θα αναμενόταν ένα όχημα. Σε δεύτερη φάση δημιουργείται ένας αλγόριθμος που συνοψίζει τα βασικά χαρακτηριστικά της κενής θέσης και αυτά συγκρίνονται με την επόμενη εικόνα. Με αυτό τον τρόπο, για σταθερή κάμερα, ελέγχεται στο σημείο όπου αναμένεται όχημα εάν μεταβλήθηκαν τα βασικά χαρακτηριστικά εκείνου του τμήματος της αρχικής φωτογραφίας. Η έρευνα που έχει πραγματοποιηθεί είναι ιδιαίτερος προνοητική, καθώς δίνει λύσεις για πολλών ειδών πιθανά προβλήματα όπως παρεμπόδιση της οπτικής της κάμερας για ορισμένες θέσεις είτε από διερχόμενα είτε από σταθμευμένα οχήματα αλλά και για άλλου είδους προβλήματα όπως στάθμευση ανάμεσα σε δύο θέσεις, ύπαρξη πεζών και άλλες ενδιαφέρουσες περιπτώσεις. Με την επίλυση προβλημάτων όπως ο ασταθής φωτισμός και η χάραξη των θέσεων εκ των προτέρων, η μεθοδολογία αυτή θα είχε πολύ καλά αποτελέσματα και σε ανοιχτούς χώρους στάθμευσης.

Τέλος, μία προσέγγιση η οποία συνδυάζει τόσο την τρισδιάστατη προσέγγιση όσο και τις ΜΔΥ αλλά και διάφορες άλλες τεχνικές που αναφέρθηκαν παραπάνω είναι η

έρευνα των Huang C.-C. et al (Huang & Tran, 2015). Συνδυάζοντας όλες αυτές τις μεθόδους οι δύο ερευνητές πέτυχαν ακρίβειες τις τάξης του 99% για πανοραμική λήψη και αξιολόγηση των θέσεων στάθμευσης, υπό 4 διαφορετικές καιρικές συνθήκες.



Εικόνα 2.2 Απεικόνιση μίας δοκιμής των (JONES & LI, 2012)

2.3.6 ΣΥΝΕΛΙΚΤΙΚΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

Ο πιο σύγχρονος τρόπος ανάλυσης των φωτογραφιών είναι ένα συγκεκριμένο είδος νευρωνικών δικτύων, τα συνελικτικά δίκτυα. Η ιδιαιτερότητά τους, οφείλεται στην ικανότητά τους να αναγνωρίζουν συγκεκριμένα (πολλά και διαφορετικά) στοιχεία εντός των εικόνων πετυχαίνοντας με αυτό τον τρόπο μεγάλες ακρίβειες στην κατηγοριοποίηση του απεικονιζόμενου στην εικόνα. Οι καθαρά τεχνικοί λόγοι που αυτό συμβαίνει εξηγούνται αναλυτικά στο Κεφάλαιο 3.6.

Στο παρόν Κεφάλαιο γίνεται μία σύντομη ανασκόπηση στην ιστορία των συνελικτικών δικτύων και στην επιτυχία και ευρεία εφαρμογή που έχουν σήμερα, χάρις την ιδιαιτερότητα που μόλις αναφέρθηκε.

Η πρώτη επιτυχημένη εφαρμογή τους έγινε από τον Yann LeCun και άλλους ερευνητές το 1998 (LeCun et al., 1998) όπου προτάθηκε για πρώτη φορά η δομή του LENET-5 η οποία αναγνώριζε ψηφία αριθμών. Από τότε, και μέχρι τις αρχές της δεύτερης δεκαετίας αυτού του αιώνα, δεν υπήρξε και μεγάλη πρόοδος. Ήταν το 2012 όταν παρουσιάστηκε το AlexNet από τον Alex Krizhevsky, ένα δίκτυο παρόμοιο με

το LUNET-5, αλλά με περισσότερα επίπεδα και περισσότερα νευρώνια ενώ χρησιμοποιήθηκαν για πρώτη φορά και διαδοχικά επίπεδα συνέλιξης (Krizhevsky et al., 2012). Το 2014 η Google φέρνει πλέον μία επαναστατική ιδέα στο προσκήνιο που την ονόμασε Inception. Το βασικό προτέρημα της λύσης αυτής ήταν οι μείωση των παραμέτρων ως και 80% για την λειτουργία του δικτύου με ίδια αποτελέσματα. Άλλα γνωστά δίκτυα είναι το VGGNet, ZFNet, ResNet κ.λ.π. (Stanford, 2016).

Τα συνελκτικά δίκτυα αποτελούν σήμερα ένα από τα βασικότερα εργαλεία των τεχνολογικών εταιρειών, με τους τεχνολογικούς κολοσσούς να πρωτοστατούν στην ανάπτυξή τους και στην εφαρμογή τους σε ποικίλα νέα προβλήματα ανοίγοντας νέους ορίζοντες για όλη την ανθρωπότητα. Ο Yann LeCun, ο Διευθυντής του τμήματος Τεχνητής Νοημοσύνης του Facebook, σε συνέντευξή του στο Fortune εξηγεί πώς η αναγνώριση προσώπων στο Facebook λειτουργεί με την χρήση των συνελκτικών δικτύων, αλλά και τον απώτερο σκοπό της εταιρείας ο οποίος δεν είναι μόνο η αναγνώριση ενός προσώπου ανάμεσα σε άλλα, αλλά η ενσυναίσθηση (Higginbotham, 2015). Συγκεκριμένα, το Facebook επιθυμεί να γνωρίζει τα συναισθήματα που προκαλεί μία συγκεκριμένη φωτογραφία και να προτρέπει ή να αποτρέπει έναν χρήστη από το να την χρησιμοποιήσει. Η Google από την άλλη, έχει δημιουργήσει λογισμικά όπως το PlaNit (Δίκτυο εύρεσης της τοποθεσίας μιας φωτογραφίας σε όλο τον πλανήτη) ή το AlphaGo (Πρόγραμμα το οποίο παίζει το κινέζικο παιχνίδι Go), τα οποία λειτουργούν είτε εξολοκλήρου με συνελκτικά δίκτυα, είτε βασίζονται σε αυτά. Το PlaNit εντοπίζει 3,6% των εικόνων σε επίπεδο δρόμου, 10,1% σε επίπεδο πόλης, 28,4% σε επίπεδο χώρας και 48% σε επίπεδο ηπείρου (Weyand et al., 2016). Το AlphaGo αποτελεί το πρώτο πρόγραμμα που κερδίζει επαγγελματία ανθρώπινο παίκτη Go σε παιχνίδι με πλήρεις διαστάσεις, ενώ με αντίστοιχα προγράμματα η υπεροχή του ανέρχεται στο 99.8% (Silver D. et al, 2016). Ακόμη, η Microsoft έχει επίσης αναπτύξει μοντέλα που βασίζονται στην χρήση συνελκτικών δικτύων. Για παράδειγμα, έχει δημιουργήσει συνελκτικά δίκτυα, τα οποία «μαθαίνουν» σημασιολογικές αναπαραστάσεις και έτσι βελτιώνουν τις μηχανές αναζήτησής τους (Shen et al., 2014). Επιπλέον, μικρότερες εταιρείες και start-ups όπως η Affectiva, όλο και περισσότερο χρησιμοποιούν συνελκτικά νευρωνικά δίκτυα. Η συγκεκριμένη, σύμφωνα με το επίσημο site της, στο προϊόν της Affectiva, χρησιμοποιεί συνελκτικά νευρωνικά δίκτυα για την μέτρηση των συναισθημάτων των ανθρώπων, από φωτογραφίες των προσώπων τους, τραβηγμένες από κινητά τηλέφωνα ή κάμερες διαδικτύου. Είναι λοιπόν εμφανές, ότι τα συνελκτικά δίκτυα αποτελούν σήμερα την πιο μοντέρνα προσέγγιση για επιτυχή υπολογιστική όραση και για το λόγο αυτό, όχι μόνο εφαρμόζονται από πολλές εταιρείες, αλλά αναπτύσσονται περαιτέρω διότι ακόμα δεν έχουν πλήρως προσδιοριστεί όλες οι δυνατότητές τους.

2.3.7 ΒΑΣΙΚΑ ΣΥΜΠΕΡΑΣΜΑΤΑ

Η χρήση της φωτογραφίας είναι δεδομένη κάτι που έγινε ήδη σαφές από το Κεφάλαιο 1. Παρόλ' αυτά, στο Κεφάλαιο 2.2 έγινε μία ανάλυση των μεθόδων που μπορούν να επιλύσουν το πρόβλημα χωρίς την χρήση της φωτογραφίας με το κεφάλαιο 2.2.4 να συνοψίζει τα μειονεκτήματά τους. Στην συνέχεια στο Κεφάλαιο 2.3 αναλύθηκαν και επεξηγήθηκαν οι μέθοδοι που έχουν χρησιμοποιηθεί ως σήμερα με την χρήση φωτογραφίας. Αν και όλες οι μέθοδοι παρουσιάζουν επιμέρους προβλήματα που αναφέρονται στην εκάστοτε παράγραφο, ακολουθεί και μια ομαδοποίηση των μειονεκτημάτων που εμφανίζουν οι περισσότερες από αυτές και ο τρόπος αποφυγής τους στην παρούσα διπλωματική εργασία.

Ένα από τα μειονεκτήματα είναι η αντιμετώπιση του προβλήματος της στάθμευσης κυρίως πανοραμικά. Παρόλ' αυτά, με εξαίρεση τους χώρους στάθμευσης, οι αρτηρίες των πόλεων δεν προσφέρονται για πανοραμική λήψη και αντιμετώπιση αυτού του ζητήματος για πολλούς λόγους. Η κίνηση πεζών και άλλων αντικειμένων από τις θέσεις θα έδιναν εσφαλμένη πληροφορία ως προς την κάλυψη των θέσεων στάθμευσης κάτι που συναντάται σε πολύ μικρότερο βαθμό στους χώρους στάθμευσης από ότι στα κέντρα πόλεων λ.χ. Αφετέρου οι συνθήκες των πόλεων μπορεί να καθιστούν την ορατότητα της κάμερας από υψόμετρο πολύ κακή διότι είναι πιθανό να μεσολαβούν πολλά εμπόδια. Τέλος, ακόμη και να ήταν δυνατή και ακριβής η κάλυψη όλων των αστικών αρτηριών με κάμερες το εγχείρημα πάλι φαντάζει αδύνατο αφού και το κόστος του θα είναι τεράστιο αλλά εγείρονται και θέματα προσωπικών δεδομένων. Επομένως επιλέχτηκε να γίνει ανάλυση φωτογραφιών από την οπτική του οχήματος του χρήστη το οποίο έχει σαφώς διερευνηθεί λιγότερο στην βιβλιογραφία και δεν δημιουργεί προβλήματα όπως τα παραπάνω που παρουσιάζουν όσες μέθοδοι υιοθετούν την πανοραμική λήψη.

Ένα ακόμη μειονέκτημα που συναντάται ευρέως στις παραπάνω μεθοδολογίες είναι η ανάγκη χάραξης των θέσεων εκ των προτέρων. Τα μοντέλα αυτά είναι αδύνατο να λειτουργήσουν σε περιοχές όπου δεν έχει γίνει χάραξη των θέσεων ή χειρότερα σε θέσεις όπου έχει γίνει χάραξη αλλά δεν τηρείται. Σε αυτές τις μεθοδολογίες ανήκει η τρισδιάστατη προσέγγιση, οι μεθοδολογίες ανίχνευσης σχημάτων όπως και οι περισσότερες από τις ΜΔΥ, οι οποίες θα λειτουργούσαν ιδανικά σε προκαθορισμένους χώρους στάθμευσης. Αντιθέτως, οι μεθοδολογίες οι οποίες επικεντρώνονται στο όχημα και στα χαρακτηριστικά του, όπως η λύση που προτείνεται σε αυτή τη διπλωματική εργασία, αποφεύγουν τον παραπάνω περιορισμό στην εφαρμογή τους.

Ανακεφαλαιώνοντας, τα συμπεράσματα από τα Υποκεφάλαια 2.2 και 2.3 αφενός επιβεβαιώνουν την ανάγκη χρησιμοποίησης της φωτογραφίας και αφετέρου αποκαλύπτουν τα κενά άλλων μεθόδων τα οποία πρέπει να καλυφθούν για την επίτευξη καλύτερων αποτελεσμάτων. Η εφαρμογή των συνελκτικών δικτύων, των πιο σύγχρονων μοντέλων για ανίχνευση συγκεκριμένης κατατομής σε φωτογραφίες, αναμένεται να προσθέσει μία νέα διάσταση στην υπάρχουσα έρευνα, καλύπτοντας πολλά από τα κενά που υπάρχουν μέχρι σήμερα και πιθανώς να προσδώσει ένα πολύ καλό αποτέλεσμα υπό τις προϋποθέσεις που επεξηγήθηκαν και αναλύθηκαν παραπάνω. Επίσης δεδομένης της εξάρσης του πληθοπορισμού (Κεφάλαιο 2.1), το εγχείρημα για συλλογή πληροφοριών από μία πληθώρα χρηστών που επιδιώκεται στην παρούσα διπλωματική εργασία, μοιάζει να εντάσσεται αρμονικά στις σημερινές τάσεις.

3. ΜΕΘΟΔΟΛΟΓΙΑ

3.1 ΔΙΑΤΥΠΩΣΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ

Στην παρούσα διπλωματική εργασία γίνεται προσπάθεια για αντιμετώπιση του προβλήματος της εύρεσης θέσης στάθμευσης κοντά στον προορισμό του χρήστη με σκοπό τόσο την μείωση του χρόνου εύρεσης θέσης για το χρήστη όσο και την ελάφρυνση του τοπικού δικτύου από την αντίστοιχη δημιουργούμενη κυκλοφορία και τις αρνητικές της επιπτώσεις στο περιβάλλον. Η επίλυση του προβλήματος αφορά στην έγκαιρη και έγκυρη πληροφόρηση του χρήστη που αναζητεί θέση στάθμευσης. Η έγκαιρη πληροφόρηση εξασφαλίζεται από το διαδίκτυο μέσω των smartphones. Η έγκυρη πληροφόρηση για την ύπαρξη ή μη κάποιας κενής θέσης είναι το θεμελιώδες πρόβλημα, το οποίο πραγματεύεται η παρούσα διπλωματική. Το πρόβλημα προς επίλυση είναι τύπου 0/1 όπου 0 ονομάστηκαν οι κενές θέσεις και 1 οι κατειλημμένες.

3.2 ΠΑΡΑΔΟΧΕΣ

Η πολυπλοκότητα των δομών των αστικών κέντρων και των συνθηκών που επικρατούν σε αυτές καθιστούν σαφές ότι ένα μοντέλο, ιδίως στα πρώιμά του στάδια απαιτεί ορισμένες παραδοχές και περιορισμούς τόσο στην κατασκευή όσο και στην εφαρμογή του. Οι παραδοχές για το μοντέλο, που αναπτύχθηκε στην παρούσα εργασία είναι οι παρακάτω:

- Αίθριες καιρικές συνθήκες (ηλιοφάνεια ή συννεφιά)
- Απουσία καιρικών φαινομένων (βροχόπτωση, χιόνι, χαλάζι, ομίχλη)
- Μέγιστη ταχύτητα οχήματος 72 χμ/ώρα
- Προβλεπόμενη απόσταση κινούμενου οχήματος από το πεζοδρόμιο 3.5 μέτρα
- Η πρόβλεψη γίνεται κατά τη διάρκεια της ημέρας (όχι νύχτα)
- Το μέσο αυτοκίνητο έχει μήκος 4.5 μέτρα
- Μήκος 5 μέτρων απαιτείται για να παρκάρει το παραπάνω τυπικό όχημα
- Η αρτηρία έχει μία μόνο λωρίδα
- Οποιοδήποτε εμπόδιο (δίκυκλο, κάδος σκουπιδιών κλπ.) καταχωρείται ως κατειλημμένη θέση (1)
- Ικανότητα φωτογραφικής μηχανής τουλάχιστον 10 φωτ./δευτ.
- Η γεωγραφική θέση (X,Y) του οχήματος άρα και των θέσεων είναι γνωστά
- Οι θέσεις όπου απαγορεύεται η στάθμευση συμπεριλαμβανομένων και των διασταυρώσεων (και τα απαιτούμενα μήκη από αυτές σύμφωνα με τους κανονισμούς) είναι ήδη γνωστές.

3.3 ΕΠΙΛΟΓΗ ΜΕΘΟΔΟΥ ΕΠΙΛΥΣΗΣ

Στο κεφάλαιο 2 παρουσιάστηκαν εκτενώς οι μέθοδοι που χρησιμοποιούνται για την ανάλυση των φωτογραφιών και την συνεισφορά τους στην επίλυση προβλημάτων κατηγοριοποίησης. Αρχικά γίνεται αναφορά στις ΜΔΥ οι οποίες αποτελούν την πρώτη προσπάθεια εκμάθησης μηχανών που έχει παρουσιαστεί. Παρόμοια λογική έχουν και τα νευρωνικά δίκτυα τα οποία παρότι υπήρχαν εξίσου παλιά, μετά το 2006 επέστρεψαν πολύ δυναμικά στην επικαιρότητα. Ο λόγος είναι η επαναστατική χρήση ενός είδους νευρωνικών δικτύων, των συνελκτικών, τα οποία αποτελούν την βέλτιστη μέθοδο για την «όραση» του ηλεκτρονικού υπολογιστή. Αυτή είναι η αιτία της κατακόρυφης διάδοσής τους και της χρήσης τους από πολλές μεγάλες τεχνολογικές εταιρείες για ποικίλα προβλήματα. Η ίδια μεθοδολογία θα χρησιμοποιηθεί στην παρούσα διπλωματική για την επίλυση του προβλήματος κατηγοριοποίησης των θέσεων στάθμευσης σε κενές ή μη.

3.4 ΑΠΑΙΤΟΥΜΕΝΟ ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

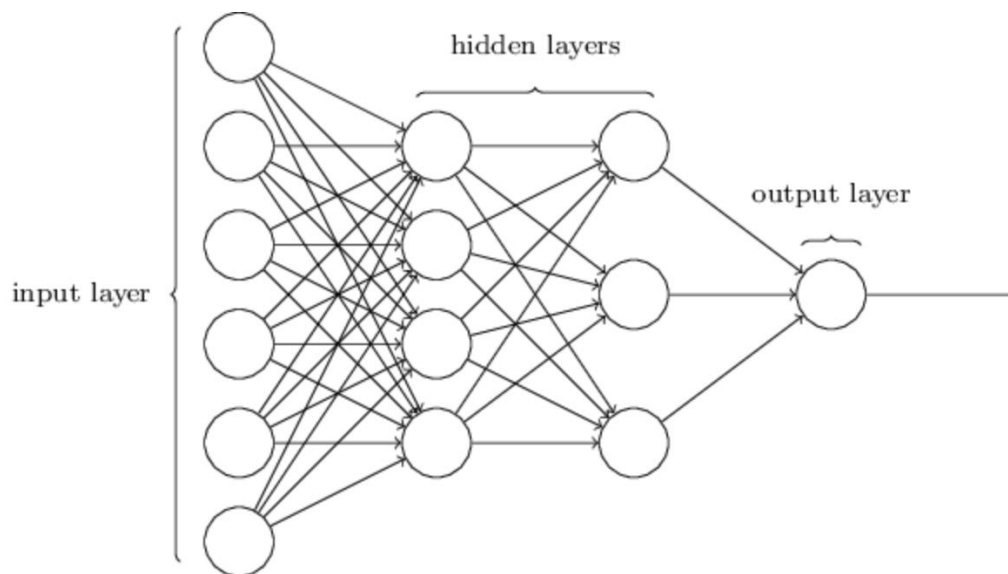
Η λειτουργία των συνελκτικών δικτύων είναι πολύ σύνθετη και ακόμη ερευνώνται πολλά χαρακτηριστικά της τα οποία παραμένουν άγνωστα. Τα συνελκτικά δίκτυα όπως έχει προαναφερθεί, ανήκουν στην οικογένεια εκμάθησης μηχανών δηλαδή ένα σύστημα το οποίο κατηγοριοποιεί μία νέα είσοδο στη σωστή έξοδο αφού πρώτα έχει εκπαιδευτεί με μεγάλο αριθμό αντιστοιχιών εισόδου-εξόδου. Η δομή του δικτύου και η εκπαίδευσή του παίζουν σημαντικό ρόλο στην επιτυχία του. Μέχρι σήμερα υπάρχουν ορισμένοι κανόνες, επεξηγήσεις και λογικές με τις οποίες δομούνται και εκπαιδεύονται τα δίκτυα αυτά.

Στο παρόν κεφάλαιο θα γίνει μία όσο γίνεται πιο σύντομη, αλλά και περιεκτική περιγραφή της λογικής των συνελκτικών δικτύων και των θεμελιωδών αρχών τους οι οποίες προέρχονται από τα απλά νευρωνικά δίκτυα. Αυτό γίνεται ώστε ο αναγνώστης να μην χρειάζεται να ανατρέξει σε άλλες πηγές, προκειμένου να αντιλαμβάνεται τις υποθέσεις της μεθοδολογίας της παρούσας διπλωματικής εργασίας. Συγκεκριμένα θα αναλυθούν τα στοιχεία τους και η δομή τους, η εκπαίδευση τους, οι παράμετροι που χρησιμοποιούνται, προβλήματα που αντιμετωπίζονται και συνήθειες πρακτικές που εφαρμόζονται από τους ερευνητές έτσι όπως θα χρησιμοποιηθούν και στην παρούσα διπλωματική εργασία για το πρόβλημα της στάθμευσης. Τα επόμενα κεφάλαια βασίζονται κυρίως σε δύο πηγές. Η πρώτη είναι το ηλεκτρονικό βιβλίο το Michael Nielsen το οποίο είναι ένα από τα λίγα τόσο περιεκτικά βιβλία με αυτή τη θεματολογία και το οποίο λόγω της δυναμικής εξέλιξης της μεθόδου είναι ακόμη υπό συγγραφή και δεν εκδίδεται σε εκτυπωμένη μορφή (Nielsen, 2015). Η δεύτερη είναι οι διαδικτυακές πληροφορίες-σημειώσεις του μαθήματος CS231n Convolutional Neural Networks for Visual Recognition του Stanford όπου υπάρχει επίσης εκτενής ανάλυση της μεθοδολογίας που ακολουθείται (Stanford, 2016). Το Κεφάλαιο 3.5 που ακολουθεί, βασίζεται σε αυτό το ηλεκτρονικό βιβλίο και οποιαδήποτε άλλη πηγή προκύπτει στο κεφάλαιο αυτό θα αναφέρεται ξεχωριστά. Το Κεφάλαιο 3.6 βασίζεται στις σημειώσεις του μαθήματος CS231n και οποιαδήποτε άλλη πηγή προκύπτει σε αυτό θα αναφέρεται ξεχωριστά (και κάποιες παραπομπές στον (Nielsen, 2015)).

3.5 ΔΟΜΗ & ΕΚΠΑΙΔΕΥΣΗ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

3.5.1 ΔΟΜΗ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

Ο θεμελιώδης λίθος των νευρωνικών δικτύων είναι ο **Νευρώνας**. Ένας νευρώνας αποτελείται από τις μεταβλητές εισόδου του, τον κορμό του και μία μεταβλητή εξόδου. Πολλοί νευρώνες παράλληλοι αποτελούν ένα επίπεδο του νευρωνικού δικτύου. Οι νευρώνες του ίδιου επιπέδου δεν συνδέονται μεταξύ τους, αλλά συνδέονται ένας προς έναν με όλους τους νευρώνες του επόμενου επιπέδου. Κάθε νευρωνικό δίκτυο έχει τουλάχιστον δύο επίπεδα, το επίπεδο εισόδου και το επίπεδο εξόδου (Perceptron). Ενδιάμεσα υπάρχουν από 0 έως πολλά κρυμμένα επίπεδα. Τα πρώτα νευρωνικά δίκτυα είχαν ένα με δύο κρυμμένα επίπεδα (Εικόνα 3.1) , τα πρώτα πιο σύνθετα δίκτυα είχαν 4-5, ενώ πλέον έχουν μέχρι και 20 χωρίς αυτό να αποτελεί αυστηρό όριο.



Εικόνα 3.1 Δομή Νευρωνικού Δικτύου με δύο κρυμμένα επίπεδα

Στην Εικόνα 3.1 τα βελάκια που εξέρχονται των νευρώνων δεν είναι πολλαπλά αλλά σχεδιάστηκε έτσι για λόγους απλότητας. Η μεταβλητή εξόδου είναι ίδια και μόνο μία και στην συνέχεια απλά διαχωρίζεται ώστε να αποτελέσει μία από τις μεταβλητές εισόδου του κάθε νευρώνα του επόμενου επιπέδου.

Για να γίνει κατανοητή η λειτουργία του δικτύου σαν σύνολο είναι σημαντική η κατανόηση της λειτουργίας του νευρώνα. Ειδικότερα, ο νευρώνας λειτουργεί όπως ο ανθρώπινος εγκέφαλος όταν θέλει να αποφασίσει για κάτι για αυτό θα χρησιμοποιηθεί και μία αναλογία. Έστω ότι θέλει κάποιος να αποφανθεί αν θα βρέχει αύριο και έχει τρία κριτήρια: την θερμοκρασία, την υγρασία και τις σημερινές καιρικές συνθήκες. Δίνοντας μία βαρύτητα σε κάθε μεταβλητή του, και αναλόγως με την τιμή των μεταβλητών του καταλήγει σε ένα αποτέλεσμα. Αν αυτό είναι μεγαλύτερο ή μικρότερο από ένα όριο που έχει θέσει ο ίδιος αποφασίζει αν θα βρέξει ή όχι. Η ίδια διαδικασία ακολουθείται και από το νευρώνα με ένα σταθμισμένο μέσο δηλαδή, όλων των μεταβλητών εισόδου ο οποίος συγκρίνεται με μία σταθερά και αναλόγως δίνει 0/1 σε διωνυμικά προβλήματα ή ποικίλες τιμές από το 0 έως το 1 σε μη διωνυμικά. Οι μεταβλητές εισόδου συμβολίζονται με X_i ενώ η βαρύτητά τους με w_i ενώ η σταθερή τιμή σύγκρισης ονομάζεται «προκατάληψη» (bias) και είναι κοινή για κάθε επίπεδο.

Οι πρώτες προσπάθειες έγιναν με νευρώνες τύπου 0/1. Βεβαίως στην πορεία φάνηκε ότι μικρές αλλαγές για την βελτίωση κάποιων αποφάσεων επηρέαζαν πάρα πολύ τις υπόλοιπες σωστές για αυτό και εγκαταλείφθηκε στην πορεία. Αντί λοιπόν να χρησιμοποιείται ένα όριο για να δοθεί μία απάντηση 0/1 από κάθε νευρώνα έπρεπε να βρεθεί μία συνάρτηση η οποία να λαμβάνει αυτές τις μεταβλητές εισόδου και πάντα να δίνει τιμές εξόδου 0 έως 1 αλλά και σωστή ποιοτικά συμπεριφορά (γίνεται έλεγχος συνήθως με τις ακραίες τιμές). Η συνάρτηση που χρησιμοποιήθηκε αρκετά είναι η συνάρτηση $\sigma = 1/(1+e^{-z})$, όπου z ο σταθμισμένος μέσος όρος συν την προκατάληψη $z=w*X+b$, ενώ στην πράξη έχουν χρησιμοποιηθεί και άλλες συναρτήσεις τέτοιου χαρακτήρα (tanh) όπως θα εξηγηθεί και στην συνέχεια. Η βέλτιστη συνάρτηση εξαρτάται από την χρήση που επιζητείται και αποτελεί ένα υπό διερεύνηση θέμα ακόμη για τα νευρωνικά δίκτυα. Εφαρμόζοντας μία τέτοια συνάρτηση το μοντέλο αποκτά το μη γραμμικό του χαρακτήρα και αυξάνει την δυσκολία πρόβλεψης της συμπεριφοράς του από τον ανθρώπινο εγκέφαλο.

Επομένως συνολικά, εισάγεται σε ένα νευρωνικό δίκτυο η πληροφορία από ένα σύνολο νευρώνων (επίπεδο εισόδου). Στην συνέχεια κάθε νευρώνας εισόδου συνδέεται με κάθε νευρώνα του επόμενου κρυμμένου επιπέδου. Κάθε νευρώνας επιτελεί τις παραπάνω λειτουργίες δηλαδή δέχεται έναν αριθμό εισερχόμενων μεταβλητών με τιμή από το 0 έως το 1, υπολογίζει το σταθμισμένο μέσο όρο τους και

εφαρμόζει μία συνάρτηση, συνήθως την σ , δίνοντας ένα αποτέλεσμα εντός του διαστήματος 0-1. Το αποτέλεσμα αυτό, αποτελεί έξοδο για αυτό το νευρώνα αλλά μεταβλητή εισόδου για όλους τους νευρώνες του επόμενου επιπέδου κ.ο.κ. σε όλα τα επίπεδα μέχρι το επίπεδο εξόδου όπου δίνεται η τελική απάντηση του προβλήματος.

Σε συνέχεια της αναλογίας που έγινε προηγουμένως, έστω ένας οργανωτής εκδηλώσεων ο οποίος προσπαθεί να αποφασίσει εάν θα πραγματοποιήσει μία εκδήλωση. Θα πρέπει να εξετάσει την πιθανότητα βροχής, την υπηρεσία του φαγητού, την διαθεσιμότητα του χώρου κ.ο.κ. Κάθε ένα από αυτά τα μικρά επιμέρους προβλήματα λύνονται (όπως εξηγήθηκε νωρίτερα) και το αποτέλεσμά τους πιθανώς να επηρεάζει και τα επόμενα επιμέρους προβλήματα. Για παράδειγμα σε περίπτωση βροχής πρέπει να εξεταστούν επιλογές πραγματοποίησης της εκδήλωσης με ενοικίαση τεντών. Το νευρωνικό δίκτυο λοιπόν, είναι σαν να διαχωρίζει ένα μεγάλο πρόβλημα σε ένα μεγάλο αριθμό επιμέρους προβλημάτων, των οποίων οι απαντήσεις επηρεάζουν τα επόμενα στάδια. Για το λόγο αυτό είναι σε θέση να αντιμετωπίσουν με μεγάλη ακρίβεια πολυσύνθετα προβλήματα, πολλών μεταβλητών που ο ανθρώπινος εγκέφαλος αδυνατεί όχι μόνο να επιλύσει αλλά και να κατανοήσει. Η δυσκολία που αντιμετωπίζει ο ανθρώπινος εγκέφαλος σε σχέση με το νευρωνικό δίκτυο είναι ότι δίνει φυσική υπόσταση στα επιμέρους προβλήματα ενώ για το νευρωνικό δίκτυο είναι απλά μία ακολουθία αριθμών και μη γραμμικών σχέσεων, η οποία δεν έχει καμία φυσική υπόσταση αλλά σκοπός ύπαρξής της είναι μόνο η βελτιστοποίηση του αποτελέσματος που δίνεται στο τελευταίο επίπεδο.

Ο τρόπος βελτιστοποίησης του τελικού αποτελέσματος γίνεται με μία **Συνάρτηση Κόστους**. Στόχος είναι η ελαχιστοποίηση αυτού του κόστους, η οποία συνήθως συνοδεύεται με αύξηση της ακρίβειας του μοντέλου. Η συνάρτηση κόστους πρέπει να είναι πάντα θετική, να μειώνεται όταν το αποτέλεσμα πλησιάζει το πραγματικό και αντιστρόφως. Ακόμη απαιτείται να είναι συνάρτηση των μεταβλητών εξόδου a των νευρώνων, άρα και των w και b που συνέβαλαν σε αυτή την τιμή του a ώστε να πραγματοποιούνται ελεγχόμενες αλλαγές και τέλος πρέπει να μπορεί να γράφεται σαν μέσος όρος άλλων συναρτήσεων κόστους (θα επεξηγηθεί σύντομα γιατί). Μία συνάρτηση που χρησιμοποιείται όχι μόνο στα νευρωνικά δίκτυα αλλά και σε άλλα μοντέλα είναι τα Ελάχιστα Τετράγωνα (1), όπου αθροίζονται τα τετράγωνα των διαφορών των τιμών που αναμένονταν με τις τιμές που υπολογίστηκαν. Επίσης διαδεδομένη είναι η συνάρτηση κόστους Cross Entropy (2), η οποία αποτελεί μία λογαριθμική προσέγγιση. Άλλες συναρτήσεις είναι το εκθετικό κόστος, η απόσταση Hellinger κ.λ.π. Οι εξισώσεις (1) και (2) επαληθεύουν ότι οι συναρτήσεις κόστους πληρούν όλα τα κριτήρια που προαναφέρθηκαν.

$$C(w, b) \equiv \frac{1}{2n} \sum_x \|y(x) - a\|^2 \quad (1)$$

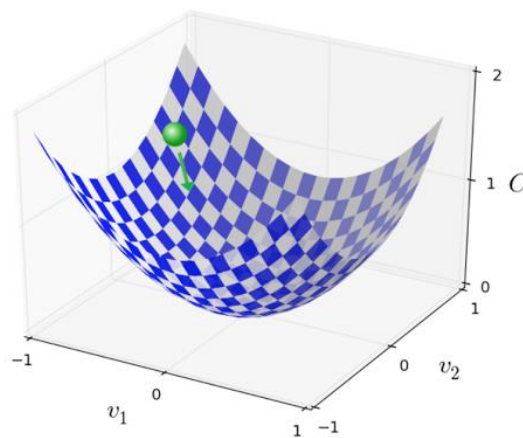
$$C = -\frac{1}{n} \sum_x [y \ln a + (1 - y) \ln(1 - a)] \quad (2)$$

3.5.2 ΕΚΠΑΙΔΕΥΣΗ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

Γενικευμένος Κανόνας Δέλτα

Αφού έγινε κατανοητή η δομή των νευρωνικών δικτύων και ο τρόπος επικοινωνίας μεταξύ των νευρώνων, είναι σημαντικό να επεξηγηθεί ο τρόπος εκπαίδευσής τους. Για το σκοπό αυτό πρέπει να επεξηγηθούν δύο έννοιες: το **(Stochastic) Gradient Descent** (Γενικευμένος Κανόνας Δέλτα) και το **Backpropagation** (Οπισθόδρομη Διάδοση Σφάλματος).

Όπως ήδη αναφέρθηκε η ελάττωση της συνάρτησης κόστους βελτιώνει το νευρωνικό δίκτυο. Για να επιτευχθεί αυτό χρησιμοποιείται η παράγωγος της συνάρτησης κόστους και ένας ρυθμός εκπαίδευσης. Για την κατανόηση του φαινομένου γίνεται μία απλούστευση και ύστερα μία γενίκευση. Έστω μία κοιλάδα η οποία περιγράφεται από την συνάρτηση κόστους, όπου έχει σε ένα σημείο ολικό ελάχιστο και κάποια τοπικά ελάχιστα. Έστω ότι η συνάρτηση C έχει δύο μεταβλητές v_1 και v_2 . Οι μεταβλητές αυτές αντικαθιστούν το βάρος w και την «προκατάληψη» b και είναι ενδεικτικές. Η αναπαράστασή της θα ήταν όπως η Εικόνα 3.2.



Εικόνα 3.2 Αναπαράσταση Συνάρτησης Κόστους με δύο μεταβλητές

Σκοπός είναι η εύρεση του ολικού ελαχίστου δηλαδή μία νοητή μπάλα να φτάσει στο χαμηλότερο σημείο. Μία πολύ μικρή μετακίνηση Δv_1 και αντίστοιχα μία πολύ μικρή μετακίνηση Δv_2 (όπου $\Delta v = (\Delta v_1, \Delta v_2)^T$) θα προκαλούσαν μία μεταβολή της συνάρτησης κόστους στην συγκεκριμένη περίπτωση κατά $\Delta C \approx \frac{\partial C}{\partial v_1} \Delta v_1 + \frac{\partial C}{\partial v_2} \Delta v_2$. (3). Σε γενικότερο πλαίσιο, για τα εκατομμύρια βάρη των νευρώνων των δικτύων και τις πολλαπλές «προκαταλήψεις», η παραπάνω εξίσωση παίρνει την εξής μορφή:

$$\Delta C \approx \nabla C \cdot \Delta v \quad (4) \quad \text{όπου} \quad \nabla C \equiv \left(\frac{\partial C}{\partial v_1}, \dots, \frac{\partial C}{\partial v_m} \right)^T \quad (5).$$

Επιλέγοντας $\Delta v = -\eta \nabla C$, (6), όπου “ η ” ο ρυθμός εκμάθησης, προκύπτει ότι η μεταβολή του κόστους θα είναι σίγουρα αρνητική, που ήταν και το αρχικό ζητούμενο, σύμφωνα με την εξίσωση:

$$\Delta C \approx -\eta \nabla C \cdot \nabla C = -\eta \|\nabla C\|^2 \quad (7).$$

Στην συνέχεια διορθώνονται οι παράμετροι $v \rightarrow v' = v - \eta \nabla C$. (8), έτσι ακριβώς όπως ορίστηκε το Δv , και επαναλαμβάνεται η διαδικασία αυτή από την αρχή μέχρι να επιτευχθεί το ολικό ελάχιστο. Η διαδικασία αυτή αφορά όλες τις παραμέτρους του δικτύου δηλαδή και τα βάρη αλλά και τις «προκαταλήψεις».

Η επιλογή του **ρυθμού εκμάθησης (η)** είναι πολύ σημαντική για την εκπαίδευση του δικτύου. Πολύ μικρός ρυθμός εκμάθησης οδηγεί σε υπερβολικά μεγάλους χρόνους εκπαίδευσης ενώ αντίθετα πολύ μεγάλος ρυθμός εκπαίδευσης οδηγεί σε υπερβολικά μεγάλα βήματα, τέτοια που μπορεί ποτέ να μην επιτευχθεί το ολικό ελάχιστο αφού θα ξεπερνιέται. Υπάρχουν πολλές μεθοδολογίες που έχουν αναπτυχθεί για την επιλογή του κατάλληλου “ η ” και θα γίνει αναφορά σε κάποιες από αυτές στο Κεφάλαιο 3.4.3 .

Πολλές φορές, σε μεγάλα νευρωνικά δίκτυα, οι νευρώνες είναι εκατομμύριοι με τις παραμέτρους να είναι πολλαπλάσιά τους. Ο υπολογιστικός φόρτος των παραγώγων μπορεί να γίνει πάρα πολύ δύσκολος. Ακόμη για να υπολογιστεί, χρησιμοποιείται ο μέσος όρος των παραγώγων κάθε δεδομένου το οποίο επεξηγεί γιατί επιλέχτηκε με αυτή την ιδιότητα η συνάρτηση κόστους. Για το λόγο αυτό επινοήθηκε μία στοχαστική προσέγγιση, η οποία είναι ανάλογη της χρήσης των polls στις εκλογές. Επιλέγοντας τυχαία έναν μικρό αριθμό δημιουργείται ένα **Αντιπροσωπευτικό Σύνολο** του οποίου υπολογίζονται οι παράγωγοι και γίνεται η υπόθεση ότι ο μέσος όρος των παραγώγων του μικρού αυτού συνόλου ισούται με τον μέσο όρο όλων των παραγώγων όλων των παραμέτρων. Αυτή η πρακτική επιταχύνει την διαδικασία της εκπαίδευσης σε μεγάλο βαθμό (όσο πιο δραστικά χρησιμοποιείται τόσο περισσότερο). Η διαδικασία αυτή επαναλαμβάνεται μέχρι να εξαντληθούν όλα τα

δεδομένα εισόδου όπου λέγεται ότι τελειώνει ο πρώτος κύκλος εκπαίδευσης, **Εποχή** (η βασική μονάδα μέτρησης της έκτασης της εκπαίδευσης).

Οι τρόποι εκπαίδευσης νευρωνικών δικτύων είναι ακόμη υπό έρευνα, αλλά αυτός είναι ο πιο διαδεδομένος και αποδεκτός μέχρι σήμερα.

Μία επιπλέον παράμετρος που χρησιμοποιείται στην εκμάθηση και βοηθά στην επίσπευσή της είναι η **Ορμή (μ)**. Στα νευρωνικά δίκτυα η ορμή έχει την ίδια σημασία που έχει και στην πραγματικότητα, δηλαδή δίνει μία ώθηση παραπάνω στο δίκτυο όταν αυτό κινείται προς κάποια συγκεκριμένη κατεύθυνση. Η ορμή βοηθά το δίκτυο να ξεπεράσει ευκολότερα και γρηγορότερα τοπικά ελάχιστα στην πορεία του προς το ολικό ελάχιστο. Η μαθηματική αλλαγή στο δίκτυο είναι αυτή της Εξίσωσης (9) όπου ουσιαστικά προστίθεται ο όρος “μ” (Ορμή) σε σχέση με την Εξίσωση (8). Η πιο ακριβής του έννοια θα ήταν η τριβή αφού για $\mu=0$ είναι ανύπαρκτη, ενώ για $\mu=1$ προστίθεται και όλος ο όρος του προηγούμενου βήματος στην διόρθωση. Ουσιαστικά το σύστημα μετά την αλλαγή των παραμέτρων υποχρεώνεται σε αυτή την περίπτωση να ‘θυμάται’ και του προηγούμενου βήματος τις παραμέτρους ώστε να εφαρμόσει την έννοια της ορμής αυξάνοντας τον υπολογιστικό φόρτο. Η προσθήκη του $\mu \cdot v$ δεν αλλάζει κάτι σε μαθηματικό επίπεδο στην εύρεση των παραγώγων και στην διαδικασία που περιγράφηκε νωρίτερα για μία τυχαία παράμετρο v (είτε είναι βάρος είτε «προκατάληψη»).

$$\begin{aligned} v &\rightarrow v' = \mu v - \eta \nabla C \\ w &\rightarrow w' = w + v'. \end{aligned} \quad (9)$$

Οπισθόδρομη Διάδοση Σφάλματος

Ένα σημαντικό θέμα είναι η εύρεση ενός γρήγορου τρόπου υπολογισμού των παραγώγων που προαναφέρθηκαν. Στο ηλεκτρονικό βιβλίο του Nielsen αποδεικνύονται οι παρακάτω 4 Βασικές Εξισώσεις για την μετάδοση των λαθών προς τα πίσω, οι οποίες παρατίθενται απλώς παρακάτω και επιτελούν αυτήν την εργασία.

$$\delta^L = \nabla_a C \odot \sigma'(z^L) \quad (10)$$

$$\delta^l = ((w^{l+1})^T \delta^{l+1}) \odot \sigma'(z^l) \quad (11)$$

$$\frac{\partial C}{\partial b_j^l} = \delta_j^l \quad (12)$$

$$\frac{\partial C}{\partial w_{jk}^l} = a_k^{l-1} \delta_j^l \quad (13)$$

Οι εξισώσεις με την σειρά που είναι γραμμένες, αφορούν :

- Μία εξίσωση η οποία υπολογίζει το λάθος στο τελευταίο επίπεδο του δικτύου κάτι λογικό αφού σε αυτό το σημείο υπολογίζεται η συνάρτηση κόστους. Εφόσον η συνάρτησή κόστους είναι «βολική» όπως περιγράφηκε, με τις κατάλληλες απλοποιήσεις η εξίσωση αυτή δεν αποτελεί παρά ένα γινόμενο Hadamard ($[1,2] \circ [3,4] = [3,8]$) δύο διανυσμάτων
- Μία εξίσωση, η οποία υπολογίζει το λάθος σε ένα επίπεδο του δικτύου ως προς το λάθος του επόμενου επιπέδου. Αυτή η εξίσωση είναι ιδιαίτερα χρήσιμη αφού το λάθος υπολογίζεται αρχικά στο τελευταίο επίπεδο. Με αυτό τον τρόπο το λάθος επιμερίζεται σε όλα τα επίπεδα του δικτύου.
- Μία εξίσωση για το ρυθμό μεταβολής της προκατάληψης ενός επιπέδου, που ήταν και ένα από τα δύο ζητούμενα
- Μία εξίσωση για το ρυθμό μεταβολής οποιουδήποτε βάρους ενός επιπέδου, που ήταν το άλλο ζητούμενο. Από αυτήν την εξίσωση παρατηρείται ότι νευρώνες με μεταβλητή εξόδου a πολύ μικρή «μαθαίνουν» αργά και αντίστροφα.

Οι παραπάνω εξισώσεις λοιπόν έχουν μετατρέψει την χρήση των νευρωνικών δικτύων σε μία γρήγορη διαδικασία αν ληφθεί υπόψη ότι με τον πατροπαράδοτο τρόπο θα έπρεπε να υπολογιστεί η συνάρτηση κόστους τόσες φορές όσες και τα δεδομένα ενώ τώρα ένα πέρασμα εμπρός και ένα πίσω στο δίκτυο όλο, επιτυγχάνει το ίδιο αποτέλεσμα, αφού σε αυτό υπολογίζονται όλες οι μερικές παράγωγοι, οι οποίες είναι συνάρτηση των μεταβλητών εξόδου a , των μεταβλητών εισόδου z και των επιμέρους παραμέτρων w, b τα οποία είναι όλα ήδη γνωστά.

Υποσύνολα του Δείγματος και Υπερπροσαρμογή

Το δείγμα που χρησιμοποιείται για την εκπαίδευση συνίσταται να χωρίζεται σε τρία μέρη: Το πρώτο μέρος, το **Σύνολο Εκπαίδευσης**, είναι το σύνολο των δεδομένων που χρησιμοποιούνται για την εκπαίδευση του δικτύου και πρέπει να είναι 60-70% του συνόλου.

Το δεύτερο μέρος, το **Σύνολο Επικύρωσης** είναι ένα σύνολο το οποίο χρησιμοποιείται για να αποφευχθεί το πρόβλημα της υπερ-προσαρμογής του μοντέλου στο Σύνολο Εκπαίδευσης και πρέπει να είναι 15-25% του συνολικού διαθέσιμου δείγματος. Με άλλα λόγια, το δίκτυο λόγω υπερπροσαρμογής, μπορεί να πάψει να μαθαίνει την σημαντική πληροφορία από τα δεδομένα, η οποία θα το βοηθήσει να γενικεύει καλύτερα και να προβλέπει καλύτερα τα νέα δεδομένα. Από

ένα σημείο δηλαδή και μετά, το δίκτυο μαθαίνει να πετυχαίνει μεγαλύτερες ακρίβειες στο Σύνολο Εκπαίδευσης, μαθαίνοντας τις ιδιομορφίες του και χάνοντας την ικανότητά του να γενικεύει.

Για να αποφευχθεί αυτό το πρόβλημα που ονομάζεται **Υπερπροσαρμογή (Overfitting)** χρησιμοποιούνται διάφορες τεχνικές. Μία από αυτές είναι η χρήση του Συνόλου Επικύρωσης το οποίο δεν χρησιμοποιείται για εκμάθηση αλλά μόνο για έλεγχο ακρίβειας με σκοπό να διερευνηθεί αν το μοντέλο συνεχίζει να αποκτά ικανότητα να γενικεύει και σε ποιο σημείο την χάνει. Ένας κανόνας υπαγορεύει ότι όσο η ακρίβεια στο Σύνολο Εκπαίδευσης μεγαλώνει ενώ ταυτόχρονα μεγαλώνει και η ακρίβεια στο Σύνολο Επικύρωσης, το δίκτυο γενικεύει καλύτερα. Αν η ακρίβεια στο Σύνολο Επικύρωσης αρχίσει να πέφτει ή σταθεροποιηθεί σημαίνει ότι βρέθηκε το σημείο αυτό. Αναλόγως με την αυστηρότητά του, ο ερευνητής μπορεί να προχωρήσει σε 1,3,5 ή 10 «Εποχές» για να αποφασίσει ότι το δίκτυό του σταμάτησε να μαθαίνει και να γενικεύει καλύτερα. Έχουν ερευνηθεί διάφορες τεχνικές για τον βέλτιστο αριθμό «Εποχών» όπου πρέπει να τερματιστεί η εκπαίδευση αλλά καμία δεν έχει δώσει κάποια ιδιαίτερος χρήσιμα αποτελέσματα. Το Σύνολο Επικύρωσης χρησιμοποιείται συνήθως για να βρεθούν οι επιμέρους παράμετροι, όπως ο ρυθμός εκμάθησης, η ορμή κ.α.

Τέλος το τρίτο μέρος, το **Σύνολο Δοκιμής**, είναι ένα νέο σύνολο εφαρμογής του μοντέλου για να μετρηθεί η ακρίβεια της πρόβλεψής του και θεωρείται το πιο αξιόπιστο. Αποτελεί συνήθως από 5-20% του συνόλου των δεδομένων. Ο λόγος ύπαρξής του είναι ουσιαστικά η προσπάθεια για να αποφευχθεί η πιθανότητα της υπερπροσαρμογής του δικτύου στο Σύνολο Επικύρωσης. Η τελική ακρίβειά του είναι συνήθως λίγο μικρότερη της ακρίβειας της τελευταίας «Εποχής» εκπαίδευσης του Συνόλου Επικύρωσης.

Συστηματοποίηση και Απόσυρση Νευρώνων

Άλλες διαδεδομένες μέθοδοι επίλυσης του προβλήματος της υπερπροσαρμογής είναι η **Συστηματοποίηση (Regularization)** και η **Απόσυρση** ορισμένων **Νευρώνων (Dropout)**.

Η Συστηματοποίηση ή αλλιώς **Μείωση Βαρών (Weight Decay)** προσθέτει έναν επιπλέον όρο στην συνάρτηση κόστους σύμφωνα με την (14) όπου “λ” η παράμετρος συστηματοποίησης. Είναι εμφανές ότι μεγαλύτερα βάρη σημαίνουν μεγαλύτερη τιμή της συνάρτησης κόστους άρα πρέπει να αποφεύγονται. Οι μερικές παράγωγοι των συναρτήσεων κόστους ως προς τα βάρη (15) και τις «προκαταλήψεις» (16)

αντίστοιχα υποδεικνύουν ότι οι «προκαταλήψεις» δεν επηρεάζονται από αυτήν την αλλαγή.

$$C = C_0 + \frac{\lambda}{2n} \sum_w w^2, \quad (14)$$

$$\frac{\partial C}{\partial w} = \frac{\partial C_0}{\partial w} + \frac{\lambda}{n} w \quad (15)$$

$$\frac{\partial C}{\partial b} = \frac{\partial C_0}{\partial b} \quad (16)$$

Ακόμη στην Εξίσωση (17) παρουσιάζεται η μεταβολή των παραμέτρων όπου είναι εμφανής η έννοια της μείωσης των βαρών από την παράμετρο “λ”.

$$\begin{aligned} w &\rightarrow w - \eta \frac{\partial C_0}{\partial w} - \frac{\eta \lambda}{n} w \\ &= \left(1 - \frac{\eta \lambda}{n}\right) w - \eta \frac{\partial C_0}{\partial w}. \end{aligned} \quad (17)$$

Όσο μεγαλύτερη είναι η παράμετρος “λ”, τόσο μικρότερα βάρη επιδιώκονται για το νευρωνικό δίκτυο αλλά τόσο επιβαρύνεται περισσότερο από αυτόν τον όρο η συνάρτηση κόστους. Όσο μικρότερα βάρη έχει ένα δίκτυο, τόσο μικρότερη επιρροή έχουν δεδομένα που αποτελούν θόρυβο για το σύστημα ενώ αντίθετα μεγάλη ισχύ έχουν όσα επαναλαμβάνονται. Πρόκειται για μία λεπτή ισορροπία, η οποία όμως είναι πολύ σημαντική για την απόδοση του δικτύου.

Η απόσυρση ορισμένων νευρώνων (dropout) λειτουργεί ακριβώς όπως λέει και το όνομά της με απόσυρση ενός ποσοστού των υπαρχόντων νευρώνων των κρυμμένων επιπέδων. Η διαδικασία αυτή επαναλαμβάνεται πολλές φορές διαλέγοντας κάθε φορά άλλο συνδυασμό νευρώνων προς απόσυρση. Στο τέλος χρησιμοποιούνται όλα οι νευρώνες με βάρη τους μέσους όρους των βαρών των επιμέρους επαναλήψεων εφαρμογής του μοντέλου, διαιρεμένων με το αντίστοιχο ποσοστό. Δηλαδή για 50% θα γινόταν διαίρεση των βαρών με το 2. Η επιτυχία της μεθόδου βασίζεται στην δύναμη του μέσου όρου. Αν εκπαιδεύονταν 10 διαφορετικά νευρωνικά δίκτυα για ένα

συγκεκριμένο πρόβλημα με 8 να κάνουν σωστή πρόβλεψη και 2 λάθος, και υιοθετείτο ο μέσος όρος των δέκα προβλέψεων θα επιτυγχανόταν σίγουρα πιο ορθή πρόβλεψη σε σχέση με την περίπτωση να χρησιμοποιηθεί το ένα από τα δύο εσφαλμένα δίκτυα. Σε αυτή τη λογική βασίζεται και η μέθοδος dropout.

Ένα ακόμη ζήτημα αποτελεί η **Αρχική Τιμή των Παραμέτρων** του δικτύου πριν αρχίσουν να μεταβάλλονται από την εκπαίδευσή τους. Μία γενική, απλή λύση είναι η επιλογή παραμέτρων μέσω της κατανομής του Gauss με μέσο όρο 0 και τυπική απόκλιση 1. Μία πιο καλή τακτική είναι να χρησιμοποιηθεί τυπική απόκλιση $1/N^{(1/2)}$ διότι με αυτό τον τρόπο οι τιμές των παραμέτρων δίνουν τιμές των μεταβλητών εισόδων z των νευρώνων πιο κοντά στο 0 με αποτέλεσμα το δίκτυο να μαθαίνει γρηγορότερα. Άλλη εναλλακτική είναι η χρήση κάποιας ομοιόμορφης κατανομής. Το ζήτημα αυτό είναι επίσης υπό διερεύνηση.

Σύντομη Σύνοψη

Συνοψίζοντας την παραπάνω ανάλυση, τα νευρωνικά δίκτυα αποτελούνται από νευρώνες, οι οποίοι δημιουργούν επίπεδα και οι οποίοι συνδέονται μεταξύ τους ανά επίπεδο με την μορφή «όλα με όλα». Επιλέγεται το μέγεθος του δικτύου, δηλαδή ο αριθμός των επιπέδων και ο αριθμός των νευρώνων ανά επίπεδο, η συνάρτηση εκπαίδευσης των νευρώνων και η συνάρτηση κόστους. Για την εκπαίδευση του νευρωνικού δικτύου απαραίτητο είναι να καθοριστούν: ο ρυθμός εκπαίδευσης “ η ”, η μέθοδος εκπαίδευσης, το μέγεθος των μικρών συνόλων εκπαίδευσης, η παράμετρος συστηματοποίησης “ λ ”, η ορμή “ μ ” (προαιρετική), η αρχική τιμή των παραμέτρων και ο αριθμός των εποχών εκπαίδευσης. Απαραίτητο είναι επίσης να έχει διαχωριστεί το δείγμα στα υποσύνολα που αναλύθηκαν.

Τακτική Εκπαίδευσης

Σε αυτό το σημείο, αφού εξηγήθηκαν η δομή και η έννοια της εκπαίδευσης των νευρωνικών δικτύων, είναι πρόπον να αναφερθεί μια τακτική προσέγγισης της εκπαίδευσης των δικτύων. Δεν υπάρχει ορισμένη ορθή μεθοδολογία και η αναζήτηση της ορθότερης είναι ίσως και άσκοπη, επειδή η μη γραμμικότητα των προβλημάτων κάνει πολύ δύσκολο έως και αδύνατο τον προσδιορισμό της.

Σύμφωνα με τον Nielsen, η πρώτη κίνηση είναι η απλούστευση του προβλήματος, αν αυτό είναι εφικτό, ώστε αφού λυθεί να χτιστεί πάνω σε αυτό το δυσκολότερο πρόβλημα. Ακόμη, για τις διάφορες δοκιμές παραμέτρων προτείνει να γίνεται έλεγχος

με λιγότερα δεδομένα από τα πραγματικά επειδή μας ενδιαφέρει η κατεύθυνση και η επιρροή τους και όχι το απόλυτο αποτέλεσμα σε αυτήν την αρχική φάση. Προτείνει αφού ξεκαθαριστούν οι παράγοντες να γίνει στην συνέχεια έλεγχος με το σύνολο των δεδομένων. Βεβαίως μερικές φορές αυτό μπορεί να είναι μοιραίο για την απόδοση του δικτύου. Επομένως ο ερευνητής πρέπει να έχει εμπειρία να κατανοήσει το ελάχιστο μέγεθος του Συνόλου Εκπαίδευσης, που απαιτείται ώστε να έχει αποτελέσματα.

Η πρώτη κίνηση είναι ο καθορισμός της μορφής του δικτύου, δηλαδή του αριθμού των επιπέδων αλλά και των νευρώνων ανά επίπεδο. Κάθε νευρωνικό δίκτυο έχει υποχρεωτικά 2 επίπεδα, ένα εισόδου με αριθμό νευρώνων ίσο με τον αριθμό της εισερχόμενης πληροφορίας ανά μονάδα εκπαίδευσης και ένα επίπεδο εξόδου με αριθμό νευρώνων κατά κανόνα ίσο με τον αριθμό των αποτελεσμάτων που υπάρχουν στο πρόβλημα. Σύμφωνα με το βιβλίο του Jeff Heaton, τα κρυφά επίπεδα πρέπει να είναι 0 για γραμμικές μόνο σχέσεις ενώ ισχυρίζεται ότι μέχρι 2 κρυφά επίπεδα αρκούν για την επίλυση φυσιολογικών προβλημάτων (Heaton, 2008). Η λογική των νευρωνικών υποδεικνύει ότι μεγαλύτερα δίκτυα έχουν μεγαλύτερη δύναμη ακρίβειας αλλά εγκυμονούν κινδύνους αργής εκμάθησης ή αδρανοποίησης μεγάλων ποσοστών του δικτύου τους. Στην ουσία πρόκειται για μία λεπτή ισορροπία της δύναμης, που απαιτείται για να επιτευχθεί μία συγκεκριμένη ακρίβεια σε ένα πρόβλημα, σε σχέση με το μέγεθος του δικτύου, που θα απαιτείται για αυτό και το χρόνο για να πραγματοποιηθεί. Έχει βρεθεί πολλές φορές στην πράξη, ότι η προσθήκη επιπλέον επιπέδων δεν προσφέρει σχεδόν καθόλου στην ακρίβεια του δικτύου επομένως αποτελεί απλή σπατάλη χώρου και χρόνου.

Όσον αφορά τον αριθμό των νευρώνων ανά επίπεδο υπάρχουν κάποιοι γενικοί κανόνες που ακολουθούνται χωρίς βέβαια να είναι τόσο δεσμευτικοί:

- Ο αριθμός των κρυφών νευρώνων πρέπει να είναι μεταξύ του αριθμού των νευρώνων του επιπέδου εισόδου και του επιπέδου εξόδου
- Ο αριθμός των κρυφών νευρώνων πρέπει να είναι ίσος με τα $2/3$ του αριθμού των νευρώνων του επιπέδου εισόδου συν τους νευρώνες του επιπέδου εξόδου
- Ο αριθμός των κρυφών νευρώνων πρέπει να είναι μικρότερος του διπλάσιου του αριθμού των νευρώνων του επιπέδου εισόδου.

Βεβαίως όλοι αυτοί οι κανόνες δεν έχουν επαρκή τεκμηρίωση διότι δεν λαμβάνουν υπόψη τον αριθμό των δεδομένων που χρησιμοποιούνται για την εκπαίδευση. Γενική προσέγγιση είναι η χρήση όσο γίνεται λιγότερων νευρώνων ανά επίπεδο γιατί αλλιώς συνήθως πραγματοποιείται μία άσκοπη απομνημόνευση στοιχείων που δεν χρειάζονται, γνωστή και ως υπερπροσαρμογή (βλέπε Παράγραφο 3.5.2 Υποσύνολα

του Δείγματος και Υπερπροσαρμογή). Από το 1975 μέχρι και σήμερα έχουν πραγματοποιηθεί και συνεχίζουν να πραγματοποιούνται πολλές έρευνες πάνω σε αυτό το θέμα. Μία από αυτές, έχει χρήσιμες συγκρίσεις λαμβάνοντας υπόψη και την παραπάνω παράμετρο, που αναφέρθηκε για τον αριθμό των δεδομένων εκπαίδευσης (Lawrence et al., 1996).

Αφού επιλεγθεί η μορφή και η δομή του νευρωνικού δικτύου ερευνώνται οι βέλτιστοι παράμετροί του. Για τύπο εκπαίδευσης με ελάττωση της παραγώγου στοχαστικά (δεν έχει αναφερθεί σε αυτήν την εργασία άλλος τρόπος) ο Nielsen, αρχικά προτρέπει στην εύρεση του κατάλληλου ρυθμού εκπαίδευσης “η”. Ξεκινώντας από μία τιμή 0.01 και πολλαπλασιάζοντας ή διαιρώντας με το 10, μπορούν να γίνουν διαδοχικές δοκιμές όπου στις πρώτες 10 «Εποχές» (εάν το δίκτυο δεν είναι μεγαλύτερο των 5-6 επιπέδων) θα πρέπει να παρατηρείται κάποια εκμάθηση. Στόχος είναι να βρεθεί το μεγαλύτερο “η”, με το οποίο το νευρωνικό δίκτυο συνεχίζει να μαθαίνει. Στην συνέχεια εάν πρόκειται να χρησιμοποιηθεί σταθερό “η” προτείνεται μία τιμή 100 φορές μικρότερη του μέγιστου “η”. Αν πρόκειται να γίνει εκπαίδευση με μεταβαλλόμενο “η”, υπάρχει ολόκληρη βιβλιογραφία που ασχολείται με την βέλτιστη μείωσή του. Παρατίθεται η έρευνα (Senior et al., 2013), η οποία συγκρίνει διάφορες τεχνικές μείωσης του “η” και προτείνει μία παραλλαγή δικιά της, AdaDec, σε ένα υπάρχον σύστημα AdaGrad. Αφού επιλεγθεί το σωστό “η” ακολουθείται παρόμοια διαδικασία για το “λ” της συστηματοποίησης. Τέλος εξετάζονται βοηθητικά μέτρα όπως η ορμή, το dropout, εναλλακτικές αρχικές τιμές των παραμέτρων κ.α. που συμβάλλουν κυρίως στην ταχύτητα του δικτύου.

Θα μπορούσε κανείς να ισχυριστεί ότι θα μπορούσε να υπάρχει ένα αυτοματοποιημένο σύστημα εύρεσης των βέλτιστων παραγόντων. Αυτό βέβαια είναι πολύ δύσκολο έως αδύνατο, διότι η μη γραμμικότητα δεν επιτρέπει στον ερευνητή να βάλει κανόνες, σύμφωνα με τους οποίους εξετάζει το δίκτυο. Επομένως οι περισσότερες από τις αποφάσεις είναι σε μεγάλο βαθμό συνδεδεμένες με το εκάστοτε πρόβλημα και τις ιδιαιτερότητές του.

3.6 ΣΥΝΕΛΙΚΤΙΚΑ ΔΙΚΤΥΑ

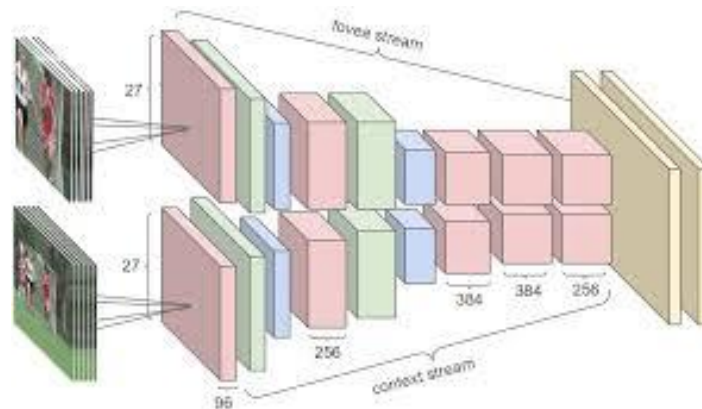
3.6.1 ΓΕΝΙΚΑ

Τα συνελικτικά δίκτυα είναι ένα είδος νευρωνικών δικτύων. Συνεπώς οι βασικές αρχές του Κεφαλαίου 3.5 ισχύουν και για τα συνελικτικά δίκτυα. Συγκεκριμένα οι βασικές διαφορές των συνελικτικών δικτύων είναι κυρίως δομικές. Η εκπαίδευση των δικτύων γίνεται με τον ίδιο τρόπο που περιγράφηκε στο Κεφάλαιο 3.5 με ορισμένες μικρές προσθήκες ή αλλαγές, οι οποίες είναι φυσικό επακόλουθο της διαφορετικής δομής αυτών των δικτύων. Για το λόγο αυτό σε αυτό το κεφάλαιο θα δοθεί ιδιαίτερη έμφαση στην δομή αυτών των δικτύων και οτιδήποτε πρόσθετο από όσα αναφέρθηκαν στο Κεφάλαιο 3.5 απαιτείται για την κατανόηση τους και την εκπαίδευσή τους.

3.6.2 ΤΡΙΣΔΙΑΣΤΑΤΗ ΔΟΜΗ

Η βασική ιδιαιτερότητα των συνελικτικών δικτύων είναι ότι είναι πολύ πιο φιλικά στην υποδοχή εικόνων ως μεταβλητή εισόδου. Οι έγχρωμες εικόνες έχουν 3 διαστάσεις: πλάτος, ύψος και βάθος όπου ως βάθος ορίζονται τα 3 επίπεδα χρωμάτων Κόκκινο-Πράσινο-Μπλε. Σε περίπτωση ασπρόμαυρης φωτογραφίας το επίπεδο είναι ένα. Κάθε εικόνα περιέχει έναν αριθμό από pixels, που υπολογίζεται απλά πολλαπλασιάζοντας κάθε της διάσταση. Τα έως τώρα νευρωνικά δίκτυα θα αποσυνθέτουν την μορφή των εικόνων αυτών από τρισδιάστατη και θα δημιουργούσαν ένα διάνυσμα μίας διάστασης παραθέτοντας γραμμικά τα στοιχεία από όλες τις διαστάσεις. Έχουν πραγματοποιηθεί στο παρελθόν πάρα πολλές έρευνες σχετικά με τον βέλτιστο τρόπο παράθεσης όλων αυτών των στοιχείων, που αλλοιώνουν ουσιαστικά την μορφή των εικόνων.

Η μεγάλη εξέλιξη των συνελκτικών δικτύων είναι ότι σαν δεδομένο εισόδου δεν δέχονται ένα διάνυσμα μίας διάστασης αλλά ένα πολυδιάστατο πίνακα. Αυτό επιτρέπει την διατήρηση της 3-διάστατης μορφής των έγχρωμων φωτογραφιών και την καλύτερη ανάλυσή τους. Η βασική διαφορά επομένως των απλών νευρωνικών δικτύων και των συνελκτικών είναι ότι τα συνελκτικά δίκτυα έχουν τρισδιάστατη μορφή και από επίπεδο σε επίπεδο αλλάζει αυτή η τρισδιάστατη μορφή (Εικόνα 3.3) ενώ προηγουμένως στα νευρωνικά δίκτυα άλλαζε μόνο η μία διάσταση.



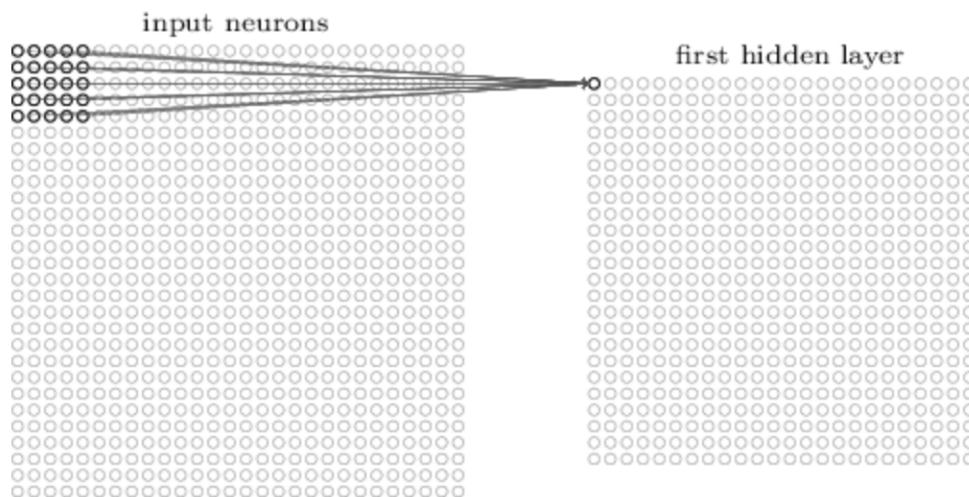
Εικόνα 3.3 Αλλαγή Τρισδιάστατης Μορφής Συνελκτικών Δικτύων

3.6.3 ΤΟΠΙΚΟΙ ΤΟΜΕΙΣ ΑΝΑΛΥΣΗΣ

Σε αυτό το σημείο πρέπει να εξεταστεί η λειτουργία του δικτύου σε 3 διαστάσεις για να επεξηγηθεί η ανωτερότητά του για την επεξεργασία εικόνων σε σχέση με τα απλά νευρωνικά δίκτυα. Οι δύο θεμελιώδεις έννοιες των συνελκτικών δικτύων που τα προσδιορίζουν και τα ξεχωρίζουν από τα απλά νευρωνικά δίκτυα είναι οι **Τοπικοί Τομείς Ανάλυσης** και τα **Κοινά Βάρη** ως αποτέλεσμά τους, και τα **Επίπεδα Συγκέντρωσης**.

Έστω ότι η φωτογραφίες εισόδου είναι έγχρωμες και διαστάσεων 28x28 pixels. Κάθε pixel αποτελεί ένα νευρώνα. Οι νευρώνες αυτοί επιχειρείται να ενωθούν με τους νευρώνες του επόμενου επιπέδου. Παρολ'αυτά, αντί να γίνει αντιστοίχιση ένας προς έναν, γίνεται αντιστοίχιση ενός πυρήνα νευρώνων (πυρήνας συνέλιξης) του προηγούμενου επιπέδου με έναν νευρώνα του επόμενου επιπέδου. Έστω ότι ο πυρήνας αυτός είναι διάστασης 5x5. Ο πυρήνας αυτός μετακινείται με βήμα 1 pixel

σε όλη την εικόνα τόσο κατά πλάτος όσο και καθ' ύψος. Το αποτέλεσμα είναι να δημιουργείται το επόμενο επίπεδο με διαστάσεις 24x24 (Εικόνα 3.4) αφού σε 28 νευρώνες αντιστοιχούν 24 κυλιόμενες πεντάδες. Ο πυρήνας των 25 νευρώνων αποτελεί τον τοπικό τομέα ανάλυσης και κάθε ένα από αυτά αποτελείται από ένα βάρους όπως και προηγουμένως. Κάθε βάρους ανακλά μία ένταση φωτεινότητας η οποία ουσιαστικά αντιπροσωπεύει μία μικρότερη εικόνα-στοιχείο η οποία αναζητείται στην μεγαλύτερη εικόνα. Όσο πιο κοντά είναι οι φωτεινότητες της μεγάλης εικόνας στις φωτεινότητες των βαρών (μικρή εικόνα) τόσο πιο πιθανό είναι να βρέθηκε το στοιχείο που αντιπροσωπεύουν αυτές οι 25 φωτεινότητες μέσα σε αυτή (Nielsen, 2015).



Εικόνα 3.4 Συνέλιξη εικόνας 28x28 για Πυρήνα 5x5 (NIELSEN, 2015)

Με άλλα λόγια, εάν οι 25 φωτεινότητες ανιχνεύουν την ανθρώπινη μύτη, εισάγοντας μία εικόνα ενός προσώπου, στο σημείο της μύτης θα υπάρχει σχετικά μεγάλη ταύτιση. Παρότι οι άνθρωποι ανιχνεύουν πράγματα με φυσική υπόσταση, τα δίκτυα, όπως προαναφέρθηκε, κάνουν την ίδια διαδικασία χωρίς να «σκέφτονται». Αυτό σημαίνει ότι οι 25 φωτεινότητες που χρησιμοποιούνται για την ανίχνευση κάποιου χαρακτηριστικού της εικόνας δεν σημαίνει ότι αντιπροσωπεύουν κάτι, αλλά απλώς μία σειρά φωτεινοτήτων. Το δίκτυο εκπαιδεύεται, δηλαδή αλλάζει τα βάρη του, με τον τρόπο που αναφέρθηκε στο Κεφάλαιο 3.5. Αλλαγή βαρών σημαίνει όμως και αλλαγή φωτεινοτήτων, που αντίστοιχα σημαίνει αλλαγή του στοιχείου που ανιχνεύεται. Με αυτό τον τρόπο ένα στοιχείο που εμφανίζεται πολλές φορές, μεταβάλλει τα βάρη του δικτύου με τέτοιο τρόπο ώστε να το αναγνωρίζουν και τελικά να γίνεται η ανίχνευσή του, η οποία και αποτελεί ένα από τα μεγαλύτερα

πλεονεκτήματα των συνελκτικών δικτύων (**Feature Detection**). Το σύνολο των 25 νευρώνων έχει και μία «προκατάληψη» άρα σύνολο 26 παραμέτρους (25 βάρη και 1 «προκατάληψη»).

Ένα επίσης πολύ μεγάλο πλεονέκτημα της μεθόδου αυτής είναι ότι η κίνηση αυτού του τοπικού τομέα ανάλυσης σε όλο το μήκος και ύψος της εικόνας διατηρεί τα ίδια βάρη (**Weight Sharing**) εφόσον ανιχνεύει το ίδιο στοιχείο σε όλο το μήκος και ύψος της εικόνας. Η χρήση των ίδιων 25 βαρών για τους νευρώνες όλου του προηγούμενου επιπέδου αποτελεί μία μεγάλη ελάφρυνση για το δίκτυο εφόσον αντί για $28 \times 28 \times 24 \times 24 = 451584$ βάρη (όλα με όλα) και δύο «προκαταλήψεις», μία για κάθε επίπεδο, χρησιμοποιούνται μόνο 25 βάρη και 1 «προκατάληψη» για εκμάθηση (ανακριβές και εξηγείται παρακάτω).

Βεβαίως σε αυτό το σημείο πρέπει να αναφερθεί ότι η ανίχνευση ενός μόνο στοιχείου δεν είναι συχνά αρκετή για να ταυτοποιηθεί κάτι. Έστω π.χ. ότι αναζητούνται οι ρόδες σαν στοιχείο για την αναγνώριση των αυτοκινήτων. Ρόδες όμως έχουν και τα παιδικά καροτσάκια και τα δίκυκλα και άλλα αντικείμενα. Αναζητούνται δηλαδή περισσότερα στοιχεία σε μία εικόνα τα οποία να ταυτοποιούν αυτό που επιδιώκεται να προσδιορισθεί. Για παράδειγμα για τα αυτοκίνητα θα μπορούσαν να χρησιμοποιούνται για αναγνώριση οι καθρέφτες, τα τζάμια, το χερούλι της πόρτας κ.α.

Όπως ήδη αναφέρθηκε, ανίχνευση διαφορετικών στοιχείων σημαίνει ανίχνευση διαφορετικών φωτεινοτήτων άρα και βαρών. Επαναλαμβάνοντας δηλαδή την ίδια διαδικασία πολλές φορές σε μία εικόνα με άλλη σειρά φωτεινοτήτων, ανιχνεύονται διαφορετικά στοιχεία. Εφόσον η ανίχνευση των στοιχείων αυτών γίνεται στις δύο διαστάσεις των εικόνων, την έννοια αυτή την δίνει η τρίτη διάσταση δηλαδή το βάθος. Στο πρώτο επίπεδο το βάθος πάντα ισούται με το 3 για τις έγχρωμες εικόνες και με το 1 για τις ασπρόμαυρες. Στα επόμενα επίπεδα με βάση το πόσα στοιχεία αναζητούνται στις εικόνες αλλάζει. Δηλαδή στο πρώτο επίπεδο συνέλιξης σύμφωνα με αυτά που περιγράφηκαν η εικόνα $28 \times 28 \times 3$ και συνέλιξη με πυρήνα $5 \times 5 \times 3$ (ο πυρήνας έχει πάντοτε βάθος ίσο με το βάθος του επιπέδου που εφαρμόζεται η συνέλιξη) για είκοσι στοιχεία αναζήτησης θα δημιουργήσει ένα νέο επίπεδο με διαστάσεις $24 \times 24 \times 20$ (24×24 όπως περιγράφηκε για τον πυρήνα και $\times 20$ όπως περιγράφηκε για τον αριθμό των στοιχείων αναζήτησης).

Εδώ λοιπόν, αφού δόθηκε και η έννοια της τρίτης διάστασης ξαναδιατυπώνεται σωστά ο αριθμός των παραμέτρων. Στο συγκεκριμένο παράδειγμα χρησιμοποιούνται όχι 25 αλλά $5 \times 5 \times 3 = 75$ βάρη και μία «προκατάληψη» αντί των 451584 βαρών και των δύο «προκαταλήψεων». Στην συνέχεια, σε δεύτερη φάση είναι πιθανό να

αναζητούνται επιμέρους στοιχεία στα στοιχεία του προηγούμενου επιπέδου. Δηλαδή, μία ρόδα π.χ. μπορεί να έχει επιμέρους στοιχεία μία ζάντα, ένα ελαστικό, μια βαλβίδα ή τα μάτια σε ένα ανθρώπινο πρόσωπο μπορεί να έχει σαν επιμέρους στοιχεία τον βολβό, τις βλεφαρίδες κ.α. (Εικόνα 3.5) Επομένως συνεχίζεται η συνέλιξη με ανίχνευση στοιχείων στα στοιχεία του προηγούμενου επιπέδου μέχρι το βαθμό της λεπτομέρειας που επιθυμεί ο ερευνητής. Όπως προαναφέρθηκε για τον ανθρώπινο εγκέφαλο πρέπει υποχρεωτικά τα στοιχεία που αναζητούνται να έχουν φυσική ερμηνεία. Αντίθετα για το δίκτυο, αρκεί να εμφανίζονται συχνά για να τα «μάθει» και να προσαρμόσει ανάλογα τα βάρη του.



Εικόνα 3.5 Ανίχνευση Επιμέρους Στοιχείων του Προσώπου και Επιμέρους Στοιχείων αυτών.
(DETTMERS, 2015)

Η παραπάνω περίπτωση που περιγράφηκε είναι μία από τις βασικές περιπτώσεις. Γενικότερα, ο πυρήνας συνέλιξης έχει τα εξής χαρακτηριστικά:

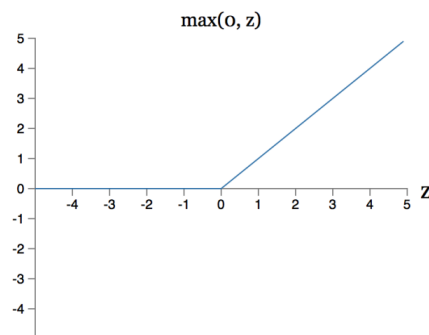
- Μετακινείται συνήθως με βήμα (S) 1 και πιο σπάνια με βήμα 2 ή ακόμα και 3. Βήματα μεγαλύτερα του 1 μικραίνουν απότομα τους όγκους και απαιτούν προσοχή.
- Οι συνηθισμένες τιμές της συνέλιξης (F) είναι μεταξύ 5x5 και 3x3 αλλά αν είναι αναγκαία μεγαλύτερη συνέλιξη χρησιμοποιείται και 7x7 ή 9x9 στα πρώτα και μόνο επίπεδα.
- Ο συνδυασμός βημάτων και τιμής συνέλιξης μερικές φορές δεν κλείνει στο μέγεθος του επιπέδου που εφαρμόζεται. Για παράδειγμα για διάσταση $W=10$ με βήμα $S=2$ και συνέλιξη $F=3$ ή δεν θα συνελιχτεί ο 10^{05} νευρώνας ή θα περισσέψουν δύο. Αυτό προκαλεί προβλήματα στο δίκτυο επειδή αφενός δημιουργεί μη ακέραιους αριθμούς νευρώνων και αφετέρου μειώνει την απόδοσή του εφόσον χάνεται η πληροφορία των άκρων. Για το λόγο αυτό χρησιμοποιούνται βοηθητικά μπλοκ (Pads) ορισμένες φορές με σκοπό να

άρουν το παραπάνω πρόβλημα. $P_{ad}=P=1$ σημαίνει 1 μπλοκ επιπλέον (1 νευρώνα επιπλέον με τιμή 0) σε κάθε πλευρά του δικτύου (Εικόνα 3.6). Επομένως π.χ. $P=1$ προσθέτει σε κάθε γραμμή/στήλη δύο επιπλέον μπλοκ. Στο παραπάνω παράδειγμα, τα δύο που περίσσευαν πλέον δεν αποτελούν πρόβλημα. Η φόρμουλα $(W-F+2P)/S+1$ δίνει απευθείας σύμφωνα με τους παραπάνω κανόνες τους νευρώνες του επόμενου επιπέδου. Για να έχει ίδιο όγκο το προηγούμενο και το επόμενο επίπεδο αρκεί να τεθεί $P=(F-1)/2$.

0	0	0	0	0	0
0	35	19	25	6	0
0	13	22	16	53	0
0	4	3	7	10	0
0	9	8	1	3	0
0	0	0	0	0	0

Εικόνα 3.6 $P=1$ και για τις δύο διαστάσεις σε Πυρήνα Συγκέντρωσης 4×4

Τέλος, πάντοτε μετά το επίπεδο συνέλιξης, ακολουθεί ένα επίπεδο δραστηριοποίησης, στο οποίο εφαρμόζεται η μη γραμμική συνάρτηση όπως και στα απλά νευρωνικά δίκτυα. Το επίπεδο αυτό δεν μεταβάλλει από καμία άποψη την δομή του δικτύου, απλά εφαρμόζει την συνάρτηση σε όλους τους νευρώνες. Η συνάρτηση που έχει επικρατήσει να χρησιμοποιείται για τα συνελκτικά δίκτυα είναι η συνάρτηση Rectified Linear Unit= $ReLU=\max(0,z)$ (Διάγραμμα 3.1).



Διάγραμμα 3.1 Συνάρτηση RELU

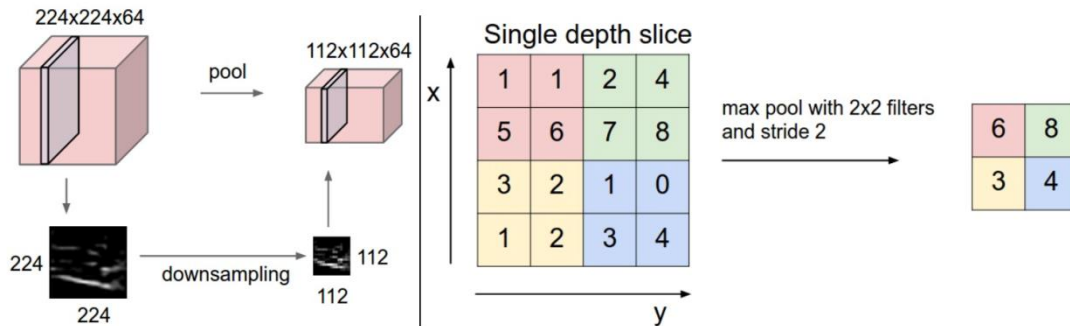
Ο λόγος της επικράτησής της είναι καθαρά εμπειρικός. Σχετικά με την ανωτερότητα της συνάρτησης αυτής για συνελκτικά δίκτυα έχουν γραφτεί αρκετά άρθρα. Συγκεκριμένα, ο Alex Krizhevsky σε έρευνα του ισχυρίζεται ότι ένα συνελκτικό δίκτυο 4 επιπέδων φτάνει στο επίπεδο λάθους 25% έξι φορές γρηγορότερα με χρήση της RELU αντί της tanh (Krizhevsky et al., 2012).

Δύο λόγοι που θα μπορούσαν να ενισχύσουν την ανωτερότητα της RELU είναι η μικρότερη πιθανότητα της εξάλειψης της παραγώγου και η αραιή αναπαράσταση. Συγκεκριμένα, εφόσον $z=wx+b$ όταν $z \geq 0$, η παράγωγος αυτής της συνάρτησης είναι σταθερά σε αντίθεση με της σ , η οποία μειώνεται όλο και γρηγορότερα όσο η απόλυτη τιμή του x μεγαλώνει. Με αυτή την ιδιότητα τα δίκτυα που χρησιμοποιούν RELU «μαθαίνουν» γρηγορότερα. Από την άλλη όταν $z \leq 0$ υπάρχει αραιότερη αναπαράσταση εφόσον οι αντίστοιχες σ θα έδιναν και κάποιες τιμές μεγαλύτερες του 0. Εμπειρικά, έχει φανεί ότι αραιότερες απεικονίσεις είναι θετικές για το δίκτυο άρα και πάλι η χρήση της RELU υπερτερεί των υπολοίπων συναρτήσεων.

3.6.4 ΕΠΙΠΕΔΑ ΣΥΓΚΕΝΤΡΩΣΗΣ

Όπως ήδη αναφέρθηκε στο Κεφάλαιο 3.6.3 τα **Επίπεδα Συγκέντρωσης** είναι η δεύτερη βασική έννοια για τα συνελκτικά δίκτυα. Τα επίπεδα συγκέντρωσης εισάγονται πολλές φορές ανάμεσα σε διαδοχικά επίπεδα συνέλιξης. Ο αποκλειστικός τους στόχος είναι η μείωση του μεγέθους του δικτύου ώστε να μειωθούν ακόμη περαιτέρω οι παράμετροι. Εφαρμόζεται ξεχωριστά σε όλα τα «βάθη» χωρίς όμως ποτέ να μεταβάλλει τον αριθμό τους. Η μείωση του όγκου γίνεται στις άλλες δύο διαστάσεις, πλάτος και ύψος. Συνηθίζεται να εφαρμόζεται συγκέντρωση με μέγεθος $F=2$ και βήμα $S=2$. Το αποτέλεσμα του είναι ο υποτετραπλασιασμός του δικτύου, εφόσον μειώνεται στο μισό η κάθε πλευρά. Η επικρατέστερη μέθοδος είναι η μέθοδος του Μεγίστου (MaxPool), η οποία επικράτησε λόγω των καλύτερων αποτελεσμάτων της από την μέθοδο του AveragePool. Η μέθοδος απλά επιλέγει να διατηρήσει στο επόμενο επίπεδο ο νευρώνας με την μέγιστη τιμή από τους $2 \times 2 = 4$ νευρώνες, που αντιμετωπίζει ο πυρήνας ανά βήμα του (Εικόνα 3.7) αντί να επιλέγει την μέση τιμή. Υπενθυμίζεται εδώ ότι εφόσον το βάθος παραμένει σταθερό σε αυτό το επίπεδο, ο πυρήνας είναι πάντα δισδιάστατος και εφαρμόζεται σε κάθε βάθος ξεχωριστά. Συγκεκριμένα, αν το δίκτυο είχε διάσταση $24 \times 24 \times 20$, όπως προηγουμένως, μετά το επίπεδο συγκέντρωσης, με τις παραπάνω τιμές των παραμέτρων, θα είχε διάσταση $12 \times 12 \times 20$ ($24/2, 24/2, 20$ =σταθερό). Άλλοι ερευνητές, προτιμούν να αποφεύγουν τα

επίπεδα συγκέντρωσης και να χρησιμοποιούν μεγαλύτερες τιμές στα επίπεδα συνέλιξης για να μειώσουν από εκεί το μέγεθος του δικτύου τους. Τα επίπεδα συγκέντρωσης μέχρι σήμερα όταν εφαρμόζονται, χρησιμοποιούν στην συντηρηπτική τους πλειοψηφία $F=2$ και $S=2$. Σπάνια χρησιμοποιείται $S=3$ διότι γίνεται πολύ απότομη συμπίκνωση της εικόνας.



Εικόνα 3.7 Μέγιστη Συγκέντρωση και Τρισδιάστατη Απεικόνισή της

3.6.5 ΑΛΛΗΛΟΥΧΙΑ ΕΠΙΠΕΔΩΝ ΚΑΙ ΕΝΑ ΠΑΡΑΔΕΙΓΜΑ

Έχοντας εξηγήσει όλες τις επιμέρους δομές και λειτουργίες των συνελκτικών δικτύων παρουσιάζεται παρακάτω μία ολοκληρωμένη απεικόνισή τους. Τα συνελκτικά δίκτυα είναι στην πλειονότητά τους της μορφής:

$$\text{Δεδομένα} \rightarrow [[\text{Ε.Συνελ.} \rightarrow \text{Ε.Δ.}] * N \rightarrow \text{Ε.Συγκ.}] * M \rightarrow \text{Επιπεδ.} \rightarrow [\text{Π.Σ.} \rightarrow \text{Ε.Δ.}] * K \rightarrow \text{Π.Σ.}$$

Όπου Ε.Συνελ. = Επίπεδο Συνέλιξης, Ε.Δ. = Επίπεδο Δραστηριοποίησης, Ε.Συγκ. = Επίπεδο Συγκέντρωσης, Επιπεδ. = Επιπεδοποίηση, Π.Σ. = Πλήρως Συνδεδεμένο και $N \geq 0$ και συνήθως ≤ 3 , $K \geq 0$ και συνήθως ≤ 3 και $M \geq 0$.

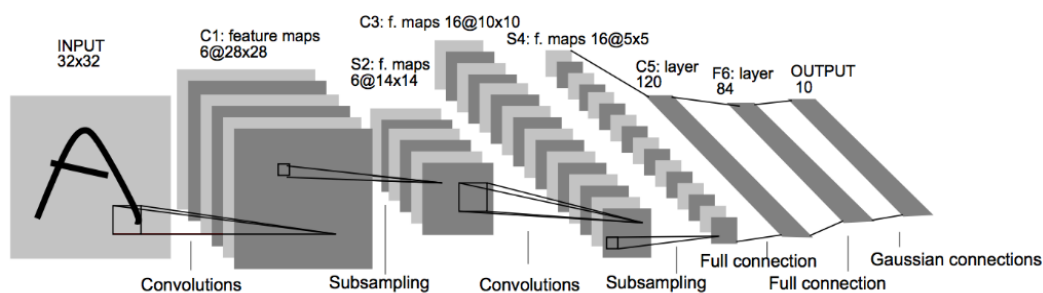
Συγκεκριμένα, τα δεδομένα εισάγονται στο δίκτυο και ύστερα ακολουθεί μία αλληλουχία επιπέδων με τους εξής συνδυασμούς: Πρώτα το επίπεδο συνέλιξης που ακολουθείται πάντα από το επίπεδο δραστηριοποίησης αποτελούν ένα πακέτο, το οποίο επαναλαμβάνεται έως και 3 φορές πριν ακολουθήσει ένα επίπεδο συγκέντρωσης, εάν θέλει ο ερευνητής να το εισάγει στο δίκτυό του. Στην συνέχεια όλο αυτό το σχήμα επαναλαμβάνεται κάποιες φορές. Επειδή η απάντηση που ζητά ο ερευνητής πρέπει να δοθεί σε μία διάσταση πρέπει να μετατραπεί το δίκτυο από

τρισδιάστατο σε μονοδιάστατο. Εφαρμόζεται επομένως, μια διαδικασία που ονομάζεται επιπεδοποίηση, η οποία παραθέτει τους νευρώνες της τελευταίας τρισδιάστατης μορφής του δικτύου και τους μετατρέπει σε ένα διάνυσμα (1 διάσταση). Στην συνέχεια σαν απλό νευρωνικό δίκτυο πια, χρησιμοποιούνται μέχρι δύο επίπεδα πλήρους σύνδεσης, μέχρι να καταλήξει το δίκτυο στο τελευταίο επίπεδο που δίνει το αποτέλεσμα.

Σε αυτό το σημείο πριν δοθεί ένα πλήρες παράδειγμα πρέπει να γίνει μία αναφορά στην συνάρτηση μεταφοράς του τελευταίου επιπέδου (Nielsen, 2015). Συνήθως, στα προβλήματα κατηγοριοποίησης και στα συνελκτικά δίκτυα χρησιμοποιείται η συνάρτηση Softmax $\alpha = e^z / \sum e^z$. Η συνάρτηση αυτή επειδή χρησιμοποιεί την e είναι πάντοτε θετική και από την μορφή της γίνεται αντιληπτό ότι είναι πάντοτε ανάμεσα στο 0 και το 1. Το μεγάλο πλεονέκτημα αυτής της συνάρτησης είναι ότι το αποτέλεσμά της δίνει στον ερευνητή την πιθανότητα με βάση το δίκτυο να έχει κάνει σωστά την εκάστοτε πρόβλεψη (ενδεικτικό είναι ότι το άθροισμα όλων των πιθανών ενδεχομένων είναι πάντοτε 1). Για τον παραπάνω λόγο προτιμάται από τους περισσότερους ερευνητές παγκοσμίως.

Επίσης, σε γενικές γραμμές προτιμάται να χρησιμοποιούνται τετράγωνες φωτογραφίες αφενός για να μειώνονται οι πράξεις και αφετέρου ώστε οι υπολογισμοί που γίνονται στην μία διεύθυνση να είναι ίδιοι με της άλλης, χωρίς βέβαια αυτό να σημαίνει ότι είναι απαραίτητο.

Ακολουθεί ένα απλό παράδειγμα νευρωνικού δικτύου και η επεξήγησή του σύμφωνα με όσα ειπώθηκαν σε αυτό το κεφάλαιο (Εικόνα 3.9). Το παράδειγμα είναι μία παραλλαγή του LeNet-5 (LeCun et al., 1998) ένα από τα πιο θεμελιώδη δίκτυα που ξεκίνησαν την επανάσταση της υπολογιστικής όρασης (Εικόνα 3.8).



CNN called LeNet by Yann LeCun (1998)

Εικόνα 3.8 Δίκτυο LUNET-5

Layer	Type	Func	Input y	Input x	Input	F (y)	F (x)	Dim F	No F	S str y	S str x	P Pad y	P Pad x	Output y	Output x	Output
0	Inp		32	32	3	-	-	-	-	-	-	-	-	32	32	3
1	conv	-	32	32	3	5	5	3	6	1	1	0	0	28	28	6
2	act	relu	28	28	6	-	-	-	-	1	1	0	0	28	28	6
3	pool	max	28	28	6	2	2	-	-	2	2	0	0	14	14	6
4	conv	-	14	14	6	5	5	6	16	1	1	0	0	10	10	16
5	act	relu	10	10	16	-	-	-	-	1	1	0	0	10	10	16
6	pool	max	10	10	16	2	2	-	-	2	2	0	0	5	5	16
7	flatten	-	5	5	16	-	-	-	-	-	-	-	-	400	1	1
8	FC	-	400	1	1	-	-	-	120	-	-	-	-	120	1	1
9	FC	-	120	1	1	-	-	-	84	-	-	-	-	84	1	1
10	FC	-	84	1	1	-	-	-	10	-	-	-	-	10	1	1
11	Softmax	-	10	1	1	-	-	-	-	-	-	-	-	10	1	1

Εικόνα 3.9 Παραλλαγή Συνελικτικού δικτύου LENET-5

Ο Yann LeCun προσπάθησε να λύσει το περιβόητο MNIST πρόβλημα, δηλαδή τον διαχωρισμό των αριθμών από το 0 έως το 9 (LeCun et al., 1998). Χρησιμοποιώντας ένα δίκτυο πολύ κοντά με αυτό της Εικόνας, πέτυχε εντυπωσιακά αποτελέσματα για το εγχείρημά του. Μία από τις διαφορές του δικτύου αυτού και του LeCun είναι ότι χρησιμοποίησε ασπρόμαυρες εικόνες ενώ εδώ τοποθετήθηκαν έγχρωμες. Η μόνη διαφορά έγκειται στο βάθος του πρώτου επίπεδου που ισούται με 3. Έστω λοιπόν μία εισερχόμενη εικόνα μεγέθους $32 \times 32 \times 3$ που υπόκειται δύο φορές στην αλληλουχία των επιπέδων συνέλιξης-δραστηριοποίησης-συγκέντρωσης με τιμές F, πέντε-μηδεν-δύο αντίστοιχα και βήμα S ένα-ένα-δύο αντίστοιχα. Επιπλέον μπλοκ δεν προστέθηκαν εφόσον δεν χρειαζόντουσαν $P=0$. Επομένως οι διαστάσεις x και y διαδοχικά μεταβάλλονται ως εξής $32-4=28$ (Συνέλιξη 5 με βήμα 1) $\rightarrow 28$ (ίδια διάσταση για το επίπεδο δραστηριοποίησης) $\rightarrow 14$ (Συγκέντρωση τιμής 2 με βήμα 2) $\rightarrow 10$ (Συνέλιξη 5 με βήμα 1) $\rightarrow 10$ (ίδια διάσταση για το επίπεδο δραστηριοποίησης) $\rightarrow 5$ (Συγκέντρωση τιμής 2 με βήμα 2). Τα βάθη επιλέγονται από τον ερευνητή στο εκάστοτε επίπεδο συνέλιξης. Στην συνέχεια ακολουθεί η επιπεδοποίηση. Παρατηρείται εδώ η ισότητα των νευρώνων πριν και μετά την επιπεδοποίηση ($5 \times 5 \times 16 = 400 \times 1 \times 1$). Έπονται δύο επίπεδα πλήρως συνδεδεμένα με αριθμό νευρώνων όσα επιλέξει ο ερευνητής και τέλος το τελικό επίπεδο στο οποίο χρησιμοποιείται η συνάρτηση Softmax έχει δέκα νευρώνες, ένα για κάθε ενδεχόμενο 0-9.

Συνοψίζοντας, όλο το παραπάνω αποτελεί την διαδικασία, στην οποία υποβάλλεται μία εικόνα όταν εισέρχεται σε αυτό το δίκτυο. Η εκπαίδευσή του είναι αντίστοιχη με των απλών νευρωνικών δικτύων. Αναφέρεται εδώ και πάλι, ότι όλες οι μέθοδοι εκπαίδευσης και οι πιθανές μορφές του δικτύου για μέγιστη απόδοση είναι υπό συνεχή διερεύνηση και προτιμητέες λύσεις με τεκμηριωμένο τρόπο πέραν κάποιων βασικών κανόνων δεν έχουν μέχρι σήμερα διατυπωθεί.

3.7 ΕΠΙΛΟΓΗ ΛΟΓΙΣΜΙΚΟΥ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΟΙ ΠΕΡΙΟΡΙΣΜΟΙ

Επιλέχτηκε να χρησιμοποιηθεί η γλώσσα προγραμματισμού R και το πιο φιλικό στον χρήστη πρόγραμμα RStudio για την εκπόνηση της παρούσας εργασίας, που αποτελεί μία από τις πιο εξελισσόμενες πλατφόρμες ανοικτού κώδικα για την ανάπτυξη προτύπων εφαρμοσμένης στατιστικής και υπολογιστικής νοημοσύνης. Ακόμη, το πακέτο `mxnet` που προσφέρει η R για ανάλυση συνελκτικών δικτύων εκδόθηκε μόλις τον Νοέμβρη του 2015 επομένως αποτελούσε μία πολύ ενημερωμένη και δυνατή επιλογή. Ακόμη, δοκιμάζοντας ποικίλες διεργασίες από αυτές που αναφέρθηκαν παραπάνω, ήταν εμφανές ότι το πρόγραμμα μπορούσε να ανταπεξέλθει για αυτά που χρειαζόταν. Επομένως, εφόσον το πρόγραμμα κάλυπτε τις ανάγκες της έρευνας, και ο ερευνητής ήταν εξοικειωμένος με αυτό επιλέχτηκε χωρίς περαιτέρω έρευνα.

Η ανάλυση των συνελκτικών δικτύων μπορεί να γίνει πολύ δύσκολη αν αναλογιστεί κανείς τους περιορισμούς του χώρου του ηλεκτρονικού υπολογιστή. Η R, αποθηκεύει τα προσωρινά της αρχεία στην μνήμη RAM του υπολογιστή για να είναι πιο προσβάσιμα και να πραγματοποιούνται οι διαδικασίες πολύ πιο γρήγορα (να μην καλούνται από αλλού). Βεβαίως όταν πρόκειται για ανάλυση ενός πολύ μεγάλου αριθμού εισερχόμενων δεδομένων, στην συγκεκριμένη περίπτωση εικόνων, δημιουργούνται θέματα χώρου εφόσον η μνήμη RAM γεμίζει και στην συνέχεια αρχίζει να γεμίζει και ο σκληρός δίσκος. Στην περίπτωση που γεμίσει και ο σκληρός δίσκος η διαδικασία σταματά και δεν μπορεί να συνεχίσει. Στην περίπτωση που απλά απαιτείται χώρος από τον σκληρό δίσκο η διαδικασία πραγματοποιείται αλλά πολύ πιο αργά. Η λύση αγοράς μεγαλύτερης RAM, έχει και αυτή περιορισμούς. Επομένως, είτε πρέπει να γίνει σταδιακή εκπαίδευση του δικτύου με μικρότερα σύνολα ενώ όλες οι φωτογραφίες έχουν «ανέβει» και καταναλώσει κάποιο ποσοστό της RAM, είτε να ανεβαίνουν και οι φωτογραφίες ανά μικρά υποσύνολα, να εκπαιδεύουν το δίκτυο και να αποσύρονται από την μνήμη RAM.

Ένα ακόμη σημαντικό θέμα προς επίλυση είναι το θέμα του χρόνου. Οι πιο σύγχρονοι υπολογιστές μπορούν να κάνουν την ανάλυση με την χρήση της κάρτας γραφικών τους GPU βελτιώνοντας πολύ την απόδοση του συστήματος από άποψη χρόνου (έως και 10 φορές γρηγορότερα). Στην συνέχεια κατάλληλη επιλογή των παραμέτρων του

δικτύου οδηγούν σε πολύ μειωμένους χρόνους ανάλυσης για την ίδια μορφή του δικτύου.

Λαμβάνοντας υπόψη αυτές τις παραμέτρους στην παρούσα διπλωματική εργασία δε χρησιμοποιούνται πολύ μεγάλα δίκτυα για να μην προκύπτουν θέματα χώρου και παράλληλα γίνεται και μία ορθή επιλογή του όγκου των δεδομένων εισαγωγής στο δίκτυο. Η επιλογή των παραμέτρων των δικτύων έλαβε υπόψη τους παραπάνω περιορισμούς, αλλά όπως αποδεικνύεται και από την παρακάτω ανάλυση ήταν ιδιαίτερα πετυχημένη.

3.8 ΜΕΘΟΔΟΣ ΑΞΙΟΛΟΓΗΣΗΣ

Η μέθοδος αξιολόγησης των συνελικτικών δικτύων γίνεται εφαρμόζοντας το μοντέλο στο Σύνολο Δοκιμής όπως έχει ήδη αναφερθεί στην παράγραφο 3.5.2. Συγκρίνοντας τις προβλέψεις με την πραγματικότητα προκύπτει το ποσοστό επιτυχίας, το οποίο είναι ο πιο βασικός δείκτης απόδοσης του μοντέλου. Παρόλ' αυτά, χρησιμοποιούνται και κάποια άλλα στατιστικά μέτρα τα οποία βοηθούν στην καλύτερη αποσαφήνιση των αποτελεσμάτων του μοντέλου αλλά και στον καλύτερο προσδιορισμό των λαθών στα οποία υποπίπτει.

Έστω ο Πίνακας Προβλέψεων/Πραγματικότητας και τα Στατιστικά Μέτρα:

	0	1		A,B,Γ,Δ: Αριθμός Δεδομένων
0	A	B	S1	S1,S2,S : Αθροίσματα ανά Γραμμή
1	Γ	Δ	S2	S3,S4,S : Αθροίσματα ανά Στήλη
	S3	S4	S	

Sensitivity : $\Delta/S4$ Specificity : $A/S3$ Λάθος Τύπου I : $\Gamma/S3$

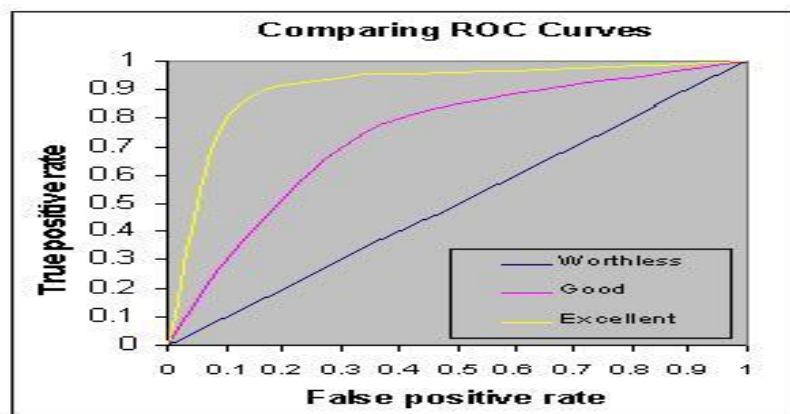
Pos. Pred. Value : $\Delta/S2$ Neg. Pred. Value : $A/S1$ Λάθος Τύπου II : $B/S4$

False Discovery Rate = $1 - \text{Pos. Pred. Value} = \Gamma/S2$

False Not Discovery Rate = $1 - \text{Neg. Pred. Value} = B/S1$

Από τα παραπάνω παρατηρείται η ανάγκη για μεγιστοποίηση όλων των όρων. Όσο μεγαλύτεροι είναι οι όροι και τείνουν στη μονάδα τόσο καλύτερο είναι το χρησιμοποιούμενο μοντέλο. Ενδεικτικά αναζητείται η μεγιστοποίηση της διαγωνίου A-Δ και η ελαχιστοποίηση της διαγωνίου B-Γ. Νοηματικά, είναι επιθυμητό να ανιχνεύει το δίκτυο τους πραγματικούς άσσους ως άσσους και τα πραγματικά μηδενικά ως μηδενικά. Αντίθετα, η κατηγοριοποίηση ενός μηδενικού ως άσσου αποτελεί το Λάθος Τύπου I (ψευδώς θετικό) ενώ η κατηγοριοποίηση ενός άσσου ως μηδενικού αποτελεί το Λάθος Τύπου II (ψευδώς αρνητικό).

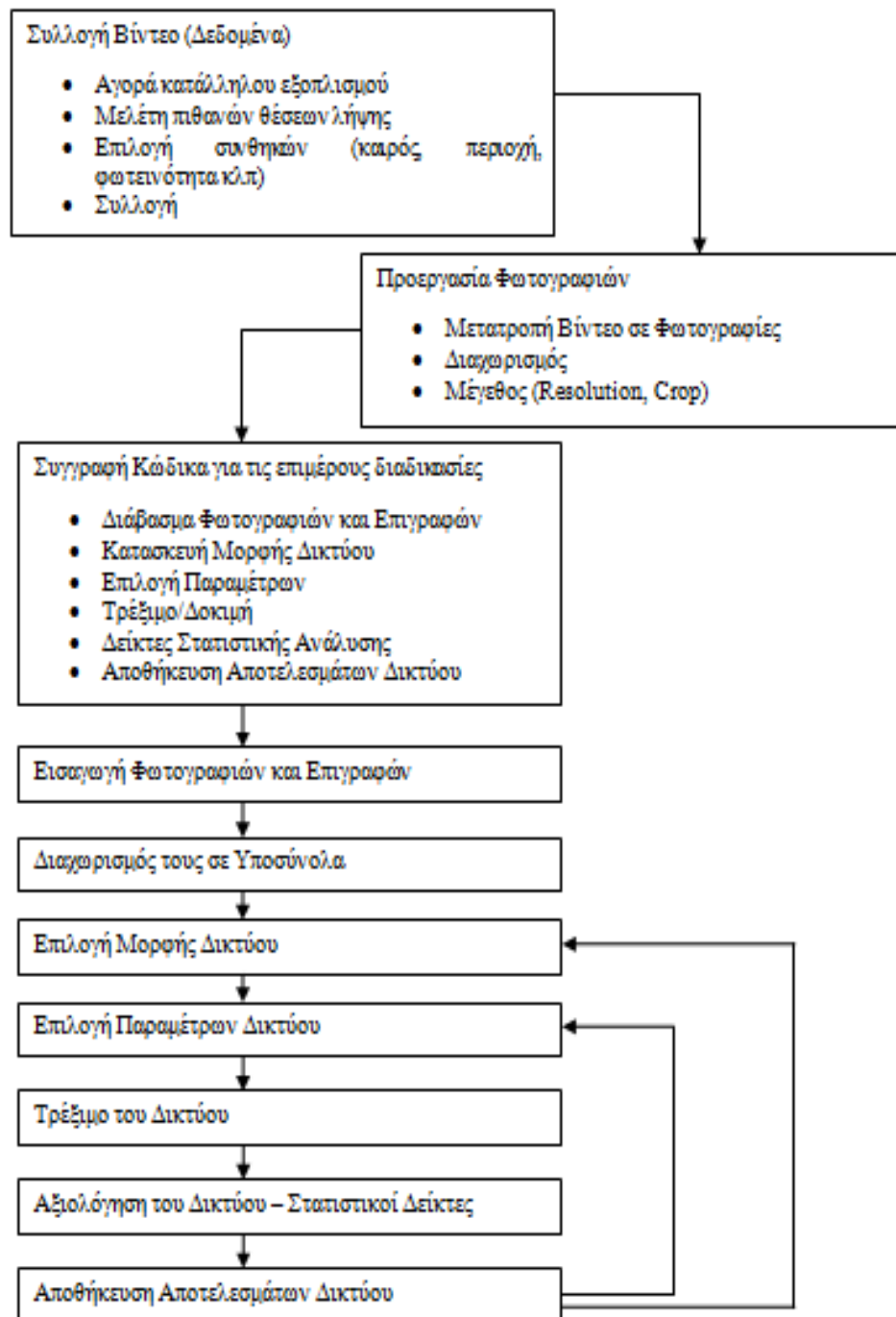
Η καμπύλη ROC αναπαριστά τις Sensitivity και Specificity σε ένα κοινό διάγραμμα. Όσο πιο κοντά στο 1 είναι οι δύο αυτές τιμές τόσο το σημείο αυτό της καμπύλης τείνει πάνω και αριστερά διαμορφώνοντας μία καμπύλη μορφής Γ που είναι το βέλτιστο σχήμα για αυτό το δείκτη. Στην Εικόνα 3.19 ως άξονας X χρησιμοποιείται ο δείκτης $1 - \text{Specificity}$. Ενδεικτικά:



Εικόνα 3.10 Καμπύλη ROC - Όσο αριστερότερα και ψηλότερα τόσο καλύτερο αποτέλεσμα

Ακόμη ενδιαφέρον έχει ο δείκτης AUC, που ισούται με το εμβαδό κάτω από τις καμπύλες ROC. Όσο πιο κοντά στο 1 είναι η τιμή του, τόσο καλύτερη είναι η πρόβλεψη του δικτύου.

3.9 ΔΙΑΓΡΑΜΜΑ ΡΟΗΣ ΠΕΙΡΑΜΑΤΟΣ



4. ΣΥΛΛΟΓΗ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ

Για την πραγματοποίηση της διπλωματικής εργασίας ήταν αναγκαία η συλλογή πολλών δεδομένων, ώστε να μπορούν να εκπαιδευτούν τα συνελκτικά δίκτυα που αναπτύχθηκαν. Για το λόγο αυτό αναπτύχθηκε μία μεθοδική προσέγγιση, η οποία έλαβε υπόψη της όλες τις παραμέτρους ώστε η συλλογή στοιχείων να επιτευχθεί με το λιγότερο κόστος, τον λιγότερο χρόνο, αλλά και τη μεγαλύτερη ακρίβεια.

4.1 ΣΥΛΛΟΓΗ ΒΙΝΤΕΟ

Καταρχήν, παρότι η διπλωματική εργασία θα ασχοληθεί με την ανάλυση φωτογραφιών, επιλέχθηκε να γίνει συλλογή στοιχείων με χρήση βίντεο. Η απόφαση αυτή βασίζεται στη δυνατότητα που παρέχει ένα βίντεο στον ερευνητή να παράγει πολλές φωτογραφίες από αυτό και να επιλέξει το χρονικό διάστημα μεταξύ τους. Με αυτόν τον τρόπο ο ερευνητής, μπορεί να παράγει έναν πολύ μεγάλο αριθμό φωτογραφιών όποια χρονική στιγμή επιθυμεί και όσο πυκνά επιθυμεί. Ακόμη, το

βίντεο επιτρέπει στον ερευνητή να έχει πιο πλήρη εικόνα των στοιχείων που θα εισάγει στο δίκτυό του και του επιτρέπει να κάνει διορθώσεις όπως αυτός επιθυμεί παράγοντας νέα δεδομένα από το αρχικό του βίντεο.

4.1.1 ΕΞΟΠΛΙΣΜΟΣ

Για την συλλογή των δεδομένων έπρεπε να επιλεγεί εξοπλισμός ο οποίος να μπορεί να στηριχτεί εξωτερικά του οχήματος στη θέση και το ύψος που θα επιλέγονταν για την έρευνα. Ακόμη θα έπρεπε να έχει αξιοπρεπείς επιδόσεις σε ανάλυση, οπτικό πεδίο και διάρκεια του βίντεο. Για τον λόγο αυτό επιλέχτηκε για κάμερα η Go Pro Hero 4 (Εικόνα 4.1) της οποίας τα χαρακτηριστικά βίντεο παρατίθενται στον Πίνακα 4.1.

Πίνακας 4.1 Χαρακτηριστικά Βίντεο Κάμερας

VIDEO SPECS (by resolution)	
4K	30, 25, 24 fps Ultra Wide FOV
4K SUPERVIEW	24 fps Ultra Wide FOV
2.7K	50, 48, 30, 25, 24 fps Ultra Wide, Medium FOV
2.7K SUPERVIEW	30, 25 fps Ultra Wide FOV
2.7K 4:3	30, 25 fps Ultra Wide FOV
1440p	80, 60, 50, 48, 30, 25, 24 fps Ultra Wide FOV
1080p	120, 90, 60, 50, 48, 30, 25, 24 fps Ultra Wide, Medium (excludes 120 and 90 fps), Narrow FOV
1080p SUPERVIEW	80, 60, 50, 48, 30, 25, 24 fps Ultra Wide FOV
960p	120, 60, 50 fps Ultra Wide FOV
720p	120, 60, 50, 30, 25 fps Ultra Wide, Medium, Narrow FOV
720p SUPERVIEW	120, 60, 50 fps Ultra Wide FOV
WVGA	240 fps Ultra Wide FOV



Εικόνα 4.1 Go Pro Hero 4

Επιπλέον κρίθηκε αναγκαίο να χρησιμοποιηθεί ένας μηχανισμός στήριξης και προστασίας της κάμερας (Εικόνα 4.2) ο οποίος αντέχει σε ταχύτητες ακόμη και 200χμ/ώρα.



Εικόνα 4.2 Μηχανισμός Στήριξης και Προστασίας της Κάμερας

4.1.2 ΕΠΙΛΟΓΗ ΘΕΣΗΣ ΛΗΨΗΣ

Ο τρόπος λήψης βίντεο είναι ιδιαίτερα σημαντικός για την επιτυχία κάθε τεχνικής που δοκιμάζεται. Σε αυτήν την διπλωματική εργασία, με την τεκμηρίωση της βιβλιογραφικής ανασκόπησης, απορρίφθηκε η πανοραμική λήψη και επιλέχθηκε η λήψη από το όχημα. Σε αυτό το υποκεφάλαιο εξετάζεται η ακριβής θέση λήψης επάνω στο όχημα και ο τρόπος που αυτή επιλέχθηκε.

Αρχικά, πραγματοποιήθηκε μια σειρά από δοκιμές με σκοπό να βρεθεί αυτό που επιδιώκεται να απεικονισθεί στο βίντεο και τις φωτογραφίες. Τα δύο θεμελιώδη συμπεράσματα αυτών των προσπαθειών ήταν τα εξής:

1. Όποια πληροφορία υπάρχει εντός της φωτογραφίας και είναι ψηλότερα από το σταθμευμένο όχημα ή το πεζοδρόμιο σε κενή θέση είναι πληροφορία άσχετη με την έρευνα που πραγματοποιείται
2. Είναι σημαντικό η ασφάλτος του δρόμου να αποτελεί ένα ικανό ποσοστό της εικόνας ώστε να ανιχνεύεται.

Το τελευταίο είναι αναγκαίο διότι αποτελεί ένα δυνατό κριτήριο που βοηθά το δίκτυο να το κατανοήσει. Μεγαλύτερο ποσοστό ασφάλτου υποδεικνύει προφανώς κενή θέση και αντιστρόφως.

Από τα παραπάνω, προκύπτει ότι η τοποθέτηση της κάμερας πρέπει να είναι αυτή κατά την οποία το άνω όριο της εξαγόμενης φωτογραφίας, είναι οριακά πάνω από το επίπεδο που περνά από την οροφή του σταθμευμένου οχήματος, είτε υπάρχει είτε όχι. Η μέση απόσταση από το πεζοδρόμιο υπολογίστηκε στα 3.5 μέτρα (Υποκεφάλαιο

3.2) επομένως μικρές μετατοπίσεις προς την μία ή την άλλη κατεύθυνση, αναλόγως με το που περνάει το εκάστοτε όχημα, επηρεάζουν λίγο την απεικόνιση που επιδιώκεται.

Η παραπάνω λογική μπορεί να εφαρμοστεί με τοποθέτηση της κάμερας μόνο στις πλάγιες πλευρές του οχήματος για δύο λόγους. Καταρχήν η χρήση της κάμερας στην εμπρός ή την πίσω πλευρά θα περιείχαν ασφάλτο σε πολύ μεγαλύτερο βαθμό. Αυτό θα είχε σαν αποτέλεσμα η «έκταση» της ασφάλτου να μην αποτελεί το κρίσιμο στοιχείο προσδιορισμού ελεύθερης θέσης στάθμευσης, εφόσον η ελαφρά μείωσή της από σταθμευμένα οχήματα δεν θα γινόταν εύκολα αντιληπτή. Ακόμη σε αυτή την περίπτωση το δίκτυο πρέπει να «μάθει» και την λογική της προοπτικής εικόνας, εφόσον όσο απομακρύνεται ένα αντικείμενο από το όχημα τόσο μικρότερο φαίνεται στην φωτογραφία. Αντίθετα, χρησιμοποιώντας τις κάμερες στις πλαϊνές πλευρές του οχήματος, το μέγεθος των οχημάτων που καταγράφεται είναι πάντοτε σταθερό και το ποσοστό της ασφάλτου στην εικόνα έχει πολύ αισθητές διαφορές από σταθμευμένη σε κενή θέση, βοηθώντας το δίκτυο να αντιληφθεί την αλλαγή.

Όσον αφορά στο μήκος που πρέπει να απεικονίζεται έγινε η υπόθεση στο Υποκεφάλαιο 3.2 ότι τα οχήματα είναι μήκους 4.5 μέτρων και ότι απαιτείται 0.5 μέτρο για την στάθμευσή τους (σύνολο 5 μέτρα). Βεβαίως η απεικόνιση μόνο 5 μέτρων δεν είναι αρκετή, διότι σε μεγάλες ταχύτητες υπάρχει ο κίνδυνος να μην καταγραφεί κάποια θέση. Για ταχύτητα 72 χμ/ώρα = 20 μ/δευτ, ταχύτητα που σπανίως έχει ένα όχημα εντός αστικής περιοχής (λαμβάνεται μία δυσμενέστατη περίπτωση), και για ικανότητα κάμερας να τραβά 10 φωτ/δευτ. προκύπτει ότι λαμβάνεται μία φωτογραφία ανά 2 μέτρα. Για να μην παραλειφθεί επομένως καμία θέση 5 μέτρων, για ταχύτητες μικρότερες των 72 χμ/ώρα και για ικανότητα κάμερας 10 φωτ/δευτ, απαιτούνται $5+2=7$ μέτρα πλάτος εικόνας.

Τα 7 μέτρα πλάτους εικόνας μπορούν να επιτευχθούν με διάφορες τοποθετήσεις της κάμερας. Με ακινητοποιημένο το όχημα και πριν την συλλογή των στοιχείων έγιναν κάποιες δοκιμές για διαφορετικά ύψη τοποθέτησης της κάμερας (h) και για διαφορετικές γωνίες λήψης ως προς την κατακόρυφο (θ), με σκοπό να μετρηθεί το πλάτος της εικόνας σε διάφορες αποστάσεις από τον φακό (d = 2, 3 και 4 μέτρα). Σκοπός ήταν να βρεθεί ένας βέλτιστος συνδυασμός ύψους και γωνίας που να δίνει πλάτος εικόνας στα 3.5 μέτρα από τον φακό (δηλαδή εκεί περίπου που αρχίζει το πεζοδρόμιο 7 μέτρα. Για βίντεο ανάλυσης 1080p Medium έγιναν οι μετρήσεις του Πίνακα 4.2 και τελικά επιλέχτηκε η σημειωμένη με κόκκινο λύση σύμφωνα με την οποία η κάμερα πρέπει να τοποθετηθεί σε ύψος 90 εκατοστά από το έδαφος και με γωνία λήψης 75 μοίρες από την κατακόρυφο.

Πίνακας 4.2 Μετρήσεις Πλάτους Φωτογραφίας για διάφορες Αποστάσεις από τον Φακό (d), διάφορες Γωνίες ως προς την Κατακόρυφο (θ) και διάφορα Ύψη (h)

Type	1080	Medium	Video	
h (cm)	θ (degrees)	2(m)	3(m)	4(m)
110	90	4,2	6,32	8,08
90	90	4,04	6	7,94
70	90	3,8	5,64	7,44
90	75	4,5	6,84	8,74
90	60	4,16	6,3	7,8
90	45	3,9	6	εως 3.3

Να σημειωθεί εδώ ότι η εξασφάλιση του πλάτους της φωτογραφίας σε κάθε χρονική στιγμή, είναι πολύ σημαντική. Αυτό ισχύει επειδή για φωτογραφίες μερικώς καλυμμένες με εμπόδιο, είναι προφανές ότι αν τα 5/7 της εικόνας είναι κενά, αυτά αντιστοιχούν σε 5 από τα 7 μέτρα του πεζοδρομίου που απεικονίζονται και απαιτούνται για την στάθμευση ενός οχήματος. Επομένως, με αυτό τον τρόπο είναι πλέον δυνατό να δοθούν οι χαρακτηρισμοί (0/1) των δεδομένων, αφού βλέποντας μια φωτογραφία μπορεί να αποφασισθεί με μεγάλη σιγουριά η ύπαρξη ή μη θέσης στάθμευσης.

Συνοψίζοντας, επιλέχθηκε λήψη από κάποια πλαϊνή πλευρά του οχήματος, σε ύψος 90 εκατοστά από το έδαφος και με γωνία λήψης 75 μοίρες ως προς την κατακόρυφο, με σκοπό να παράγονται εικόνες-βίντεο που να απεικονίζουν πλάτος 7 μέτρων στο ύψος του πεζοδρομίου και ύψος φωτογραφίας που να εμπεριέχει όσο το δυνατόν γίνεται λιγότερο την περιττή πληροφορία, που αντιστοιχεί στο χώρο επάνω από την οροφή των σταθμευμένων οχημάτων.

4.1.3 ΕΠΙ ΤΟΠΟΥ ΕΡΕΥΝΑ

Αφού μελετήθηκε ο τρόπος λήψης του βίντεο και εξασφαλίστηκε ο κατάλληλος εξοπλισμός πραγματοποιήθηκε η συλλογή δεδομένων.

Ακολουθήθηκαν όλα όσα αναφέρονται στο Υποκεφάλαιο 3.2 για τις παραδοχές και τις υποθέσεις της διπλωματικής εργασίας.

Συγκεκριμένα η συλλογή των στοιχείων έλαβε χώρα κατά τη διάρκεια της ημέρας με καιρικές συνθήκες ηλιοφάνειας ή ελαφριάς συννεφιάς. Επιδιώχθηκε να υπάρχει ρεαλιστική απόσταση από το πεζοδρόμιο, η οποία κυμάνθηκε κατά μέσο όρο στα 3.5 μέτρα. Η συλλογή πραγματοποιήθηκε μόνο σε αστικά οδικά τμήματα μίας λωρίδας ανά κατεύθυνση με επιτρεπόμενη στάθμευση. Η ταχύτητα του οχήματος επιβεβαιώθηκε ότι ήταν πάντοτε μικρότερη από 72 χμ/ώρα.

Επιλέχθηκαν περιοχές της Αθήνας διαφορετικού τύπου και άνεσης οδήγησης λόγω χαρακτηριστικών των οδών όπως η Νέα Ερυθραία, το Γαλάτσι, το Χαλάνδρι και το Ψυχικό.

Το συνολικό βίντεο που τελικά αποκτήθηκε, αφού πρώτα εξετάστηκε η επάρκειά του, ήταν διάρκειας περίπου 1000 δευτερολέπτων.

4.2 ΠΡΟΕΤΟΙΜΑΣΙΑ ΔΕΔΟΜΕΝΩΝ

4.2.1 ΜΕΤΑΤΡΟΠΗ ΒΙΝΤΕΟ ΣΕ ΦΩΤΟΓΡΑΦΙΕΣ

Έχει ήδη εξηγηθεί στο Υποκεφάλαιο 4.1, ο λόγος που επιλέχθηκε η χρήση βίντεο αντί της φωτογραφίας για την απόκτηση πολλών δεδομένων. Έχοντας πλέον όλο το βίντεο σε ηλεκτρονική μορφή, χρησιμοποιήθηκε ένα πρόγραμμα επεξεργασίας βίντεο για τον υπολογιστή, στο οποίο σημειωνόταν η χρονική στιγμή της πρώτης φωτογραφίας και το χρονικό διάστημα ανάμεσα στις διαδοχικές φωτογραφίες. Δημιουργούταν στην συνέχεια, ένας φάκελος με όλες τις φωτογραφίες που προέκυπταν από τις παραπάνω ρυθμίσεις. Η χρονική στιγμή της πρώτης φωτογραφίας ήταν πάντοτε το 0.00 και το χρονικό διάστημα για διαδοχικές φωτογραφίες 0.10 δευτερόλεπτα (10 φωτογραφίες ανά δευτερόλεπτο). Από την μία, δεν πρέπει το

χρονικό διάστημα να είναι πολύ μικρό γιατί οι φωτογραφίες είναι σχεδόν οι ίδιες. Από την άλλη αν χρησιμοποιηθεί πολύ μεγάλο χρονικό διάστημα χάνεται πληροφορία αξιοποιήσιμη και επομένως η επιλογή αυτή δεν αποτελεί πρακτική λύση. Η χρήση του 0.1 δευτερολέπτου είναι μία καλή τιμή με βάση τις παραπάνω παρατηρήσεις. Στην συνέχεια βέβαια αν αυτή οδηγούσε σε ανεπαρκές δείγμα θα χρησιμοποιούνταν αρχική χρονική στιγμή της φωτογραφίας το 0.05 δευτερόλεπτο και πάλι 0.10 χρονικό διάστημα μεταξύ των διαδοχικών φωτογραφιών με αποτέλεσμα να λαμβάνονταν οι ενδιάμεσες φωτογραφίες των προηγούμενων. Αντίθετα αν κρινόταν ότι το 0.1 είναι ήδη πολύ μικρό και το δείγμα πολύ μεγάλο, αρκεί να διαγραφόταν κάθε δεύτερη φωτογραφία, ουσιαστικά διπλασιάζοντας το χρονικό διάστημα μεταξύ διαδοχικών φωτογραφιών.

4.2.2 ΔΙΑΧΩΡΙΣΜΟΣ ΦΩΤΟΓΡΑΦΙΩΝ ΚΑΙ ΥΠΟΠΡΟΒΛΗΜΑΤΑ

Τα 1000 περίπου δευτερόλεπτα βίντεο παρήγαγαν περίπου 10000 φωτογραφίες. Αυτές οι φωτογραφίες διαχωρίστηκαν σε πρώτο στάδιο σε εύκολες και δύσκολες, όπου δύσκολες ορίζονται οι φωτογραφίες οι οποίες δεν είναι ξεκάθαρα ούτε κενές θέσεις, ούτε έχουν κάποιο σταθμευμένο όχημα, αλλά περιέχουν άλλα αντικείμενα ή πρόκειται για ειδικές περιπτώσεις που συναντώνται σπάνια στην οδήγηση (όπως κάδοι απορριμμάτων, κώνοι, εξαρτήματα κλπ.). 7413 φωτογραφίες θεωρήθηκαν ως εύκολες και 1539 ως δύσκολες. Οι υπόλοιπες φωτογραφίες (1048) απορρίφθηκαν από την ανάλυση επειδή δεν πληρούσαν τις υποθέσεις που έγιναν (π.χ. απόσταση από το πεζοδρόμιο κατά την έναρξη ή λήξη της μέτρησης κ.α.).

Σύμφωνα με την μεθοδολογία, ενδείκνυται η επίλυση υποπροβλημάτων πριν την επίλυση του μεγάλου προβλήματος. Για το λόγο αυτό επιδιώχθηκε να επιλυθεί πρώτα το πρόβλημα χρησιμοποιώντας μόνο τις εύκολες φωτογραφίες και ύστερα να ξαναλυθεί με την αποκτηθείσα εμπειρία εισάγοντας και τις δύσκολες. Από τις 7413 φωτογραφίες, 5000 χρησιμοποιήθηκαν ως Σύνολο Εκπαίδευσης, 1500 ως Σύνολο Επικύρωσης και 913 ως Σύνολο Δοκιμής (Γενίκευση).

Στην συνέχεια έπρεπε να προστεθούν και οι δύσκολες φωτογραφίες προς ανάλυση γιατί μόνο τότε έχει εφαρμοσιμότητα το μοντέλο, εφόσον τέτοιες συνθήκες σίγουρα θα υπάρχουν στην πράξη. Όπως παρατηρήθηκε, η αναλογία εύκολων-δύσκολων φωτογραφιών στην πραγματικότητα είναι 5 προς 1. Αυτή περίπου η αναλογία διατηρείται σε όλα τα σύνολα του πειράματος για να ανταποκρίνονται στην πραγματικότητα και η γενίκευση να αποκτά ρεαλιστική χροιά. Από τις

7413+1539=8952 φωτογραφίες, 5000+1000=6000 χρησιμοποιήθηκαν ως Σύνολο Εκπαίδευσης, 1500+300=1800 ως Σύνολο Επικύρωσης και 913+183=1096 ως Σύνολο Δοκιμής.

Τα δεδομένα «καθαρίστηκαν» και από τις περιπτώσεις γεωμετρικής χάραξης που δεν θα μπορούσαν να θεωρηθούν ως θέσεις στάθμευσης, για παράδειγμα οι περιοχές των κόμβων. Στην ανάλυση γίνεται θεώρηση ότι η περιοχή ενός κόμβου είναι δεδομένη χωρικά και ως εκ τούτου το σύστημα δεν θα κληθεί να αποφασίσει για αυτά τα σημεία. Επομένως οι δύσκολες φωτογραφίες με κόμβους έπρεπε να εξαχθούν από τα σύνολα εκπαίδευσης, επικύρωσης και δοκιμής. Τελικά οι 1539 δύσκολες φωτογραφίες περιείχαν 616 φωτογραφίες διασταυρώσεων και 923 φωτογραφίες δύσκολων περιπτώσεων, οι οποίες βαφτίστηκαν «Rest». Θεωρώντας μαζί τις Rest φωτογραφίες με τις εύκολες, με αναλογία εμφάνισης πλέον περίπου 8 προς 1 δημιουργήθηκαν τα σύνολα: 5000+600=5600 φωτογραφίες το Σύνολο Εκπαίδευσης, 1500+180=1680 φωτογραφίες το Σύνολο Επικύρωσης και 913+143=1056 φωτογραφίες το Σύνολο Δοκιμής. Σε αυτά τα σύνολα εκπαιδεύτηκαν τα συνελκτικά δίκτυα που δοκιμάστηκαν στην παρούσα διπλωματική εργασία για πραγματικές συνθήκες.

4.2.3 ΤΕΛΙΚΟ ΣΤΑΔΙΟ ΕΠΕΞΕΡΓΑΣΙΑΣ

Αφού αποφασίστηκε ο διαχωρισμός των φωτογραφιών και τα υποπροβλήματα που θα οδηγήσουν στην επίλυση του μεγαλύτερου προβλήματος, παραμέναν ορισμένες λεπτομέρειες που έπρεπε να καθορισθούν για να ξεκινήσει η επίλυση αλλά και για να βελτιωθεί η απόδοση των μοντέλων.

Καταρχήν, έπρεπε να αποφασιστεί το μέγεθος-ανάλυση της φωτογραφίας που θα χρησιμοποιηθεί γιατί πιθανώς να δημιουργηθεί θέμα χώρου στη RAM του υπολογιστή. Για δίκτυα με 3-4 κρυφά επίπεδα όπως αυτά που αναπτύσσονται σε αυτήν την διπλωματική εργασία, 5000 εικόνες για εκπαίδευση θεωρούνται αρκετές. Αναλόγως με την ανάλυση της εικόνας που επιλέγεται αλλάζει και το βάρος της, άρα και η μνήμη που καταναλώνει. Ενδεικτικά, το βίντεο και οι αρχικές φωτογραφίες ήταν σε διάσταση 1280x720. Αυτό σημαίνει 921600 pixels άρα και νευρώνες στο πρώτο επίπεδο για κάθε φωτογραφία που εισάγεται. Ο αριθμός αυτός είναι υπερβολικά μεγάλος για πρώτο επίπεδο. Συνηθίζεται, όπως φαίνεται και από όλες τις άλλες έρευνες που έχουν αναφερθεί ως τώρα στην παρούσα διπλωματική, να χρησιμοποιούνται φωτογραφίες έως και 256*256, δηλαδή 65536 νευρώνες στο πρώτο

επίπεδο. Ακολουθώντας αυτή τη λογική, και διατηρώντας τις αναλογίες της φωτογραφίας, επιλέχθηκε να γίνει ανάλυση με μέγεθος εικόνας 128x72, δηλαδή 9216 νευρώνες. Θα μπορούσε να γίνει ανάλυση και με φωτογραφίες 320x180, αλλά διπλασιαζόταν το βάρος και ο χρόνος ανάλυσης χωρίς αξιοσημείωτα καλύτερα αποτελέσματα. Από την άλλη, περαιτέρω μείωση σε μέγεθος 32x18, έχανε κάθε σημαντική πληροφορία της εικόνας, καθιστώντας την μη αξιοποιήσιμη.

Με αυτήν την προσέγγιση οι 5600 φωτογραφίες εκπαίδευσης καταλάμβαναν περίπου 1GB μνήμης του υπολογιστή με τις υπόλοιπες 2500 περίπου να καταλαμβάνουν άλλο 0.5 GB. Για υπολογιστή με RAM 8 GB, του οποίου οι λειτουργίες απαιτούν 1.5 GB, η παραπάνω επιλογή είναι αρκετά καλή εφόσον του δίνει περιθώριο 5 GB για την ανάλυση (μεγαλύτερη σαφώς από όσο απαιτείται).

Σε δεύτερη φάση έπρεπε να γίνει μία περικοπή των φωτογραφιών. Συγκεκριμένα, όπως αναφέρθηκε στη Παράγραφο 4.1.2 οποιαδήποτε πληροφορία αφορά τον χώρο πάνω από την οροφή του σταθμευμένου οχήματος αποτελεί θόρυβο για το συνελκτικό δίκτυο. Με άλλα λόγια δεν προσφέρει κάποια χρήσιμη πληροφορία στο δίκτυο. Αυτό επιβεβαιώθηκε και από δοκιμές με λιγότερη ή περισσότερη περικοπή στις εύκολες φωτογραφίες, όπου το δίκτυο έφτανε στην ίδια ακριβώς απόδοση, απλά σε περισσότερο χρόνο, μέχρι να καταφέρει να αποδώσει βάρη μηδενικά στους επιπλέον περιττούς νευρώνες που δεν περικοπήθηκαν.

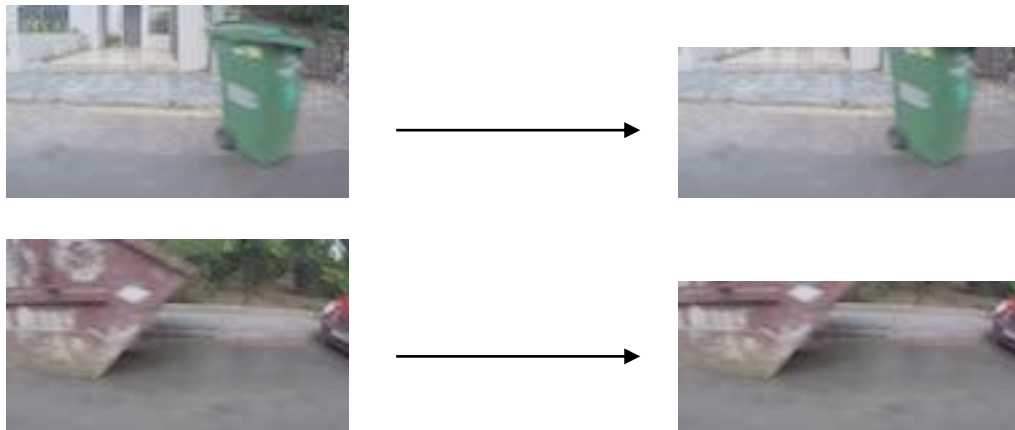
Βεβαίως, επειδή ανάλογα με την απόσταση του κινούμενου οχήματος από το πεζοδρόμιο η απόσταση των 3.5 μέτρων αλλάζει, η περικοπή αυτή πρέπει να είναι λίγο επιεικής. Όσο πιο αυστηρή είναι, μειώνοντας δηλαδή περισσότερο το επάνω μέρος της εικόνας, τόσο μεγαλώνει η πιθανότητα να περικοπούν και χρήσιμα μέρη της εικόνας αν το κινούμενο όχημα απέχει περισσότερο από 3.5 μέτρα από το πεζοδρόμιο. Αντιθέτως, πολύ επιεικής περικοπή αυξάνει τους νευρώνες που αποτελούν θόρυβο για το δίκτυο επιβαρύνοντας τόσο χωρικά όσο και χρονικά το συνελκτικό δίκτυο. Βεβαίως, στην πράξη φάνηκε ότι όταν χρησιμοποιούνται και οι δύσκολες φωτογραφίες, οι οποίες περιέχουν πιο ιδιαίτσες περιπτώσεις, η μικρότερη περικοπή επιφέρει καλύτερα αποτελέσματα διότι έτσι αναγνωρίζονται καλύτερα κάποια εμπόδια, τα οποία δεν είναι σταθμευμένα οχήματα.

Επομένως, κατά την εκπόνηση αυτής της διπλωματικής εργασίας έγινε κι μία μικρή έρευνα σχετικά με την κατάλληλη περικοπή των εικόνων για στάθμευση και λήψη από πλαϊνή κάμερα. Για τις εύκολες φωτογραφίες, οι φωτογραφίες 128x72 μετατρέπονταν σε 128x48 (περικοπή των 28 επάνω γραμμών της κάθε εικόνας, Εικόνα 4.3), ενώ στις φωτογραφίες όπου αναλύονταν μαζί οι εύκολες και οι Rest, επιλέχθηκε τελικά 128x56 (περικοπή των 20 επάνω γραμμών της κάθε εικόνας,

(Εικόνα 4.4) για καλύτερα αποτελέσματα. Στο Κεφάλαιο 5 παρατίθενται τα σχετικά αποτελέσματα μαζί με άλλες επιπλέον αναλύσεις.



Εικόνα 4.3 Εύκολες Εικόνες και οι Περικοπές τους (128x72-128x48)



Εικόνα 4.4 Δύσκολες-Rest Εικόνες και οι Περικοπές τους (128x72-128x56)

5. ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΑΞΙΟΛΟΓΗΣΗ ΔΙΚΤΥΩΝ

Ο τρόπος που πραγματοποιείται η εκπαίδευση στα νευρωνικά δίκτυα και η αξιολόγησή τους εξηγήθηκε αναλυτικά στο Κεφάλαιο 3. Στο Κεφάλαιο αυτό, παρουσιάζονται τα αποτελέσματα της έρευνας της παρούσας διπλωματικής εργασίας, η αξιολόγησή τους και ο σχολιασμός τους. Το Κεφάλαιο είναι δομημένο ως προς τα υποπροβλήματα που έχουν ήδη αναφερθεί από το πιο απλό στο πιο σύνθετο. Σε κάθε υποπρόβλημα αναγράφονται τα δεδομένα που χρησιμοποιήθηκαν, η δομή και οι παράμετροι του βέλτιστου δικτύου και τέλος ακολουθεί ένας σχολιασμός. Παρότι έχει γίνει μεγάλος αριθμός δοκιμών για την επίτευξη αυτών των ακριβειών στη βέλτιστη λύση κάθε υποπροβλήματος, δεν θα παρατεθούν όλες αυτές οι δοκιμές, αλλά η συλλογιστική πορεία που οδήγησε στα παραπάνω δίκτυα και όσες από αυτές τις δοκιμές κρίνεται ότι απαιτείται να σχολιαστούν για ακαδημαϊκούς λόγους. Ακόμη, γίνονται οι απαραίτητες συγκρίσεις μεταξύ των δικτύων αλλά και των συμπεριφορών που παρουσιάζουν τα δίκτυα σε σχέση με αυτήν που αναμένονταν για λόγους πληρότητας. Το Κεφάλαιο κλείνει με την παρουσίαση μίας μεθοδολογίας για την εφαρμογή του μοντέλου που αναπτύχθηκε, υπό πραγματικές συνθήκες, εντοπίζοντας τις πιθανές προκλήσεις και προτείνοντας τρόπους αντιμετώπισής τους.

5.1 ΠΡΟΒΛΗΜΑ ΜΕΙΩΜΕΝΗΣ ΠΟΛΥΠΛΟΚΟΤΗΤΑΣ

5.1.1 ΔΕΔΟΜΕΝΑ

Εύκολες εικόνες ορίζονται όσες εικόνες περιέχουν καθαρά είτε μία κενή θέση στάθμευσης, είτε ένα σταθμευμένο όχημα σε ένα ποσοστό της εικόνας τους. Όσες εικόνες περιέχουν και άλλα εμπόδια ή απρόβλεπτα αντικείμενα κατηγοριοποιούνται ως δύσκολες. Από το σύνολο των δεδομένων, 7413 φωτογραφίες κατηγοριοποιήθηκαν ως εύκολες. Αφού προστέθηκε η επιγραφή 0/1 σε κάθε φωτογραφία, αυτές χωρίστηκαν στα απαραίτητα υποσύνολα για την ανάλυση. Συγκεκριμένα από τις 7413 φωτογραφίες, 5000 χρησιμοποιήθηκαν ως Σύνολο Εκπαίδευσης (67%), 1500 ως Σύνολο Επικύρωσης (20%) και 913 ως Σύνολο Δοκιμής (13%).

5.1.2 ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΔΟΜΗΣ ΚΑΙ ΕΚΠΑΙΔΕΥΣΗΣ

ΔΟΜΗ

Επιλέχθηκε το δίκτυο του Πίνακα 5.1, επειδή ήταν παραπλήσιο του LeNet-5 και αποτελούσε καλή αρχή για την έρευνα.

Πίνακας 5.1 Δομή Συνελκτικού Δικτύου Προβλήματος Μειωμένης Πολυπλοκότητας

Layer	Type	Func	Input y	Input x	Input	F (y)	F (x)	Dim F	No F	S str y	S str x	P Pad y	P Pad x	Output y	Output x	Output
0	Inp		128	40	3	-	-	-	-	-	-	-	-	128	40	3
1	conv	-	128	40	3	5	5	3	20	1	1	0	0	124	36	20
2	act	relu	124	36	20	-	-	-	-	1	1	0	0	124	36	20
3	pool	max	124	36	20	2	2	-	-	2	2	0	0	62	18	20
4	conv	-	62	18	20	5	5	20	50	1	1	0	0	58	14	50
5	act	relu	58	14	50	-	-	-	-	1	1	0	0	58	14	50
6	pool	max	58	14	50	2	2	-	-	2	2	0	0	29	7	50
7	flatten	-	29	7	50	-	-	-	-	-	-	-	-	10150	1	1
8	FC	-	10150	1	1	-	-	-	50	-	-	-	-	50	1	1
9	FC	-	50	1	1	-	-	-	2	-	-	-	-	2	1	1
10	Softmax	-	2	1	1	-	-	-	-	-	-	-	-	2	1	1

Αν και το δίκτυο του Πίνακα 5.1 ήταν το πρώτο που δοκιμάστηκε αποδείχτηκε και το αποδοτικότερο, πιθανώς λόγω της επιτυχίας του και σε άλλα προβλήματα κατηγοριοποίησης. Παρόλ' αυτά έγινε μία προσπάθεια αλλαγής της μορφής του για βελτίωση των αποτελεσμάτων. Για κάθε αλλαγή της μορφής του, ακολουθούσε όλη η διαδικασία αλλαγής παραμέτρων εκπαίδευσης, που περιγράφεται στην επόμενη παράγραφο, για την εύρεση των βέλτιστων παραμέτρων της εκάστοτε μορφής. Από τις δοκιμές που έγιναν κανένα δίκτυο δεν ξεπέρασε την απόδοση και την ταχύτητα του δικτύου που απεικονίζεται στον Πίνακα 5.1.. Προσπάθειες βελτίωσής του αποτέλεσαν η αλλαγή συνάρτησης δραστηριοποίησης, η προσθήκη επιπλέον κρυφών επιπέδων ή η αλλαγή της δυναμικότητας της συνέλιξης επιλέγοντας μικρότερους πυρήνες. Η συνάρτηση δραστηριοποίησης επιβεβαίωσε την θεωρία, σύμφωνα με την οποία η ReLu υπερτερεί των tanh και σ συναρτήσεων. Η προσθήκη επιπλέον επιπέδων δεν βελτίωνε την απόδοση του δικτύου, επομένως απλώς επιβάρυνε την διαδικασία εκμάθησης. Η αλλαγή του πυρήνα συνέλιξης δεν έδωσε σε αυτό το υποπρόβλημα καλύτερα αποτελέσματα από τον ομοιόμορφο πυρήνα 5x5 και στα δύο κρυφά επίπεδα, επομένως διατηρήθηκε η αρχική δομή που επιλέχτηκε (Πίνακας 5.1).

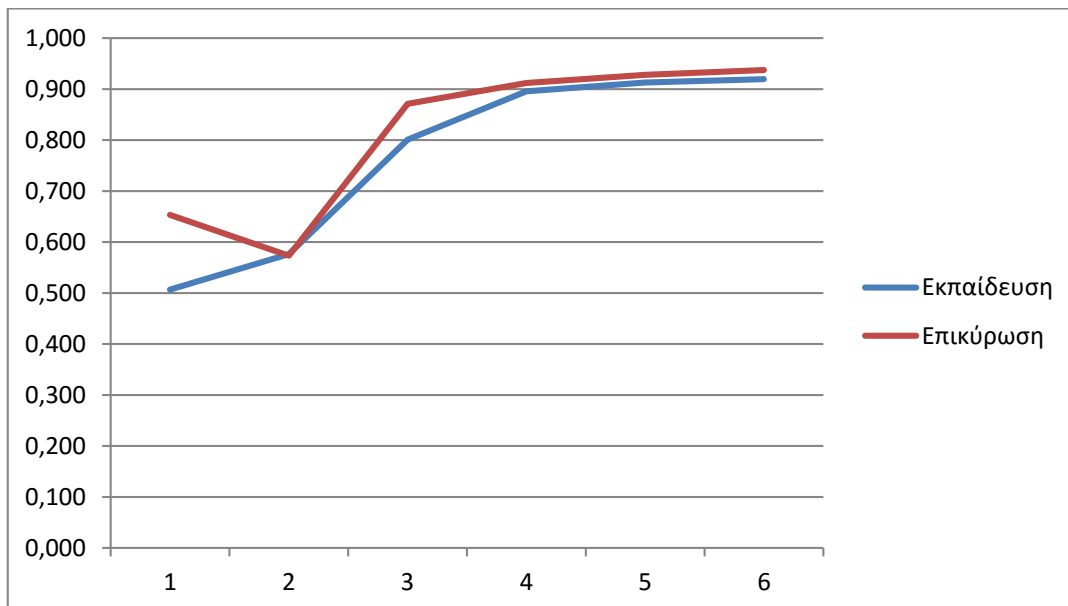
ΕΚΠΑΙΔΕΥΣΗ

Η μορφή εισόδου των δεδομένων στο συνελκτικό δίκτυο, καθώς και οι παράμετροι του δικτύου που επιλέχτηκαν τελικά για την εκπαίδευσή του, φαίνονται στον Πίνακα 5.2. Για την επιλογή των παραμέτρων ακολουθήθηκε σύμφωνα με την θεωρία, μία διαδικασία δοκιμών, η οποία περιγράφεται εκτενώς στα επόμενα εδάφια.

Πίνακας 5.2 Δεδομένα και Παράμετροι Συνελκτικού Δικτύου Προβλήματος Μειωμένης Πολυπλοκότητας

ΔΕΔΟΜΕΝΑ		ΠΑΡΑΜΕΤΡΟΙ	
Ύψος Εικόνας	40	Εποχές	6
Πλάτος Εικόνας	128	Αντιπροσωπευτικό Σύνολο Εκπαίδ.	10
Βάθος Εικόνας	3	Ρυθμός Εκμάθησης	0,01
Αριθμός Εικόνων Εκπαίδευσης	5000	Ορμή	0,9
Αριθμός Εικόνων Επικύρωσης	1500	Συστηματοποίηση	$5 \cdot 10^{-6}$

Πρώτο βήμα αποτέλεσε η εύρεση του κατάλληλου ρυθμού εκμάθησης. Για σταθερά τα υπόλοιπα στοιχεία, έγινε ένας αριθμός δοκιμών όπου ξεκινώντας από “ η ”=0.01 και πολ/ζοντας ή διαιρώντας με το 10 κάθε φορά, βρέθηκε το ανώτατο όριο εκμάθησης. Το όριο αυτό ήταν το 0.01. Στην συνέχεια έγινε προσπάθεια να προσδιοριστεί με μεγαλύτερη ακρίβεια το “ η ”, ξεφεύγοντας από την τάξη μεγέθους, αλλά οι δοκιμές έδειξαν ότι το 0.01 ήταν η καλύτερή του προσέγγιση. Αν και χρησιμοποιούνταν 30 Εποχές σε αυτήν την φάση, η πράξη έδειξε ότι μετά τις 3-5 πρώτες Εποχές το συνελκτικό δίκτυο ξεκινά να «μαθαίνει» (Διάγραμμα 5.1) και ως τις 10 Εποχές έχει φτάσει τις περισσότερες φορές το μέγιστό του σημείο με το επιλεγμένο « η ». Για το λόγο αυτό στις επόμενες δοκιμές οι Εποχές δεν ξεπερνούσαν τις 10, παρά μόνο ορισμένες φορές για επιβεβαίωση αυτής της υπόθεσης.



Διάγραμμα 5.1 Ακρίβεια Συνόλου Εκπαίδευσης και Συνόλου Επικύρωσης ως προς τον αριθμό των Εποχών Εκπαίδευσης του Συνελκτικού Δικτύου

Στην συνέχεια ακολουθήθηκε η ίδια διαδικασία και για την παράμετρο συστηματοποίησης “ λ ”. Ξεκινώντας με την τιμή 10^{-4} και με την ίδια ακριβώς διαδικασία τόσο για την τάξη μεγέθους του, όσο και για την ακριβή του τιμή, βρέθηκε ότι το καταλληλότερο “ λ ” ισούται με $5 \cdot 10^{-6}$. Οι τιμές 0.9 για την ορμή “ μ ” και 10 για το υποσύνολο εκπαίδευσης αποδείχτηκαν οι καλύτερες ύστερα από αντίστοιχες διαδικασίες οι οποίες δεν προσέδωσαν κάποια καλύτερη τιμή που να αυξάνει την αποδοτικότητα του δικτύου ή την ταχύτητά του. Ορμή μεγαλύτερη του

0.9 επιτάχυνε τόσο την εκμάθηση που την κατέστρεφε, ενώ ορμή μικρότερη του 0.9 επιβράδυνε εμφανώς την διαδικασία.

Τα υποσύνολα εκπαίδευσης, αν και ερευνήθηκαν εξίσου, δεν έδωσαν κάποιο χαρακτηριστικό γνώρισμα και επομένως η επιλογή του 10 αποτελεί μία περίπτωση καλής προσαρμογής στο συγκεκριμένο μοντέλο.

Τέλος πραγματοποιήθηκε μία προσπάθεια αλλαγής της τιμής των αρχικών παραμέτρων (π.χ. κανονική κατανομή με τ.α. 0.01) χωρίς όμως να έχει ευεργετικά αποτελέσματα για την απόδοση ή την ταχύτητα του δικτύου. Σε δεύτερο στάδιο, εκπαίδευση με μικρότερο “η” δεν βελτίωνε την απόδοση του δικτύου, πιθανώς επειδή δεν απαιτούνταν μεγαλύτερη λεπτομέρεια για την κατηγοριοποίηση των εύκολων φωτογραφιών.

Υπενθυμίζεται ότι δυνητικά, μπορούν να δοκιμαστούν εκατομμύρια συνδυασμοί μορφών και παραμέτρων του δικτύου βελτιώνοντας συνεχώς έστω και ελάχιστα την απόδοση του δικτύου. Βεβαίως, στα πλαίσια της παρούσας διπλωματικής εργασίας η απόδοση που επιτεύχθηκε με το συνελκτικό δίκτυο του Πίνακα 5.1 επιτρέπει τον τερματισμό της περαιτέρω ανάλυσης, της οποίας το μέγεθος της βελτίωσης της απόδοσης του δικτύου σε σχέση με τον όγκο δουλειάς που απαιτείται για αυτήν να την καθιστούν μη αναγκαία.

5.1.3 ΑΠΟΤΕΛΕΣΜΑΤΑ

Τα αποτελέσματα πρόβλεψης του συνελκτικού δικτύου στο Σύνολο Δοκιμής παρατίθενται στον Πίνακα 5.3 μαζί με τα στατιστικά κατηγοριοποίησης που προκύπτουν. Δημιουργώντας τον πίνακα Προβλέψεων/Πραγματικότητας ο ερευνητής μπορεί να μετρήσει την απόδοση του δικτύου του και να εκτιμήσει ποιοτικά τα λάθη στα οποία υποπίπτει. Στο Υποκεφάλαιο 3.8 αναφέρθηκαν όλα τα στατιστικά μέτρα που χρησιμοποιούνται, πώς αυτά προκύπτουν αλλά και τι δηλώνουν για το υπό εξέταση συνελκτικό δίκτυο.

Πίνακας 5.3 Πίνακας Προβλέψεων/Πραγματικότητας και Στατιστικά Μέτρα Κατηγοριοποίησης του Συνελκτικού Δικτύου για το Πρόβλημα Μειωμένης Πολυπλοκότητας

		Πραγματικά Δεδομένα	Πραγματικά Δεδομένα
	π/π	Κενή	Μη Κενή
Πρόβλεψη	Κενή	184	59
Πρόβλεψη	Μη Κενή	20	650
AUC	0,909	Accuracy	0,913
Λάθος Τύπου I	0,083	Λάθος Τύπου II	0,098
Specificity	0,917	Sensitivity	0,902
False Empty Rate	0,243	False Not Empty Rate	0,030
Positive Pred. Value	0,757	Negative Pred. Value	0,970

Αξιολόγηση

Το βασικότερο μέτρο αξιολόγησης είναι η ακρίβεια του δικτύου στο Σύνολο Δοκιμής, η οποία ισούται με το άθροισμα όλων των σωστά κατηγοριοποιημένων φωτογραφιών (0-0,1-1) προς το σύνολο των φωτογραφιών. Στο συγκεκριμένο δίκτυο το ποσοστό επιτυχούς κατηγοριοποίησης άγγιξε το 91,34%. Το νούμερο αυτό επιτεύχθηκε μία φορά με την δεδομένη μορφή και παραμετροποίηση. Για το ίδιο δίκτυο και τις ίδιες παραμέτρους θα επιτευχθεί ακρίβεια κοντά στην προαναφερθείσα, χωρίς ωστόσο να είναι ποτέ η ίδια. Αυτό οφείλεται στην τυχαιότητα των αρχικών τιμών των βαρών, η οποία εξαλείφεται σχεδόν εξολοκλήρου κατά τη διάρκεια της εκπαίδευσης γι' αυτό και η τελική απόδοση τείνει στην ίδια τιμή αλλά δεν είναι ποτέ ακριβώς ίδια.

Η καμπύλη ROC, η οποία σχηματίζεται από τις Sensitivity και Specificity βελτιώνεται όσο οι τιμές αυτές προσεγγίζουν το 1. Εφόσον και οι δύο τιμές στο συγκεκριμένο δίκτυο είναι πάνω από 0.9 είναι εμφανής η επιτυχία του δικτύου. Ακόμη την επιτυχία της καμπύλης άρα και την επιτυχία του συνελκτικού δικτύου, την υποδηλώνει και ο όρος AUC, ο οποίος υπολογίζει το εμβαδό υπό την καμπύλη. Τιμές του AUC μεγαλύτερες του 0.9, όπως η συγκεκριμένη, υποδηλώνουν δίκτυο πολύ καλά εκπαιδευμένο.

Τέλος, υπολογίζονται τα επιμέρους ποσοστά λάθους στα οποία υποπίπτει το συνελκτικό δίκτυο. Γενικός κανόνας που ακολουθείται είναι η μείωση της διαγωνίου 2-3 και η αύξηση της διαγωνίου 1-4 για την βελτίωση όλων των ποσοστών. Επιδιώκεται να μειώνεται το λάθος τύπου I (ψευδώς θετικό), που αντιστοιχεί σε

κατηγοριοποίηση ενός μηδενικού ως άσσου. Έτσι όπως δομήθηκε το πείραμα, «1» θεωρήθηκε η κατειλημμένη θέση, επομένως στην πράξη ισχύει το αντίθετο εφόσον το μοντέλο που αναπτύχθηκε ανιχνεύει κενές θέσεις. Όλα τα στατιστικά μέτρα που επεξηγήθηκαν στο Κεφάλαιο 3, προσαρμόζονται στην παραπάνω παραδοχή και υπολογίζονται αναλόγως ως προς το 0-0 του αντίστοιχου πίνακα και όχι ως προς το 1-1. Εφόσον το δίκτυο που χτίζεται απαιτείται να εντοπίζει τις κενές θέσεις, το λάθος που πρέπει να μειωθεί δραστικά, είναι η κατηγοριοποίηση μίας κατειλημμένης θέσης ως κενή (Τύπου I). Το αντίστροφο λάθος (Τύπου II), δηλαδή η κατηγοριοποίηση μίας κενής θέσης ως κατειλημμένη να μεν μειώνει την ακρίβεια του μοντέλου, αλλά αποτελεί ευχάριστη έκπληξη στον οδηγό που αναζητά θέση στάθμευσης. Ενδιαφέρον βέβαια, θα είχε να παρατεθεί αντ' αυτού, το ποσοστό των προβλεπόμενων θέσεων που κατηγοριοποιούνται εσφαλμένα ως κενές θέσεις. Στο συγκεκριμένο δίκτυο, το ποσοστό των κατειλημμένων θέσεων που κατηγοριοποιήθηκαν εσφαλμένα ως κενές είναι το 8,3% (False Negative Rate - Λάθος Τύπου I), ενώ το ποσοστό των προβλεπόμενων ως κενών θέσεων που κατηγοριοποιούνται εσφαλμένα ως κενές θέσεις είναι το 24,27% (False Empty Rate).

Άλλες Παρατηρήσεις

Για τα επίπεδα συγκέντρωσης χρησιμοποιήθηκε πάντοτε η γνωστή μορφή από το Υποκεφάλαιο 3.6 όπου $F=2$ και $S=2$ με χρήση της maxpooling. Τα επίπεδα συνέλιξης είχαν πάντοτε $S=1$ και δεν δοκιμάστηκε κάποια άλλη περίπτωση (δεν ενδείκνυται και από την θεωρία). Τέλος, έγινε μία προσπάθεια βελτιστοποίησης του αριθμού των πλήρως συνδεδεμένων επιπέδων και του αριθμού των νευρώνων τους, χωρίς όμως να παρατηρηθεί κάτι αξιοσημείωτο.

5.2 ΠΡΟΒΛΗΜΑ ΑΥΞΗΜΕΝΗΣ ΠΟΛΥΠΛΟΚΟΤΗΤΑΣ

5.2.1 ΔΕΔΟΜΕΝΑ

Στην συνέχεια έπρεπε να προστεθούν και οι δύσκολες φωτογραφίες προς ανάλυση γιατί μόνο τότε έχει εφαρμοσιμότητα το μοντέλο, εφόσον τέτοιες συνθήκες θα αντιμετωπίσει στην πράξη. Δύσκολες εικόνες ορίζονται όσες εικόνες περιέχουν απρόβλεπτα εμπόδια ή αντικείμενα, τα οποία ξεφεύγουν από την συνηθισμένη εικόνα μίας κενής θέσης ή μίας θέσης με ένα σταθμευμένο όχημα. Όπως παρατηρείται, η αναλογία εύκολων-δύσκολων φωτογραφιών στην πράξη είναι 5 προς 1. Από το σύνολο των δεδομένων, 1539 φωτογραφίες κατηγοριοποιήθηκαν ως δύσκολες. Από τις $7413+1539=8952$ φωτογραφίες, $5000+1000=6000$ χρησιμοποιήθηκαν ως Σύνολο Εκπαίδευσης (67%), $1500+300=1800$ ως Σύνολο Επικύρωσης (20%) και $913+183=1096$ ως Σύνολο Δοκιμής (13%).

5.2.2 ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΔΟΜΗΣ ΚΑΙ ΕΚΠΑΙΔΕΥΣΗΣ

Επιλέχθηκε το δίκτυο του Πίνακα 5.4, επειδή ήταν παραπλήσιο του LeNet-5 και αποδείχτηκε από το προηγούμενο πρόβλημα ότι δίνει καλά αποτελέσματα.

Πίνακας 5.4 Δομή Συνελικτικού Δικτύου Προβλήματος Αυξημένης Πολυπλοκότητας

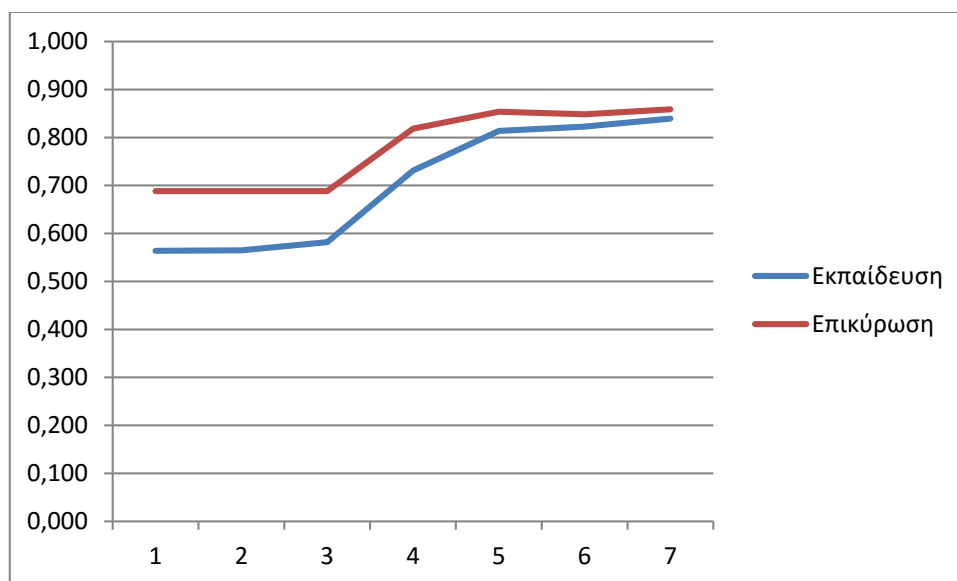
Layer	Type	Func	Input y	Input x	Input	F (y)	F (x)	Dim F	No F	S str y	S str x	P Pad y	P Pad x	Output y	Output x	Output
0	Inp		128	40	3	-	-	-	-	-	-	-	-	128	40	3
1	conv	-	128	40	3	5	5	3	20	1	1	0	0	124	36	20
2	act	relu	124	36	20	-	-	-	-	1	1	0	0	124	36	20
3	pool	max	124	36	20	2	2	-	-	2	2	0	0	62	18	20
4	conv	-	62	18	20	5	5	20	50	1	1	0	0	58	14	50
5	act	relu	58	14	50	-	-	-	-	1	1	0	0	58	14	50
6	pool	max	58	14	50	2	2	-	-	2	2	0	0	29	7	50
7	flatten	-	29	7	50	-	-	-	-	-	-	-	-	10150	1	1
8	FC	-	10150	1	1	-	-	-	50	-	-	-	-	50	1	1
9	FC	-	50	1	1	-	-	-	2	-	-	-	-	2	1	1
10	Softmax	-	2	1	1	-	-	-	-	-	-	-	-	2	1	1

Η εκπαίδευση του δικτύου αυτού ακολούθησε ακριβώς την ίδια λογική που περιγράφηκε στο Κεφάλαιο 5.1.2 και χρησιμοποίησε τις παραμέτρους που φαίνονται στον Πίνακα 5.5.

Πίνακας 5.5 Δεδομένα και Παράμετροι Συνελκτικού Δικτύου Προβλήματος Αυξημένης Πολυπλοκότητας

ΔΕΔΟΜΕΝΑ		ΠΑΡΑΜΕΤΡΟΙ	
Ύψος Εικόνας	40	Εποχές	7
Πλάτος Εικόνας	128	Αντιπροσωπευτικό Σύνολο Εκπαίδ.	10
Βάθος Εικόνας	3	Ρυθμός Εκμάθησης	0,01
Αριθμός Εικόνων Εκπαίδευσης	6000	Ορμή	0,9
Αριθμός Εικόνων Επικύρωσης	1800	Συστηματοποίηση	$1 \cdot 10^{-6}$

Παρολ' αυτά, πρέπει να σημειωθεί ότι κανένα συνελκτικό δίκτυο δεν ξεπερνούσε την ακρίβεια του 85% (Διάγραμμα 5.2).



Διάγραμμα 5.2 Ακρίβεια Συνόλου Εκπαίδευσης και Συνόλου Επικύρωσης ως προς τον αριθμό των Εποχών Εκπαίδευσης του Συνελκτικού Δικτύου

5.2.3 ΑΠΟΤΕΛΕΣΜΑΤΑ

Η μειωμένη απόδοση του συνελκτικού δικτύου ήταν προβληματική, επειδή συγκρίνοντας το αποτέλεσμα του προβλήματος μειωμένης πολυπλοκότητας με αυτό της αυξημένης, προέκυπτε ότι όχι μόνο το δίκτυο ουσιαστικά είχε απόδοση $\frac{1}{2}$ στις δύσκολες φωτογραφίες, αλλά ουσιαστικά επηρέαζε (ορισμένες φορές, όχι πάντα) αρνητικά και την απόδοση των εύκολων εικόνων. Ενδεικτικά, οι 913 εύκολες φωτογραφίες είχαν απόδοση 91,34% (834/913 φωτογραφίες σωστά κατηγοριοποιημένες), ενώ οι εύκολες (913) και οι δύσκολες (183) μαζί είχαν απόδοση της τάξης του 85% (Πίνακας 5.6), που ισούται περίπου με τον σταθμισμένο μέσο $91,34\% \cdot 913 + 183 \cdot 50\%$.

Πίνακας 5.6 Πίνακας Προβλέψεων/Πραγματικότητας και Στατιστικά Μέτρα Κατηγοριοποίησης του Συνελκτικού Δικτύου για το Πρόβλημα Αυξημένης Πολυπλοκότητας

		Πραγματικά Δεδομένα	Πραγματικά Δεδομένα
	π/π	Κενή	Μη Κενή
Πρόβλεψη	Κενή	184	129
Πρόβλεψη	Μη Κενή	36	747
AUC	0,845	Accuracy	0,849
Λάθος Τύπου I	0,147	Λάθος Τύπου II	0,164
Specificity	0,853	Sensitivity	0,836
False Empty Rate	0,412	False Not Empty Rate	0,046
Positive Pred. Value	0,588	Negative Pred. Value	0,954

Αυτό επιβεβαιώνεται και από τον Πίνακα 5.7, όπου 93 από τις συνολικά 183 δύσκολες φωτογραφίες του Συνόλου Δοκιμής χαρακτηρίστηκαν με αποτέλεσμα εσφαλμένο (πολύ κοντά στο 50% που υποτέθηκε). Η απόδοση των εύκολων φωτογραφιών στο Σύνολο Δοκιμής, όταν στην εκπαίδευση συμπεριλαμβάνονταν και δύσκολες φωτογραφίες, βελτιώθηκε από 99 λάθη στα 72 αλλά ποιοτικά τα νέα λάθη δεν είχαν μεγάλη επικάλυψη με τα παλαιότερα. Με άλλα λόγια, η προσθήκη των δύσκολων φωτογραφιών δεν βελτίωσε το μοντέλο που θεωρητικά είχε πριν εκπαιδευτεί. Στην πραγματικότητα, μετέβαλε τα βάρη με τέτοιο τρόπο ώστε γενικά να ανιχνεύουν καλύτερα τα εμπόδια εντός των εικόνων, άρα και τα απλά οχήματα.

Για το λόγο αυτό υπήρξε βελτίωση της απόδοσής τους. Παρολ' αυτά, το μοντέλο που συζητήθηκε δεν είναι επαρκές, εφόσον όπως εξηγήθηκε, ουσιαστικά βελτιώνει ελαφρώς την κατηγοριοποίηση των εύκολων φωτογραφιών και τοποθετεί σχεδόν τυχαία τις δύσκολες, πράγμα μη αποδεκτό.

Πίνακας 5.7 Αύξων Αριθμός των Λαθών του Συνελκτικού Δικτύου Εύκολων και Δύσκολων Φωτογραφιών (με Κίτρινο Υπογραμμισμένες οι Δύσκολες Φωτογραφίες)

6	96	294	366	569	714	858	1000
12	100	296	372	602	720	864	1002
17	102	297	375	606	726	870	1005
18	108	300	376	615	768	876	1008
19	112	301	377	622	774	882	1014
24	115	306	390	634	777	888	1020
30	119	312	396	636	778	891	1021
36	126	318	402	637	779	894	1022
42	132	324	414	642	780	899	1023
46	138	330	419	648	781	900	1024
48	144	336	434	654	786	906	1025
54	175	354	435	660	791	912	1026
60	176	355	444	666	792	918	1032
61	177	356	450	672	793	924	1038
62	207	357	456	676	815	930	1044
66	214	358	462	678	817	933	1045
72	239	360	468	684	818	936	1053
75	246	361	474	690	840	947	1055
76	252	362	480	696	841	949	1064
80	273	363	531	702	846	950	
	288	364	532	708	852	985	

Για να εντοπιστεί η πηγή του λάθους, πραγματοποιήθηκε μία έρευνα σχετικά με τις φωτογραφίες που κατηγοριοποιούνταν λάθος. Τα δύο κυριότερα είδη λαθών ήταν τα εξής:

- Οι διασταυρώσεις - οδικό κόμβοι, μετά και την περικοπή, κατηγοριοποιούνταν από το δίκτυο, ορθώς, ως κενές θέσεις 0. Παρόλ' αυτά είχαν από τον χρήστη περιγραφτεί σαν 1, εφόσον δεν αποτελούν θέσεις στάθμευσης. Ενώ θα μπορούσε να γίνει ξεχωριστό συνελκτικό δίκτυο, το οποίο με την διατήρηση του επάνω τμήματος των φωτογραφιών, θα εντόπιζε ποιες από αυτές αποτελούν κόμβο και στην συνέχεια μόνο όσες δεν αποτελούν κόμβο να αναλύονται για την ανίχνευση θέσης στάθμευσης, προτιμήθηκε στην λογική της εφαρμογής του μοντέλου, τα σημεία των κόμβων να είναι δοσμένα εξαρχής, ώστε το συνελκτικό δίκτυο να μην εξετάζει φωτογραφίες με συντεταγμένες x,y, εκείνες των κόμβων (είναι εκ των προτέρων γνωστές σε ένα οδικό δίκτυο).
- Η αυστηρή περικοπή μέρους της φωτογραφίας, πολλές φορές, παρέλειπε σημαντική πληροφορία σχετικά με το εικονιζόμενο εμπόδιο, κυρίως στις δύσκολες φωτογραφίες, όπου είναι απρόβλεπτο τί εμπόδιο θα συναντηθεί. Απαιτούνταν λοιπόν, μία πιο προσεκτική επιλογή στην περικοπή, για να εξασφαλίζεται αυτή η χρήσιμη πληροφορία, χωρίς όμως το δίκτυο να «βαραίνει» υπερβολικά από την επέκταση της εισερχόμενης πληροφορίας.

5.3 ΣΥΝΟΛΙΚΟ ΠΡΑΓΜΑΤΙΚΟ ΠΡΟΒΛΗΜΑ

5.3.1 ΔΕΔΟΜΕΝΑ

Όπως ήδη αναφέρθηκε ορισμένες από τις φωτογραφίες που κατηγοριοποιήθηκαν ως δύσκολες εμπεριείχαν και κόμβους. Επειδή όμως κρίθηκε ότι οι θέσεις των κόμβων είναι δεδομένες στο οδικό δίκτυο, αποφασίστηκε το σύστημα να μην κληθεί να αποφασίσει για αυτά τα σημεία. Επομένως οι δύσκολες φωτογραφίες με κόμβους έπρεπε να εξαχθούν από τα σύνολα εκπαίδευσης, επικύρωσης και δοκιμής. Τελικά, όπως ήδη αναφέρθηκε στο Κεφάλαιο 4, οι 1539 δύσκολες φωτογραφίες περιείχαν 616 φωτογραφίες διασταυρώσεων και 923 φωτογραφίες δύσκολων περιπτώσεων, οι οποίες ονομάστηκαν Rest. Συνδυάζοντας λοιπόν, τις Rest φωτογραφίες με τις εύκολες, με αναλογία εμφάνισης πλέον περίπου 8 προς 1 δημιουργήθηκαν τα σύνολα: 5000+600=5600 φωτογραφίες το Σύνολο Εκπαίδευσης (67%), 1500+180=1680 φωτογραφίες το Σύνολο Επικύρωσης (20%) και 913+143=1056 φωτογραφίες το Σύνολο Δοκιμής (13%).

5.3.2 ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΔΟΜΗΣ ΚΑΙ ΕΚΠΑΙΔΕΥΣΗΣ

Επιλέχτηκε το δίκτυο του Πίνακα 5.8 για επίλυση με δύο κρυφά επίπεδα. Η επιλογή της μορφής του, ήταν αποτέλεσμα των αλλαγών που επήλθαν στο δίκτυο του Πίνακα 5.1. Συγκεκριμένα, η αρχική προσέγγιση ήταν η χρήση του ίδιου δικτύου για την επίλυση και του συνολικότερου προβλήματος. Παρόλ' αυτά, το δίκτυο αυτό, δεν ανταποκρίθηκε στις προσδοκίες που είχε δημιουργήσει η εφαρμογή του στα απλούστερα προβλήματα, με αποτέλεσμα να ξεκινήσουν δοκιμές παραπλήσιων μορφών. Κάποιες από τις επιλογές πυρήνων συνέλιξης για 2 κρυφά επίπεδα ήταν οι:

- (5,5)-(5,5) που εξαρχής απέτυχε
- (3,3)-(2,2) που απέτυχε

- (7,3)-(6,2) που έδωσε μέτρια αποτελέσματα
- (7,3)-(2,2) που απέτυχε
- (7,3)-(4,2) που έδωσε κακά αποτελέσματα
- (7,3)-(6,4) που αποδείχτηκε η βέλτιστη (Πίνακας 5.8)
- (9,5)-(5,5) που απέτυχε
- (9,5)-(3,3) που απέτυχε

Πίνακας 5.8 Δομή Συνελκτικού Δικτύου Συνολικού Πραγματικού Προβλήματος με 2 Κρυφά Επίπεδα

Layer	Type	Func	Input y	Input x	Input	F (y)	F (x)	Dim F	No F	S str y	S str x	P Pad y	P Pad x	Output y	Output x	Output
0	Inp		128	56	3	-	-	-	-	-	-	-	-	128	56	3
1	conv	-	128	56	3	7	3	3	20	1	1	0	0	122	54	20
2	act	relu	122	54	20	-	-	-	-	1	1	0	0	122	54	20
3	pool	max	122	54	20	2	2	-	-	2	2	0	0	61	27	20
4	conv	-	61	27	20	6	4	20	50	1	1	0	0	56	24	50
5	act	relu	56	24	50	-	-	-	-	1	1	0	0	56	24	50
6	pool	max	56	24	50	2	2	-	-	2	2	0	0	28	12	50
7	flatten	-	28	12	50	-	-	-	-	-	-	-	-	16800	1	1
8	FC	-	16800	1	1	-	-	-	50	-	-	-	-	50	1	1
9	FC	-	50	1	1	-	-	-	2	-	-	-	-	2	1	1
10	Softmax	-	2	1	1	-	-	-	-	-	-	-	-	2	1	1

Η συνάρτηση δραστηριοποίησης που χρησιμοποιήθηκε και πάλι ήταν η ReLu, η οποία και πάλι επιβεβαίωσε την σχετική θεωρία.

Στην συνέχεια, και ενώ υπήρχε ήδη πολύ καλή απόδοση στο δίκτυο, ερευνήθηκε η πιθανότητα ύπαρξης δικτύου με περισσότερα κρυμμένα επίπεδα με σκοπό την αισθητή βελτίωση του δικτύου. Αναλόγως με τους συνδυασμούς πυρήνων συνέλιξης που αναφέρθηκαν παραπάνω, το μοντέλο προσέθετε κάθε φορά ένα επίπεδο συνέλιξης ακόμα, με έναν πυρήνα, ο οποίος να αφήνει ακέραιο αριθμό νευρώνων στο δίκτυο όπως εξηγήθηκε στο Υποκεφάλαιο 3.6. Τελικά και πάλι βέλτιστο δίκτυο ήταν ο συνδυασμός πυρήνων (7,3)-(6,4) με την προσθήκη ενός ακόμη πυρήνα (7,3), όπως φαίνεται αναλυτικά και στον Πίνακα 5.9 με συνάρτηση δραστηριοποίησης και πάλι την ReLu.

Πίνακας 5.9 Δομή Συνελκτικού Δικτύου Συνολικού Πραγματικού Προβλήματος με 3 Κρυφά Επίπεδα

Layer	Type	Func	Input y	Input x	Input	F (y)	F (x)	Dim F	No F	S str y	S str x	P Pad y	P Pad x	Output y	Output x	Output
0	Inp		128	56	3	-	-	-	-	-	-	-	-	128	56	3
1	conv	-	128	56	3	7	3	3	20	1	1	0	0	122	54	20
2	act	relu	122	54	20	-	-	-	-	1	1	0	0	122	54	20
3	pool	max	122	54	20	2	2	-	-	2	2	0	0	61	27	20
4	conv	-	61	27	20	6	4	20	50	1	1	0	0	56	24	50
5	act	relu	56	24	50	-	-	-	-	1	1	0	0	56	24	50
6	pool	max	56	24	50	2	2	-	-	2	2	0	0	28	12	50
7	conv	-	28	12	50	7	3	50	100	1	1	0	0	22	10	100
8	act	relu	22	10	100	-	-	-	-	1	1	0	0	22	10	100
9	pool	max	22	10	100	2	2	-	-	2	2	0	0	11	5	100
10	flatten	-	11	5	100	-	-	-	-	-	-	-	-	5500	1	1
11	FC	-	5500	1	1	-	-	-	100	-	-	-	-	100	1	1
12	FC	-	100	1	1	-	-	-	2	-	-	-	-	2	1	1
13	Softmax	-	2	1	1	-	-	-	-	-	-	-	-	2	1	1

Η εκπαίδευση του δικτύου των δύο κρυφών επιπέδων (Πίνακας 5.8) ακολούθησε ακριβώς την ίδια λογική που περιγράφηκε στο Κεφάλαιο 5.1.3. Πρώτο βήμα αποτέλεσε η εύρεση του ρυθμού εκμάθησης. Για δίκτυα παρόμοιου μεγέθους και χαρακτηριστικών παρέμεινε η ίδια εικόνα, με την τιμή “η”=0.01 να αποτελεί για άλλη μία φορά το ανώτατο όριο εκμάθησης. Στην συνέχεια ακολουθήθηκε η ίδια διαδικασία και για την παράμετρο συστηματοποίησης “λ” και για την ορμή “μ”, για τις οποίες επιλέχθηκαν οι τιμές $5 \cdot 10^{-6}$ και 0.9 αντίστοιχα. Όλες οι παράμετροι, συμπεριλαμβανομένων και των υποσυνόλων εκπαίδευσης, ελέγχθηκαν όπως και στο 5.1.2 και συνέτειναν στις τιμές που παρουσιάζονται στον Πίνακα 5.10. Παρατηρείται ότι οι παράμετροι του δικτύου είναι σχεδόν ίδιες, με αυτές του Κεφαλαίου 5.1. Παρολ’ αυτά, δεν έχουν καμία σχέση εφόσον η δομή του δικτύου έχει αλλάξει και η ταύτισή τους (με εξαίρεση των αριθμό των Εποχών) είναι τυχαία και πιθανώς οφείλεται στην παρόμοια συμπεριφορά που αναπτύσσουν παρομοίου μεγέθους και χαρακτηριστικών συνελκτικά δίκτυα.

Σε αντίθεση βέβαια με την εκπαίδευση μόνο των εύκολων φωτογραφιών, το συνελκτικό δίκτυο του 5.3.2 επιδεχόταν περαιτέρω βελτίωση της απόδοσής του με χρήση μικρότερου ρυθμού εκμάθησης. Συγκεκριμένα αφού το δίκτυο εκπαιδεύτηκε για 4 Εποχές με τις παραπάνω παραμέτρους, στην συνέχεια πραγματοποιήθηκε μία αλλαγή των παραμέτρων και κυρίως του ρυθμού εκμάθησης. Σύμφωνα με το Υποκεφάλαιο 3.5, ο ρυθμός εκμάθησης σε δεύτερο στάδιο εκπαίδευσης πρέπει περίπου να υποδεκαπλασιαστεί. Με αυτήν την λογική και ακολουθώντας την ίδια ακριβώς διαδικασία για κάθε παράμετρο βρέθηκε ότι για υποδεκαπλάσιο ρυθμό εκμάθησης, μηδενική συστηματοποίηση και ορμή, και ανά μία μόνο φωτογραφία για

1 Εποχή Εκπαίδευσης (Πίνακας 5.10), το σύστημα βελτιώνει την ακρίβειά του από 85,60% στο 90,43%.

Πίνακας 5.10 Δεδομένα και Παράμετροι Συνελκτικού Δικτύου Συνολικού Πραγματικού Προβλήματος με 2 Κρυφά Επίπεδα για 2 Στάδια Εκπαίδευσης

ΔΕΔΟΜΕΝΑ		ΠΑΡΑΜΕΤΡΟΙ	1ο Στάδιο	2ο Στάδιο
Ύψος Εικόνας	56	Εποχές	4	1
Πλάτος Εικόνας	128	Αντιπροσωπευτικό Σύνολο Εκπαιδ.	10	1
Βάθος Εικόνας	3	Ρυθμός Εκμάθησης	0,01	0,001
Αριθμός Εικόνων Εκπαίδευσης	5600	Ορμή	0,9	0
Αριθμός Εικόνων Επικύρωσης	1680	Συστηματοποίηση	$5 \cdot 10^{-6}$	0

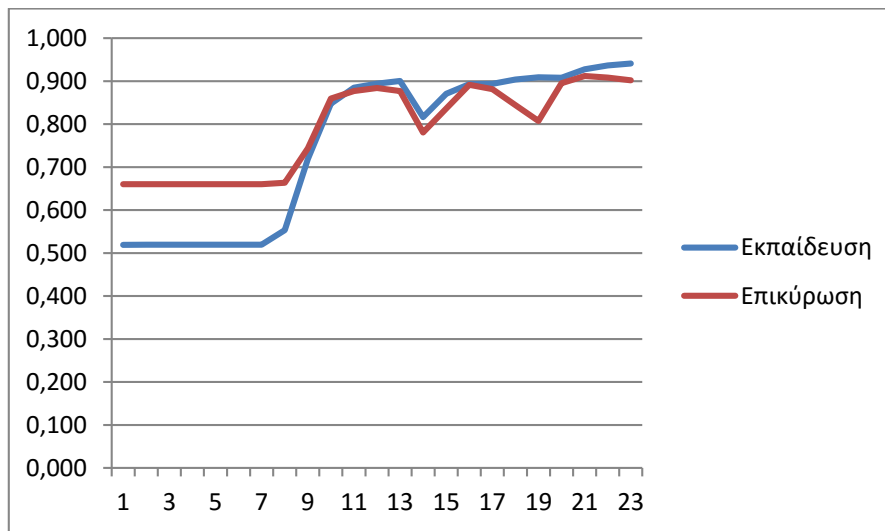
Ενδεικτικά παρατίθενται οι Πίνακες Προβλέψεων/Πραγματικότητας και των δύο σταδίων εκπαίδευσης για να είναι εμφανής η βελτίωση του δικτύου (Πίνακας 5.11). Τα στατιστικά μέτρα όμως αφορούν το τελικό στάδιο. Η αισθητή αυτή αλλαγή της απόδοσης, πιθανώς να οφείλεται στην «ανάγκη» του συστήματος να εντυπώσει σε πιο συγκεκριμένα χαρακτηριστικά των εικόνων για να κατηγοριοποιήσει σωστά τις δυσκολότερες φωτογραφίες, οι οποίες εκ των πραγμάτων περιέχουν πιο σύνθετες απεικονίσεις. Θα μπορούσε η διαδικασία αυτή να επαναληφθεί και για “η”=0.0001, αλλά παύει από ένα σημείο και μετά να έχει αποτέλεσμα αυτή η σταδιακή μείωση (τουλάχιστον στο επίπεδο της ανάλυσης που πραγματοποιείται στην παρούσα διπλωματική εργασία).

Υπενθυμίζεται ότι δυναμικά, μπορούν να δοκιμαστούν εκατομμύρια συνδυασμοί μορφών και παραμέτρων του δικτύου βελτιώνοντας συνεχώς έστω και ελάχιστα την απόδοση του δικτύου. Βεβαίως, στα πλαίσια της παρούσας διπλωματικής εργασίας η απόδοση που επιτεύχθηκε με το συνελκτικό δίκτυο του Πίνακα 5.8 επιτρέπει τον ικανοποιητικό τερματισμό της περαιτέρω ανάλυσης, της οποίας το κέρδος σε απόδοση του δικτύου σε σχέση με τον όγκο δουλειάς που απαιτείται για αυτήν εκτιμάται από την αποκτηθείσα εμπειρία ότι θα την καθιστούν μη αναγκαία.

Πίνακας 5.11 Πίνακας Προβλέψεων/Πραγματικότητας και Στατιστικά Μέτρα Κατηγοριοποίησης του Συνελικτικού Δικτύου για το Συνολικό Πραγματικό Πρόβλημα (2 κρυφά επίπεδα)

	1ο Στάδιο	Πραγματικά Δεδομένα	Πραγματικά Δεδομένα
	π/π	Κενή	Μη Κενή
Πρόβλεψη	Κενή	187	110
Πρόβλεψη	Μη Κενή	42	717
	2ο Στάδιο	Πραγματικά Δεδομένα	Πραγματικά Δεδομένα
	π/π	Κενή	Μη Κενή
Πρόβλεψη	Κενή	202	74
Πρόβλεψη	Μη Κενή	27	753
AUC	0,896	Accuracy	0,904
Λάθος Τύπου I	0,089	Λάθος Τύπου II	0,118
Specificity	0,911	Sensitivity	0,882
False Empty Rate	0,268	False Not Empty Rate	0,035
Positive Pred. Value	0,732	Negative Pred. Value	0,965

Αυτό αποδεικνύεται και από το παράδειγμα του Πίνακα 5.9, όπου προσθέτοντας ένα ακόμη κρυφό επίπεδο με τα χαρακτηριστικά που ήδη περιγράφηκαν και αφού επαναληφθούν όλες οι διαδικασίες που πραγματοποιήθηκαν και για το συνελικτικό δίκτυο του Πίνακα 5.8 για την εύρεση των παραμέτρων του, η απόδοση που επιτυγχάνεται είναι 90.53%, μόλις 0.1% αύξηση σε σχέση με το προηγούμενο δίκτυο (Πίνακας 5.13). Αν και στην πράξη, ενδιαφέρει μόνο η απόδοση του συνελικτικού δικτύου, διότι τα δίκτυα μία φορά εκπαιδεύονται και στην συνέχεια απλώς εφαρμόζονται, δυσανάλογα καταναλισκόμενοι πόροι και χρόνος για πολύ μικρή βελτίωση του δικτύου, δεν παύουν να αποτελούν άσκοπη σπατάλη. Έτσι λοιπόν, το δίκτυο του Πίνακα 5.9 εκπαιδεύεται για 20 Εποχές (αντί για 4) στο πρώτο επίπεδο εκπαίδευσης και για 3 Εποχές στο δεύτερο επίπεδο εκπαίδευσης (αντί για 1) (Διάγραμμα 5.3).



Διάγραμμα 5.3 Ακρίβεια Συνόλου Εκπαίδευσης και Συνόλου Επικύρωσης ως προς τον αριθμό των Εποχών Εκπαίδευσης του Συνελικτικού Δικτύου

Οι υπόλοιπες παράμετροι του πρώτου επιπέδου εκπαίδευσης είναι ίδιες με του δικτύου του Πίνακα 5.8 με εξαίρεση το “λ” που ισούται με $5 \cdot 10^{-4}$, ενώ οι υπόλοιπες παράμετροι του δεύτερου επιπέδου εκπαίδευσης είναι όμοιες με του δικτύου του Πίνακα 5.8 και παρουσιάζονται όλες συνοπτικά στον Πίνακα 5.12.

Πίνακας 5.12 Δεδομένα και Παράμετροι Συνελικτικού Δικτύου Συνολικού Πραγματικού Προβλήματος με 2 Κρυφά Επίπεδα για 2 Στάδια Εκπαίδευσης

ΔΕΔΟΜΕΝΑ		ΠΑΡΑΜΕΤΡΟΙ	1ο Στάδιο	2ο Στάδιο
Ύψος Εικόνας	56	Εποχές	20	3
Πλάτος Εικόνας	128	Αντιπροσωπευτικό Σύνολο Εκπαιδ.	10	1
Βάθος Εικόνας	3	Ρυθμός Εκμάθησης	0,01	0,001
Αριθμός Εικόνων Εκπαίδευσης	5600	Ορμή	0,9	0
Αριθμός Εικόνων Επικύρωσης	1680	Συστηματοποίηση	$5 \cdot 10^{-4}$	0

5.3.3 ΑΠΟΤΕΛΕΣΜΑΤΑ

Από τους Πίνακες Προβλέψεων/Πραγματικότητας προκύπτουν όλα τα χρήσιμα συμπεράσματα για τα δίκτυα που αναλύθηκαν. Συγκεκριμένα πέρα από τις ακρίβειες και των δύο συνελκτικών δικτύων που ήδη αναφέρθηκαν (90,434%-90,53%), το λάθος τύπου I, που είναι αυτό που ενδιαφέρει όπως επεξηγήθηκε και στο Υποκεφάλαιο 5.1.3, βελτιώνεται στο δίκτυο με τα 3 κρυφά επίπεδα από 8.9% (δίκτυο με 2 κρυφά επίπεδα) στο 7.9% (Πίνακας 5.13) εφόσον τα λάθη που ενδιαφέρουν μειώνονται από 74 (Πίνακας 5.11) στα 65 (Πίνακας 5.13). Επειδή όμως τα συνολικά λάθη παραμένουν σχεδόν σταθερά (100 αντί για 101) στο δίκτυο με τα 3 κρυφά επίπεδα αυξάνονται τα λάθη Τύπου II, που δεν ενοχλούν τόσο (από 27 σε 35). Άμεσο αποτέλεσμα αυτού είναι η μείωση του Sensitivity στο 84,7%, γεγονός που να μην είναι επιθυμητό, αλλά δεν αποτελεί ιδιαίτερο πρόβλημα στην πράξη (Πίνακας 5.13).

Πίνακας 5.13 Πίνακας Προβλέψεων/Πραγματικότητας και Στατιστικά Μέτρα Κατηγοριοποίησης του Συνελκτικού Δικτύου για το Συνολικό Πραγματικό Πρόβλημα (3 κρυφά επίπεδα)

	1ο Στάδιο	Πραγματικά Δεδομένα	Πραγματικά Δεδομένα
	π/π	Κενή	Μη Κενή
Πρόβλεψη	Κενή	175	84
Πρόβλεψη	Μη Κενή	54	743
	2ο Στάδιο	Πραγματικά Δεδομένα	Πραγματικά Δεδομένα
	π/π	Κενή	Μη Κενή
Πρόβλεψη	Κενή	194	65
Πρόβλεψη	Μη Κενή	35	762
AUC	0,884	Accuracy	0,905
Λάθος Τύπου I	0,079	Λάθος Τύπου II	0,153
Specificity	0,921	Sensitivity	0,847
False Empty Rate	0,251	False Not Empty Rate	0,044
Positive Pred. Value	0,749	Negative Pred. Value	0,956

Η καμπύλη ROC, η οποία σχηματίζεται από τις Sensitivity και Specificity βελτιώνεται όσο οι τιμές αυτές προσεγγίζουν το 1. Εφόσον και οι δύο τιμές στα

συγκεκριμένα δίκτυα είναι κοντά στο 0.9 είναι εμφανής η επιτυχία του δικτύου. Ακόμη την επιτυχία της καμπύλης, άρα και την επιτυχία του συνελκτικού δικτύου, την υποδηλώνει και το στατιστικό μέτρο AUC, το οποίο υπολογίζει το εμβαδό υπό την καμπύλη. Τιμές του μέτρου AUC κοντά στο 0.9, όπως η συγκεκριμένη, υποδηλώνουν δίκτυο πολύ καλά εκπαιδευμένο.

Άλλες Παρατηρήσεις

Παρατίθεται ο Πίνακας 5.14 στον οποίο καταγράφονται τα λάθη των δικτύων που αναλύθηκαν, με δύο και τρία κρυφά επίπεδα και εντοπίζεται επικάλυψη λαθών 52%.

Πίνακας 5.14 Θέσεις Λαθών των Συνελκτικών Δικτύων Με 2 και 3 Κρυφά Επίπεδα.
Τα υπογραμμισμένα με Κίτρινο Αντιστοιχούν στα Κοινά Λάθη

36	277	476	874	17	332	539	874
38	279	504	887	36	336	558	886
43	280	513	888	45	337	564	888
45	282	522	895	54	338	565	890
54	330	558	896	56	339	574	892
56	336	565	897	57	340	582	893
57	339	576	898	58	341	585	900
58	340	595	900	63	351	594	927
69	341	598	923	70	352	595	928
70	342	610	927	75	353	596	936
71	351	621	936	94	360	597	942
105	352	633	938	107	368	598	949
106	353	634	942	108	369	608	957
107	359	684	949	117	405	618	958
108	360	693	958	155	407	684	959
112	369	700	959	165	431	693	960

117	373	728	960	166	432	702	987
126	396	730	969	238	477	728	989
165	405	731	987	245	489	730	994
166	413	732	997	249	496	731	995
174	420	738	999	256	499	732	1035
202	423	742	1034	277	500	742	1036
256	432	799	1035	279	501	766	1049
257	439	835	1036	330	502	788	
270	474	843		331	534	789	
	475	872			536	835	

Ακόμη, παρατηρήθηκε ότι κατά το δεύτερο στάδιο εκπαίδευσης των συνελκτικών νευρωνικών δικτύων, η ορμή αποτελεί αρνητικό παράγοντα εφόσον επιδιώκεται η αργή εκμάθηση. Παρατηρήσεις έγιναν και ως προς το βαθμό συστηματοποίησης “λ”, ο οποίος ποτέ δεν φάνηκε να βελτιώνει την απόδοση του συστήματος. Τέλος, παρατηρήθηκε επίσης ότι στα αντιπροσωπευτικά υποσύνολα εκπαίδευσης, η μείωσή τους, επιφέρει βελτίωση στην απόδοση του δικτύου.

Τέλος, αξίζει να σημειωθεί μία επιπρόσθετη ενδιαφέρουσα παρατήρηση, που έρχεται να επιβεβαιώσει την θεωρία ότι τα συνελκτικά δίκτυα, των οποίων ο πυρήνας συνέλιξης δεν διαιρείται ακριβώς με την διάσταση της εικόνας, χάνει την πληροφορία των άκρων με αποτέλεσμα να μειώνεται η απόδοση του δικτύου. Σε μία παρόμοια δοκιμή του δικτύου του Πίνακα 5.4, με πυρήνες συνέλιξης ακριβώς με την ανάποδη σειρά, δηλαδή (6,4)-(7,3) χειρότερε η δομή του συνελκτικού δικτύου. Για το λόγο αυτό, η απόδοσή του δεν ξεπερνούσε ποτέ το 84% και ήταν σαφώς χαμηλότερη από την αρχική προσέγγιση. Μία ένσταση σε αυτή τη δοκιμή πάντως είναι η αλλαγή της μορφής του δικτύου. Θεωρήθηκε καλύτερη για σύγκριση με την αρχική περίπτωση, η αλλαγή της σειράς των πυρήνων συνέλιξης και όχι η διατήρηση της σειράς τους και η αλλαγή των επιμέρους διαστάσεων τους π.χ. (7,4)-(5,4). Η επιλογή αυτή οφείλεται στο ότι με αυτό τον τρόπο διατηρείται η ένταση της συνέλιξης και όπως προέκυψε με την διαφορετική σειρά καταστρέφεται η δομή. Αντίθετα διατηρώντας την ίδια δομή και αλλάζοντας τις επιμέρους διαστάσεις, πραγματοποιείται διαφορετική συνέλιξη που μπορεί να αποτελεί την αιτία της χειρότερης απόδοσης του δικτύου.

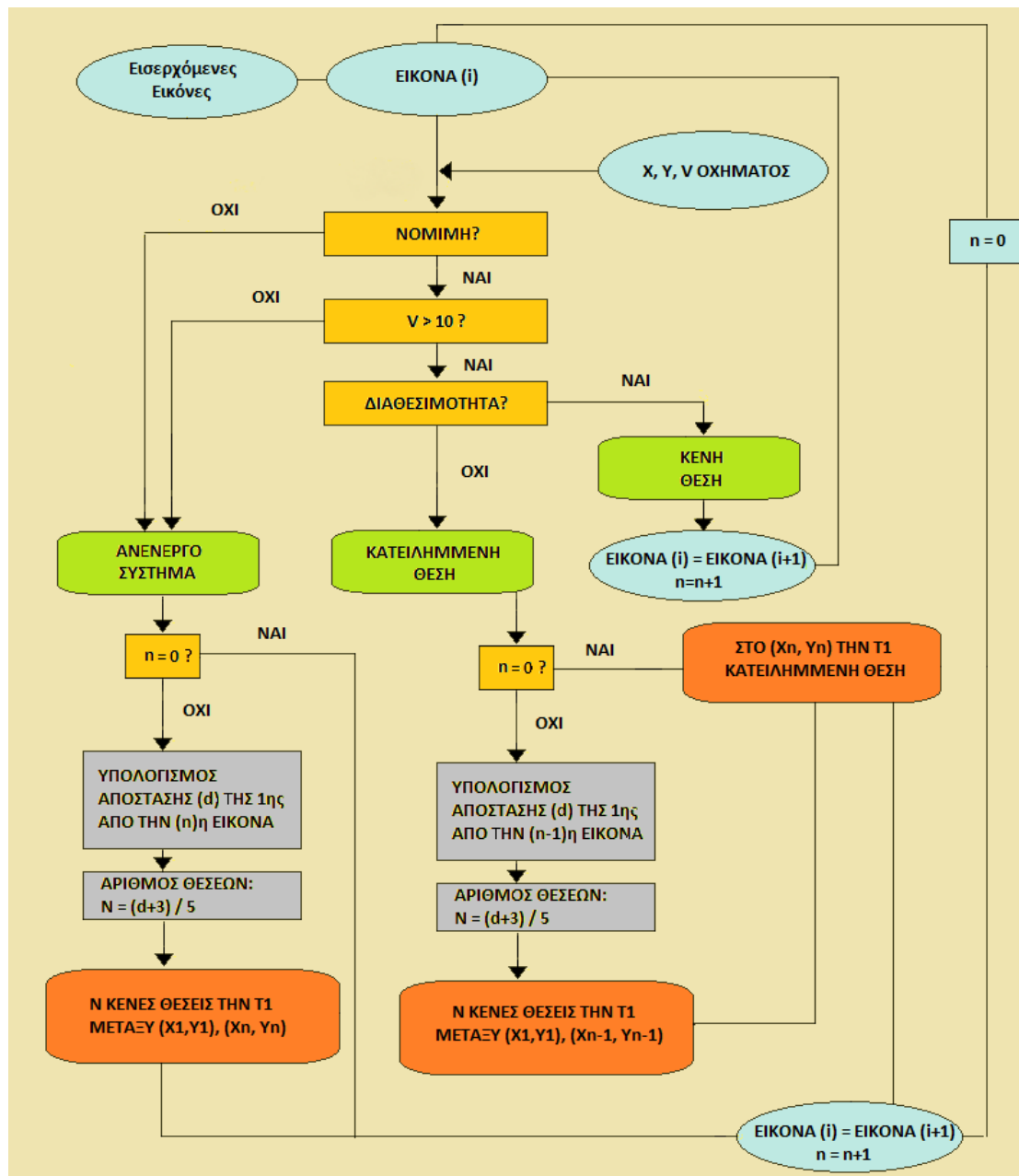
5.4 ΕΝΑ ΕΥΦΥΕΣ ΣΥΣΤΗΜΑ ΣΤΑΘΜΕΥΣΗΣ

Για να αξιοποιηθεί η μεγάλη ακρίβεια των συνελκτικών δικτύων που επιτεύχθηκε στην ανάλυση των Υποκεφαλαίων 5.1-5.3, πρέπει να γίνει ένας σχεδιασμός για την εφαρμογή του προτύπου σε πραγματικές συνθήκες. Λαμβάνοντας υπόψη τις παραδοχές που αναφέρονται στο Υποκεφάλαιο 3.2, σε αυτό το υποκεφάλαιο παρουσιάζεται μία ολοκληρωμένη προσέγγιση για την διαχείριση της εισερχόμενης πληροφορίας, που χάρις την μαζικότητά της λόγω πληθοπορισμού, μετατρέπεται σε χρήσιμη πληροφορία για τους μετακινούμενους σε πραγματικό χρόνο. Συγκεκριμένα, εξετάζεται η εύρεση των πρακτικών προβλημάτων της πραγματικότητας που προκύπτουν από την εφαρμογή του προτύπου και προτείνεται σε αδρές γραμμές μία μεθοδολογία η οποία τα αντιμετωπίζει με τον βέλτιστο τρόπο. Ακολουθεί ένα Διάγραμμα Ροής (Εικόνα 6.1), το οποίο παρουσιάζει συνοπτικά όλες τις φάσεις που πρέπει να υιοθετηθούν για να αποκτήσει το πρότυπο του Κεφαλαίου 5 ικανοποιητική εφαρμογή σε συνθήκες πραγματικότητας.

Η εφαρμογή του μοντέλου, για τις συνθήκες που αναφέρονται στις παραδοχές, απαιτεί την επίλυση και των παρακάτω προβλημάτων:

1. Αναγνώριση Κόμβου, Εισόδων-Εξόδων Κτηρίων, Εισόδων Ιδιωτικών Χώρων Στάθμευσης και άλλες θέσεις τις οποίες συναντά ένα όχημα και απαγορεύεται η στάθμευση.
2. Το κόκκινο του φωτεινού σηματοδότη, η κυκλοφοριακή συμφόρηση ή οποιαδήποτε άλλη κατάσταση που σχεδόν ακινητοποιεί το όχημα.
3. Η ύπαρξη πολλών συνεχόμενων κενών θέσεων και ο προσδιορισμός του αριθμού τους.

Αναλύοντας την προτεινόμενη μέθοδο της Εικόνας 6.1, θα αναδειχθεί και ο τρόπος επίλυσης όλων των παραπάνω προβλημάτων. Η ανάλυση γίνεται για την μία πλευρά της οδικής αρτηρίας και τα ίδια ακριβώς, ισχύουν και για την άλλη.



Εικόνα 5.1 Διάγραμμα Ροής Εφαρμογής του Προτύπου

Η εφαρμογή του μοντέλου βασίζεται στην ανανέωση ενός ηλεκτρονικού οδικού χάρτη, ο οποίος θα εμπεριέχει όλη την εισερχόμενη πληροφορία από όλα τα κινούμενα οχήματα του δικτύου σε πραγματικό χρόνο και θα την επιστρέφει στον χρήστη, για να την χρησιμοποιεί όπως αυτός επιθυμεί.

Αρχικά, εισάγεται μία εικόνα στο σύστημα προς αναγνώριση. Σε πρώτη φάση, πρέπει να ερευνηθεί η νομιμότητα της στάθμευσης στο σημείο αυτό. Για το λόγο αυτό, θα χρησιμοποιηθούν τα αρκετά ακριβή σήμερα συστήματα εντοπισμού θέσης, GPS.

Συγκεκριμένα, εάν κάθε όχημα έχει εγκατεστημένο στο κέντρο του ένα GPS είναι ικανή η μετατόπιση του σημείου αυτού κατά 3.5 μέτρα (ή όσο έχει γίνει παραδοχή ότι απέχει το κέντρο της θέσης στάθμευσης από το κέντρο μάζας του κινούμενου οχήματος) αριστερά ή δεξιά (από όποια πλευρά αναζητείται η θέση) για τον προσδιορισμό των συντεταγμένων της υπό μελέτης θέσης. Εφόσον η τεχνολογία έχει εξελιχτεί σε τέτοιο βαθμό που να υπάρχουν διαθέσιμοι οδικοί χάρτες πολύ μεγάλης ακρίβειας, θα μπορούσε σε αυτούς τους χάρτες να προστεθεί η πληροφορία της μη-νόμιμης θέσης στάθμευσης σε διάφορα σημεία του οδικού συστήματος. Κόμβοι, Είσοδοι-Εξοδοι Κτηρίων, Είσοδοι Ιδιωτικών Χώρων Στάθμευσης, Θέσεις ειδικά διαμορφωμένες για Α.Μ.Ε.Α. και άλλες πολλές θέσεις που υπάρχουν στις δομές των αστικών πόλεων και στα οποία παρότι υπάρχει διαθέσιμος χώρος για στάθμευση απαγορεύεται, θα μπορούσαν να ενταχθούν σαν πληροφορία στους προαναφερθέντες χάρτες. Στα σημεία αυτά, προτείνεται το σύστημα να μην λειτουργεί και να παραμένει ανενεργό, μέχρι να του δοθούν συντεταγμένες οι οποίες να αντιστοιχούν σε νόμιμη θέση στάθμευσης (Λύση Προβλήματος 1). Ακόμη, υποθέτοντας ότι η κάμερα είναι στο κέντρο του οχήματος και η εικόνα απεικονίζει 3.5 μέτρα από την κάθε πλευρά, το μοντέλο θα πρέπει να σταματά/ξεκινά να ανιχνεύει διαθέσιμες θέσεις, όχι στην αρχή/στο τέλος των απαγορευμένων θέσεων, αλλά και 1.5 μέτρο πριν την έναρξη/μετά το πέρας τους, ώστε η διαθέσιμη θέση που ανιχνεύεται να είναι εξολοκλήρου στην επιτρεπόμενη περιοχή.

Στην περίπτωση που οι συντεταγμένες υποδηλώνουν νόμιμη θέση, η εικόνα είναι διαθέσιμη προς αναγνώριση. Πριν από την αναγνώριση όμως, προτείνεται ένας έλεγχος της ταχύτητας του οχήματος ώστε να ελαφρυνθεί η φόρτωση του δικτύου. Συγκεκριμένα, ταχύτητες μικρότερες των 10 χμ/ώρα θεωρούνται ταχύτητες, οι οποίες υποδηλώνουν είτε κυκλοφοριακή συμφόρηση είτε κόκκινο σε φωτεινό σηματοδότη, είτε κάποια άλλη κατάσταση που σχεδόν ακινητοποιεί το όχημα με αποτέλεσμα η εισερχόμενη πληροφορία να παραμένει ουσιαστικά στάσιμη και επαναλαμβανόμενη. Επομένως για ταχύτητες μικρότερες των 10 χμ/ώρα προτείνεται και πάλι το σύστημα να παραμένει ανενεργό (Λύση Προβλήματος 2). Μόλις η ταχύτητα ξεπεράσει τα 10 χμ/ώρα, και επομένως εφόσον υπάρχει κάποια κίνηση, υπάρχει νόημα στην αναζήτηση θέσης στάθμευσης.

Αφού λοιπόν πραγματοποιηθούν οι δύο έλεγχοι για νομιμότητα και ταχύτητα, εφαρμόζεται το μοντέλο που αναπτύχθηκε και ανιχνεύει με ακρίβεια 90,34%, κενές θέσεις στάθμευσης. Στην περίπτωση που η θέση υπό εξέταση είναι κατειλημμένη, θα δίνεται σε ένα κεντρικό σύστημα η πληροφορία αυτή, μαζί με τις συντεταγμένες της θέσης αλλά και την ώρα που στάλθηκε, ώστε να ανανεώνεται αυτή η πληροφορία για την θέση αυτή, με κάθε διέλευση οχήματος, στον ηλεκτρονικό χάρτη που διατίθεται σε κάθε χρήση.

Εάν η ένδειξη είναι κενή θέση, η διαδικασία αυτή είναι λίγο πιο πολύπλοκη, εφόσον το σύστημα θα πρέπει να μπορεί να μετρήσει τον αριθμό των διαθέσιμων κενών θέσεων που προκύπτουν από έναν αριθμό συνεχόμενων φωτογραφιών που αναγνωρίζονται ως κενές. Για παράδειγμα, πόσες κενές θέσεις αντιστοιχούν σε 23 συνεχόμενες φωτογραφίες που «κρίνονται» ως κενές? Η απάντηση συσχετίζεται άμεσα με τον αριθμό των φωτογραφιών που λαμβάνονται ανά δευτερόλεπτο αλλά και την ταχύτητα του οχήματος. Στη συγκεκριμένη διπλωματική εργασία έγινε η παραδοχή για 10 φωτ. / δευτ. και η ανάλυση που ακολουθεί αφορά αυτή την δυνατότητα. Ο τρόπος που προτείνεται να αντιμετωπιστεί το παραπάνω θέμα είναι ο υπολογισμός της απόστασης μεταξύ της πρώτης κενής θέσης που εντοπίζεται μέχρι την εύρεση της τελευταίας συνεχόμενης κενής, η οποία προσδιορίζεται ως η αμέσως προηγούμενη της πρώτης κατειλημμένης που θα ανιχνευτεί ή του πρώτου σημείου που η στάθμευση καθίσταται παράνομη. Στην συνέχεια υπολογίζεται ο αριθμός των διαθέσιμων κενών θέσεων από την διαίρεση της απόστασης αυτής, αφού προστεθεί δύο φορές το ενάμισι μέτρο (μία για την πρώτη και μία για την τελευταία θέση) το οποίο χάνεται κατά τον υπολογισμό αυτής της απόστασης, προς τα 5 μέτρα που έχει γίνει παραδοχή στο Υποκεφάλαιο 3.2 ότι απαιτούνται για την στάθμευση ενός οχήματος (Λύση Προβλήματος 3). Ο αριθμός αυτός στρογγυλοποιείται προς τα κάτω και η περαιτέρω ανανέωσή του από επόμενα διερχόμενα οχήματα θα δώσει τελικά την πληροφορία αν κακώς στρογγυλοποιήθηκε προς τα κάτω, μόλις κάποια οχήματα αρχίσουν να σταθμεύουν σε αυτές. Τέλος, δίνεται αυτή η πληροφορία στον ηλεκτρονικό χάρτη, μαζί με τα όρια των συντεταγμένων και την χρονική στιγμή που αναλύθηκαν και έτσι εξασφαλίζεται η ανανέωση της παρεχόμενης πληροφορίας που φτάνει στον χρήστη.

Η διαδικασία αυτή επαναλαμβάνεται για κάθε φωτογραφία ή αριθμό φωτογραφιών που εισέρχονται στο σύστημα από ένα όχημα από την στιγμή της εκκίνησης της διαδρομής του μέχρι και το τέλος της στο εξεταζόμενο οδικό δίκτυο. Αν κάθε όχημα διαθέτετε ένα τέτοιο σύστημα, γίνεται αντιληπτό ότι από όπου διέρχονταν οχήματα, θα υπήρχε και αυτόματα η πληροφορία για την διαθεσιμότητα της στάθμευσης, σε συνδυασμό με την ακριβή τοποθεσία της αλλά και την ώρα της ενημέρωσης.

Μία τέτοια εφαρμογή του συστήματος θα μπορούσε ύστερα από ειδική μελέτη να δώσει και άλλες χρήσιμες πληροφορίες στο χρήστη. Για παράδειγμα, θα μπορούσε να υπολογίζει τον ρυθμό κατάληψης θέσεων ανά κάποιο χρονικό διάστημα και ανά περιοχή, ώστε να ενημερώνει τους χρήστες που πιθανόν αναζητούν θέση στάθμευσης σε κάποιο συγκεκριμένο μέρος, κάποια συγκεκριμένη στιγμή, για την ζήτηση που υπάρχει και την πιθανότητα να καταλειφθούν οι ανιχνευόμενες κενές θέσεις πριν ο χρήστης φτάσει σε αυτές. Στην ίδια λογική, θα μπορούσαν να αναπτυχθούν μοντέλα και για τα χρονικά διαστήματα στα οποία δεν κινούνται πολλά οχήματα και υπάρχει

γενικά μία στασιμότητα, αν και σύμφωνα με όσα ήδη αναφέρθηκαν, η στασιμότητα στο δίκτυο, υπονοεί την διατήρηση της διαθεσιμότητας των θέσεων όπως ήταν στην τελευταία ενημέρωση του ηλεκτρονικού οδικού χάρτη. Παρ' όλα αυτά, μία εμπειριστατωμένη μελέτη για την εναλλαγή των θέσεων στάθμευσης, τόσο χωρικά όσο και χρονικά, θα μπορούσε να ενισχύσει το υπάρχον μοντέλο.

Για την βέλτιστη εφαρμογή του μοντέλου, απαιτείται η επίλυση και άλλων προβλημάτων, τα οποία όμως ξεπερνούν τα όρια της ανάλυσης της παρούσας διπλωματικής εργασίας.

6. ΣΥΜΠΕΡΑΣΜΑΤΑ

6.1 ΓΕΝΙΚΑ

Είναι σημαντικό σε αυτό το σημείο, να πραγματοποιηθεί μία σύνοψη των σημαντικότερων σημείων που τονίστηκαν στην παρούσα διπλωματική εργασία και να αξιολογηθούν όλα τα συμπεράσματα που προέκυψαν, τόσο από την ανάλυση, όσο και από όλη την μελέτη που εκπονήθηκε για την επίλυση του προβλήματος της εύρεσης κενής θέσης στάθμευσης.

Ο σκοπός αυτής της διπλωματικής εργασίας ήταν η επίλυση του προβλήματος της εύρεσης κενής θέσης στάθμευσης στις οδικές αρτηρίες των αστικών κέντρων, όπου η αναζήτηση για θέση στάθμευσης προκαλεί επιπτώσεις τόσο στο χρήστη, όσο και στο σύστημα και στο περιβάλλον. Για το λόγο αυτό, αποτελεί υψηλή προτεραιότητα η επίλυσή του, αλλά και η δυνατότητα εφαρμογής του, τόσο στα σύγχρονα, όσο και στα μελλοντικά αστικά κέντρα.

6.2 ΜΕΘΟΔΟΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ

Η ανάγκη για την επίλυση των προβλημάτων που δημιουργεί η δυσκολία εύρεσης κενής θέσης στάθμευσης για τους μετακινούμενους στα αστικά κέντρα αποτέλεσε το κίνητρο για την πραγματοποίηση της παρούσας διπλωματικής εργασίας. Η μεθοδολογική προσέγγιση που ακολουθήθηκε έλαβε υπόψη τόσο την εξέλιξη της τεχνολογίας και τις δυνατότητες που αυτή προσφέρει, όσο και τις σύγχρονες τάσεις του πληθοπορισμού και της υιοθέτησης των «Ευφυών Συστημάτων» στα αστικά κέντρα.

Αρχικά, διερευνήθηκε η σχέση της θέσης της κάμερας πάνω στο όχημα (ύψος και γωνία λήψης) με το μήκος παρά το κράσπεδο που αποτυπώνεται στις παραγόμενες φωτογραφίες, και κατ' επέκταση, την ύπαρξη ή μη ελεύθερης θέσης παρά το κράσπεδο. Η χρήση της εικόνας κρίθηκε αναγκαία, εφόσον αποτελεί τον πιο διαδεδομένο και αποτελεσματικό τρόπο για την βελτίωση της απόδοσης της υπολογιστικής όρασης.

Στην συνέχεια, αναπτύχθηκε μία πολυεπίπεδη μεθοδολογία για την επίλυση του προβλήματος, με την χρήση συνελκτικών νευρωνικών δικτύων, χάρις στην ικανότητά τους να ανιχνεύουν χαρακτηριστικά εντός των εικόνων. Αρχικά επιλύθηκαν μικρότερα προβλήματα τα οποία θέσαν τις βάσεις για την επίλυση και του συνθετότερου, μεγαλύτερου προβλήματος. Σε πρώτο στάδιο, επιλύθηκε το πρόβλημα με ένα υποσύνολο δεδομένων στο οποίο οι εικόνες που χρησιμοποιήθηκαν ήταν πιο ξεκάθαρες ως προς την κατηγοριοποίησή τους ως 0/1. Ουσιαστικά, άλλου τύπου εμπόδια εντός των εικόνων, θεωρήθηκαν συνθετότερα ως προς την αναγνώρισή τους. Σε δεύτερο στάδιο, χρησιμοποιήθηκαν όλες οι φωτογραφίες και αφού έγινε αντιληπτό ένα συστηματικό λάθος που υπονόμει την απόδοση του δικτύου, αφαιρέθηκαν κάποια από τα δεδομένα και το δίκτυο εκπαιδεύτηκε με ακρίβεια 90%. Πραγματοποιήθηκε μια σειρά από δοκιμές, για πολλές διαφορετικές μορφές του συνελκτικού δικτύου ή παραμέτρους του, ώστε να επιβεβαιωθεί η βέλτιστη απόδοσή του.

Στο τέλος, διερευνήθηκε μία στρατηγική εφαρμογής του προτύπου, η οποία κλήθηκε να αντιμετωπίσει πολλά από τα προβλήματα που θα αντιμετωπίσει το πρότυπο υπό

πραγματικές συνθήκες. Συγκεκριμένα, σχεδιάστηκε η αρχιτεκτονική αυτού του συστήματος και διατυπώθηκαν επιμέρους προβλήματα που πρέπει να επιλυθούν με αντίστοιχους αλγόριθμους για την υλοποίησή του.

6.3 ΒΑΣΙΚΑ ΣΥΜΠΕΡΑΣΜΑΤΑ

Η επί τόπου έρευνα και η διερεύνηση για την βέλτιστη τοποθέτηση της κάμερας στο όχημα, οδήγησαν στο συμπέρασμα ότι η ασφαλτος αποτελεί ένα πολύ καλό κριτήριο διαχωρισμού των θέσεων και συνεπώς αναγνώρισης της ύπαρξης ή μη οχήματος ή εμποδίου σε κάποιο οδικό τμήμα. Ακόμη διαπιστώθηκε ότι η επιλογή της κάμερας και των δυνατοτήτων της αποτελούν καθοριστικό παράγοντα στην επιτυχή ανάλυση, εφόσον για παράδειγμα η δυνατότητα της κάμερας για λήψη περισσότερων φωτογραφιών ανά δευτερόλεπτο, μειώνει την απαιτούμενη απόσταση που προστίθεται πέραν των πέντε μέτρων στο πλάτος της εισερχόμενης φωτογραφίας για να μην χαθεί κάποια θέση στάθμευσης από το διερχόμενο όχημα, με αποτέλεσμα να μειώνονται τόσο χωρικά όσο και χρονικά οι απαιτήσεις του υπολογιστή για την επίλυση του προβλήματος. Τέλος, παρατηρήθηκε ότι η μεγαλύτερη ταχύτητα, τουλάχιστον σε αποδεκτά όρια, δεν επηρεάζει την απόδοση της ανάλυσης του συνελκτικού δικτύου παρότι μειώνεται η καθαρότητά της (κουνημένη φωτογραφία).

Ένα επιπλέον συμπέρασμα ως προς την μορφή των εισερχόμενων εικόνων, αποτελεί η διαπίστωση ότι η υιοθέτηση μίας προσέγγισης κατανόησης της υπολογιστικής όρασης μπορεί να βελτιώσει πολύ το τελικό αποτέλεσμα. Ειδικότερα, ο ερευνητής πρέπει να αφιερώσει χρόνο στα δεδομένα του και να αντιληφθεί αυτό που «βλέπει» ο ηλεκτρονικός υπολογιστής μέσα από αυτά. Μετά από ελαφρύ πειραματισμό, μπορεί να κατανοήσει ποια χαρακτηριστικά των εισερχόμενων δεδομένων επηρεάζουν πιο πολύ το δίκτυό του και ποια αποτελούν θόρυβο. Με την κατάλληλη διαδικασία μπορεί να τα περικόψει ή να τα μεταβάλλει όπως κρίνει κάθε φορά σκόπιμο για να βελτιώσει τα αποτελέσματα του συνελκτικού του δικτύου. Η ίδια λογική ισχύει και αντιστρόφως. Μελετώντας τα λάθη στα οποία υποπίπτει το μοντέλο που αναπτύχθηκε

και αναγνωρίζοντας μοτίβα λαθών ή κάποια συγκεκριμένη αιτία που τα προκαλεί, ο ερευνητής είναι σε θέση να βελτιώσει σημαντικά την απόδοση του δικτύου.

Το πιο βασικό συμπέρασμα της παρούσας διπλωματικής εργασίας αφορά την μέθοδο που χρησιμοποιήθηκε και την αποτελεσματικότητά της. Ειδικότερα, μετά από την ανάλυση του Κεφαλαίου 5, έγινε σαφές ότι τα συνελκτικά νευρωνικά δίκτυα είναι πολύ ικανά μοντέλα για την ανίχνευση χαρακτηριστικών στις φωτογραφίες, γεγονός που επιβεβαιώνει την αντίστοιχη θεωρία. Ακόμη, διαπιστώθηκε ότι εφαρμόζεται πολύ ικανοποιητικά και στην ανίχνευση οχημάτων για την εύρεση κενής θέσης στάθμευσης και συγκεκριμένα ότι τα 2 κρυφά επίπεδα εκπαίδευσης είναι ήδη αρκετά για να επιτευχθούν πολύ υψηλές ακρίβειες (της τάξης του 90%).

Τα συμπεράσματα που προκύπτουν από την εκπαίδευση του δικτύου παρουσιάζουν επίσης ιδιαίτερο ενδιαφέρον και συνοπτικά αναφέρονται κάποια από τα πιο σημαντικά:

- Για συνελκτικά νευρωνικά δίκτυα με 2-3 κρυφά επίπεδα και Σύνολο Εκπαίδευσης ως 5000 φωτογραφίες η εύρεση του ρυθμού εκμάθησης γίνεται εμφανής ως τις πρώτες 6-8 Εποχές.
- Η χρήση ολοένα και μικρότερων ρυθμών εκμάθησης βελτιώνουν την απόδοση του συστήματος κυρίως όταν αυτό έχει ένα σημαντικό επίπεδο πολυπλοκότητας.
- Η ορμή και η παράμετρος συστηματοποίησης στο Δεύτερο Στάδιο Εκπαίδευσης προτιμάται να είναι αρκετά μικρές έως ανύπαρκτες για την βελτίωση της απόδοσης του συνελκτικού δικτύου.
- Η απλούστευση του προβλήματος στη πρώτη φάση της μελέτης του και η επίλυσή του, δίνει τις βάσεις για την καλύτερη κατανόηση και αντιμετώπιση του συνολικού προβλήματος.

Τέλος, διαπιστώνεται ότι η πρόοδος της τεχνολογίας και η επικράτηση του πληθοπορισμού έχουν θέσει τα απαραίτητα θεμέλια για την ανάπτυξη των «ευφών» συστημάτων στις «ευφυείς» πόλεις και για την άριστη εφαρμογή τους με την βοήθεια των χρηστών (από το χρήστη, για το χρήστη). Οι δυνατότητες που προσφέρονται σε αυτήν την κατεύθυνση ξεπερνούν ήδη την ικανότητα των συστημάτων που αναπτύσσονται για την επίλυση διαφόρων ανθρωπίνων προβλημάτων και για το λόγο αυτό τα αποτελέσματα της παρούσας διπλωματικής εργασίας αποκτούν ιδιαίτερη σημασία.

6.4 ΠΡΟΤΑΣΕΙΣ ΓΙΑ ΠΕΡΑΙΤΕΡΩ ΕΡΕΥΝΑ

Το πρόβλημα προς επίλυση της παρούσας διπλωματικής είναι πολυσύνθετο και έχουν αναπτυχθεί πολλοί τρόποι προσέγγισής του. Η χρήση των συνελκτικών δικτύων που χρησιμοποιήθηκε στην παρούσα διπλωματική αποτελεί μία καινοτόμα λύση, η οποία όχι μόνο πετυχαίνει μεγάλη ακρίβεια πρόβλεψης των κενών θέσεων, αλλά επίσης εναρμονίζεται πλήρως με τις σύγχρονες τάσεις και τις εξελίξεις της τεχνολογίας. Βέβαια, στα πλαίσια της έρευνας που πραγματοποιήθηκε τέθηκαν κάποιοι περιορισμοί και έγιναν κάποιες παραδοχές. Ο αυστηρός ορισμός του προβλήματος, περιορίζει την εφαρμογή του και, επομένως, επισημαίνει σημεία που απαιτούν περαιτέρω έρευνα.

Σημαντικό στοιχείο που θα πρέπει να διερευνηθεί συστηματικά είναι η επιρροή των καιρικών συνθηκών στην ακρίβεια των προτεινόμενων τεχνικών πρόβλεψης ελεύθερης θέσης στάθμευσης. Θεωρητικά, το μοντέλο μπορεί να εκπαιδευτεί στην αναγνώριση κενών θέσεων στάθμευσης και υπό κακές καιρικές συνθήκες (βροχοπτώση, χιόνι, ομίχλη κ.α.). Η μεγάλη αλλαγή στις φωτεινότητες των εισερχόμενων εικόνων, αλλά και στον περιβάλλοντα χώρο των απεικονιζόμενων οχημάτων (π.χ. χιονισμένο τοπίο) αποτελούν σημαντικά θέματα προς επίλυση. Αντίστοιχα, θα ήταν επιθυμητό να διερευνηθεί η χρήση του μοντέλου και κατά τη διάρκεια της νύχτας. Οι εισερχόμενες εικόνες την νύχτα, έχουν εντελώς διαφορετικά χαρακτηριστικά και σημεία αναγνώρισης, με αποτέλεσμα να αποτελούν ένα καινούριο πρόβλημα προς επίλυση από το μηδέν.

Ακολουθεί μία σειρά από προτάσεις για περαιτέρω έρευνα, που θα συμπλήρωναν ή θα βελτίωναν την υπάρχουσα έρευνα και το μοντέλο που αναπτύχθηκε, αφού πρώτα διερευνηθεί η εφαρμογή του κατά τη διάρκεια της νύχτας και υπό κακές καιρικές συνθήκες όπως προαναφέρθηκε.

1. Αρχικά, θα ήταν σημαντικό να πραγματοποιηθεί μία ανάλυση για το μήκος των ανιχνευόμενων ελεύθερων θέσεων. Συγκεκριμένα, τα οχήματα έχουν διάφορα μεγέθη και κάθε χρήστης θα επιθυμούσε να γνωρίζει αν το δικό του όχημα χωρά σε μία θέση στάθμευσης. Για το λόγο αυτό, μία πρόσθετη έρευνα, θα μπορούσε να λύσει το πρόβλημα θεωρώντας ως λύση τέσσερα

πιθανά αποτελέσματα (κατειλημμένη θέση, μικρή θέση, μεσαία θέση, μεγάλη θέση). Αρκεί να αποδειχτεί ότι μπορεί να ελεγχθεί το μέγεθος της θέσης που προβλέπεται –πράγμα που από τα επιμέρους συμπεράσματα της παρούσας εργασίας φαίνεται εφικτό- και τότε με την κατάλληλη εκπαίδευση, το δίκτυο θα μπορεί να προβλέψει οποιοδήποτε μέγεθος θέσης του ζητείται.

2. Ακόμη, σημαντικό θα ήταν να επιλυθεί το πρόβλημα της περίπτωσης που οι οδικές αρτηρίες διαθέτουν παραπάνω από μία λωρίδα. Συγκεκριμένα, θα πρέπει να διερευνηθεί ένας τρόπος διαχωρισμού από το μοντέλο και κατανόησής του, σχετικά με την θέση του κινούμενου οχήματος στην αρτηρία και αν αυτό βρίσκεται στην εξωτερική λωρίδα (προς το πεζοδρόμιο) να γίνονται αποδεκτές οι εισερχόμενες εικόνες, ενώ αν βρίσκεται στις εσωτερικές να απορρίπτονται, επειδή παρεμβάλλονται τα οχήματα της εξωτερικής λωρίδας. Θα μπορούσε βέβαια να διερευνηθεί τρόπος επίλυσης και αυτού του προβλήματος.
3. Επιπλέον, είναι σκόπιμο να ερευνηθεί η εφαρμογή του μοντέλου στην πράξη, σε πολύ μεγαλύτερο βάθος από αυτό που παρουσιάστηκε στην παρούσα διπλωματική εργασία. Ειδικότερα, μπορούν να μελετηθούν αναλυτικότερα τρόποι καταμέτρησης των συνεχόμενων διαθέσιμων θέσεων στάθμευσης, οι κανόνες παροχής και λήψης πληροφοριών του κάθε χρήστη, η εναλλαγή των θέσεων σε συγκεκριμένες περιοχές και συγκεκριμένες ώρες, όπου δεν κυκλοφορούν πολλά οχήματα άρα που το σύστημα δεν ενημερώνεται τακτικά κ.α.
4. Θα μπορούσε επίσης να γίνει μία λεπτομερέστερη ανάλυση στην διαχείριση και διόρθωση των αποτελεσμάτων του συνελκτικού δικτύου, ώστε αναζητώντας μοτίβα και συνήθειες ακολουθίες στις προβλέψεις του για συγκεκριμένα χαρακτηριστικά οδήγησης (π.χ. ταχύτητα), να βελτιωθεί η απόδοση του δικτύου αισθητά (από 90% στο 95%).

Τέλος, προχωρώντας σε μία γενικότερη θεώρηση, παρατίθενται ενδεικτικά και άλλα συγκοινωνιακά (ή σχετικά με αυτά) προβλήματα, τα οποία θα μπορούσαν να επιλυθούν με την προσέγγιση που χρησιμοποιήθηκε στην παρούσα διπλωματική εργασία:

- ✓ Αναγνώριση Πινακίδων Αυτοκινήτων κυρίως σε χώρους στάθμευσης
- ✓ Αναγνώριση του Οδηγού ενός οχήματος κατά την είσοδό του σε κάποιο χώρο
- ✓ Αναγνώριση Πεζών ή άλλων εμποδίων επί της οδού
- ✓ Αναγνώριση Άλλων Οχημάτων
- ✓ Αναγνώριση Οδικής Σήμανσης ή του Χρώματος των Φωτεινών Σηματοδοτών

7. ΒΙΒΛΙΟΓΡΑΦΙΑ

- Ahmed, K. (2016, 04 25). *Ford: 'We assume Apple is working on a car'*. Retrieved 06 2016, from BBC NEWS: <http://www.bbc.com/news/business-36126591>
- Birch, A. (2016, 02 12). *The 2016 Ford F-150 Has An Optional 360° Camera*. Retrieved from FORD AUTHORITY: <http://fordauthority.com/2016/02/the-2016-ford-f-150-has-an-optional-360-camera/>
- C.I.A. (2015). *THE WORLD FACTBOOK*. Retrieved from Central Intelligence Agency: <https://www.cia.gov/library/publications/the-world-factbook/index.html>
- Chen, X., Santos-Neto, E., & Ripeanu, M. (2012). Crowd-based Smart Parking: A Case Study for Mobile Crowdsourcing.
- Coric, V., & Gruteser, M. (2013). Crowdsensing Maps of On-Street Parking Spaces. *Proceedings of the 2013 IEEE International Conference on Distributed Computing in Sensor Systems*, 115-122.
- D'Aloia, M., Rizzi, M., Russo, R., & Pellicani, M. N. (2015). A Marker-Based Image Processing Method for Detecting Available Parking Slots from UAVs. *Springer International Publishing Switzerland 2015*, 275-281.
- Dettmers, T. (2015). Deep Learning in a Nutshell: Core Concepts.

- E.E.A. (2012). *Passenger car ownership in the EEA*. Retrieved 06 2016, from Europa Environment Agency: <http://www.eea.europa.eu/data-and-maps/figures/passenger-car-ownership-in-the-eea>
- Estellés-Arolas, E., & González Ladrón-de-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information Science*.
- European Commission. (2015). *EU transport in figures, Statistical Pocketbook 2015*. Retrieved 06 2016, from European Commission, Transport: http://ec.europa.eu/transport/facts-fundings/statistics/pocketbook-2015_en.htm
- FORD(a). (2016). *No Lights? No Problem! Ford Fusion Autonomous Research Vehicles use LIDAR Snsor Technology to see in the Dark*. Retrieved from THE FORD MOTOR COMPANY MEDIA CENTER: <https://media.ford.com/content/fordmedia/fna/us/en/news/2016/04/11/no-lights--no-problem--ford-fusion-autonomous-research-vehicles-.html>
- FORD(b). (2016). *2016 EDGE, Front 180-Degree Camera*. Retrieved 06 2016, from FORD: <http://www.ford.com/crossovers/edge/gallery/videos-and-demos/Videos/front-180-camera/FMFT0701000H/>
- FORD(c). (2016). *2017 EXPLORER, Front 180-Degree Camera*. Retrieved 06 2016, from FORD: <http://www.ford.com/suvs/explorer/gallery/videos-and-demos/Videos/front-180-degree-camera/FMFT0850000H/>
- Gantelet, E., & Lefauconnier, A. (2006). The time looking for a parking space:strategies, associated nuisances and stakes of parking management in France. *SARECO research report, Association for European Transport and Contributors*.
- Google. (2016). *Technology and Costs*. Retrieved 06 2016, from Google's Autonomous Vehicle: <http://googlesautonomousvehicle.weebly.com/technology-and-costs.html>
- Grassi, G., Sammarco, M., Bahl, P., & Pau, K. J. (2015). Poster: Park Master - Leveraging Edge Computing in Visual analytics.
- Heaton, J. (2008). *Introduction to Neural Networks for Java, 2nd Edition*. Heaton Research.
- Higginbotham, S. (2015). How Facebook is teaching computers to see. *Fortune*.

- Huang, C.-C., & Tran, H. V. (2015). A multi-layer discriminative framework for parking space detection. *IEE International Workshop on Machine Learning for Signal Processing*.
- Huang, S. C., & Chen, B. H. (2013). Highly accurate moving object detection in variable bit rate video-based traffic monitoring systems. *IEEE Transactions on neural networks and learning systems* 24(12), (pp. 1920-1931).
- INFINITI. (2016). *AROUND VIEW MONITOR*. Retrieved 06 2016, from INFINITI: <http://www.infinitiusa.com/now/technology/around-view-monitor.html>
- JAGUAR. (2016). *XJ*. Retrieved 06 2016, from JAGUAR: <http://www.jaguarusa.com/all-models/xj/xj-models/xj-r-sport-awd.html>
- Jones, M., & Li, L. (2012). Vacant Parking Spot Detection.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems 25, NIPS Proceedings*.
- Lawrence, S., Giles, C., & Tsoi, A. C. (1996). What Size Neural Network Gives Optimal Generalization? Convergence Properties of Backpropagation. *Institute for Advanced Computer Studies, University of Maryland*.
- LeCun, Y., Bengio, Y., & Haffner, P. (1998). Gradient-Based learning applied to document recognition. *Proc of the IEEE*.
- Liao, R., Roman, C., Ball, P., Ou, S., & Chen, L. (2016). Crowdsourcing On-street Parking Space Detection.
- Marthur, S., T. Jin, N. K., Chandrasekaran, J., Xue, W., Grutese, M., & W.Trappe. (2010). "Parknet: drive-by sensing of road-side parking statistics". *Proceedings of the 8th international conference on Mobile systems, applications, and services*, ACM DL 123-136.
- Medvedev, A., Zaslvsky, A., Khoruzhnikov, S., & Grudinin, V. (2015). Reporting Road Problems in smart Cities Using Open IoT FrameWork. *Springer International Publishing Switzerland 2015*, 169-182.
- Nawaz, S., Efstratiou, C., & Mascolo, C. (2013). ParkSense: a smartphone based sensing system for on-street parking. *Proceedings of the 19th annual*

international conference on Mobile computing & networking, ACM DL 75-86.

Newcomb, D. (2016, 04 27). *Ford, Google, Uber, Lyft, Volvo Create Coalition To Push For Federal Self-Driving Car Policy*. Retrieved from Forbes/Autos/#TestDrive:
<http://www.forbes.com/sites/dougnewcomb/2016/04/27/ford-google-uber-lyft-volvo-create-coalition-to-push-for-federal-self-driving-car-policy/#4ecfd66b4180>

NHTSA. (2014, 03 14). *NHTSA Announces Final Rule Requiring Rear Visibility Technology*. Retrieved 05 2016, from NHTSA:
<http://www.nhtsa.gov/About+NHTSA/Press+Releases/2014/NHTSA+Announces+Final+Rule+Requiring+Rear+Visibility+Technology>

Nielsen, M. A. (2015). *Neural Networks and Deep Learning* . Determination Press.

Ommeren, J. N., Wentink, D., & Rietveld, P. (2012). Empirical evidence on cruising for parking. *Transportation Research Part A* 46, 123-130.

Schaller Consulting. (2006). *Curbing cars: shopping, parking and pedestrian space in SoHo*. (S. Consulting, Producer) Retrieved 06 2016, from Transportation Alternatives: <https://transalt.org>

Senior, A., Heigold, G., & Marc'Aurelio Ranzato, K. y. (2013). An empirical study of learning rates in deep neural networks for speech recognition. *Google Inc*.

SFpark. (2012). *Sensors*. Retrieved 06 23, 2016, from SFpark: <http://sfpark.org/>

Shen, Y., He, X., Gao, J., Deng, L., & Mesnil, G. (2014). Learning Semantic Representations Using Convolutional Neural Networks for Web Search. WWW 2014. Retrieved from Microsoft.

Sherman, D. (2016, 02). *The Physiology of Semi-Autonomy and Test Results*. Retrieved 06 2016, from CAR AND DRIVER:
<http://www.caranddriver.com/features/semi-autonomous-cars-compared-tesla-vs-bmw-mercedes-and-infiniti-feature-the-physiology-of-semi-autonomy-and-test-results-page-6>

Shoup, D. C. (2006). Cruising for parking. *Transport Policy*, 13(6), 479-486.

- SIEMENS. (2015). *Smarter parking*. Retrieved 06 23, 2016, from SIEMENS/Mobility: <https://www.siemens.com/customer-magazine/en/home/mobility/smarter-parking.html>
- Smart Parking. (2016). *Smart Eye*. Retrieved 06 23, 2016, from Smart Parking: <http://www.smartparking.com/>
- Smart Santander. (2016). Retrieved 06 23, 2016, from Smart Santander: <http://www.smartsantander.eu/>
- Stanford. (2016). CS231n Convolutional Neural Networks for Visual Recognition.
- Suhr, J. K., Jung, H. G., Bae, K., & Kim, J. (2010). Automatic free parking space detection by using motion stereo-based #D construction. *Machine Vision and Applications*, 21:163-176.
- The Ford Motor Company Media Center. (2015, 01 06). *Mobility Experiment : Parking Spotter Atlanta*. Retrieved 06 23, 2016, from Ford: <https://media.ford.com/content/fordmedia/fna/us/en/news/2015/01/06/mobility-experiment-parking-spotter-atlanta.html>
- Tian, B., Morris, B. T., Tang, M., Liu, Y., Yao, Y., Gou, C., & Tang, S. (2015). Hierarchical and networked vehicle surveillance in ITS. *IEEE Transactions on Intelligent Transportation Systems* (2), (pp. 557-580).
- Tillemann, L., & McCormick, C. (2016, 04 14). *Will the Tesla Model 3 be the First Truly Self-driving Car?* Retrieved 06 2016, from THE NEW YORKER: <http://www.newyorker.com/business/currency/will-the-tesla-model-3-be-the-first-truly-self-driving-car>
- Transportation Alternatives. (2007). *No vacancy. Park Slope's parking problem and how to fix it*. Retrieved 06 2016, from Transportation Alternatives: <http://transalt.org>
- True, N. (2007). Vacant Parking Space Detection in Static Images.
- U.S.E.P.A. (2008). *Average Annual Emissions and Fuel Consumption for Gasoline-Fueled Passenger Cars and Light Trucks*. Retrieved 06 2016, from United States Environmental Protection Agency: <https://www3.epa.gov/otaq/consumer/420f08024.pdf>

- VOLVO. (2015). *Volvo Cars presents a unique solution for integrating self-driving cars into real traffic*. Retrieved 06 2016, from VOLVO CAR GROUP GLOBAL NEWSROOM: <https://www.media.volvocars.com/global/en-gb/media/pressreleases/158276/volvo-cars-presents-a-unique-system-solution-for-integrating-self-driving-cars-into-real-traffic>
- Weyand, T., Kostrikov, I., & Philbin, J. (2016). PlaNet - Photo Geolocation with Convolutional Neural Networks. *arXiv:1602.05314*.
- Worldsensing. (n.d.). Retrieved 06 23, 2016, from Fastprk: <http://www.fastprk.com/>
- Wu, Q., Huang, C., Wang, S.-y., Chiu, W.-c., & Chen, T. (2009). Robust Parking Space Detection Considering Inter-Space Correlation.