



Πανεπιστήμιο Πειραιώς – Τμήμα Πληροφορικής
Πρόγραμμα Μεταπτυχιακών Σπουδών
«Πληροφορική»

Μεταπτυχιακή Διατριβή

Τίτλος Διατριβής	Αναγνώριση πεζών από αυτόνομο ρομποτικό όχημα με χρήση τεχνητής νοημοσύνης και τεχνικών βαθιάς μάθησης. Autonomous robotic vehicle pedestrian detection using artificial intelligence and deep learning techniques.
Όνοματεπώνυμο Φοιτητή	Ευάγγελος Διαμαντής
Πατρώνυμο	Γεώργιος
Αριθμός Μητρώου	ΜΠΠΛ13016
Επιβλέπων	Θεμιστοκλής Παναγιωτόπουλος, καθηγητής

31/10/2018

Οκτώβριος 2018

Τριμελής Εξεταστική Επιτροπή

(υπογραφή)

(υπογραφή)

(υπογραφή)

Θεμιστοκλής Παναγιωτόπουλος
Καθηγητής

Δημήτριος Αποστόλου
Αναπληρωτής Καθηγητής

Άγγελος Πικράκης
Επίκουρος Καθηγητής

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες στον Δρ. Νικολάου Γρηγόρη για την ευχάριστη συνεργασία μας, τις χρήσιμες και ουσιαστικές υποδείξεις του και τη γενικότερη συμβολή και βοήθεια του κατά τη συγγραφή της παρούσας Μεταπτυχιακής Διατριβής. Ακόμα, θα ήθελα να ευχαριστήσω θερμά τον καθηγητή μου κύριο Παναγιωτόπουλο Θέμη για την πολύτιμή συνδρομή του και τη υποστήριξη στο πρόσωπο μου ώστε να συγγραφεί η παρούσα διατριβή, βοηθώντας ουσιαστικά και απλόχερα στην αντιμετώπιση των όποιων δυσκολιών παρουσιάστηκαν.

Επίσης, θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες και σε όλους τους υπόλοιπους καθηγητές του Προγράμματος Μεταπτυχιακών Σπουδών όπου παρακολούθησα, καθώς μου έδωσαν πολύτιμες γνώσεις, τόσο σε θεωρητικό όσο και σε πρακτικό επίπεδο, οι οποίες με βοήθησαν σε μεγάλο βαθμό, όχι μόνο στην υλοποίηση της παρούσας Μεταπτυχιακής Διατριβής αλλά και στην γενικότερη επαγγελματική μου πορεία. Παράλληλα, μου έδωσαν σημαντικό έναυσμα για την έναρξη των ερευνητικών μου δραστηριοτήτων και την ενασχόληση μου με τον τομέα της Τεχνητής Νοημοσύνης.

Τέλος, δεν θα ήθελα να παραλείψω τις θερμότερες ευχαριστίες μου, και σε όλους τους ανθρώπους του οικογενειακού μου και ευρύτερου περιβάλλοντος, που με στήριξαν ενεργά τόσο κατά την διάρκεια της μεταπτυχιακής μου διατριβής, όσο των Μεταπτυχιακών μου σπουδών εν γένει.

ΣΚΟΠΟΣ ΤΗΣ ΔΙΑΤΡΙΒΗΣ

Η παρούσα Μεταπτυχιακή Διατριβή, υλοποιήθηκε στα πλαίσια λειτουργίας του Μεταπτυχιακού Προγράμματος Σπουδών «Πληροφορική» του τμήματος Πληροφορικής, του Πανεπιστημίου Πειραιώς, με σκοπό την κάλυψη των απαιτήσεων για την κτήση του τίτλου Μεταπτυχιακών Σπουδών.

Εντός της Μεταπτυχιακής Διατριβής, επιχειρείται αρχικά η ανάλυση του πλαισίου και του τρόπου λειτουργίας των *Αυτόνομων Ρομποτικών Οχημάτων*, ενώ στη συνέχεια πραγματοποιείται μία προσπάθεια υλοποίησης ενός εξελιγμένου αλγορίθμου *Τεχνητής Νοημοσύνης* που υλοποιεί αναγνώριση πεζών από κάποιο *Αυτόνομο Ρομποτικό Όχημα*, με χρήση τεχνικών *Βαθιάς Μάθησης*.

Για την συγγραφή της παρούσας Μεταπτυχιακής Διατριβής, αξιοποιήθηκε πλήθος επιστημονικών συγγραμμάτων και πηγών, τόσο της εγχώριας όσο και της διεθνούς επιστημονικής βιβλιογραφίας. Σε κάθε τμήμα της Μεταπτυχιακής Διατριβής αυτής όπου χρησιμοποιείται επιστημονική γνώση ή αυτούσιο τμήμα προερχόμενα από τρίτο επιστημονικό σύγγραμμα ή πηγή, αυτά αναφέρονται ευκρινώς στο αντίστοιχο πεδίο και σε καμία περίπτωση δεν επιχειρείται οικειοποίηση πνευματικής ιδιοκτησίας τρίτου.

ΠΕΡΙΛΗΨΗ

Η υιοθέτηση σύγχρονων τεχνολογιών με προσανατολισμό τη βελτίωση και την εξάλειψη όλων των βασικών προβλημάτων που ταλανίζουν την σημερινή κοινωνία, αποτελεί αδιαπραγμάτευτη ανάγκη για τις τη σύγχρονη επιστημονική κοινότητα. Στο πλαίσιο υλοποίησης αυτής της φιλοσοφίας, η έρευνα και ανάπτυξη των αυτόνομων ρομποτικών οχημάτων θα οδηγούσε δυνητικά στην εξάλειψη μίας πληθώρας προβλημάτων που απασχολούν την ανθρωπότητα, ενώ παράλληλα θα συνεισέφερε ενεργά και καθοριστικά στην προαγωγή του βιοτικού επιπέδου καθώς και στην προστασία του περιβάλλοντος.

Η ραγδαία εξέλιξη που έχει σημειωθεί κατά τα τελευταία έτη, τόσο στον τομέα της τεχνολογίας των αυτόνομων ρομποτικών οχημάτων, όσο και της τεχνολογίας εν γένει, οφείλεται κατά μεγάλο βαθμό στην εξίσου ραγδαία εξέλιξη που έχει σημειώσει ο τομέας της *Τεχνητής Νοημοσύνης*. Η *Τεχνητή Νοημοσύνη*, περιλαμβάνοντας ένα πλήθος εξελιγμένων αλγορίθμων, τεχνικών και μεθόδων, επιτυγχάνει την εφαρμογή της ανθρώπινης εγκεφαλικής συλλογιστικής διαδικασίας, μέσω αλγοριθμικών, σε κοινά υπολογιστικά συστήματα.

Ως εκ τούτου, η μίμηση της εγκεφαλικής συλλογιστικής κατά την διαδικασία ελέγχου ενός οχήματος, από κάποιο υπολογιστικό σύστημα, προϋποθέτει την υλοποίηση και ενσωμάτωση πλήρων και εξελιγμένων αλγορίθμων μοντέλων που υλοποιούν το σύνολο των απαιτούμενων δραστηριοτήτων με χρήση αλγορίθμων, τεχνικών και μεθόδων *Τεχνητής Νοημοσύνης*, και ειδικότερα, *Βαθιάς Μάθησης*, η οποία αποτελεί το πλέον εξελιγμένο τομέα της *Τεχνητής Νοημοσύνης* που συντελεί στον αποτελεσματικό εντοπισμό προτύπων σε οπτικά περιβάλλοντα.

Σκοπό της παρούσας μεταπτυχιακής διατριβής, αποτελεί αφενός η παρουσίαση και ανάλυση της τεχνολογίας που φέρουν τα αυτόνομα ρομποτικά οχήματα και αφετέρου, η υλοποίηση ενός εξελιγμένου αλγορίθμου *Βαθιάς Μάθησης* ο οποίος είναι σε θέση να εντοπίσει - μέσω των κατάλληλων οπτικών συστημάτων που φέρει ένα αυτόνομο ρομποτικό όχημα – τους πεζούς που βρίσκονται εντός του πεδίου ανάλυσης του οπτικού συστήματος.

Τα πρώτα κεφάλαια της διατριβής αυτής, έχουν υλοποιηθεί εντός βιβλιογραφικού και αναφορικού πλαισίου ανάπτυξης, και περιλαμβάνουν τις θεματικές ενότητες που παρουσιάζουν τις βασικές έννοιες, τα ιστορικά στοιχεία εξέλιξης και τα δομικά συστήματα που αφορούν ένα αυτόνομο ρομποτικό όχημα. Εν συνεχεία, τα επόμενα κεφάλαια υλοποιούνται μέσα από ένα συνδυαστικό πλαίσιο βιβλιογραφικού περιεχομένου και περιεχομένου παρουσιαζόμενου μέσα από καταλλήλως δομημένα παραδείγματα που συντελούν στην περεταίρω κατανόηση των πραγματευομένων εννοιών.

Τέλος, στο τελευταίο τμήμα της παρούσας διατριβής παρουσιάζεται η ανάπτυξη ενός αλγορίθμου *Βαθιάς Μάθησης* που πραγματοποιεί αναγνώριση πεζών καθώς και η παρουσίαση των αποτελεσμάτων εργαστηριακών δοκιμών αυτού αλλά και πρόταση πιθανών περιθωρίων περεταίρω βελτίωσης του αλγορίθμου αυτού με χρήση των σύγχρονων εργαλείων *Τεχνητής Νοημοσύνης*.

ABSTRACT

The adoption of modern technologies with the aim of improving and eliminating all the basic problems that afflict today's society, is a non-negotiable need for the modern scientific community. In the context of implementation of this philosophy, research and development of autonomous robotic vehicles will potentially result in the elimination of a multitude of problems facing humanity, while actively contributed decisively to the standard of living and the environment protection.

The rapid progress made in recent years, both in the technology of autonomous robotic vehicles, and technology in general, due mainly to the equally rapid evolution that has made the field of *Artificial Intelligence*. *Artificial Intelligence*, including a multitude of sophisticated algorithms, techniques and methods, achieves the application of the human brain operation through algorithms to common computer systems.

Therefore, the imitation of human cerebral function during a vehicle driving procedure, by a computer, requires the implementation and integration of complete and sophisticated modeling algorithms which implement all the required activities using algorithms, techniques and methods of *Artificial Intelligence*, particularly *Deep Learning*, which is the most advanced field of *Artificial Intelligence* that contributes to the effective identification of prototypes in visual environments.

The aim of this dissertation is both the presentation and analysis of technology used in autonomous robotic vehicles, and secondly, the implementation of an advanced *Deep Learning* algorithm that is able to detect - through appropriate optical sensors used into an autonomous robotic vehicle - pedestrians within the analysis range of optical sensors.

The first chapters of this dissertation have been implemented at bibliographic level, and include topics that present the basic concepts, historical evolution elements and structure systems for an autonomous robotic vehicle. Thereinafter, the following chapters are implemented through a combination of bibliographic and example-based content, that contributes to the further understanding of the presented concepts.

In conclusion, the last part of this dissertation presents the development of a pedestrian detection *Deep Learning* algorithm, the results produced during algorithm testing procedure and also the proposal of possible further improvement points of this algorithm, based on the modern *Artificial Intelligence* tools.

Περιεχόμενα

ΕΙΣΑΓΩΓΗ	1
1. Η Έννοια του Αυτόνομου Οχήματος	3
1.1 Τι Είναι το Αυτόνομο Όχημα;	3
1.2 Ο Βαθμός Αυτονομίας των Αυτόνομων Οχημάτων	3
1.3 Τα Πλεονεκτήματα Χρήσης των Αυτόνομων Οχημάτων	6
1.3.1 Ασφάλεια	6
1.3.2 Οικονομικά και κοινωνικά οφέλη	6
1.3.3 Αποτελεσματικότητα και συμβολή	6
1.3.4 Βελτίωση καθημερινών μετακινήσεων	6
1.4 Η Ισχύων Νομικό Καθεστώς των Αυτόνομων Οχημάτων	7
1.4.1 Η ανάγκη της νομικής αδειοδότησης κυκλοφορίας	7
1.4.2 Το ισχύων νομικό καθεστώς στις Η.Π.Α.	7
1.5 Πεδία Εφαρμογής των Αυτόνομων Οχημάτων	10
1.5.1 Τα υφιστάμενα πεδία εφαρμογής των αυτόνομων οχημάτων	10
1.5.2 Τα πιθανά πεδία μελλοντικής εφαρμογής αυτόνομων οχημάτων	11
2 Η Ιστορική Εξέλιξη των Αυτόνομων Αυτοκινήτων	13
2.1 Σύνομη Αναδρομή στην Ιστορία των Αυτόνομων Αυτοκινήτων	13
2.2 Τα Πρώτα Οράματα για μία Νέα Εποχή στην Αυτοκίνηση	13
2.3 Τα Πρώτα Οχήματα Αυτόνομων Λειτουργιών	16
2.4 Οι Αρχικές Προσεγγίσεις των Σύγχρονων Αυτόνομων Οχημάτων	17
2.5 Τα Αυτόνομα Αυτοκίνητα Σήμερα	19
3 Η Δομή και οι Λειτουργίες των Αυτόνομων Οχημάτων	25
3.1 Η Δομή των Αυτόνομων Οχημάτων	25
3.1.1 Το Δίκτυο των Αισθητηρίων Διατάξεων	26
3.1.2 Η μονάδα ελέγχου του αυτόνομου οχήματος	32
3.2 Οι Εφαρμοσμένες Τεχνολογίες των Αυτόνομων Οχημάτων	36
3.2.1 Οι Λειτουργίες πλοήγησης του αυτόνομου οχήματος	36
3.3 Οι Τεχνολογίες Επικοινωνιών των Αυτόνομων Οχημάτων	40
3.3.1 Η ανάγκη εφαρμογής συστημάτων επικοινωνιών στα αυτόνομα οχήματα	40
3.3.2 Οι τεχνολογίες επικοινωνίας μεταξύ οχημάτων και υποδομών	40
3.3.3 Η τεχνολογία V2V	40
3.3.4 Η τεχνολογία V2I	41
3.3.5 Η έννοια του V2X	42
4 Η Τεχνητή Νοημοσύνη και η Μηχανική Μάθηση	43
4.1 Εισαγωγή στη Τεχνητή Νοημοσύνη	43
4.1.1 Βασικοί ορισμοί για την Τεχνητή Νοημοσύνη	43
4.1.2 Η έννοια της νοημοσύνης	44
4.1.3 Η έννοια της Τεχνητής Νοημοσύνης	45
4.2 Μηχανική Μάθηση	46
4.2.1 Ορισμός μάθησης και Μηχανικής Μάθησης	46
4.2.2 Τα είδη της μάθησης	47
4.2.3 Οι κατηγορίες της Μηχανικής Μάθησης	47

4.2.4	Η Επιβλεπόμενη Μάθηση.....	47
4.2.5	Η μη Επιβλεπόμενη Μάθηση.....	56
4.2.6	Άλλες μορφές μάθησης.....	59
5	Εισαγωγή στις Τεχνικές Βαθιάς Μάθησης	63
5.1	Ανίχνευση Προτύπου Εικόνας.....	63
5.2	Tunnel Vision.....	65
5.2.1	Αναζήτηση τη μέθοδο συρόμενου πλαισίου.....	66
5.2.2	Χειρισμός περισσότερων δεδομένων εισόδου.....	66
5.3	Συνελκτικά Νευρωνικά Δίκτυα.....	68
5.3.1	Πώς λειτουργούν τα Συνελκτικά Νευρωνικά Δίκτυα.....	69
5.4	Σχεδίαση Νευρωνικού Δικτύου Εντοπισμού Πτηνών.....	72
5.4.1	Υλοποίηση Νευρωνικού Δικτύου αναγνώρισης πτηνών.....	73
5.4.2	Έλεγχος του δικτύου.....	75
5.4.3	Αξιολόγηση της ακρίβειας.....	75
5.5	Εισαγωγή στις Σύγχρονες Τεχνικές Υπολογιστικής Όρασης.....	77
5.6	R-CNN.....	78
5.6.1	Προβλήματα του R-CNN.....	80
5.7	Fast R-CNN.....	81
5.7.1	ROI Pooling.....	81
5.7.2	Συνδυασμός όλων των μοντέλων σε ένα ενιαίο δίκτυο.....	82
5.8	Faster R-CNN.....	83
5.9	Mask R-CNN.....	86
5.9.1	RoiAlign.....	87
5.10	YOLO.....	89
5.11	Άλλες Τεχνικές Μηχανικής Όρασης.....	89
6	Εφαρμογή Τεχνικών Βαθιάς Μάθησης στα Αυτόνομα Ρομποτικά Οχήματα	91
6.1	Ο Αλγόριθμος Μάθησης BP για εκπαίδευση MLPs και RNNs.....	92
6.2	Εφαρμογή Βαθιάς Μάθησης σε MLPs και RNNs για την Ενίσχυση της Μάθησης.....	94
6.2.1	Υλοποίηση RNNs με Deep CAPs μέσω εφαρμογής μεθόδου RL σε NNS.....	94
6.2.2	Deep MLPs σε διαδικασίες λήψης αποφάσεων RL και Markov.....	96
6.3	Υλοποίηση της Διαδικασίας Βαθιάς Εκπαίδευσης σε MLPs στην πλατφόρμα MATLAB.....	98
6.3.1	Αρχιτεκτονική των CNN.....	98
6.3.2	Παράδειγμα 1: αναγνώριση αντικειμένων με εφαρμογή Βαθιάς Μάθησης.....	99
6.3.3	Παράδειγμα 2: αναγνώριση αντικειμένου με εφαρμογή Βαθιάς Μάθησης και Faster R-CNN.....	109
7	Εφαρμογή Τεχνικών Βαθιάς Μάθησης για τον Εντοπισμό Πεζών	117
7.1	Σχετικές Εργασίες.....	119
7.2	Μεθοδολογία.....	121
7.3	Υλοποίηση Αλγόριθμου Εντοπισμού Πεζών.....	124
7.4	Αποτελέσματα.....	133
	Προτάσεις Περαιτέρω Βελτίωσης.....	137
	Επίλογος - Συμπεράσματα.....	139
	Βιβλιογραφία – Αναφορές.....	141
	Ευρετήριο Αρκτηκόλεξων.....	165

ΕΙΣΑΓΩΓΗ

Στις μέρες μας, η έκρυθμη και γεωμετρική πρόοδος που σημειώνει καθημερινά η τεχνολογία, οδηγεί στη δημιουργία ολοένα και περισσότερων και αποτελεσματικότερων λύσεων σε μία πληθώρα προβλημάτων που αντιμετωπίζουν οι σύγχρονοι πολιτισμοί εδώ και πολλές δεκαετίες. Η ραγδαία εξέλιξη που σημειώνουν τομείς όπως αυτοί της πληροφορικής, της μηχανικής, της μικροηλεκτρονικής, της βιοιατρικής, της διαστημικής και άλλων πολλών, οδηγεί με μαθηματική ακρίβεια στην εξάλειψη όλων βασικών προβλημάτων που ταλανίζουν τις σύγχρονες κοινωνίες. Τέτοια προβλήματα αντιμετωπίζουν οι τομείς της υγείας, των μεταφορών, των επικοινωνιών, των υποδομών, της μηχανικής, της οικονομίας κτλ. Έτσι, καθίσταται βαθμιαία δυνατός ο καθολικός αυτοματισμός υπηρεσιών με εφαρμογή σύγχρονων μεθόδων και εργαλείων, αναβαθμίζοντας έτσι την καθημερινότητα των σύγχρονων πολιτών. Η συμβολή της επιστήμης της πληροφορικής στην οικοδόμηση σύγχρονων λύσεων που αντιμετωπίζουν αποτελεσματικά τα υφιστάμενα προβλήματα, είναι μείζονος σημασίας καθώς, δύναται να αποτελέσει για τις υπόλοιπες επιστήμες ένα υψίστης σημασίας εργαλείο που αντιμετωπίζει καθολικά τις όποιες προκλήσεις στον όγκο, την ταχύτητα και την ασφάλεια στην επεξεργασία δεδομένων.

Ωστόσο, πολλά από τα προβλήματα που καλείται να λύσει ο σύγχρονος πολιτισμός, χαρακτηρίζονται από πολυμεταβλητότητα, μη γραμμική συμπεριφορά και ασαφή λογική, που καθιστούν την μοντελοποίηση και παραγωγή πρότυπων λύσεων τους, ανέφικτη για ένα υπολογιστικό σύστημα. Παράλληλα, τέτοιας αρχιτεκτονικής προβλήματα μπορούν με απόλυτη επιτυχία να αντιμετωπίζονται από την ανθρώπινη συλλογιστική σκέψη, η οποία με τον ιδιαίτερο τρόπο λειτουργίας της, είναι σε θέση να αντιμετωπίσει επιτυχώς προβλήματα τα οποία μπορούν να είναι πρωτότυπα και ιδιαίτερα σύνθετα. Το συγκεκριμένο κενό στην ψηφιακή αντιμετώπιση αντίστοιχων προβλημάτων και προκλήσεων, έρχεται να αντιμετωπίσει η επιστήμη της *Τεχνητής Νοημοσύνης*, η οποία καλείται να μοντελοποιήσει την δομή και την συλλογιστική μέθοδο του ανθρώπινου εγκεφάλου, και να εφαρμόσει αυτές τις αρχές ως αλγοριθμικό μοτίβο στα υπολογιστικά συστήματα. Η επιστήμη της *Τεχνητής Νοημοσύνης*, έχει ήδη οικοδομήσει ιδιαίτερα αξιόλογες λύσεις σε μία πληθώρα μη γραμμικών και πολυμεταβλητών προβλημάτων, κυρίως στους τομείς της ιατρικής, της οικονομίας, των μεταφορών, τις υπηρεσίες κοινωνικής εξυπηρέτησης κτλ.

Ένα από τα πεδία των οποίων η πλήρης αυτοματοποίηση κρίνεται επιτακτικής ανάγκης, είναι αυτό των μετακινήσεων και μεταφορών. Αποτελεί πλέον γεγονός ότι ένα από τα αίτια θνησιμότητας που φέρει ύψιστο ποσοστό παγκοσμίως, είναι αυτό των τροχαίων δυστυχημάτων, ενώ παράλληλα εξίσου υψηλά ποσοστά φέρουν και τα μη θανατηφόρα και σοβαρά ατυχήματα που προκαλούν μόνιμες παραλύσεις και λοιπές αναπηρίες. Παράλληλα, η αφενός αλόγιστη και αφετέρου λαθεμένη χρήση των μεταφορικών οχημάτων και ιδίως των αυτοκινήτων, επιφέρει μία σωρεία λοιπών σημαντικών κοινωνικών επιπτώσεων όπως η κυκλοφοριακή συμφόρηση, η μόλυνση του περιβάλλοντος, η αλόγιστη σπατάλη ενέργειας και φυσικών πόρων, οι κωλυσιεργίες σε λοιπές καθημερινές δραστηριότητες, η επιβάρυνση υγείας πολιτών λόγω αυξημένου επιπέδου ρύπων αλλά και οικονομικές απώλειες που επιφέρει με τη σειρά της η αλόγιστη σπατάλη ενέργειας και φυσικών πόρων. Είναι επομένως κοινώς αποδεκτό ότι το μείζον αυτό θέμα των μετακινήσεων και μεταφορών εν γένει, θα πρέπει άμεσα να καταστεί αντιμετώπισιμο από την ανθρωπότητα και ιδίως από την επιστημονική κοινότητα.

Στο πλαίσιο αυτής της φιλοσοφίας, εδώ και αρκετές δεκαετίες, η επιστημονική κοινότητα άρχισε να οραματίζεται την οικοδόμηση σύγχρονων αστικών και μη κέντρων, τα οποία θα χρησιμοποιούσαν για την εύρυθμη μετακίνηση των πολιτών ρομποτικά οχήματα ή αλλιώς τα αυτόνομα αυτοκίνητα. Τα οχήματα αυτό δεν θα διέφεραν σε τίποτα ουσιαστικό από τα συμβατικά οχήματα και αυτοκίνητα, πέραν του γεγονότος ότι η πλοήγησή τους δεν θα ήταν αποτέλεσμα του αντίστοιχου ανθρώπινου χειρισμού, άλλα αποτέλεσμα της αυτόνομης λειτουργίας τους η οποία θα δομείτο από προηγμένα αισθητήρια συστήματα και αντίστοιχο λογισμικό πλοήγησης που θα εκτελούσε η μονάδα μικροϋπολογιστή η οποία θα έλεγχε και στο σύστημα μηχανικής πλοήγησης του οχήματος. Με την ύπαρξη ενός τέτοιου οχήματος, η λειτουργία της οδήγησης θα πραγματοποιείτο αυτόνομα, με τον βέλτιστο δυνατό τρόπο και βεβαίως με την απουσία του ενδεχομένου των λανθασμένων ανθρώπινων χειρισμών. Έτσι, όλα τα προβλήματα που παρουσιάστηκαν στην προηγούμενη παράγραφο, θα μπορούσαν να εξαλειφθούν.

Αναγνώριση πεζών από αυτόνομο ρομποτικό όχημα με χρήση τεχνητής νοημοσύνης και τεχνικών βαθιάς μάθησης

Ως αποτέλεσμα αυτού του οράματος, οι διεθνής επιστημονικές ομάδες, προσπάθησαν να μοντελοποιήσουν σε αλγόριθμο την ανθρώπινη συλλογιστική του εγκέφαλου ενός οδηγού. Έτσι, η πολυμεταβλητότητα του συγκεκριμένου προβλήματος, η ανάγκη της ύπαρξης Μηχανικής Όρασης, ο χειρισμός των αβέβαιων ή αγνώστων καταστάσεων, η ασαφής – εμπειρική διαδικασία χειρισμού που ασκεί ο ανθρώπινος εγκέφαλος πολλές φορές και άλλες λοιπές ιδιαιτερότητες, οδήγησαν στην ανάγκη εφαρμογής των μεθόδων *Τεχνητής Νοημοσύνης* για την αλγοριθμική προσέγγιση του συγκεκριμένου προβλήματος. Έτσι, μετά την εφαρμογή των μεθόδων *Τεχνητής Νοημοσύνης* κατέστησαν πραγματικότητα πολλά αυτόνομα οχήματα, μικρού ωστόσο βαθμού αυτονομίας, τα οποία μπορούσαν παρά τα όποια προβλήματα που παρουσίαζαν, να πλοηγηθούν επιτυχώς εντός ορισμένης διαδρομής, ικανοποιώντας σε μικρό έως ικανοποιητικό βαθμό, τη μεγαλύτερη πρόκληση που είχαν να αντιμετωπίσουν, αυτή της Μηχανικής Όρασης και της αναγνώρισης προτύπων.

Είναι ευνόητο, ότι η σπουδαιότερη λειτουργία του ανθρώπινου εγκέφαλου κατά την διαδικασία της οδήγησης ενός οχήματος, δεν είναι αυτή της πλοήγησης του αλλά αυτή της όρασης. Βάσει των εικόνων που λαμβάνει ο εγκέφαλος διαμέσου της οράσεως, πραγματοποιεί ταυτοποίηση όλων των περιβαλλόντων, οντοτήτων και αντικειμένων, συνεκτιμά τις όποιες πιθανότητες άμεσων ή έμμεσων κινδύνων και λαμβάνει τις κατάλληλες αποφάσεις για την πλοήγηση του οχήματος. Η παραπάνω λειτουργία είναι εξαιρετικά δύσκολο να μοντελοποιηθεί σε κάποιο υπολογιστικό σύστημα και απαιτείται η εφαρμογή και αξιοποίηση παράλληλων συστημάτων επεξεργασίας καθώς και τεχνικών *Βαθιάς Μάθησης* οι οποίες αποτελούν την πλέον αποδοτικότερη μέθοδο της *Τεχνητής Νοημοσύνης* για την μοντελοποίηση της ανθρώπινης συλλογιστικής πάνω σε οπτικά δεδομένα όπως είναι οι εικόνες. Η σύγχρονη αντιμετώπιση του προβλήματος αυτού από την επιστημονική κοινότητα περιλαμβάνει την εφαρμογή εξελιγμένων *Βαθιών Συνελκτικών Νευρωνικών Δικτύων* καθώς και εκπαίδευση αυτών μέσα από συγκεκριμένες διαδικασίες ώστε να καταστεί δυνατή η βέλτιστη και αλάνθαστη αναγνώριση των προβαλλόμενων περιβαλλόντων, οντοτήτων και αντικειμένων από το σύστημα, σε πραγματικό χρόνο.

Οι διαδικασίες αναγνώρισης όπως αναφέρθηκαν άνω, αποτελούν τον ακρογωνιαίο λίθο για την αυτόνομη λειτουργία ενός ρομποτικού οχήματος. Τα αντικείμενα που καλείται να αναγνωρίσει και να αξιολογήσει ένα αυτόνομο όχημα δεν περιλαμβάνουν οτιδήποτε βρίσκεται στο οπτικό πεδίο του αυτοκινήτου, αλλά κυρίως αυτά που κατέχουν ζωτική σημασία για την αυτόνομη πλοήγηση του. Για παράδειγμα, ίσως να μην είναι πολύ σημαντικό για ένα αυτόνομο όχημα να αναγνωρίσει τα σύννεφα του ουρανού, τα πουλιά ή τις στέγες των κτηρίων. Είναι όμως πολύ σημαντικό να είναι σε θέση να αναγνωρίσει τους πεζούς, τα άλλα οχήματα, τους φωτεινούς σηματοδότες, τις οδικές σημάσεις, τα όρια του δρόμου και τις λωρίδες του, τα όποια οδικά εμπόδια κτλ. Κάθε ένα αντικείμενο ή οντότητα από τα παραπάνω κατέχει διαφορετικές ιδιότητες και χαρακτηριστικά, και αποτελεί διαφορετικό στοιχείο για το αυτόνομο όχημα, το οποίο και θα πρέπει να αξιολογηθεί αντιμετωπιστεί διαφορετικά από αυτό.

Η παρούσα εργασία επιχειρεί την προσέγγιση του ζητήματος της αναγνώρισης ενός πεζού, με χρήση ενός σύγχρονου, εξελιγμένου και αποδοτικού αλγορίθμου *Τεχνητής Νοημοσύνης*, που εφαρμόζει μερικές μεθόδους *Βαθιάς Μάθησης* πάνω σε *Προεκπαιδευμένα Βαθιά Συνελκτικά Νευρωνικά Δίκτυα* και ειδικότερα την πλέον αποδοτικότερη μέθοδο των *Faster R-CNN* που επιτυγχάνει ταχύτατη αναγνώριση του προτύπου εκπαίδευσης σε πραγματικό χρόνο. Παράλληλα, αξιοποιούνται και τεχνικές ανάλυσης και πρόβλεψης της κίνησης του προτύπου εκπαίδευσης – πεζού, ενώ στο τέλος πραγματοποιείται εργαστηριακή δοκιμή των εν λόγω αλγορίθμου, αξιολογείται η αποτελεσματικότητά του, εντοπίζονται τα πιθανά περιθώρια βελτίωσης του και προτείνονται οι αντίστοιχες θεματικές περιοχές περεταίρω έρευνας και βελτίωσης του.

Εν κατακλείδι, τα αυτόνομα ρομποτικά οχήματα ή κοινώς αποκαλούμενα αυτόνομα αυτοκίνητα, αποτελούν μία πολλά υποσχόμενη τεχνολογική εφαρμογή της *Τεχνητής Νοημοσύνης* η οποία προσδοκείται να αλλάξει ριζικά και οριστικά τον τομέα των μετακινήσεων και μεταφορών και να εξαλείψει δεκάδες προβλήματα των σύγχρονων κοινωνιών που παρέμεναν άλυτα για δεκαετίες. Ωστόσο, η ανάπτυξη και εφαρμογή των βέλτιστων τεχνολογιών και μεθόδων *Τεχνητής Νοημοσύνης* είναι η μόνη πραγματικά επιτακτική ανάγκη που θα οδηγήσει με την σειρά της στην έμμεση επίλυση και λοιπών προβλημάτων ασαφούς και πολυμεταβλητής αρχιτεκτονικής, όλων των επιστημονικών πεδίων.

1. Η Έννοια του Αυτόνομου Οχήματος

1.1 Τι Είναι το Αυτόνομο Όχημα;

Σε μία εποχή όπου η τεχνολογική επανάσταση κινείται με εκθετικούς ρυθμούς, εφαρμογή των σύγχρονων τεχνολογικών εργαλείων σε κάθε πτυχή της καθημερινότητας είναι ήδη προδιαγεγραμμένη. Έτσι, η εκσυγχρόνιση της σύγχρονης ζωής θα πρέπει να έχει σαν βάση πρωτίστως την βελτιστοποίηση του βιοτικού επιπέδου και την εκμηδένιση κάθε παράγοντα που αποτελεί κίνδυνο για την ανθρώπινη ύπαρξη.

Κινούμενοι σε μία τέτοια φιλοσοφία, θα ήταν αναπόφευκτη η εφαρμογή των σύγχρονων τεχνολογικών εργαλείων στον τομέα της μετακίνησης των ανθρώπων μιας και αυτός αποτελεί μία από τις βασικότερες πτυχές της καθημερινότητας του σύγχρονου ανθρώπου, η οποία όμως ενέχει μεγάλο βαθμό επικινδυνότητας λόγω των κατά κύριο λόγο λαθεμένων ανθρώπινων αντιδράσεων και χειρισμών. Έτσι, η ανάπτυξη των αυτόνομων οχημάτων αποτελεί μία τεχνολογική επανάσταση η οποία όχι μόνο θα βελτιώσει την ανθρώπινη καθημερινότητα σε πολλές από τις πτυχές της, αλλά κυρίως θα εξασφαλίσει την ασφάλεια της ανθρώπινης ζωής.

Τα αυτόνομα οχήματα έχουν ιστορία πολύ παλαιότερη από όσο πιστεύουμε. Από τις αρχές κιόλας της δεκαετίας του 1920 αρκετοί επιστήμονες της εποχής είχαν αρχίσει να οραματίζονται και να σχεδιάζουν το όχημα του μέλλοντος που θα μπορούσε να μετακινηθεί σχεδόν οπουδήποτε χωρίς την ανάγκη ανθρώπινου χειρισμού. Οι όροι «Αυτόνομο όχημα» και «Αυτοκινούμενο όχημα», συνηθίζεται να χρησιμοποιούνται εναλλακτικά για να περιγράψουν ένα ειδικά σχεδιασμένο όχημα το οποίο χρησιμοποιεί μία τεχνολογία η οποία του επιτρέπει να μπορεί να κινηθεί μεταξύ δύο σημείων με απόλυτη ασφάλεια, χρησιμοποιώντας στο έπακρο τα συστήματα επιτάχυνσης, πέδησης, αλλαγής πορείας κτλ, χωρίς ανάγκη ανθρώπινη παρέμβαση αφού θα εκτελεί αυτόνομα όλες τις απαιτούμενες λειτουργίες. Αρκετά υφιστάμενα οχήματα είναι σε θέση να πάρουν εντελώς τον έλεγχο πλοήγησης εκτελώντας διάφορες λειτουργίες όπως το *Cruise Control* το οποίο αναλαμβάνει την ρύθμιση της ταχύτητας. Μερικά από τα μελλοντικά σχέδια για τις λειτουργίες των αυτόνομων οχημάτων περιλαμβάνουν πέραν των άλλων, αυτόματη απάντηση του αυτοκινήτου σε εισερχόμενο γραπτό μήνυμα, αυτόματη εκκίνηση κινητήρα, αυτόματη μη επανδρωμένη πλοήγηση του οχήματος προς τον κάτοχο του κτλ.

Στην ουσία πρόκειται για ένα ρομποτικό όχημα, εξοπλισμένο με πληθώρα αισθητήρων Μηχανικής Όρασης, αισθητήρων μέτρησης λοιπών φυσικών μεγεθών, συστημάτων προσανατολισμού και πλοήγησης, λογισμικών Τεχνητής Νοημοσύνης και Μηχανικής Μάθησης κ.α. Αποσκοπεί στην πλήρη και αδιάληπτη αυτόνομη πλοήγηση του ίδιου οχήματος όχι μόνο στον βαθμό αποτελεσματικότητας του φυσικού οδηγού αλλά και σε βαθμό πολύ αποτελεσματικότερο από αυτόν.

Για τον λόγο αυτό, εκατοντάδες κατηγορίες αλγορίθμων Τεχνητής Νοημοσύνης έχουν αναπτυχθεί, αποσκοπώντας στην όσο το δυνατόν βέλτιστη πρόβλεψη πιθανών κυκλοφοριακών συμβάντων και την άμεση αντιμετώπιση τους. [4], [10], [17], [25], [54], [56], [76], [203]

1.2 Ο Βαθμός Αυτονομίας των Αυτόνομων Οχημάτων

Η αυτοματοποίηση της λειτουργίας ενός οχήματος, περιλαμβάνει τον συντονισμό και την αυτοματοποίηση δεκάδων ανεξάρτητων λειτουργιών του. Έτσι, όσο περισσότερες λειτουργίες ενός οχήματος είναι αυτοματοποιημένες και πλήρως ανεξάρτητες από τον οδηγό, τόσο και μεγαλύτερο βαθμό αυτονομίας φέρει το όχημα αυτό.

Σύμφωνα με την υπηρεσία *NHTSA* των Η.Π.Α. και το *SAE*, ο βαθμός αυτονομίας ελέγχου ενός αυτοκινήτου κατηγοριοποιείται σε έξι διαφορετικά επίπεδα:

Επίπεδο 0:**Απουσία Αυτοματισμού Λειτουργιών
(No Automation)**

Ο οδηγός του οχήματος είναι ο μοναδικός παράγοντας ελέγχου και πλοήγησης του. Ελέγχει διαρκώς όλα τα συστήματα πλοήγησης όπως σύστημα το πέδησης, το σύστημα κατεύθυνσης, το σύστημα μετάδοσης, το σύστημα κίνησης κτλ καθ' όλη τη διάρκεια της διαδρομής. Στο επίπεδο αυτό κανένα σύστημα αυτοματισμού ή / και αυτονομίας δεν εφαρμόζεται και η όποια αυτονομία του οχήματος περιορίζεται αποκλειστικά στην παραγωγή ειδοποιήσεων και ενδείξεων.

Επίπεδο 1^ο:**Αυτοματισμός Συγκεκριμένων Λειτουργιών
(Driver Assisted / Function Specific Automation)**

Σε αυτό το επίπεδο ο έλεγχος του οχήματος βρίσκεται στον οδηγό. Ωστόσο υπάρχει και η παρουσία ενός προηγμένου συστήματος υποβοήθησης του οδηγού (ADAS) το οποίο περιλαμβάνει διάφορες λειτουργίες όπως όπως το ACC (*Adaptive Cruise Control*), η υποβοήθηση στο παρκάρισμα με αυτόματη κίνηση του τιμονιού, το LKA (*Lane Keeping Assistance*) Type II κτλ. Οι λειτουργίες αυτές αποσκοπούν στην διευκόλυνση του οδηγού, ενώ δεν υπάρχει η δυνατότητα εφαρμογής τους συνδυαστικά.

Επίπεδο 2^ο:**Αυτοματισμός Συνδυαζόμενων Λειτουργιών
(Partial Automation / Combined Autonomous Functions)**

Στο επίπεδο αυτό ο έλεγχος του οχήματος παραμένει στον οδηγό ενώ επίσης υπάρχει η παρουσία του προηγμένου συστήματος υποβοήθησης του οδηγού (ADAS). Ωστόσο, στο επίπεδο αυτό δίνεται η δυνατότητα ταυτόχρονης αξιοποίησης των λειτουργιών ελέγχου του τιμονιού και του συστήματος πέδησης / επιτάχυνσης κάτω από συγκεκριμένες συνθήκες. Οι παραπάνω λειτουργίες επιτελούνται αποκλειστικά υποβοηθητικά και η ευθύνη ασφαλούς ελέγχου του οχήματος παραμένει στον οδηγό.

Επίπεδο 3^ο:**Περιορισμένος Αυτοματισμός Λειτουργίας Αυτό-Οδήγησης
(Conditional Automation / Combined Autonomous Functions)**

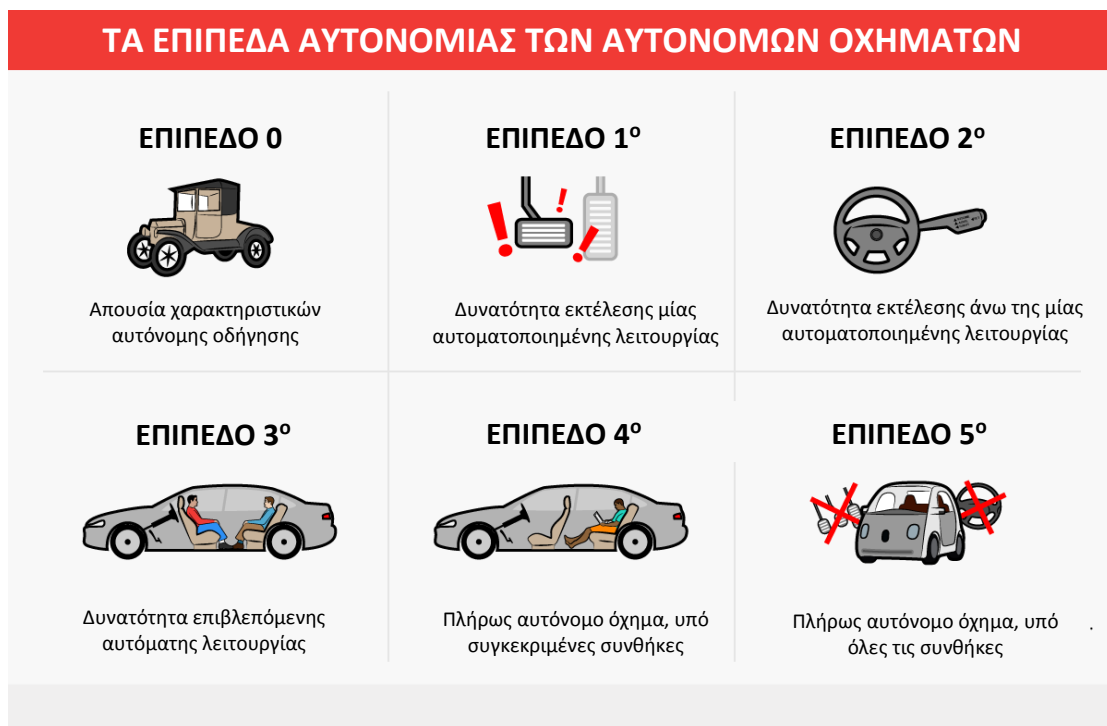
Τα οχήματα που χαρακτηρίζονται σε αυτό το επίπεδο αυτοματοποίησης, διαθέτουν ένα προηγμένο σύστημα οδήγησης (ADS) το οποίο κάτω από συγκεκριμένες και περιορισμένες συνθήκες, είναι σε θέση να επιτελέσει το σύνολο των λειτουργιών και διαδικασιών που απαιτούνται για την αυτόνομη οδήγηση του οχήματος από το σημείο εκκίνησης προς το σημείο προορισμού. Κατά τη διάρκεια της διαδρομής αυτής, ο οδηγός θα πρέπει να βρίσκεται σε πλήρη ετοιμότητα να αναλάβει τον έλεγχο του οχήματος μόλις το σύστημα ADS το απαιτήσει. Σε κάθε περίπτωση όπου οι επικρατούσες συνθήκες δεν επιτρέπουν την χρήση του ADS, ο οδηγός του οχήματος πραγματοποιεί την πλήρη λειτουργία της οδήγησης του.

Επίπεδο 4^ο:**Υψηλό Επίπεδο Αυτοματοποίησης Λειτουργίας Αυτό-οδήγησης
(High Automation)**

Σε αυτό το επίπεδο, το όχημα είναι σχεδιασμένο ώστε να παρέχει υψηλού επιπέδου αυτοματοποίηση των λειτουργιών του. Το σύστημα ADS είναι σε θέση να πραγματοποιήσει όλες τις λειτουργίες οδήγησης αλλά και να παρακολουθεί όλες τις κυκλοφοριακές συνθήκες του περιβάλλοντος στο οποίο κινείται. Στην ουσία πραγματοποιεί μία πλήρη λειτουργία οδήγησης, ομοίως με κάποιον οδηγό. Το σύστημα είναι σε θέση να ανταποκριθεί άριστα κάτω από όλες σχεδόν τις συνθήκες, με ελάχιστες εξαιρέσεις (π.χ. ακραία καιρικά φαινόμενα). Έτσι ο επιβάτης/οδηγός θα πρέπει να διασφαλίσει ότι πληρούνται όλες οι αναγκαίες συνθήκες και προϋποθέσεις προτού θέσει το σύστημα σε λειτουργία.

Επίπεδο 5°:**Πλήρης Αυτοματισμός Λειτουργίας Αυτό-Οδήγησης
(Full Automation)**

Σε αυτό το επίπεδο, το σύστημα *ADS* του οχήματος είναι σε θέση να πραγματοποιήσει όλες τις λειτουργίες οδήγησης αλλά και να παρακολουθεί όλες τις κυκλοφοριακές συνθήκες του περιβάλλοντος στο οποίο κινείται κάτω από οποιοδήποτε συνθήκες και προϋποθέσεις αντικαθιστώντας απόλυτα τον οδηγό. Το επίπεδο ασφαλείας και αποτελεσματικότητας του επιπέδου αυτού είναι το μέγιστο δυνατό και ο επιβάτης του οχήματος δεν απαιτείται να έχει την οποιαδήποτε συνδρομή στη διαδικασία της οδήγησης πέραν από τον ορισμό του επιθυμητού προορισμού. [16], [17], [18], [19], [193]

**Εικόνα 1.1**

Τα επίπεδα αυτονομίας των αυτοκινούμενων οχημάτων.

1.3 Τα Πλεονεκτήματα Χρήσης των Αυτόνομων Οχημάτων

1.3.1 Ασφάλεια

Αν αναλογιστούμε το γεγονός ότι το 94% των σοβαρών κυκλοφοριακών ατυχημάτων οφείλεται σε ανθρώπινο λάθος, καταλαβαίνουμε ότι τα οφέλη που πηγάζουν από την ιδιαίτερα αυξημένη ασφάλεια που παρέχουν τα αυτοματοποιημένα οχήματα, είναι υψίστης σημασίας. Μόνο το 2015, 35.092 άτομα στις Η.Π.Α. έχασαν τη ζωή τους σε συντριβές μηχανοκίνητων οχημάτων. Τα αυτοματοποιημένα οχήματα δίνουν τη δυνατότητα εξάλειψης του οποιουδήποτε ανθρώπινου λάθους θα οδηγούσε ενδεχομένως σε κάποια σύγκρουση, σώζοντας έτσι ανθρώπινες ζωές εκατομμυρίων οδηγών, επιβατών και πεζών.

[1], [10], [13], [78], [79]

1.3.2 Οικονομικά και κοινωνικά οφέλη

Η εκτεταμένη χρήση αυτοματοποιημένων οχημάτων δύναται να αποφέρει επιπρόσθετα οικονομικά και πρόσθετα κοινωνικά οφέλη στις σύγχρονες κοινωνίες. Ο *NHTSA* δημοσίευσε μία μελέτη από την οποία όπως προκύπτει οι καταγεγραμμένες συντριβές των μηχανοκίνητων οχημάτων για το 2010, ανήλθαν στο ύψος των 242 δισεκατομμυρίων δολαρίων, συμπεριλαμβανομένων 57,6 δισεκατομμυρίων δολαρίων που αφορούσαν απώλεια παραγωγικότητας στο χώρο εργασίας και 594 δισεκατομμύρια δολάρια που αφορούσαν απώλεια ζωής και μειωμένης ποιότητας ζωής λόγω τραυματισμών. Έτσι, η μαζική και αποκλειστική χρήση των αυτοματοποιημένων οχημάτων, θα ήταν δυνατό να εξαλείψει τα κόστη αυτά και να οδηγήσει σε προαγωγή λοιπών κοινωνικών αξιών. [9], [18], [19], [27], [39], [50], [53], [78]

1.3.3 Αποτελεσματικότητα και συμβολή

Από την καθολική και αποκλειστική χρήση αυτοματοποιημένων οχημάτων, θα ήταν δυνατή η ανάπτυξη μοντέλου επικοινωνίας και συνεργασίας τους με σκοπό την ομαλή ροή της κυκλοφορίας και τη κατακόρυφη μείωση της κυκλοφοριακής συμφόρησης. Όπως προκύπτει από τα επίσημα καταγεγραμμένα στοιχεία του έτους 2014, οι Αμερικανοί πολίτες δαπάνησαν περισσότερο από 6,9 δισεκατομμύρια ώρες σε καθυστερήσεις λόγω κυκλοφοριακής συμφόρησης, κάτι που οδηγεί αυτόματα σε μείωση του διαθέσιμου χρόνου εργασίας και προσωπικού χρόνου ενώ παράλληλα αυξάνεται το κόστος των καυσίμων, το κόστος συντήρησης οχημάτων και οι εκπομπές οχημάτων. Με τη καθολική χρήση των αυτοματοποιημένων οχημάτων, ο χρόνος και το κόστος που απαιτούνται για την καθημερινή μετακίνηση θα μπορούσαν να μειωθούν σε μεγάλο βαθμό. Μια πρόσφατη μελέτη του *NHTSA* έδειξε πως τα αυτοματοποιημένα οχήματα θα μπορούσαν να εξοικονομήσουν έως και 50 λεπτά της ώρα καθημερινά για τον κάθε επιβάτη. [10], [14], [19], [50], [69], [73], [74], [189]

1.3.4 Βελτίωση καθημερινών μετακινήσεων

Αν και το σύνολο των οφελών για την κοινωνική ζωή, που απορρέουν από την χρήση των αυτόνομων οχημάτων, είναι κάτι σύνθετο που χρήζει ιδιαίτερης ανάλυσης, μία σημαντική κοινωνική πτυχή που ανοικοδομείται και βελτιώνεται δραστικά, αφορά την δυνατότητα αυτόνομης μετακίνησης κάθε κοινωνικής ομάδας. Όπως προκύπτει από την ανασκόπηση των αμερικανικών στατιστικών και δημογραφικών στοιχείων, σήμερα υπάρχουν 49 εκατομμύρια Αμερικανοί πολίτες ηλικίας άνω των 65 και 53 εκατομμύρια πολίτες που έχουν κάποια μορφή αναπηρίας. Έτσι, τα αυτόνομα οχήματα μπορούν να δώσουν την δυνατότητα σε εκατομμύρια πολιτών να μετακινηθούν πλέον ανεξάρτητα. Στις μέρες μας, σε πολλές περιπτώσεις η απασχόληση ή η ανεξάρτητη διαβίωση εξαρτάται άμεσα από την ικανότητα οδήγησης. Επομένως, τα αυτόνομα οχήματα θα μπορούσαν να εξασφαλίσουν άνετη και ομαλή μετακίνηση σε πολίτες που το έχουν ανάγκη και μέχρι πρότινος δεν είχαν την δυνατότητα.

[14], [15], [19], [83]

1.4 Η Ισχύων Νομικό Καθεστώς των Αυτόνομων Οχημάτων

1.4.1 Η ανάγκη της νομικής αδειοδότησης κυκλοφορίας

Τα συμβατικά οχήματα δημόσιας κυκλοφορίας διέπονται από τις αντίστοιχες νομικές ρυθμιστικές διατάξεις αναφορικά με την κυκλοφορία τους στους δημόσιους δρόμους και τους εκάστοτε περιορισμούς. Έτσι, είναι δυνατή η καθολική εφαρμογή ενός ενιαίου πλαισίου κανονισμών και νόμων ώστε να εξασφαλίζεται αφενός η μέγιστη ασφάλεια για τους πολίτες και αφετέρου ένας κοινός κώδικας συμπεριφοράς των οχημάτων αυτών.

Ωστόσο, μετά την πρώτη εμφάνιση των αυτόνομων οχημάτων και την αίτηση αδειοδότησης κυκλοφορίας τους στους δημόσιους δρόμους, υπήρξαν τα πρώτα νομικά κολλήματα. Τα αυτόνομα οχήματα, σύμφωνα με την φιλοσοφία τους δε φέρουν οδηγό με την έννοια του χειριστή, και συνεπώς όλοι αυτοί οι κανόνες όπου ισχύουν για τον οδηγό ενός συμβατικού οχήματος, εδώ δεν θα μπορούσαν να έχουν άμεση και αυτόματη εφαρμογή. Παράλληλα, με την είσοδο των αυτόνομων οχημάτων στους δημόσιους δρόμους κυκλοφορίας, έρχονται στην επιφάνεια διάφορα ηθικά ζητήματα φιλοσοφικού και νομικού περιεχομένου.

Γίνεται συνεπώς απαραίτητη η καθιέρωση μίας νομικής βάσης και ενός πλαισίων κανόνων λειτουργίας και οδικής συμπεριφοράς, αλλά και τεχνολογικών δικλείδων και προτύπων που θα πρέπει να ακολουθούν τα αυτόνομα οχήματα ώστε να καταστεί δυνατή η αδειοδότηση τους στους δημόσιους δρόμους κυκλοφορίας.

Πάνω σε αυτή τη βάση, έχουν διατυπωθεί τα πρώτα νομικά και λειτουργικά πλαίσια, α οποία προδιαγράφουν το σύνολο των προϋποθέσεων που θα πρέπει να πληροί ένα αυτόνομο όχημα ώστε να εκδοθεί η άδεια κυκλοφορίας του στους δημόσιους αυτοκινητόδρομους. Έτσι, αρκετές πολιτείες των Η.Π.Α. έχουν προχωρήσει στην δημοσίευση των αντίστοιχων ρυθμιστικών πλαισίων, ενώ άλλες δεν το έχουν πράξει ακόμα. Καθώς βρισκόμαστε στα πρώτα στάδια της εξέλιξης αυτής, δεν υφίσταται προς στιγμής ένα ενιαίο και δια-πολιτειακό νομικό πλαίσιο αδειοδότησης της κυκλοφορίας των αυτόνομων οχημάτων, και συνεπώς η κάθε πολιτεία αποφασίζει αυτόνομα για την αδειοδότηση ή μη της κυκλοφορίας των αυτόνομων οχημάτων. [10], [13], [18], [56]

1.4.2 Το ισχύων νομικό καθεστώς στις Η.Π.Α.

1.4.2.1 Η ισχύουσα νομική κατάσταση στην Washington

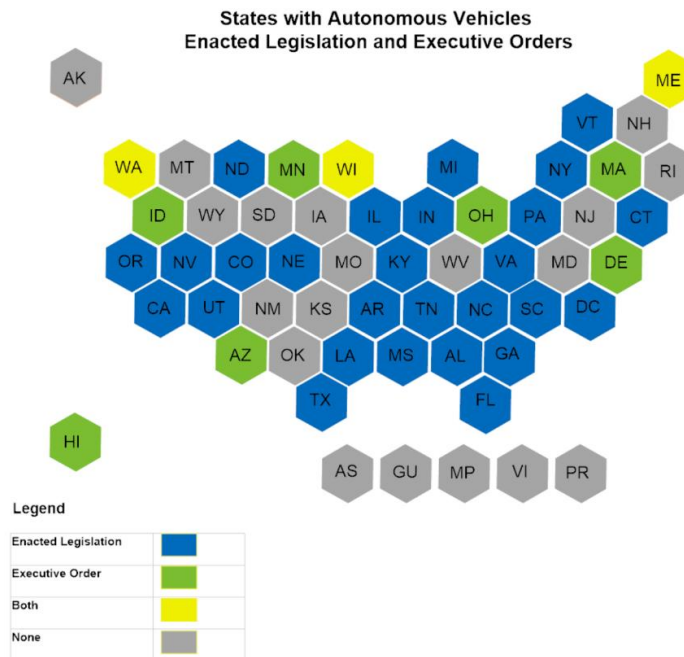
Τα αυτόνομα οχήματα, τα τελευταία 7 έτη έχουν αρχίσει να κυκλοφορούν στους δημόσιους αυτοκινητόδρομους των Η.Π.Α. ανάμεσα στα υπόλοιπα συμβατικά οχήματα. Δεδομένης της πλήρους ανεξαρτησίας και αυτονομίας της κάθε μίας πολιτείας από την άλλη, η αδειοδότηση κυκλοφορίας των αυτόνομων οχημάτων επιτελείται υπό διαφορετικό νομικό πλαίσιο και συνεπώς διαφορετικούς όρους και προϋποθέσεις σε κάθε πολιτεία.

Η Washington έχει ήδη προχωρήσει στα πρώτα νομικά βήματα για την νομιμοποίηση κυκλοφορίας των αυτόνομων οχημάτων. Στην τακτική συνεδρίαση του 2013, εισήχθη το πλαίσιο (κανονισμός) *HB1649*, το οποίο πλέον βρίσκεται υπό αναθεώρηση στο υπουργείο μεταφορών και συγκοινωνιών. Βάσει αυτού του πλαισίου, επιτρέπεται η δοκιμή των αυτόνομων οχημάτων στους δημόσιους δρόμους των Η.Π.Α. προτού καν ολοκληρωθεί η ανάπτυξη αντίστοιχων επίσημων προτύπων ασφαλείας. Ωστόσο, το πλαίσιο αυτό απαιτεί την ύπαρξη ενός οδηγού εντός του αυτόνομου οχήματος καθ' όλη τη διάρκεια της δοκιμαστικής φάσης και οι όποιες ποινικές ευθύνες προκύψουν από τυχούσα πρόκληση ζημιάς ή ατυχήματος, βαραινούν αποκλειστικά τον οδηγό του συγκεκριμένου οχήματος.

Είναι σημαντικό να αναφερθεί ότι η τρέχουσα έκδοση του πλαισίου *HB1649* καλύπτει μόνο τα οχήματα που διαθέτουν «αυτόνομη τεχνολογία» το οποίο σημαίνει ότι το όχημα θα είναι σε θέση να κινηθεί πλήρως αυτόνομα χωρίς την οποιαδήποτε ανθρώπινη παρέμβαση (επίπεδα αυτονομίας 4 και 5) και δεν περιλαμβάνει οχήματα των οποίων ο βαθμός αυτονομίας τους απαιτεί ανθρώπινή

παρέμβαση, καθώς για τις περιπτώσεις αυτές ισχύει ο συμβατικός Κ.Ο.Κ. Στη συνέχεια η Washington προχώρησε στην σχεδίαση και υλοποίηση ενός συνόλου προτύπων και κανόνων τα οποία θα πρέπει να πληρούν τα παραπάνω αυτόνομα οχήματα. [18] , [56]

1.4.2.2 Η παρούσα νομική κατάσταση σε ανά τις πολιτείες των Η.Π.Α.



Εικόνα 1.2

Η ισχύουσα νομική κατάσταση αδειοδότησης κυκλοφορίας των αυτόνομων οχημάτων, στις πολιτείες των Η.Π.Α.

[209]

1.4.2.3 Το γενικό νομικό πλαίσιο

Όλα τα νομικά πλαίσια που θεσπίστηκαν στις υπόλοιπες πολιτείες μετά από το *HB1649* της Ουάσιγκτον, επίσης αποκλείουν από την αδειοδότηση όλες τις υπόλοιπες βοηθητικές τεχνολογίες οι οποίες απαιτούν την προσοχή και ενδεχόμενη παρέμβαση του οδηγού για την οδήγηση του οχήματος. Όλα αυτά τα πλαίσια, έχουν σχεδόν ταυτόσημη γλώσσα με το *HB1649*. Ως εκ τούτου, το σύνολο υποστηρικτικών τεχνολογιών αυτοκίνησης, διέπεται από πλαίσιο των υφιστάμενων κωδικών συμβατικής κυκλοφορίας και όχι στο πλαίσιο της αυτοκίνησης.

Ακόμα, ορισμένες πολιτείες που έχουν θεσπίσει νομικό πλαίσιο για τα αυτοκινούμενα οχήματα, όπως η California και η Florida, ορίζουν με σαφήνεια εντός αυτού την υποχρεωτική ύπαρξη ενός οδηγού – χειριστή του οχήματος καθ' όλη την διαδικασία της λειτουργίας του οχήματος. Ως εκ τούτου, η προσέγγιση αυτή θα μεταβιβάζει το σύνολο των ευθυνών της ορθής κυκλοφορίας του οχήματος, στον οδηγό – χειριστή του. Ωστόσο στο πλαίσιο της Ουάσιγκτον, δεν έχουμε κάποια άμεση αναφορά στο πρόσωπο που φέρει την ευθύνη του οχήματος. [18] , [56]

1.4.2.4 Η σχεδίαση των προτύπων και κανόνων

Μία παράμετρος που διαφοροποιεί το *HB1649* έναντι των άλλων πλαισίων, είναι ότι προϋποθέτει τη συνεργασία του *Washington Department of Licensing* και του *Washington State Patrol* ώστε να αναπτυχθούν οι απαιτούμενοι κανονισμοί ασφαλείας, οι οποίοι σε επόμενο στάδιο εφαρμόζονται και

σε διάφορες άλλες πολιτείες καθιερώνοντας γενικού τύπου νόρμες και πρότυπα ασφαλείας για τα αυτόνομα οχήματα. [10], [56], [70]

1.4.2.5 Η ισχύουσα νομοθεσία αναφορικά με τις δοκιμαστικές φάσεις

Όλες οι ισχύουσες νομικές διατάξεις, επιτρέπουν τη δοκιμή των αυτόνομων οχημάτων χωρίς οδηγό σε δημόσιους δρόμους. Ωστόσο, η κάθε πολιτεία εφαρμόζει διαφορετικούς περιορισμούς αναφορικά με τις εκτελούμενες λειτουργίες κατά της διάρκεια της δοκιμής. Όταν μία δοκιμαστική διαδρομή εκτελείται χωρίς την φυσική παρουσία ενός οδηγού – χειριστή εντός του οχήματος, τότε η συμπερίληψη μίας αυτοματοποιημένης λειτουργίας του οχήματος στα πλαίσια της δοκιμής αυτής, απαιτεί και ξεχωριστή άδεια καθώς και κάλυψη προαπαιτούμενων ώστε να καταστεί εφικτή η δοκιμή της. [18], [56]

1.4.2.6 Ειδικές άδειες οδήγησης

Μόνο η Nevada, μέχρι στιγμής, έχει απαιτήσει ξεχωριστή ειδική άδεια - δίπλωμα για τη λειτουργία ενός αυτόνομου οχήματος. Σε όλες τις υπόλοιπες πολιτείες, το καθεστώς απαιτεί απλώς την ύπαρξη ενός έγκυρου διπλώματος οδήγησης για τον χειριστή του οχήματος. [18]

1.4.2.7 Αυτόματος και χειροκίνητος έλεγχος του οχήματος

Σχεδόν κάθε νομοθετική διάταξη που αφορά τη λειτουργία των αυτόνομων οχημάτων, αναφέρεται ρητώς η ανάγκη ύπαρξης τεχνολογικής υποδομής η οποία θα επιτρέπει στον οδηγό να πάρει άμεσα στα χέρια του τον έλεγχο του οχήματος εφόσον κριθεί αναγκαίο ώστε να αποφευχθεί η όποια σύγκρουση, ατύχημα η επιπλοκή προκύψει ενδεχομένως από λάθος χειρισμό του συστήματος αυτόνομης οδήγησης. Έτσι γίνεται άμεσα αντιληπτό ότι οι όποιες ευθύνες κατά τη λειτουργία του οχήματος βαραίνουν αποκλειστικά τον οδηγό. [18], [56]

1.4.2.8 Απόδοση ποινικών ευθυνών

Σχεδόν κάθε νομική προσέγγιση που εξετάζεται, η ποινική ευθύνη για την όποια έκνομη συμπεριφοράς του οχήματος, βαραίνει αποκλειστικά στον χειριστή αυτού. Ωστόσο, στο πλαίσιο SB1298 της California, δεν αποδίδονται οι ποινικές ευθύνες με τον ίδιο τρόπο καθώς το χρέος ρύθμισης του ζητήματος αυτού, μετακυλιέται στο τμήμα Μηχανοκινήτων Οχημάτων το οποίο και θα αναπτύξει τις αντίστοιχες νομικές διατάξεις. Η Washington και η Florida αναφέρουν στα πλαίσια τους ότι ο κατασκευαστής του οχήματος απαλλάσσεται από τις όποιες ποινικές ευθύνες προκύψουν από δυσλειτουργία του οχήματος, στην περίπτωση που ένα όχημα φέρει σύστημα αυτόνομης οδήγησης. [10], [13], [18], [44], [56]

1.4.2.9 Αστικές ευθύνες

Το πλαίσιο SB1298 της California, δεν αναφέρει οτιδήποτε σχετικό με τις ποινικές ευθύνες που αφορούν την χρήση των αυτόνομων οχημάτων και το χρέος ρύθμισης του ζητήματος αυτού, μετακυλιέται στο Τμήμα Μηχανοκινήτων Οχημάτων το οποίο και θα αναπτύξει τις αντίστοιχες νομικές διατάξεις. Η Florida και η Washington, προδιαγράφουν στα αντίστοιχα πλαίσια τους, την απαλλαγή αστικής ευθύνης του κατασκευαστή του οχήματος σε περιπτώσεις όπου αυτές προκύψουν. [10], [13], [18], [44], [56]

1.4.2.10 Λοιπά νομικά ζητήματα

Η California στο πλαίσιο της, απαιτεί από τους κατασκευαστές των οχημάτων να αναφέρουν στους αγοραστές το είδος των πληροφοριών που συλλέγονται από τα οχήματα τους. Η Nevada εισήγαγε στο πλαίσιο της, αντισταθμιστικές διατάξεις για τα αυτοκίνητα με υψηλά και χαμηλά επίπεδα ρύπων.

Ακόμα, η πλειοψηφία των Πολιτειών, έχουν προδιαγράψει εντός των πλαισίων τους, την αναγκαιότητα ύπαρξης ενός μαύρου κουτιού εντός των αυτόνομων οχημάτων, έτσι ώστε να είναι δυνατή η απρόσκοπτη καταγραφή όλων των δεδομένων και των συμβάντων που λαμβάνουν χώρα κατά της λειτουργίας του αυτόνομου οχήματος. [8], [18], [44], [56]

1.4.2.11 Συμπεράσματα από τις ισχύουσες νομικές διατάξεις στις Η.Π.Α.

Κάθε πολιτεία έχει εφαρμόσει στα αντίστοιχα ρυθμιστικά πλαίσια τις ίδιες ουσιαστικά διατάξεις που έχουν ορίσει και οι περισσότερες πολιτείες. Η ανάγκη φυσικής παρουσίας του οδηγού, η δυνατότητα μετάβασης του αυτόματου ελέγχου του οχήματος στον οδηγό, οι κανόνες και οι προϋποθέσεις δοκιμής των οχημάτων στους δημόσιους αυτοκινητόδρομους καθώς και η πολιτική ποινικών και αστικών ευθυνών, αποτελούν σε γενικές γραμμές κοινό παρονομαστή στα ρυθμιστικά πλαίσια των περισσότερων πολιτειών. [18]

1.5 Πεδία Εφαρμογής των Αυτόνομων Οχημάτων

1.5.1 Τα υφιστάμενα πεδία εφαρμογής των αυτόνομων οχημάτων

1.5.1.1 Στρατιωτικές εφαρμογές

Μία από τις πρώτες εφαρμογές των αυτόνομων οχημάτων, υπήρξε η σχεδίαση αυτόνομων οχημάτων από τα σώματα ασφαλείας των ενόπλων δυνάμεων των Η.Π.Α, ενώ πολλά ακόμη ερευνητικά προγράμματα σχετικά με τα αυτόνομα οχήματα συνεχίζουν να χρηματοδοτούνται από τον στρατό των Η.Π.Α. Στις εφαρμογές αυτές σχεδιάστηκαν αυτόνομα οχήματα τα οποία κινούνταν ως ομάδα εντός περιορισμένης ακτίνας, σε ιδιαίτερα δυσπρόσιτα εδάφη. Ωστόσο, το σύνολο των εφαρμογών αυτών αφορά δοκιμαστικές αποστολές και όχι πραγματικές. [23], [52]

1.5.1.2 Διαστημικές εφαρμογές

Όπως θα ήταν αναμενόμενο, οι πρώτες εφαρμογές αυτόνομων οχημάτων, πραγματοποιήθηκαν σε αεροδιαστημικές εφαρμογές από την NASA. Ένα μεγάλο πλήθος εφαρμογών που περιλαμβάνουν δειγματοληψίες αέρα και εδάφους, χαρτογράφηση περιοχών, δειγματοληψίες, χημικές αναλύσεις, αποστολές μεταφοράς εξοπλισμού κ.α. , υλοποιείται διαρκώς από την NASA σε σχεδόν κάθε τομέα έρευνας της. [23]

1.5.1.3 Εφαρμογές γεωλογικών εφαρμογών

Σε αρκετές γεωλογικές εφαρμογές, δεν είναι δυνατή η προσέγγιση του ερευνώμενου χώρου από τον άνθρωπο, καθώς λαμβάνουν χώρα ιδιαίτερα βεβαρυμμένες συνθήκες που αφορούν την επικρατούσα θερμοκρασία, το επιβλαβές χημικό περιβάλλον, την δυσπρόσιτη γεωγραφική μορφολογία κ.α. Σε πολλές από αυτές τις περιπτώσεις, αποστέλλονται αυτόνομα οχήματα που επιτελούν τις διαδικασίες χαρτογράφησης και δειγματοληψίας. [23]

1.5.1.4 Εφαρμογές έκτακτων αναγκών

Σε ορισμένες καταστάσεις έκτακτης ανάγκης, ειδικά σχεδιασμένα αυτόνομα ρομποτικά οχήματα αναλαμβάνουν την υποβοήθηση των ανθρώπινων ενεργειών. Ένα τέτοιο παράδειγμα αποτελούν τα ρομποτικά οχήματα που αποστέλλονται σε σεισμόπληκτες περιοχές για το εντοπισμό εγκλωβισμένων.

1.5.2 Τα πιθανά πεδία μελλοντικής εφαρμογής αυτόνομων οχημάτων

1.5.2.1 Εφαρμογές οχημάτων Ι.Χ.

Όπως αναλύεται εκτενώς στην παρούσα διατριβή, παρόλο που τα πρώτα βήματα για την δημιουργία αυτόνομων οχημάτων δημοσίας κυκλοφορίας έχουν ήδη πραγματοποιηθεί, απαιτείται αρκετή ακόμη προσπάθεια ώστε να επιτευχθεί η καθολική εφαρμογή των αυτόνομων οχημάτων επιπέδου 5. Ωστόσο, η ολοκληρωτική αλλαγή της συμβατικής αυτοκίνησης σε αυτόματα, θα επιφέρει πρωτόγνωρες αλλαγές και βελτιώσεις στην κοινωνική ζωή. [23]

1.5.2.2 Εφαρμογές μέσων μαζικής μεταφοράς

Με την ανάπτυξη της τεχνολογίας των αυτόνομων οχημάτων, είναι αναπόφευκτη και η εισαγωγή της στον τομέα των καθημερινών μετακινήσεων με χρήση των *Μέσων Μαζικής Μεταφοράς*. Έτσι, ολοένα και περισσότεροι πληθυσμός θα μπορούσε να μετακινείται ταυτόχρονα, ασφαλέστερα, ταχύτερα και σε περισσότερους προορισμούς. [23]

1.5.2.3 Βιομηχανικές εφαρμογές

Το βιομηχανικό περιβάλλον αποτελεί ένα ιδιαίτερο περιβάλλον που περιλαμβάνει διαρκείς μεταφορές φορτίων και υλικών, κάτω από επικίνδυνες συνθήκες, με μεγάλη ακρίβεια τοποθέτησης και με ανελαστικά χρονοδιαγράμματα ολοκλήρωσης. Οι συγκεκριμένες εφαρμογές αποτελούν ένα ιδανικό παράδειγμα υλοποίησης τους με χρήση κατάλληλα διαμορφωμένων αυτόνομων ρομποτικών οχημάτων. [23]

1.5.2.4 Εφαρμογές διανομών

Εκατομμύρια οχημάτων καθημερινά εκτελούν οδικά δρομολόγια παραλαβής και μεταφοράς και παράδοσης εμπορευμάτων, όπου τα χρονοδιαγράμματα ολοκλήρωσης τους είναι ιδιαίτερα περιορισμένα, τα αυτόνομα αυτοκίνητα θα μπορούσαν να παρέχουν μία επαναστατική λύση. Ακόμα, ζητήματα όπως η βέλτιστη επιλογή διαδρομής και ο έγκαιρος αναπρογραμματισμός διαδρομής βασιζόμενος σε κυκλοφοριακά δεδομένα, που παραμένουν άλυτα για τα συμβατικά οχήματα, θα μπορούσαν αν αποτελούν παρελθόν με την εφαρμογή των αυτόνομων οχημάτων και έτσι να υπάρξουν ολοκληρωτικές αλλαγές στις διαδικασίες διανομών. [23]

2 Η Ιστορική Εξέλιξη των Αυτόνομων Αυτοκινήτων

2.1 Σύντομη Αναδρομή στην Ιστορία των Αυτόνομων Αυτοκινήτων

Παρότι, η εμφάνιση των αυτόνομων αυτοκινήτων έγινε αισθητή τα τελευταία χρόνια, η πρώτη απόπειρα επαφής με αυτό το είδος τεχνολογίας χρονολογείται πολύ παλιότερα απ' ό,τι θα περίμενε κανείς. Ήδη από το 1920 παρουσιάζονται οι πρώτες αναφορές στην αυτοματοποίηση των αυτοκινήτων, ξεκινώντας κατ' αυτό τον τρόπο μια εποχή στην οποία ακολούθησαν πολλά σχετικά πειράματα οδηγώντας στην σταδιακή σημερινή εξέλιξη.

Τα πρώτα πραγματικά αυτόνομα αυτοκίνητα εμφανίζονται τη πρώτη δεκαετία του '80 με το *Navlab* του Πανεπιστημίου *Carnegie Mellon*, το *ALV* (1984), το ρομποτικό βαν της *Mercedes-Benz*, το *Prometheus Project* και το *Bundeswehr* του Πανεπιστημίου του Μονάχου (1987). Από εκείνη τη περίοδο, πληθώρα ερευνητικών οργανισμών και μεγάλων τεχνολογικών εταιρειών, ξεκινούν να εργάζονται πάνω στην ανάπτυξη πρωτότυπων αυτόνομων οχημάτων, όπως: η *Mercedes-Benz*, η *General Motors*, η *Continental Automotive Systems*, η *Autoliv Inc.*, η *Bosch*, η *Nissan*, η *Toyota*, η *Audi*, η *Volvo*, η ομάδα *Vislab* από το Πανεπιστήμιο της Πάρμα, το Πανεπιστήμιο της Οξφόρδης, η *Google* και η *Tesla Motors* αποτελώντας τους κυρίαρχους ερευνητικούς οργανισμούς αυτής της τεχνολογίας. [204], [67], [72]

Στη παρούσα ενότητα, παρουσιάζεται μία πιο αναλυτική ιστορική αναδρομή.

2.2 Τα Πρώτα Οράματα για μία Νέα Εποχή στην Αυτοκίνηση

Το 1925 στη Νέα Υόρκη παρουσιάζεται το «*Linrrican Wonder*» από τη *Houdina Radio Control*, το οποίο μέσω ραδιοσυχνοτήτων μπορούσε να ελέγχεται από απόσταση. Για την ακρίβεια, εξοπλίστηκε με κεραίες εκπομπής ένα *Chandler*, ώστε να μπορεί να ελέγχεται μέσω τηλεχειρισμού από άλλο αυτοκίνητο στο οποίο ακολουθούσε ο τηλεχειριστής. Το όχημα δοκιμάστηκε πειραματικά επιτυχώς στην 5th Avenue του Broadway, κάτω από συνθήκες κανονικής κυκλοφορίας.

Το 1926, η *Aachen Motor*, λάνσαρε στους δρόμους του *Milwaukee*, το "*Phantom Auto*", όνομα το οποίο προέκυψε από το ότι αφού δεν υπάρχει οδηγός, είναι σα να οδηγείται από κάποιο φάντασμα. Τέλος, το 1932 το "*Phantom Auto*" λανσαρίστηκε εκ νέου στη πόλη *Fredericksburg*.

Το 1939, πραγματοποιείται η διεθνής έκθεση *Futurama* στη Νέα Υόρκη, με την υποστήριξη της *General Motors*. Εκεί παρουσιάστηκε μια αναπαράσταση του τεχνολογικού κόσμου για το έτος 1960 και τα μελλοντικά αυτοκίνητα. Πιο συγκεκριμένα, υπήρξε μια προσομοίωση μελλοντικής πόλης με δρόμους πολλών λωρίδων, πολύ υψηλών ταχυτήτων κυκλοφορίας, όπου υπήρχαν πολλά αυτόνομα ηλεκτρικά αυτοκίνητα, ο χειρισμός των οποίων πραγματοποιούνταν μέσω συστήματος ραδιοσυχνοτήτων τηλεκατεύθυνσης, ενώ η ηλεκτροδότηση τους θα πραγματοποιείται μέσω επαγωγής από ηλεκτρομαγνητικά πεδία που θα προκαλούνταν από κυκλώματα ενσωματωμένα στο οδόστρωμα των πόλεων.



Εικόνα 2.1

Η διεθνής έκθεση Futurama.

Το 1936 στο εξώφυλλο του περιοδικού *Modern Mechanix* παρουσιάζεται το όραμα για το όχημα "*Electric Eye Automobile*", το οποίο θα χαρακτηριζόταν από ένα ειδικό σύστημα ελέγχου, όπου με τη χρήση ενός συνόλου φωτοκύτταρων, θα είχε τη δυνατότητα να ακολουθεί μια δέσμη φωτός, η οποία θα εκπέμπεται από το αυτοκίνητο και θα αντανακλάται από κάτοπτρα στην επιφάνεια του δρόμου, με αποτέλεσμα να κινείται χωρίς την αναγκαιότητα οδηγού.

Το 1953, η *RCA Labs* κατασκεύασε μια μινιατούρα αυτοκινήτου η οποία ελεγχόταν και κατευθυνόταν μέσω καλωδίων που ήταν συνδεδεμένα σε μία ειδική δομή ενός εργαστηριακού δαπέδου, επίτευγμα το οποίο ενθουσίασε τους Leland M Hancock, μηχανικός κυκλοφορίας στη πολιτεία της Nebraska και L. N. Ress, διευθυντής του Hancock, οι οποίοι αποφάσισαν να προχωρήσουν στην πειραματική λειτουργία ενός κανονικού αυτοκινήτου στους τοπικούς αυτοκινητόδρομους, σε πραγματικές συνθήκες κυκλοφορίας.

Έτσι το 1958 κατασκευάστηκε επιτυχώς από την *RCA Labs* ένα αυτοκίνητο κανονικών διαστάσεων και τέθηκε προς δοκιμή σε μια λωρίδα μήκους 400 ποδιών, σε κεντρικό αυτοκινητόδρομο, λίγο έξω από το Lincoln, κατά μήκος των πεζοδρομίων του οποίου είχαν τοποθετηθεί αισθητήρες παρουσίας του αυτοκινήτου. Κατά συνέπεια, μόλις το όχημα γινόταν αντιληπτό, το σύστημα ελέγχου, το οποίο βρισκόταν κατά μήκος του πεζοδρομίου, έπαιρνε υπό τον έλεγχο του το σύστημα πλοήγησης του αυτοκινήτου.

Το σύστημα αυτό αναπτύχθηκε σε συνεργασία με την *General Motors*, και περιλάμβανε πέραν των αισθητήρων, πομποδέκτες ασύρματης μετάδοσης, ηχητική ειδοποίηση, οπτική ειδοποίηση καθώς και τα απαιτούμενα ηλεκτρικά μέρη ελέγχου του αυτοκινήτου.

Το 1960, στην έδρα της *RCA Lab* στο Princeton του New Jersey, έλαβε χώρα και η επίσημη παρουσίαση του επιτεύγματος, κατά το οποίο οι παρευρισκόμενοι δημοσιογράφοι είχαν τη δυνατότητα να πάρουν τον έλεγχο των αυτοκινήτων. Η ολοκλήρωση του ερευνητικού σταδίου και η έναρξη παραγωγής φιλοδοξούσε να πραγματοποιηθεί κοντά στο 1975, ενώ στα τέλη της δεκαετίας η *General Motors* παρουσίασε την νέα σειρά αυτοκινήτων *Firebirds*. Τα *Firebirds* που αποτελούσαν μια σειρά πειραματικών αυτοκινήτων που σε συνδυασμό με έναν αυτόματο αυτοκινητόδρομο, μπορούσαν να ξεκουράσουν για κάποιες στιγμές τον οδηγό.

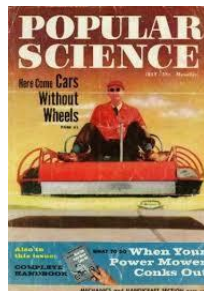


Εικόνα 2.2

Το αυτόνομο αυτοκίνητο Firebird.

Το 1960, το *Εργαστήριο Επικοινωνιών και Συστημάτων Ελέγχου* του Πανεπιστημίου του Οχάιο, ξεκινά τις έρευνες για την ανάπτυξη αυτοκινήτων χωρίς οδηγό, τα οποία θα ελέγχονταν από ηλεκτρονικές διατάξεις τοποθετημένες εντός του οδοστρώματος ενώ το 1966 ο Δρ. Robert L. Cosgriff, επικεφαλής της συγκεκριμένης έρευνας, υποστήριξε πως το σύστημα θα μπορούσε να είναι έτοιμο για εγκατάσταση στους δημόσιους δρόμους σε 15 χρόνια.

Παράλληλα, το *Γραφείο Δημοσίων Οδών* εξέταζε το ενδεχόμενο της κατασκευής ενός πειραματικού ηλεκτρονικά ελεγχόμενου αυτοκινητοδρόμου. Τέσσερις πολιτείες – το Οχάιο, η Μασαχουσέτη, η Νέα Υόρκη και η Καλιφόρνια - πλειοδότησαν το έργο της κατασκευής. Λίγο αργότερα, τον Αύγουστο του 1961, το περιοδικό *Popular Science* παρουσίασε ένα αφιέρωμα για το *Aeromobile 35B*, ένα αεριωθούμενο όχημα (ACV), δημιούργημα του William Bertelsen σχεδιασμένο με σκοπό να φέρει την επανάσταση στις μεταφορές αφού, όχι μόνο θα αποτελούσε ένα αυτοκινούμενο όχημα, αλλά θα μπορούσε να επιταχύνει σε ταχύτητες της τάξης των 2400 km/h.



Εικόνα 2.3

Το αεριωθούμενο όχημα Aeromobile 35B.

Κατά τη διάρκεια της δεκαετίας του 1960, το *Ερευνητικό Ινστιτούτο Μεταφορών και Οδών* του Ηνωμένου Βασιλείου δοκίμασε το αυτοκινούμενο Citroen DS, το οποίο ελεγχόταν ηλεκτρομαγνητικά από κατάλληλα διαμορφωμένα καλώδια που βρίσκονταν τοποθετημένα κάτω από το οδόστρωμα ενώ δοκιμάστηκε επιτυχώς εντός πίστας στα 130km/h, χωρίς την παραμικρή απόκλιση ταχύτητας ή κατεύθυνσης.



Εικόνα 2.4

Το αυτόνομο αυτοκίνητο Citroen DS.

Οι έρευνες για τον συγκεκριμένο τρόπο ελέγχου οχήματος συνεχίστηκαν καθ' όλη τη δεκαετία του '70. Οι αναλύσεις κόστους-οφέλους που πραγματοποιήθηκαν, έδειξαν πως η εφαρμογή του συγκεκριμένου συστήματος θα έδινε τα πρώτα της αποτελέσματα στο τέλος του αιώνα ενώ τα αυτοκινητιστικά ατυχήματα θα μπορούσαν να ελαττωθούν έως και 40%. Ωστόσο, στα μέσα της δεκαετίας του 1970, η χρηματοδότηση του όλου εγχειρήματος διακόπηκε.

Τις δεκαετίες 1960 και 1970, η *Bendix Corporation* προχώρησε στην επιτυχή κατασκευή αυτοκινούμενου οχήματος που ελεγχόταν μέσω καλωδίων τοποθετημένα εντός του οδοστρώματος, ενώ παράλληλα πομποδέκτες παρείχαν επικοινωνία του οχήματος και ενός υπολογιστή, μέσω μηνυμάτων. Παράλληλα, το *εργαστήριο Τεχνητής Νοημοσύνης* του *Πανεπιστημίου Στάνφορντ*, κατασκεύασε και παρουσίασε ένα μικρό αυτοκινούμενο τροχοφόρο ρομπότ, ενώ στο πρώτο μισό της δεκαετίας του 1970 το *Πανεπιστήμιο του Ιλινόις* πραγματοποίησε εντατικές μελέτες πάνω στην *Ευφυή Αυτοματοποιημένη Λογική* των αυτόνομων οχημάτων. [204]

2.3 Τα Πρώτα Οχήματα Αυτόνομων Λειτουργιών

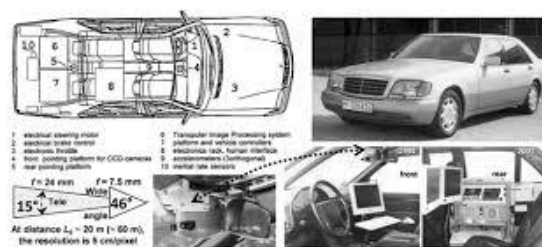
Το 1980, ο Ernst Dickmanns με την ομάδα του από το πανεπιστήμιο του Μονάχου, σχεδίασαν επιτυχώς ένα ρομποτικό φορτηγό όχημα *Mercedes-Benz*, το οποίο έφερε για πρώτη φορά στοιχεία Μηχανικής Όρασης. Το όχημα δοκιμάστηκε επιτυχώς στους δρόμους του Μόναχο, πετυχαίνοντας την ταχύτητα των 63 km/h.

Επιπρόσθετα, από το 1987 έως το 1995, ο οργανισμός *EUREKA*, διεξήγαγε το project με το όνομα *Prometheus* συνολικής δαπάνης €749.000.000, πάνω στα αυτόνομα οχήματα ενώ την ίδια δεκαετία, το πειραματικό όχημα με το όνομα *Autonomous Land Vehicle (ALV)*, που σχεδιάστηκε από το *DARPA* των Η.Π.Α. , υιοθέτησε ένα πλήθος νέων τεχνολογιών αυτόνομων αυτοκινήτων, οι οποίες αναπτύχθηκαν στα πανεπιστήμια *Μέριλαντ*, *Carnegie Mellon*, *Martin Marietta*, *SRI International* και το *Ινστιτούτο Ερευνών Περιβάλλοντος του Μίσιγκαν*. Το *ALV* ήταν ένα αυτόνομο ρομποτικό όχημα το οποίο χρησιμοποιούσε *LIDAR* και τεχνικές Υπολογιστικής Όρασης και δοκιμάστηκε επιτυχώς ακολουθώντας αλάνθαστα τον αυτοκινητόδρομο πάνω στον οποίο τοποθετήθηκε, αναπτύσσοντας την ταχύτητα των 31 km/h.

Παράλληλα, τα Εργαστήρια *HRL* (πρώην *Hughes Research Labs*) προχώρησαν στην σχεδίαση και εφαρμογή μιας καινοτόμου λύσης πλοήγησης μέσω off-road χαρτών και σύγχρονων οπτικών αισθητήρων στο *AVL*. Έτσι, το *AVL* διένυσε 610m με ταχύτητα 3,1 km/h πάνω σε ανώμαλο έδαφος με απότομες πλαγιές, χαράδρες, μεγάλες πέτρες, και τη βλάστηση ενώ έως το 1989, το πανεπιστήμιο *Carnegie Mellon* κατάφερε να είναι πρωτοπόρο στην έρευνα και εφαρμογή των Νευρωνικών Δικτύων αλλά και την εφαρμογή τους στον έλεγχο αυτόνομων αυτοκινήτων, γεγονός το οποίο αποτελεί θεμελιώδη τεχνική ελέγχου έως και σήμερα.

Το 1991, το Κογκρέσο των Ηνωμένων Πολιτειών ψήφισε το *ISTEA*, βάσει του οποίου η *USDOT* έπρεπε μέχρι το 1997 να παρουσιάσει ένα πρωτότυπο σύστημα αυτοματοποιημένου οχήματος και οδικής κυκλοφορίας. Την υλοποίηση του έργου ανέλαβε η *Federal Highway Administration*, ιδρύοντας την *NAHSC (National Automated Highway System Consortium)* και συγχρηματοδοτήθηκε από την *FHWA*, την *General Motors*, την *Caltrans*, την *Delco*, την *Parsons Brinkerhoff*, την *Bechtel*, το *UC-Berkeley*, το Πανεπιστήμιο *Carnegie Mellon*, και το *Lockheed Martin*. Το 1997 στο Σαν Ντιέγκο της Καλιφόρνια έλαβε χώρα η τεχνολογική έκθεση *Demo*, στην οποία παρουσιάστηκαν 20 περίπου αυτοματοποιημένα οχήματα, συμπεριλαμβανομένων αυτοκινήτων, λεωφορείων και φορτηγών, προσελκύοντας χιλιάδες επισκέπτες και δημοσιογράφους.

Το 1994, τα δίδυμα ρομποτικά οχήματα *Vamp* και *Vita-2* της *Daimler-Benz* διένυσαν απόσταση 1.000 χιλιομέτρων σε αυτοκινητόδρομο τριών λωρίδων στο Παρίσι, όπου σε λειτουργία ημι-αυτονομίας ανέπτυξαν ταχύτητες της τάξης των 130 χλμ/ώρα.



Εικόνα 2.5

Το αυτόνομο αυτοκίνητο Vamp.

Το 1995, το πειραματικό όχημα *Navlab* του Πανεπιστημίου *Carnegie Mellon* διένυσε 5.000 χιλιόμετρα κατά μήκος όλης τη χώρας, εκ των οποίων το 98,2% σε αυτόνομη λειτουργία. Ωστόσο το όχημα αυτό, ενώ ήταν από τη φύση του ημι-αυτόνομο, δηλαδή χρησιμοποιούσε Νευρωνικά Δίκτυα για τον έλεγχο του τιμονιού, η επιτάχυνση και η πέδηση για λόγους ασφαλείας πραγματοποιούνταν από τον οδηγό.



Εικόνα 2.6

Το πειραματικό αυτόνομο όχημα Navlab.

Το ίδιο έτος, η ανασχεδιασμένη ημι-αυτόνομη *Mercedes-Benz S-Class* διένυσε 1.590 χιλιόμετρα από το Μόναχο στη Δανία και πάλι πίσω, χρησιμοποιώντας εξοπλισμό Μηχανικής Όρασης. Το ρομποτικό αυτό όχημα ξεπέρασε την ταχύτητα των 175 km/h, ενώ στο 95% του χρόνου, η οδήγηση ήταν αυτόνομη ενώ αξιοσημείωτο είναι πως χρησιμοποιούσε τεχνικές ελιγμών και αποφυγής των υπολοίπων οχημάτων.

Το 1996, ο *Alberto Broggi* από το Πανεπιστήμιο της Πάρμα, ξεκίνησε τη σχεδίαση του *ARGO*, με στόχο να μπορεί να ακολουθεί τις λωρίδες του συμβατικού αυτοκινητόδρομου, από μία ειδικά τροποποιημένη *Lancia Thema*. Το εγχείρημα στέφθηκε από απόλυτη επιτυχία και το όχημα διένυσε μία απόσταση 1.900 χιλιομέτρων, διάρκειας 6 ημερών στους Ιταλικούς αυτοκινητοδρόμους, με μέση ταχύτητα τα 90 km/h. Κατά το 94% του χρόνου λειτουργίας του, το αυτοκίνητο βρισκόταν σε πλήρως αυτόματη λειτουργία και ήταν εξοπλισμένο με δύο ασπρόμαυρες κάμερες χαμηλού κόστους, χρησιμοποιώντας αλγόριθμους στερεοσκοπικής όρασης για την αντίληψη του γύρω περιβάλλοντος.

Το 1998, ο *Willie Jones* δήλωσε ότι πολλές αυτοκινητοβιομηχανίες έχουν εντάξει τον τομέα της αυτόνομης οδήγησης στον κορμό της ετήσιας τους έρευνας. Ανέφερε επίσης, ότι τον Μάιο του 1998, η Τοyota ήταν η πρώτη αυτοκινητοβιομηχανία που εισήγαγε το σύστημα *ACC (Adaptive Cruise Control)* σε όχημα παραγωγής, το οποίο για το σκοπό αυτό φέρει ένα εξελιγμένο σύστημα λέιζερ. [204], [56]

2.4 Οι Αρχικές Προσεγγίσεις των Σύγχρονων Αυτόνομων Οχημάτων

Στις αρχές της δεκαετίας του 2000, τίθεται σε λειτουργία στην Ολλανδία το *ParkShuttle*, ένα αυτόνομο μέσο μαζικής μεταφοράς.



Εικόνα 2.7

Το αυτόνομο όχημα μαζικής μεταφοράς ParkShuttle.

Παράλληλα, η κυβέρνηση των ΗΠΑ χρηματοδοτεί τρία στρατιωτικά προγράμματα ερευνητικού χαρακτήρα στοχεύοντας στη σχεδίαση μίας ομάδας μη επανδρωμένων οχημάτων που θα είχαν τη δυνατότητα να κινούνται σε ανώμαλο και δύσβατο έδαφος αποφεύγοντας εμπόδια όπως βράχους και δέντρα. Τα τρία αυτά προγράμματα είναι γνωστά ως *Demo I* (του Στρατού των ΗΠΑ), *Demo II* (του *DARPA*) και *Demo III* (του Στρατού των ΗΠΑ). Συγκεκριμένα, το *Demo III* στέφθηκε από απόλυτη επιτυχία και ο *James Albus* από το *Εθνικό Ινστιτούτο Προτύπων και Τεχνολογίας*, παρουσίασε ένα σύστημα ελέγχου πραγματικού χρόνου, το οποίο λειτουργούσε ιεραρχικά και ήταν σε θέση να ελέγχει όχι μόνο

μεμονωμένα οχήματα αλλά και ομάδες αυτών συντεταγμένα με σκοπό την επίτευξη σύνθετων στόχων που απαιτούσαν συνεργασία.

Επίσης, τον Μάρτιο του 2004 ο *DARPA (Defense Advanced Projects Agency)* διεξάγει το πρώτο *Grand Challenge*, στο οποίο και προσέφερε ως βραβείο ένα εκατομμύριο δολάρια σε όποια ομάδα μηχανικών ρομποτικής θα μπορούσε να καταφέρει να κατασκευάσει ένα αυτόνομο αυτοκίνητο ικανό να διασχίσει μια πορεία 150 μιλίων στην έρημο *Mojave*, αλλά καμία ομάδα δεν έφερε εις πέρας τη πρόκληση.

Στη συνέχεια τον Οκτώβριο του 2005, πραγματοποιήθηκε το *Grand Challenge II* του *DARPA* που αφορούσε και πάλι το περιβάλλον της ερήμου, όμως αυτή τη φορά τοποθετήθηκαν δέκτες *GPS* που βοήθησαν στη χαρτογράφηση των σημείων που περιείχαν σημαντικά εμπόδια. Πέντε οχήματα έφεραν εις πέρας την αποστολή.

Την ίδια χρονιά, η *BMW* πραγματοποιεί δοκιμές στα αυτόνομα αυτοκίνητα της από το 2005.

Τον Ιανουάριο του 2006, η "*Foresight*" του Ηνωμένου Βασιλείου, δημοσίευσε μία έκθεση η οποία προέβλεπε ότι μέχρι το 2056, αυτόνομα οχήματα *RFID* θα βρίσκονται στους δρόμους του Ηνωμένου Βασιλείου, ενώ παράλληλα η *Ακαδημία Μηχανικών του Ηνωμένου Βασιλείου* υποστήριξε ότι μέχρι το 2019, θα βρίσκονται στους Βρετανικούς αυτοκινητόδρομους αυτόνομα φορτηγά οχήματα.

Τον Νοέμβριο του 2007, ο *DARPA* διεξήγαγε το *Grand Challenge III*, αλλά αυτή τη φορά η πρόκληση αφορούσε στο αστικό περιβάλλον. Σε αυτόν τον διαγωνισμό, το *Chevy Tahoe*, ένα αυτόνομο αυτοκίνητο του Πανεπιστημίου *Carnegie Mellon*, κέρδισε την 1η θέση. Έτσι, οι διαγωνισμοί βραβείων του *DARPA*, έδωσαν την ευκαιρία σε φοιτητές και ερευνητές να ερευνήσουν και να σχεδιάσουν νέες τεχνολογίες για τα αυτόνομα αυτοκίνητα, ελαττώνοντας σημαντικά προβλήματα μετακινήσεων, όπως η κυκλοφοριακή συμφόρηση και τα τροχαία ατυχήματα, ειδικά στα μεγάλα αστικά κέντρα.



Εικόνα 2.8

Το αυτόνομο όχημα Chevy Tahoe του Πανεπιστημίου Carnegie Mellon

Τον Δεκέμβριο του 2008 στα ορυχεία σιδήρου *Pilbara* της Δυτικής Αυστραλίας, η *Rio Tinto Alcan* πραγματοποίησε δοκιμές με το *Komatsu Autonomous Haulage System* – το πρώτο στον κόσμο αυτόνομο όχημα μεταφορών μεταλλουργίας και προϊόντων εξόρυξης. Η *Tinto Alcan* παρατήρησε ότι το εγχείρημα αυτό θα προσέφερε μεγάλα οφέλη στους τομείς της υγείας, της ασφάλειας και την παραγωγικότητας με αποτέλεσμα τον Νοέμβριο του 2011, η *Rio Tinto* να υπογράψει συμφωνία επέκτασης του στόλου αυτόνομων φορτηγών της.

Τέλος, το 2009, η *Google* προχώρησε στην ανάπτυξη των αυτόνομων αυτοκινήτων της, ωστόσο κράτησε την έρευνα σε ιδιωτικό επίπεδο, αποφεύγοντας τη δημοσιοποίηση της μέχρι τον ερχομό της κατάλληλης στιγμής. [204], [6], [18]

2.5 Τα Αυτόνομα Αυτοκίνητα Σήμερα

Το 2010 η Audi απέστειλε ένα αυτόνομο *Audi TTS* στην κορυφή του όρους *Peak Pike* του Κολοράντο, αναπτύσσοντας αρκετά υψηλές ταχύτητες.

Το ίδιο έτος, το εργαστήριο *VisLab* του Πανεπιστημίου της Πάρμα, με επικεφαλή τον καθηγητή Alberto Broggi, διεξήγαγε το *VisLab Intercontinental Autonomous Challenge (VIAC)*, το πρώτο διηπειρωτικό δοκιμαστικό ταξίδι αυτόνομων αυτοκινήτων, συνολικής απόστασης 15.900 χιλιομέτρων. Αυτό το ταξίδι διάρκειας 100 ημερών, το ολοκλήρωσαν με επιτυχία τέσσερα αυτόνομα ηλεκτρικά βαν, φεύγοντας στις 20 Ιουλίου 2010 από την Πάρμα της Ιταλίας, και καταλήγοντας στις 28 Οκτωβρίου στη Shanghai Expo στην Κίνα. Το ερευνητικό αυτό πρόγραμμα συγχρηματοδοτήθηκε από το Ευρωπαϊκό πρόγραμμα *CORDIS*.

Παράλληλα, το Ινστιτούτο Αυτομάτου Ελέγχου του *Technische Universität Braunschweig* παρουσίασε το πρώτο ερευνητικό αυτόνομο όχημα, το *Leonie*, το οποίο αποτέλεσε και το πρώτο αυτόνομο όχημα που έλαβε άδεια κυκλοφορίας στους δρόμους της Γερμανίας.

Το 2011, η *GM* δημιούργησε το *EN-V (Electric Networked Vehicle)*, ένα αυτόνομο ηλεκτρικό αυτοκίνητο πόλη, ενώ παράλληλα το *Freie Universität Berlin* αναπτύσσει δύο αυτόνομα αυτοκίνητα σχεδιασμένα για κυκλοφορία σε συνθήκες πόλης. Τα δύο αυτόνομα αυτοκίνητα κατάφεραν επιτυχώς να κινηθούν εντός της πόλης, ανταπεξερχόμενα πλήρως σε σηματοδότηση δρόμων, κυκλικούς κόμβους κ.α.

Το 2012, η *Volkswagen* ξεκίνησε τις δοκιμές του συστήματος "*Temporary Auto Pilot*" (TAP), ένα σύστημα το οποίο επέτρεπε την αυτοκίνηση ενός οχήματος με ταχύτητες έως και 130km/h σε αυτοκινητόδρομο.

Την 1η Μαΐου του 2012, μηχανικοί δοκίμασαν τις δυνατότητες ένα αυτόνομου αυτοκινήτου της *Google*, κατά τη διάρκεια μίας δοκιμαστικής διαδρομής μήκους 22 χιλιομέτρων στο Λας Βέγκας. Το αυτόνομο αυτό αυτοκίνητο ολοκλήρωσε με επιτυχία τη διαδρομή, χωρίς όμως να εξεταστεί η συμπεριφορά του σε συνθήκες κυκλοφορίας κυκλικών κόμβων, μη επισημασμένων σιδηροδρομικών και σχολικών διαβάσεων.

Τέλος, τη περίοδο αυτή, η *Ford* ολοκληρώνει μία εκτεταμένη έρευνα σε τεχνολογίες και συστήματα επικοινωνίας μεταξύ αυτόνομων οχημάτων.

Τον Ιανουάριο του 2013, η *Toyota* παρουσίασε ένα μερικώς αυτόνομο όχημα εξοπλισμένο με πολυάριθμους αισθητήρες και συστήματα επικοινωνίας.

Παράλληλα, εξελίχθηκαν κι άλλα ερευνητικά προγράμματα που αφορούσαν τα επιβατικά οχήματα όπως το *2GetThere* στην Ολλανδία, ενώ πραγματοποιήθηκαν και άλλες μελέτες που αφορούσαν την εφαρμογή της τεχνολογίας αυτόνομων οχημάτων στον τομέα των δημόσιων μεταφορών.

Από το 2013 και μετά, πολλοί μεγάλοι κατασκευαστές αυτοκινήτων, συμπεριλαμβανομένου των *General Motors, Ford, Mercedes Benz, Volkswagen, Audi, Nissan, Toyota, BMW* και *Volvo*, ξεκινούν την πραγματοποίηση μελετών και δοκιμών στα αυτόνομα αυτοκίνητα.

Στις 12 Ιουλίου της ίδιας χρονιάς, η *VisLab* διεξάγει άλλη μία πρωτοποριακή δοκιμή, κατά την οποία ένα μη επανδρωμένο ρομποτικό όχημα κινήθηκε στο κέντρο της πόλης Πάρμα με απόλυτη επιτυχία, ενώ η οδήγηση περιλάμβανε κυκλικούς κόμβους, φανάρια, διαβάσεις πεζών και άλλους κίνδυνους.

Λίγες μέρες μετά, τον Αύγουστο του 2013, η *Daimler E&A* μαζί με το *Karlsruhe Institute of Technology/PZI*, εγκαθιστώντας σε μία *Mercedes-Benz S-Class* στέρεο κάμερες και ραντάρ, κατάφεραν να δημιουργήσουν ένα πλήρως αυτόνομο όχημα το οποίο διένυσε επιτυχώς απόσταση περίπου 100 χλμ από το Mannheim στην Pforzheim, της Γερμανίας, μέσω της ιστορικής διαδρομής *Bertha Benz Memorial*.

Τον Αύγουστο του 2013, η *Nissan* ανακοινώνει τα σχέδιά της για έναρξη παραγωγής μη επανδρωμένων αυτόνομων αυτοκινήτων από το 2020. Η *Nissan* εγκαθιστά ένα σύστημα αυτόνομης οδήγησης σε ένα ηλεκτρικό αυτοκίνητο τύπου *Nissan Leaf*, το οποίο παρουσιάστηκε στο έκθεση επίδειξης *Nissan 360 test drive*, που έλαβε χώρα στην Καλιφόρνια, τον Αύγουστο του 2013. Τον Σεπτέμβριο του 2013, το *Nissan Leaf* ακολουθεί το πρότυπο *Advanced Driver Assistance System* κάτι το οποίο του επιτρέπει να κινηθεί στους Ιαπωνικούς αυτοκινητόδρομους. Το όχημα αυτό χρησιμοποιείται από μηχανικούς της *Nissan* ώστε να αξιολογήσουν την αποδοτικότητα του λογισμικού αυτόνομης οδήγησης που δημιούργησαν, στον πραγματικό κόσμο. Έτσι το λογισμικό θα μπορέσει να βελτιωθεί σημαντικά αντιμετωπίζοντας κάθε πιθανό χαρακτηριστικό που απαιτεί βελτίωση. Κατά συνέπεια, τον Νοέμβριο του 2013, το αυτόνομο *Nissan Leaf* λανσαρίστηκε στους δημόσιους δρόμους της Ιαπωνίας κατά την διάρκεια μίας εκδήλωσης MME. Το *Nissan Leaf* κινήθηκε στην λεωφόρο Sagami Expressway του νομού Kanagawa, κοντά στο Τόκιο. Ο Toshiyuki Shiga, αντιπρόεδρος της *Nissan* και ο Yuji Kuroiwa, κυβερνήτης του νομού, επέβαιναν στο αυτοκίνητο κατά τη διάρκεια της δοκιμής.



Εικόνα 2.9

Το αυτόνομο αυτοκίνητο Nissan Leaf.

Η Mercedes S-Class του 2014, διατίθεται επιλεκτικά με την δυνατότητα αυτόνομης πλοήγησης, αυτόνομης διατήρησης λωρίδας κυκλοφορίας, αυτόματης επιτάχυνσης / πέδησης, αυτόματου συστήματος πάρκινγκ, αυτόματης αποφυγής ατυχημάτων και συστήματος ανίχνευση κόπωσης του οδηγού, ενώ οι ταχύτητες που μπορεί να αναπτύξει εντός και εκτός πόλης, φτάνουν τα 200 χιλιόμετρα ανά ώρα.

Το *Infiniti Q50* της ίδιας χρονιάς, εξοπλισμένο με κάμερες, ραντάρ και λοιπές άλλες τεχνολογίες, εφαρμόζει τεχνικές διατήρησης λωρίδων κυκλοφορίας, αποφυγής σύγκρουσης και ελέγχου ταξιδιού. Ένας κριτικός δήλωσε, «Με το *Q50* να διαχειρίζεται τη δική του ταχύτητα και προσαρμογή, θα μπορούσατε να καθίσετε πίσω και να απλά να παρακολουθείτε, ακόμη και σε αυτοκινητόδρομους με ήπιες στροφές, για τρία ή περισσότερα μίλια, και μάλιστα χωρίς στάση», προσθέτοντας ότι δεν θα απαιτούταν η όποια επαφή με το τιμόνι ή τα πετάλια.



Εικόνα 2.10

Το αυτόνομο αυτοκίνητο Nissan Infiniti Q50.

Παρά το γεγονός ότι από το 2013, τα πλήρως αυτόνομα οχήματα δεν είναι ακόμα διαθέσιμα στο κοινό, πολλά σύγχρονα μοντέλα αυτοκινήτων έχουν χαρακτηριστικά που προσφέρουν περιορισμένη αυτόνομη λειτουργικότητα. Αυτά περιλαμβάνουν το *Adaptive Cruise Control*, ένα σύστημα που παρακολουθεί τις αποστάσεις σε παρακείμενα οχήματα στην ίδια λωρίδα, ρυθμίζοντας την ταχύτητα με την ροή της κυκλοφορίας και *Lane Assist*, το οποίο επιτηρεί τη θέση του οχήματος στην λωρίδα, και είτε προειδοποιεί τον οδηγό όταν το όχημα αφήνει τη λωρίδα του, ή λιγότερο συχνά, αναλαμβάνει Αναγνώριση πεζών από αυτόνομο ρομποτικό όχημα με χρήση τεχνητής νοημοσύνης και τεχνικών βαθιάς μάθησης

διορθωτικά μέτρα και υποβοήθηση στάθμευσης, βοηθώντας τον οδηγό στο έργο της παράλληλης στάθμευσης.

Τον Ιανουάριο του 2014, η υπηρεσία μεταφοράς με λεωφορείο το *Navia*, έγινε η πρώτη εμπορική υπηρεσία μεταφοράς με αυτοκινούμενο όχημα. Με σημαντικό περιορισμό την ταχύτητα των 20km/h, το υπαίθριο ηλεκτρικό όχημα διαθέτει καθίσματα για έως και οκτώ άτομα και προορίζεται για να κινείται γύρω από «πεζοδρομημένα» κέντρα πόλεων, μεγάλες βιομηχανικές εγκαταστάσεις, αεροδρόμια, θεματικά πάρκα, πανεπιστήμια και νοσοκομεία.

Στις 27 Μαΐου 2014, η *Google* ανακοίνωσε τα σχέδιά της να παρουσιάσει 100 αυτόνομα πρωτότυπα αυτοκίνητα που δημιουργήθηκαν από το μηδέν στο εσωτερικό μυστικό Χ εργαστήριο της, τρέχοντας μέσα στα επόμενα δύο χρόνια ένα πιλοτικό πρόγραμμα παρόμοιο με εκείνο που χρησιμοποιήθηκε για την *Cr-48 Chromebook* το 2010.

Τον Οκτώβριο του 2014 η *Tesla Motors* ανακοίνωσε την πρώτη έκδοση του αυτόματου πιλότου εν ονόματι *Model S*, όπου αυτοκίνητα εξοπλισμένα με το σύστημα αυτό, έχουν τη δυνατότητα αυτόματου ελέγχου λωρίδας, αυτόματου ελέγχου, πέδησης και αυτόματη προσαρμογή του ορίου ταχύτητας βασιζόμενα στην αναγνώριση σημάτων κυκλοφορίας. Το σύστημα αυτό παρέχει επίσης τη δυνατότητα αυτόματης στάθμευσης, ενώ παράλληλα δύναται να λαμβάνει ενημερώσεις λογισμικού ώστε να βελτιώνει διαρκώς την οδική του συμπεριφορά. Τον Μάρτιο του 2015, η *Tesla* λανσάρει στα νέα της μοντέλα ένα νέο προηγμένο σύστημα αυτόνομης οδήγησης με χρήση αυτόματου πιλότου σε μεγάλους αυτοκινητόδρομους, όπως και εφαρμόστηκε αντίστοιχο παράδειγμα στην εθνική οδό μεταξύ του Σαν Φρανσίσκο και του Σιάτλ, με έναν οδηγό εντός του αυτοκινήτου που κατείχε αποκλειστικά τον ρόλο του επιβάτη.

Τον ίδιο μήνα του 2015, η *Tesla Motors* ανακοίνωσε ότι θα εισαγάγει την τεχνολογία αυτόματου πιλότου, μέσα από μια ενημερωμένη έκδοση λογισμικού για αυτοκίνητα εξοπλισμένα με συστήματα που υποστηρίζουν την αυτόνομη οδήγηση, και έτσι στα μέσα του Οκτώβρη του 2015 παρουσιάζει την έκδοση 7 του λογισμικού της στις Η.Π.Α., όπου περιλάμβανε τη δυνατότητα αυτόματου πιλότου. Στις 9 Ιανουαρίου του 2016, η *Tesla* εκδίδει την ενημέρωση 7.1 του λογισμικού της, προσθέτοντας ένα νέο χαρακτηριστικό το οποίο δίνει τη δυνατότητα στα αυτοκίνητα να παρκάρουν αυτόματα σε θέσεις στάθμευσης, ακόμη και χωρίς οδηγό εντός του αυτοκινήτου.



Εικόνα 2.11

Το αυτόνομο αυτοκίνητο Tesla Model S.

Τον Φεβρουάριο του 2015, η *Volvo Cars* ανακοίνωσε τα σχέδιά της για την εκμίσθωση 100 από τα *SUV XC90* εξοπλισμένα με το σύστημα *Drive Me*.

Τον Απρίλιο του 2015, ένα αυτοκίνητο που σχεδιάστηκε από την *Delphi Automotive*, έγινε το πρώτο αυτοματοποιημένο όχημα που είχε τη δυνατότητα να πραγματοποιήσει ένα παράκτιο ταξίδι σε όλη τη Βόρεια Αμερική, ταξιδεύοντας από το Σαν Φρανσίσκο στη Νέα Υόρκη, υπό τον έλεγχο του αυτόματου πιλότου για ποσοστό μεγαλύτερο του 99% της συνολικής απόστασης.

Τον Ιούλιο του 2015, η *Google* ανακοίνωσε ότι τα πειραματικά της αυτόνομα οχήματα, είχαν εμπλακεί σε 14 μικροατυχήματα από την έναρξη του ερευνητικού της έργου μέχρι και το 2009, ενώ ο Chris Urmsom, επικεφαλής του έργου, δήλωσε ότι όλα τα ατυχήματα προκλήθηκαν από άνθρωπο που οδηγούσε άλλο αυτοκίνητο. Πιο συγκεκριμένα 11 από τα ατυχήματα αυτά αφορούσαν οπίσθιες συγκρούσεις ενώ σημείωσε ότι στα έξι συνολικά έτη διάρκειας του ερευνητικού προγράμματος, τα

πειραματικά αυτόνομα οχήματα της, είχαν διανύσει σχεδόν 2 εκατομμύρια μίλια στους δημόσιους αυτοκινητόδρομους. Το Δεκέμβριο του 2016, η *Google* προχώρησε στην ίδρυση της *Waymo*, της θυγατρικής της εταιρεία που πλέον έχει αναλάβει εξολοκλήρου την μελέτη των αυτόνομων οχημάτων



Εικόνα 2.12

Το πειραματικό αυτόνομο αυτοκίνητο της Google.

Τον Απρίλιο του 2016, η *Volvo* ανακοίνωσε τα σχέδιά της για την ανάπτυξη 100 αυτόνομων οχημάτων τύπου *XC90*, ώστε να τα δοκιμάσει σε καθημερινές συνθήκες οδήγησης στην Κίνα το 2017. Επίσης, τον Απρίλιο του 2016, η αυτοκινητοβιομηχανία ανακοίνωσε τα σχέδιά της για την κυκλοφορία άλλων 100 υβριδικών αυτόνομων οχημάτων τελευταίας τεχνολογίας, τύπου *XC90* στους αυτοκινητόδρομους του Λονδίνου το 2017. Πιο συγκεκριμένα τα 100 *XC90* θα μισθωθούν σε ιδιώτες, και έτσι τα αυτόνομα οχήματα θα καταγράφουν λεπτομερώς τις καθημερινές τους μετακινήσεις άλλα και διάφορες παραμέτρους τους που θα σχετίζονται με αυτές. Σε επόμενο στάδιο, θα αξιολογηθούν όλα τα δεδομένα από την *Volvo*, ώστε να επιτευχθεί η βελτιστοποίηση του οχήματος.



Εικόνα 2.13

Το αυτόνομο αυτοκίνητο XC90 της Volvo.

Τέλος, το πρώτο γνωστό θανατηφόρο για τον οδηγό ατύχημα στο οποίο εμπλέκεται αυτόνομο όχημα, συνέβη στο Γουίλιστον της Φλόριντα στις 7 Μαΐου του 2016, όπου ένα ηλεκτρικό αυτοκίνητο τύπου *Tesla Model S*, λειτουργούσε με αυτόματο πιλότο. Το αυτοκίνητο συγκρούστηκε κάθετα με διερχόμενο 18-τροχη νταλικά και ο οδηγός του *Tesla* σκοτώθηκε ακαριαία. Έτσι, στις 28 Ιουνίου του 2016, ο *NHTSA* κίνησε επίσημη διαδικασία έρευνας σχετικά με τα αίτια του δυστυχήματος και οι εκθέσεις έδειξαν ότι το δυστύχημα σημειώθηκε μετά από απότομο ελιγμό της νταλίκας και αφού το *Tesla* δεν κατάφερε να σταματήσει εγκαίρως.

Τον Αύγουστο του 2016, ιδρύθηκε στη Σιγκαπούρη η πρώτη υπηρεσία μίσθωσης αυτόνομου ταξί, η οποία και παρέχεται από την εταιρεία *nuTonomy*.



Εικόνα 2.14

Το αυτόνομο ταξί της nuTonomy.

Τον Οκτώβριο του 2016, η *Tesla Motors* ανακοίνωσε ότι όλα τα νέα μοντέλα της θα είναι εφοδιασμένα με όλο τον απαραίτητο εξοπλισμό που θα τους επιτρέπει την πλήρως αυτόνομη πλοήγηση σε υψηλό επίπεδο ασφάλειας και αυτονομίας (SAE Level 5). Ωστόσο, λόγω έλλειψης συστήματος εφεδρείας και αισθητήρων ικανών να λειτουργούν αξιόπιστα σε συνθήκες χιονιού, ομίχλης και βροχής, δεν είναι σαφές πώς η *Tesla Motors* θα το επιτύχει αυτό. Ο εξοπλισμός που υλοποιεί την αυτόνομη λειτουργία, αποτελείται από οκτώ στέρεοκάμερες, δώδεκα αισθητήρες προσέγγισης / υπερήχων και εμπρόσθιο ραντάρ που φέρει προηγμένες δυνατότητες επεξεργασίας. Το σύστημα θα λειτουργεί σε "shadow mode" (επεξεργασία χωρίς να αναλαμβάνει δράση) και θα αποστέλλει δεδομένα στους servers της *Tesla Motors*, έτσι ώστε το λογισμικό διαρκώς να βελτιώνεται μέχρι ότου καταστεί έτοιμο για λειτουργία μέσω ασύρματων ενημερώσεων που θα μεταφορτωθούν. Η πλήρης αυτονομία θα είναι εφικτή μόνο μετά από δοκιμές εκατομμυρίων μιλίων και αντίστοιχη έγκριση από τις αρχές. Την ίδια περίοδο η *Tesla Motors* ανακοίνωσε ότι προσβλέπει στην επίτευξη της πλήρους αυτονομίας μέχρι τα τέλη του 2017, ωστόσο αυτό δεν συνέβη.

Στις 4 Ιουνίου 2017, η *Audi* ανακοίνωσε ότι το νέο της μοντέλο *A8* θα είναι πλήρως αυτόνομο για ταχύτητες μέχρι και 60 km/h χρησιμοποιώντας το πρότυπο λογισμικό της *Audi AI*. Σε αντίθεση με άλλα αυτόνομα αυτοκίνητα, ο οδηγός δεν χρειάζεται να πραγματοποιεί ελέγχους ασφαλείας αγγίζοντας το τιμόνι κάθε 15 δευτερόλεπτα. Έτσι, το *Audi A8* θα είναι το πρώτο αυτοκίνητο παραγωγής αυτόνομης οδήγησης επιπέδου 3, κάτι που συνεπάγεται ότι ο οδηγός μπορεί με ασφάλεια να εστιάσει τη προσοχή του σε οτιδήποτε άλλο επιθυμεί πέραν της οδήγησης. Ακόμα, η *Audi* θα είναι ο πρώτος κατασκευαστής που θα χρησιμοποιεί ένα σύστημα 3D LIDAR, πέραν από τις στέρεο κάμερες και τους αισθητήρες υπερήχων του *Audi AI*.



Εικόνα 2.15

Το αυτόνομο αυτοκίνητο A8 της Audi.

Τον Μάρτιο του 2018, ένας πεζός στην Αριζόνα χτυπήθηκε θανάσιμα από ένα αυτόνομο όχημα της *Uber*. Τα αποτελέσματα της έρευνας έδειξαν αυτό ήταν και το πρώτο καταγεγραμμένο θανατηφόρο ατύχημα μεταξύ αυτόνομου οχήματος και πεζού στις Η.Π.Α. Λίγο αργότερα τον ίδιο μήνα, η αστυνομία του Σαν Φρανσίσκο επέβαλε διοικητικό πρόστιμο στον επιβάτη ενός αυτόνομου οχήματος το οποίο δεν κατάφερε να εντοπίσει έναν διερχόμενο πεζό σε μια διασταύρωση. [204], [60], [6], [57], [58], [59]



Εικόνα 2.16

Το αυτόνομο αυτοκίνητο Tesla Model S υπό την αιγίδα της Uber.

3 Η Δομή και οι Λειτουργίες των Αυτόνομων Οχημάτων

3.1 Η Δομή των Αυτόνομων Οχημάτων

Τα αυτόνομα οχήματα αποτελούν συμβατικά οχήματα τα οποία διαθέτουν κοινές μηχανολογικές διατάξεις όμως διαφοροποιούνται με αυτά στο γεγονός ότι είναι εξοπλισμένα επιπλέον με ένα σύστημα αυτονομίας και ρομποτικής. Ένα σύγχρονο αυτόνομο ρομποτικό όχημα εφαρμόζει τεχνολογίες που του παρέχουν δυνατότητα Μηχανικής Όρασης, αυτόματης πλοήγησης και συστημάτων ελέγχου και λήψης αποφάσεων, βασισμένα σε τεχνολογίες Τεχνητής Νοημοσύνης. Συνεπώς όπως γίνεται αντιληπτό, η δομή ενός σύγχρονου αυτόνομου ρομποτικού οχήματος αποτελείται αφενός από το δίκτυο προηγμένων αισθητήρων και τεχνολογικού υλικού που διαθέτει και αφετέρου από το σύνολο των λογισμικών που επιτρέπουν να διαχειρίζεται πιθανοκρατικά και αφαιρετικά τα δεδομένα που προέρχονται από τους αισθητήρες, και να εφαρμόζει προηγμένους αλγόριθμους ώστε παρέχει αξιόπιστα και ασφαλή μοντέλα πρόβλεψης και αποφάσεων στον χειρισμό του οχήματος.

Το προηγμένο δίκτυο αισθητηρίων διατάξεων που φέρει ένα αυτόνομο όχημα, είναι και αυτό που του επιτρέπει να αντιλαμβάνεται τα ερεθίσματα του εξωτερικού περιβάλλοντος, ενώ η μονάδα ελέγχου του οχήματος είναι αυτή που αναλαμβάνει την επεξεργασία των προερχόμενων από τις αισθητήριες διατάξεις δεδομένα, της ερμηνείας τους βάσει συγκεκριμένων αλγόριθμων αναγνώρισης προτύπων και την παραγωγή των λογικών αποφάσεων για τον χειρισμό του οχήματος βασιζόμενες σε τεχνικές και αλγορίθμους Τεχνητής Νοημοσύνης.

Το Αυτόνομο όχημα φέρει ένα πρότυπο δίκτυο αισθητηρίων διατάξεων και μονάδων ελέγχου και λήψης αποφάσεων έτσι ώστε να καταστεί δυνατή η ολοένα και περισσότερο αυτόνομη λειτουργία του.

Παρά το γεγονός ότι κάθε κατασκευάστρια εταιρεία ενός τέτοιου οχήματος, εφαρμόζει και διαφορετικές αισθητήριες και λογικές μονάδες για την ανάπτυξη του συστήματος αυτονομίας, υπάρχει ένα σύνολο καθορισμένων απαιτήσεων που θα πρέπει το σύστημα αυτονομίας να ικανοποιεί και συνεπώς ένας καθορισμένος τύπος δεδομένων τα οποία θα πρέπει να λαμβάνει από το εξωτερικό περιβάλλον του οχήματος. Επομένως, παρά τις μικρές διαφορές που εντοπίζονται στις διατάξεις του υλικού από εταιρεία σε εταιρεία, το μέσο αυτόνομο όχημα καλύπτει το δίκτυο των αισθητηρίων διατάξεων του όπως αυτός που εφαρμόζεται στην παρακάτω εικόνα.



Εικόνα 3.1
Η δομή ενός αυτόνομου οχήματος.
[26]

Τα παραπάνω αισθητήρια συστήματα συνεργάζονται με τα τμήματα λογισμικού της μονάδας ελέγχου ώστε να αποτελέσουν τα συστήματα Μηχανικής Όρασης και πλοήγησης. Το σύστημα Μηχανικής Όρασης, αναλαμβάνει να αντικαταστήσει το ανθρώπινο μάτι και τον ανθρώπινο εγκέφαλο. Είναι υπεύθυνο για την αντίληψη, αναπαράσταση, μοντελοποίηση και ερμηνεία του περιβάλλοντος περιφερειακά του οχήματος, όπως ακριβώς θα έκανε και ένας οδηγός. Το σύστημα πλοήγησης αναλαμβάνει με τη σειρά του την καταχώρηση του επιθυμητού προορισμού, την ανάλυση των δυνατών διαδρομών προσέγγισης του, την επιλογή της βέλτιστης διαδρομής συναρτήσει σταθερών και δυναμικών / μεταβλητών δεδομένων, και εν γένει την πλοήγηση του οχήματος προς τον τελικό προορισμό.

Το σύστημα της Μηχανικής Όρασης και το σύστημα της πλοήγησης, είναι άρρηκτα συνδεδεμένα μεταξύ τους ώστε η πληροφορία που λαμβάνεται και αναλύεται στο σύστημα Μηχανικής Όρασης, να αποτελέσει αναπαράσταση γνώσης και συγκεκριμένου συμβάντος βάσει το οποίου το σύστημα πλοήγησης θα εντολοδοτηθεί για τον κατάλληλο χειρισμό.

Πέραν των παραπάνω, τα σύγχρονα αυτόνομα οχήματα σχεδιάζονται ώστε να περιλαμβάνουν και ένα σύνολο επιπρόσθετου υλικού ή λογισμικού εξοπλισμού όπως είναι τα μαύρα κουτιά καταγραφής δραστηριότητας, οι αλγόριθμοί Τεχνητής Νοημοσύνης ελαχιστοποίησης κυκλοφοριακής αποσυμφόρησης, οι επικοινωνίες μεταξύ διαφόρων οχημάτων κτλ. [11] , [26] , [7] , [29] , [43] , [47] , [48] , [79]

3.1.1 Το Δίκτυο των Αισθητηρίων Διατάξεων

3.1.1.1 RADAR

Το RADAR χρησιμοποιείται στα αυτοκινούμενα οχήματα έτσι ώστε να καταστεί δυνατός ο εντοπισμός διαφόρων αντικειμένων σε αποστάσεις έως 300 μέτρων. Παράλληλα, το RADAR έχει την δυνατότητα να εντοπίζει αντικείμενα ακόμα και κάτω από συνθήκες ομίχλης, υγρασίας, συννεφιάς, βροχής και γενικότερα μειωμένης ορατότητας.

Αποτελεί ένα σύστημα εντοπισμού αντικειμένων, το οποίο χρησιμοποιεί ηλεκτρομαγνητικά κύματα (ραδιοκύματα) ώστε να προσδιορίσει εντός της περιοχής εύρους του, το υψόμετρο, την κατεύθυνση και την ταχύτητα αντικειμένων που βρίσκονται μέσα στην ακτίνα αυτή. Χρησιμοποιείται εκτεταμένα σε αεροσκάφη, πλοία, διαστημόπλοια, βαλιστικούς πυραύλους κ.α.

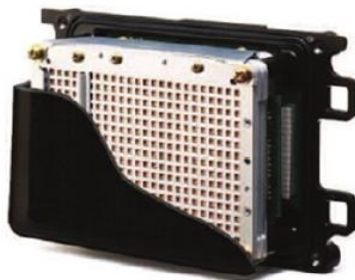
Ένα σύστημα RADAR φέρει έναν πομπό ο οποίος εκπέμπει παλμούς ραδιοκυμάτων ή μικροκυμάτων οι οποίοι προσκρούουν πάνω σε οποιοδήποτε αντικείμενο βρεθεί στην πορεία τους, το οποίο στη συνέχεια απορροφά ένα μικρό ποσοστό της ενέργειας του κύματος και το υπόλοιπο ανακλάται και διαχέεται στο περιβάλλον και κατ' επέκταση στο πιάτο και την κεραία του πομπού. Τα σήματα του RADAR τα οποία ανακλώνται πίσω στον πομπό είναι και αυτά που εξασφαλίζουν τον εντοπισμό του αντικειμένου. Όσο περισσότερο ηλεκτρικά αγωγίμο είναι το αντικείμενο που βρίσκεται εντός του εύρους του RADAR, τόσο καλύτερα είναι και τα αποτελέσματα του εντοπισμού του. Ακόμη, λόγω του φαινομένου Doppler, όσο το αντικείμενο – στόχος κινείται, τόσο μεταβάλλεται και η συχνότητα επιστροφής των ανακλώμενων σημάτων, και έτσι επιτυγχάνεται ο εντοπισμός κίνησης και πορείας του αντικειμένου.

Στις περισσότερες περιπτώσεις, οι δέκτες των RADAR βρίσκονται εγκατεστημένοι στο ίδιο σημείο που βρίσκονται και οι πομποί των σημάτων. Λόγω του ότι τα ανακλώμενα σήματα που λαμβάνονται από τους δέκτες των RADAR πολλές φορές είναι ιδιαίτερα εξασθενημένα, όλα τα RADAR ενσωματώνουν ηλεκτρονικούς ενισχυτές σήματος οι οποίοι και βοηθούν στην ευκολότερη ανάλυση και αναγνώριση των επιστρεφόμενων σημάτων. Η ασθενής απορρόφηση των ραδιοκυμάτων από το μέσο διέλευσης τους, είναι αυτό που επιτρέπει στα RADAR την ανίχνευση αντικειμένων σε σχετικά μεγάλες αποστάσεις στις οποίες άλλα ηλεκτρομαγνητικά μήκη κύματος, όπως το ορατό φως, υπέρυθρο φως, και η υπεριώδης ακτινοβολία είναι πολύ μικρά ώστε να επηρεάσουν. Έτσι, η ομίχλη, τα σύννεφα, η βροχή, η χιονόπτωση, και το χιονόνερο αφήνουν ανεπηρέαστα τα ραδιοκύματα του RADAR και ο εντοπισμός αντικειμένων συνεχίζει να λειτουργεί απροβλημάτιστα. Ακόμη, κατά τη σχεδίαση και κατασκευή συστημάτων RADAR αποφεύγεται η χρήση των ραδιοσυχνοτήτων του φάσματος εκείνου όπου αυτές απορροφώνται

ανακλώνται ή διαχέονται από την υγρασία, τη βροχή και τα αέρια του περιβάλλοντος. Το RADAR βασίζεται αποκλειστικά στις δικές του μεταδόσεις σήματος, και όχι στο ηλιακό φως, στις μικροκυμματικές εκπομπές τρίτων αντικειμένων κτλ.

Τα σύγχρονα συστήματα RADAR είναι σχεδιασμένα να εφαρμόζουν τεχνικές ψηφιακής επεξεργασίας σήματος ώστε να ανιχνεύουν αντικείμενα ακόμα και κάτω από συνθήκες πολύ υψηλών επιπέδων ηλεκτρομαγνητικού θορύβου.

Τα RADAR πλέον χρησιμοποιούνται ευρέως σε μία πληθώρα εφαρμογών όπως αεροπορικές, αστρονομικές, συστήματα αεράμυνας, αντιπυραυλικά συστήματα, συστήματα αποφυγής σύγκρουσης αεροπλάνων, συστήματα αποφυγής σύγκρουσης πλοίων, συστήματα ωκεανογράφησης, συστήματα εποπτείας χώρων, συστήματα ανάλυσης εδάφους, και άλλα. [81], [2], [5], [7], [11]



Εικόνα 3.2

Η μονάδα RADAR ενός αυτόνομου οχήματος.

3.1.1.2 LIDAR

Το σύστημα LIDAR (Light Detection And Ranging) (συχνά αποκαλούμενο και 3D σάρωση με laser), αποτελεί μία τεχνολογία τηλεμετρίας, με την οποία φωτίζοντας ένα αντικείμενο με χρήση παλμικών δεσμών laser, καθίσταται δυνατή η πληροφόρηση μας σχετικά με την θέση του, την απόσταση του και διάφορες άλλες ιδιότητες του, καταγράφοντας και αναλύοντας την οπισθοσκεδαζόμενη ακτινοβολία λέιζερ. Αποτελεί μία ευρέως διαδεδομένη λύσης στις εφαρμογές Μηχανικής Όρασης των αυτόνομων ρομποτικών οχημάτων, λόγω τις ακριβούς πληροφορίας που μπορεί να μας παρέχει, το πλήθος πληροφοριών πεδίου που μας εξασφαλίζει, την ανεπηρέαστη από ηλεκτρομαγνητικό θόρυβο λειτουργία του κ.α.

Το LIDAR χρησιμοποιεί υπεριώδη, ορατή ή υπέρυθη ακτινοβολία για τον εντοπισμό των μίας μεγάλης γκάμας αντικειμένων όπως για παράδειγμα μη-μεταλλικά αντικείμενα, βράχοι, βροχή, χημικές ουσίες, αερολύματα, σύννεφα, ακόμα και μεμονωμένα μόρια. Έτσι μία δέσμη laser δίνει την δυνατότητα για εντοπισμό των φυσικών χαρακτηριστικών ενός αντικείμενου με πολύ υψηλή ανάλυση.

Το LIDAR σχεδιάστηκε με σκοπό την αξιοποίηση του σε εφαρμογές ανάλυσης της ατμόσφαιρας μιας και μέσα από την ανάλυση των οπισθοσκεδαζόμενων σημάτων που προέρχονται από την αλληλεπίδραση των συστατικών της ατμόσφαιρας με την ακτινοβολία λέιζερ, είναι δυνατός ο καθορισμός της κατακόρυφης κατανομής των κυριότερων ρύπων και συστατικών της ατμόσφαιρας με μεγάλη χωρική ακρίβεια (~3-7 μέτρα) και χρονική ακρίβεια (από 10-30 δευτερόλεπτα έως μερικά λεπτά).

Κατά τη λειτουργία του, ο πομπός του LIDAR εκπέμπει μία παλλόμενη ακτίνα λέιζερ πάνω σε έναν καθρέπτη και την προβάλλει κάθετα της του φορέα του το οποίο είναι συνήθως κάποιο εναέριο όχημα. Η laser δέσμη σαρώνει από άκρη σε κατά τη στιγμή εφαρμογής της με ταχύτητα δειγματοληψίας από 20.000 έως 150.000 δείγματα ανά δευτερόλεπτο. Μόλις η δέσμη λέιζερ προσκρούσει σε ένα αντικείμενο αντανακλάται πίσω στον καθρέπτη του δέκτη. Το χρονικό διάστημα που μεσολαβεί από την εκπομπή έως την επιστροφή του laser παλμού, μετριέται και καταγράφεται. Παράλληλα, το GPS καταγράφει με ακρίβεια τις γεωγραφικές συντεταγμένες και το υψόμετρο του οχήματος. Ο αισθητήρας LIDAR συλλέγει έναν τεράστιο όγκο δεδομένων, τέτοιο όπου αν αναλύσουμε μια ενιαία καταγραφή του αισθητήρα, θα

μπορούσαμε καταμετρήσουμε μερικά δισεκατομμύρια σημεία καταγραφής όπου αντιστοιχούν σε αρκετά TB (terabytes) αποθηκευτικού χώρου.

Τα δεδομένα του αισθητήρα LIDAR, στη συνέχεια επεξεργάζονται, αναλύονται και ερμηνεύονται κατάλληλα από εξειδικευμένα για την κάθε χρήση και εφαρμογή λογισμικά. Έτσι, το εξειδικευμένο λογισμικό αναλαμβάνει τον χειρισμό και την ερμηνεία των δεδομένων αυτών, αναλύει τις πολλαπλές επιστροφές ενός εκπεμπόμενου παλμού και του χρονικού διαστήματος μεταξύ εκπομπής και επιστροφών, εφαρμόζει τα όποια προεπιλεγμένα φίλτρα συχνοτήτων και εν τέλει συνθέτει τα παραπάνω για να παράγει την ψηφιακή εικόνα.

Ένα ακόμα πολύ σημαντικό σημείο της τεχνολογίας LIDAR είναι η δυνατότητα επιλεκτικής σάρωσης αντικειμένων, κάτι το οποίο χρησιμοποιείται εκτενώς για την χαρτογράφηση περιοχών, μη λαμβάνοντας υπόψιν τα διάφορα κτίρια, οχήματα, δένδρα, κτλ. Παρόμοιες εφαρμογές της τεχνολογίας αυτής συναντάμε σε ένα ευρύ φάσμα εφαρμογών όπως στη γεωμετρία, στην αρχαιολογία, στη γεωγραφία (συστήματα χαρτογράφησης και DTM), στη γεωλογία, στη γεω-μορφολογία, στη σεισμολογία, στη δασοκομία, στην αστροφυσική, όπως επίσης και στην δορυφορική χαρτογράφηση (ALSM), τις εφαρμογές μέτρησης υψόμετρου και της εφαρμογές υψομετρία λέιζερ και χαρτογράφησης περιγραμμάτων. Το αρκτικόλεξο LADAR (Light Detection And Ranging) εμφανίζεται συχνά στην στρατιωτικές αναφορές. Ο όρος «laser RADAR» μερικές φορές χρησιμοποιείται, ακόμη και αν το LIDAR δεν χρησιμοποιεί μικροκύματα ή ραδιοκύματα και στην πραγματικότητα δεν σχετίζεται με RADAR.

Ακόμα, το LIDAR έχει χρησιμοποιηθεί εκτενώς για την ατμοσφαιρική έρευνα και την μετεωρολογία. Το LIDAR χρησιμοποιείται από τη NASA με σκοπό την ανάπτυξη μίας τεχνολογίας που επιτρέπει στα μελλοντικά αυτόνομα ρομποτικά οχήματα να προσγειωθούν σε διάφορες επιφάνειες με υψηλό βαθμό ακρίβειας και ασφάλειας. Τυπικά σε ένα σύστημα LIDAR χρησιμοποιούνται δέσμες φωτός μήκους κύματος από μερικά μικρόμετρα μέχρι και ακτινοβολίας UV (περίπου 250nm) ώστε να στοχευθεί και αναλυθεί η στόχος.

Αν και το LIDAR έχει χρησιμοποιηθεί εκτενώς σε εναέριες εφαρμογές με την εγκατάσταση του να λαμβάνει χώρα σε αεροπλάνα, ελικόπτερα κτλ, πλέον αποτελεί το βασικότερο αισθητήριο ενός αυτόνομου οχήματος λόγω των ισχυρών πλεονεκτημάτων που φέρει και θα αναλύσουμε παρακάτω. Μοναδική διαφορά του συστήματος LIDAR που φέρουν τα εναέρια οχήματα με αυτό που φέρουν τα αυτόνομα οχήματα εδάφους, είναι ότι το σύστημα LIDAR εδάφους έχει τη δυνατότητα να σαρώνει αντικείμενα με γωνία θέασης των 360°.

Τα σημαντικότερα υποσυστήματα που συνθέτουν ένα σύστημα LIDAR, είναι τα παρακάτω:

1. Laser

Τα laser με μήκος κύματος μεταξύ 600 και 1000 nm, αποτελούν τα πιο κοινά χρησιμοποιούμενα laser σε μη επιστημονικές εφαρμογές. Αποτελούν μία οικονομική λύση, η οποία όμως επειδή διαθέτει μήκος κύματος το οποίο μπορεί εύκολα να απορροφηθεί από τον οφθαλμό και να προκαλέσει βλάβες, διατίθεται σε χαμηλές ισχύς για λόγους ασφαλείας. Η ασφάλεια για το ανθρώπινο μάτι είναι πάντα ένα ζήτημα μείζονος σημασίας. Έτσι, μία εναλλακτική των laser αυτών αποτελούν τα laser στα 1550 nm τα οποία είναι ασφαλή για τον οφθαλμό σε πολύ υψηλότερα επίπεδα ισχύος δεδομένου ότι συγκεκριμένο το μήκος κύματος δεν απορροφάται από το μάτι. Ωστόσο η τεχνολογία τους δεν είναι πολύ ανεπτυγμένη και ως επακόλουθο οι τα συγκεκριμένα μήκη κύματος χρησιμοποιούνται περισσότερο και εφαρμογές που απαιτούν μικρή ακρίβεια και κάλυψη μεγάλων αποστάσεων. Μία ακόμα χρήση των συγκεκριμένων laser είναι οι στρατιωτικές εφαρμογές μίας και τα laser στα 1550 nm δεν είναι ορατά τη νύκτα από τα κιάλια νυχτερινής όρασης, σε αντίθεση με τα laser στα 1000 nm.

2. Σαρωτής εικόνας και οπτικά μέσα

Το πόσο γρήγορα μπορεί να αποτυπωθεί μία εικόνα από το περιβάλλον, στο εσωτερικό του αισθητήρα LIDAR, είναι κάτι άμεσα εξαρτώμενο με την ταχύτητα με την οποία μπορεί το σύστημα να σαρώσει και να αποθηκεύσει τα οπισθοσκεδαζόμενα σήματα. Υπάρχουν διάφορες τεχνικές και εργαλεία που βοηθούν στη σάρωση των τριών διαστάσεων του χώρου, όπως το σύστημα διπλών καθρεπτών του Αναγνώριση πεζών από αυτόνομο ρομποτικό όχημα με χρήση τεχνητής νοημοσύνης και τεχνικών βαθιάς μάθησης

αισθητήρα, ο συνδυασμός ενός ευρυγώνιου φακού, ο σαρωτής δύο καρτεσιανών αξόνων κ.α. Έτσι οι διάφορες τεχνικές που και τα οπτικά μέσα που φέρει ο αισθητήρας LIDAR, έχουν άμεση σχέση και με την ποιότητα και ταχύτητα αποτύπωσης μίας εικόνας από το περιβάλλον.

3. Αισθητήρας εικόνας και μικρο-ηλεκτρονικά του δέκτη

Δύο είναι οι βασικές τεχνολογίες όπου χρησιμοποιούνται στα κυκλώματα των δεκτών των συστημάτων LIDAR: οι solid-state ανιχνευτές εικόνας, όπως οι είναι οι φωτοδίοδοι πυριτίου, ή οι photo-multipliers. Ο βαθμός ευαισθησίας του δέκτη αποτελεί μία ακόμα παράμετρο όπου πρέπει να ληφθεί υπόψιν κατά τη σχεδίαση ενός συστήματος LIDAR.

4. Θέση και συστήματα πλοήγησης

Οι αισθητήρες LIDAR που οι οποίοι βρίσκονται τοποθετημένοι σε οχήματα, απαιτούν την ύπαρξη ενός αισθητηρίου προσδιορισμού της απόλυτης θέσης έτσι ώστε να πραγματοποιούνται οι απαραίτητες αντισταθμίσεις στα μετρούμενα μεγέθη. Για την λειτουργία αυτή, στον αισθητήρα LIDAR ενσωματώνεται κάποιος δέκτης GPS (Global Positioning System). [2], [5], [7], [11], [24], [37], [42], [85]



Εικόνα 3.3

Η μονάδα LIDAR ενός αυτόνομου οχήματος.

3.1.1.3 GPS

Το GPS (Global Positioning System) αποτελεί ένα δορυφορικό σύστημα πλοήγησης (GNSS) το οποίο παρέχει τις γεωγραφικές του συντεταγμένες και τα δεδομένα τοπικής ώρας, κάτω από οποιοσδήποτε καιρικές συνθήκες, σε οποιοδήποτε σημείο της γης, εφόσον υπάρχει ανεμπόδιστη οπτική επαφή με τέσσερις τουλάχιστον δορυφόρους GPS, ώστε να υπολογιστούν σε πραγματικό χρόνο οι ισχύουσες γεωγραφικές συντεταγμένες και τα χρονικά δεδομένα.

Κάθε δορυφόρος μεταδίδει συνεχώς πακέτα δεδομένων τα οποία περιέχουν:

- 1) Το χρονικό στίγμα όπου το πακέτο μηνύματος μεταδόθηκε.
- 2) Ακριβής τροχιακή πληροφορία.
- 3) Την ισχύουσα κατάσταση λειτουργίας και τις τρέχουσες τροχιές όλων των δορυφόρων GPS.

Ο δέκτης χρησιμοποιεί τα μηνύματα που λαμβάνει για τον προσδιορισμό του χρόνου διέλευσης του κάθε μηνύματος, και υπολογίζει την απόσταση με κάθε δορυφόρο. Οι αποστάσεις αυτές μαζί μαζί με τις θέσεις των δορυφόρων που χρησιμοποιούνται, συνυπολογίζονται βάσει τριγωνισμού και δίνουν την γεωγραφική θέση του δέκτη. Στη συνέχεια η θέση αυτή εμφανίζεται ως κινούμενο γεωγραφικό στίγμα σε κάποιον ψηφιακό χάρτη, ενώ παράλληλα μπορούν να παρέχονται και άλλες πληροφορίες πέραν από το στίγμα στον ψηφιακό χάρτη, όπως οι γεωγραφικές συντεταγμένες, το υψόμετρο, η τοπική ώρα κ.α. Ωστόσο πολλές μονάδες GPS παρέχουν και ένα σύνολο άλλων πληροφοριών όπως κατεύθυνση και η ταχύτητα, του δέκτη, η λωρίδα κυκλοφορίας κτλ. Για να επιτευχθούν οι υπολογισμοί αυτών των

πληροφοριών, απαιτούνται τρεις δορυφόροι δειγματοληψίας και ένας αλγόριθμος παρακολούθησης μεταβολών θέσης στον ψηφιακό γεωγραφικό χάρτη. [2], [5], [7], [11], [37], [51]



Εικόνα 3.4

Η μονάδα GPS ενός αυτόνομου οχήματος.

3.1.1.4 Αισθητήρας γωνιακής θέσης

Σαν αισθητήρας θέσης ορίζεται οποιαδήποτε συσκευή που πραγματοποιεί μέτρηση θέσης. Σε μηχανικά περιστροφικά συστήματα όπως είναι ο άξονας μετάδοσης κίνησης στους τροχούς ενός αυτοκινήτου κτλ, υπάρχει η ανάγκη πληροφόρησης σχετικά με την τρέχουσα θέση του περιστρεφόμενου αντικειμένου. Έτσι, αν γνωρίζουμε την εξωτερική περίμετρο μιας ρόδας, τον αριθμό των φορών που έχει περιστρέψει ολοκληρωμένα (κατά 360°) και τον αριθμό των μοιρών που έχει περιστρέψει κατά την τελευταία της περιστροφή (η οποία δεν έχει ολοκληρώσει ακόμα 360°), είναι δυνατός ο και ο υπολογισμός της συνολικής απόστασης που η ρόδα αυτή έχει καλύψει.

Για να επιτευχθεί αυτό, εφαρμόζεται ένας κωδικοποιητής απολύτου θέσεως στον άξονα, ο οποίος παράγει έναν συγκεκριμένο αριθμό παλμών (συνήθως μερικές εκατοντάδες χιλιάδες) ανά πλήρη περιστροφή 360° . Έτσι είναι εύκολα υπολογιζόμενο το ποσοστό μίας μοίρας που αντιστοιχεί σε έναν παλμό και συνεπώς και η διανυθήσα απόσταση που αντιστοιχεί. Οι κωδικοποιητές άξονα απολύτου θέσεως εφαρμόζονται και σε περιπτώσεις που θέλουμε να γνωρίζουμε την ακριβή ταχύτητα περιστροφής του (RPM), αναλύοντας την συχνότητα καταμέτρησης των παλμών που προκύπτουν από την περιστροφή του άξονα.

Οι περιστροφικοί κωδικοποιητές άξονα απολύτου θέσεως, αποτελούν ένα από τα βασικότερα αισθητήρια συστήματα σε εφαρμογές ρομποτικής, βιομηχανικών συστημάτων, φωτογραφικών φακών, ποντίκια trackball, περιστροφικές πλατφόρμες RADAR κ.α. [2], [5], [7], [11], [37], [51]



Εικόνα 3.5

Η μονάδα εντοπισμού γωνιακής θέσης (encoder) ενός αυτόνομου οχήματος.

3.1.1.5 Κάμερα

Η στέρεο κάμερα αποτελεί έναν τύπο κάμερας που φέρει δύο ή και περισσότερους φακούς, καθένας εκ των οποίων διαθέτει ξεχωριστό αισθητήρα εικόνας CMOS. Έτσι, η κάμερα έχει την δυνατότητα να καταγράφει βίντεο και εικόνες τριών διαστάσεων, προσομοιώνοντας έτσι την ανθρώπινη διοπτρική όραση. Η στέρεο κάμερα, χρησιμοποιείται εκτεταμένα για την καταγραφή τρισδιάστατου περιεχομένου, στέρεο-εικόνων υψηλής ανάλυσης κτλ. Η απόσταση μεταξύ των φακών σε μια τυπική στέρεο κάμερα είναι περίπου η απόσταση μεταξύ των ματιών, δηλαδή περίπου 6,35 cm.

Όταν μία στέρεο κάμερα καταγράφει μία τρισδιάστατη εικόνα, ουσιαστικά ο κάθε ένας αισθητήρας της καταγράφει την ίδια εικόνα από ένα διαφορετικό πεδίο λήψης. Στη συνέχεια οι εικόνες αυτές αποθηκεύονται σαν μία συνδυασμένη όπου φέρουν μεταξύ τους κάποια εκατοστά διαφορά προς τα αριστερά ή τα δεξιά, δημιουργώντας ένα θολό ή «κουνημένο» μοτίβο. Σε επόμενο στάδιο, το σύστημα προβολής της εικόνας, της επεξεργάζεται και την προβάλλει με έναν τέτοιο τρόπο ώστε το κάθε ένα μάτι να βλέπει μόνο μία συγκεκριμένη εικόνα από έναν συγκεκριμένο φακό. Έτσι επιτυγχάνεται η αίσθηση των τριών διαστάσεων. Αυτή η μέθοδος έχει προβλήματα με αντικείμενα που κινούνται στις διαφορετικές προβολές, αν και λειτουργεί καλά με νεκρή φύση.

Οι στέρεο κάμερες χρησιμοποιούνται εκτεταμένα σε εφαρμογές αυτόνομων οχημάτων, μίας και είναι απαραίτητη η δυνατότητα καταγραφής και προσδιορισμού του βάθους της εικόνας σε αυτές τις εφαρμογές. Η σύνηθες χρήση των στέρεο καμερών στα αυτόνομα οχήματα, περιλαμβάνει τόσο την αναγνώριση των σημάτων οδικής κυκλοφορίας (φανάρια, οδοσημάνσεις, διαγραμμίσεις κτλ) όσο και την αναγνώριση κινούμενων και στατικών αντικειμένων. Οι σύγχρονες κάμερες που προορίζονται για χρήση σε ένα αυτόνομο όχημα, διαθέτουν φακούς ελεγχόμενης παραμόρφωσης, πολύ υψηλής ευκρίνειας. Πολλές φορές, ένα αυτόνομο όχημα φέρει πολλές κάμερες περιμετρικά του, ώστε μετά από επεξεργασία από κατάλληλο λογισμικό, να παράγεται μια ενιαία – συνδυασμένη εικόνα με γωνία κάλυψης 360°. Η σύνηθες θέση εγκατάστασης της βασικής στέρεο κάμερας ενός αυτόνομου αυτοκινήτου είναι ακριβώς μπροστά από τον καθρέπτη επιτήρησης οπίσθιας κυκλοφορίας. [2], [5], [7], [11], [28], [37], [64]



Εικόνα 3.6

Η μονάδα κάμερας ενός αυτόνομου οχήματος.

3.1.1.6 Αισθητήρας υπολογισμού απόστασης

Οι αισθητήρες υπολογισμού απόστασης ή αλλιώς αισθητήρες υπερήχων, χρησιμοποιούν την ίδια αρχή λειτουργίας με τον αισθητήρα RADAR, με την διαφορά ότι εφαρμόζουν εκπομπές υπερηχητικών κυμάτων υψηλής συχνότητας αντί για μικροκύματα. Οι εκπομπές υπερήχων εκπέμπουν αποτελεσματικά κύματα με συχνότητες πολύ υψηλότερες από αυτές που ακούγονται στο ανθρώπινο αυτί και είναι έτσι κατάλληλες για εφαρμογές μικρής και μεσαίας κλίμακας με χαμηλή ταχύτητα. Οι αισθητήρες απόστασης διαθέτουν έναν πομπό και έναν δέκτη υπερηχητικών κυμάτων οι οποίοι συγκλίνουν μεταξύ τους κατά μία μικρή γωνία κλίσης (συνήθως 15°), ενώ το πεδίο εύρους τους κυμαίνεται από μερικά εκατοστά, έως και μερικές δεκάδες μέτρα, ενώ η ακρίβεια τους ξεπερνάει αυτή του 1mm.

Ο πομπός παράγει υψηλής συχνότητας ηχητικά κύματα, τα οποία εάν προσκρούσουν σε ένα αντικείμενο (εντός του πεδίου εύρους τους), τότε αυτά ανακλώνται προς όλες τις κατευθύνσεις, όπου και ο δέκτης τις λαμβάνει. Μετά την λήψη τους, αναλύονται και βάσει του χρόνου που μεσολάβησε μεταξύ της Αναγνώριση πεζών από αυτόνομο ρομποτικό όχημα με χρήση τεχνητής νοημοσύνης και τεχνικών βαθιάς μάθησης

εκπομπής και της λήψης τους, υπολογίζεται και η απόσταση του εντοπισμένου αντικειμένου. Οι αισθητήρες υπερηχητικών κυμάτων υπολογισμού απόστασης, γνωστοί και ως ultrasonic sensors, έχουν πολύ μεγάλη εφαρμογή σε όλα τα πεδία της βιομηχανίας, της τεχνολογίας και των συστημάτων ασφαλείας.

Οι αισθητήρες μέτρησης απόστασης βρίσκουν εφαρμογή πέραν των άλλων, και στα αυτοκινούμενα οχήματα αφού οι αισθητήρες μπορούν να εντοπίσουν πόσο μακριά είναι ένα εντοπιζόμενο αντικείμενο από το όχημα, και να ειδοποιήσουν τον οδηγό όσο η απόσταση των δύο μειώνεται. Οι συγκεκριμένοι αισθητήρες χρησιμοποιούνται στα αυτόνομα οχήματα αποκλειστικά για εφαρμογές μικρής εμβέλειας, ώστε να λειτουργήσουν ομοίως όπως το σύστημα του RADAR, αλλά προς τις υπόλοιπες κατευθύνσεις του οχήματος. Πολλά σύγχρονα αυτόνομα οχήματα, φέρουν περιμετρικά τους ένα πλήθος από αισθητήρες μέτρησης απόστασης μεγάλης εμβέλειας, οι οποίοι παρέχουν όραση 360° ώστε να είναι δυνατός ο εντοπισμός οποιουδήποτε αντικειμένου και οποιουδήποτε μεγέθους ακόμα και μόλις μερικά χιλιοστά δίπλα από το όχημα. [2], [5], [7], [11], [24], [37], [38]



Εικόνα 3.7

Η μονάδα υπολογισμού απόστασης (ultrasonic) ενός αυτόνομου οχήματος.

3.1.2 Η μονάδα ελέγχου του αυτόνομου οχήματος

3.1.2.1 Η κεντρική μονάδα επεξεργασίας

Τα αυτόνομα οχήματα, πέραν από τους λοιπούς σύγχρονους και ακριβείς αισθητήρες που φέρουν ώστε να λαμβάνουν σε πραγματικό χρόνο τα τρέχοντα πραγματικά δεδομένα του περιβάλλοντος τους, φέρουν και την κεντρική μονάδα επεξεργασίας η οποία επεξεργάζεται τα παραπάνω δεδομένα και σε συνδυασμό με τα λογισμικά και τους αλγόριθμους Τεχνητής Νοημοσύνης που εκτελεί, παράγει τις κατάλληλες αποφάσεις για τον ορθό χειρισμό του οχήματος. Η αρχιτεκτονική που φέρει η κεντρική μονάδα επεξεργασίας των αυτοκινούμενων οχημάτων, είναι παρόμοια με αυτή των σύγχρονων ηλεκτρονικών υπολογιστών.

Τα πρώτα έτη σχεδίασης και δοκιμών των αυτόνομων οχημάτων, συμβατικοί πολύ-πύρινοι και πολύ-νηματικοί επεξεργαστές εφαρμόζονταν στις μονάδες ελέγχου των αυτόνομων οχημάτων. Στη συνέχεια, παρατηρήθηκε εφαρμογή πολλαπλών επεξεργαστών στην μονάδα ελέγχου του αυτόνομου οχήματος, έτσι ώστε να υπάρχει δυνατότητα ανεξάρτητης επεξεργασίας του δεδομένων που προέρχονται από τα αισθητήρια όργανα του οχήματος (δεδομένα εισόδου), και ανεξάρτητη επεξεργασία των δεδομένων πλοήγησης και χειρισμού του οχήματος (δεδομένα εξόδου).

Οστόσο, λόγω της ραγδαία εξελισσόμενης τεχνολογίας των αυτόνομων οχημάτων, του τεράστιου όγκου δεδομένων που χρίζουν ταχύτατης επεξεργασίας, αλλά και λόγω της φύσης της επεξεργασίας του χρίζουν τα δεδομένα αυτά, στην ανωτέρα κεντρική μονάδα επεξεργασίας των δεδομένων εισόδου από τα αισθητήρια συστήματα του οχήματος, ξεκίνησε να εφαρμόζεται παράλληλη αρχιτεκτονική επεξεργασίας και όχι σειριακή. Αυτό συνεπάγεται πως τα σύγχρονα αυτόνομα οχήματα, φέρουν κεντρική μονάδα επεξεργασίας που διαθέτει επεξεργαστές τύπου GPU για τον χειρισμό των δεδομένων των αισθητηρίων οργάνων και επεξεργαστές τύπου CPU για τον έλεγχο του οχήματος.

Ακόμα, μεγάλες εταιρείες του κλάδου παραγωγής των κεντρικών μονάδων επεξεργασίας τύπου CPU και GPU, έχουν αρχίσει να σχεδιάζουν και να παράγουν ολοκληρωμένες υπολογιστικές μονάδες που ειδικά βελτιστοποιημένες για τον έλεγχο των αυτοκινούμενων οχημάτων. Η βελτιστοποίηση αυτή αφορά το γεγονός πως η αρχιτεκτονική του υλικού είναι η βέλτιστη δυνατή όχι μόνο για τις ταχύτητες χειρισμού των δεδομένων, αλλά και για την εκτέλεση των απαραίτητων αλγορίθμων Τεχνητής Νοημοσύνης.

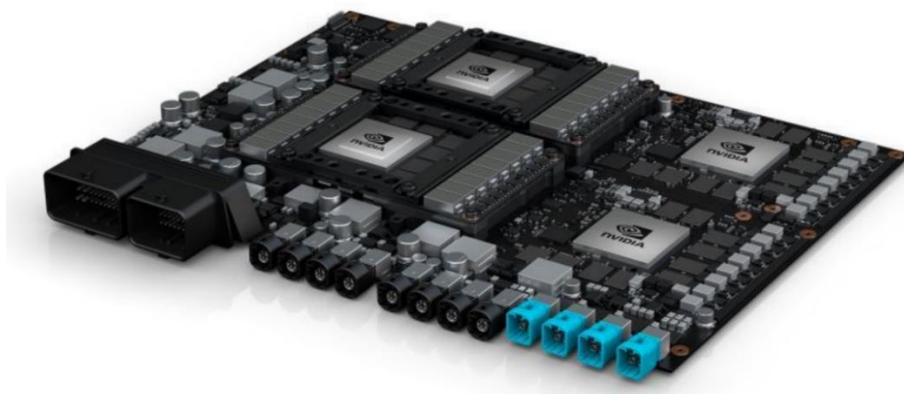
Έτσι, μία σύγχρονη μονάδα επεξεργασίας ειδικά κατασκευασμένη για τον έλεγχο ενός αυτόνομου οχήματος, δύναται τυπικά να φέρει τα κάτωθι χαρακτηριστικά αρχιτεκτονικής και επιδόσεων:

	DRIVE PX Pegasus
SoCs	2x Xavier
Discrete GPUs	2x Post-Volta
CPU Cores	16x NVIDIA Custom ARM
GPU Cores	2x Xavier Volta iGPU & 2x Post-Volta dGPUs
DL TOPS	320 TOPS
FP32 TFLOPS	N/A
TDP	500W

Εικόνα 3.8

Οι επιδόσεις του επεξεργαστή αυτόνομων οχημάτων DRIVE PX Pegasus της NVIDIA.

Οι επεξεργαστές της μονάδας φέρουν τεχνολογία των 16nm, ενώ η ταχύτητα διαύλου μπορεί να προσεγγίσει το 1TB/s. Τα συγκεκριμένα συστήματα είναι ιδιαίτερα αποδοτικά καθώς μπορούν να επιτελέσουν έως 320×10^{11} λειτουργίες deep learning ανά δευτερόλεπτο, κάτι που δίνει την δυνατότητα ανάπτυξης ενός αυτόνομου οχήματος επιπέδου 5 (NHTSA). [7], [11], [12], [37]



Εικόνα 3.9

Ο επεξεργαστής αυτόνομων οχημάτων DRIVE PX Pegasus της NVIDIA.

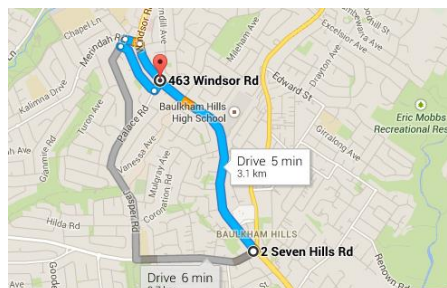
3.1.2.2 Τα λογισμικά της μονάδας ελέγχου

3.1.2.2.1 Λογισμικό πλοήγησης

Το λογισμικό πλοήγησης, αποτελεί ένα από τα βασικότερα συστήματα του αυτόνομου οχήματος και την καρδιά του συστήματος πλοήγησης του οχήματος. Το όχημα, με την χρήση του πομποδέκτη GPS που φέρει, λαμβάνει σε πραγματικό χρόνο την γεωγραφική του θέση την οποία το λογισμικό πλοήγησης αναλαμβάνει να απεικονίσει τους εγκατεστημένους χάρτες πλοήγησης. Εφόσον υπάρχει ομαλή και αδιάλειπτη επικοινωνία του δέκτη GPS με τους δορυφόρους GPS, ο οδηγός / επιβάτης του οχήματος ορίζει τον επιθυμητό προορισμό ώστε να ξεκινήσει η διαδικασία της μετάβασης και η λειτουργία του αυτόνομου οχήματος.

Τα σύγχρονα λογισμικά πλοήγησης, ενσωματώνουν μία πληθώρα εξελιγμένων χαρακτηριστικών και δυνατοτήτων όπως τρισδιάστατη προβολή, υποστήριξη φωνητικών εντολών και καθοδηγήσεων / ενημερώσεων, πληροφορίες σχετικά με τα κοντινότερα σημεία ενδιαφέροντος (νοσοκομεία, υπηρεσίες, πρατήρια υγρών καυσίμων, κοινότητες, καταστήματα κτλ), ενημερώσεις / συμβουλές και αναπροσαρμογή της διαδρομής βάσει στοιχείων από το διαδίκτυο κτλ.

Η πιο αναγνωρίσιμη αντίστοιχη εφαρμογή στις μέρες μας είναι το *Google Maps* το οποίο φέρει πολλές από τις παραπάνω δυνατότητες και αρκετές επιπλέον. Μία πολύ δημοφιλής και χρήσιμη δυνατότητα του λογισμικού *Google Maps*, είναι το *Google Street View*. Το *Google Street View* αποτελεί μια τεχνολογία που προσφέρει στο λογισμικό *Google Maps* πανοραμική και ρεαλιστική εικόνα από σχεδόν όλα τα οδικά δίκτυα του κόσμου (υπεραστικά, αστικά ή επαρχιακά). Το λογισμικό βασίστηκε στην εισαγωγή τρισεκατομμυρίων ενοποιημένων φωτογραφιών (εικονική περιήγηση) που λήφθηκαν από κατάλληλα διαμορφωμένα οχήματα, σε σχεδόν κάθε αυτοκινητόδρομο του κόσμου. [5], [7], [10], [12], [30], [37], [43], [51]



Εικόνα 3.10

Το λογισμικό πλοήγησης Google Maps της Google.

3.1.2.2.2 Αλγόριθμοι Τεχνητής Νοημοσύνης

Η Τεχνητή Νοημοσύνη αποτελεί την τεχνολογία στον χώρο της επιστήμης των υπολογιστών, που προσπαθεί να προσομοιώσει την λειτουργία του ανθρώπινου εγκεφάλου πάνω σε έναν συγκεκριμένο τομέα, για την λύση ενός συγκεκριμένου προβλήματος. Πολλές από τις καθημερινές ενέργειες και πράξεις του ανθρώπου, συμπεριλαμβανομένης της οδήγησης ενός οχήματος, δεν χαρακτηρίζονται από αυστηρή ακολουθιακή λογική ώστε να αντικατασταθούν από ένα κατάλληλα δομημένο αλγόριθμο λογισμικού ηλεκτρονικού υπολογιστή.

Αυτό συμβαίνει γιατί σε πολλές από τις πράξεις του ο άνθρωπος λειτουργεί με βάση τα όσα είδη γνωρίζει από προηγούμενες εμπειρίες, με βάση τα συναισθήματα του, με βάση απρόβλεπτους εξωτερικούς παράγοντες, με βάση το ένστικτο του, με βάση την κοινή λογική και με βάση την ικανότητα του να μαθαίνει κάτι να το κάνει καλύτερα με την πάροδο του χρόνου. Συνεπώς, οι παραπάνω λειτουργίες θα ήταν αδύνατο να αναπαραριστούν αλγοριθμικά σε έναν ηλεκτρονικό υπολογιστή με τη δομή ενός ακολουθιακού τρόπου σχεδίασης.

Όπως θα δούμε και στο κεφάλαιο 5 της παρούσης, ένα σύνολο αλγορίθμων ανθρωπο-εγκεφαλικής προσέγγισης, επιτρέπουν την εν μέρει μίμηση της λειτουργίας του ανθρώπινου εγκεφάλου

Αναγνώριση πεζών από αυτόνομο ρομποτικό όχημα με χρήση τεχνητής νοημοσύνης και τεχνικών βαθιάς μάθησης

από ένα υπολογιστικό σύστημα. Έτσι, με χρήση των αλγορίθμων Τεχνητής Νοημοσύνης, δίνεται η δυνατότητα σε ένα υπολογιστικό σύστημα της μίμησης της λειτουργίας του ανθρώπινου εγκεφάλου ως προς την όραση ενός οδηγού, την σκέψη και τις αποφάσεις ενός οδηγού και τέλος το χειρισμό που ασκεί σε ένα όχημα με βάση τα παραπάνω.

Η κεντρική μονάδα επεξεργασίας κάθε αυτόνομου οχήματος, είναι εξοπλισμένη με ένα σύνολο αλγορίθμων Τεχνητής Νοημοσύνης οι οποίοι επεξεργάζονται με κατάλληλο τρόπο τα δεδομένα που προέρχονται από τις αισθητήριες διατάξεις του οχήματος, ώστε να συνθέσουν μία ολοκληρωμένη εικόνα του εξωτερικού περιβάλλοντος και των συνθηκών του, και βάσει αυτών να ληφθούν οι κατάλληλες αποφάσεις πλοήγησης και χειρισμού του οχήματος και εν συνεχεία αυτές να εφαρμοστούν στο σύστημα ελέγχου των μηχανικών μερών του. Εν γένει, οι αλγόριθμοι Τεχνητής Νοημοσύνης που εκτελούνται σε ένα αυτόνομο όχημα, αποτελούν και τον οδηγό του. [5], [12], [31], [37], [45], [46]



Εικόνα 3.11

Η εικόνα που προκύπτει στην εικόνα που λαμβάνεται από μία κάμερα αυτόνομου οχήματος, μετά την εφαρμογή ψηφιακής επεξεργασίας με χρήση αλγορίθμων αναγνώρισης προτύπων.

3.1.2.2.3 Βοηθητικά λογισμικά

Υπάρχει ένα σύνολο λογισμικών πέραν αυτών της πλοήγησης και της Τεχνητής Νοημοσύνης, που συναντώνται στα αυτόνομα οχήματα και επιτελούν λοιπές χρήσιμες και βοηθητικές λειτουργίες. Τέτοια λογισμικά αφορούν την βελτιστοποίηση λειτουργίας των μηχανικών μερών, τις προστασίες ενδεδειγμένων ενεργειών (service, πλήρωσης καυσίμου), την πληροφόρηση σχετικά με τα αναγνωριζόμενα εξωτερικά αντικείμενα, τα λογισμικά ρυθμίσεων και παραμετροποιήσεων, το λογισμικό που καταγράφει των δεδομένων κουτιού ασφαλείας (μαύρο κουτί), τα λογισμικά εκτέλεσης στατιστικών προβλέψεων και προβολής δεδομένων κ.α. [5], [11], [12]



Εικόνα 3.12

Η κονσόλα ενδείξεων του αυτόνομου οχήματος Tesla Motors Model S.

3.2 Οι Εφαρμοσμένες Τεχνολογίες των Αυτόνομων Οχημάτων

3.2.1 Οι Λειτουργίες πλοήγησης του αυτόνομου οχήματος

Υπάρχει ένα ευρύ φάσμα τεχνολογιών που χρησιμοποιούνται για την ανάπτυξη των αυτόνομων οχημάτων. Τα συστήματα χαμηλότερου επιπέδου αυτονομίας (επίπεδα 0 έως 3) απαιτούν πάντα κάποιου είδους αλληλεπίδραση με τον οδηγό. Σε όλα τα επίπεδα αυτονομίας των αυτόνομων και ημιαυτόνομων οχημάτων, εφαρμόζονται διάφορες τεχνολογίες πλοήγησης του. [20], [66], [68], [78]

3.2.1.1 Συστήματα πλευρικού ελέγχου

Ένα μεγάλο μέρος των συστημάτων πλευρικού ελέγχου, αναλαμβάνει να διατηρήσει το όχημα εντός της επιλεγμένης λωρίδας κυκλοφορίας και εντός του δρόμου. Τα πιο δημοφιλή συστήματα πλευρικού ελέγχου, είναι το *LDWS* (*Lane Departure Warning Systems*), και το *LKA* (*Lane Keeping Assist Systems*). Επίσης, πολλά σύγχρονα αυτόνομα οχήματα, φέρουν και το νέο σύστημα παράλληλης στάθμευσης *PAP* (*Parallel Assist Parking*). [20]

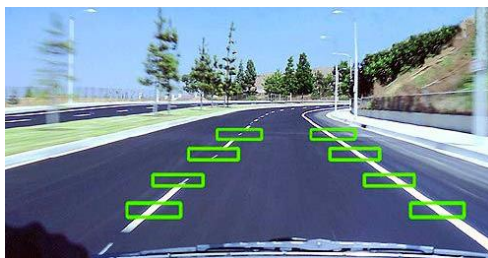
3.2.1.1.1 Σύστημα LDWS

Σκοπός του συστήματος *LDWS* (*Lane Departure Warning System*), είναι η διατήρηση του οχήματος εντός του δρόμου, η διατήρηση του εντός της λωρίδας κυκλοφορίας και η αποφυγή πλευρικών συγκρούσεων με άλλα οχήματα ή αντικείμενα. Για να γίνει αυτό, το όχημα πρέπει να έχει την δυνατότητα να αντιληφθεί τα όρια της λωρίδας του, τα όρια του δρόμου και την ακριβή του θέση πάνω σε αυτόν. Υπάρχουν διάφορες μέθοδοι για την υλοποίηση της παραπάνω μεθόδου, όπως ενσωματωμένοι μαγνητικοί δείκτες στο δρόμο, ή πολύ ακριβείς αισθητήρες GPS και ψηφιακοί χάρτες που εφαρμόζουν ψηφιακή επεξεργασία εικόνας.

Η πιο αξιόπιστη μέθοδος ταυτοποίησης λωρίδας, είναι αυτή των ενσωματωμένων μαγνητικών δεικτών στο δρόμο, οι οποίοι εκπέμπουν ένα μαγνητικό πεδίο το οποίο δύναται να εντοπιστεί από αισθητήρες του οχήματος. Ωστόσο η υλοποίηση του παραπάνω εγχειρήματος είναι ιδιαίτερα δύσκολη και πολύπλοκη, καθώς θα απαιτούσε την ενσωμάτωση κατάλληλων μαγνητών σε όλους τους υφιστάμενους και μελλοντικούς δρόμους. Μία άλλη λύση θα μπορούσε να ήταν ο συνδυασμός πολύ υψηλής ακρίβειας αισθητήρων GPS και αντίστοιχα διαμορφωμένους ψηφιακούς χάρτες.

Μία άλλη μέθοδος η οποία επικρατεί για την υλοποίηση του παραπάνω, είναι η χρήση μίας κάμερας και σε επόμενο στάδιο η ψηφιακή επεξεργασία της εικόνας ώστε να μοντελοποιηθούν και αναγνωριστούν οι λωρίδες στον δρόμο, τα όρια αυτού καθώς και η θέση του οχήματος. Ωστόσο, η μέθοδος αυτή αντιμετωπίζει προβλήματα σε περιπτώσεις όπου τα όρια των λωρίδων και του δρόμου δεν είναι ευδιάκριτα λόγω χιονιού, βροχόπτωσης, μειωμένου φωτισμού κτλ. Στις περιπτώσεις αυτές συνήθως εφαρμόζονται αλγόριθμοι αναγνώρισης των λωρίδων του δρόμου ακόμα και κάτω από αυτές τις συνθήκες.

Τέλος, μία ακόμη μέθοδος αναγνώρισης των λωρίδων, είναι η εφαρμογή αισθητήρων υπέρυθρης ακτινοβολίας στο κάτω μέρος του οχήματος. Έτσι ο αισθητήρας αναλύει τις εισερχόμενες ανακλάσεις και αντιλαμβάνεται την ύπαρξη λωρίδας. [39]

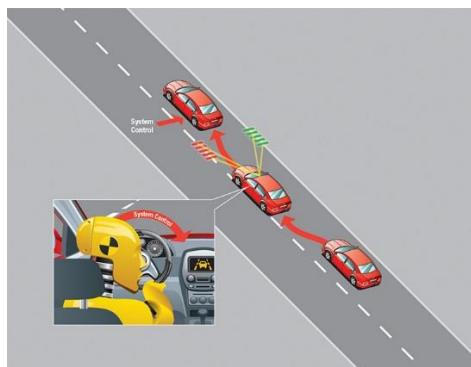


Εικόνα 3.13

Αναγνώριση των λωρίδων κυκλοφορίας από κάμερα συστήματος LDWS. Αναγνώριση πεζών από αυτόνομο ρομπωτικό όχημα με χρήση τεχνητής νοημοσύνης και τεχνικών βαθιάς μάθησης

3.2.1.1.2 Σύστημα LKA

Το Σύστημα Υποβοήθησης Διατήρησης Λωρίδας ή *LKA (Lane Keeping Assist System)* χρησιμοποιείται ως έτσι ώστε να πραγματοποιεί μικρές κινήσεις στο τιμόνι του οχήματος, ώστε να βοηθήσει το όχημα να παραμείνει εντός λωρίδας. Αρκετά σύγχρονα συστήματα αξιοποιούν της πληροφορίες που προέρχονται από τα αισθητήρια συστήματα (RADAR, LIDAR, ultrasonics κτλ) ώστε να υπολογίζουν με ακρίβεια την απαιτούμενη ροπή στρέψης ανά χρονική στιγμή, αναλόγως των εξωτερικών συνθηκών (τύπος οδοστρώματος, κλίση οδοστρώματος, άνεμοι κτλ). Μάλιστα λόγω της ευκολίας που παρέχει το σύστημα κατά τη διάρκεια ενός ταξιδιού, αρκετοί οδηγοί αποσπούσαν την προσοχή τους από τον χειρισμό του οχήματος με αποτέλεσμα να εκτεθούν σε κίνδυνο. Για το λόγω αυτό το σύστημα αναπροσαρμόστηκε έτσι ώστε όταν δεν έχει λάβει ανθρώπινο χειρισμό για κάποιο χρονικό διάστημα, να ενεργοποιεί αντίστοιχες ηχητικές ειδοποιήσεις. Μια εταιρεία, η *IrisBus*, ανέπτυξε το σύστημα CIVIS, το οποίο εφαρμόζει επεξεργασία εικόνας στις ειδικές σημάσεις στο δρόμο. Η κάμερα παρακολουθεί μια διπλή διακεκομμένη γραμμή και επεξεργάζεται τις πληροφορίες στο σύστημα LKA ώστε να κατευθύνει το τιμόνι ανάλογα. [39]

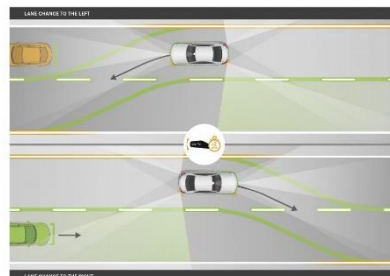


Εικόνα 3.14

Το σύστημα LKS κατά την διαδικασία υποβοήθησης διατήρησης λωρίδας κυκλοφορίας.

3.2.1.1.3 Σύστημα LCA

Το σύστημα *LCA (Lane Change Assist)* ή *ALCA (Active Lane Change Assist)*, βοηθάει τον οδηγό να αλλάξει λωρίδα κυκλοφορίας όποτε το χρειαστεί, αποφεύγοντας το κίνδυνο σύγκρουσης με άλλα οχήματα. Για την αποφυγή συγκρούσεων, χρησιμοποιούνται οπτικές και ηχητικές ειδοποιήσεις προς τον οδηγό, ενώ παράλληλα εφαρμόζονται μικρή ένταση πέδηση και αλλαγή κατεύθυνσης του οχήματος. Το σύστημα αποτελείται από δύο RADAR τα οποία βρίσκονται τοποθετημένα στις δύο γωνίες το εμπρόσθιου προφυλακτήρα του οχήματος. [39]

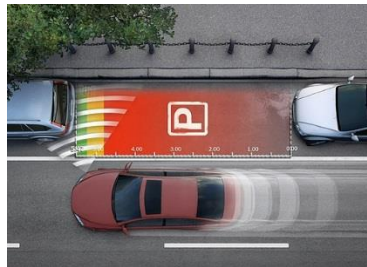


Εικόνα 3.15

Το σύστημα LCA κατά την διαδικασία υποβοήθησης αλλαγής λωρίδας κυκλοφορίας.

3.2.1.1.4 Σύστημα PAP

Το σύστημα *PAP (Parallel Assist Parking)* εντοπίζει και προτείνει στον οδηγό διαθέσιμες θέσεις για στάθμευση. Έπειτα, ο οδηγός υποδεικνύει στο σύστημα την θέση στην οποία επιθυμεί να σταθμεύσει, και η διαδικασία στάθμευσης ξεκινάει. Τότε, ο οδηγός βλέπει από την κάμερα οπισθοπορείας ώστε να μπορεί να ελέγξει καλύτερα στην πέδηση και την επιτάχυνση του αυτοκινήτου. Το σύστημα *Parallel Assist Parking* αναλαμβάνει τον χειρισμό του τιμονιού ώστε να ολοκληρωθεί με επιτυχία και ευκολία η διαδικασία της στάθμευσης. [39]



Εικόνα 3.16

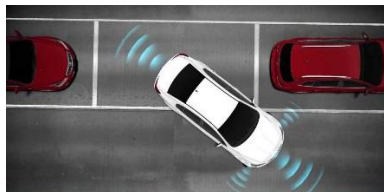
Το σύστημα PAP κατά την διαδικασία εντοπισμού θέσης στάθμευσης.

3.2.1.2 Σύστημα διαμήκη ελέγχου

Το σύστημα διαμήκη ελέγχου περιλαμβάνει τις εμπρόσθιες και οπίσθιες κατευθύνσεις του οχήματος. Διαθέτει εφαρμογές για τον έλεγχο της ταχύτητας του οχήματος και την βοηθάει τον οδηγό σχετικά με τις ενέργειες εμπρόσθιας και οπίσθιας οδήγησης. Αυτές οι εφαρμογές περιλαμβάνουν το σύστημα οπίσθιου ελέγχου στάθμευσης, το *ACC (Adaptive Cruise Control)* και το *pre-crash break assist*. [39]

3.2.1.2.1 Σύστημα υποβοήθησης οπίσθιου παρκαρίσματος

Υπάρχουν πολλά συστήματα που βοηθούν τον οδηγό στην στάθμευση. Το πιο συνηθισμένο είναι σύστημα αισθητήρων υπερήχων το οποίο ενσωματώνει ανιχνευτές υπερήχων στον πίσω προφυλακτήρα του αυτοκινήτου, οι οποίοι δίνουν στον οδηγό έναν μεταβαλλόμενο ακουστικό ήχο καθώς η απόσταση μειώνεται. Το μειονέκτημα του συστήματος αυτού είναι ότι οι αισθητήρες αυτοί λειτουργούν μόνο σε πολύ μικρές αποστάσεις, μερικών μόλις εκατοστών. Τα νεότερα συστήματα περιλαμβάνουν κάμερες οπισθοπορείας που εμφανίζουν μία εικόνα με την πίσω οπτική του οχήματος στην οποία έχει εφαρμοστεί ένα μοτίβο κάθετων γραμμών οδηγών ώστε να κατευθύνουν τον οδηγό στη σωστή διεύθυνση και τον ειδοποιούν ώστε να σταματήσει μόλις σταθμεύσει το όχημα επιτυχώς. Ακόμα, πολλά άλλα συστήματα φέρουν αισθητήρες *RADAR* στο εμπρός ή / και στο πίσω τμήμα του οχήματος, οι οποίοι είναι πολύ πιο ακριβείς αλλά αυξάνουν κατά πολύ το κόστος του συστήματος. [39]



Εικόνα 3.17

Το σύστημα υποβοήθησης οπίσθιου παρκαρίσματος, κατά την διαδικασία στάθμευσης.

3.2.1.2.2 Σύστημα ACC

Το σύστημα ACC (*Adaptive Cruise Control*) χρησιμοποιείται για να βοηθήσει τον οδηγό, ελέγχοντας την ταχύτητα με την οποία κινείται το όχημα σε σχέση με τα οχήματα που βρίσκονται μπροστά του. Υπάρχουν διάφορα συστήματα ACC, όπως το ACC υψηλής ταχύτητας, το ACC χαμηλής ταχύτητας και το πλήρες ACC. Υπάρχουν διάφοροι αισθητήρες μπορούν να χρησιμοποιηθούν για την σύνθεση ενός συστήματος ACC.

Ο πιο συχνά χρησιμοποιούμενος αισθητήρας για το σύστημα ACC, είναι το RADAR. Το RADAR χρησιμοποιείται ώστε να εντοπιστεί η απόσταση από το προπορευόμενο όχημα καθώς επίσης ο ρυθμός μεταβολής της ανάμεσα τους απόστασης. Έτσι, το σύστημα ρυθμίζει την ταχύτητα του οχήματος ώστε να διατηρήσει μια καθορισμένη απόσταση από το προπορευόμενο όχημα. Στην περίπτωση όπου κάποιο άλλο όχημα βρεθεί στην λωρίδα, το ACC θα ρυθμίσει την ταχύτητα ώστε να διατηρηθεί η απαιτούμενη απόσταση ασφαλείας. Επίσης, υπάρχουν και συστήματα ACC τα οποία χρησιμοποιούν αισθητήρα LIDAR.

Το ACC υψηλής ταχύτητας προορίζεται για οδικά δίκτυα με κεντρικούς αυτοκινητόδρομους και μεγάλες αποστάσεις. Το ACC χαμηλής ταχύτητας προορίζεται για χρήση σε περιοχές με υψηλή κυκλοφοριακή συμφόρηση. Το συγκεκριμένο σύστημα δεν θα σταματήσει τελείως το όχημα, αλλά αν η ταχύτητα φτάσει στην ελάχιστη δυνατή, τότε θα προειδοποιήσει τον οδηγό για το αντικείμενο ώστε να σταματήσει αυτός το όχημα.

Το πλήρες ACC αποτελεί έναν συνδυασμό ACC υψηλής και χαμηλής ταχύτητας. Το σύστημα αφού ενεργοποιηθεί, μπορεί να προσαρμόσει την συμπεριφορά του ως υψηλής ή χαμηλής ταχύτητας, αναλόγως της κυκλοφορίας του οχήματος σε κάποιο αυτοκινητόδρομο ή η αστικό κέντρο αντίστοιχα. Αυτό προσφέρει στον οδηγό να μην ελέγχει την ταχύτητα του οχήματος του παρά μόνο την κατεύθυνση του και το εξωτερικό περιβάλλον. [39]



Εικόνα 3.18

Το σύστημα ACC κατά τη διαδικασία ελέγχου της ταχύτητας του οχήματος, σε έναν αυτοκινητόδρομο.

3.2.1.3 Πλήρως αυτόνομα συστήματα

Προκειμένου να επιτευχθεί ένα πλήρως αυτόνομο όχημα, πρέπει να υπάρξει ομαλή ενσωμάτωση τόσο των πλευρικών όσο και των διαμηκών συστημάτων ελέγχου, καθώς και ο απαραίτητος συγχρονισμός μεταξύ τους. Τα σύγχρονα αυτόνομα οχήματα, βρίσκονται ακόμη σε ερευνητικό στάδιο και ο βαθμός αυτονομίας τους δεν τους επιτρέπει να κινηθούν με ταχύτητες πολύ μεγαλύτερες από αυτές που επιτρέπονται σε έναν αστικό αυτοκινητόδρομο. Ακόμα, πέραν την εξέλιξης των συστημάτων Αυτό θα παρείχε στο όχημα ACC, το επόμενο βήμα για την αυτόνομη οδήγηση είναι η ανάπτυξη των δικτύων επικοινωνίας των αυτόνομων αυτοκινήτων ώστε να καταστεί δυνατή η βέλτιστη συμπεριφορά τους στους δρόμους συναρτήσει των εξωτερικών μη ορατών παραγόντων και συνθηκών. [39]

3.3 Οι Τεχνολογίες Επικοινωνιών των Αυτόνομων Οχημάτων

3.3.1 Η ανάγκη εφαρμογής συστημάτων επικοινωνιών στα αυτόνομα οχήματα

Αν και η τεχνολογία των αυτόνομων οχημάτων αναπτύσσεται ραγδαία, είναι κατά γενική ομολογία γεγονός πως η εποχή όπου το σύνολο των συμβατικών αυτοκινήτων θα αντικατασταθεί με πλήρως αυτόνομα, θα αργήσει. Αν και οι αλγόριθμοι ελέγχου των αυτόνομων οχημάτων προσφέρουν ολοένα και μεγαλύτερη ασφάλεια για τον έλεγχο του οχήματος, η τεχνολογία του αυτόνομου οχήματος αυτού καθ' αυτού δεν είναι και ο μόνος τομέας που θα πρέπει να λάβουμε υπόψιν μας προς διερεύνηση και ανάπτυξη.

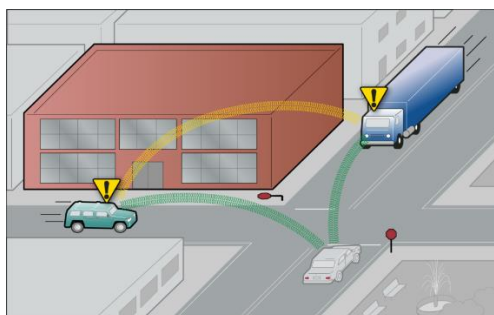
Ο διευθύνων σύμβουλος της *Tesla Motors*, Elon Musk, προβλέπει ότι το 90% όλων των οδηγών θα μπορούσε να αυτοματοποιηθεί μέχρι το 2017, αλλά η πλήρης αυτοματοποίηση δεν είναι εφικτή λόγω της πολυπλοκότητας που συνεπάγεται η διαμόρφωση των αυτοματοποιημένων ενεργειών για κάθε πιθανή κατάσταση. Ακόμα, υπάρχει ένα σύνολο γεγονότων που καθιστά απαραίτητη την ύπαρξη τεχνολογίας, δυνατότητα και υποδομής ενημέρωσης ενός αυτόνομου οχήματος σχετικά με ένα συμβάν. Προκύπτει λοιπόν η ανάγκη ύπαρξης μίας υποδομής κεντρικού ελέγχου και επικοινωνιών σε ένα μελλοντικό περιβάλλον καθολικής αυτόνομης οδήγησης. [20], [21], [53], [71], [75]

3.3.2 Οι τεχνολογίες επικοινωνίας μεταξύ οχημάτων και υποδομών

Ο διεθνής Κ.Ο.Κ. υποχρεώνει τον οδηγό ενός οχήματος να ενημερώνει μέσω κατάλληλων φωτεινών διατάξεων (φλας) τους υπολοίπους οδηγούς σχετικά με την πρόθεση του να κατευθυνθεί το όχημα προς μία συγκεκριμένη κατεύθυνση. Αυτό είναι υποχρεωτικό επειδή όσο πιο συντονισμένοι είναι οι οδηγοί, τόσο ασφαλέστερη θα είναι και η κυκλοφορία. Επίσης, ο καλύτερος συντονισμός μεταξύ των επιμέρους οδηγών καθιστά την κυκλοφορία πιο αποτελεσματική συνολικά, όσον αφορά τόσο την κατανάλωση ενέργειας και την ταχύτητα μεταφοράς. Ακόμα, το ίδιο το οδικό δίκτυο ενημερώνει τον οδηγό σχετικά με την προτεραιότητα κίνησης (φωτεινοί σηματοδότες), τις τρέχουσες οδικές και κυκλοφοριακές συνθήκες και κανονισμούς (οδικές σημάσεις), και σχετικά με έκτακτες εξελίξεις (ενημερωτικές φωτεινές πινακίδες προβολής μηνυμάτων). Συνεπώς απαιτείται η ύπαρξη ενός σύγχρονου συστήματος επικοινωνίας τόσο μεταξύ των διερχόμενων οχημάτων που βρίσκονται στην ίδια περιοχή, όσο και μεταξύ του κάθε οχήματος με τις οδικές και κυκλοφοριακές υποδομές. [20], [21], [53], [68]

3.3.3 Η τεχνολογία V2V

Η τεχνολογία οχήματος προς όχημα V2V (*Vehicle to Vehicle*), η οποία επιτρέπει στα οχήματα να ανταλλάσσουν πληροφορίες μεταξύ τους, αναμένεται να αυξήσει σημαντικά την ποσότητα και την ποιότητα των πληροφοριών ή των δεδομένων που μπορούν να μοιράζονται μεταξύ των οχημάτων. Εάν για παράδειγμα, ένα όχημα που συνδέεται μέσω V2V περάσει από ένα ατύχημα, θα μεταδώσει αυτές τις πληροφορίες σε άλλα αυτοκίνητα της περιοχής, προειδοποιώντας τους αποτελεσματικά ώστε να επιστήσουν την προσοχή τους. [9], [20], [25], [53]



Εικόνα 3.19

Ένα δίκτυο επικοινωνίας αυτόνομων οχημάτων τεχνολογίας V2V.

3.3.4 Η τεχνολογία V2I

Η τεχνολογία οχήματος προς υποδομές V2I (*Vehicle to Infrastructure*), επιτρέπει στα οχήματα όχι μόνο να ανταλλάσσουν πληροφορίες με άλλα οχήματα (ως έμμεσος κόμβος ανακατεύθυνσης πληροφορίας), αλλά και επιτρέπει και την επικοινωνία των υποδομών με το όχημα. Έτσι ένα τέτοιο παράδειγμα επικοινωνίας θα μπορούσε να ήταν η ενημέρωση του οχήματος για την κατάσταση του φωτεινού σηματοδότη, των οδικών σημάτων των έκτακτων κυκλοφοριακών ρυθμίσεων και ενημερωτικών μηνυμάτων κτλ. Η συγκεκριμένη τεχνολογία θα μπορούσε εν δυνάμει να οδηγήσει σε εντελώς νέες δυνατότητες στα οδικά δίκτυα όπως ολική απαλοιφή των φωτεινών σηματοδοτών, οδικών σημάτων και ενημερωτικών πινακίδων (στη περίπτωση καθολικής αντικατάστασης των συμβατικών οχημάτων με αυτόνομα / ευφυή αυτοκίνητα). Άλλες χρήσιμες δυνατότητες που προσφέρει η τεχνολογία αυτή είναι η άμεση αλλαγή των κυκλοφοριακών ρυθμίσεων, η απαλοιφή του φαινομένου κυκλοφοριακής συμφόρησης μέσω κατάλληλου αλγορίθμου ελέγχου της ταχύτητας των οχημάτων, ο περιορισμός και έλεγχος των οχημάτων ώστε να μην παραβιάζουν τον Κ.Ο.Κ. , η άμεση εξυπηρέτηση έκτακτων καταστάσεων ή ακόμα και ο κεντρικός έλεγχος πλοήγησης των οχημάτων από ένα υπέρ-δίκτυο δικτύων V2I. [20], [25]

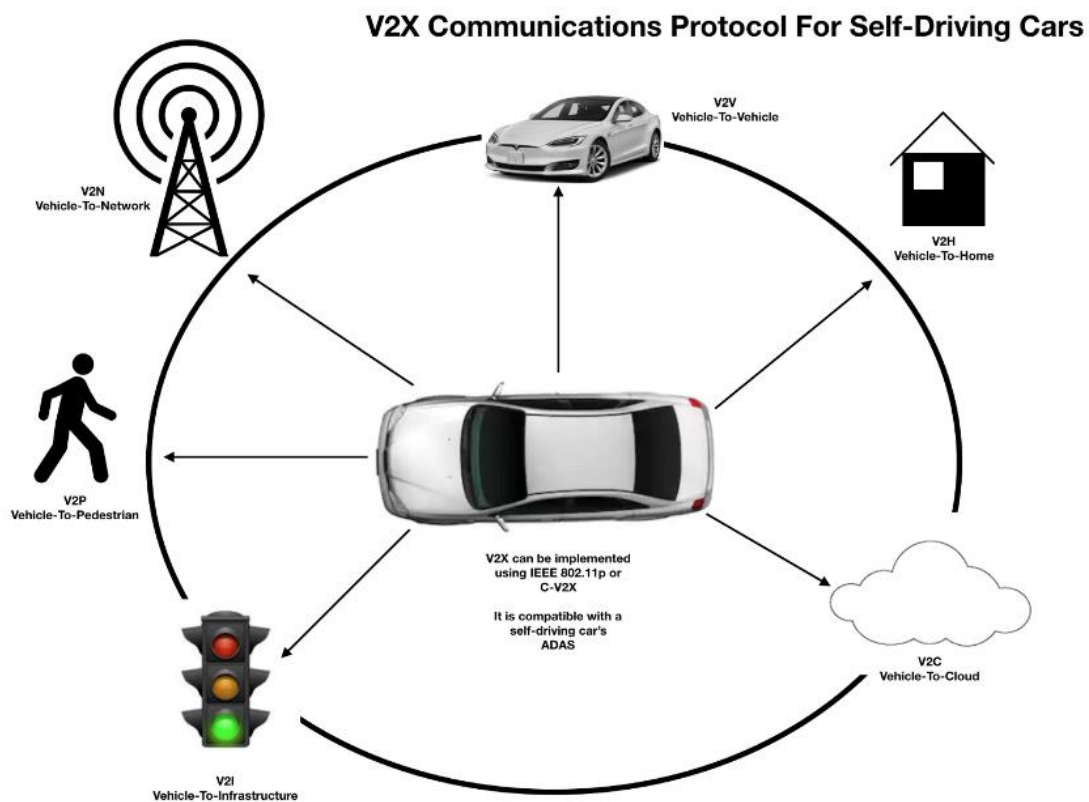


Εικόνα 3.20

Ένα δίκτυο επικοινωνίας αυτόνομων οχημάτων τεχνολογίας V2I.

3.3.5 Η έννοια του V2X

Ο όρος V2X (*Vehicle to Everything*) είναι ένας γενικός, ο οποίος περιλαμβάνει τόσο την έννοια του V2V όσο και αυτή του V2I. Επίσης και άλλες πολλές αντίστοιχες τεχνολογίες ενδέχεται να αναπτυχθούν όπως οι V2P (*Vehicle to Pedestrian*), V2D (*Vehicle to Device*) και V2G (*Vehicle to Grid*). Η τεχνολογία V2X αναμένεται να αποφέρει σημαντικά οφέλη στην κυκλοφορία. Μόλις συμβεί κάτι αναπάντεχο, όπως ένα ατύχημα ή ένα ζώο που προσεγγίζει το δρόμο, όλα τα αυτοκίνητα που βρίσκονται σε ακτίνα 300 μέτρων θα ενημερωθούν για αυτό και, ως εκ τούτου, θα μπορέσουν να λάβουν τα κατάλληλα μέτρα. Το πιο σημαντικό είναι ότι η τεχνολογία V2X θα βοηθήσει τους οδηγούς αυτοκινήτων να αποφύγουν τα ατυχήματα, αν είναι σε θέση να «βλέπουν» άλλα οχήματα όπου ο άνθρωπος δεν μπορεί (για παράδειγμα όταν η οπτική περιορίζεται από ένα κτίριο). Δημιουργώντας ένα συλλογικό δίκτυο όπου όλα τα αυτοκίνητα μοιράζονται πληροφορίες μεταξύ τους και μεταξύ των υποδομών, η τεχνολογία V2X είναι πιθανό να καταστήσει την συμβατική κυκλοφορία, καθιστώντας την ασφαλέστερη και πιο αποτελεσματική αναφορικά τόσο με τον χρόνο των μετακινήσεων, όσο και την κατανάλωση ενέργειας. Τα δίκτυα V2X σχεδιάζονται ώστε να υλοποιούνται εντός κάποιας WLAN επικοινωνίας. Ωστόσο με την ταυτόχρονη ραγδαία ανάπτυξη του IoT και την σχεδίαση και εφαρμογή των ασύρματων δικτύων επικοινωνίας τύπου 5G, ενδέχεται να δοθούν και άλλες δυνατότητες υλοποίησης των δικτύων V2X καθώς και νέες δυνατότητες εξέλιξης τους. [21], [25]



Εικόνα 3.21

Ένα δίκτυο επικοινωνίας αυτόνομων οχημάτων τεχνολογίας V2X.

4 Η Τεχνητή Νοημοσύνη και η Μηχανική Μάθηση

4.1 Εισαγωγή στη Τεχνητή Νοημοσύνη

Παρά την ραγδαία ανάπτυξη των σύγχρονων υπολογιστικών συστημάτων, και των πλεονεκτημάτων ασφάλειας και ταχύτητας διαχείρισης της πληροφορίας που απορρέουν από την ένταξη τους στη σύγχρονη καθημερινότητα, ο ανθρώπινος εγκέφαλος παραμένει ισχυρότερο εργαλείο λήψης αποφάσεων καθώς φέρει τη δυνατότητα να επεξεργάζεται νέα δεδομένα και να παράγει νέους κανόνες ώστε να λαμβάνει αποφάσεις για γεγονότα που είναι πρωτοφανή και μη καθορισμένα. Η δυνατότητα αυτή του ανθρώπινου εγκεφάλου ονομάζεται νοημοσύνη και αποτελεί ένα από τα ισχυρότερα πλεονεκτήματα του ανθρώπινου εγκεφάλου έναντι της προγραμματιζόμενης λογικής που φέρει ένα ψηφιακό υπολογιστικό σύστημα.

Έτσι στις αρχές τις δεκαετίας του 80', οι επιστήμονες θέλησαν να αναπτύξουν μία τεχνολογία στους ηλεκτρονικούς υπολογιστές η οποία θα τους έδινε τη δυνατότητα να μιμούνται την συμπεριφορά του ανθρώπινου εγκεφάλου και εν δυνάμει να τον αντικαθιστούν σε μία πληθώρα αυτοματοποιημένων και μη διαδικασιών. Παρά τις επίμονες προσπάθειες και τους ανυπολόγιστους πόρους που επενδύθηκαν διεθνώς στην ανάπτυξη της τεχνολογίας αυτής, κάτι τέτοιο δεν έχει επιτευχθεί μέχρι και σήμερα καθώς στην πραγματικότητα η λειτουργία του ανθρώπινου εγκεφάλου είναι ιδιαίτερα πολύπλοκη και δεν έχει αποκρυπτογραφηθεί μέχρι και σήμερα.

Οστόσο, το σύνολο των προσπαθειών αυτών έχει οδηγήσει στην ανάπτυξη πρότυπων τεχνολογιών λογισμικού οι οποίες βασίζονται στην μίμηση της συλλογιστικής του ανθρώπινου εγκεφάλου και στην λήψη αποφάσεων για νέα προβλήματα, οδηγούμενη από συνδυαστική λογική των κανόνων και των γεγονότων που έχουν καταγραφεί από το σύστημα. Έτσι δίνεται η δυνατότητα σε ένα σύστημα, έχοντας καθορισμένους τους κανόνες και τα γεγονότα μίας κατάστασης, να παρακολουθεί τα νέα δεδομένα που παράγονται καθώς και τα νέα γεγονότα που προκύπτουν, να τα αναλύει μαθηματικά, να τα εντάσσει σε νέα μοντέλα και έτσι να αναπροσαρμόζει την συμπεριφορά του βάση αυτής της συνδυαστικής λογικής. Η συγκεκριμένη τεχνολογία μίμησης της ανθρώπινης συλλογιστικής ονομάστηκε, και όχι αδικώς, Τεχνητή Νοημοσύνη - TN (Artificial Intelligence - AI).

Η Τεχνητή Νοημοσύνη αποτελεί έναν ιδιαίτερο και πολυσύνθετο διεπιστημονικό κλάδο ο οποίος πέρα από την πληροφορική εμπλέκει και άλλες λοιπές επιστήμες για την ολοκληρωμένη ανάπτυξη ενός συστήματος όπως Μαθηματικά, Φυσικές Επιστήμες, Μηχανική (Ηλεκτρολογία, Μηχανολογία, Ηλεκτρονική, Αυτοματισμός), Βιολογία, Φιλοσοφία, Ψυχολογία, Γλωσσολογία, κ.α.

Όπως είναι γνωστό, ο ανθρώπινος εγκέφαλος προκειμένου να επεξεργαστεί τα δεδομένα που λαμβάνει και να λάβει τις απαραίτητες αποφάσεις αξιολογεί τόσο τις τρέχουσες συνθήκες, καταστάσεις και δεδομένα όσο και τις υφιστάμενες εμπειρίες και τη γνωσιακή του βάση, όπως αυτή έχει προκύψει από την ακατάπαυτη διαδικασία της μάθησης. Η επεξεργασία όλων των παραπάνω δεδομένων, η λειτουργία της μάθησης καθώς και η λήψη των αποφάσεων διενεργούνται από τους κατάλληλους νευρώνες του ανθρώπινου εγκεφάλου.

Ακριβώς με παρόμοιο τρόπο, τα σύγχρονα υπολογιστικά συστήματα Τεχνητής Νοημοσύνης, προχωρούν στη συλλογή των απαραίτητων δεδομένων από το περιβάλλον, εφαρμόζουν τις αρχές αναγνώρισης προτύπων και Μηχανικής Μάθησης ώστε να διαχειρίζονται βέλτιστα τα δεδομένα αυτά και μαθαίνουν από τις υφιστάμενες καταστάσεις και τέλος χρησιμοποιούν ένα εκτεταμένο δίκτυο τεχνητών νευρώνων για να καταστεί δυνατή η υλοποίηση των προαναφερθέντων λειτουργιών. [10], [82], [84], [197], [199], [200], [201]

4.1.1 Βασικοί ορισμοί για την Τεχνητή Νοημοσύνη

Ο ορισμός της νοημοσύνης και της Τεχνητής Νοημοσύνης, αποτελεί επιστημονικό αντικείμενο έρευνας φιλοσοφικού περιεχομένου αφού η ίδια η έννοια της νοημοσύνης είναι πολυδιάστατη, ασαφής και χαώδης όπως επίσης και ο ανθρώπινος εγκέφαλος. Εκατοντάδες επιστήμονες ανά τις προηγούμενες δεκαετίες έχουν προσπαθήσει να ορίσουν τις δύο αυτές συγγενείς έννοιες, διαμορφώνοντας κάθε φορά

και μία νέα οπτική θεώρησης και οδηγώντας στην ανάπτυξη ενός πολυδιάστατου μοντέλου θεώρησης την ανθρώπινη νοημοσύνης. [82], [84], [197], [199], [200], [201]

4.1.2 Η έννοια της νοημοσύνης

Για να αντιληφθούμε ορθά την έννοια της Τεχνητής Νοημοσύνης, θα πρέπει πρώτα να έχουμε ορίσει επαρκώς την έννοια της νοημοσύνης. Ωστόσο, έχουν δοθεί εκατοντάδες ορισμοί για την έννοια της νοημοσύνης, οι οποίοι πραγματεύονται τις ίδιες έννοιες, χωρίς όμως να έχουμε κάποιον επικρατέστερο ορισμό.

Σύμφωνα με τον D. Wechsler, ως νοημοσύνη ορίζεται ως «Μία γενική και σύνθετη ικανότητα του ατόμου να ενεργεί με βάση του σκοπούς που θέτει, να σκέπτεται με λογικό τρόπο και να ανταποκρίνεται στις εκάστοτε απαιτήσεις του κοινωνικού και φυσικού περιβάλλοντος που το περιστοιχίζει».

Ενώ σύμφωνα με τον Douglas Hofstadter (βραβείο Pulitzer), ως νοημοσύνη ορίζεται η ικανότητα του ανθρώπου να:

- Ανταποκρίνεται σε καταστάσεις με ελαστικότητα (όχι μηχανική συμπεριφορά).
- Κατανοεί τα ασαφή ή αντιφατικά μηνύματα από τα συμφραζόμενα.
- Αναγνωρίζει και να ιεραρχεί τα διάφορα δεδομένα με βάση τη σπουδαιότητα τους.
- Βρίσκει ομοιότητες μεταξύ καταστάσεων οι οποίες μοιάζουν διαφορετικές.
- Βρίσκει διαφορές μεταξύ καταστάσεων οι οποίες μοιάζουν παρόμοιες.

Οι παραπάνω ικανότητες έχουν τον παρακάτω κοινό μεταξύ τους χαρακτηριστικό: αποκτώνται εύκολα από τους ανθρώπους και βασίζονται συνήθως σε ένα σύνολο σταθερών και στερεότυπων απόψεων/γνώσεων που κατέχει οποιοσδήποτε άνθρωπος και αποκαλείται κοινή λογική (common sense).

Ακόμα, πέραν από την προσπάθεια ορισμού της πολύπλοκης έννοιας της νοημοσύνης, ο Howard Gardner διακρίνει τους κάτωθι 8 τύπους νοημοσύνης:

- Γλωσσική
- Λογική και Μαθηματική
- Μουσική
- Χωρική
- Σωματική
- Διαπροσωπική
- Ενδοπροσωπική
- Φυσιολογική

Οι παραπάνω τύποι νοημοσύνης, βάσει του Howard Gardner, υπάρχουν σε κάθε άνθρωπο ανεξαιρέτως, όμως σε διαφορετικό βαθμό ανεπτυγμένη η κάθε μια σε κάθε άνθρωπο, κάτι που συντελεί στην μοναδική ανθρώπινη φύση και προσωπικότητα του κάθε ανθρώπου.

Είναι άμεσα αντιληπτό πως δεν υπάρχει σαφής τοποθέτηση σχετικά με την έννοια της νοημοσύνης, ενώ παράλληλα όλοι κατανοούμε την βαθύτερη φύση και έννοια της. Ακόμα, είναι ευρέως αντιληπτά τα αντικειμενικά χαρακτηριστικά γνωρίσματα της νοημοσύνης, και αυτά είναι:

- A) Η ικανότητα της συλλογιστικής και συνειρμικής σκέψης.
- B) Η ικανότητα αναγνώρισης προτύπων και μοντελοποίησης δεδομένων.
- Γ) Η ικανότητα μάθησης.
- Δ) Η ικανότητα συνδυαστικής λογικής.

Ωστόσο, η λειτουργία της νοημοσύνης βασίζεται σε τρεις δυνατότητες των έμβιων όντων, όπου χωρίς αυτές δεν θα ήταν δυνατή η λειτουργία της, και αυτές είναι:

- A) Η ικανότητα των αισθήσεων.
- B) Η ικανότητα της απομνημόνευσης.
- Γ) Η ικανότητα επιλογής της βέλτιστης λύσης, βασιζόμενη σε μη αντικειμενικούς παράγοντες.

Επομένως, κατανοούμε πως η έννοια της νοημοσύνης είναι ένα αντικείμενο φιλοσοφικής αναζήτησης δεκάδων αιώνων που έχει σαν αρχή της την εξερεύνηση του τρόπου λειτουργίας του ανθρώπινου εγκεφάλου ως προς τα ερεθίσματα που λαμβάνει ώστε να παράγει μοναδικές και λογικές αποφάσεις, στηριζόμενος τόσο σε αξιολόγηση των τρεχουσών αντιλαμβανόμενων ερεθισμάτων, όσο και σε προηγούμενες εμπειρίες ή γνώσεις που κατέχει. [197], [199], [200], [201]

4.1.3 Η έννοια της Τεχνητής Νοημοσύνης

Ο όρος Τεχνητή Νοημοσύνη αναφέρεται στην τεχνολογία ανάπτυξης λογισμικού η οποία αποσκοπεί να μοντελοποιήσει, να αντιγράψει και να εφαρμόσει την ανθρώπινη λογική και συμπεριφορά σε ένα συγκεκριμένο πρόβλημα που χρήζει επίλυσης. Παρότι συνήθως πιστεύεται ότι η επιστήμη της Τεχνητής Νοημοσύνης αποτελεί μία νέα επιστήμη, στην πραγματικότητα οι ρίζες της χρονολογούνται από τα μέσα του προηγούμενου αιώνα. Αυτό συνέβη το 1956 σε μία συνάντηση επιφανών επιστημόνων της εποχής μεταξύ των οποίων οι John McCarthy, Marvin Minsky, Claude Shannon, κ.α. , τέθηκε για πρώτη φορά το ζήτημα της τεχνητής μίμησης της ανθρώπινης νοημοσύνης. Τα λογισμικά που ενσωματώνουν αλγορίθμους Τεχνητής Νοημοσύνης, συνήθως αποσκοπούν στην εφαρμογή μίας ή και περισσότερων εκ των τεσσάρων ακόλουθων ανθρώπινων νοητικών γνωρισμάτων.

- Μίμηση της ανθρώπινης σκέψης
- Μίμηση της ανθρώπινης συμπεριφοράς
- Μίμηση της ανθρώπινης λογικής και συλλογιστικής
- Μίμηση των ανθρωπίνως αντιδράσεων

Έτσι, ένας συμπληρωματικός ορισμός που θα μπορούσαμε να δώσουμε για την Τεχνητή Νοημοσύνη είναι πως ορίζεται ως ο τομέας της Επιστήμης των Υπολογιστών που ασχολείται με τη σχεδίαση και την υλοποίηση προγραμμάτων τα οποία είναι ικανά να μιμηθούν τις ανθρώπινες γνωστικές ικανότητες, εμφανίζοντας έτσι χαρακτηριστικά που αποδίδουμε συνήθως σε ανθρώπινη συμπεριφορά, όπως η επίλυση προβλημάτων, η αντίληψη μέσω της όρασης, η μάθηση, η εξαγωγή συμπερασμάτων, η κατανόηση φυσικής γλώσσας, κλπ.

Η Τεχνητή Νοημοσύνη έχει μία πληθώρα περιοχών εφαρμογής στις μέρες μας οι οποίες μέρα με την μέρα αυξάνονται εκθετικά. Μερικές μόνο από τις καθιερωμένες περιοχές εφαρμογής της Τεχνητής Νοημοσύνης είναι:

- Μοντέλα Προβλέψεων (καιρικών, οικονομικών, κοινωνικών, πολιτικών κτλ)
- Επεξεργασία Φυσικής Γλώσσας (αυτόματοι τηλεφωνητές, chat bots, κτλ)
- Σχεδιασμός Ενεργειών και Χρονοπρογραμματισμός (προηγμένα συστήματα παραγωγής)
- Προσαρμοζόμενα και εξελισσόμενα ευφυή συστήματα (ευφυής διαχείριση πληροφορίας)
- Αυτόνομα Ρομποτικά Οχήματα (self-driving cars, quadcopters swarm, κτλ)
- Έμπειρα Συστήματα και Συστήματα Γνώσης (βέλτιστη διαχείριση γνώσης)
- Ευφυείς πράκτορες (agents)
- Τεχνητή Όραση (αναγνώριση και μοντελοποίηση φυσικών χώρων, αντικειμένων κτλ)
- Ευφυείς υπηρεσίες διαδικτύου και σημασιολογικό διαδίκτυο (semantic web)
- Αυτόνομες Ρομποτικές Εφαρμογές (εφαρμογές βιομηχανίας, εργαστηριακών δοκιμών κτλ)
- Απόδειξη Θεωρημάτων (πολύπλοκα θεωρήματα υψηλής μεταβλητότητας δεδομένων)
- Επίλυση γενικής Φύσεως προβλημάτων (σύγχρονα προβλήματα υψηλής αβεβαιότητας)

Ο Alan Turing (1913-1954) ο οποίος θεωρείται ο πατέρας της Τεχνητής Νοημοσύνης, διατύπωσε το 1950 μία δοκιμασία (Turing test), για την αναγνώριση ευφύων μηχανών. Η δοκιμασία περιλαμβάνει ένα σύνολο ερωτημάτων τα οποία τίθενται παράλληλα σε έναν άνθρωπο και σε ένα υπολογιστικό σύστημα. Τα ερωτήματα αυτά εμπλέκουν ένα ευρύ φάσμα επιστημονικών πεδίων όπως επεξεργασία φυσικής γλώσσας, αναπαράσταση γνώσης, αυτοματοποιημένη συλλογιστική, Μηχανική Μάθηση, Αναγνώριση πεζών από αυτόνομο ρομποτικό όχημα με χρήση τεχνητής νοημοσύνης και τεχνικών βαθιάς μάθησης

αναγνώριση εικόνων και αντικειμένων, κα. Στόχος της δοκιμασίας είναι η αναγνώριση του υπολογιστικού συστήματος και του ανθρώπου. Σε περίπτωση που η αναγνώριση του υπολογιστικού συστήματος δεν επιτευχθεί, τότε το υπολογιστικό σύστημα αναγνωρίζεται ως ευφυές.

Μία από τις πιο σύνθετες εφαρμογές της Τεχνητής Νοημοσύνης στα σύγχρονα υπολογιστικά περιβάλλοντα, είναι όπως διατυπώθηκε και προηγουμένως, η εξομίωση της ανθρώπινης συλλογιστικής, κάτι που συχνά συναντάται και ως Υπολογιστική Νοημοσύνη. Η Υπολογιστική Νοημοσύνη κατ' ουσία αποσκοπεί στην έρευνα και εφαρμογή των ανθρώπινων νοητικών και γενετικών λειτουργιών, σε ένα υπολογιστικό σύστημα. Σε αυτή τη κατηγορία της Τεχνητής Νοημοσύνης συναντάμε δύο από τους πιο μελετημένους και δημοφιλείς τομείς της Τεχνητής Νοημοσύνης, τους Γενετικούς Αλγόριθμους και τα Νευρωνικά Δίκτυα. Οι Γενετικοί Αλγόριθμοι και τα Νευρωνικά Δίκτυα αποτελούν δύο από τα ισχυρότερα εργαλεία της Τεχνητής Νοημοσύνης τα οποία μας παρέχουν τις μεθόδους και τις τεχνικές εφαρμογής των αρχών Τεχνητής Νοημοσύνης στα υπολογιστικά συστήματα.

Όπως είναι αντιληπτό, ουσιαστικά το αντικείμενο της Τεχνητής Νοημοσύνης είναι η επίλυση προβλημάτων των οποίων η φύση χαρακτηρίζεται από ιδιαίτερη αβεβαιότητα και των οποίων τα δεδομένα φέρουν υψηλό βαθμό μεταβλητότητας. Έτσι ένα συμβατικό πληροφοριακό σύστημα θα αδυνατούσε να προσφέρει λύση σε ένα τόσο σύνθετο, πολυμεταβλητό και δυναμικό περιβάλλον και η μόνη δυνατότητα επίλυσης θα ήταν η ανθρώπινη συλλογιστική, η οποία έχει τη δυνατότητα να συνδυάζει τα παρόντα δεδομένα, τη κεκτημένη γνώση και εμπειρία καθώς και την «κοινή αίσθηση» ώστε να προχωρήσει στην παραγωγή μίας λύσης. Το κενό αυτό έρχεται να καλύψει η επιστήμη της Τεχνητής Νοημοσύνης η οποία αν και βρίσκεται ακόμα πάρα πολύ μακριά από την πλήρη μοντελοποίηση και εξομίωση του ανθρώπινου εγκεφάλου, έχει ήδη κάνει με απόλυτη επιτυχία τα πρώτα της βήματα στην τμηματική μίμηση ορισμένων λειτουργιών του ανθρώπινου εγκεφάλου. [82], [84], [197], [199], [200], [201]

4.2 Μηχανική Μάθηση

4.2.1 Ορισμός μάθησης και Μηχανικής Μάθησης

Η **Μάθηση** αποτελεί μία από τις σημαντικότερες ικανότητες του ανθρώπου. Είναι το εργαλείο που του επιτρέπει, πέραν των άλλων, να απομνημονεύει και να χρησιμοποιεί δεδομένα που λαμβάνει από το περιβάλλον ώστε να βελτιώνεται σταδιακά κατά την εκτέλεση μιας λειτουργίας. Αποτελεί ένα από τα σημαντικότερα χαρακτηριστικά της νοημοσύνης και μπορούμε να αναλογιστούμε την σπουδαιότητα της αν εξετάσουμε τη συμβολή της σε νοητικές λειτουργίες της καθημερινότητας όπως η ομιλία, η κατανόηση της γλώσσας, η αναγνώριση αντικειμένων και προσώπων, η προσαρμογή της συμπεριφοράς κ.α.

Όπως γίνεται αντιληπτό, όσο σπουδαίος είναι ο ρόλος της μάθησης για την ανθρώπινη νοημοσύνη, άλλο τόσο σπουδαία είναι η Μηχανική Μάθηση για την Τεχνητή Νοημοσύνη. Αυτό συμβαίνει διότι η παραδοσιακή Τεχνητή Νοημοσύνη η οποία δεν βασίζεται σε τεχνικές Μηχανικής Μάθησης αλλά σε κανόνες και γεγονότα μόνο, παρουσιάζει αδυναμία προσαρμογής σε νέα δεδομένα και εν γένει σε πολύ χαμηλή προσαρμοστικότητα και ευελιξία. Έτσι, η Μηχανική Μάθηση αποτελεί ένα αναπόσπαστο κομμάτι των εφαρμογών Τεχνητής Νοημοσύνης. Σύμφωνα λοιπόν με τα παραπάνω, θα μπορούσαμε να ορίσουμε την Μηχανική Μάθηση ως *το φαινόμενο κατά το οποίο ένα σύστημα βελτιώνει την απόδοσή του κατά την εκτέλεση μιας συγκεκριμένης εργασίας, χωρίς να υπάρχει ανάγκη να προγραμματιστεί εκ νέου.*

Όστόσο, για να καταστεί εφικτή η διαδικασία της μάθησης, είναι απαραίτητη ύπαρξη ενός συστήματος μοντελοποίησης της καταγραφόμενης γνώσης. Για το λόγο αυτό, στα σύγχρονα συστήματα ευφυούς λειτουργίας, εφαρμόζονται αλγόριθμοι αναγνώρισης προτύπων. Ως Αναγνώριση Προτύπων ορίζεται το επιστημονικό πεδίο που έχει σκοπό την κατάταξη αντικειμένων σε κατηγορίες οι οποίες ονομάζονται κλάσεις. Κατά την αρχική σχεδίαση ενός τέτοιου ευφυούς συστήματος, μοντελοποιούνται και καταχωρούνται τα βασικά πρότυπα, ως προς τα χαρακτηριστικά που τα διέπουν. [199], [200], [201]

4.2.2 Τα είδη της μάθησης

Ο όρος **Επαγωγική Μάθηση** (Inductive Learning), αναφέρεται στο είδος της μάθησης όπου επιτυγχάνεται δια μέσου της διαδικασίας της **επαγωγής** (induction). Σε αυτό το είδος της μάθησης, ο άνθρωπος μαθαίνει κατανοώντας το περιβάλλον του μέσω παρατηρήσεων, δημιουργώντας έτσι μία απλοποιημένη εκδοχή αυτού που ονομάζεται **νοητικό μοντέλο** (mental model), ενώ νοητικά πρότυπα ονομάζονται οι δομές που μπορεί να δημιουργήσει ο ανθρώπινος εγκέφαλος βασιζόμενος στη οργάνωση και συσχέτιση των εμπειριών και παρατηρήσεων του.

Ένα ακόμη είδος μάθησης το οποίο εφαρμόζεται επίσης από τον τομέα της Μηχανικής Μάθησης, είναι η **Εξελικτική Μάθηση** (Evolutionary Learning), η οποία χρησιμοποιεί κατά κύριο λόγο γενετικούς αλγόριθμους, και εφαρμόζεται συνήθως σε προβλήματα βελτιστοποίησης.

[199] , [200] , [201]

4.2.3 Οι κατηγορίες της Μηχανικής Μάθησης

Στον τομέα της Μηχανικής Μάθησης επικρατούν τρεις κατηγορίες μάθησης, οι οποίες διαφοροποιούνται με βάση της λειτουργία του μοντέλου μάθησης του αλγόριθμου. Οι τρεις αυτές κατηγορίες μάθησης είναι:

- **Επιβλεπόμενη Μάθηση** (Supervised Learning): Κατά τη διαδικασία αυτή, το σύστημα (ο αλγόριθμος) καλείται να «μάθει» μία έννοια ή μία συνάρτηση από ένα σύνολο δεδομένων, η οποία αποτελεί περιγραφή ενός μοντέλου. Το μοντέλο αυτού του τύπου μάθησης, χρησιμοποιείται κυρίως για επίλυση προβλημάτων:
 - Ταξινόμησης (Classification)
 - Πρόγνωσης (Prediction)
 - Διερμηνείας (Interpretation)
- **Μη Επιβλεπόμενη Μάθηση** (Unsupervised Learning): Κατά τη διαδικασία αυτή, το σύστημα (ο αλγόριθμος) καλείται να ανακαλύψει μόνο του συσχετίσεις ή ομάδες σε ένα σύνολο δεδομένων, δημιουργώντας το ίδιο πρότυπα, χωρίς να είναι γνωστό εάν υπάρχουν και πόσα είναι αυτά. Το μοντέλο αυτού του τύπου μάθησης, χρησιμοποιείται κυρίως σε προβλήματα:
 - Ανάλυσης Συσχετισμών (Association Analysis) , και
 - Ομαδοποίησης (Clustering)

Άλλες Μορφές Μάθησης

Πέραν των μεθόδων μάθησης που εντάσσονται στις κατηγορίες της επιβλεπόμενης και μη επιβλεπόμενης μάθησης, όπως αυτές παρουσιάστηκαν παραπάνω, υπάρχουν και μέθοδοι μάθησης οι οποίες δεν εντάσσονται στις δύο κατηγορίες. Οι μέθοδοι αυτές αποτελούν διαφορετικές προσεγγίσεις στο θέμα της Τεχνητής Νοημοσύνης και ενίοτε εφαρμόζονται είτε ως μορφές επιβλεπόμενης μάθησης είτε ως μορφές μη επιβλεπόμενης μάθησης. [199] , [200] , [201]

4.2.4 Η Επιβλεπόμενη Μάθηση

Στην επιβλεπόμενη μάθηση, ορίζουμε ως συνάρτηση στόχο (target function) τη συνάρτηση που αποτελεί το αντικείμενο εκμάθησης από το σύστημα. Η συνάρτηση στόχος χρησιμοποιείται για την πρόβλεψη της τιμής μιας μεταβλητής, η οποία ονομάζεται εξαρτημένη μεταβλητή ή μεταβλητή εξόδου, βάσει των τιμών ενός συνόλου μεταβλητών, όπου ονομάζονται ανεξάρτητες μεταβλητές, μεταβλητές εισόδου ή χαρακτηριστικά.

Η επαγωγική μάθηση στηρίζεται στην "υπόθεση επαγωγικής μάθησης" (inductive learning hypothesis), βάσει της οποίας κάθε υπόθεση που προσεγγίζει καλά τη συνάρτηση στόχο για ένα αρκετά μεγάλο σύνολο παραδειγμάτων, θα προσεγγίζει το ίδιο καλά τη συνάρτηση στόχο και για περιπτώσεις που δεν έχει εξετάσει.

Στην μάθηση με επίβλεψη διακρίνονται δυο είδη προβλημάτων (learning tasks), τα προβλήματα ταξινόμησης και τα προβλήματα παρεμβολής. Τα προβλήματα ταξινόμησης (classification) αφορούν στη δημιουργία μοντέλων πρόβλεψης διακριτών τάξεων (κλάσεων/κατηγοριών). Τα προβλήματα παρεμβολής (regression) αφορούν τη δημιουργία μοντέλων πρόβλεψης αριθμητικών τιμών.

Στις εφαρμογές επιβλεπόμενης Μηχανικής Μάθησης, χρησιμοποιούνται ορισμένα συγκεκριμένα εργαλεία και μέθοδοι εφαρμογής του μοντέλου μάθησης στο σύστημα. Τα βασικότερα πρότυπα μάθησης που ανήκουν στο μοντέλο της επιβλεπόμενης μάθησης, είναι τα κάτωθι:

- 1) Η Μάθηση Εννοιών
- 2) Τα Δένδρα Απόφασης
- 3) Η Μάθηση Κανόνων Ταξινόμησης
- 4) Η Μάθηση κατά Περίπτωση
- 5) Η Μάθηση κατά Bayes
- 6) Παρεμβολή ή Ταξινόμηση
- 7) Τεχνητά Νευρωνικά Δίκτυα
- 8) Μηχανές Διανυσμάτων Υποστήριξης [201]

4.2.4.1 Η Μάθηση Εννοιών (Concept Learning)

Η «έννοια» αποτελεί ένα υποσύνολο αντικειμένων ενός μεγαλύτερου συνόλου.

Κατά την μάθηση των εννοιών, αρχικά το σύστημα ενημερώνεται με παραδείγματα που ανήκουν ή δεν ανήκουν στη συγκεκριμένη έννοια. Στη συνέχεια σχεδιάζεται το αφαιρετικό μοντέλο της έννοιας αυτής ώστε να καταστεί δυνατός ο έλεγχος ταιριάσματος αυτής της έννοιας με κάποια νέα περίπτωση.

Ο Αλγόριθμος Απαλοιφής Υποψηφίων (AAY) - Candidate Elimination Algorithm, αποτελεί τον πιο δημοφιλή αλγόριθμο απαλοιφής μάθησης εννοιών. Η βασική του λειτουργία έγκειται:

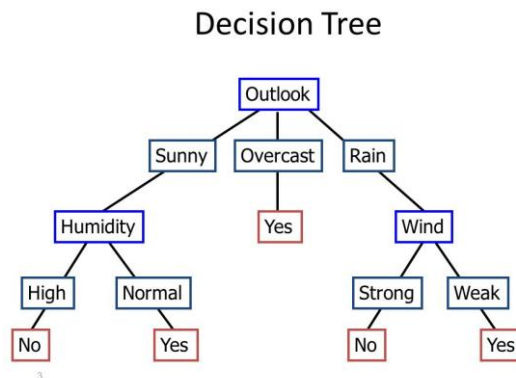
1. Στον περιορισμό του χώρου αναζήτησης επιτελώντας γενικεύσεις ή εξειδικεύσεις σε κάποιες αρχικές έννοιες βάσει των δεδομένων εκπαίδευσης.
2. Στην διατήρηση δύο συνόλων G και S , όπου περιγράφουν το χώρο αναζήτησης. Το σύνολο G περιγράφει το σύνολο των πιο γενικών υποψήφιων εννοιών ενώ το σύνολο S περιγράφει το σύνολο των πιο εξειδικευμένων υποψήφιων εννοιών. [201]

4.2.4.2 Τα Δένδρα Απόφασης / Ταξινόμησης

Τα Δένδρα Απόφασης / Ταξινόμησης αποτελούν μία δενδροειδή σχηματικά δομή η οποία περιγράφει με γραφικό τρόπο τα δεδομένα. Χρησιμοποιούνται ώστε προβλέψουν με ακρίβεια τη τιμή που φέρει μία μεταβλητή που μοντελοποιούν, βάσει των τιμών των χαρακτηριστικών.

Η δομή αυτή αποτελείται από:

1. Τον κόμβο ο οποίος ορίζει την συνθήκη ελέγχου της τιμής κάποιου χαρακτηριστικού των περιπτώσεων.
2. Το κλαδί το οποίο φεύγει από ένα κόμβο αντιστοιχεί σε μια διαφορετική διακριτή τιμή του χαρακτηριστικού που σχετίζεται με τον κόμβο και,
3. Τα κλαδιά-φύλλα στα οποία έχουμε το τι συνέβη.



Εικόνα 4.1

Δέντρο απόφασης σε εφαρμογή καιρικής πρόβλεψης.

[201]

A) Αναπαράσταση με Κανόνες:

Το δέντρο αποφάσεων μπορεί να αναπαρασταθεί και με χρήση των συνθηκών if-then (**κανόνες ταξινόμησης - classification rules**).

B) Ο Αλγόριθμος ID3:

Ο αλγόριθμος ID3 αποτελεί τον πιο γνωστό αλγόριθμο μάθησης δένδρων ταξινόμησης, είναι αναδρομικός και η γενική του μορφή έχει ως εξής:

1. Βρες την ανεξάρτητη μεταβλητή η οποία αν χρησιμοποιηθεί ως κριτήριο διαχωρισμού των δεδομένων εκπαίδευσης θα οδηγήσει σε κόμβους κατά το δυνατό διαφορετικούς σε σχέση με την εξαρτημένη μεταβλητή.
2. Κάνε το διαχωρισμό.
3. Επανάλαβε τη διαδικασία για κάθε έναν από τους κόμβους που προέκυψαν μέχρι να μην είναι δυνατός περαιτέρω διαχωρισμός.

Ο αλγόριθμος ID3 κατασκευάζει το δένδρο κάθετα, ξεκινώντας από το καταλληλότερο χαρακτηριστικό για έλεγχο στη ρίζα. Η επιλογή βασίζεται σε κάποιο στατιστικό μέτρο που υπολογίζεται από τα δεδομένα. Στη συνέχεια, για κάθε δυνατή τιμή του χαρακτηριστικού δημιουργούνται οι αντίστοιχοι απόγονοι της ρίζας και τα δεδομένα μοιράζονται στους νέους κόμβους ανάλογα με την τιμή που έχουν για το χαρακτηριστικό που ελέγχεται στη ρίζα. Μόλις η διαδικασία ολοκληρωθεί, επαναλαμβάνεται εκ νέου για κάθε νέο κόμβο μέχρις ότου οι κόμβοι γίνουν τερματικοί (ή φύλλα). [35]

4.2.4.3 Η Μάθηση Κανόνων Ταξινόμησης

Οι κανόνες if-then αποτελούν από τους πιο κατανοητούς κανόνες για τον άνθρωπο. Κυριότερες κατηγορίες των κανόνων αυτών είναι:

1. Προτασιακοί κανόνες (propositional rules)
2. Κατηγορηματικοί κανόνες πρώτης τάξης (first order predicate rules)

Οι προτασιακοί κανόνες προκύπτουν είτε άλλες μορφές αναπαράστασης, είτε από απ' ευθείας μάθηση με αλγόριθμους σειριακής κάλυψης. Δεν περιλαμβάνουν μεταβλητές και έτσι δεν μπορεί να αναπαραστήσουν γενικές σχέσεις ανάμεσα στις τιμές των χαρακτηριστικών.

Οι κατηγορηματικοί κανόνες πρώτης τάξης περιέχουν μεταβλητές (μεγάλη εκφραστική ικανότητα) και προκύπτουν με απευθείας μάθηση μέσω αλγορίθμων μάθησης κανόνων 1ης τάξης.

Μάθηση Προτασιακών Κανόνων Ταξινόμησης

Η Μάθηση Προτασιακών Κανόνων Ταξινόμησης ενδείκνυται για προβλήματα που δεν απαιτούν αναπαράσταση σχέσεων ανάμεσα στις τιμές των διαφόρων χαρακτηριστικών. Ένας παράδειγμα τέτοιου αλγόριθμου, αποτελεί **αλγόριθμος σειριακής κάλυψης** (Sequential Covering Algorithm), ο οποίος δημιουργεί ένα σύνολο προτασιακών κανόνων σταδιακά (incrementally) μαθαίνοντας έναν κανόνα κάθε φορά. Ο κάθε κανόνας καλύπτει ένα σύνολο θετικών παραδειγμάτων, που είναι ξένο προς τα σύνολα των υπολοίπων κανόνων και καλύπτει ένα παράδειγμα όταν οι τιμές των χαρακτηριστικών της συνθήκης του κανόνα συμφωνούν με τις αντίστοιχες τιμές του παραδείγματος. Το σύνολο των κανόνων καλύπτει το σύνολο των θετικών παραδειγμάτων. Τα βήματα του παραπάνω αλγορίθμου συνοψίζονται ως εξής:

Αλγόριθμος Σειριακής Κάλυψης

1. Αρχικοποίησε το $\Sigma\upsilon\nu\omicron\lambda\omicron\upsilon\kappa\alpha\iota\ \text{Κανόνων}$ με το κενό σύνολο.
2. Μάθε $\acute{\epsilon}\nu\alpha\iota\ \text{Κανόνα}$ (Εξαρτημένη Μεταβλητή, Μεταβλητές, Παραδείγματα).
3. Αν ο Κανόνας ικανοποιεί το Κριτήριο Απόδοσης:
 - 3 α. Αφαίρεσε τα θετικά παραδείγματα που κάλυψε ο Κανόνας αυτός.
 - β. Πρόσθεσε τον Κανόνα στο $\Sigma\upsilon\nu\omicron\lambda\omicron\upsilon\kappa\alpha\iota\ \text{Κανόνων}$.
4. Επανάλαβε από το 2, όσο ικανοποιείται το Κριτήριο Απόδοσης.

Η συνάρτηση **Μάθε_έναν_Κανόνα** είναι κρίσιμη για την απόδοση του αλγορίθμου και έχει την παρακάτω μορφή:

Συνάρτηση Μάθε_έναν_Κανόνα

Εξαρτημένη Μεταβλητή, Χαρακτηριστικά, Παραδείγματα

//Αναζήτηση "Γενικό προς Ειδικό" (General-to-Specific Search)

1. Έστω η βέλτιστη υπόθεση (αρχικά, ο πιο γενικός κανόνας) που ταιριάζει με όλα τα παραδείγματα του συνόλου εκπαίδευσης.
2. Επανάλαβε, όσο υπάρχουν υποψήφια υποθέσεις:

Εξειδίκευσε τη βέλτιστη υπόθεση, προσθέτοντας το ζεύγος

χαρακτηριστικού-τιμής που βελτιστοποιεί το κριτήριο απόδοσης.

Το Κριτήριο Απόδοσης είναι ένα μέτρο της απόδοσης του κανόνα που καθορίζει ο χρήστης, όπως η Εντροπία, η Σχετική Συχνότητα ή ο m-Εκτιμητής Ακρίβειας. [36]

4.2.4.4 Η Μάθηση κατά Περίπτωση

Σε αυτό το τύπο μάθησης τα δεδομένα εκπαίδευσης διατηρούνται αυτούσια, σε αντίθεση με τις άλλες μεθόδους Μηχανικής Μάθησης οι οποίες κωδικοποιούν τα παραδείγματα εκπαίδευσης σε μια συμπαγή περιγραφή. Όταν ένα τέτοιο σύστημα κληθεί να αποφασίσει για την κατηγορία μιας νέας περίπτωσης, εξετάζει τη σχέση αυτής με τα ήδη αποθηκευμένα παραδείγματα. Ένα παράδειγμα τέτοιου αλγορίθμου αποτελεί ο αλγόριθμος των k-κοντινότερων γειτόνων (k-Nearest Neighbors) ο οποίος κάνει την παραδοχή ότι τα διάφορα παραδείγματα μπορούν να αναπαρασταθούν ως σημεία σε κάποιον n-διάστατο Ευκλείδειο χώρο R^n , όπου n ο αριθμός των χαρακτηριστικών (ανεξάρτητων μεταβλητών). Έτσι, κάθε νέα περίπτωση τοποθετείται σε αυτό το χώρο ως νέο σημείο και η τιμή του προσδιορίζεται με βάση το χαρακτηρισμό των k γειτονικών σημείων. Οι κοντινότεροι γείτονες μιας περίπτωσης υπολογίζονται βάση της Ευκλείδειας απόστασης τους. [201]

$$d(x, x') = \sqrt{\sum_{r=1}^n (a_r(x) - a_r(x'))^2}$$

4.2.4.5 Η Μάθηση κατά Bayes

Στη *Μάθηση κατά Bayes* (*Bayesian learning*), κάθε παράδειγμα εκπαίδευσης σταδιακά μειώνει ή αυξάνει την πιθανότητα ορθότητας μίας υπόθεσης. Για την εφαρμογή αυτής της μάθησης, απαιτείται η γνώση πολλών τιμών πιθανοτήτων. Στην περίπτωση όπου οι τιμές αυτές δεν είναι δυνατό να υπολογιστούν με ακρίβεια, υπολογίζονται κατ' εκτίμηση από προηγούμενη γνώση. Μια απλουστευμένη εκδοχή της μάθησης κατά Bayes, είναι ο **απλός ταξινομητής Bayes**, στον οποίο γίνεται η παραδοχή ότι τα χαρακτηριστικά είναι ανεξάρτητα μεταξύ τους.

Απλός ταξινομητής Bayes (simple/naive Bayes classifier)

Ο απλός ταξινομητής Bayes αποτελεί μια πρακτική μέθοδο μάθησης που στηρίζεται σε στατιστικά στοιχεία (κατανομές πιθανότητας). Η ποσότητα P που περιγράφει έναν απλό ταξινομητή Bayes για ένα σύνολο παραδειγμάτων, εκφράζει την πιθανότητα να είναι c η τιμή της εξαρτημένης μεταβλητής C με βάση τις τιμές $x=(x_1, x_2, \dots, x_n)$ των χαρακτηριστικών $X=(X_1, X_2, \dots, X_n)$ και δίνεται από τη σχέση:

$$P(c|x) = P(c) \cdot \prod_i P(x_i|c), \text{ όπου τα χαρακτηριστικά } X_i \text{ θεωρούνται ανεξάρτητα μεταξύ τους.}$$

Ο υπολογισμός της παραπάνω ποσότητας για ένα σύνολο N παραδειγμάτων βάσει των ακόλουθων σχέσεων.

1. $P(c) = N(c) / N$,
2. $P(x_i|c) = N(x_i, c) / N(c)$, για χαρακτηριστικό X_i με διακριτές τιμές,
3. $P(x_i|c) = g(x_i, \mu_c, \sigma_c^2)$, για χαρακτηριστικό X_i με αριθμητικές τιμές,

Όπου $N(c)$, ο αριθμός των παραδειγμάτων που έχουν στην εξαρτημένη μεταβλητή την τιμή c ,

Όπου $N(x_i, c)$, ο αριθμός των παραδειγμάτων που έχουν για το χαρακτηριστικό X_i και την εξαρτημένη μεταβλητή, τιμές x_i και c αντίστοιχα.

Όπου $g(x_i, \mu_c, \sigma_c^2)$, η συνάρτηση πυκνότητας πιθανότητας Gauss με μέσο όρο μ_c και διασπορά σ_c για το χαρακτηριστικό X_i . [201]

4.2.4.6 Παρεμβολή ή Παλινδρόμηση

Παρεμβολή ή παλινδρόμηση (regression), ονομάζεται η διαδικασία προσδιορισμού της σχέσης μιας μεταβλητής y (εξαρτημένη μεταβλητή ή έξοδος) με μια ή περισσότερες άλλες μεταβλητές x_1, x_2, \dots, x_n (ανεξάρτητες μεταβλητές ή εισόδοι), και έχει σαν σκοπό την πρόβλεψη της τιμής της εξόδου υπό δεδομένες εισόδους. Το πιο διαδεδομένο μοντέλο είναι το γραμμικό (linear) όπου η αναμενόμενη τιμή της εξόδου μοντελοποιείται με μία γραμμική συνάρτηση ή σταθμισμένο άθροισμα (weighted sum) των παραμέτρων εισόδου.

$$y_j = \beta_0 + \beta_1 x_{1j} + \beta_2 x_{2j} + \dots + \beta_n x_{nj} \quad j = 1, 2, \dots, m$$

όπου m , ο αριθμός των δεδομένων (παραδειγμάτων) εκπαίδευσης, και

όπου β_i , οι συντελεστές που χρήζουν υπολογισμού.

Μία άλλη διαδεδομένη μέθοδος επίλυσης είναι η μέθοδος των ελαχίστων τετραγώνων (least squares), η οποία αποσκοπεί στην ελαχιστοποίηση του σφάλματος μεταξύ εκτιμώμενης συνάρτησης και πραγματικών δεδομένων. [201]

4.2.4.7 Νευρωνικά Δίκτυα

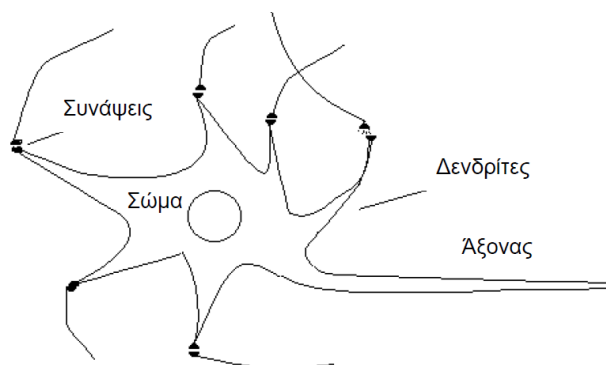
Τα *Νευρωνικά Δίκτυα* (*Neural Networks*) παρέχουν ένα πρακτικό τρόπο για την εκμάθηση αριθμητικών και διανυσματικών συναρτήσεων ορισμένων σε συνεχή ή διακριτά μεγέθη. Χρησιμοποιούνται τόσο για παρεμβολή (γραμμική και μη γραμμική) όσο και για ταξινόμηση. Έχουν το μεγάλο πλεονέκτημα της ανοχής που παρουσιάζουν σε δεδομένα εκπαίδευσης με θόρυβο, δηλαδή δεδομένα που περιστασιακά έχουν λανθασμένες τιμές. Αδυνατούν όμως να εξηγήσουν ποιοτικά τη γνώση που μοντελοποιούν.

Φυσικά Νευρωνικά Δίκτυα

Ο Νευρώνας αποτελεί τη δομική μονάδα του εγκεφάλου, η οποία επιτελεί όλες τις νοητικές διεργασίες. Ο Νευρώνας αποτελείται από τα κάτωθι δομικά μέρη:

1. Σώμα: Ο πυρήνας του νευρώνα
2. Δενδρίτες: Υλοποιούν τη λήψη των σημάτων από γειτονικούς νευρώνες (σημεία εισόδου).
3. Άξονας: Υλοποιεί την διασύνδεση του Νευρώνα με τους γειτονικούς Νευρώνες (σημείο εξόδου).

Σε κάθε δενδρίτη υπάρχει ένα απειροελάχιστο κενό, η *σύναψη*. Η ικανότητα μάθησης και μνήμης που παρουσιάζει ο εγκέφαλος οφείλεται στην ικανότητα των συνάψεων να μεταβάλουν την αγωγιμότητά τους.



Εικόνα 4.2

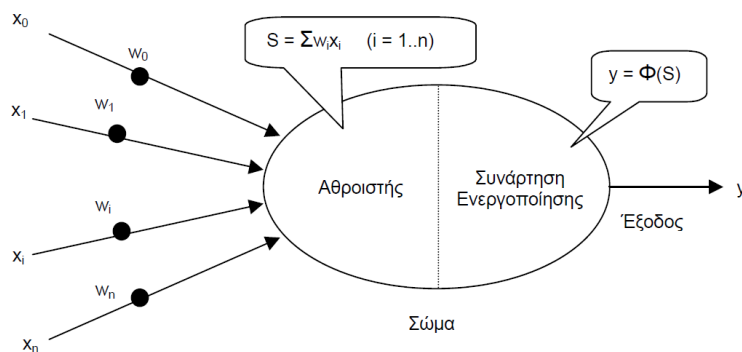
Η βιολογική δομή του εγκεφαλικού νευρώνα.

[201]

Μοντέλο Τεχνητού Νευρώνα

Ένα σύστημα Τεχνητού νευρώνα λαμβάνει *Σήματα εισόδου* εντός του συνεχούς πεδίου ορισμού x_0, x_1, \dots, x_n . Η *Τιμή Βάρους* w_i (*weight*) αποτελεί τμήμα ανάλογο της σύναψης των βιολογικών νευρώνων. Τα δομικά στοιχεία του σώματος του Τεχνητού Νευρώνα είναι τα ακόλουθα:

1. *Αθροιστής (sum)*: προσθέτει τα επηρεασμένα από τα βάρη σήματα εισόδου και παράγει την ποσότητα S .
2. *Συνάρτηση ενεργοποίησης ή κατωφλίου (activation ή threshold function)*: μη γραμμικό φίλτρο που διαμορφώνει το σήμα εξόδου y , σε συνάρτηση με την ποσότητα S .

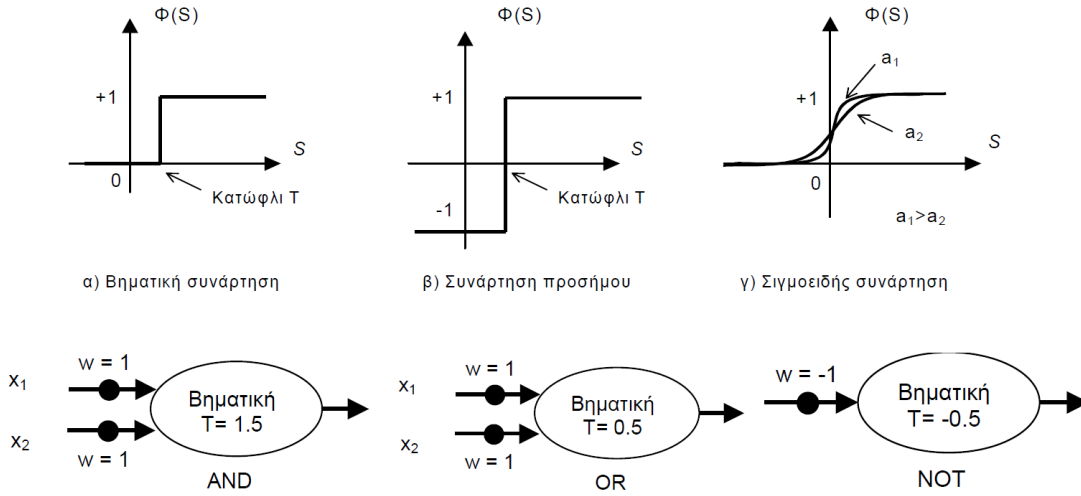


Εικόνα 4.3

Η δομή του μοντέλου Τεχνητού Νευρώνα.

[201]

Συναρτήσεις Ενεργοποίησης

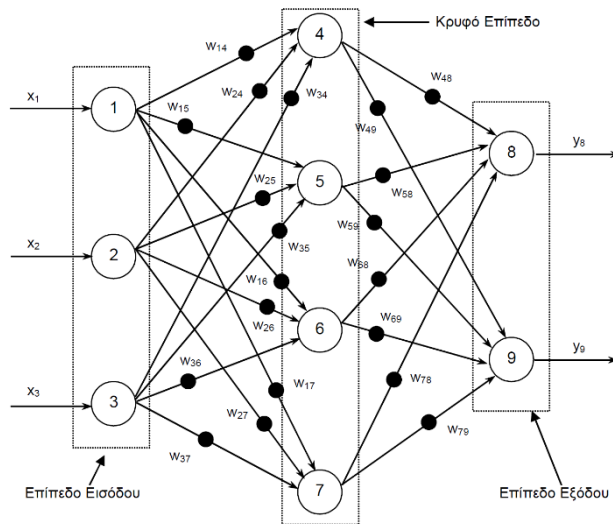


Εικόνα 4.4
Οι συναρτήσεις ενεργοποίησης του Τεχνητού Νευρώνα.
[201]

Τεχνητά Νευρωνικά Δίκτυα

Τα *Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ)*, αποτελούνται από στρώματα ή επίπεδα (*layers*) νευρώνων, τον οποίων η δομή περιλαμβάνει:

1. Το επίπεδο εισόδου (*input layer*)
2. Τα ενδιάμεσα ή κρυφά επίπεδα (*hidden layers*)
3. Το επίπεδο εξόδου (*output layer*)



Εικόνα 4.5
Παράδειγμα Τεχνητού Νευρωνικού Δικτύου.
[201]

Οι Νευρώνες που συμμετέχουν σε ένα Δίκτυο μπορούν να είναι είτε *Πλήρως συνδεδεμένοι (fully connected)* μεταξύ τους (συνδέονται με όλους τους νευρώνες του επόμενου επιπέδου), είτε *Μερικώς συνδεδεμένοι (partially connected)*. Τα Τεχνητά Νευρωνικά Δίκτυα, χωρίζονται σε:

1. Δίκτυα με *απλή τροφοδότηση (feedforward)*, στα οποία δεν υπάρχουν συνδέσεις μεταξύ νευρώνων ενός επιπέδου και νευρώνων προηγούμενου επιπέδου, και
2. Δίκτυα με *ανατροφοδότηση (feedback ή recurrent)*, στα οποία υπάρχουν συνδέσεις μεταξύ νευρώνων ενός επιπέδου και νευρώνων προηγούμενου επιπέδου.

Μέθοδοι Μάθησης και Ανάκλησης των Τεχνητών Νευρωνικών Δικτύων

- *Μάθηση (learning)* σε ένα Τεχνητό Νευρωνικό Δίκτυο, ορίζεται ως η διαδικασία της τροποποίησης της τιμής των βαρών του δικτύου, ώστε για συγκεκριμένο δοσμένο διάνυσμα εισόδου, να παράγεται συγκεκριμένο διάνυσμα εξόδου.
- *Ως Ανάκληση (recall)* σε ένα Τεχνητό Νευρωνικό Δίκτυο, ορίζεται η διαδικασία του υπολογισμού ενός διανύσματος εξόδου συναρτήσει συγκεκριμένου διανύσματος εισόδου και τιμές βαρών.
- *Μάθηση υπό επίβλεψη (supervised learning)* σε ένα Τεχνητό Νευρωνικό Δίκτυο, είναι η διαδικασία κατά την οποία στο Δίκτυο δίνονται ζευγάρια διανυσμάτων *εισόδου - επιθυμητής εξόδου*. Αντίστοιχα μοντέλα μάθησης είναι τα εξής:
 - *Κανόνας Δέλτα (Delta rule learning)*
 - *Αλγόριθμος ανάστροφης μετάδοσης λάθους (back propagation)*
 - *Ανταγωνιστική μάθηση (competitive learning)*
 - *Τυχαία μάθηση (random learning)*

Δίκτυα με Απλή Τροφοδότηση

Τα Δίκτυα Απλής Τροφοδότησης, αποτελούνται από:

1. Το επίπεδο εισόδου
2. Το επίπεδο εξόδου
3. Κανένα, ένα ή περισσότερα ενδιάμεσα, κρυφά επίπεδα

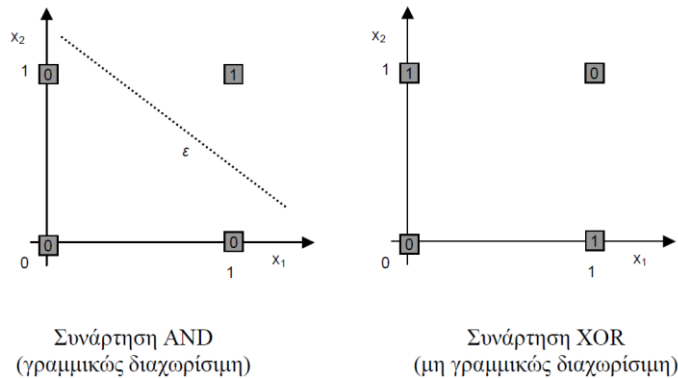
Η τοπολογία των Τεχνητών Νευρωνικών Δικτύων με απλή τροφοδότηση καθορίζεται από το πόσα κρυφά επίπεδα θα έχει το Δίκτυο και από πόσους νευρώνες θα αποτελείται το κάθε επίπεδο. Η πιο απλή τοπολογία δικτύου με τροφοδότηση, είναι το *Perceptron*, ο οποίος αποτελεί έναν μοναδικό τεχνητό νευρώνα, ο οποίος χρησιμοποιεί ως συνάρτηση κατωφλίου τη βηματική συνάρτηση.

Ο αλγόριθμος *Perceptron*, μεταβολής των βαρών περιγράφεται ως εξής:

1. Εάν για τη συγκεκριμένη είσοδο παράγεται το επιθυμητό αποτέλεσμα τότε δε γίνεται καμία μεταβολή.
2. Εάν το αποτέλεσμα είναι 1, ενώ θα έπρεπε να είναι 0, τότε μειώνονται τα βάρη των ενεργών γραμμών (εκείνες που στο συγκεκριμένο πρότυπο έχουν είσοδο 1) κατά μια τιμή d , η οποία ονομάζεται *ρυθμός μάθησης (learning rate)*.
3. Εάν το αποτέλεσμα είναι 0 ενώ θα έπρεπε να είναι 1 τότε αυξάνονται τα βάρη των ενεργών γραμμών κατά d .

Ένα *perceptron* με n γραμμές εισόδου μπορεί να θεωρηθεί ότι αναπαριστά ένα υπερεπίπεδο $n-1$ διαστάσεων που διαχωρίζει τα διανύσματα εισόδου σε δύο ομάδες, τοποθετώντας από τη μία μεριά όσα παράγουν έξοδο 1 και από την άλλη όσα παράγουν έξοδο 0.

Γραμμικώς διαχωρίσιμα προβλήματα (*linearly separable*)



Εικόνα 4.6
Οι γραμμικώς διαχωρίσιμες συναρτήσεις AND και XOR.
[201]

Ο Κανόνας Δέλτα περιλαμβάνει:

- Τη συνάρτηση ενεργοποίησης
- Την έξοδο ενός νευρώνα *i*
- Τη μεταβολή του βάρους *w_j*
- Το συνολικό σφάλμα

$$g(z) = \frac{1}{1 + e^{-z}}$$

$$a_i = g \left(\sum_{j=0}^n w_j a_j \right)$$

$$w_j = w_{j_old} - d \cdot (a_i - a) a_j$$

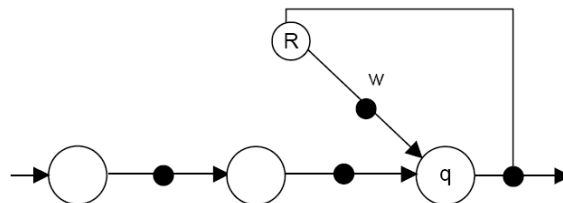
$$Err = \frac{1}{2} \sum (a_i - a)^2$$

Ανάστροφη Μετάδοση Λάθους

Η φιλοσοφία της ανάστροφης μετάδοσης λάθους, βασίζεται στον καθορισμό του ποσοστού του συνολικού σφάλματος που αντιστοιχεί στα βάρη του κάθε νευρώνα. Έτσι, αρχικά υπολογίζεται το σφάλμα για τους νευρώνες του επιπέδου εξόδου και αμέσως μετά, τα σφάλματα στο τελευταίο κρυμμένο επίπεδο. Η παραπάνω διαδικασία συνεχίζεται αναδρομικά μέχρι το πρώτο επίπεδο. Έπειτα υπολογίζεται η συνεισφορά κάθε βάρους των νευρώνων στο ολικό σφάλμα. Τα σφάλματα κάθε επιπέδου χρησιμοποιούνται για να μεταβάλλουν τα βάρη του κάθε νευρώνα. Η διαδικασία επαναλαμβάνεται μέχρι το σφάλμα να πάρει τιμή στα όρια ανοχής που έχει θέσει ο χρήστης. Τέλος πραγματοποιείται αναζήτηση τύπου *αναρρίχησης λόφου* και εύρεση των τοπικών ελαχίστων.

Νευρωνικά Δίκτυα με Ανατροφοδότηση

Βρόχοι στην συνδεσμολογία των νευρώνων.

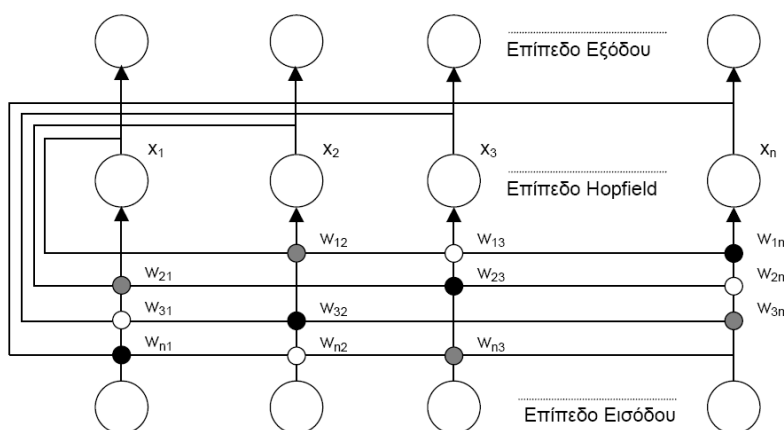


Εικόνα 4.7
Νευρωνικό Δίκτυο με ανατροφοδότηση.

Δίκτυα Hopfield

Στα δίκτυα Hopfield, ο κάθε νευρώνας συνδέεται με όλους τους υπόλοιπους. Οι συνδέσεις μεταξύ των νευρώνων είναι αμφίδρομες και έχουν συμμετρικά βάρη ($w_{ij}=w_{ji}$). Όλοι οι νευρώνες λειτουργούν ταυτόχρονα σαν εισοδοί και έξοδοι του δικτύου. Τα δίκτυα Hopfield αποτελούν απλά perceptrons με συνάρτηση ενεργοποίησης τη συνάρτηση πρόσημου και έξοδο -1 ή 1. Η εκπαίδευση του δικτύου βασίζεται στη μεταβολή της τιμής των βαρών βάσει της σχέσης:

$$\Delta w_{ij} = (2 \cdot x_i - 1) \cdot (2 \cdot x_j - 1), \text{ όπου τα } x_i \text{ και } x_j \text{ έχουν τιμές } 0 \text{ ή } 1.$$



Εικόνα 4.8

Η τοπολογία του δικτύου Hopfield.

[201]

Δίκτυο Kohonen

Στα δίκτυα Kohonen οι νευρώνες βρίσκονται σε γεωμετρική τοπολογία. Κάθε νευρώνας συνδέεται μέσω βαρών με την είσοδο (η οποία αποτελείται από M στοιχεία) και λαμβάνει ένα πλήρες αντίγραφο του διανύσματος εισόδου τροποποιημένο από τα βάρη. Για μια δεδομένη είσοδο, ο νευρώνας του οποίου τα συνδεδετικά βάρη είναι πλησιέστερα σε αυτήν, επιλέγεται και μεταβάλλει τα βάρη του έτσι ώστε να την πλησιάζει περισσότερο. Λόγω της συνδεσμολογίας, μεταβάλλονται και τα βάρη στους νευρώνες που βρίσκονται στην άμεση γειτονιά του επιλεγμένου νευρώνα. Ένα εκπαιδευμένο δίκτυο Kohonen μπορεί να χρησιμοποιηθεί για να κατατάξει στοιχεία σε μια από τις κατηγορίες που ήδη έχουν σχηματιστεί.

[201]

4.2.4.8 Μηχανές Διανυσμάτων Υποστήριξης

Οι Μηχανές Διανυσμάτων Υποστήριξης (ΜΔΥ) ή Support Vector Machines (SVMs), στηρίζονται στη Θεωρία Στατιστικής Μάθησης (Statistical Learning Theory) και στα Νευρωνικά Δίκτυα τύπου Perceptron. Αποτελούν μια από τις πιο διαδεδομένες μεθόδους παρεμβολής και ταξινόμησης, με πλήθος εφαρμογών όπως αναγνώριση γραφής, ταξινόμηση κειμένων, ταξινόμηση δεδομένων έκφρασης γονιδίων κ.α. Στην περίπτωση εφαρμογής τους στην ταξινόμηση, οι ΜΔΥ προσπαθούν να βρουν μια υπερ-επιφάνεια (hypersurface) που να διαχωρίζει στο χώρο των παραδειγμάτων τα αρνητικά από τα θετικά παραδείγματα. Η υπερ-επιφάνεια αυτή επιλέγεται έτσι ώστε να απέχει όσο το δυνατόν περισσότερο από τα κοντινότερα θετικά και αρνητικά παραδείγματα (maximum margin hypersurface). [201], [8], [14], [22]

4.2.5 Η μη Επιβλεπόμενη Μάθηση

Στη μη επιβλεπόμενη μάθηση, το σύστημα έχει στόχο να ανακαλύψει συσχετίσεις ή / και ομάδες από τα δεδομένα, βασιζόμενο αποκλειστικά στις ιδιότητές τους. Έτσι, προκύπτουν πρότυπα, κάθε ένα από τα οποία περιγράφει ένα μέρος των δεδομένων.

Παραδείγματα προτύπων πληροφόρησης είναι:

Αναγνώριση πεζών από αυτόνομο ρομποτικό όχημα με χρήση τεχνητής νοημοσύνης και τεχνικών βαθιάς μάθησης

1. οι κανόνες συσχέτισης (association rules) και
2. οι ομάδες (clusters), οι οποίες προκύπτουν από τη διαδικασία της ομαδοποίησης (clustering). [201]

4.2.5.1 Οι κανόνες συσχέτισης

Η ανακάλυψη ή εξόρυξη κανόνων συσχέτισης (association rule mining) αποτελεί ένα ερευνητικό πεδίο αρκετά νεότερο της Μηχανικής Μάθησης. Σκοπός του είναι η ανακάλυψη των συσχετίσεων που φέρει μεταξύ τους ένα σύνολο αντικειμένων.

Έστω $I = \{i_1, i_2, \dots, i_n\}$ ένα σύνολο από διακριτά στοιχεία, που αποκαλούνται items (αντικείμενα) και ότι $D = \{t_1, t_2, \dots, t_m\}$ ένα σύνολο από δοσοληψίες (transactions), όπου κάθε δοσοληψία T αποτελεί ένα σύνολο αντικειμένων τα οποία ονομάζονται *itemset*, και όπου ισχύει $T \subseteq I$. Κάθε δοσοληψία ταυτίζεται με ένα μοναδικό αναγνωριστικό που καλείται TID.

Ένας κανόνας είναι μία συσχέτιση της μορφής $X \Rightarrow Y$ όπου $X \subseteq I$, $Y \subseteq I$ και $X \cap Y = \emptyset$. Το πρώτο μέλος του κανόνα ονομάζεται *υπόθεση* και το δεύτερο *συμπέρασμα*. Υπάρχουν δύο βασικές μετρικές στον συσχετισμό κανόνων. Η υποστήριξη (support) που σημαίνει ότι ο κανόνας $X \Rightarrow Y$ έχει υποστήριξη s , αν το $s\%$ των δοσοληψιών στο D περιέχουν το $(X \cup Y)$ και η εμπιστοσύνη που σημαίνει ότι ο κανόνας $X \Rightarrow Y$ ισχύει στο D , αν το $c\%$ των δοσοληψιών στο D που περιέχουν το X , περιέχουν επίσης και το Y . Από τα παραπάνω προκύπτει ότι ο κανόνας $X \Rightarrow Y$ έχει υποστήριξη s , όταν $\text{supp}(X \cup Y)$ και εμπιστοσύνη c , όταν $\text{conf}(X \Rightarrow Y) = \text{supp}(X \cup Y) / \text{supp}(X)$.

Για την ανακάλυψη κανόνων συσχέτισης χρησιμοποιείται η ιδιότητα της μονοτονίας (monotonicity property) ή αλλιώς ιδιότητα a priori, η οποία περιγράφεται ως εξής:

"Αν ένα σύνολο αντικειμένων S είναι συχνό, τότε όλα τα υποσύνολα του S είναι επίσης συχνά".

Ο αλγόριθμος Apriori

Αποτελεί τον κλασικότερο αλγόριθμο ανακάλυψης κανόνων συσχέτισης. Περιλαμβάνει δυο βασικά βήματα, τη δημιουργία των συχνών συνόλων αντικειμένων και τη δημιουργία των κανόνων συσχέτισης. Η διαδικασία της δημιουργίας συχνών συνόλων αντικειμένων συνοψίζεται ως εξής:

1. Δημιουργείται ένα σύνολο υποψήφιων συχνών αντικειμένων C_i και μετά δημιουργείται το σύνολο των συχνών συνόλων αντικειμένων L_i .
2. Η διαδικασία επαναλαμβάνεται πραγματοποιώντας διαδοχικά περάσματα στα δεδομένα έως ότου βρεθούν είτε τα συχνά σύνολα αντικειμένων ενός προκαθορισμένου επιπέδου ή τα μέγιστα συχνά σύνολα αντικειμένων.

Το πρώτο στάδιο αποτελείται από ακόμα ένα βήμα συνένωσης (join step) και ένα βήμα κλαδέματος (prune step). Για τη εξαγωγή κανόνων συσχέτισης, ελέγχεται η εμπιστοσύνη (confidence) όλων των πιθανών κανόνων (βάση των μέγιστων συχνών συνόλων αντικειμένων) και στο τέλος μένουν μόνο αυτοί που έχουν βαθμό εμπιστοσύνης μεγαλύτερο από το όριο που τέθηκε.

[201]

4.2.5.2 Η ομαδοποίηση

Η ομαδοποίηση (clustering) αφορά τη διαδικασία διαχωρισμού δεδομένων σε ομάδες, που εφαρμόζεται σε ένα σύνολο δεδομένων έτσι ώστε να επιτευχθεί η βέλτιστη δυνατή ομοιογένεια μεταξύ των δεδομένων που υφίστανται σε κάποια ομάδα.

Αλγόριθμοι Ομαδοποίησης

Οι τρεις κατηγορίες των αλγορίθμων ομαδοποίησης είναι:

1. Οι διαχωριστικοί (partition based) αλγόριθμοι, οι οποίοι αποσκοπούν στον βέλτιστο δυνατό διαχωρισμό ενός συνόλου δεδομένων σε ένα συγκεκριμένο αριθμό ομάδων.
2. Οι ιεραρχικοί (hierarchical) αλγόριθμοι, οι οποίοι αποσκοπούν στην που ιεραρχική εύρεση του αριθμού και της δομής των ομάδων.
3. Οι πιθανοκρατικοί (probabilistic) αλγόριθμοι, οι οποίοι στηρίζονται σε πιθανοκρατικά μοντέλα.

Όπως είναι κατανοητό, η διαδικασία της ομαδοποίησης προϋποθέτει την ύπαρξη ενός μέτρου ομοιότητας ή διαφοράς μεταξύ των δεδομένων. Σαν το μέτρο αυτό συνήθως επιλέγεται "απόσταση" μεταξύ των δεδομένων, η οποία και πρωτίστως υπολογίζεται. Τυπικά μέτρα υπολογισμού απόστασης μεταξύ δύο δεδομένων, είναι η απόσταση *Μανχάταν* και η *Ευκλείδεια* απόσταση:

$$d(x, y) = \sum_i |x_i - y_i| \quad d(x, y) = \sqrt{\sum_i (x_i - y_i)^2}$$

Απόσταση *Μανχάταν* *Ευκλείδεια* απόσταση

Αλγόριθμοι Βασισμένοι σε Διαχωρισμούς

Ο αλγόριθμος των K-μέσων (K-means)

Ο αλγόριθμος K-means αποτελεί έναν από τους πιο γνωστούς αλγόριθμους ομαδοποίησης αυτής της κατηγορίας. Ο αριθμός K των ομάδων καθορίζεται πριν την εκτέλεση του αλγορίθμου. Αμέσως μετά τον καθορισμό του αριθμού K, ο αλγόριθμος ξεκινά διαλέγοντας K τυχαία σημεία από τα δεδομένα ως τα κέντρα των ομάδων και τα αναθέτει στην ομάδα αυτή της οποίας το κέντρο είναι πιο κοντά στο σημείο αυτό. Στη συνέχεια, υπολογίζεται ο μέσος όρος των στοιχείων κάθε ομάδας, και ορίζεται σαν κέντρο της. Η διαδικασία επαναλαμβάνονται για ένα προκαθορισμένο αριθμό βημάτων ή μέχρις ότου να μην υπάρχει αλλαγή στο διαχωρισμό των σημείων σε ομάδες.

Ο αλγόριθμος των K-μέσων δίνεται σε ψευδογλώσσα:

```

Αλγόριθμος K-μέσων
είσοδος:
Σύνολο δεδομένων D = {x1, ..., xn}
Αριθμός Ομάδων k
έξοδος:
Ομάδες Ci
1.//ανάθεση τυχαίων κέντρων
για i= 1,...,k κάνε:
θεώρησε mi ως ένα τυχαίο στοιχείο από το D;
2.//ομαδοποίηση
όσο υπάρχουν αλλαγές στις ομάδες Ci κάνε:
2α.//δημιουργία ομάδων
για i = 1, ..., k κάνε
Ci = {x ∈ D | d(mi, x) ≤ d(mj, x) για όλα τα j = 1, ..., k, j ≠ k};
2β.//υπολογισμός νέων κέντρων
για i = 1,...,k κάνε
mi = το μέσο διάνυσμα των σημείων που ανήκουν στην ομάδα Ci;

```

Αλγόριθμοι Ιεραρχικής Ομαδοποίησης

Οι αλγόριθμοι ιεραρχικής ομαδοποίησης είτε συνδυάζουν μικρές ομάδες σε μεγαλύτερες, είτε διαιρούν μεγαλύτερες ομάδες σε μικρότερες, έτσι ώστε να προκύψει μια ιεραρχία από διαφορετικές ομαδοποιήσεις των δεδομένων, στο ένα άκρο της οποίας βρίσκεται μια μόνο ομάδα με όλα τα δεδομένα, και στο άλλο τόσες ομάδες όσες και ο αριθμός των δεδομένων.

Με κριτήριο την κατεύθυνση της αναπτυσσόμενης ιεραρχίας, οι ιεραρχικοί αλγόριθμοι ομαδοποίησης χωρίζονται στους αλγόριθμους συγχώνευσης (agglomerative) και στους αλγόριθμους διαίρεσης (divisive).

Οι αλγόριθμοι συγχώνευσης βασίζονται σε μετρικές απόστασης ανάμεσα σε ομάδες. Στόχος τους είναι η εύρεση των δύο πιο πλησιέστερων ομάδων και η συγχώνευση τους.

Ο γενικός αλγόριθμος ιεραρχικής ομαδοποίησης, είναι:

Ιεραρχικός Αλγόριθμος Ομαδοποίησης

είσοδος:

Σύνολο δεδομένων $D = \{x_1, \dots, x_n\}$

Συνάρτηση $d(C_i, C_j)$ απόστασης δύο ομάδων C_i, C_j

έξοδος:

Ομάδες C_i

1. για $i = 1, \dots, n$ κάνε:

θεώρησε $C_i = \{x_i\}$;

2. όσο ο αριθμός των ομάδων είναι μεγαλύτερος από 1 κάνε:

2α. Υπολόγισε την απόσταση μεταξύ όλων των ομάδων ανά δύο

2β. $C_i = C_i \cup C_j$, όπου C_i και C_j οι δυο πιο κοντινές ομάδες

2γ. Αφαίρεσε την ομάδα C_j από το σύνολο των ομάδων

[201], [8]

4.2.6 Άλλες μορφές μάθησης

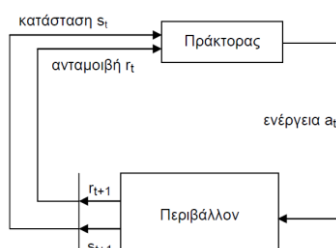
Εκτός από τις μεθόδους που παρουσιάστηκαν, υπάρχουν και άλλες προσεγγίσεις στο πρόβλημα της Μηχανικής Μάθησης. Τρεις από αυτές, είναι η ενισχυτική μάθηση, οι γενετικοί αλγόριθμοι και η Βαθιά Μάθηση. [201]

4.2.6.1 Η Ενισχυτική Μάθηση

Η Ενισχυτική Μάθηση (Reinforcement Learning) αποτελεί γενική περιγραφή οικογένειας τεχνικών στις οποίες το σύστημα μάθησης προσπαθεί να μάθει μέσω άμεσης αλληλεπίδρασης με το περιβάλλον. Για να καταστεί αυτό δυνατό, το σύστημα προσπαθεί να μεγιστοποιήσει μια συνάρτηση του αριθμητικού σήματος ενίσχυσης. Σε αυτό το είδος μάθησης, το σύστημα είναι πλήρως αυτόνομο και δεν καθοδηγείται από κάποιον εξωτερικό επιβλέποντα για το ποια ενέργεια πρέπει να ακολουθήσει. Αντιθέτως, θα πρέπει να ανακαλύψει μόνο του ποιες ενέργειες είναι αυτές που θα του αποφέρουν το μεγαλύτερο κέρδος. Αυτό το είδος μάθησης έχει αρκετά συχνή εφαρμογή σε ένα πλήθος εφαρμογών όπως έλεγχο ρομποτικού συστήματος, βελτιστοποίηση βιομηχανικών διεργασιών κ.α.

Το Πλαίσιο της Ενισχυτικής Μάθησης

Στη τεχνική της ενισχυτικής μάθησης, το αντικείμενο που καλείται να μάθει ονομάζεται *πράκτορας* (agent), και το εξωτερικό του περιβάλλον ονομάζεται *περιβάλλον*. Ο πράκτορας και το περιβάλλον αλληλοεπιδρούν συνεχώς, με την ίδια σταθερή σχέση. Ο πράκτορας αναλύει τα δεδομένα και λαμβάνει αποφάσεις ενώ το περιβάλλον αποκρίνεται σε αυτές επιστρέφοντας στον πράκτορα αφενός τις νέες καταστάσεις και αφετέρου τις ανταμοιβές (rewards), οι οποίες αποτελούν ειδικές αριθμητικές τιμές τις οποίες ο πράκτορας προσπαθεί να μεγιστοποιήσει. [201], [78]



Εικόνα 4.9

Η αλληλεπίδραση του πράκτορα με το περιβάλλον του. [201]

4.2.6.2 Οι Γενετικοί Αλγόριθμοι

Οι Γενετικοί Αλγόριθμοι (Genetic Algorithms) αποτελούν μία μέθοδο μάθησης η οποία βασίζεται στην προσομοίωση της βιολογικής εξέλιξης (evolution).

Σε αυτό το τύπο μάθησης, οι υποθέσεις αναπαρίστανται από ακολουθίες συμβολοσειρών bit (bit-strings). Η διαδικασία αναζήτησης της βέλτιστης κατάλληλης υπόθεσης, ξεκινάει με έναν τυχαίο πληθυσμό P αρχικών υποθέσεων, τα μέλη του οποίου παράγουν τη νέα "γενιά" P' μέσω διαδικασιών αναπαραγωγής αντίστοιχων των βιολογικών, όπως η διασταύρωση (crossover) και η τυχαία μετάλλαξη (random mutation). Ο πληθυσμός της επόμενης γενιάς ορίζεται σαν P'. Κάθε επόμενη γενιά φέρει κατά μέσο όρο καλύτερα χαρακτηριστικά από της προηγούμενη, όπως συμβαίνει και με τις βιολογικές γενιές. Σε κάθε βήμα, οι υποθέσεις του τρέχοντος πληθυσμού αξιολογούνται βάσει μιας προκαθορισμένης συνάρτησης καταλληλότητας (fitness function), βάσει της οποίας επιλέγονται για το αν θα υφίστανται ή όχι στην επόμενη γενιά. [201]

4.2.6.3 Η Βαθιά Μάθηση

Τι είναι η Βαθιά Μάθηση

Η Βαθιά Μάθηση (Deep Learning) αποτελεί μια τεχνική μάθησης, η οποία αποσκοπεί να διδάξει τους υπολογιστές πως μπορούν να μαθαίνουν αυτόνομα μέσα από παραδείγματα, όπως ακριβώς και ο άνθρωπος – χωρίς την ανάγκη παροχής προτυποποιημένων και καθορισμένων μοντέλων γνώσης. Ο ανθρώπινος εγκέφαλος λειτουργεί με τρόπο βασιζόμενο στην πολύ-επίπεδη εξελικτική μάθηση και έτσι έχει τη δυνατότητα να ανταποκρίνεται με επιτυχία σε πρωτόγνωρα και ακαθόριστα δεδομένα που λαμβάνει. Στη Βαθιά Μάθηση, ένα υπολογιστικό σύστημα μαθαίνει να εκτελεί εργασίες ταξινόμησης απευθείας από τα εισερχόμενα δεδομένα (εικόνες, κείμενο, ήχος κτλ). Τα μοντέλα Βαθιάς Μάθησης μπορούν να επιτύχουν ακρίβεια που πολλές φορές υπερβαίνει την ανθρώπινη απόδοση. Μέσα από τη Βαθιά Μάθηση, τα μοντέλα εκπαιδεύονται αξιοποιώντας ένα μεγάλο σύνολο ετικετών δεδομένων και αρχιτεκτονικών πολύ-στρωματικών Νευρωνικών Δικτύων.

Η τεχνολογία της Βαθιάς Μάθησης, αν και έχει διατυπωθεί από τις αρχές τις δεκαετίας του 1980, η εφαρμογή της κατέστη δυνατή μόλις στις μέρες μας καθώς η τεχνική της Βαθιάς Μάθησης απαιτεί :

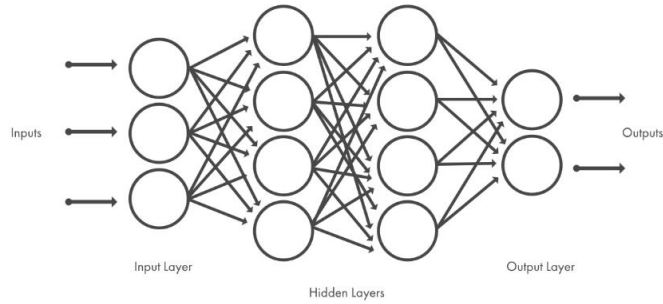
1. Αξιοποίηση τεράστιου πλήθους δεδομένων, τα οποία είναι διαθέσιμα στις μέρες μας λόγω της ραγδαίας ανάπτυξης των αισθητήριων συστημάτων, της διεύρυνσης του αποθηκευτικού χώρου στα ψηφιακά συστήματα και της αύξησης της ταχύτητας μετάδοσης των δεδομένων.

2. Σημαντική υπολογιστική ισχύ ώστε ένα τεράστιο πλήθος δεδομένων να επεξεργάζεται παράλληλα και με πολύ υψηλές ταχύτητες. Η ανάγκη παράλληλης επεξεργασίας δεδομένων, οδηγεί στην ανάγκη αξιοποίησης συστημάτων παράλληλης επεξεργασίας έναντι της σειριακής. Αυτό πρακτικά σημαίνει ότι η μονάδα επεξεργασίας που λαμβάνει χώρα σε συστήματα Βαθιάς Μάθησης είναι αποτελεσματικότερο να είναι αρχιτεκτονικής GPU έναντι CPU. Ειδικά όταν η συγκεκριμένη αρχιτεκτονική συνδυάζεται με clusters ή cloud computing, δίνεται η δυνατότητα μείωσης του απαιτούμενου χρόνου εκπαίδευσης ενός δικτύου Βαθιάς Μάθησης.

Πώς λειτουργεί η Βαθιά Μάθηση

Η τεχνική της Βαθιάς Μάθησης συνίσταται στην αξιοποίηση πολύ-επίπεδων Νευρωνικών Δικτύων, γι 'αυτό τα μοντέλα Βαθιάς Μάθησης πολλές φορές αναφέρονται και ως Βαθιά Νευρωνικά Δίκτυα.

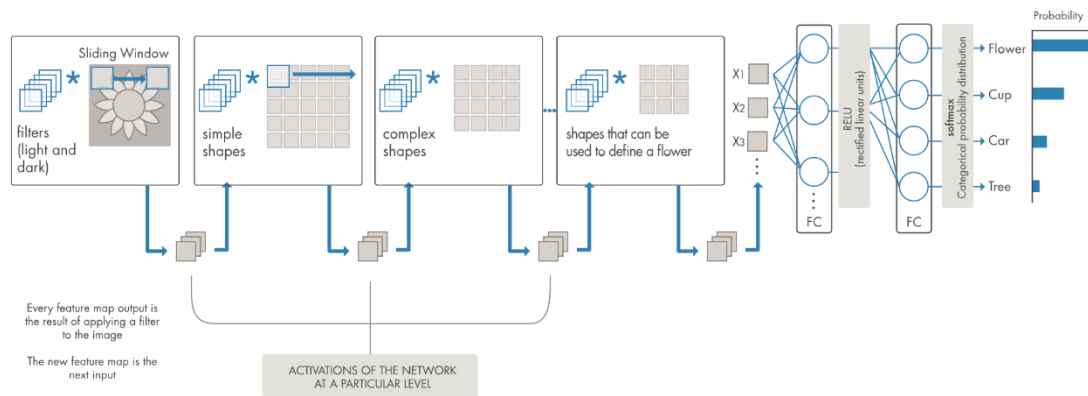
Ο όρος "Βαθιά", συνήθως αναφέρεται στον αριθμό των κρυφών επιπέδων (hidden layers) στο Νευρωνικό Δίκτυο. Τα παραδοσιακά Νευρωνικά Δίκτυα περιέχουν μόνο 2-3 κρυμμένα στρώματα, ενώ τα Βαθιά δίκτυα μπορούν να έχουν έως και 150. Τα μοντέλα Βαθιάς Μάθησης εκπαιδεύονται με τη χρήση μεγάλων συνόλων ετικετών δεδομένων και αρχιτεκτονικών Νευρωνικών Δικτύων που μαθαίνουν τις λειτουργίες απευθείας από τα δεδομένα χωρίς να χρειάζεται χειροκίνητη εξαγωγή χαρακτηριστικών.



Εικόνα 4.10
 Η δομή ενός πολύ-επίπεδου Νευρωνικού Δικτύου.

Ένας από τους πιο δημοφιλείς τύπους βαθιών Νευρωνικών Δικτύων είναι γνωστός ως Συνελικτικά Νευρωνικά Δίκτυα (CNN ή ConvNet). Το CNN περιγράφει τις γνωστές λειτουργίες με δεδομένα εισόδου και χρησιμοποιεί 2D στρώματα περιστροφής, κάνοντας αυτή την αρχιτεκτονική κατάλληλη για την επεξεργασία δεδομένων 2D, όπως είναι οι εικόνες.

Τα CNN εξαλείφουν την ανάγκη για χειροκίνητη εξαγωγή χαρακτηριστικών, οπότε δεν απαιτείται προσδιορισμός των λειτουργιών που απαιτούνται για την ταξινόμηση των εικόνων, μια και το CNN λειτουργεί εξαγοντας χαρακτηριστικά απευθείας από τις εικόνες. Τα χαρακτηριστικά αυτά δεν απαιτείται να προ-καταχωρηθούν στο σύστημα καθώς εξαγονται καθώς το Δίκτυο εκπαιδεύεται πάνω σε ένα σύνολο δεδομένων που λαμβάνεται διαρκώς. Η δυνατότητα αυτόματης εξαγωγής χαρακτηριστικών, καθιστά τα μοντέλα Βαθιάς Μάθησης ιδιαίτερα αξιόπιστα και κατάλληλα για λειτουργίες Υπολογιστικής Όρασης και ταξινόμησης αντικειμένων.



Εικόνα 4.11
 Παράδειγμα εφαρμογής αυτόματης εξαγωγής χαρακτηριστικών και προτύπων ενός Συνελικτικού Νευρωνικού Δικτύου σε εικόνα.

Τα CNNs μαθαίνουν να ανιχνεύουν και να ανιχνεύουν διαφορετικά χαρακτηριστικά μιας δομής χρησιμοποιώντας δεκάδες ή εκατοντάδες κρυμμένα στρώματα. Κάθε κρυμμένη στρώση αυξάνει την πολυπλοκότητα των χαρακτηριστικών γνώσης της εικόνας. Για παράδειγμα, το πρώτο κρυφό στρώμα θα μπορούσε να μάθει να ανιχνεύει τις άκρες και το τελευταίο να μάθει να ανιχνεύει πιο περίπλοκα και ακαθόριστα σχήματα, ειδικά στις περιπτώσεις αναγνώρισης αντικειμένων κτλ.

Στο επόμενο κεφάλαιο, πραγματοποιείται μία εκτενής ανάλυση των Συνελικτικών Νευρωνικών Δικτύων και της Βαθιάς Μάθησης εν γένει, εμβαθύνοντας και αναλύοντας όλες τις βασικές πτυχές του τεχνικού υπόβαθρου και της αρχιτεκτονικής τους. [201], [199], [191]

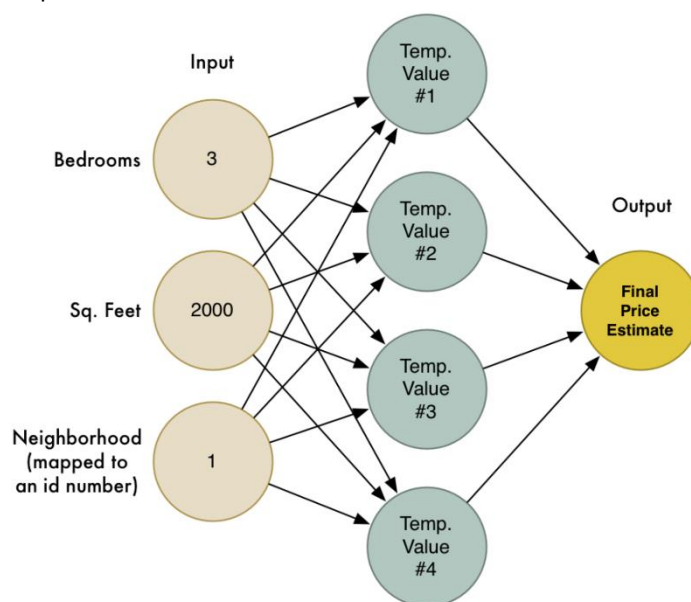
5 Εισαγωγή στις Τεχνικές Βαθιάς Μάθησης

Τα τελευταία χρόνια, αναπτύχθηκε μια ενδιαφέρουσα προσέγγιση για την αναγνώριση αντικειμένων, με χρήση των *Deep Convolutional Neural Networks*. Η αναγνώριση των αντικειμένων από ένα υπολογιστικό σύστημα, βασίζεται στις αρχές της ανάλυσης και επεξεργασίας των εικόνων/frames που παρέχει μία κάμερα (κλειστό κύκλωμα τηλεόρασης). Επομένως, η εφαρμογή των Νευρωνικών Δικτύων στα δομικά στοιχεία μίας εικόνας (εικονοστοιχεία / pixels) μπορεί να επιφέρει σπουδαία αποτελέσματα στον εντοπισμό και την παρακολούθηση αντικειμένων.

Στα επόμενα τμήματα του παρόντος κεφαλαίου, θα επιχειρηθεί η παρουσίαση και ανάλυση της δομής ενός Βαθιού Νευρωνικού Δικτύου που είναι σε θέση μετά από αντίστοιχη εκπαίδευση, να αναγνωρίσει το πρότυπο εκπαίδευσης σε κάποια εικόνα. [206], [173], [201]

5.1 Ανίχνευση Προτύπου Εικόνας

Όπως παρουσιάστηκε στο προηγούμενο κεφάλαιο, τα Νευρωνικά Δίκτυα είναι σε θέση να παρέχουν λύση σε προβλήματα των οποίων οι συναρτήσεις μεταφοράς δεν είναι δεδομένες, χρησιμοποιώντας πολλούς απλούς νευρώνες. Έτσι, ένα απλό γραμμικό πρόβλημα υπολογισμού το του κόστους ενός διαμερίσματος με άγνωστη τη συνάρτηση μεταφοράς, θα μπορούσε να υλοποιηθεί με εφαρμογή του ακόλουθου Νευρωνικού Δικτύου και την εξαγωγή των κατάλληλων βαρών των νευρώνων μετά από κατάλληλη εκπαίδευση αυτού.



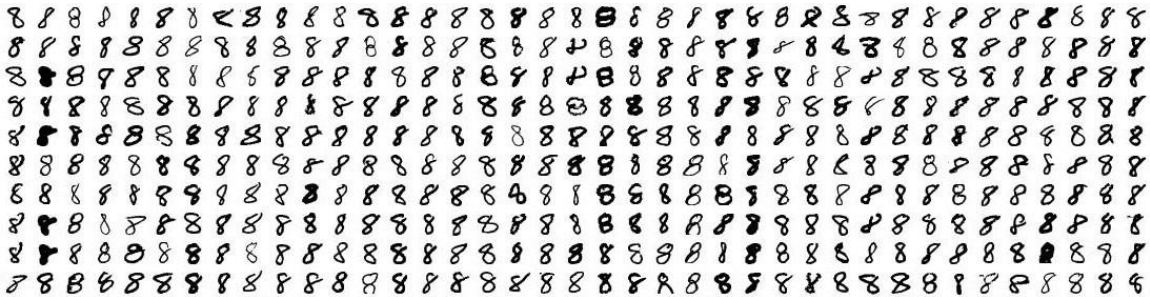
Εικόνα 5.1

Το Νευρωνικό Δίκτυο επίλυσης ενός προβλήματος κοστολόγησης διαμερίσματος.

[206]

Η κεντρική ιδέα της Μηχανικής Μάθησης έγκειται στο γεγονός ότι οι γενικοί αλγόριθμοι των απλών Νευρωνικών Δικτύων, μπορούν να επαναχρησιμοποιηθούν με διαφορετικά όμως δεδομένα για την επίλυση προβλημάτων διαφορετικού τύπου. Επομένως, για να καταστεί εφικτός ο εντοπισμός μίας εικόνας, θα μπορούσε να χρησιμοποιηθεί κάποιο Νευρωνικό Δίκτυο όπως αυτό που παρουσιάστηκε παραπάνω για την επίλυση ενός προβλήματος κοστολόγησης διαμερίσματος. Σε επόμενο βήμα, θα επιχειρηθεί η τροποποίηση του παραπάνω Νευρωνικού Δικτύου έτσι ώστε να χρησιμοποιηθεί για τον εντοπισμό ενός προτύπου εικόνας, του χειρόγραφου αριθμού "8".

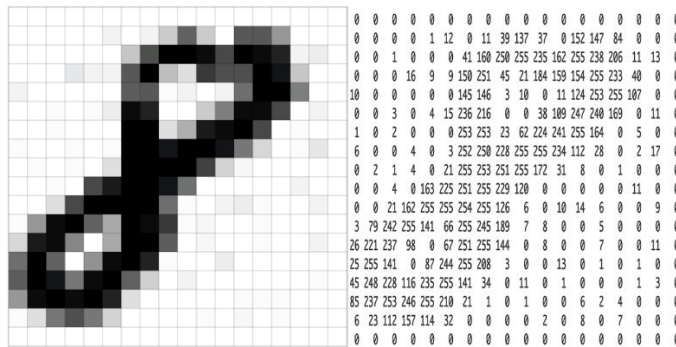
Όπως είναι γνωστό, η Μηχανική Μάθηση απαιτεί ένα πολύ μεγάλο πλήθος δεδομένων προς εκμάθηση ώστε να λειτουργήσει αποτελεσματικά. Έτσι, για την εφαρμογή του παραπάνω Νευρωνικού Δικτύου, απαιτείται η ύπαρξη ενός μεγάλου data set, δηλαδή η παροχή προς το Δίκτυο πολλών δεδομένων εκπαίδευσης, άρα πολλών χειρόγραφων αριθμών "8".



Εικόνα 5.2

Το data set εκπαίδευσης του δικτύου, αποτελούμενο από 60.000 εικόνες χειρόγραφων ψηφίων "8", διαστάσεων 18x18 ανά εικόνα.

Το Νευρωνικά Δίκτυα λαμβάνουν στην είσοδο τους τα αριθμητικά δεδομένα που είναι γνωστά από το περιβάλλον του προβλήματος. Ωστόσο, τα δεδομένα που προέρχονται από μία εικόνα δεν είναι αριθμητικά αλλά οπτικά. Η τροφοδότηση του Νευρωνικού Δικτύου μίας εικόνας, απαιτεί και πάλι την είσοδο αριθμητικών δεδομένων. Τα δεδομένα αυτά είναι ο χρωματικός αριθμός που χαρακτηρίζει κάθε ένα pixel της εικόνας.

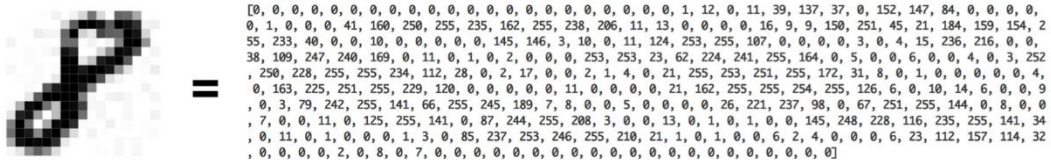


Εικόνα 5.3

Η αντικατάσταση των pixel που αποτελούν κάποιον χειρόγραφο αριθμό «8», με αριθμητικές τιμές χρωματικής διαβάθμισης.

[206]

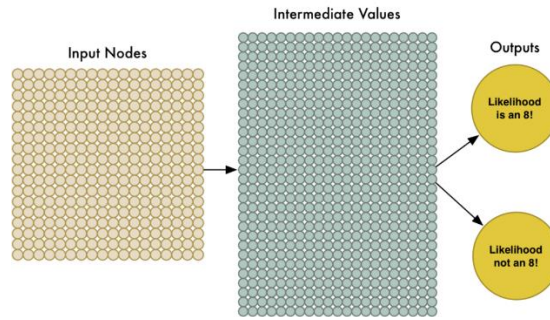
Για να τροφοδοτήσουμε μια εικόνα στο Νευρωνικό Δίκτυο, απλά αντιμετωπίζουμε την εικόνα 18x18 pixel ως μια σειρά / πίνακα, από 324 αριθμούς:



Εικόνα 5.4

Ο χαρακτηριστικός αριθμητικός πίνακας του εξεταζόμενου χειρόγραφου αριθμού.

[206]



Εικόνα 5.5

Δόμηση του Νευρωνικού Δικτύου βάσει των απαιτούμενων εισόδων, μετά την αριθμητική αναπαράσταση των pixel της [206]

Επομένως, για να χειριστούμε τις 324 εισόδους του δικτύου, μεγεθύνουμε το Νευρωνικό Δίκτυο έτσι ώστε να διαθέτει 324 κόμβους εισόδου. Ακόμα, το Νευρωνικό Δίκτυο θα διαθέτει δύο εξόδους, μία για την πρόβλεψη της πιθανότητας η εικόνα είναι "8" και μία για την πρόβλεψη της πιθανότητας η εικόνα να μην είναι "8". Έχοντας μια ξεχωριστή έξοδο για κάθε τύπο αντικείμενου που θέλουμε να αναγνωρίσουμε, μπορούμε να χρησιμοποιήσουμε ένα Νευρωνικό Δίκτυο για να ταξινομήσουμε αντικείμενα σε ομάδες.

Ακολούθως, σειρά έχει η υλοποίηση της εκπαίδευσης του δικτύου. Η εκπαίδευση αυτή, βασιζόμενη στη αρχιτεκτονική του δικτύου, θα πρέπει να παρέχει στο Δίκτυο εικόνες και των δύο κατηγοριών (εικόνες "8" και εικόνες "όχι-8"), παρέχοντας παράλληλα μαζί με κάθε εκπαιδευτική εικόνα τροφοδότησης, και την αντίστοιχη πιθανότητα η εικόνα αυτή ένα "8" (πιθανότητα 100%) και την αντίστοιχη πιθανότητα η εικόνα αυτή να μην είναι "8" (πιθανότητα 0%). Για τις εικόνες που δεν παρουσιάζουν κάποιο "8", ισχύουν τα αντίστροφα ποσοστά.

Έτσι ένα σύνολο δεδομένων εκπαίδευσης (dataset), θα μπορούσε να περιλαμβάνει τα ακόλουθα δεδομένα. [206], [195], [32], [49], [55], [61], [62], [62]



Εικόνα 5.6

Το dataset εκπαίδευσης του Νευρωνικού Δικτύου αναγνώρισης του χειρόγραφου αριθμού «8».

5.2 Tunnel Vision

Με εφαρμογή της παραπάνω τεχνικής καθίσταται δυνατή η αναγνώριση ενός προτύπου, εφόσον το αντίστοιχο Νευρωνικό Δίκτυο έχει εκπαιδευτεί κατάλληλα και επαρκώς. Έτσι, όταν ο χειρόγραφος αριθμός «8» βρίσκεται ακριβώς στο κέντρο της εξεταζόμενης εικόνας, μπορεί με επιτυχία να εντοπιστεί.



Εικόνα 5.7

Η αναγνώριση του αριθμού 8 και η μη αναγνώριση του αντίστοιχα, σε μία εξεταζόμενη εικόνα, μετά από κατάλληλη εκπαίδευση του Νευρωνικού Δικτύου. [206]

Αναγνώριση πεζών από αυτόνομο ρομποτικό όχημα με χρήση τεχνητής νοημοσύνης και τεχνικών βαθιάς μάθησης

Ωστόσο, το Νευρωνικό Δίκτυο δεν δύναται να προχωρήσει σε αναγνώριση ή μη αναγνώριση ενός αριθμού, όταν ο αριθμός αυτός δεν είναι απόλυτα κεντραρισμένος στην εικόνα.



Εικόνα 5.8

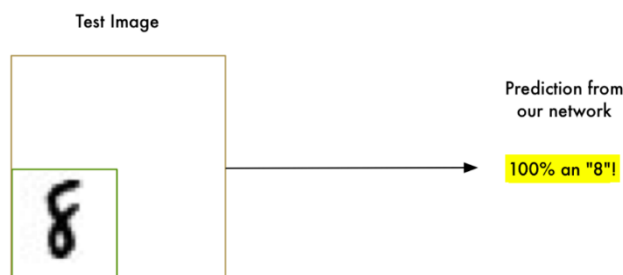
Η μη-αναγνώριση του προτύπου όταν αυτό εμφανίζεται σε διαφορετική θέση από αυτήν των δεδομένων εκπαίδευσης.

[206]

Αυτό συμβαίνει επειδή το Νευρωνικό Δίκτυο εκπαιδεύτηκε μόνο την περίπτωση ενός τέλεια κεντραρισμένου "8". Έτσι, γίνεται άμεσα αντιληπτό ότι το υφιστάμενο Νευρωνικό Δίκτυο θα πρέπει να τροποποιηθεί με κάποιον τρόπο έτσι ώστε να αναγνωρίζει το πρότυπο της εκπαίδευσης, ακόμα και στην περίπτωση όπου αυτό δεν βρίσκεται τέλεια κεντραρισμένο στην εικόνα ή όταν βρίσκεται σε κάποια άλλη προοπτική. [206]

5.2.1 Αναζήτηση τη μέθοδο συρόμενου πλαισίου

Η προηγούμενη μέθοδος, βασιζόταν στην ανίχνευση ενός μοναδικού προτύπου στο κέντρο της εικόνας. Ωστόσο, χρησιμοποιώντας μία παραλλαγή αυτής, θα μπορούσαμε να ελέγξουμε την δοθείσα εικόνα για την περιεκτικότητα της σε περισσότερα από ένα πρότυπα ταυτόχρονα, σε διαφορετικές θέσεις της.



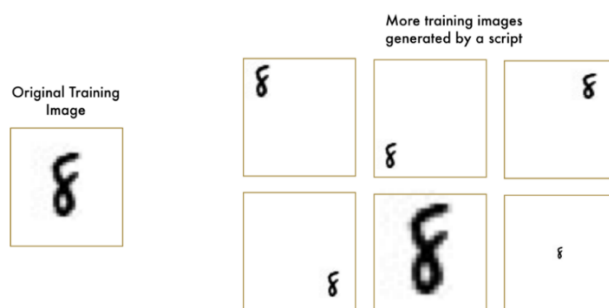
Εικόνα 5.9

Η αναγνώριση του προτύπου όταν αυτό εμφανίζεται σε διαφορετική θέση και διαφορετικό μέγεθος από τα αντίστοιχα των δεδομένων εκπαίδευσης, με χρήση της τεχνικής του συρόμενου πλαισίου. [206]

Η παραπάνω προσέγγιση ονομάζεται *συρόμενο πλαίσιο*. Λειτουργεί καλά σε ορισμένες περιορισμένες περιπτώσεις, αλλά σε γενικές γραμμές θεωρείται αναποτελεσματική καθώς θα πρέπει η ίδια εικόνα να ελέγχεται πολλαπλές φορές προκειμένου να εντοπιστούν ενδεχομένως πολλαπλά πρότυπα διαφορετικών μεγεθών. [206]

5.2.2 Χειρισμός περισσότερων δεδομένων εισόδου

Κατά την εκπαίδευση του δικτύου προηγουμένως, τροφοδοτήθηκαν εικόνες που περιλάμβαναν το πρότυπο τοποθετημένο στο κέντρο της εικόνας. Όπως έχει ήδη ειπωθεί, αυτό που είναι πολύ σημαντικό είναι το να εκπαιδευτεί το Δίκτυο έτσι ώστε να είναι ικανό να εντοπίζει το πρότυπο σε διάφορες θέσεις και σε διάφορα μεγέθη εντός της εικόνας. Έτσι, με χρήση της ακόλουθης μεθοδολογίας, είναι εφικτή η δημιουργία ενός σεναρίου το οποίο δημιουργεί νέες εικόνες οι οποίες περιέχουν το πρότυπο σε πολλαπλές θέσεις και σε πολλαπλά μεγέθη στην εικόνα.



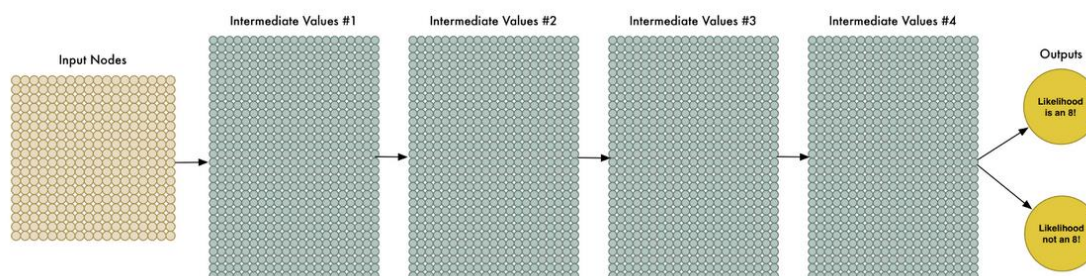
Εικόνα 5.10

Τα διαφορετικά δεδομένα εκπαίδευσης που παράγονται από τον αλγόριθμο, για το ίδιο πρότυπο σε διαφορετικά μεγέθη και θέσεις.

[206]

Χρησιμοποιώντας αυτή την τεχνική, μπορούμε εύκολα να δημιουργήσουμε μια ατέρμονη γεννήτρια δεδομένων εκπαίδευσης.

Όσο αυξάνονται τα δεδομένα εκπαίδευσης, τόσο πολυπλοκότερη και μη αποδοτική καθίσταται η λειτουργία του Νευρωνικού Δικτύου. Όμως, μεγαλώνοντας αντίστοιχα το Νευρωνικό Δίκτυο, επιλύεται η όποια υπερφόρτωση του και καθίσταται δυνατή η εκπαίδευση του δικτύου σε πολυπλοκότερα πρότυπα. Η αύξηση του δικτύου πραγματοποιείται τοποθετώντας στρώμα πάνω από το στρώμα των κόμβων:



Εικόνα 5.11

Η δομή διαστρωμάτωσης ενός Βαθιού Νευρωνικού Δικτύου.

[206]

Η παραπάνω αρχιτεκτονική χαρακτηρίζει την ίσως ταχύτερα εξελισσόμενη κατηγορία Νευρωνικών Δικτύων σήμερα, τα "Βαθιά Νευρωνικά Δίκτυα". Καλούνται έτσι επειδή φέρουν πολλαπλά επίπεδα συγκριτικά με ένα παραδοσιακό Νευρωνικό Δίκτυο.

Η αρχική προσέγγιση της συγκεκριμένης φιλοσοφίας, πραγματοποιήθηκε στα τέλη της δεκαετίας του '60. Έκτοτε, πραγματοποιήθηκαν αναρίθμητες προσπάθειες εκπαίδευσης Νευρωνικών Δικτύων, όμως η πολύ αργή εκπαίδευση του δικτύου, καθιστούσε τη μέθοδο αναποτελεσματική. Ωστόσο, η ανάπτυξη και αξιοποίηση των Μονάδων Επεξεργασίας Γραφικών των σύγχρονων Η/Υ στις εφαρμογές Νευρωνικών Δικτύων, επέφερε δραματική μείωση του χρόνου εκπαίδευσης. Η αρχιτεκτονική υλικού των Μονάδων Επεξεργασίας Τρισδιάστατων Γραφικών, υλοποιούν πολλαπλασιασμούς πινάκων πολύ γρήγορα συγκριτικά με μία Κεντρική Μονάδα Επεξεργασίας, και έτσι η εργασία με μεγάλα Νευρωνικά Δίκτυα κατέστη εφικτή και ιδιαίτερα αποτελεσματική.

Ωστόσο, αυτό που είναι πραγματικά μείζονος σημασίας, δεν είναι τόσο η δυνατότητα του υλικού να επεξεργάζεται τάχιστα τα δεδομένα αλλά η αποδοτική αξιοποίηση των δεδομένων και εικόνων αυτών από το Νευρωνικό Δίκτυο. Έτσι, η απαίτηση ξεχωριστής εκπαίδευσης ενός δικτύου έτσι ώστε να αναγνωρίσει ένα πρότυπο στο πάνω μέρος μίας εικόνας από μία ξεχωριστή εκπαίδευση του δικτύου ώστε να αναγνωρίζει το ίδιο πρότυπο στο κάτω μέρος της εικόνας, είναι κάτι σίγουρα αναποτελεσματικό.

Επομένως, θα πρέπει να εφαρμοστεί μία ενιαία μέθοδος εκπαίδευσης του Δικτύου έτσι ώστε να το καταστήσει ικανό να αναγνωρίζει το ίδιο πρότυπο σε διαφορετικά μεγέθη και σε διαφορετικές θέσεις στην δοθείσα εικόνα. [206]

5.3 Συνελικτικά Νευρωνικά Δίκτυα

Οι εικόνες που λαμβάνει, αξιολογεί και επεξεργάζεται ο ανθρώπινος εγκέφαλος, χαρακτηρίζονται από *ιεραρχική ή εννοιολογική δομή*. Έτσι, παρατηρώντας μία εικόνα, ο ανθρώπινος εγκέφαλος είναι σε θέση προβεί στην ανάλυση της *ιεραρχικής ή εννοιολογικής* τους δομής.



Εικόνα 5.12

Εικόνα που περιέχει πρότυπα προς αναγνώριση. [206]

Επομένως, ο άνθρωπος παρατηρώντας της παραπάνω εικόνα, καταλήγει στα κάτωθι συμπεράσματα, τα οποία βασίζονται στην ιεραρχική ανάλυση της εικόνας.

- Το έδαφος καλύπτεται από γρασίδι και τσιμέντο.
- Υπάρχει ένα παιδί.
- Το παιδί κάθεται πάνω σε ένα παιχνίδι.
- Το παιχνίδι βρίσκεται πάνω στο γρασίδι.

Εστιάζοντας τώρα στα πρότυπα αναγνώρισης, είναι αντιληπτό ότι από την παραπάνω εικόνα ο ανθρώπινος εγκέφαλος προβαίνει στην αναγνώριση ενός παιδιού.

Μέχρι στιγμής (βάσει της μέχρι και έως εδώ ανάλυσης), το Νευρωνικό Δίκτυο που αναλύθηκε προηγουμένως, δεν είναι σε θέση να πραγματοποιήσει μία τέτοια αναγνώριση. Αυτό συμβαίνει λόγω του ότι με την μέχρι στιγμής μεθοδολογία, το Νευρωνικό Δίκτυο αναγνωρίζει σαν τελείως ξεχωριστό ένα αντικείμενο στην πάνω θέση της εικόνας από το ίδιο αντικείμενο όταν βρεθεί στο κάτω μέρος της εικόνας. Επομένως δεν υπάρχει μέχρι στιγμή μία κοινή μέθοδος ταυτοποίησης του ίδιου και μοναδικού προτύπου σε κάθε θέση της εξεταζόμενης εικόνας.

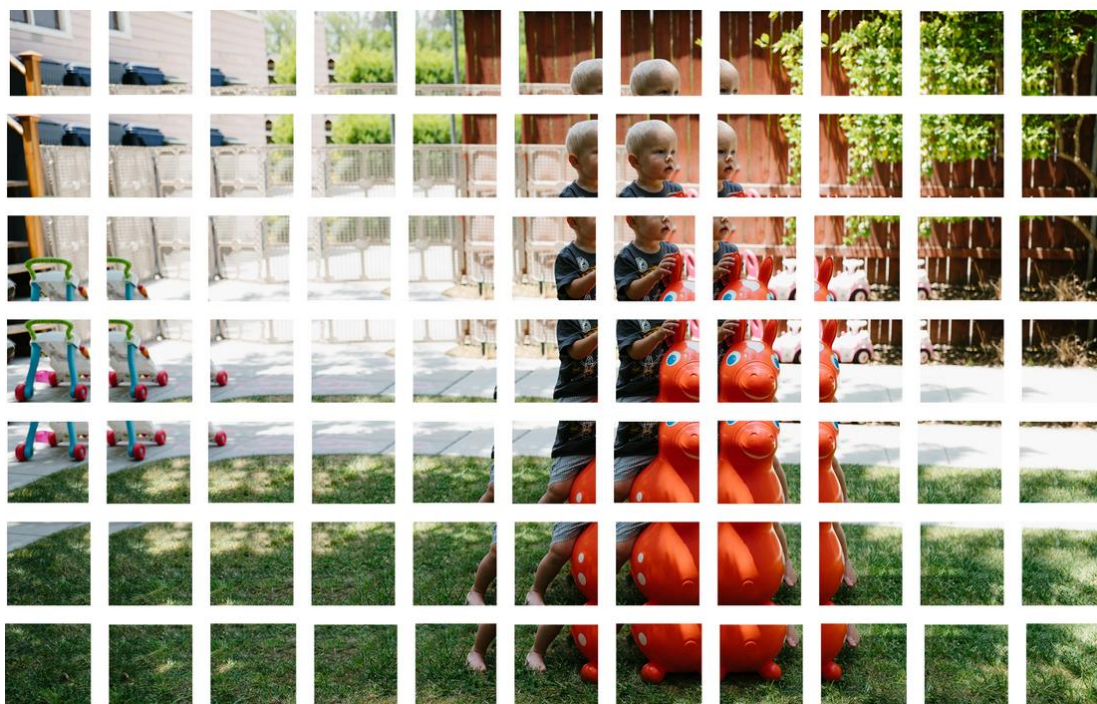
Οστόσο, πλέον έχει καταστεί εφικτή η αναγνώριση του κοινού προτύπου σε κάθε πιθανή θέση αυτού, με χρήση της τεχνικής που ονομάζεται *συνέλιξη (convolution)*. Η ιδέα της συνέλιξης προέρχεται εν μέρει από την επιστήμη των υπολογιστών και εν μέρει από τη βιολογία (παρακολούθηση βιολογικών μηχανισμών που ενεργοποιούνται κατά την διαδικασία της εγκεφαλικής αναγνώρισης). [206]. [3]

5.3.1 Πώς λειτουργούν τα Συνελικτικά Νευρωνικά Δίκτυα

Με την εφαρμογή της μεθόδου της συνέλιξης, αντί να προβούμε σε τροφοδότηση ενός πλήθους εικόνων ως ένα πλέγμα αριθμών στο Νευρωνικό Δίκτυο, θα προβούμε σε μια διαφορετική και πολύ αποτελεσματικότερη υλοποίηση η οποία στηρίζεται στην ιδέα ότι ένα αντικείμενο είναι το ίδιο ανεξάρτητα από την θέση εμφάνισής του στην εικόνα. Παρακάτω ακολουθεί η διαδικασία λειτουργίας των Συνελικτικών Δικτύων, εφαρμοζόμενα στην εικόνα που παρατέθηκε στην προηγούμενη ενότητα.

Βήμα 1: Τμηματοποίηση της εικόνας σε επικαλυπτόμενα αποσπάσματα

Ομοιοτρόπως με την μέθοδο των συρόμενων πλαισίων που αναλύθηκε προηγουμένως, εφαρμόζουμε ένα συρόμενο πλαίσιο σε ολόκληρη την επιφάνεια της αρχικής εικόνας και αποθηκεύουμε κάθε στιγμήτυπο ως ξεχωριστό απόσπασμα εικόνας.



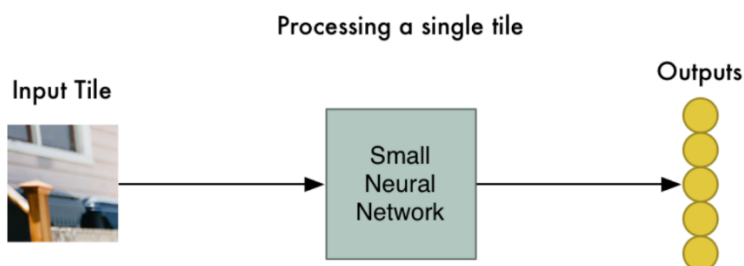
Εικόνα 5.13

Η εικόνα του παραδείγματος μετά την εφαρμογή της τεχνικής της *Συνέλιξης*. [206]

Με αυτό το τρόπο, η αρχική εικόνα μετατράπηκε σε 77 αποσπάσματα μικρού που συνδυαζόμενα αποτελούν την πλήρους μεγέθους αρχική εικόνα.

Βήμα 2: Τροφοδότηση του κάθε αποσπάσματος σε ένα μικρό Νευρωνικό Δίκτυο

Προηγουμένως, το Νευρωνικό Δίκτυο έλαβε στην είσοδο του μία εικόνα έτσι ώστε να παράγει πρόβλεψη για το αν η εικόνα αυτή ήταν ένα «8» ή όχι. Η ίδια διαδικασία ακολουθείται και εδώ, όμως για κάθε ένα μεμονωμένο απόσπασμα εικόνας.



Εικόνα 5.14

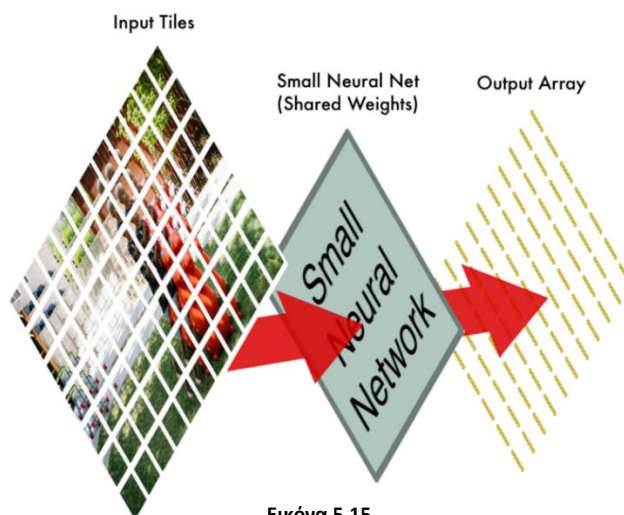
Η διαδικασία τροφοδότησης του *Νευρωνικού Δικτύου* με αποσπάσματα της αρχικής εικόνας. [206]

Αναγνώριση πεζών από αυτόνομο ρομπωτικό όχημα με χρήση τεχνητής νοημοσύνης και τεχνικών βαθιάς μάθησης

Ωστόσο, μια μεγάλη διαφορά αποτελεί το γεγονός ότι στο στάδιο αυτό θα διατηρηθούν τα ίδια βάρη Νευρωνικού Δικτύου για κάθε ένα ξεχωριστό απόσπασμα της πρωτότυπης εικόνας. Έτσι, κάθε απόσπασμα εικόνας αντιμετωπίζεται ακριβώς με τον ίδιο τρόπο. Σε περίπτωση όπου εντοπιστεί κάτι «ενδιαφέρον» σε οποιοδήποτε δεδομένο απόσπασμα εικόνας, θα το «σημειώσουμε» ως ενδιαφέρον.

Βήμα 3: Αποθήκευση Αποτελεσμάτων Κάθε Αποσπάσματος Εικόνας Σε Μια Νέα Συστοιχία

Για να διατηρηθεί αναλλοίωτη η καταγραφή της διάταξης των αρχικών αποσπασμάτων, πραγματοποιείται αποθήκευση του αποτελέσματος επεξεργασίας κάθε μεμονωμένου αποσπάσματος εικόνας, σε ένα πλέγμα που φέρει την ίδια διάταξη με αυτήν της αρχικής εικόνας.



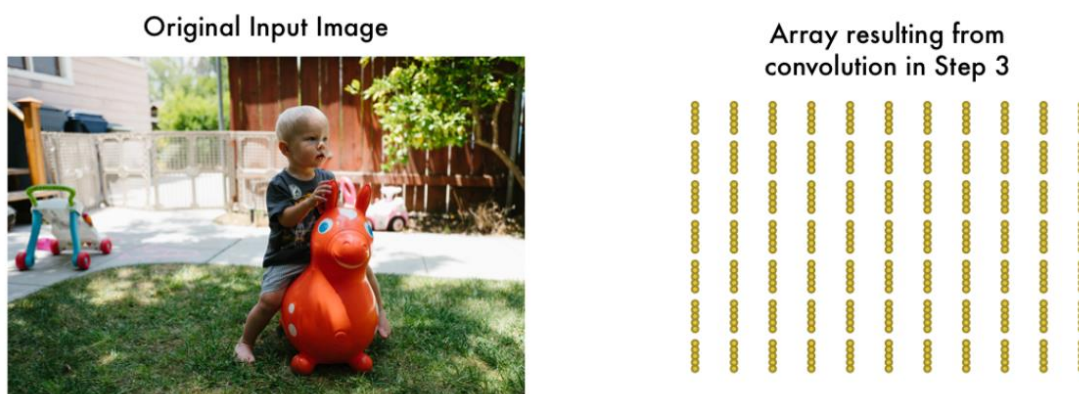
Εικόνα 5.15

Η διαδικασία αποθήκευσης των αποτελεσμάτων των *Μικρών Νευρωνικών Δικτύων* σε έναν νέο ενιαίο πίνακα. [206]

Έτσι, η συνολική διαδικασία της συνέλιξης ενώ ξεκίνησε με μια μεγάλη εικόνα, ολοκληρώθηκε με έναν ελαφρώς μικρότερο πίνακα ο οποίος καταγράφει ποια αποσπάσματα της αρχικής εικόνας ήταν περισσότερο ενδιαφέροντα.

Βήμα 4: Υποδειματοληψία

Το αποτέλεσμα του Βήματος 3, είναι μια συστοιχία που καταγράφει ποια αποσπάσματα της αρχικής εικόνας είναι τα πιο ενδιαφέροντα. Ωστόσο, αυτή η διάταξη παραμένει ακόμα αρκετά μεγάλη:

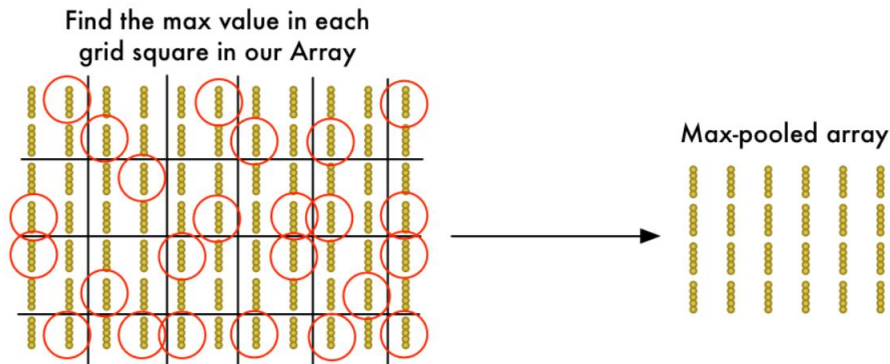


Εικόνα 5.16

Η αναπαράσταση την αρχικής εικόνας με τον πίνακα εξόδου του *Συνελικτικού Νευρωνικού Δικτύου*. [206]

Για να επιτευχθεί η περαιτέρω μείωση του μεγέθους της συστοιχίας, πραγματοποιείται *δειγματοληψία* βάσει ενός αλγορίθμου ο οποίος ονομάζεται *Max Pooling* (μέγιστη συγκέντρωση).

Κατά την διαδικασία της δειγματοληψίας, εξετάζεται κάθε υποπίνακας 2x2 του συνολικού πίνακα, και διατηρείται μόνο το στοιχείο που φέρει τη μεγαλύτερη τιμή.



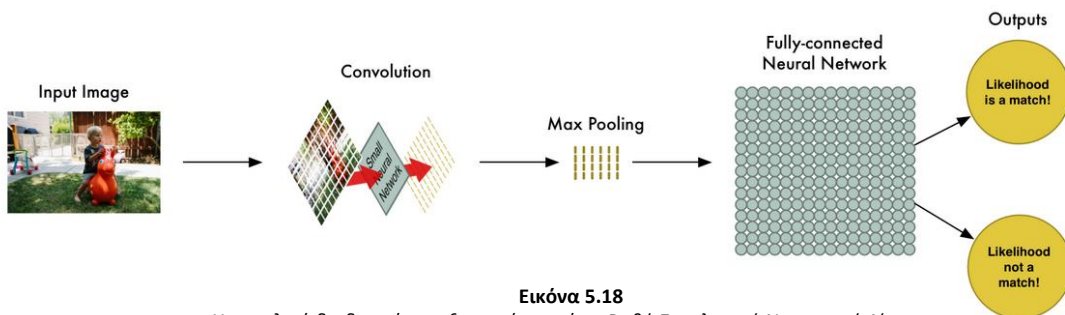
Εικόνα 5.17
Η διαδικασία του *Max Pooling*.
[206]

Με χρήση της παραπάνω μεθόδου διατηρείται μόνο ένα εκ των τεσσάρων στοιχείων κάθε υπό-πίνακα 2x2 και συνεπώς η συνολική διάσταση του τελικού πίνακα, είναι κατά 75% μικρότερη συναρτήσει του αρχικού. Με αυτό το τρόπο αφενός μειώνονται σημαντικά οι διαστάσεις του τελικού πίνακα και αφετέρου, διατηρούνται τα «σημαντικά» τμήματα της αρχικής εικόνας.

Τελικό Βήμα: Πρόβλεψη

Μέχρι στιγμής, έχει παραχθεί μία τελική εικόνα κατά πολύ μικρότερη από την αρχική. Η εικόνα αυτή ουσιαστικά αποτελεί έναν πίνακα αριθμών. Αυτός ο μικρός πίνακας αριθμών, θα αποτελέσει την είσοδο ενός άλλου *Νευρωνικού Δικτύου*. Αυτό το τελικό Νευρωνικό Δίκτυο, θα αποφασίσει το αν η εικόνα ταιριάζει ή όχι στο πρότυπο εκπαίδευσης. Το συγκεκριμένο στάδιο, αποκαλείται ως "πλήρως συνδεδεμένο" Δίκτυο.

Έτσι, η συνολική διαδικασία αποτελείται από πέντε συνολικά στάδια, τα οποία παρουσιάζονται στην ακόλουθη εικόνα.



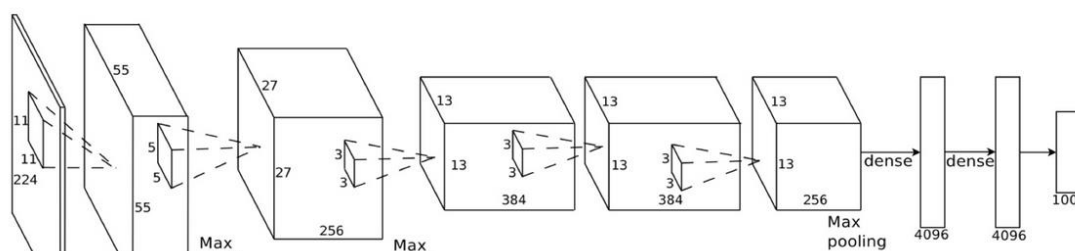
Πολυστρωματικά Συνελκτικά Νευρωνικά Δίκτυα

Η διαδικασία ανάλυσης και ταυτοποίησης μίας εικόνας, αποτελείται από τα βήματα που παρουσιάστηκαν προηγουμένως: συνέλιξη, μέγιστη συγκέντρωση και τελικά ένα πλήρως συνδεδεμένο Δίκτυο.

Κατά την επίλυση προβλημάτων στον πραγματικό κόσμο, αυτά τα βήματα μπορούν να συνδυάζονται και να στοιβάζονται πολλές φορές, δημιουργώντας πολλαπλά στρώματα συνέλιξης. Έτσι, μία αρχικά μεγάλη εικόνα μπορεί να επεξεργαστεί σε πολλαπλά επίπεδο μέχρις ότου παραχθεί το τελικό επιθυμητό αποτέλεσμα. Σε κάθε στρώμα συνέλιξης, το Νευρωνικό Δίκτυο εκπαιδεύεται ώστε να αναγνωρίζει κάποιο διαφορετικό στοιχείο του προτύπου.

Έτσι, στο πρώτο στρώμα συνέλιξης, το Νευρωνικό Δίκτυο θα μπορούσε να εκπαιδευτεί ώστε να αναγνωρίζει τις αιχμηρές άκρες, το δεύτερο στρώμα θα μπορούσε να εκπαιδευτεί ώστε να αναγνωρίζει κάποιο συγκεκριμένο χαρακτηριστικό του προτύπου (πχ στην αναγνώριση πτηνών, τα ράμφη), χρησιμοποιώντας τις «γνώσεις» του από αιχμηρές άκρες του προηγούμενου στρώματος, το τρίτο στρώμα θα μπορούσε να αναγνωρίσει ενδεχομένως ολόκληρα το πρότυπο (πχ πτηνά) αξιοποιώντας τις «γνώσεις του» από το προηγούμενο στρώμα (πχ ράμφη) κλπ.

Στην ακόλουθη εικόνα παρουσιάζεται ένα πραγματικό Βαθύ Συνελικτικό Δίκτυο.



Εικόνα 5.19

Η πραγματική αρχιτεκτονική ενός Βαθιού Συνελικτικού Νευρωνικού Δικτύου. [206]

Στο Βαθύ Συνελικτικό Δίκτυο που παρουσιάζεται στην παραπάνω εικόνα, η είσοδος του δικτύου είναι μία εικόνα διάστασης 224x224 εικονοστοιχείων. Στο Δίκτυο αυτό αρχικά εφαρμόζεται συνέλιξη δύο στρωμάτων, σε επόμενο επίπεδο εφαρμόζεται τεχνική μέγιστης συγκέντρωσης δύο στρωμάτων, έπειτα τεχνική συνέλιξης τριών στρωμάτων, στη συνέχεια εφαρμόζεται τεχνική μέγιστης συγκέντρωσης και τέλος τα δύο πλήρως συνδεδεμένα επίπεδα οδηγούν στην ταξινόμηση της εικόνας εισόδου σε μία από τις 1000 κατηγορίες προτύπων.

Δόμηση του Κατάλληλου Δικτύου

Η δόμηση του κατάλληλου δικτύου αποτελεί μία διαδικασία μείζονος σημασίας για την επίλυση του προβλήματος ταξινόμησης. Η επιλογή της κατάλληλης αρχιτεκτονικής του δικτύου απαιτεί αρκετή εμπειρία, υψηλή ανάλυση των παραμέτρων ταξινόμησης και πολλαπλές δοκιμές. [206], [3]

5.4 Σχεδίαση Νευρωνικού Δικτύου Εντοπισμού Πτηνών

Στην ακόλουθη ενότητα, θα επιχειρηθεί η εφαρμογή των παραπάνω τεχνικών, για την δόμηση ενός Νευρωνικού Δικτύου που θα είναι σε θέση να επιτελέσει αναγνώριση πτηνών σε μία εισερχόμενη στο Δίκτυο εικόνα.

Προτού υλοποιηθεί ο σχεδιασμός του Νευρωνικού Δικτύου, απαιτείται η ύπαρξη των δεδομένων εκπαίδευσης του Νευρωνικού Δικτύου (data set). Έτσι, θα χρησιμοποιηθεί το γνωστό δωρεάν πακέτο δεδομένων εκπαίδευσης *CIFAR10 data set*, το οποίο περιέχει 6.000 εικόνες πτηνών και 52.000 φωτογραφίες από λοιπά αντικείμενα που δεν είναι πτηνά. Για την βελτιστοποίηση της εκπαίδευσης του Νευρωνικού Δικτύου, θα εισάγουμε στα υπάρχοντα πρότυπα εκπαίδευσης και ακόμα ένα πακέτο προτύπων εκπαίδευσης, το *Caltech-UCSD Birds-200-2011 data set*, το οποίο περιλαμβάνει ακόμα 12.000 φωτογραφίες πτηνών.

Στην ακόλουθη εικόνα παρουσιάζονται μερικά πρότυπα από το data set (εικόνες πτηνών).



Εικόνα 5.20

Εικόνες που αποτελούν το θετικό πρότυπο του Πακέτου Προτύπων Εκπαίδευσης. [206]

Στην ακόλουθη εικόνα παρουσιάζονται μερικά πρότυπα από το data set (εικόνες μη-πτηνών).



Εικόνα 5.21

Εικόνες που αποτελούν το αρνητικό πρότυπο του Πακέτου Προτύπων Εκπαίδευσης. [206]

^[206] Το πακέτο προτύπων εκπαίδευσης όπου χρησιμοποιήθηκε, αν και θα μπορούσε να υλοποιήσει την εκπαίδευση του δικτύου, σε γενικές γραμμές δεν θεωρείται επαρκές καθώς αποτελείται από εικόνες χαμηλής ανάλυσης και επίσης το πλήθος των 72.000 εικόνων είναι πολύ μικρό για τις εφαρμογές του πραγματικού κόσμου. Η εκπαίδευση ενός Νευρωνικού Δικτύου που θα παρείχε ύψιστες ακρίβειες και σχεδόν μηδενική πιθανότητα σφάλματος, θα απαιτούσε από μερικά εκατομμύρια έως και πολλά δισεκατομμύρια πρότυπων εικόνων εκπαίδευσης. Στη Μηχανική Μάθηση, οι βελτιστοποίηση των αλγορίθμων είναι άρρηκτα συνδεδεμένη με την κατοχή μεγάλου πλήθους και υψηλής ανάλυσης πακέτων προτύπων εκπαίδευσης.

Για την δόμηση του Νευρωνικού Δικτύου αναγνώρισης πτηνών, θα χρησιμοποιηθεί το *TFLearn*. Το *TFLearn* αποτελεί ένα πλαίσιο λειτουργικής υποστήριξης της βιβλιοθήκη Βαθιάς Μάθησης *TensorFlow* της *Google* που περιλαμβάνει ένα απλοποιημένο API η χρήση του οποίου απλοποιεί σε μεγάλο βαθμό την σχεδίαση Συνελκτικών Νευρωνικών Δικτύων.

5.4.1 Υλοποίηση Νευρωνικού Δικτύου αναγνώρισης πτηνών

Ακολούθως παρατίθεται το πρόγραμμα σχεδιασμού και εκπαίδευσης του Βαθιού Συνελκτικού Νευρωνικού Δικτύου που επιτελεί αναγνώριση πτηνού στη εισερχόμενη στο Δίκτυο εικόνα.

```

1. # -*- coding: utf-8 -*-
2.
3. """
4. Based on the tflearn example located here:
5. https://github.com/tflearn/tflearn/blob/master/examples/images/convnet_cifar10.py
6. """
7. from __future__ import division, print_function, absolute_import
8.
9. # Import tflearn and some helpers
10. import tflearn
11. from tflearn.data_utils import shuffle
12. from tflearn.layers.core import input_data, dropout, fully_connected
13. from tflearn.layers.conv import conv_2d, max_pool_2d
14. from tflearn.layers.estimator import regression
15. from tflearn.data_preprocessing import ImagePreprocessing
16. from tflearn.data_augmentation import ImageAugmentation

```

Αναγνώριση πεζών από αυτόνομο ρομποτικό όχημα με χρήση τεχνητής νοημοσύνης και τεχνικών βαθιάς μάθησης

```

17. import pickle
18.
19. # Load the data set
20. X, Y, X_test, Y_test = pickle.load(open("full_dataset.pkl", "rb"))
21.
22. # Shuffle the data
23. X, Y = shuffle(X, Y)
24.
25. # Make sure the data is normalized
26. img_prep = ImagePreprocessing()
27. img_prep.add_featurewise_zero_center()
28. img_prep.add_featurewise_stdnorm()
29.
30. # Create extra synthetic training data by flipping, rotating and blurring the
31. # images on our data set.
32. img_aug = ImageAugmentation()
33. img_aug.add_random_flip_leftright()
34. img_aug.add_random_rotation(max_angle=25.)
35. img_aug.add_random_blur(sigma_max=3.)
36.
37. # Define our network architecture:
38.
39. # Input is a 32x32 image with 3 color channels (red, green and blue)
40. network = input_data(shape=[None, 32, 32, 3],
41. data_preprocessing=img_prep,
42. data_augmentation=img_aug)
43.
44. # Step 1: Convolution
45. network = conv_2d(network, 32, 3, activation='relu')
46.
47. # Step 2: Max pooling
48. network = max_pool_2d(network, 2)
49.
50. # Step 3: Convolution again
51. network = conv_2d(network, 64, 3, activation='relu')
52.
53. # Step 4: Convolution yet again
54. network = conv_2d(network, 64, 3, activation='relu')
55.
56. # Step 5: Max pooling again
57. network = max_pool_2d(network, 2)
58.
59. # Step 6: Fully-connected 512 node neural network
60. network = fully_connected(network, 512, activation='relu')
61.
62. # Step 7: Dropout - throw away some data randomly during training to prevent over-
63. # fitting
64. network = dropout(network, 0.5)
65.
66. # Step 8: Fully-connected neural network with two outputs (0=isn't a bird, 1=is a
67. # bird) to make the final prediction
68. network = fully_connected(network, 2, activation='softmax')
69.
70. # Tell tflearn how we want to train the network
71. network = regression(network, optimizer='adam',
72. loss='categorical_crossentropy',
73. learning_rate=0.001)
74.
75. # Wrap the network in a model object
76. model = tflearn.DNN(network, tensorboard_verbose=0, checkpoint_path='bird-
77. classifier.tfl.ckpt')
78.
79. # Train it! We'll do 100 training passes and monitor it as it goes.
80. model.fit(X, Y, n_epoch=100, shuffle=True, validation_set=(X_test, Y_test),
81. show_metric=True, batch_size=96,
82. snapshot_epoch=True,
83. run_id='bird-classifier')
84.
85. # Save model when training is complete to a file
86. model.save("bird-classifier.tfl")
87.
88. print("Network trained and saved as bird-classifier.tfl!")

```

Η ταχύτητα εκπαίδευσης του συγκεκριμένου Νευρωνικού Δικτύου, εξαρτάται αποκλειστικά από την ταχύτητα επεξεργασίας του υπολογιστικού συστήματος. Μία σύγχρονη Μονάδα Επεξεργασίας Τρισδιάστατων Γραφικών, είναι σε θέση να υλοποιήσει την εκπαίδευση του συγκεκριμένου δικτύου σε λιγότερο από μία ώρα. Αντιθέτως, αν η εκπαίδευση του συγκεκριμένου δικτύου πραγματοποιηθεί από μία σύγχρονη Κεντρική Μονάδα Επεξεργασίας, ενδεχομένως απαιτηθούν αρκετές ώρες εκπαίδευσης.

Κατά την διάρκεια της εκπαίδευσης του δικτύου, η ακρίβεια αυτού αυξάνεται. Έτσι, μετά από μία εκτέλεση η ακρίβεια του δικτύου ήταν 75,4%, μετά από δέκα επαναλήψεις η ακρίβεια του δικτύου αυξήθηκε σε 91,7%, ενώ μετά από 50 επαναλήψεις η ακρίβεια του δικτύου ήταν 95,5%. [206]

5.4.2 Έλεγχος του δικτύου

Η πραγματική αποτελεσματικότητα και ακρίβεια ενός δικτύου, μπορεί να κριθεί μόνο κατόπιν πολλαπλών δοκιμών. Το σύνολο προτύπων εκπαίδευσης της εφαρμογής που δημιουργήθηκε προηγουμένως, επέλεξε 15.000 εικόνες προς επικύρωση. Μετά από έλεγχο αυτών των εικόνων από το Δίκτυο, το Δίκτυο προέβλεψε αληθώς τη σωστή απάντηση στο 95% των δοκιμών. [206]

5.4.3 Αξιολόγηση της ακρίβειας

Το δίκτυο φαίνεται πως παρουσιάζει ακρίβεια της τάξης του 95%. Ωστόσο η ακρίβεια κατά 95% ενδεχομένως να μην θεωρείται επαρκής. Αυτό συμβαίνει διότι όλα τα ποσοστά είναι σχετικά. Επομένως, αν το Δίκτυο ακρίβειας κατά 95%, εξέταζε κάποιο σύνολο προτύπων στο οποίο η περιεκτικότητα πτηνών ήταν 5% και η περιεκτικότητα μη-πτηνών ήταν 95%, τότε το Δίκτυο θα ταυτοποιούσε ασφάλεια μόνο τα δεδομένα που δεν περιλάμβαναν το πρότυπο (μη-πτηνό) ενώ θα ήταν αναξιόπιστο στην ταυτοποίηση προτύπων (πτηνών).

Επομένως αυτό που αποτελεί τον ουσιαστικό δείκτη ακρίβειας του δικτύου, δεν είναι το συνολικό ποσοστό ακρίβειας των αληθών προβλέψεων, το συνολικό ποσοστό ακρίβειας των αληθών προβλέψεων συναρτήσει των ποσοστών κατανομής των προτύπων.

Παρακάτω ακολουθεί η κατηγοριοποίηση των προβλέψεων που παράγει το Δίκτυο, σε τέσσερις βασικές κατηγορίες.

1. Παρακάτω παρατίθενται μερικά δείγματα (πτηνά) όπου το δίκτυο αναγνώρισε ορθά ως πτηνά. Αυτά καλούνται **Αληθώς Θετικά**:



2. Παρακάτω παρατίθενται μερικά δείγματα (πτηνά) όπου το δίκτυο αναγνώρισε ορθά ως μη-πτηνά. Αυτά καλούνται **Αληθώς Αρνητικά**:



3. Παρακάτω παρατίθενται μερικά δείγματα (πτηνά) όπου το δίκτυο αναγνώρισε εσφαλμένα ως πρότυπα (πουλιά) ενώ δεν ήταν. Αυτά καλούνται **Ψευδώς Θετικά**:



4. Και τέλος, παρακάτω παρατίθενται μερικά δείγματα (πτηνά) όπου το δίκτυο αναγνώρισε εσφαλμένα ως μη πρότυπα (μη-πτηνά) ενώ ήταν. Αυτές καλούνται **Ψευδώς Αρνητικά**:



Χρησιμοποιώντας το σύνολο επικύρωσης των 15.000 εικόνων, παρακάτω παρουσιάζεται το πόσες φορές το Δίκτυο προέβλεψε ορθά το αποτέλεσμα.

Results for 15,000 Validation Images

(6000 images are birds, 9000 images are not birds)

	Predicted 'bird'	Predicted 'not a bird'
Bird	5,450 <small>True Positives</small>	550 <small>False Negatives</small>
Not a Bird	162 <small>False Positives</small>	8,838 <small>True Negatives</small>

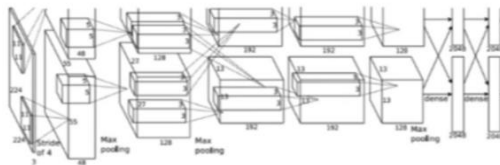
Στην κάτωθι εικόνα παρουσιάζονται τα ποσοστά επιτυχίας των προβλέψεων της ακρίβειας και της ανάκλησης. Τα ποσοστά αυτά παρέχουν μία σαφέστερη εικόνα της ορθότητας του Νευρωνικού Δικτύου.

Precision <small>If we predicted 'bird', how often was it really a bird?</small>	97.11% <small>(True Positives ÷ All Positive Guesses)</small>
Recall <small>What percentage of the actual birds did we find?</small>	90.83% <small>(True Positives ÷ Total Birds in Dataset)</small>

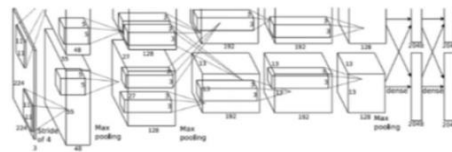
Από τον παραπάνω πίνακα προκύπτει ότι στο 97% των περιπτώσεων ταυτοποιήθηκε αληθώς το εξεταζόμενο αντικείμενο ως "Πτηνό". Ακόμα, προκύπτει ότι το εντοπίστηκε το 90% του συνολικού πλήθους των προτύπων (πτηνά). Επομένως, αν και το Δίκτυο δεν λειτουργεί άριστα, παρέχει ένα αρκετά υψηλό ποσοστό ορθών προβλέψεων εντοπισμού του προτύπου. [206]

5.5 Εισαγωγή στις Σύγχρονες Τεχνικές Υπολογιστικής Όρασης

Η Υπολογιστική Όραση, αποτελεί ένα διεπιστημονικό πεδίο που έχει κερδίσει τεράστια προσοχή τα τελευταία χρόνια, ειδικά μετά την έναρξη ερευνών γύρω από τα αυτόνομα ρομποτικά οχήματα. Ένα άλλο αναπόσπαστο μέρος της Υπολογιστικής Όρασης, είναι η ανίχνευση αντικειμένων. Η διαφορά μεταξύ των αλγορίθμων ανίχνευσης αντικειμένων και αλγορίθμων ταξινόμησης αντικειμένων, είναι ότι στους αλγόριθμους ανίχνευσης, προσπαθούμε να σχεδιάσουμε ένα πλαίσιο οριοθέτησης γύρω από το αντικείμενο ενδιαφέροντος, έτσι ώστε να το εντοπίσουμε εντός της εικόνας. Επίσης, δεν είναι απαραίτητο να σχεδιάζεται απαραίτητα μόνο ένα πλαίσιο οριοθέτησης ανά εικόνα. Σε περίπτωση ανίχνευσης πολλαπλών αντικειμένων, θα μπορούσαν να υπάρχουν πολλά πλαίσια οριοθέτησης που εκπροσωπούν διαφορετικά αντικείμενα ενδιαφέροντος μέσα στην εικόνα.



CAT: (x, y, w, h)



DUCK: (x, y, w, h)
DUCK: (x, y, w, h)
....

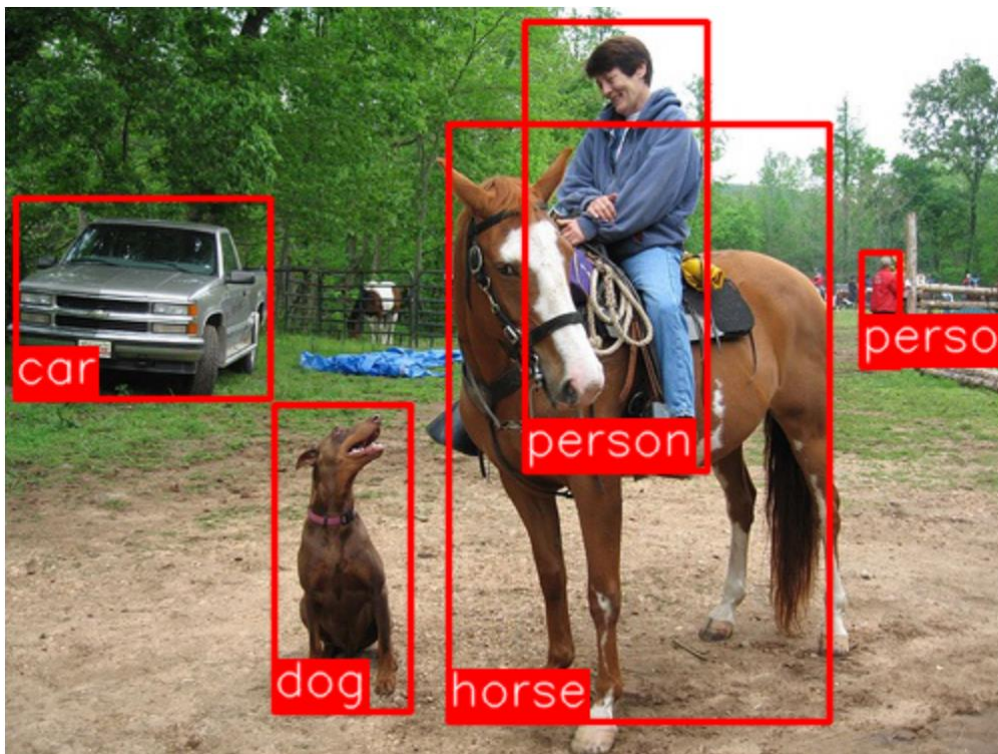
Εικόνα 5.22

Τα μοντέλα Συνελκτικών Νευρωνικών Δικτύων κατηγοριοποίησης ενός και πολλαπλών αντικειμένων. [208]

Ο κυριότερος λόγος για τον οποίο δεν είναι εφικτή η λύση αυτού του προβλήματος με την σχεδίαση ενός προτύπου Συνελκτικού Νευρωνικού Δικτύου, που ακολουθούμενου από ένα στρώμα πλήρους διασύνδεσης, είναι ότι το μήκος του στρώματος εξόδου είναι μεταβλητό και όχι σταθερό, και αυτό οφείλεται στο γεγονός ότι ο αριθμός των εμφανίσεων των αντικειμένων που μας ενδιαφέρει είναι δεν είναι δεδομένος.

Μία απλοϊκή προσέγγιση για την επίλυση αυτού του προβλήματος θα μπορούσε να ήταν η λήψη διαφόρου μεγέθους υπό-περιοχών της εικόνας και η ταξινόμηση τους μέσω του CNN για να χαρακτηριστεί η παρουσία του αντικειμένου μέσα σε αυτή την περιοχή. Το πρόβλημα με αυτήν την προσέγγιση είναι ότι τα αντικείμενα ενδιαφέροντος μπορεί να έχουν διαφορετικές χωρικές θέσεις μέσα στην εικόνα και διαφορετικές αναλογίες διαστάσεων. Ως εκ τούτου, θα πρέπει να επιλεχθεί ένας τεράστιος αριθμός των υπό-περιοχών και αυτό θα ήταν πρακτικά αδύνατο. Ως εκ τούτου, αλγόριθμοι όπως οι R-CNN, YOLO κλπ, έχουν αναπτυχθεί για αυτόν ακριβώς το λόγο. [208]

5.6 R-CNN



Εικόνα 5.23

Αλγόριθμοι ανίχνευσης αντικειμένων όπως ο R-CNN, λαμβάνουν μια εικόνα και ταξινομούν τις κατηγορίες και τις ακριβείς θέσεις των αντικειμένων που περιλαμβάνονται στην εικόνα αυτή.
[207]

Μια πρώιμη εφαρμογή των CNNs για ανίχνευση αντικειμένων

Εμπνευσμένη από το ερευνητικό εργαστήριο Hinton στο Πανεπιστήμιο του Τορόντο, το 2015 μια μικρή ομάδα στο πανεπιστήμιο του Berkeley, με επικεφαλής τον καθηγητή *Jitendra Malik*, έθεσαν στους εαυτούς τους το ακόλουθο ερώτημα:

Μέχρι ποιο βαθμό γενικεύεται η ανίχνευση αντικειμένων;

Ως ανίχνευση αντικειμένων νοείται η διαδικασία του εντοπισμού των διαφόρων αντικειμένων που περιλαμβάνονται σε μία εικόνα και η ταξινόμηση τους. Η ομάδα, αποτελούμενη από τους Ross Girshick, Jeff Donahue και Trevor Darrel, διαπίστωσε ότι αυτό το πρόβλημα μπορεί να λυθεί με τα αποτελέσματα της έρευνας Krizhevsky, κατά τον έλεγχο στο PASCAL VOC Challenge, ένα δημοφιλή διαγωνισμό ανίχνευσης αντικειμένων. Όπως δήλωσαν οι ίδιοι οι ερευνητές,

«Αυτό το σύγγραμμα, αποτελεί το πρώτο που παρουσιάζει ότι ένα CNN μπορεί να οδηγήσει σε δραματικά υψηλότερη απόδοση ανίχνευσης αντικειμένου στο PASCAL VOC σε σύγκριση με τα συστήματα που βασίζονται σε απλούστερη HOG-όπως χαρακτηριστικά γνωρίσματα.

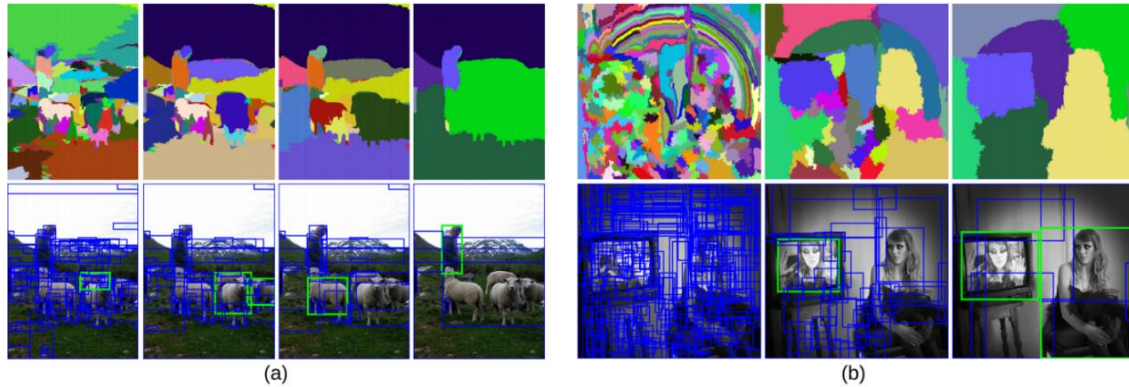
Κάτωθι ακολουθεί μία συνοπτική ανάλυση της αρχιτεκτονικής των R-CNN.

Κατανόώντας τα R-CNN

Ο στόχος του R-CNN είναι να λάβει μια εικόνα, και να προσδιορίσει σωστά, που ακριβώς βρίσκονται τα κύρια αντικείμενα στην εικόνα, μέσω εφαρμογής ενός πλαισίου οριοθέτησης.

- **Είσοδοι:** Εικόνα
- **Έξοδοι:** Πλαίσια οριοθέτησης + ετικέτες για κάθε αντικείμενο στην εικόνα.

Αλλά πώς ακριβώς επιτυγχάνεται η εξεύρεση των ορίων των αντικειμένων; Αυτό επιτυγχάνεται μέσω εφαρμογής ενός προτεινόμενου συνόλου πλαισίων σε λοιπά σημεία της εικόνας, και σε επόμενο βήμα εξετάζει τα πλαίσια αυτά ένα προς ένα ώστε να εντοπίσει πιθανά αντικείμενα εντός τους.

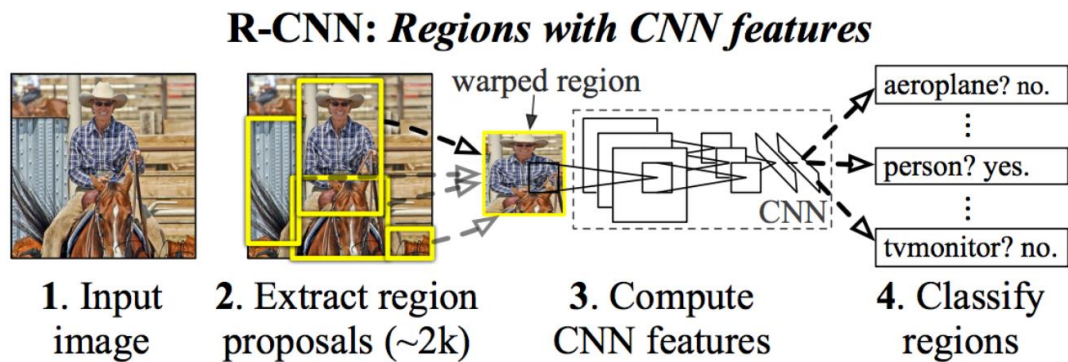


Εικόνα 5.24

Δύο παραδείγματα επιλεκτικής αναζήτησης όπου παρουσιάζουν την ανάγκη επεξεργασίας αντικειμένων διαφόρων μεγεθών. Αριστερά παρουσιάζονται αντικείμενα σε διαφορετικά μεγέθη. Δεξιά παρουσιάζονται αντικείμενα διαφόρων μεγεθών όπως η κοπέλα περιέχεται στην τηλεόραση. Η Επιλεκτική Αναζήτηση ελέγχει μέσα στα πλαίσια διαφόρων μεγεθών, και αναζητά γειτονικά pixels που φέρουν πανομοιότυπες υφές, χρώματα, ή εντάσεις.

[207]

Ο R-CNN δημιουργεί τα πλαίσια οριοθέτησης (ή προτεινόμενου εύρους), χρησιμοποιώντας μια διαδικασία που ονομάζεται **Selective Search** (επιλεκτική αναζήτηση). Σε ένα υψηλότερο επίπεδο, η διαδικασία αυτή (η οποία παρουσιάζεται στην παραπάνω εικόνα) εξετάζει την εικόνα μέσα από τα πλαίσια διαφόρων μεγεθών, και για κάθε μέγεθος πλαισίου, προσπαθεί να ομαδοποιήσει παρακείμενα εικονοστοιχεία (pixels) με γνώμονα την υφή, το χρώμα, ή η ένταση τους, έτσι ώστε να διευκολυνθεί ο εντοπισμός των αντικειμένων.



Εικόνα 5.25

Μετά τη δημιουργία ενός συνόλου πλαισίων οριοθέτησης, ο R-CNN εφαρμόζει την εικόνα σε μια τροποποιημένη έκδοση του AlexNet, ώστε να καθοριστεί εάν η περιοχή περιέχει ανιχνεύσιμο αντικείμενο ή όχι. [207]

Μόλις δημιουργηθούν οι παραπάνω προτάσεις, οι οποίες ανέρχονται σε περίπου 2.000 πλαίσια προτάσεων περιοχών, ο R-CNN συμπιέζει την κάθε περιοχή σε ένα πρότυπο τετραγωνικό μέγεθος και τις τροφοδοτεί με μία τροποποιημένη έκδοση του AlexNet, ο οποίος με τη σειρά του παράγει ένα 4096-διάστατο διάνυσμα χαρακτηριστικών ως έξοδο όπως φαίνεται και στην παραπάνω εικόνα. Στο τελευταίο στρώμα του CNN, ο R-CNN προσθέτει μία Support Vector Machine (SVM) η οποία αποφασίζει αν αυτό αποτελεί κάποιο αντικείμενο ή όχι, και αν ναι, τι αντικείμενο είναι αυτό. Εκτός από την πρόβλεψη της παρουσίας ενός αντικειμένου εντός των προτεινόμενων περιοχών, ο αλγόριθμος προβλέπει επίσης τέσσερις τιμές αντιστάθμισης για την αύξηση της ακρίβειας του πλαισίου οριοθέτησης. Για παράδειγμα, σε μια προτεινόμενη περιοχή, ο αλγόριθμος θα μπορούσε να προβλέψει την παρουσία ενός ανθρώπου, αλλά το πρόσωπο του εν λόγω ανθρώπου θα μπορούσε να βρίσκεται κατά το ήμισυ μόνο εντός της

προτεινόμενης περιοχής. Ως εκ τούτου, οι τιμές αντιστάθμισης θα βοηθούσαν στην αναπροσαρμογή του πλαισίου οριοθέτησης της προτεινόμενης περιοχής. Η διαδικασία αυτή υλοποιείται κατά το τέταρτο στάδιο της παραπάνω εικόνας.

Βελτιώνοντας τα Πλαίσια Οριοθέτησης

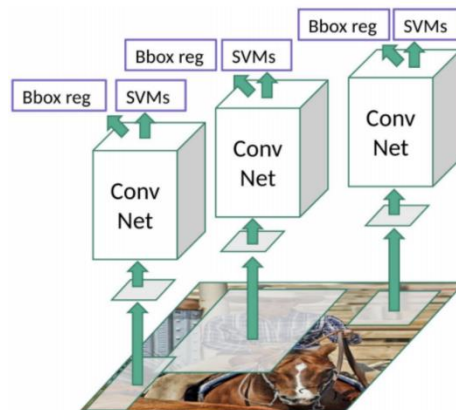
Έχοντας ανιχνεύσει το αντικείμενο εντός του πλαισίου, είναι δυνατή η επαναφορά του πλαισίου στις αρχικές διαστάσεις του (προτού πραγματοποιηθεί η συμπίεση), μέσω μίας αντιστρόφως ανάλογης διαδικασίας, αυτή της συρρίκνωσης. Αυτό επιτυγχάνεται στο τελικό βήμα του R-CNN. Ο R-CNN εκτελεί μια απλή γραμμική παλινδρόμηση στο προτεινόμενο εύρος έτσι ώστε να συρρικνώσει τις διευρυμένες διαστάσεις του πλαισίου οριοθέτησης και να το επαναφέρει στις αρχικές διαστάσεις και συντεταγμένες του. Παρακάτω παρουσιάζονται οι είσοδοι και οι έξοδοι του μοντέλου γραμμικής παλινδρόμησης:

- **Είσοδοι:** Περιοχές της εικόνας που αντιστοιχούν σε αντικείμενα.
- **Έξοδοι:** Πλαίσιο οριοθέτησης νέων διαστάσεων, για το αντικείμενο της περιοχής.

Συνοψίζοντας, τα βήματα του R-CNN είναι:

1. Δημιουργία ενός συνόλου προτάσεων πλαισίων οριοθέτησης.
2. Τροφοδότηση των εικόνων στα πλαίσια οριοθέτησης, με χρήση ενός προ-εκπαιδευμένου AlexNet και τέλος, εφαρμογή μίας SVM για την αναγνώριση του αντικειμένου εντός του πλαισίου.
3. Εφαρμογή ενός μοντέλου γραμμικής παλινδρόμησης στο πλαίσιο οριοθέτησης, έτσι ώστε να επιτευχθεί η εφαρμογή των αρχικών διαστάσεων και συντεταγμένων στο πλαίσιο.

[207], [208], [161], [164]



Εικόνα 5.26

Η λειτουργική δομή του R-CNN

[208]

5.6.1 Προβλήματα του R-CNN

Αν και ο R-CNN λειτουργεί πολύ καλά, αλλά είναι πραγματικά πολύ αργός για τους δύο εξής λόγους:

- Η εκπαίδευση του δικτύου διαρκεί πάρα πολύ ώρα λόγω της ταξινόμησης των 2000 προτεινόμενων περιοχών ανά εικόνα.
- Δεν δύναται να εφαρμοστεί σε πραγματικό χρόνο καθώς απαιτεί περίπου 47 δευτερόλεπτα για κάθε εφαρμοζόμενη εικόνα.
- Ο αλγόριθμος επιλεκτικής αναζήτησης δεν αποτελεί έναν τυποποιημένο αλγόριθμο. Ως εκ τούτου, σε αυτό το στάδιο δεν υλοποιείται εκπαίδευση. Αυτό θα μπορούσε να οδηγήσει στη δημιουργία κακών προτάσεων υποψήφιων περιοχών.
- Πρέπει να εκπαιδεύσει τρία διαφορετικά μοντέλα ξεχωριστά - τον CNN για την εξαγωγή των χαρακτηριστικών της εικόνας, τον ταξινομητή που προβλέπει την κατηγορία και τέλος, το μοντέλο γραμμικής παλινδρόμησης για να συρρικνώσει τα πλαίσια οριοθέτησης. Αυτό έχει σαν αποτέλεσμα τον πολύ υψηλό βαθμό δυσκολίας εκπαίδευσης. [207], [208], [162]

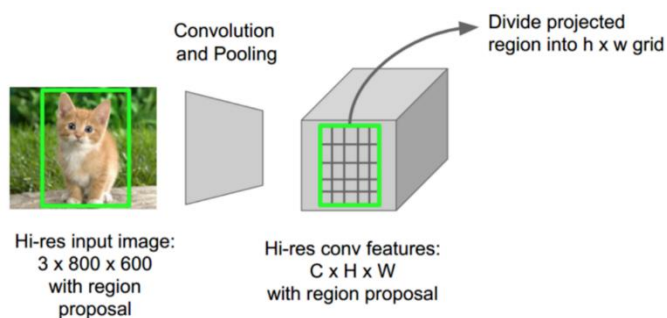
5.7 Fast R-CNN

Για την αντιμετώπιση των προβλημάτων που έφερε ο R-CNN, το 2015 ο Ross Girshick, δημιουργός του R-CNN, υλοποίησε την δεύτερη και βελτιωμένη έκδοση του αλγορίθμου αυτού, τον Fast R-CNN. Η προσέγγιση είναι παρόμοια με αυτήν του αλγορίθμου του R-CNN. Ωστόσο, αντί να τροφοδοτούνται οι προτεινόμενες περιοχές στο CNN, τροφοδοτείται η εικόνα εισόδου στο CNN ώστε να δημιουργηθεί ένας Συνελκτικός χάρτης χαρακτηριστικών. Από τον Συνελκτικό χάρτη χαρακτηριστικών, εξάγονται οι προτεινόμενες περιοχές και εν συνεχεία παραμορφώνονται / στρεβλώνονται σε πλαίσια σταθερού μεγέθους τετραγωνικής μορφής έτσι ώστε να μπορούν να τροφοδοτηθούν σε ένα συνδεδεμένο στρώμα. Η διεργασία αυτή πραγματοποιείται μέσω της τεχνικής ROI Pooling. [207], [208], [162]

5.7.1 ROI Pooling

Ο Girshick συνειδητοποίησε ότι κατά την εμπρόσθια τροφοδότηση του CNN στον R-CNN, για κάθε εικόνα, πολλές από τις προτεινόμενες περιοχές επικαλύπτονται οδηγώντας στην εκτέλεση του ίδιου CNN επαναλαμβανόμενα (περίπου 2000 φορές). Έτσι, εισήγαγε το παρακάτω ερώτημα:

«Γιατί να μην εκτελεστεί το CNN μόνο μία φορά ανά εικόνα και στη συνέχεια να βρει έναν τρόπο για να μοιραστεί αυτόν τον υπολογισμό σε όλες τις ~2000 προτάσεις;»

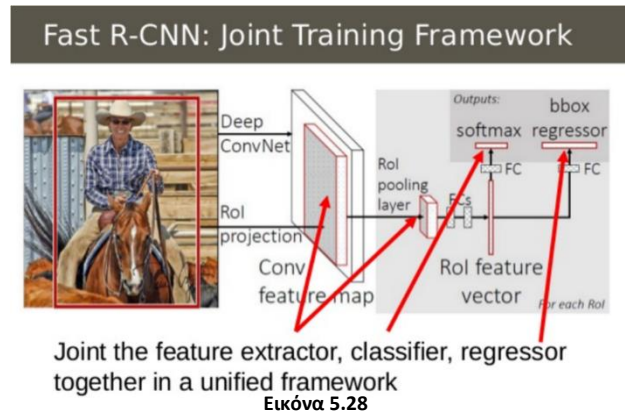


Εικόνα 5.27

Στη ROI Pool, δημιουργείται μια πλήρης εμπρόσθια τροφοδότηση της εικόνας, και τα χαρακτηριστικά κάθε περιοχής ενδιαφέροντος εξάγονται από το αποτέλεσμα της εμπρόσθιας τροφοδότησης. [207]

Αυτή αποτελεί και τη διαφοροποίηση του Fast R-CNN, ο οποίος εφαρμόζει μία τεχνική γνωστή ως ROI Pool (Region of Interest Pooling). Στον πυρήνα του, ο ROI Pool μοιράζεται την εμπρόσθια τροφοδότηση του CNN για μια εικόνα, με όλες τις περιοχές της. Στη παραπάνω εικόνα, φαίνεται πώς τα χαρακτηριστικά του CNN για κάθε περιοχή, λαμβάνονται επιλέγοντας την αντίστοιχη περιοχή από τον χάρτη χαρακτηριστικών του CNN. Στη συνέχεια, τα χαρακτηριστικά σε κάθε περιοχή ομαδοποιούνται (συνήθως εφαρμόζοντας max pooling). Έτσι, το μόνο που απαιτείται, είναι ένα πέρασμα από την αρχική εικόνα, σε αντίθεση με τις 2000 τροφοδοτήσεις του R-CNN αντίστοιχα. [207], [208], [162]

5.7.2 Συνδυασμός όλων των μοντέλων σε ένα ενιαίο δίκτυο

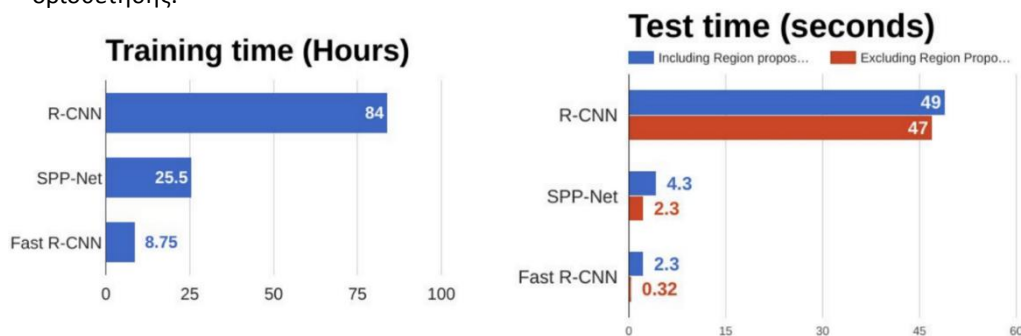


Ο Fast R-CNN συνδυάζεται με το CNN, τον ταξινομητή, και το πλαίσιο οριοθέτησης παλινδρόμησης, σε ένα, ενιαίο δίκτυο. [207]

Η δεύτερη λειτουργία του Fast R-CNN είναι να εκπαιδεύσει από κοινού τους CNN, ταξινομητή, και πλαίσιο οριοθέτησης παλινδρόμησης εντός ενός ενιαίου μοντέλου. Έτσι, από εκεί που προηγουμένως αξιοποιούνταν διαφορετικά μοντέλα για την εξαγωγή χαρακτηριστικών της εικόνας (CNN), την ταξινόμηση (SVM) και την συρρίκνωση του πλαισίου οριοθέτησης παλινδρόμησης, ο Fast R-CNN χρησιμοποιεί ένα ενιαίο πλαίσιο για τον υπολογισμό τους αντί ένα ενιαίο δίκτυο για τον υπολογισμό τους αντί για τρία.

Στην παραπάνω εικόνα διακρίνεται αυτός ο τρόπος λειτουργίας. Ο Fast R-CNN αντικατέστησε τον ταξινομητή SVM με ένα στρώμα Softmax στην κορυφή του CNN ώστε να παράγει ένα αποτέλεσμα ταξινόμησης. Ακόμα, πρόσθεσε ένα στρώμα γραμμικής παλινδρόμησης, παράλληλο προς το στρώμα Softmax, ώστε να παράγει τις συντεταγμένες του πλαισίου οριοθέτησης. Έτσι, όλες οι απαιτούμενες έξοδοι, προήλθαν από ένα ενιαίο δίκτυο. Οι εισοδοί και οι έξοδοι αυτού του ενιαίου μοντέλου, είναι:

- **Είσοδοι:** Εικόνες με προτάσεις περιοχών.
- **Έξοδοι:** Ταξινομήσεις αντικειμένων κάθε περιοχής, σε συνδυασμό με τον ορισμό νέων πλαισίων οριοθέτησης.



Εικόνα 5.29

Σύγκριση των ταχύτητας και χρόνου δοκιμής των αλγορίθμων ανίχνευσης αντικειμένων. [208]

Από τα παραπάνω γραφήματα, προκύπτει ότι ο Fast R-CNN είναι σημαντικά ταχύτερος από τον R-CNN, τόσο κατά την διαδικασία της εκπαίδευσής του, όσο και κατά την διαδικασία δοκιμών. Ο λόγος που ο Fast R-CNN είναι ταχύτερος του R-CNN, είναι επειδή χρειάζεται να τροφοδοτήσει 2000 προτεινόμενες προτάσεις περιοχή στο Συνελικτικό Νευρωνικό Δίκτυο κάθε φορά. Αντ' αυτού, η λειτουργία συνέλιξης γίνεται μόνο μία φορά ανά εικόνα και ένα χαρακτηριστικό χάρτη παράγεται από αυτό. [207], [208], [162]

5.8 Faster R-CNN

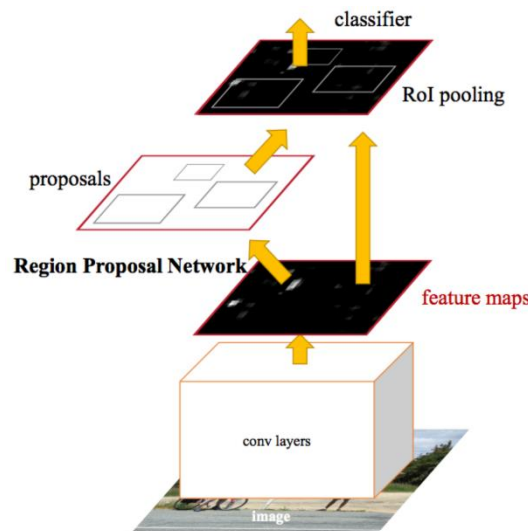
Επιτάχυνση Πρότασης Περιοχών

Παρά το πλήθος των εξελίξεων στα R-CNN, παρέμενε ακόμα ένα σημαντικό εμπόδιο στον αλγόριθμο του Fast R-CNN, η διαδικασία της πρότασης περιοχών. Όπως ειπώθηκε προηγουμένως, το πρώτο στάδιο του εντοπισμού των ορίων ενός αντικειμένου, είναι η εξαγωγή ενός πλήθους πλαισίων πιθανών ορίων αντικειμένων. Στον αλγόριθμο Fast R-CNN, οι προτάσεις αυτές δημιουργήθηκαν χρησιμοποιώντας την Επιλεκτική Αναζήτηση, μια αρκετά αργή διαδικασία, που ουσιαστικά αποτελούσε και την βασικό μειονέκτημα του αλγορίθμου.

Στα μέσα του 2015, μια ομάδα στον τομέα έρευνας της Microsoft, αποτελούμενη από τους Shaoqing Ren, Kaiming He, Ross Girshick, και Jian Sun, ανακάλυψε μία μέθοδο για να κάνει τη διαδικασία πρότασης υπό-περιοχής πολύ απλούστερη και ταχύτερη, χάρις σε μία νέα αρχιτεκτονική που ονομάστηκε Faster R-CNN.

Η βάση της θεωρίας του Faster R-CNN ήταν ότι οι προτάσεις υπό-περιοχών εξαρτώνται από τα χαρακτηριστικά της εικόνας που είχαν ήδη υπολογιστεί με την χρήση μίας εμπρόσθιας τροφοδότησης του CNN (πρώτο στάδιο της ταξινόμησης).

Γιατί λοιπόν να μην επαναχρησιμοποιηθούν ξανά τα ίδια αποτελέσματα του CNN για τις προτάσεις υπό-περιοχών, αντί να εκτελείται ένας ξεχωριστός αλγόριθμος επιλεκτικής αναζήτησης;



Εικόνα 5.30

Στον Faster R-CNN, για τις προτάσεις περιοχών και για την ταξινόμηση, χρησιμοποιείται ένα και μόνο CNN.

[207]

Στην παραπάνω εικόνα, παρουσιάζεται το πώς ακριβώς επιτυγχάνει ένα και μόνο CNN να χρησιμοποιείται τόσο για τη διεξαγωγή προτάσεων περιοχών όσο και για την λειτουργία ταξινόμησης. Με αυτό τον τρόπο, χρειάζεται να εκπαιδευτεί μόνο ένα CNN, και οι προτάσεις υπό-περιοχών πραγματοποιούνται πολύ ταχύτερα.

Οι δημιουργοί του Faster R-CNN, δήλωσαν:

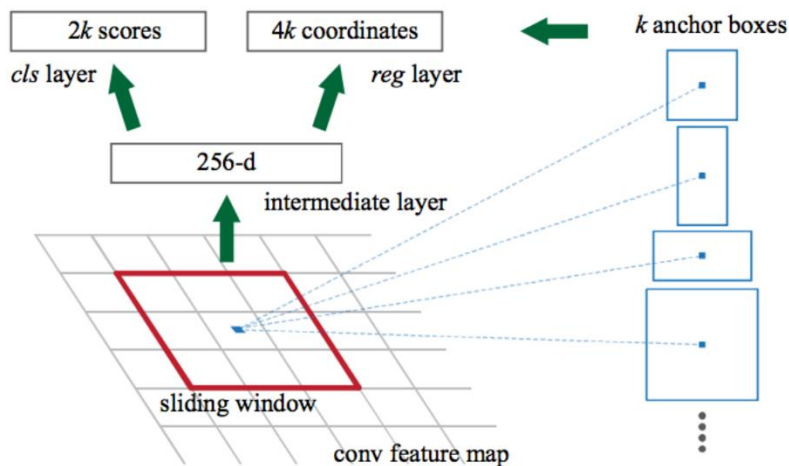
Η παρατήρησή μας είναι ότι οι Συνελικτικοί χάρτες χαρακτηριστικών που χρησιμοποιούνται από τους ανιχνευτές υπό-περιοχής, όπως ο Fast R-CNN, δύναται επίσης να χρησιμοποιηθούν για τη δημιουργία προτάσεων υπό-περιοχών.

Οι είσοδοι και έξοδοι του συγκεκριμένου μοντέλου, είναι:

- **Είσοδοι:** Εικόνες (δεν χρειάζονται προτάσεις υπό-περιοχών).
- **Έξοδοι:** Ταξινομήσεις και συντεταγμένες πλαισίου οριοθέτησης των αντικειμένων στις εικόνες.

Πως δημιουργούνται οι υπό-περιοχές;

Ας ρίξουμε μια ματιά για να δούμε πώς ο Faster R-CNN δημιουργεί αυτές τις προτάσεις υπό-περιοχών από τα χαρακτηριστικά του CNN. Ο Faster R-CNN προσθέτει ένα πλήρως Συνελικτικό Δίκτυο πάνω από τα χαρακτηριστικά του CNN, δημιουργώντας αυτό που είναι γνωστό ως **Region Proposal Network**.



Εικόνα 5.31

Το *Region Proposal Network* μετακινεί ένα πλαίσιο πάνω από τα χαρακτηριστικά του CNN. Σε κάθε θέση του πλαισίου, το δίκτυο εξάγει ένα σκορ και ένα πλαίσιο οριοθέτησης ανά anchor.

[207]

Το **Region Proposal Network** λειτουργεί μεταφέροντας ένα συρόμενο πλαίσιο πάνω από το χάρτη χαρακτηριστικών του CNN και σε κάθε πλαίσιο, εξάγοντας k πιθανά πλαίσια οριοθέτησης και αξιολογεί το πόσο καλό αναμένεται να είναι κάθε ένα από αυτά τα πλαίσια. Τι είναι αυτά τα k πλαίσια και τι αντιπροσωπεύουν;

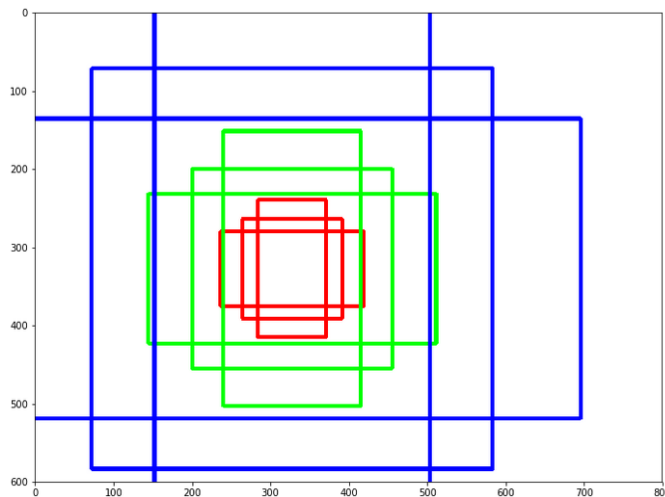


Εικόνα 5.32

Γνωρίζουμε ότι τα πλαίσια οριοθέτησης για τους ανθρώπους συνήθως είναι κάθετα ορθογώνια. Μπορούμε να χρησιμοποιήσουμε αυτή την υπόθεση ώστε να καθοδηγήσουμε τα Region Proposal networks, για τη δημιουργία anchor αντίστοιχων διαστάσεων.

[207]

Διαισθητικά, γνωρίζουμε ότι τα αντικείμενα μίας εικόνας θα πρέπει να ταιριάζουν σε συγκεκριμένους λόγους αναλογιών και διαστάσεων. Για παράδειγμα, γνωρίζουμε ότι θέλουμε κάποια ορθογώνια κουτιά που ταιριάζουν με τα σχήματα των ανθρώπων. Επίσης, ξέρουμε ότι δεν θα δούμε πολλά πλαίσια που θα πολύ μικρού πλάτους. Με τον τρόπο αυτό, δημιουργούμε k τέτοιους κοινούς λόγους αναλογιών τους οποίους καλούμε **anchor boxes**. Ουσιαστικά τα **anchor boxes** αποτελούν κάποια συγκεκριμένα -9 σε προεπιλεγμένη κατάσταση – πλαίσια οριοθέτησης που εφαρμόζονται στην εικόνα. Έτσι, για κάθε **anchor box**, εξάγουμε ένα πλαίσιο οριοθέτησης όπου «σκοράρει» ανά θέση στην εικόνα.



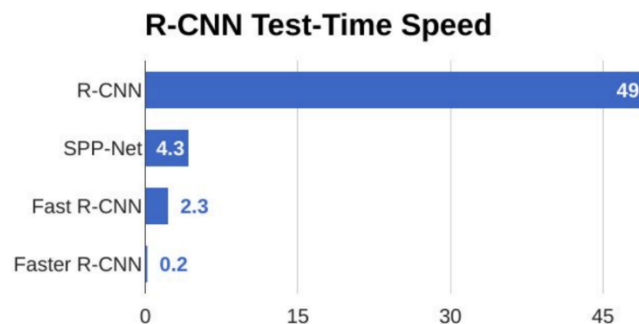
Εικόνα 5.33

Τα 9 anchor boxes που εμφανίζονται στις προεπιλεγμένες θέσεις (320, 320), μίας εικόνας διαστάσεων (600, 800).

Με τα **anchor boxes**, κατά νου, ας ρίξουμε μια ματιά στις εισόδους και εξόδους του **Region Proposal Network**:

- **Είσοδοι:** Χαρακτηριστικά χάρτη CNN.
- **Έξοδοι:** Ένα πλαίσιο οριοθέτησης ανά anchor. Ένα σκορ που αντιπροσωπεύει το πόσο πιθανό είναι η εικόνα σε αυτό το πλαίσιο οριοθέτησης, να είναι ένα αντικείμενο.

Στη συνέχεια, κάθε τέτοιο πλαίσιο οριοθέτησης, που είναι πιθανό να είναι ένα αντικείμενο, περνάει στο Fast R-CNN έτσι ώστε να δημιουργηθεί μια ταξινόμηση καθώς και τα συρρικνωμένα πλαίσια οριοθέτησης.



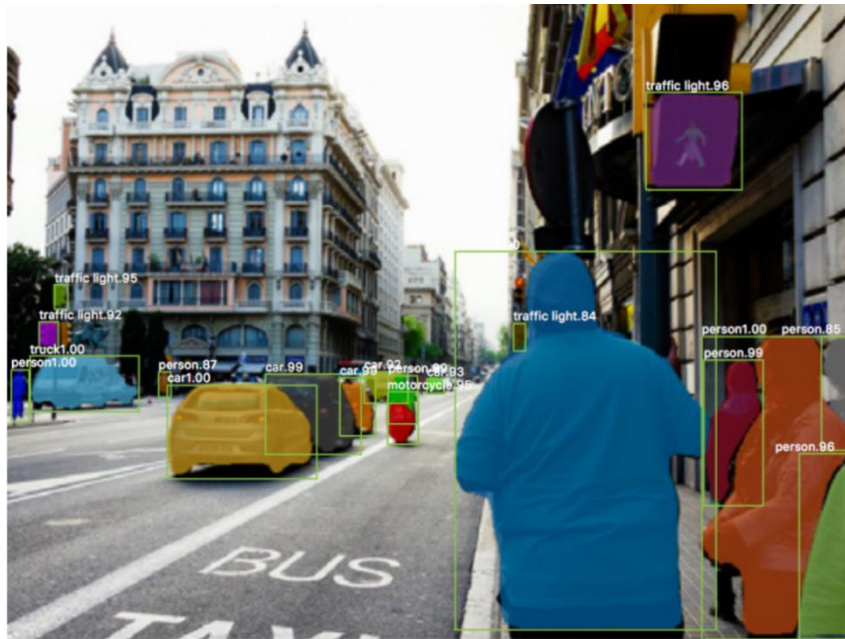
Εικόνα 5.34

Σύγκριση Αλγορίθμων Εντοπισμού Αντικειμένων

[208]

Από την παραπάνω εικόνα, προκύπτει το συμπέρασμα ότι ο αλγόριθμος Faster R-CNN είναι ο ταχύτερος και αποδοτικότερος αλγόριθμος εντοπισμού αντικειμένων. [207], [208], [163], [175]

5.9 Mask R-CNN



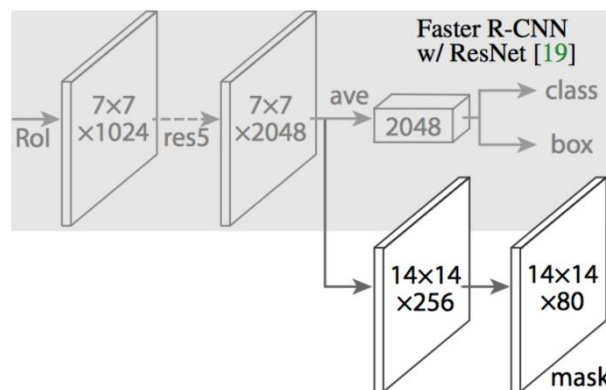
Εικόνα 5.35

Ο σκοπός της τεχνικής κατάτμησης της εικόνας, είναι να εντοπίσει σε επίπεδο pixel, ποιά είναι τα διαφορετικά αντικείμενα μίας εικόνας.

[207]

Με βάση τις τεχνικές που αναλύθηκαν προηγουμένως, είναι δυνατή η λειτουργική επέκταση τους έτσι ώστε όχι μόνο να αναγνωρίζονται τα αντικείμενα βάσει του πλαισίου οριοθέτησης τους, αλλά να εντοπίζονται με ακρίβεια και το σύνολο των pixel που απαρτίζουν το κάθε αντικείμενο. Το 2017, μία ομάδα ερευνητών του τομέα Τεχνητής Νοημοσύνης του Facebook, δημιούργησαν έναν αλγόριθμο που υλοποιεί αυτή τη λειτουργία, ο οποίος ονομάζεται **Mask R-CNN**.

Ομοιοτρόπως με τους Fast R-CNN και Faster R-CNN, ο Mask R-CNN βασική διαίσθηση είναι κατευθείαν προς τα εμπρός. Δεδομένου ότι ο Faster R-CNN επιτελεί ικανοποιητικά ανιχνεύσεις αντικειμένων, είναι δυνατή και η επέκταση της λειτουργίας του ώστε να επιτελεί την τμηματοποίηση των αντικειμένων σε επίπεδο εικονοστοιχείου (pixel).



Εικόνα 5.36

Στον Mask R-CNN, ένα Πλήρως Συνέλικτικό Δίκτυο (FCN) προστίθεται πάνω από τα χαρακτηριστικά του CNN του Faster R-CNN έτσι ώστε να δημιουργηθεί μια μάσκα (έξοδος κατάτμησης). Αυτή η μάσκα βρίσκεται παράλληλα του ταξινομητή και του πλαισίου οριοθέτησης του Faster R-CNN.

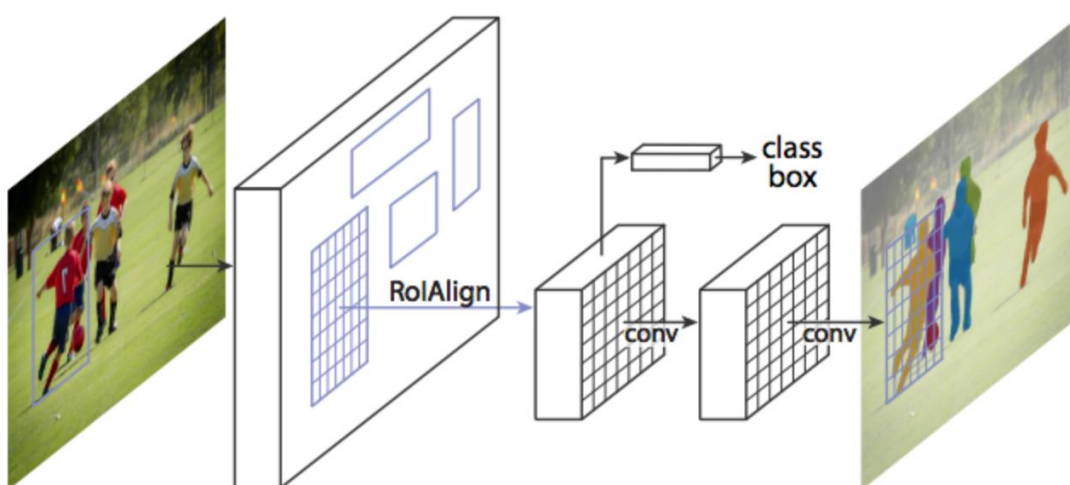
[207]

Ο Mask R-CNN το επιτυγχάνει αυτό, με τη προσθήκη μίας διακλάδωσης στον Faster R-CNN όπου εξάγει μία δυαδική μάσκα που υποδεικνύει το αν ένα συγκεκριμένο ρίxel είναι μέρος ενός αντικειμένου ή όχι. Η διακλάδωση, ομοίως με πριν, είναι ουσιαστικά ένα πλήρως Συνελικτικό Δίκτυο, πάνω σε ένα χάρτη χαρακτηριστικών CNN. Οι είσοδοι και έξοδοι αντίστοιχα, είναι:

- **Είσοδοι:** Χάρτης χαρακτηριστικών CNN.
- **Έξοδοι:** Πίνακας που φέρει το «1» σε όλες τις περιοχές όπου το ρίxel ανήκει στο αντικείμενο και «0» σε όλες τις υπόλοιπες περιοχές. (γνωστό και ως δυαδική μάσκα).

Ωστόσο, το Mask R-CNN, έπρεπε να τροποποιηθεί σε ένα μικρό βαθμό έτσι ώστε να λειτουργήσει. [207]

5.9.1 RoiAlign

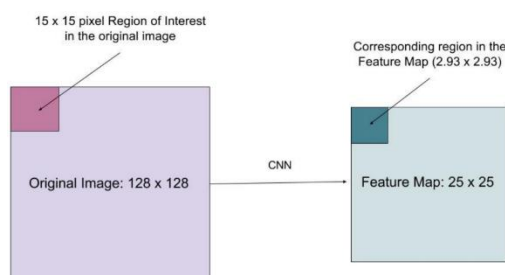


Εικόνα 5.37

Σε αντίθεση με την RoiPool, η εικόνα περνάει από το RoiAlign έτσι ώστε οι περιοχές του χάρτη χαρακτηριστικών που επιλέγει η RoiPool να αντιστοιχούν με μεγαλύτερη ακρίβεια στις περιοχές της αρχικής εικόνας. Αυτό είναι απαραίτητο, διότι η ακρίβεια επιπέδου pixel, απαιτεί πιο λεπτομερή ευθυγράμμιση συγκριτικά με τα πλαίσια οριοθέτησης. [207]

Κατά την εκτέλεση του Mask R-CNN πάνω στην πρωτότυπη και χωρίς τροποποιήσεις αρχιτεκτονική του Faster R-CNN, οι ερευνητές εντόπισαν πως που επιλέγονται στον χάρτη χαρακτηριστικών από την RoiPool, ήταν ελαφρά μη-ευθυγραμμισμένες ως προς τις περιοχές της πρωτότυπης εικόνας. Εφόσον η διαδικασία τμηματοποίησης της εικόνας απαιτεί ακρίβεια επιπέδου ρίxel, σε αντίθεση με τα πλαίσια οριοθέτησης, αυτό φυσικά οδήγησε σε σημαντικές αποκλίσεις.

Η λύση του προβλήματος ήρθε με την προσαρμογή και εφαρμογή της μεθόδου RoiPool έτσι ώστε να πετύχουν ευθυγράμμιση μεγαλύτερης ακρίβειας. Η τεχνική αυτή ονομάζεται RoiAlign.



Εικόνα 5.38

Η διαδικασία διασύνδεσης της αρχικής εικόνας και του χάρτη χαρακτηριστικών, στον Mask R-CNN, είναι σε θέση να τμήματος καθώς και ταξινομήσουν τα αντικείμενα σε μια εικόνα. [207]

Πώς καθίσταται δυνατή η χαρτογράφηση μίας περιοχής ενδιαφέροντος από την αρχική εικόνα, στον χάρτη χαρακτηριστικών;

Ας υποθέσουμε ότι υπάρχει μια εικόνα διαστάσεων 128x128 και ένας χάρτης χαρακτηριστικών διαστάσεων 25x25. Υποθέτουμε επίσης ότι χρειαζόμαστε τα αντίστοιχα χαρακτηριστικά της περιοχής διαστάσεων 15x15 pixels που βρίσκεται επάνω και αριστερά της αρχικής εικόνας. Πως θα ήταν δυνατός ο εντοπισμός αυτών των pixel στον χάρτη χαρακτηριστικών;

Γνωρίζουμε ότι κάθε pixel στην αρχική εικόνα αντιστοιχεί σε $\sim 25/128$ pixels στον χάρτη χαρακτηριστικών. Για την επιλογή 15 pixel από την αρχική εικόνα, απλώς επιλέγονται $15 * 25/128 \sim 2.93$ pixels. Στη RoIPool, με εφαρμογή στρογγυλοποίησης προς τα κάτω, θα επιλέγαμε 2 pixel προκαλώντας μια μικρή απόκλιση. Ωστόσο, στον RoIAlign, αποφεύγονται στρατηγικά τέτοιου είδους στρογγυλοποιήσεις. Αντ' αυτού, εφαρμόζεται διηλεκτρική παρεμβολή ώστε να πλαισιώναμε την έννοια των 2.93 pixel. Αυτό, σε ένα υψηλό επίπεδο, θα επέτρεπε την αποφυγή αποκλίσεων που προκλήθηκαν από την RoIPool.

Μετά την δημιουργία αυτών των масκών, ο Mask R-CNN τις συνδυάζει με τις ταξινομήσεις και τα πλαίσια οριοθέτησης του Faster R-CNN, ώστε να δημιουργηθούν οι ακριβείς οριοθετήσεις ακριβείας επιπέδου pixel. [207]



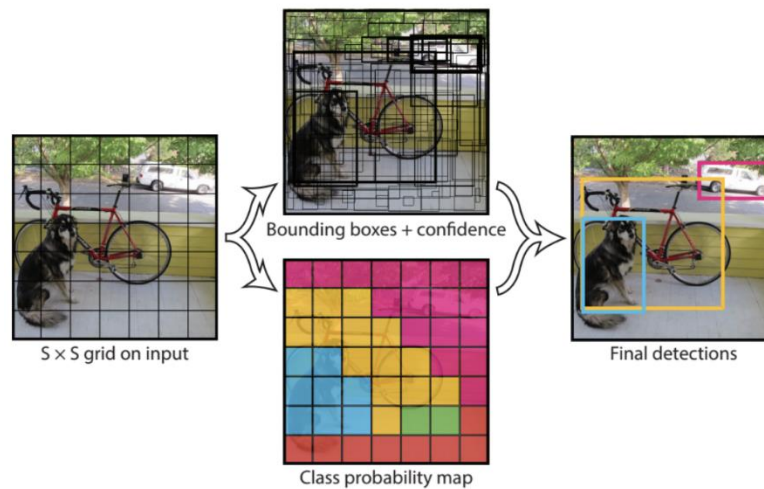
Εικόνα 5.39

Η διαδικασία εντοπισμού αντικειμένων με ακρίβεια pixel, από τον αλγόριθμο Mask R-CNN.

[207]

5.10 YOLO

Σε όλους τους παραπάνω αλγόριθμους ανίχνευσης αντικειμένων, εφαρμόστηκε ο εντοπισμός περιοχών για τον προσδιορισμό των αντικειμένων εντός μίας εικόνας. Έτσι, το Δίκτυο δεν εξέτασε ποτέ την συνολική εικόνα, αλλά μόνο περιοχές της εικόνας οι οποίες συγκέντρωναν υψηλές πιθανότητες να περιέχουν κάποιο αντικείμενο. Ο YOLO (You Only Look Once) αποτελεί έναν αλγόριθμο εντοπισμού αντικειμένων πολύ διαφορετικότερης φιλοσοφίας από τους αλγόριθμους που εξετάστηκαν προηγουμένως. Στον αλγόριθμο YOLO, ένα μονό Συνελκτικό Δίκτυο δημιουργεί τα πλαίσια οριοθέτησης και τις πιθανότητες ταξινόμησης αντικειμένου για τα πλαίσια αυτά.



Εικόνα 5.40

Εντοπισμός και ταξινόμηση εικόνων με χρήση της μεθόδου YOLO.

[208]

Ο YOLO επιτυγχάνει τη λήψη μιας εικόνας και την εφαρμογή ενός πλέγματος 5x5 συνολικών τετραγώνων, και εντός του κάθε τετραγώνου εφαρμόζονται η πλαίσια οριοθέτησης. Για κάθε πλαίσιο οριοθέτησης, το Δίκτυο εξάγει μία πιθανότητα ύπαρξης αντικειμένου καθώς και τις αντίστοιχες τιμές αντιστάθμισης για το πλαίσιο οριοθέτησης.

Ο YOLO είναι κλάσεις ταχύτερος (ταχύτητα επιπέδου 45 καρτέ ανά δευτερόλεπτο) από άλλους του υπόλοιπους αλγόριθμους εντοπισμού αντικειμένων. Ο περιορισμός του εν λόγω αλγόριθμου, είναι πως δεν είναι ιδιαίτερα αποδοτικός στον εντοπισμό μικρών αντικειμένων. Για παράδειγμα, ίσως αντιμετωπίσει δυσκολία στην ανίχνευση ενός σμήνους πτηνών σε μία εικόνα. Αυτό συμβαίνει λόγω των χωρικών περιορισμών που υπάρχουν στον αλγόριθμο. [174], [208]

5.11 Άλλες Τεχνικές Μηχανικής Όρασης

Τα μόλις πέντε τελευταία χρόνια, έχουν σημειωθεί ραγδαίες εξελίξεις στους αλγόριθμους εντοπισμού αντικειμένων. Από τις συμβατικές υλοποιήσεις των CNN, περάσαμε στις σύνθετες και ιδιαίτερα εξελιγμένες υλοποιήσεις αλγορίθμων εντοπισμού αντικειμένων με πολύ υψηλά ποσοστά προβλέψεων και ακρίβεια εντοπισμού σε επίπεδο εικονοστοιχείου. Η τεχνολογία εξέλιξης των συγκεκριμένων αλγορίθμων κινείται με ταχύτατους ρυθμούς, δημιουργώντας ένα πλήθος νέων υλοποιήσεων σε ετήσια βάση.

Αν και οι σημαντικότεροι αλγόριθμοι εντοπισμού αναλύθηκαν εντός της παρούσας ενότητας, υπάρχουν πολλοί ακόμα αλγόριθμοι, ιδιαίτερα δημοφιλής ανά τύπο εφαρμογής που έχουν επιδείξει πολύ υψηλά ποσοστά ακρίβειας και επιτυχίας. Έτσι, στην παγκόσμια επιστημονική βιβλιογραφία μπορούν να βρεθούν αναφορές εφαρμογών πολλών άλλων αλγορίθμων εντοπισμού αντικειμένων όπως οι RFCN, SPP NET, YOLOv2, YOLOv3, Deeplab Xception, SSD (Single Shot Detector), RetinaNet κα.

[207], [208]

6 Εφαρμογή Τεχνικών Βαθιάς Μάθησης στα Αυτόνομα Ρομποτικά Οχήματα

Ποια τροποποιήσιμα στοιχεία ενός συστήματος μάθησης είναι υπεύθυνα για την επιτυχία ή την αποτυχία του; Ποιες αλλαγές τους βελτιώνουν την απόδοση; Αυτό το πρόβλημα ονομάστηκε το βασικό πρόβλημα εκχώρησης πίστωσης σφάλματος των βαρών Νευρωνικού Δικτύου (credit assignment) (Minsky, 1963). Υπάρχουν γενικές μέθοδοι εκχώρησης πίστωσης σφάλματος βαρών για τους καθολικούς επιλυτές προβλημάτων που είναι βέλτιστες στο χρόνο από διαφορετικές θεωρητικές απόψεις. Η παρούσα επισκόπηση όμως, θα επικεντρωθεί στο στενότερο, αλλά τώρα εμπορικά σημαντικότερο, υποπεδίο της Deep Learning (DL) δηλ. στα Τεχνητά Νευρωνικά Δίκτυα (NN). Ένα τυποποιημένο Νευρωνικό Δίκτυο (NN) αποτελείται από πολλούς απλούς συνδεδεμένους επεξεργαστές που ονομάζονται νευρώνες, όπου ο καθένας παράγει μια ακολουθία πραγματικών τιμών ενεργοποιήσεων. Οι νευρώνες εισόδου ενεργοποιούνται μέσω αισθητήρων που αντιλαμβάνονται το περιβάλλον, οι άλλοι νευρώνες ενεργοποιούνται μέσω σταθμισμένων συνδέσεων από προηγούμενως ενεργούς νευρώνες. Μερικοί νευρώνες μπορεί να επηρεάσουν το περιβάλλον ενεργοποιώντας δράσεις. Η εκμάθηση ή η ανάθεση πίστωσης αφορά την εύρεση βαρών που κάνουν το NN να επιδεικνύει την επιθυμητή συμπεριφορά, όπως η οδήγηση αυτοκινήτου. Ανάλογα με το πρόβλημα και τον τρόπο σύνδεσης των νευρώνων, μια τέτοια συμπεριφορά μπορεί να απαιτήσει μακρές αιτιακές αλυσίδες υπολογιστικών φάσεων, όπου κάθε στάδιο μετασχηματίζει (συχνά με μη γραμμικό τρόπο) τη συνολική ενεργοποίηση του δικτύου. Η Βαθιά Μάθηση είναι ακριβής αποδίδοντας πίστωση σε πολλά τέτοια στάδια.

Τα ρηχά (shallow) μοντέλα που μοιάζουν με NN με λίγα επίπεδα υπάρχουν στην βιβλιογραφία πολλές δεκαετίες, αν όχι αιώνες, και είναι περισσότερο γραμμικά συστήματα (π.χ πρώιμες NN αρχιτεκτονικές (McCulloch and Pitts, 1943)). Τα μοντέλα με διάφορα διαδοχικά μη γραμμικά στρώματα νευρώνων χρονολογούνται τουλάχιστον μέχρι τη δεκαετία του 1960 και 1970. Μια αποτελεσματική μέθοδος βασισμένη στην πτώση κλίσης της συνάρτησης σφάλματος (gradient descent/η και βέλτιστη πτώση κλίσης (steepest descent)) κατάλληλη για δίκτυα NN που εκπαιδεύονται με επίβλεψη, είναι η λεγόμενη εκπαίδευση με οπισθοδρομική διάδοση σφάλματος (backpropagation error). Αυτή η μέθοδος εκμάθησης με επίβλεψη (Supervised Learning - SL) σε διακριτά, με διαφορίσιμες συναρτήσεις ενεργοποίησης, δίκτυα αυθαίρετου βάθους που ονομάζεται αλγόριθμος οπισθοδρομικής διάδοσης σφάλματος (backpropagation error (BP)), που αναπτύχθηκε στη δεκαετίες 1960 και 1970, εφαρμόστηκε στα NN το 1981.

Αυτός ο τύπος εκπαίδευσης Βαθιάς Μάθησης (DL) σε NNs με πολλά στρώματα, ωστόσο, βρέθηκε να είναι ιδιαίτερα δύσκολο να εφαρμοστεί στην πράξη ήδη από τα τέλη της δεκαετίας του 1980, και είχε καταστεί ρητά ερευνητικό θέμα στις αρχές της δεκαετίας του 1990. Το DL έγινε πρακτικά εφικτό σε κάποιο βαθμό μέσω της βοήθειας της μη εποπτευόμενης μάθησης (Unsupervised Learning (UL)), (1991, 2006). Οι δεκαετίες του 1990 και του 2000 σημείωσαν επίσης πολλές βελτιώσεις στο καθαρά εποπτευόμενο DL. Στη νέα χιλιετία, τα Βαθιά NN (Deep NN) έχουν προσελκύσει τελικά την ευρεία προσοχή, κυρίως με την καλύτερη απόδοση σε μηχανικές μεθόδους μάθησης όπως μηχανές πυρήνα Support Vector Machines (Vapnik, 1995, Scholkopf et al., 1998) σε πολλές σημαντικές εφαρμογές. Στην πραγματικότητα, από το 2009, τα εποπτευόμενα Βαθιά NNs έχουν κερδίσει πολλούς επίσημους διεθνείς διαγωνισμούς αναγνώρισης προτύπων, επιτυγχάνοντας καλύτερα σχετικά αποτελέσματα ως προς τον άνθρωπο σε αναγνώριση οπτικών προτύπων σε περιορισμένους τομείς. Deep NNs μέθοδοι επίσης έχουν γίνει σημαντικές για το γενικότερο πεδίο της Ενίσχυσης Μάθησης (RL) όπου δεν υπάρχει επόπτης.

Τόσο τα προσωτροφοδοτούμενα (ακυκλικά) πολυστρωματικά NN (Multilayer Perceptrons - MLPs) όσο και τα επαναλαμβανόμενα (κυκλικά) NNs (RNNs) έχουν κερδίσει πολλούς διαγωνισμούς. Κατά μία έννοια, τα RNNs είναι τα πιο Βαθιά από όλα τα NN - αυτά είναι γενικοί υπολογιστές πιο ισχυροί από τους MLP και μπορούν καταρχήν να δημιουργήσουν και να επεξεργαστούν μνήμες των αυθαίρετων ακολουθιών των προτύπων εισαγωγής (π.χ., Siegelmann and Sontag, 1991, Schmidhuber, 1990).

Σε αντίθεση με τις παραδοσιακές μεθόδους για αυτόματη διαδοχική σύνθεση προγράμματος (π.χ., Waldinger and Lee, 1969; Balzer, 1985; Soloway, 1986; Deville and Lau, 1994), τα RNN μπορούν να

μάθουν προγράμματα που αναμειγνύουν διαδοχική και παράλληλη επεξεργασία πληροφοριών με φυσικό και αποτελεσματικό τρόπο, αξιοποιώντας τον μαζικό παραλληλισμό και θεωρείται μεθοδολογία κρίσιμη για τη διατήρηση της ταχείας μείωσης του κόστους υπολογισμού που παρατηρήθηκε τα τελευταία 80 χρόνια. [88], [89], [90], [91], [92], [93], [94], [95], [96], [97], [98], [99], [100], [101]

6.1 Ο Αλγόριθμος Μάθησης BP για εκπαίδευση MLPs και RNNs

Χρησιμοποιώντας τον αλγόριθμο που φαίνεται στο παρακάτω σχήμα, είναι εφικτή η εκπαίδευση προσωποδοτούμενων πολυστρωματικών Νευρωνικών Δικτύων (MLP) είτε Νευρωνικών Δικτύων με επαναλαμβανόμενη αρχιτεκτονική (αναδρομικά η επαναλαμβανόμενα (Recurrent) RNN) διαμοιρασμού βαρών, μετά από το βήμα προσωποδοδότησης (εξάπλωσης της ενεργοποίησης των νευρώνων με κατεύθυνση από την είσοδο προς την έξοδο του δικτύου). Ο αλγόριθμος αυτός μπορεί να χρησιμοποιηθεί στην περίπτωση που οι συναρτήσεις ενεργοποίησης είναι διαφορίσιμες. Στο κατωτέρω σχήμα βλέπουμε μία μόνο επανάληψη της διαδικασίας υπολογισμού των μεταβολών των βαρών μέσω του αλγόριθμου BP πτώσης κλίσης της συνάρτησης κόστους, στην περίπτωση της offline (batch) εκδοχής του αλγόριθμου οπισθοδρομικής διάδοσης σφάλματος BP. Οι τροφοδοτήσεις των νευρώνων προς τα εμπρός και προς τα πίσω επαναλαμβάνονται μέχρις ότου επιτευχθούν επαρκείς επιδόσεις στην βελτιστοποίηση της συνάρτησης κόστους.

Αλγόριθμος 6.1: Μια επανάληψη της διαδικασίας BP για εκπαίδευση με επίβλεψη MLPs ή RNNs αρχιτεκτονικών διαμοιρασμού βαρών.

για $t = T, \dots, 1$ do (for loop)

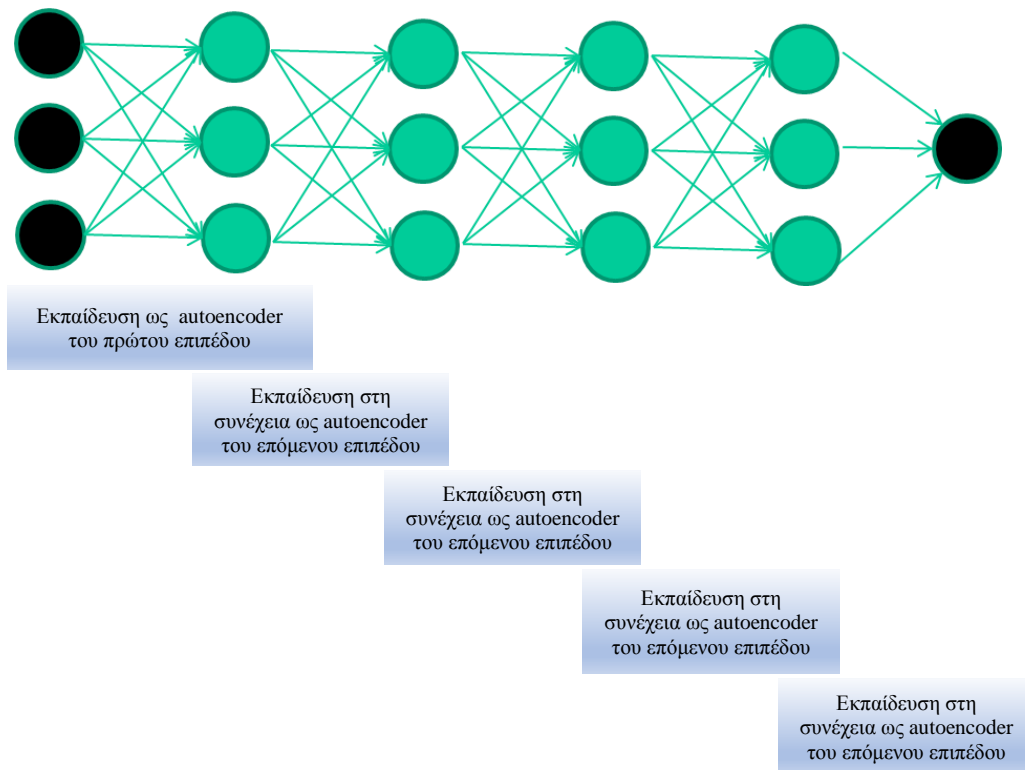
- Για τον υπολογισμό του $\frac{\partial E}{\partial NET_t}$, όπου E η συνάρτηση σφάλματος και NET_t η συνάρτηση ενεργοποίησης νευρώνα, αρχικοποίησε την μεταβλητή σήματος σφάλματος πραγματικών τιμών, $\delta_t = 0$;
- Αν x_t είναι συμβάν εισόδου, συνεχίστε με την επόμενη επανάληψη;
- Αν υπάρχει σφάλμα e_t τότε $\delta_t = x_t - d_t$;
- Προσθέστε στην τιμή δ_t την τιμή $\sum_{k \in E \rightarrow \Delta O_NN_t} Wv(t, k) \delta_k$; (αυτός είναι ο κομψός και αποτελεσματικός κανόνας εφαρμογής της παραγωγίσιμης αναδρομικής αλυσίδας σύνθετων συναρτήσεων, συλλέγοντας τις επιδράσεις της συνάρτησης ενεργοποίησης net_t σε μελλοντικά γεγονότα σφαλμάτων στην έξοδο του δικτύου);
- Πολλαπλασιάστε την τιμή δ_t με $f_t'(net_t)$;
- Για όλους τους νευρώνες $k \in E \rightarrow \Delta O_NN_t$ προσθέστε την τιμή $x_k * \delta_t$ στην μεταβολή βαρών $\Delta_{Wv(t,k)}$

Τέλος της για (for loop)

Τέλος, αλλάξτε κάθε βάρος w_i σε αναλογία με την μεταβολή του Δ_i και μια μικρή πραγματική τιμή που αντιστοιχεί στον συντελεστή εκμάθησης

Μέχρι και σήμερα, αυτή η απλή μέθοδος BP εξακολουθεί να είναι ο κεντρικός αλγόριθμος μάθησης για τα Deep MLPs και RNNs. Ιδιαίτερα, τα περισσότερα NNs που κέρδισαν τους διαγωνισμούς βασίστηκαν σε αυτό τον αλγόριθμο και δεν επαύξησαν τις δυνατότητές τους με κάποιο είδος μη επιτηρούμενης μάθησης όπως είναι εφικτό από σύγχρονους υβριδικούς αλγόριθμους.

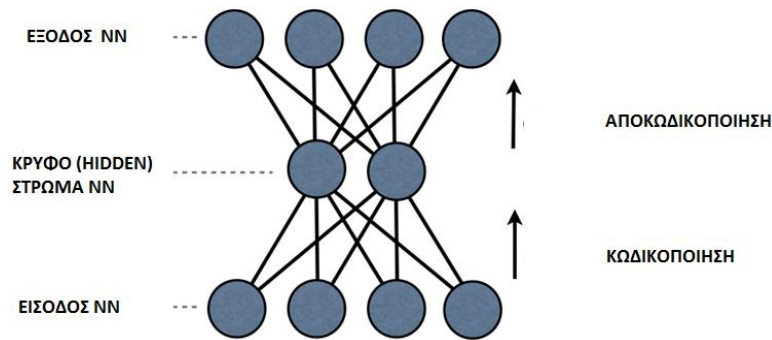
Στα παρακάτω σχήματα εμφανίζεται η κυρίαρχη αρχή εκμάθησης MLP μέσω του αλγόριθμου BP και της αρχής της αυτοκωδικοποίησης. [91], [92], [93], [94], [95]



Εικόνα 6.1

Τα στάδια εκπαίδευσης ενός πολυστρωματικού NN.

Η κεντρική ιδέα της εκπαίδευσης Deep NN με βάση την αρχή αυτοκωδικοποίησης
 Ο αλγόριθμος αυτοκωδικοποίησης (auto-encoder) συνίσταται σε εκπαίδευση με τον κλασικό BP αλγόριθμο ή κάποια standard παραλλαγή του ώστε να αναπαράγεται στην έξοδο το διάνυσμα της εισόδου για όλα τα πρότυπα εισόδου. Κάνοντας αυτό να συμβαίνει με λιγότερες μονάδες επεξεργασίας από τις μονάδες εισόδου, αυτό αναγκάζει τις μονάδες 'κρυμμένου στρώματος' να γίνουν καλοί ανιχνευτές χαρακτηριστικών



Εικόνα 6.2
Η δομή ενός πολυστρωματικού NN.

6.2 Εφαρμογή Βαθιάς Μάθησης σε MLPs και RNNs για την Ενίσχυση της Μάθησης

Μέχρι στιγμής έχουμε επικεντρωθεί στη Βαθιά Μάθηση (DL) σε επιτηρούμενα κυρίως και επιγραμματικά μόνο σε μη επιτηρούμενα NNs. Αυτά τα NN μαθαίνουν να αντιλαμβάνονται / να κωδικοποιούν / να προβλέπουν / ταξινομούν μοτίβα/patterns ή ακολουθίες μοτίβων/patterns, αλλά δεν μαθαίνουν να ενεργούν με τη γενικότερη έννοια της ενίσχυσης της μάθησης (Reinforcement Learning - RL) σε άγνωστα περιβάλλοντα (βλ. έρευνες, π.χ., Kaelbling et al., 1996; Sutton and Barto, 1998, Wiering και van Otterlo, 2012). Στο παρόν τμήμα του κεφαλαίου αυτού επιχειρείται μια συζήτηση των DL MLP και RNNs για RL. Χωρίς επόπτη, αποκλειστικά από περιστασιακά σήματα ανταμοιβής και τιμωρίας, οι πράκτορες RL πρέπει να ανακαλύψουν πώς να υπάρξει επιτυχής αλληλεπίδραση με ένα δυναμικό, αρχικά άγνωστο περιβάλλον για να μεγιστοποιηθούν τα αναμενόμενα σωρευτικά σήματα ανταμοιβής. Μπορεί να υπάρχουν αυθαίρετες, a priori άγνωστες καθυστερήσεις μεταξύ των ενεργειών και των αντιληπτών συνεπειών/αποτελεσμάτων. Το πρόβλημα είναι τόσο δύσκολο όσο οποιοδήποτε πρόβλημα της επιστήμης των υπολογιστών, δεδομένου ότι οποιοσδήποτε το έργο με μια αξιόπιστη περιγραφή μπορεί να διαμορφωθεί στο πλαίσιο RL (π.χ., Hutter, 2005). Για παράδειγμα, μια απάντηση στο κλασικό ερώτημα για το εάν το πρόβλημα $P = NP$ (Levin, 1973 · Cook, 1971) θα μπορούσε επίσης να θέσει όρια για το τι είναι εφικτό από τη γενική RL μεθοδολογία. Μπορούμε επίσης να συγκρίνουμε τους περιορισμούς της μεθοδολογίας RL με βάση πιο συγκεκριμένους περιορισμούς, π.χ., (Blondel και Tsitsiklis, 2000, Madani et al., 2003; Vlassis et al, 2012). Τα ακόλουθα εδάφια του παρόντος κεφαλαίου επικεντρώνονται κυρίως σε ορισμένες προφανείς διασταυρώσεις μεταξύ DL και RL - δεν μπορούν να χρησιμεύσουν ως γενική επισκόπηση στην έρευνα RL. [102], [103], [104], [105], [105], [107], [108], [109], [110]

6.2.1 Υλοποίηση RNNs με Deep CAPs μέσω εφαρμογής μεθόδου RL σε NNs

Στην ειδική περίπτωση ενός ελεγκτή RN MLP που αλληλοεπιδρά με ένα ντετερμινιστικό, προβλέψιμο περιβάλλον, ένα ξεχωριστό MLP που καλούμε M μπορεί να μάθει να ανταποκρίνεται ως παγκόσμιο μοντέλο της C μέσω της αναγνώρισης του συστήματος (system identification), (Werbos, 1989, Munro, 1987, Robinson and Fallside, 1989; Schmidhuber, 1990; Narendra and Parthasarathy, 1990; Cochocki και Unbehauen, 1993; Levin and Narendra, 1995; Ljung, 1998; Prokhorov et al., 2001; Ge et al., 2010). Για παράδειγμα, υποθέτουμε ότι ο M έχει μάθει να παράγει ακριβείς προβλέψεις. Μπορούμε να χρησιμοποιήσουμε το M για να το αντικαταστήσουμε το περιβάλλον.

Στη συνέχεια, οι M και C σχηματίζουν ένα RNN όπου οι έξοδοι του M γίνονται είσοδοι του C, του οποίου οι έξοδοι (ενέργειες) με τη σειρά τους γίνονται είσοδοι του M. Τώρα, η BP για RNNs (αλγόριθμος 6.1) μπορεί να χρησιμοποιηθεί για την επίτευξη των επιθυμητών συμβάντων εισόδου, όπως τα υψηλά πραγματικών τιμών σήματα ανταμοιβής: Ενώ τα βάρη M παραμένουν σταθερά, οι πληροφορίες κλίσης για τα βάρη του C μεταδίδονται πίσω από το M προς τα κάτω στο C και πίσω μέσω του M κ.λπ. Σε κάποιο βαθμό, η προσέγγιση εφαρμόζεται επίσης σε πιθανοτικά ή αβέβαια περιβάλλοντα, για όσο διάστημα τα εσωτερικά γινόμενα των εκτιμήσεων κλίσης με βάση το C της M και των "πραγματικών" κλίσεων του M τείνουν να είναι θετικές. Γενικά, αυτή η προσέγγιση συνεπάγεται βαθιά CAPs για το C, αντίθετα με την παραδοσιακή RL που βασίζεται σε DP (Δυναμικό Προγραμματισμό, Dynamic Programming).

Πριν από δεκαετίες, η μέθοδος χρησιμοποιήθηκε για να μάθει να υποστηρίζει την οδήγηση ενός μοντέλου φορτηγού (Nguyen and Widrow, 1989). Ένα ενεργό σύστημα όρασης RL χρησιμοποιήθηκε για να μάθει διαδοχικές μετατοπίσεις (saccades) της ωχράς κηλίδας (fovea), για την ανίχνευση στόχων σε οπτικές σκηνές (Schmidhuber και Huber, 1991), μαθαίνοντας έτσι τον έλεγχο της επιλεκτικής προσοχής. Μπορεί να συγκριθεί επίσης η βασισμένη στην RL επιλεκτική μάθηση χωρίς την χρήση NN (Whitehead, 1992). Για να επιτρέπεται η ανάμνηση προηγούμενων γεγονότων σε μερικώς παρατηρούμενους κόσμους, η πιο γενική παραλλαγή αυτής της τεχνικής χρησιμοποιεί RNNs αντί για MLPs για την εφαρμογή και των δύο M και C (Schmidhuber, 1990, Feldkamp & Puskorius, 1998). Αυτό μπορεί να προκαλέσει βαθιά CAPs όχι μόνο για το C, αλλά επίσης για M.

Το Βαθύ νευρωνικό σύστημα M μπορεί επίσης να χρησιμοποιηθεί για τη βελτιστοποίηση της αναμενόμενης ανταμοιβής με το σχεδιασμό μελλοντικών αλληλουχιών δράσης (Schmidhuber, 1990). Στην πραγματικότητα, οι νικητές του Παγκόσμιου Πρωταθλήματος RoboCup 2004, στην γρήγορη κατηγορία (Egozua et al., 2004) εκπαίδευσαν NN για την πρόβλεψη των αποτελεσμάτων των σημάτων διεύθυνσης των γρήγορων ρομπότ με 4 κινητήρες για 4 διαφορετικούς τροχούς. Κατά τη διάρκεια του παιχνιδιού, τέτοια μοντέλα NN χρησιμοποιήθηκαν για την επίτευξη επιθυμητών υποστόχων (subgoals), βελτιστοποιώντας τις ακολουθίες δράσης με γρήγορο προγραμματισμό των μελλοντικών κινήσεων του συστήματος.

Αυτή η προσέγγιση ήταν επίσης κατάλληλη για χρήση για τη δημιουργία ρομπότ αυτό-αντιστάθμισης σφαλμάτων (self-healing) ικανών να αντισταθμίσουν ελαττωματικούς κινητήρες των οποίων τα αποτελέσματα δεν ταιριάζουν με τις προβλέψεις των μοντέλων NN (Gloye et al., 2005, Schmidhuber, 2007). Συνήθως το σύστημα M δεν δίνεται εκ των προτέρων.

Στη συνέχεια, ένα ουσιαστικό ερώτημα είναι: σε ποια πειράματα πρέπει να υποβληθεί το σύστημα C για τη γρήγορη βελτίωση της συμπεριφοράς του M;

Η τυπική θεωρία της διασκέδασης και της δημιουργικότητας (π.χ. Schmidhuber, 2013) παρέχει τον φορμαλισμό των κινητήριων δυνάμεων και των συναρτήσεων αξίας που βρίσκονται πίσω από μια τέτοια περίεργη και εξερευνητική συμπεριφορά:

Ένα μέτρο της προόδου της μάθησης του M γίνεται η εγγενής ανταμοιβή του C (Oudeyer et al., 2013). Αυτό παρακινεί το C να δημιουργήσει ακολουθίες δράσης (πειράματα) έτσι ώστε το M να κάνει ταχεία πρόοδο.

[111], [112], [113], [114], [115], [116], [117], [118], [119], [120], [121], [122], [123], [124], [125], [126], [127], [128], [129], [130], [131]

6.2.2 Deep MLPs σε διαδικασίες λήψης αποφάσεων RL και Markov

Η κλασική προσέγγιση του RL (Samuel, 1959, Bertsekas και Tsitsiklis, 1996) χρησιμοποιεί την απλουστευτική υπόθεση των Markov Decision Processes (MDPs): η τρέχουσα είσοδος του πράκτορα RL μεταφέρει όλες τις πληροφορίες που είναι απαραίτητες για τον υπολογισμό ενός βέλτιστου επόμενου γεγονότος ή απόφασης εξόδου. Αυτό επιτρέπει την σημαντική μείωση του βάρους CAP (Credit Assignment Path) σε RL NNs χρησιμοποιώντας το τέχνασμα του Δυναμικού Προγραμματισμού Dynamic Programming (DP) (Bellman, 1957). Το τελευταίο εξηγείται συχνά σε ένα πιθανοτικό πλαίσιο (π.χ., Sutton and Barto, 1998), αλλά η βασική του ιδέα μπορεί ήδη να μεταφερθεί σε ένα ντετερμινιστικό πλαίσιο.

Για λόγους απλότητας, ως υποθέσουμε ότι τα γεγονότα εισόδου x_t κωδικοποιούν ολόκληρη την τρέχουσα κατάσταση του περιβάλλοντος, συμπεριλαμβανομένης μιας *realvalued* ανταμοιβής πραγματικής τιμής r_t (δεν χρειάζεται να εισαχθούν επιπλέον συμβολισμοί, δεδομένου ότι οι πραγματικές τιμές μπορούν να κωδικοποιούν οποιαδήποτε διανύσματα πραγματικών τιμών). Ο αρχικός στόχος RL (να βρεθούν βάρη που μεγιστοποιούν το άθροισμα όλων των ανταμοιβών ενός επεισοδίου στην όλη ακολουθία γεγονότων του προβλήματος) αντικαθίσταται από ένα ισοδύναμο σύνολο εναλλακτικών στόχων που καθορίζονται από μια συνάρτηση πραγματικών τιμών V που ορίζεται με βάση τα συμβάντα εισόδου. Ας θεωρήσουμε οποιαδήποτε δύο επόμενα συμβάντα εισόδου x_t, x_k . Αναδρομικά ορίζουμε $V(x_t) = r_t + V(x_k)$, όπου $V(x_k) = r_k + V(x_{k+1})$ είναι το τελευταίο συμβάν εισόδου. Στην συνέχεια ψάχνουμε για τα βάρη που μεγιστοποιούν το V όλων των συμβάντων εισόδου, προκαλώντας κατάλληλα γεγονότα ή ενέργειες εξόδου. Λόγω της υποθέσεως Markov, ένα νευρωνικό σύστημα MLP αρκεί για να εφαρμόσει την πολιτική που χαρτογραφεί τον χώρο γεγονότων εισόδου στον χώρο γεγονότων εξόδου.

Τα σχετικά CAPs δεν είναι βαθύτερα από αυτό του MLP. Το ίδιο το V συχνά μοντελοποιείται από ένα ξεχωριστό MLP, (που επίσης αποδίδει τυπικά όχι πολύ Βαθιά CAPs) και μαθαίνει να προσεγγίζει το $V(x_t)$ μόνο από τοπικές πληροφορίες $r_t, V(x_k)$. Υπάρχουν πολλές παραλλαγές της παραδοσιακής RL (π.χ., Tsitsiklis et al., 1996; Brafman and Tenenbaum, 2002; Abounadi et al., 2002; Sutton et al., 2008; Maei και Sutton, 2010, van Hasselt, 2012). Οι περισσότεροι φορμαλισμοί διατυπώνονται σε ένα πιθανοτικό πλαίσιο και αξιολογούν τα ζεύγη γεγονότων εισόδου και εξόδου (δράσης) (αντί για αξιολόγηση μόνο των γεγονότων εισόδου). Για να διευκολυνθούν ορισμένες μαθηματικές διατυπώσεις, χρησιμοποιούνται σήματα ανταμοιβής με καθυστέρηση, αλλά αυτές οι στρεβλώσεις του αρχικού προβλήματος RL είναι προβληματικές.

Ίσως το πιο γνωστό RL NN είναι ο παγκοσμίου επιπέδου παίκτης τάβλι RL (Tesauro, 1994), που πέτυχε να ανταγωνιστεί στο επίπεδο των παγκόσμιων πρωταθλητών παίζοντας εναντίον τους. Το μη γραμμικό, μάλλον ρηχό MLP, που χρησιμοποιήθηκε χαρτογραφεί έναν μεγάλο αλλά πεπερασμένο αριθμό διακριτών καταστάσεων του πίνακα σε τιμές. Πιο πρόσφατα, ένα μάλλον βαθύ (deep) GPU-CNN χρησιμοποιήθηκε σε ένα παραδοσιακό πλαίσιο RL για να παίξει με αρκετούς υπολογιστές Atari 2600 παιχνίδια απευθείας από την είσοδο, που είναι βίντεο των 84x84 εικονοστοιχείων στα 60 Hz (Mnih et al., 2013), χρησιμοποιώντας επανάληψη εμπειρίας (Lin, 1993), εργασία που επεκτείνει τις προηγούμενες έρευνες σχετικά με την νευρο-προσαρμοσμένη Q-Learning (NFQ) (Riedmiller, 2005) μεθοδολογία.

Ακόμη καλύτερα αποτελέσματα επιτυγχάνονται με τη χρήση (αργού) δενδρικού σχεδιασμού Monte Carlo για την εκπαίδευση συγκριτικά γρήγορων και βαθιών NNs (Guo κ.ά., 2014). Η μεθοδολογία αυτή συγκρίνεται με την μεθοδολογία RL με βάση RBM (Sallans and Hinton, 2004) με υψηλή διάσταση διανύσματος εισόδου (Elfwing et al., 2010), με αυτή παλαιότερων παικτών της RL Atari (Gruttner et al., 2010), καθώς και μία παλαιότερη μεθοδολογία βασισμένη σε πρώιμο ακατέργαστο βίντεο (Koutnik et al., 2013) που εκπαιδεύεται από Έμμεση Αναζήτηση Πολιτικής. Παρακάτω εμφανίζεται ακριβώς ο αλγόριθμος Deep Q-RL στην εφαρμογή του σε επανάληψη εμπειρίας για το πρόβλημα των παιχνιδιών ATARI που αναφέρθηκε ανωτέρω (Mnih et al., 2013). [\[132\]](#), [\[133\]](#), [\[134\]](#), [\[135\]](#), [\[136\]](#), [\[137\]](#), [\[138\]](#)

Algorithm 1: deep Q-learning with experience replay.Initialize replay memory D to capacity N Initialize action-value function Q with random weights θ Initialize target action-value function \hat{Q} with weights $\theta^- = \theta$ **For** episode = 1, M **do**Initialize sequence $s_1 = \{x_1\}$ and preprocessed sequence $\phi_1 = \phi(s_1)$ **For** $t = 1, T$ **do**With probability ϵ select a random action a_t otherwise select $a_t = \operatorname{argmax}_a Q(\phi(s_t), a; \theta)$ Execute action a_t in emulator and observe reward r_t and image x_{t+1} Set $s_{t+1} = s_t, a_t, x_{t+1}$ and preprocess $\phi_{t+1} = \phi(s_{t+1})$ Store transition $(\phi_t, a_t, r_t, \phi_{t+1})$ in D Sample random minibatch of transitions $(\phi_j, a_j, r_j, \phi_{j+1})$ from D Set $y_j = \begin{cases} r_j & \text{if episode terminates at step } j+1 \\ r_j + \gamma \max_{a'} \hat{Q}(\phi_{j+1}, a'; \theta^-) & \text{otherwise} \end{cases}$ Perform a gradient descent step on $(y_j - Q(\phi_j, a_j; \theta))^2$ with respect to the network parameters θ Every C steps reset $\hat{Q} = Q$ **End For****End For**

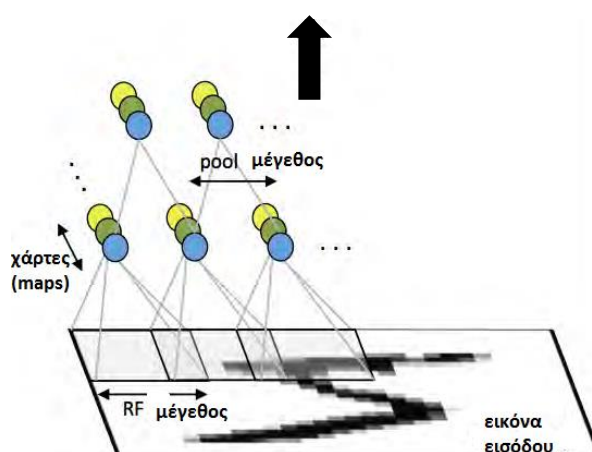
6.3 Υλοποίηση της Διαδικασίας Βαθιάς Εκπαίδευσης σε MLPs στην πλατφόρμα MATLAB

Αυτό το μέρος του κεφαλαίου αναφέρεται στην επίδειξη της δυνατότητας και τεχνικής λειτουργίας των κλασικών Convolutional Deep NN που αναλύθηκαν στο προηγούμενο κεφάλαιο. Τα Convolutional Deep NN, θα χρησιμοποιηθούν στο επόμενο κεφάλαιο 7 για την αναγνώριση πεζών, μέσω MATLAB, σε προβλήματα αναγνώρισης αντικειμένων μέσω της τεχνικής του Deep Learning. Προς τούτο μελετάται εδώ η διαδικασία που ακολουθείται στην πλατφόρμα MATLAB, χρησιμοποιώντας παραδείγματα της MATHWORKS. Προτού παρουσιαστούν τα εν λόγω παραδείγματα σε MATLAB, θα παρουσιαστεί μία περισσότερη τεχνική περιγραφή της αρχιτεκτονικής των CNN.

6.3.1 Αρχιτεκτονική των CNN

Όπως αναφέραμε και προηγουμένως, ένα CNN αποτελείται από ένα ή περισσότερα επίπεδα συνέλιξης (convolutional layers) συχνά μαζί με ένα επίπεδο υποδειγματοληψίας (πχ pooling) ακολουθούμενο από ένα ή περισσότερα fully connected επίπεδα όπως συμβαίνει και σε ένα κλασικό πολύ-επίπεδο (πολυστρωματικό) NN τύπου MLP. Η αρχιτεκτονική του CNN σχεδιάζεται έτσι ώστε να εκμεταλλεύεται την 2D δομή των εικόνων εισόδου ή άλλα 2D σήματα όπως σήματα ήχου. Αυτό επιτυγχάνεται με τοπικές συνδέσεις και κατάλληλα βάρη ακολουθούμενα από pooling προκειμένου να δημιουργηθούν ανεξάρτητα μετατοπίσεων (translation invariant) χαρακτηριστικά. Άλλο ένα προσόν των CNNs είναι ότι είναι ευκολότερα στην εκπαίδευση και έχουν πολύ λιγότερες παραμέτρους από τα fully connected NN με τον ίδιο αριθμό κρυφών επιπέδων.

Η είσοδος σε ένα επίπεδο συνέλιξης είναι μια $M \times M \times R$ εικόνα όπου το M είναι το ύψος και το πλάτος της εικόνας ενώ το R ο αριθμός των καναλιών πχ για RGB, $R=3$. Το επίπεδο συνέλιξης έχει K φίλτρα (kernels) μεγέθους $N \times N \times Q$ όπου N είναι μικρότερο από τη διάσταση της εικόνας και Q μπορεί να είναι ίδιου μεγέθους με τα κανάλια ή μικρότερου και μπορεί να ποικίλει για κάθε kernel. Το μέγεθος των φίλτρων προκαλεί τοπικά συνδεδεμένη δομή όπου καθένα συνελίσσεται με κάθε εικόνα για να παράγουν χάρτες K χαρακτηριστικών (feature maps) μεγέθους $M-N+1$. Κάθε χαρακτηριστικό υποδειγματοληπτείται τυπικά με mean ή max pooling σε $P \times P$ συνεχείς περιοχές όπου το P παίρνει τυπικά τιμές μεταξύ 2 και κοντά στο 5 για μεγάλες εικόνες εισόδου. Πριν ή μετά το pooling layer ακολουθεί μια προσθήκη πόλωσης (bias) και σιγμοειδής μη γραμμικότητα σε κάθε χάρτη χαρακτηριστικών. Το σχήμα της κατωτέρω εικόνας παρουσιάζει ένα CNN αποτελούμενο από επίπεδα συνέλιξης και pooling. Οι νευρώνες με το ίδιο χρώμα έχουν παρόμοια βάρη.



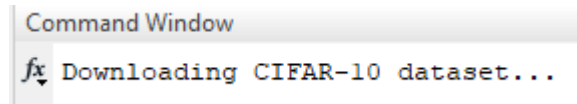
Εικόνα 6.3
Η δομή ενός CNN.

6.3.2 Παράδειγμα 1: αναγνώριση αντικειμένων με εφαρμογή Βαθιάς Μάθησης

Σε πρώτο βήμα θα πρέπει να ακολουθήσουμε τις εντολές της MATHWORKS step by step ώστε να εγκαταστήσουμε (σε ένα temporary file) σωστά το CIFAR-10 dataset το οποίο είναι απαραίτητο για την υλοποίηση των πειραμάτων:

```
%----Neural Networks
clear all
close all
clc
% katevazume to CIFAR-10 data set se ena prosorino directory
cifar10Data = tempdir;
url = 'https://www.cs.toronto.edu/~kriz/cifar-10-matlab.tar.gz';
helperCIFAR10Data.download(url, cifar10Data);
```

Μετά την εκτέλεση του παραπάνω, το CIFAR-10 dataset μεταφορτώνεται:



Ακολούθως, πραγματοποιείται μεταφόρτωση των δεδομένων του CIFAR-10:

```
% Load the CIFAR-10 training and test data.
[trainingImages, trainingLabels, testImages, testLabels] = helperCIFAR10Data.load(cifar10Data);

% Each image is a 32x32 RGB image and there are 50,000 training samples.
size(trainingImages)
```

Όπως διακρίνεται στο workspace, η μεταφόρτωση των δεδομένων ολοκληρώθηκε επιτυχώς:

Workspace	
Name ▲	Value
ans	[32 32 3 50000]
cifar10Data	'C:\Users\vangelis\AppData\Lo...
testImages	4-D uint8
testLabels	10000x1 categorical
trainingImages	4-D uint8
trainingLabels	50000x1 categorical
url	'https://www.cs.toronto.edu...

Εικόνα 6.4

Το Workspace του MATLAB μετά την φόρτωση του dataset.

Κάθε εικόνα αποτελείται από 32x32 RGB pixels και υπάρχουν 50.000 εκπαιδευτικά samples:

```
ans =
    32    32     3   50000
```


Το dataset θα έχει επίσης τις ακόλουθες 10 κατηγορίες:

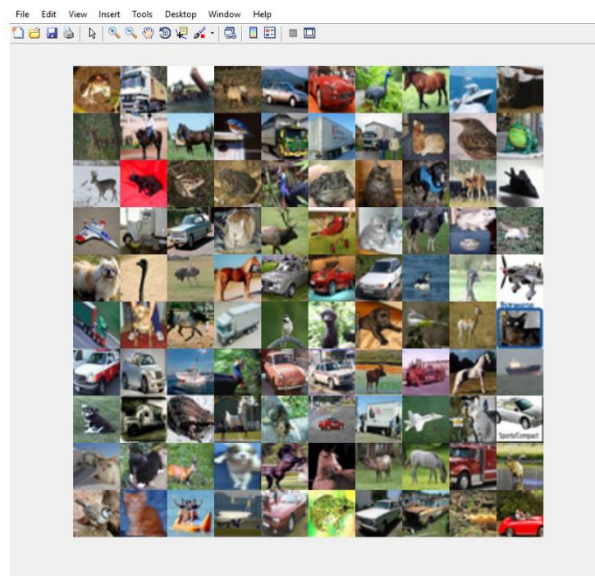
```
% το CIFAR-10 έχει 10 κατηγορίες εικόνων τις οποίες θα δούμε σε μια lista
numImageCategories = 10;
categories(trainingLabels)
```

ans =

```
10x1 cell array

    {'airplane' }
    {'automobile'}
    {'bird' }
    {'cat' }
    {'deer' }
    {'dog' }
    {'frog' }
    {'horse' }
    {'ship' }
    {'truck' }
```

Παρακάτω, παρουσιάζονται μερικές από τις εκπαιδευτικές εικόνες:



Εικόνα 6.5

Οι εκπαιδευτικές εικόνες μετά από αντίστοιχο resizing στο MATLAB, ώστε να καταστεί δυνατή η προβολή τους.

Στην συνέχεια δομείται το Δίκτυο:

Το Νευρωνικό Δίκτυο είναι πολυστρωματικό (multilayer) και σε κάθε layer γίνεται και από ένας υπολογισμός. Στην περίπτωση μας τα layers είναι:

- **imageInputLayer**
- **convolution2dLayer**
- **reluLayer**
- **maxPooling2dLayer**
- **fullyConnectedLayer**
- **softmaxLayer**
- **classificationLayer**

Η δόμηση του δικτύου ξεκινάει από το image input layer και προς τα κάτω:

```
% dimiourgoume to image input layer
[height, width, numChannels, ~] = size(trainingImages);

imageSize = [height width numChannels];
inputLayer = imageInputLayer(imageSize)
```

Η έξοδος στο command window είναι η εξής:

```
inputLayer =
```

[ImageInputLayer](#) with properties:

```
      Name: ''
  InputSize: [32 32 3]
```

Hyperparameters

```
DataAugmentation: 'none'
  Normalization: 'zerocenter'
```

Εν συνεχεία, προχωράμε την υλοποίηση με τα middle layers τα οποία είναι τα εξής:

- **convolution2dLayer**
- **reluLayer**
- **maxPooling2dLayer**

```
%sinexeizoyme kai me ta middle layers
filterSize = [5 5];
numFilters = 32;
middleLayers = [
% The first convolutional layer
convolution2dLayer(filterSize, numFilters, 'Padding', 2)
% Next add the ReLU layer:
reluLayer()
% adding maxPooling2dLayer too
maxPooling2dLayer(3, 'Stride', 2)
% Repeat the 3 core layers to complete the middle of the network.
convolution2dLayer(filterSize, numFilters, 'Padding', 2)
reluLayer()
maxPooling2dLayer(3, 'Stride', 2)
convolution2dLayer(filterSize, 2 * numFilters, 'Padding', 2)
reluLayer()
maxPooling2dLayer(3, 'Stride', 2)
]
```

Παρακάτω παρατίθενται τα αποτελέσματα από το command window:

```
middleLayers =
```

```
9x1 Layer array with layers:
```

```
1 '' Convolution 32 5x5 convolutions with stride [1 1] and padding [2 2 2 2]
2 '' ReLU ReLU
3 '' Max Pooling 3x3 max pooling with stride [2 2] and padding [0 0 0 0]
4 '' Convolution 32 5x5 convolutions with stride [1 1] and padding [2 2 2 2]
5 '' ReLU ReLU
6 '' Max Pooling 3x3 max pooling with stride [2 2] and padding [0 0 0 0]
7 '' Convolution 64 5x5 convolutions with stride [1 1] and padding [2 2 2 2]
8 '' ReLU ReLU
9 '' Max Pooling 3x3 max pooling with stride [2 2] and padding [0 0 0 0]
```

Ομοιοτρόπως δομούμε και τα final layers:

- **fullyConnectedLayer**
- **softmaxLayer**
- **classificationLayer**

Κώδικας MATLAB:

```
%final layers
finalLayers = [
% fully connected layer
fullyConnectedLayer(64)
% Add an ReLU non-linearity.
reluLayer
% Add the last fully connected layer
fullyConnectedLayer(numImageCategories)
% Add the softmax loss layer and classification layer.
softmaxLayer
classificationLayer
]
```

Αποτελέσματα command window:

```
finalLayers =
```

```
5x1 Layer array with layers:
```

```
1 '' Fully Connected 64 fully connected layer
2 '' ReLU ReLU
3 '' Fully Connected 10 fully connected layer
4 '' Softmax softmax
5 '' Classification Output crossentropyex
```

Εν συνέχεια, θέτοντας βάρη στο σύστημα είμαστε σε θέση να εκπαιδύσουμε το σύστημα:

```

% Combine the input, middle, and final layers.
layers = [
    inputLayer
    middleLayers
    finalLayers
]

%%
% Initialize the first convolutional layer weights using normally
% distributed random numbers with standard deviation of 0.0001. This helps
% improve the convergence of training.

layers(2).Weights = 0.0001 * randn([filterSize numChannels numFilters]);

%% Train CNN Using CIFAR-10 Data

opts = trainingOptions('sgdm', ...
    'Momentum', 0.9, ...
    'InitialLearnRate', 0.001, ...
    'LearnRateSchedule', 'piecewise', ...
    'LearnRateDropFactor', 0.1, ...
    'LearnRateDropPeriod', 8, ...
    'L2Regularization', 0.004, ...
    'MaxEpochs', 40, ...
    'MiniBatchSize', 128, ...
    'Verbose', true);

% Train the network using the |trainNetwork| function.
doTraining = false;

if doTraining
    % Train a network.
    cifar10Net = trainNetwork(trainingImages, trainingLabels, layers, opts);
else
    % Load pre-trained detector for the example.
    load('rcnnStopSigns.mat','cifar10Net')
end

```

Στη συνέχεια, ακολουθεί ο έλεγχος της εκπαίδευσης του δικτύου:

```

%% Validate CIFAR-10 Network Training
w = cifar10Net.Layers(2).Weights;
% rescale and resize the weights for better visualization
w = mat2gray(w);
w = imresize(w, [100 100]);
figure
montage(w)

%%
% Run the network on the test set.
YTest = classify(cifar10Net, testImages);
% Calculate the accuracy.
accuracy = sum(YTest == testLabels)/numel(testLabels)

```

Σημείωση: Όσο πιο πολύ εκπαιδεύεται ένα Δίκτυο, αυξάνεται η ακρίβεια του.

Σε επόμενο βήμα, σειρά έχει η εκτέλεση του αλγορίθμου και η εξαγωγή των αποτελεσμάτων:

```

%% Load Training Data
data = load('stopSignsAndCars.mat', 'stopSignsAndCars');
stopSignsAndCars = data.stopSignsAndCars;

% Update the path to the image files to match the local file system
visiondata = fullfile(toolboxdir('vision'),'visiondata');
stopSignsAndCars.imageFilename = fullfile(visiondata, stopSignsAndCars.imageFilename);

% Display a summary of the ground truth data
summary(stopSignsAndCars)

stopSigns = stopSignsAndCars(:, {'imageFilename','stopSign'});
% Display one training image and the ground truth bounding boxes
I = imread(stopSigns.imageFilename{1});
I = insertObjectAnnotation(I, 'Rectangle', stopSigns.stopSign{1}, 'stop sign', 'LineWidth', 8);

figure
imshow(I)

doTraining = false;

if doTraining

    % Set training options
    options = trainingOptions('sgdm', ...
        'MiniBatchSize', 128, ...
        'InitialLearnRate', 1e-3, ...
        'LearnRateSchedule', 'piecewise', ...
        'LearnRateDropFactor', 0.1, ...
        'LearnRateDropPeriod', 100, ...
        'MaxEpochs', 100, ...
        'Verbose', true);

    % Train an R-CNN object detector. This will take several minutes.
    rcnn = trainRCNNObjectDetector(stopSigns, cifar10Net, options, ...
        'NegativeOverlapRange', [0 0.3], 'PositiveOverlapRange',[0.5 1])
else
    % Load pre-trained network for the example.
    load('rcnnStopSigns.mat','rcnn')
end

```

```

%% Test R-CNN Stop Sign Detector
testImage = imread('stopSignTest.jpg');
% Detect stop signs
[bboxes, score, label] = detect(rcnn, testImage, 'MiniBatchSize', 128)

%%
% Display the detection results
[score, idx] = max(score);
bbox = bboxes(idx, :);
annotation = sprintf('%s: (Confidence = %f)', label(idx), score);
outputImage = insertObjectAnnotation(testImage, 'rectangle', bbox, annotation);
figure
imshow(outputImage)

%% Debugging Tips
rcnn.Network

%%
featureMap = activations(rcnn.Network, testImage, 'softmax', 'OutputAs', 'channels');
% The softmax activations are stored in a 3-D array.
size(featureMap)

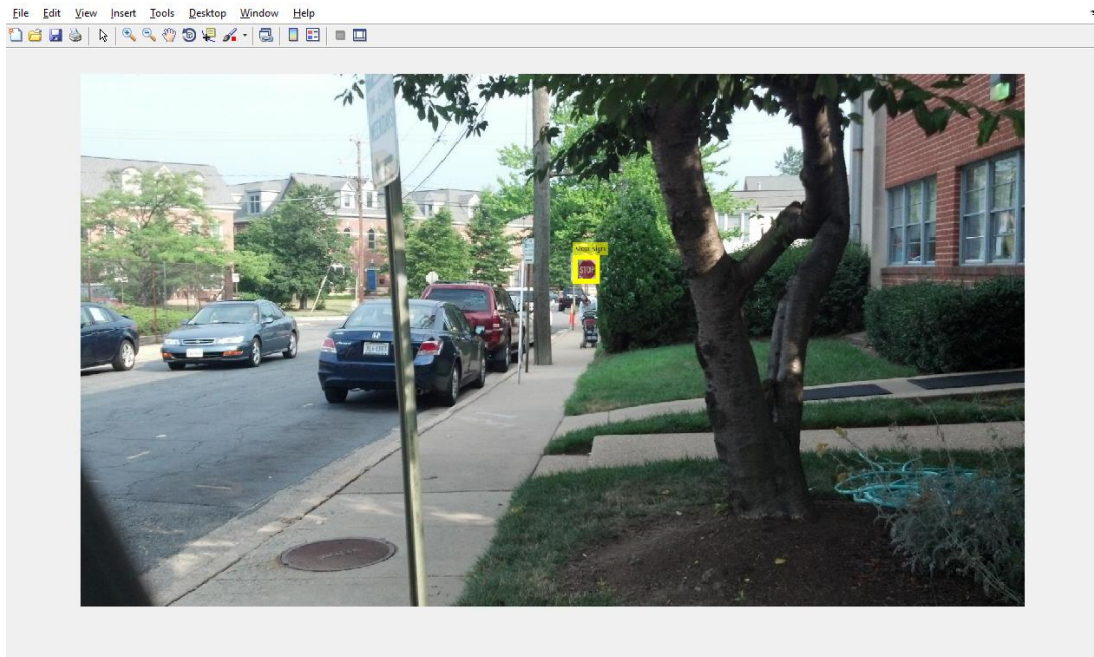
%%
% The 3rd dimension in featureMap corresponds to the object classes
rcnn.ClassNames

%%
% The stop sign feature map is stored in the first channel
stopSignMap = featureMap(:, :, 1);

%%
[height, width, ~] = size(testImage);
stopSignMap = imresize(stopSignMap, [height, width]);
% Visualize the feature map superimposed on the test image.
featureMapOnImage = imfuse(testImage, stopSignMap);
figure
imshow(featureMapOnImage)

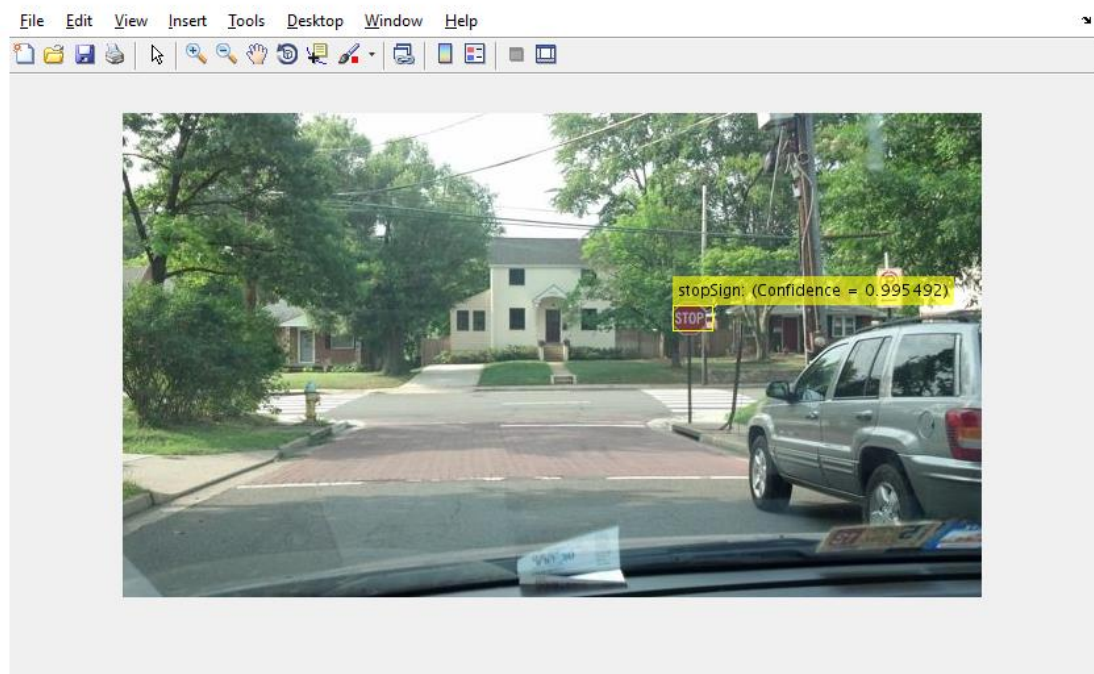
```


Στις ακόλουθες εικόνες παρουσιάζονται τα αποτελέσματα αναγνώρισης του αλγορίθμου.



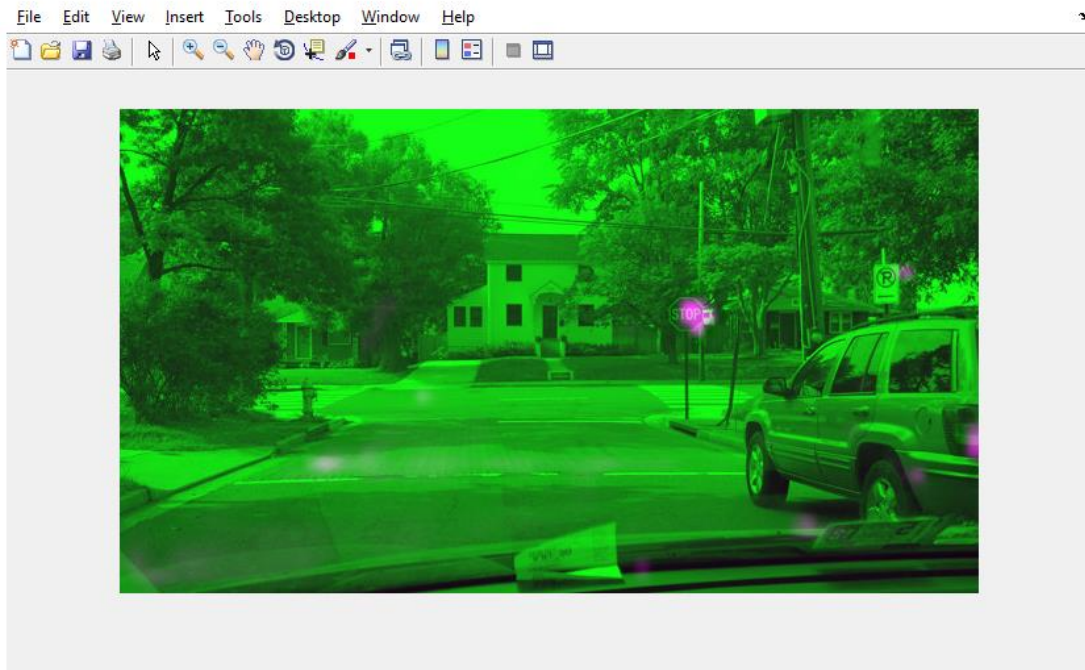
Εικόνα 6.6

Αναγνώριση του οδικού σήματος STOP σε εικόνα, από τον αλγόριθμο.



Εικόνα 6.7

Αναγνώριση του οδικού σήματος STOP σε εικόνα, από τον αλγόριθμο.



Εικόνα 6.8
Αναγνώριση του οδικού σήματος STOP σε εικόνα, από τον αλγόριθμο.

Όπως προκύπτει, ο αλγόριθμος αναγνώρισε επιτυχώς τις οδικές σημάσεις (STOP) μέσα από την φωτογραφία οπότε και εκπλήρωσε το σκοπό του:

```
score =          label =
single          categorical
0.9955          stopSign
```

Συμπέρασμα: Η μέθοδος *deep learning* είναι μία ιδιαίτερα αποτελεσματική μέθοδος όπου μπορεί να αναγνωρίσει αντικείμενα που οι άλλες μέθοδοι δεν θα μπορούσαν.

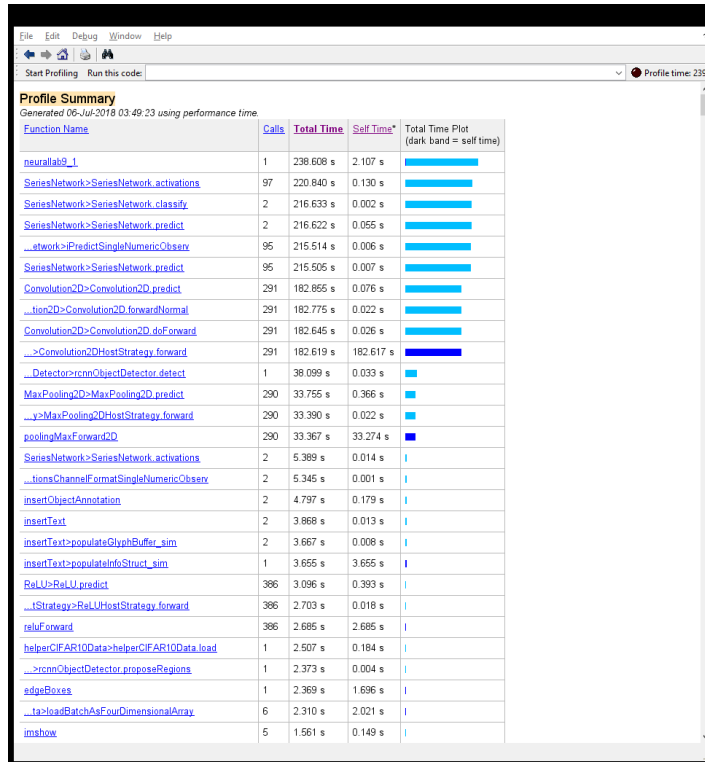
Ωστόσο, το μειονέκτημα της είναι ότι απαιτεί τεράστια επεξεργαστική ισχύ:

Διεργασίες	Επιδόσεις	Ιστορικό εφαρμογών	Εκκίνηση	Χρήστες	Λεπτομέρειες	Υπηρεσίες
Όνομα	Κατάσταση	97% CPU	56% Μνήμη	0% Δίσκος	0% Δίκτυο	
Εφαρμογές (5)						
> Google Chrome (14)		2,1%	799,1 MB	0,1 MB/s	0 Mbps	
> MATLAB R2017b (3)		87,9%	976,3 MB	0 MB/s	0 Mbps	

Εικόνα 6.9
Παρουσίαση της ποσοστιαίας κατανάλωσης πόρων του Η/Υ από τον αλγόριθμο στο MATLAB. Αναγνώριση πεζών από αυτόνομο ρομποτικό όχημα με χρήση τεχνητής νοημοσύνης και τεχνικών βαθιάς μάθησης

Όπως βλέπουμε στο task manager του Η/Υ, η μέθοδος απαιτούσε την χρήση σχεδόν του 90% της επεξεργαστικής ισχύς του.

Άλλο ένα μειονέκτημα της μεθόδου αυτής, είναι ότι είναι μια αρκετά χρονοβόρα διαδικασία που μπορεί να μας δημιουργήσει σοβαρά προβλήματα σε εφαρμογές όπως π.χ. η αναγνώριση προσώπου σε είσοδο δημοσίου χώρους:



Εικόνα 6.10

Ο συνολικός χρόνος που απαιτήθηκε από το MATLAB για να εκτελέσει τον αλγόριθμο. Ο χρόνος αυτός ήταν 238,608 seconds δηλαδή χρειαστήκαν σχεδόν 4 λεπτά για την ολοκλήρωση όλων των διεργασιών.

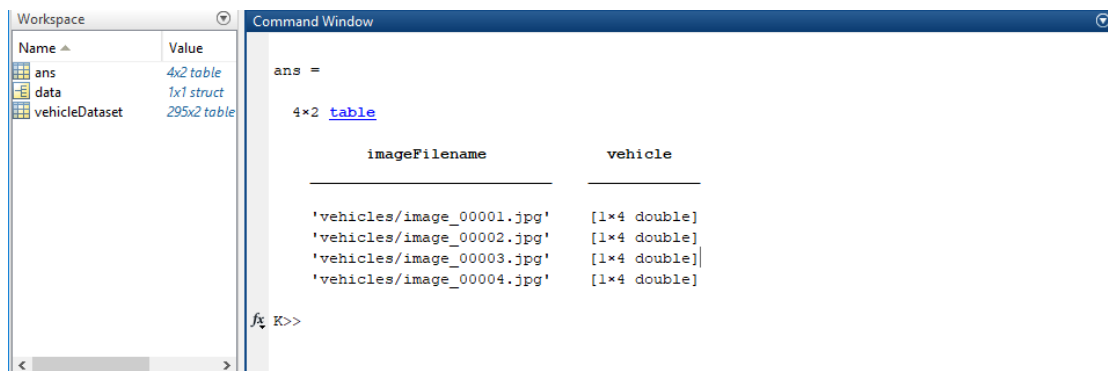
6.3.3 Παράδειγμα 2: αναγνώριση αντικειμένου με εφαρμογή Βαθιάς Μάθησης και Faster R-CNN

Πρώτο βήμα της υλοποίησης, αποτελεί η διασύνδεση του περιβάλλοντος με το dataset, καθώς και η προβολή της πρώτης γραμμής του dataset:

```
%----Neural Networks
clear all
close all
clc
%-----
%% Load Dataset
% kanoume loading to vehicle data set
data = load('fasterRCNNVehicleTrainingData.mat');
vehicleDataset = data.vehicleTrainingData;

%%
%ta dedomena apothikevonte se ena pinaka (h proth stili einai to pach tu
%image file oi ipolipes stiles einai ta ROI labels ton oximatwn
vehicleDataset(1:4,:) % di3e mou thn proth grammh tu dataset
```

Και η προβολή της πρώτης γραμμής του dataset, οδηγεί στην ακόλουθη εικόνα:



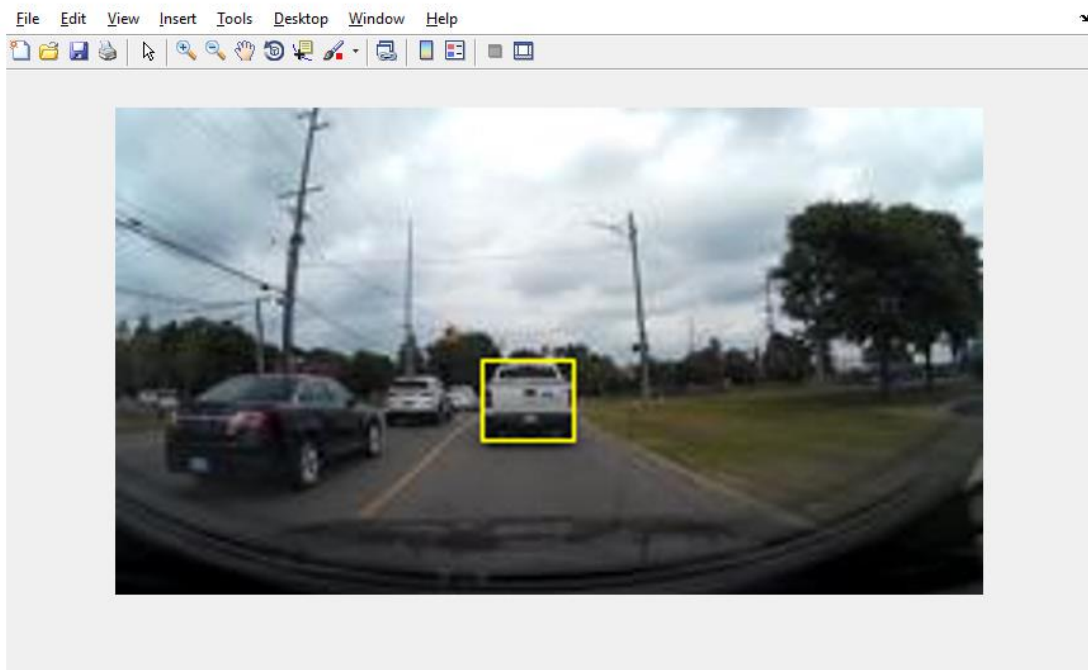
Εικόνα 6.11

Τα στοιχεία της πρώτης γραμμής του dataset.

Με εκτέλεση του παρακάτω τμήματος, προβάλλεται μια εικόνα από το dataset:

```
% Display one of the images from the data set
dataDir = fullfile(toolboxdir('vision'),'visiondata');
vehicleDataset.imageFilename = fullfile(dataDir, vehicleDataset.imageFilename);
% Read one of the images.
I = imread(vehicleDataset.imageFilename{10});
% Insert the ROI labels.
I = insertShape(I, 'Rectangle', vehicleDataset.vehicle{10});
% Resize and display image.
I = imresize(I, 3);
figure
imshow(I)
```

Η εικόνα που θα προβληθεί από το dataset, είναι η ακόλουθη:



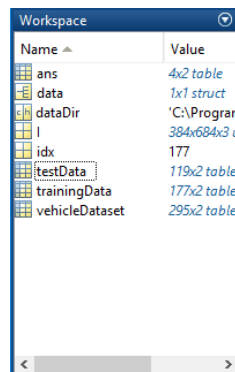
Εικόνα 6.12

Η εικόνα όπου προβάλλεται από το dataset μετά από εκτέλεση του αντίστοιχου τμήματος κώδικα.

Εν συνεχεία θα ακολουθήσει ο διαχωρισμός του σετ δεδομένων σε 2 μέρη. Το πρώτο μέρος θα αποτελέσει το 60% προς εκπαίδευση και αυτό που απομένει θα χρησιμοποιηθεί για τον λειτουργικό έλεγχο του δικτύου:

```
% Split the data set into a training set for training the detector, and a
% test set for evaluating the detector. Select 60% of the data for
% training. Use the rest for evaluation.
idx = floor(0.6 * height(vehicleDataset));
trainingData = vehicleDataset(1:idx,:);
testData = vehicleDataset(idx:end,:);
```

Πράγματι, στο workspace διακρίνεται πως τα δεδομένα έχουν χωριστεί σε 2 κατηγορίες δεδομένων, την trainData και την testData:



Εικόνα 6.13

Το workspace του MATLAB, μετά από τον διαχωρισμό των δεδομένων σε δύο μέρη.

Επομένως, είναι πλέον δυνατή η δόμηση του Νευρωνικού Δικτύου το οποίο θα είναι multilayer, και τα layers του θα είναι τα εξής:

Input layers

- **image input layers**

Middle layers

- **convolution2dLayer**
- **reluLayer**
- **maxpooling2dLayer**

Final layers

- **fully connected layer**
- **reluLayer**
- **softmaxLayer**
- **classificationLayer**

Η δόμηση του δικτύου ξεκινάει από το image input layer και προς τα κάτω

```
%% Create a Convolutional Neural Network (CNN)
% A CNN is the basis of the Faster R-CNN object detector. Create the CNN
% layer by layer using Neural Network Toolbox(TM) functionality.
% Create image input layer.
inputLayer = imageInputLayer([32 32 3]);

%%
% Next, define the middle layers of the network. The middle layers are made
% up of repeated blocks of convolutional, ReLU (rectified linear units),
% and pooling layers. These layers form the core building blocks of
% convolutional neural networks.

% Define the convolutional layer parameters.
filterSize = [3 3];
numFilters = 32;

% Create the middle layers.
middleLayers = [

    convolution2dLayer(filterSize, numFilters, 'Padding', 1)
    reluLayer()
    convolution2dLayer(filterSize, numFilters, 'Padding', 1)
    reluLayer()
    maxPooling2dLayer(3, 'Stride',2)

];
```



```

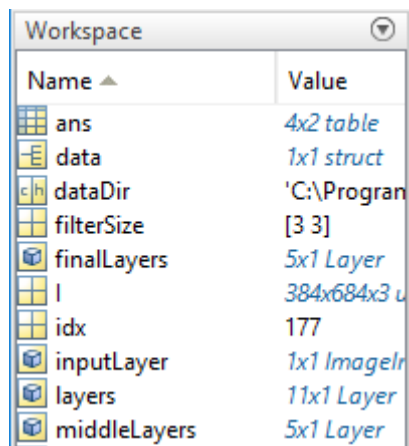
% The final layers of a CNN are typically composed of fully connected
% layers and a softmax loss layer.
finalLayers = [

    fullyConnectedLayer(64)
    % Add a ReLU non-linearity.
    reluLayer()
    % Add the last fully connected layer.
    fullyConnectedLayer(width(vehicleDataset))
    % Add the softmax loss layer and classification layer.
    softmaxLayer()
    classificationLayer()
];

% Combine the input, middle, and final layers.
layers = [
    inputLayer
    middleLayers
    finalLayers
]

```

Τα layers είναι πλέον έτοιμα και εμφανίζονται στο workspace:



Name	Value
ans	4x2 table
data	1x1 struct
dataDir	'C:\Program
filterSize	[3 3]
finalLayers	5x1 Layer
I	384x684x3 u
idx	177
inputLayer	1x1 ImageLr
layers	11x1 Layer
middleLayers	5x1 Layer

Εικόνα 6.14

Το workspace του MATLAB, μετά την προσθήκη των layers.

Στην συνέχεια, ακολουθεί η ρύθμιση της εκπαίδευσης του δικτύου σε 4 βήματα:

```
%% Configure Training Options
% |trainFasterRCNNObjectDetector| trains the detector in four steps.
% Options for step 1.
optionsStage1 = trainingOptions('sgdm', ...
    'MaxEpochs', 10, ...
    'InitialLearnRate', 1e-5, ...
    'CheckpointPath', tempdir);

% Options for step 2.
optionsStage2 = trainingOptions('sgdm', ...
    'MaxEpochs', 10, ...
    'InitialLearnRate', 1e-5, ...
    'CheckpointPath', tempdir);

% Options for step 3.
optionsStage3 = trainingOptions('sgdm', ...
    'MaxEpochs', 10, ...
    'InitialLearnRate', 1e-6, ...
    'CheckpointPath', tempdir);

% Options for step 4.
optionsStage4 = trainingOptions('sgdm', ...
    'MaxEpochs', 10, ...
    'InitialLearnRate', 1e-6, ...
    'CheckpointPath', tempdir);

options = [
    optionsStage1
    optionsStage2
    optionsStage3
    optionsStage4
];
```

Επομένως ακολουθεί η εκπαίδευση του Νευρωνικού Δικτύου:

```
%% Train Faster R-CNN
doTrainingAndEval = false;
if doTrainingAndEval
    % Set random seed to ensure example training reproducibility.
    rng(0);

    % Train Faster R-CNN detector. Select a BoxPyramidScale of 1.2 to allow
    % for finer resolution for multiscale object detection.
    detector = trainFasterRCNNObjectDetector(trainingData, layers, options, ...
        'NegativeOverlapRange', [0 0.3], ...
        'PositiveOverlapRange', [0.6 1], ...
        'BoxPyramidScale', 1.2);
else
    % Load pretrained detector for the example.
    detector = data.detector;
end
```

Σε επόμενο στάδιο, ακολουθεί ο έλεγχος της ποιότητας του δικτύου:

```
% To quickly verify the training, run the detector on a test image.
% Read a test image.
I = imread('highway.png');
% Run the detector.
[bboxes, scores] = detect(detector, I);
% Annotate detections in the image.
I = insertObjectAnnotation(I, 'rectangle', bboxes, scores);
figure
imshow(I)

%% Evaluate Detector Using Test Set
if doTrainingAndEval
    % Run detector on each image in the test set and collect results.
    resultsStruct = struct([]);
    for i = 1:height(testData)

        % Read the image.
        I = imread(testData.imageFilename(i));

        % Run the detector.
        [bboxes, scores, labels] = detect(detector, I);

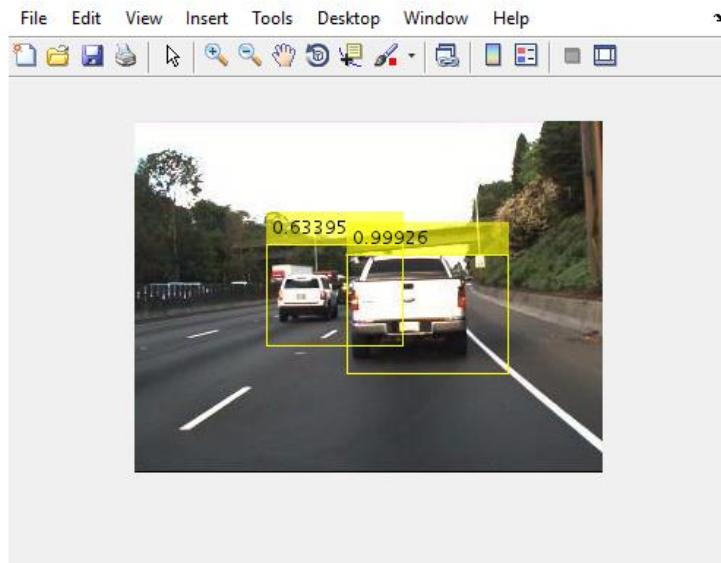
        % Collect the results.
        resultsStruct(i).Boxes = bboxes;
        resultsStruct(i).Scores = scores;
        resultsStruct(i).Labels = labels;
    end

    % Convert the results into a table.
    results = struct2table(resultsStruct);
else
    % Load results from disk.
    results = data.results;
end

% Extract expected bounding box locations from test data.
expectedResults = testData(:, 2:end);

% Evaluate the object detector using Average Precision metric.
[ap, recall, precision] = evaluateDetectionPrecision(results, expectedResults);
```

Στην ακόλουθη εικόνα, διακρίνεται πως ο αλγόριθμος κατάφερε να αναγνωρίσει επιτυχώς τα πρότυπα (αυτοκίνητα):



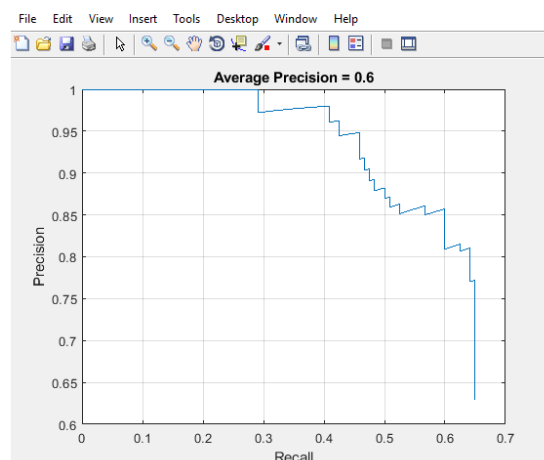
Εικόνα 6.15

Η αναγνώριση ενός αυτοκινήτου από τον αλγόριθμο.

Τέλος, προχωρούμε στην υλοποίηση ενός κώδικα ο οποίος θα δημιουργήσει μια καμπύλη ακριβείας / ανάκλησης:

```
% The precision/recall (PR) curve
% Plot precision/recall curve
figure
plot(recall, precision)
xlabel('Recall')
ylabel('Precision')
grid on
title(sprintf('Average Precision = %.1f', ap))
```

Το αποτέλεσμα της εκτέλεσης του παραπάνω τμήματος κώδικα, είναι το ακόλουθο:



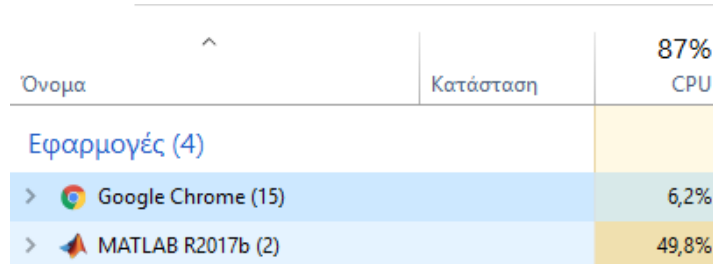
Εικόνα 6.16

Η καμπύλη ακριβείας / ανάκλησης.

Αναγνώριση πεζών από αυτόνομο ρομποτικό όχημα με χρήση τεχνητής νοημοσύνης και τεχνικών βαθιάς μάθησης

Όπως προκύπτει από το παραπάνω γράφημα, η ακρίβεια βρίσκεται στο 100% μέχρι το 0,3 recall και μετά μειώνεται, ενώ ο μέσος όρος της ακρίβειας βρίσκεται στο 60%.

Συμπέρασμα: ο αλγόριθμος εκπαιδεύτηκε και έκανε αναγνώριση προτύπων αλλά όπως και στο προηγούμενο παράδειγμα έτσι και εδώ απαιτούσε σχετικά μεγάλη επεξεργαστική ισχύ (όπως φαίνεται και στο task manager) ενώ η διεργασία (σύμφωνα με το timer του MATLAB) διήρκησε 12 seconds.



Όνομα	Κατάσταση	87% CPU
Εφαρμογές (4)		
> Google Chrome (15)		6,2%
> MATLAB R2017b (2)		49,8%

Εικόνα 6.17

Παρουσίαση της ποσοστιαίας κατανάλωσης πόρων του Η/Υ από τον αλγόριθμο στο MATLAB.

7 Εφαρμογή Τεχνικών Βαθιάς Μάθησης για τον Εντοπισμό Πεζών

Σε αυτό το κεφάλαιο χρησιμοποιούνται τα αυτοματοποιημένα Βαθιά Συνελκτικά Νευρωνικά Δίκτυα που αναλύθηκαν στα δύο προηγούμενα κεφάλαια, και ιδιαίτερα ο αλγόριθμος Faster R-CNN, για τη διαδικασία εξαγωγής χαρακτηριστικών από εικόνες (video) κλειστού κυκλώματος τηλεόρασης, που αφορούν την αναγνώριση πεζών σε δρόμους αστικού κέντρου. Όπως γίνεται αντιληπτό, μείζονος σημασίας ζήτημα για ένα αυτόνομο όχημα δεν είναι μόνο η επίτευξη αναγνώρισης του πεζού αλλά, εξίσου μεγάλη σημασία έχει η επίτευξη πρόβλεψης της κίνησης του πεζού. Σε αυτό το σκοπό αποσκοπεί αυτή η μερική προσπάθεια που επιχειρείται εδώ. Τα εξαγόμενα χαρακτηριστικά χρησιμεύουν ως ισχυρή βάση για μια ποικιλία εργασιών αναγνώρισης αντικειμένων και όχι μόνο για αναγνώριση πεζών. Η προσέγγιση που ακολουθείται εδώ είναι να ταιριάζονται τα εξαγόμενα χαρακτηριστικά μεμονωμένων ανιχνεύσεων από πλαίσια (frames) video στα χαρακτηριστικά και τις ανιχνεύσεις επόμενων πλαισίων, δημιουργώντας έτσι μια αντιστοιχία ανιχνεύσεων σε πολλαπλά πλαίσια. Η προσέγγιση αυτή φαίνεται να είναι σε θέση να αντιμετωπίσει προκλήσεις όπως το γεμάτο σκηνές πλάνο, την αλλαγή φωτισμού, τις σκιές, την αλλαγή στις εμφανίσεις πεζών ακόμη και εμφανίσεις πεζών που ένα μικρό έστω τμήμα τους κρύβεται από το πλάνο.

Ωστόσο, δεν είναι εφικτή η αναγνώριση όταν σημαντικό τμήμα του πεζού κρύβεται στο πλάνο του πλαισίου. Προκειμένου να υλοποιηθούν οι εργασίες αυτού του κεφαλαίου χρησιμοποιήθηκε video κλειστού κυκλώματος τηλεόρασης, το οποίο στη συνέχεια με κατάλληλη εφαρμογή μετατροπής video ανι σε εικόνες jpeg (η εφαρμογή που επιλέχθηκε είναι η *Free Video to JPG Converter*), μετατράπηκε σε frames (πλαίσια). Ενδεικτικά αξιοποιήθηκαν 250 πλαίσια/καρέ από αυτό το video και δόθηκε μεγαλύτερη βαρύτητα στον κώδικα MATLAB για την ανίχνευση της συμπεριφοράς του δικτύου στην παρακολούθηση των πεζών σε αυτά τα πλάνο/καρέ που εξήχθησαν από το video. Τα αποτελέσματα είναι μεν ενθαρρυντικά αλλά απαιτείται περαιτέρω έρευνα και εργασία, τόσο στον αλγόριθμο, όσο και στην εκπαίδευση του δικτύου. Η εφαρμογή που δημιουργήθηκε για την παρακολούθηση της κίνησης των πεζών σε MATLAB, τροποποιώντας μια πιο απλή στη βάση του *Piotr Dollar toolbox* στο MATLAB, όπου χρησιμοποιήθηκαν από αυτό το toolbox και προ-εκπαιδευμένα δίκτυα CNN, εκτελείται απρόσκοπτα σε μεγάλη κατηγορία video κλειστού κυκλώματος τηλεόρασης που δοκιμάστηκε.

Προτού παρουσιαστεί η ανάλυση της μεθοδολογίας που ακολουθήθηκε, η οποία βασίζεται στις εργασίες [139] και [146] και στο *Piotr Dollar MATLAB toolbox* [140], θα εξεταστεί συνοπτικά το ζήτημα του εντοπισμού και παρακολούθησης πεζών και οι δυσκολίες του. Η αυτοματοποιημένη παρακολούθηση των πεζών έχει αποκτήσει σημαντικό ενδιαφέρον τις τελευταίες δύο δεκαετίες. Το αυξανόμενο αυτό ενδιαφέρον οφείλεται στη διαθεσιμότητα υψηλής ποιότητας οικονομικών καμερών κλειστού κυκλώματος τηλεόρασης και την ανάγκη για αυτοματοποιημένη ανάλυση των αντίστοιχων βίντεο.

Η αναγνώριση των ανθρώπινων ενεργειών σε περιβάλλοντα πραγματικού κόσμου βρίσκει εφαρμογή σε έξυπνη παρακολούθηση βίντεο, ανάλυσης συμπεριφοράς πελατών (Chen et al., 2016), εθνικής ασφάλειας, πρόληψης εγκληματικότητας, παρακολούθησης συμπεριφοράς ασθενών σε νοσοκομεία (ιδιαίτερα αυτών σε κρίσιμη κατάσταση), παρακολούθησης των ηλικιωμένων αλλά και χώρων και ατόμων ασχολούμενων με την παιδική φροντίδα (Wang, 2013) και μπορεί προφανώς να χρησιμοποιηθεί επίσης για τη διαχείριση δημόσιων χώρων και αντιμετώπιση καταστάσεων έκτακτης ανάγκης.

Υπάρχει μια πλούσια βιβλιογραφία (Yilmaz et al., 2006, Smeulders et al., 2014) για την αντιμετώπιση του προβλήματος που ακολουθεί το συμβατικό παράδειγμα της αναγνώρισης προτύπων, περιλαμβάνοντας την εξαγωγή χειροποίητων χαρακτηριστικών (προκαθορισμένα χαρακτηριστικά όπως το ιστόγραμμα του (HOG)) από τις εικόνες για την ανίχνευση πεζών σε μια σκηνή και την επακόλουθη ταξινόμησή τους, χρησιμοποιώντας ταξινομητές. Το μειονέκτημα της χρήσης τέτοιων χειροποίητων χαρακτηριστικών για μια εργασία παρακολούθησης πεζών είναι η περιορισμένη ικανότητα των χειροποίητων αυτών χαρακτηριστικών για να προσαρμοστούν στις παραλλαγές της εμφάνισης αντικειμένων που είναι πολύπλοκες και εξαιρετικά μη γραμμικές (Yilmaz et al., 2006, Chen κ.ά., 2016).

Επιπλέον, για να επιτευχθεί ακριβής αναγνώριση και παρακολούθηση πεζών υπάρχουν μεγάλες προκλήσεις που πρέπει να αντιμετωπιστούν και οι οποίες περιλαμβάνουν την μερική εμφάνιση και μόνο των πεζών, συσσωματωμένα υπόβαθρα, παραλλαγές όψεων, αλλαγές σε εμφάνιση (κλίμακα, στάση και σχήμα), παρόμοιους πεζόδρομους ως υπόβαθρο, παραλλαγές φωτισμού αλλά και απρόβλεπτες κινήσεις πεζών (Ji et al., 2013, Chen κ.ά., 2016).

Τα περισσότερα από τα state-of-the-art εργαλεία τέτοιας παρακολούθησης δεν έχουν επαρκώς τις απαιτούμενες ικανότητες γενίκευσης (Feris et al., 2013). Ο επαναπροσδιορισμός των πεζών (σε προβολές μίας κάμερας) παραμένει μια ανοιχτή πρόκληση. Παρόλα αυτά το πρόβλημα είναι ενδιαφέρον και η χρήση των CNN έχει όλες τις δυνατότητες για να βελτιώσει σημαντικά τις αποδόσεις.

Σε αυτό το κεφάλαιο, οι πεζοί ανιχνεύονται σε κάθε πλαίσιο εικόνων κλειστού κυκλώματος τηλεόρασης χρησιμοποιώντας ένα state-of-the-art πλαίσιο ανίχνευσης αντικειμένων, δηλ. τα πιο γρήγορα R-CNN (Faster R-CNN), (Ren et al., 2015). Στη συνέχεια, για να ξεπεραστούν οι περιορισμοί της χρήσης των χειροποίητων (προκαθορισμένων) χαρακτηριστικών (HOG descriptors), οι οποίοι χρησιμοποιούνται και στη παρούσα εργασία, ενισχύεται η αυτόματη εξαγωγή χαρακτηριστικών για τους ανιχνευμένους πεζούς μέσω της χρήσης των Βαθιά Συνελικτικών Νευρωνικών Δικτύων (CNNs). Οι Donahue et al. (2014) διατύπωσαν πως οι ενεργοποιήσεις των νευρώνων στα τελευταία στρώματα δικτύων Βαθιάς Μάθησης, είναι σε θέση να εξαγάγουν ισχυρά χαρακτηριστικά για μια ποικιλία εργασιών αναγνώρισης αντικειμένων.

Βασική υπόθεση της παρούσας εργασίας, είναι ότι οι εξαγόμενες ενεργοποιήσεις από τα τελευταία στρώματα ενός βαθιού CNN, δύναται να χρησιμοποιηθούν για να διακρίνουν ανιχνευμένους πεζούς μέσα από διαφορετικά frames (πλαίσια), και μπορεί αυτό να χρησιμοποιηθεί για την ακριβή αντιμετώπιση ενός προβλήματος ανίχνευσης με εντοπισμό. Έτσι, με ένα νέο τρόπο οι τυπικές λειτουργίες της αναγνώρισης χρησιμοποιούνται για την αντιμετώπιση ενός προβλήματος παρακολούθησης. Η παρακολούθηση διατυπώνεται ως η αντιστοιχία των ανιχνεύσεων σε πολλαπλά πλαίσια και επιτυγχάνεται με την αντιστοιχία των εξαγόμενων χαρακτηριστικών των μεμονωμένων ανιχνεύσεων στα επόμενα πλαίσια. Βέβαια πρέπει να τονιστεί ότι τα αποτελέσματα είναι απλά ενθαρρυντικά.

Οι βασικές συνεισφορές του παρόντος κεφαλαίου και της εργασίας εν γένει, είναι:

- Δημιουργήθηκε μέσα από μετατροπές και προσθήκες στο Piotr Dollar Toolbox ένα πλαίσιο για την ανίχνευση και παρακολούθηση πεζών σε πραγματικό χρόνο χρησιμοποιώντας CNN/RCNN που είτε εκπαιδεύονται από τον χρήστη, είτε είναι ήδη εκπαιδευμένα (προεκπαιδευμένα). Τα προεκπαιδευμένα νευρωνικά που χρησιμοποιήθηκαν στη παρούσα εργασία, είναι και αυτά που παρουσίασαν τις βέλτιστες αποδόσεις.
- Μέσω μικρών τροποποιήσεων ήδη υπάρχοντος αλγορίθμου, γίνεται προσπάθεια για να καθοριστεί η αντιστοιχία μεταξύ των ανιχνεύσεων στην ακολουθία των πλαισίων.

Το πλαίσιο αντιμετώπισης του προβλήματος που δοκιμάστηκε, αντιμετωπίζει προκλήσεις όπως η μερική απόκρυψη πεζών στα πλάνα, οι διακυμάνσεις στον φωτισμό, οι αλλαγές στη στάση, το σχήμα, την κλίμακα αλλά και την κίνηση των πεζών, το γεμάτο φόντο και οι συνολικές αποκρύψεις τους για σύντομες περιόδους. Το πλαίσιο αυτό δεν είναι σε θέση να χειριστεί ολικές αποκρύψεις φιγούρας πεζών μακρών περιόδων, ενώ αποτυγχάνει να αντιμετωπίσει το πρόβλημα της εμφάνισης παρόμοιας εμφάνισης στο ίδιο πλαίσιο. [141], [142], [143], [144], [145], [146], [147], [150], [151]

7.1 Σχετικές Εργασίες

Η παρακολούθηση αντικειμένων ορίζεται ως η δημιουργία τροχιάς ενός αντικειμένου σε ένα επίπεδο εικόνας, και ένας αλγόριθμος εντοπισμού ο οποίος τοποθετεί σωστά ετικέτες στα αντικείμενα που παρακολουθούνται σε διαφορετικά πλαίσια ενός βίντεο. Υπάρχουν τρεις θεμελιώδεις πτυχές ενός συστήματος παρακολούθησης πεζού που είναι ανάλογες με την παρακολούθηση αντικειμένων:

- 1) ανίχνευση του πεζού στο πλαίσιο βίντεο,
- 2) παρακολούθηση της ανίχνευσης, και
- 3) ανάλυση των διαδρομών του συγκεκριμένου πεζού (Yilmaz et al., 2006).

Στη βιβλιογραφία, έχουν προταθεί διάφοροι αλγόριθμοι για ανίχνευση και παρακολούθηση αντικειμένων, για μεθοδολογίες αφαίρεσης υποβάθρου, για τμηματοποίηση (segmentation) και επίβλεψη μάθησης. Για την ακριβή παρακολούθηση, η επιλογή κατάλληλων χαρακτηριστικών παίζει μείζονα ρόλο και σχετίζεται με την *αντιπροσώπηση μέσω κεντροειδών*. Ακολουθώντας, εκτελείται η αποστολή δημιουργίας αντιστοιχίας των ανιχνεύσεων. Αυτό έχει γίνει στο παρελθόν χρησιμοποιώντας ντετερμινιστικά ή πιθανοτικά μοντέλα κίνησης και πυρήνα, βασισμένο στην εμφάνιση μοντέλων παρακολούθησης. Επιπλέον, έχουν προταθεί μέθοδοι προσαρμογής on-line για την προσαρμογή των ανιχνευτών στις παραλλαγές της εμφάνισης των αντικειμένων που έχουν εντοπιστεί με την πάροδο του χρόνου. Οι ανιχνευτές είναι εκπαιδευμένοι και ενημερωμένοι σε απευθείας σύνδεση κατά τη διάρκεια της παρακολούθησης, ωστόσο αυτές συνήθως απαιτούν μεγάλο αριθμό περιπτώσεων μάθησης, κάτι που μπορεί να μην είναι πάντα διαθέσιμο. (Chen et al., 2016, Feris et al., 2013).

Πρόσφατα, σημειώθηκε σημαντική βελτίωση της απόδοσης στον τομέα της κατηγοριοποίησης κατηγοριών εικόνων και αναγνώριση με την εκπαίδευση ενός βαθιού CNN με εκατομμύρια εικόνες διαφορετικών τάξεων (Krizhevsky et al., 2012). Από αυτή τη συζήτηση προφανώς προκύπτει και η ανάγκη των προεκπαιδευμένων βαθιών Νευρωνικών Δικτύων, και αυτή η δυνατότητα χρησιμοποιείται στο παρόν κεφάλαιο.

Τα Συνελκτικά Νευρωνικά Δίκτυα CNNs (Lecun et al., 1998) είναι μια μέθοδος Μηχανικής Μάθησης, όπως αναφέρθηκε στα κεφάλαια 5 και 6, που εκμεταλλεύεται τις τοπικές και χωρικές πληροφορίες σε μια εικόνα και μαθαίνει μια ιεραρχία όλο και πιο περίπλοκων χαρακτηριστικών, αυτοματοποιώντας έτσι τη διαδικασία κατασκευής χαρακτηριστικών. Τα CNN είναι σχετικά μη ευαίσθητα σε ορισμένες παραλλαγές των εισόδων (Ji et al., 2013). Με γνώμονα την επιτυχία της ταξινόμησης και της αναγνώρισης των εικόνων, έγιναν προσπάθειες εκμετάλλευσης των CNN για την παρακολούθηση εργασιών, Fan et al (2010). Σχεδιάστηκε ένας αλγόριθμος ανίχνευσης CNN με μεταβλητή αρχιτεκτονική, όπου τα χαρακτηριστικά μαθαίνονται κατά τη διάρκεια της εξάσκησης εκτός γραμμής, η οποία αφαιρεί τόσο τις χωρικές όσο και τις χρονικές πληροφορίες, θεωρώντας ζευγάρια εικόνων δύο διαδοχικών πλαισίων αντί για ένα μόνο πλαίσιο.

Ο αλγόριθμος εξάγει τόσο τοπικά όσο και γενικότερα χαρακτηριστικά για την αντιμετώπιση μερικών περιπτώσεων και αλλαγών στις εμφανίσεις των αντικειμένων/προσώπων υπό παρακολούθηση. Οι Ji et al (2013) χρησιμοποιούν ένα μοντέλο 3D CNN για αναγνώριση πεζών. Το μοντέλο χαρακτηρίζεται τόσο από τη χωρική όσο και από τις χρονικές διαστάσεις με την εκτέλεση 3D συνελίξεων και συλλαμβάνει πληροφορίες κίνησης σε πολλαπλά πλαίσια. Στην Jin et al. (2013) εισάγεται Βαθύ CNN για το έργο της παρακολούθησης, το οποίο εξάγει χαρακτηριστικά και μετατρέπει τις εικόνες σε διανύσματα υψηλής διαστάσεως.

Ο *χάρτης εμπιστοσύνης* δημιουργείται με τον υπολογισμό των ομοιοτήτων των δύο αντιστοιχιών χρησιμοποιώντας την λειτουργία και τις επιδόσεις δικτύων ακτινικής βάσης. Οι Hong et al. (2015) προτείνουν τη χρήση εξόδων από το τελευταίο στρώμα ενός προ-εκπαιδευμένου CNN για να μάθουν τη διάκριση μοντέλων που χρησιμοποιούν μηχανή διανυσμάτων υποστήριξης on-line (SVM). Στη συνέχεια, η παρακολούθηση πραγματοποιείται χρησιμοποιώντας διαδοχικά Bayesian φίλτρα με ένα ειδικό χάρτη στόχου, το οποίο υπολογίζεται από την οπισθόδρομη προβολή των εξόδων από το τελευταίο στρώμα. Οι Wang et al. (2015) χρησιμοποιούν τα χαρακτηριστικά που έχουν μάθει από προ-εκπαιδευμένο CNN για την on-line παρακολούθηση.

Το CNN που προτείνουν ρυθμίζεται με ακρίβεια κατά την on-line παρακολούθηση για να καταγράψει την εμφάνιση ενός αντικειμένου που καθορίζεται στο πρώτο πλαίσιο της ακολουθίας frames και παράγεται ένας χάρτης πιθανοτήτων αντί να παράγονται απλές ετικέτες κλάσης. Οι Wang και Yeung (2013) διερευνούν την εκπαίδευση ενός ολισθαίνοντα αυτόματου κωδικοποιητή για την μεταφορά γνώσεων από την εκπαίδευση off-line σε on-line παρακολούθηση για την προσαρμογή των αλλαγών εμφάνισης ενός κινούμενου στόχου.

Οι Nam και Han (2015) προτείνουν την παρακολούθηση μέσω αλγόριθμου που μαθαίνει ανεξάρτητες αναπαραστάσεις τομέων των πλάνων κατά την προεκπαίδευση και καταγράφει συγκεκριμένες πληροφορίες για τους συγκεκριμένους τομείς, μέσω της μάθησης on-line κατά τη διάρκεια της παρακολούθησης. Το Δίκτυο έχει μια απλή αρχιτεκτονική σε σχέση με αυτό που είναι σχεδιασμένο για εργασίες ταξινόμησης εικόνων. Το σύνολο του δικτύου είναι προ-εκπαιδευμένο off-line, και αργότερα πλήρως συνδεδεμένο για online λειτουργία.

Οι Li et al. (2016) προτείνουν ένα νέο αλγόριθμο παρακολούθησης που χρησιμοποιεί το CNN για να μάθει αυτόματα την πιο χρήσιμη αναπαράσταση χαρακτηριστικών ενός συγκεκριμένου αντικειμένου στόχου. Μια στρατηγική εντοπισμού ακολουθίας αναπτύσσεται για να διακρίνει το αντικείμενο στόχο από το φόντο του.

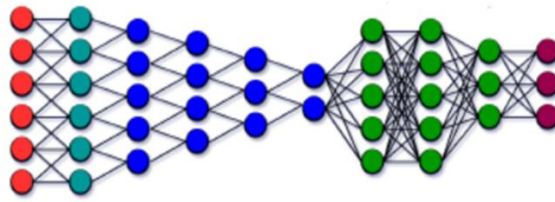
Το CNN δημιουργεί από όλες τις πιθανές υποθέσεις θέσεων αντικειμένων ένα χάρτη για το πλαίσιο. Ο αλγόριθμος παρακολούθησης μαθαίνει τα ληφθέντα δείγματα από την τρέχουσα ακολουθία εικόνων. Οι Chen et al. (2016) εκπαιδεύουν ένα βαθύ CNN για να μεταφέρει τις μαθησιακές παραμέτρους για την εντολή παρακολούθησης και την κατασκευή ενός μοντέλου εμφάνισης αντικειμένων.

Η αρχική και η on-line εκπαίδευση χρησιμοποιούνται για την ενημέρωση του μοντέλου εμφάνισης. Παρά την επιτυχία των CNNs, υπάρχει μόνο ένας περιορισμένος αριθμός αλγορίθμων παρακολούθησης (όπως συζητείται παραπάνω) που εκμεταλλεύονται τις μεγάλες δυνατότητες των CNNs, που προτείνονται μέχρι τώρα στη βιβλιογραφία. Επιπλέον, τα προηγούμενα ερευνητικά έργα δεν έχουν ενσωματώσει το μοντέλο προσέγγισης της ανίχνευσης και παρακολούθησης ταυτόχρονα στα ίδια CNN.

Η μελέτη της αναφερθείσας βιβλιογραφίας, οδήγησε στο συμπέρασμα πως το πρόβλημα είναι πολύ ενδιαφέρον, οι λύσεις που θα προταθούν εν συνεχεία πολύ σημαντικές, και γι' αυτό στο παρόν κεφάλαιο πραγματοποιήθηκε η προσπάθεια δοκιμής της σημαντικότητας των CNN και ιδιαίτερα των προ-εκπαιδευμένων CNN, τα οποία έχουν προ-εκπαιδευτεί μέσα από πολύ μεγάλα datasets, σε νέα κλειστού κυκλώματος τηλεόρασης video frames παρακολούθησης πεζών.

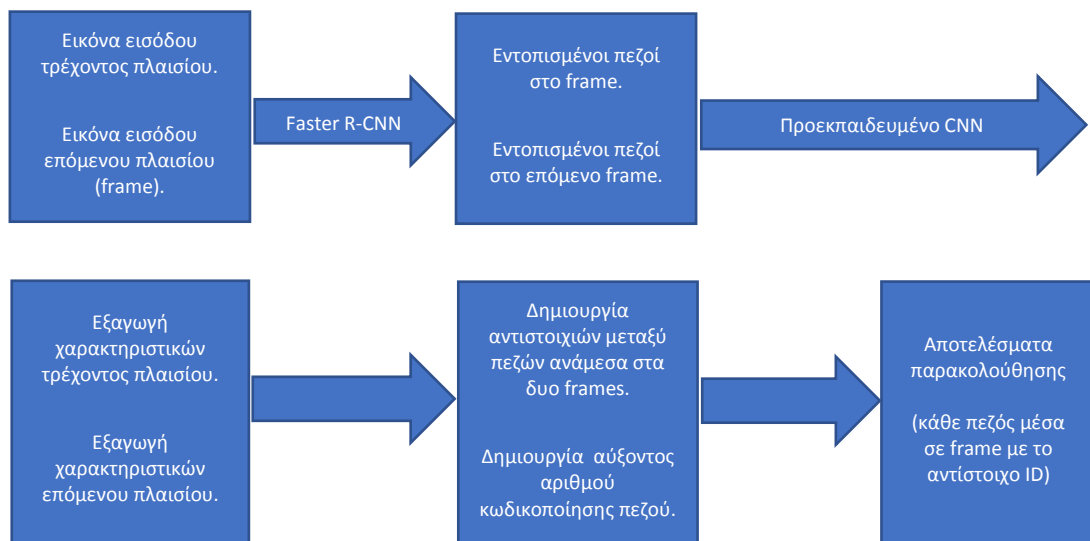
[142], [143], [144], [148], [149], [150], [151], [152], [153], [154]

7.2 Μεθοδολογία



Εικόνα 7.1

Μια απλοποιημένη αρχιτεκτονική του βαθιού (deep) CNN. Κόκκινοι είναι οι νευρώνες εισόδου, γαλάζιοι οι νευρώνες πυρήνα, μπλε οι Συνελκτικοί νευρώνες, πράσινοι όλοι οι πλήρως συνδεδεμένοι νευρώνες και μπορντό οι νευρώνες εξόδου. Οι νευρώνες του προτελευταίου επιπέδου είναι αυτοί που παράγουν τα ανώτερα χαρακτηριστικά.



Εικόνα 7.2

Μια απλοποιημένη αρχιτεκτονική του συστήματος που δοκιμάστηκε στο MATLAB για την παρακολούθηση

Η προσέγγιση που αναπτύσσεται εδώ, χρησιμοποιεί CNN τόσο για τον εντοπισμό πεζών μέσα στα πλαίσια, όσο και για την παρακολούθηση των πεζών εντός των επόμενων πλαισίων. Χρησιμοποιείται ένα προσφάτως Βαθύ προ-εκπαιδευμένο CNN για την ανίχνευση αντικειμένων, το Faster R-CNN (Ren et al., 2015), το οποίο όπως αργότερα παρουσιάζεται, μπορεί με σχετική επιτυχία να εντοπίζει και να παρακολουθεί πεζούς. Τα χαρακτηριστικά που χρησιμοποιούνται για την παρακολούθηση, προέρχονται από αυτό το προ-εκπαιδευμένο CNN (Εικόνα 7.1) και χρησιμεύουν ως μια ισχυρή βάση για την αναγνώριση αντικειμένων γενικότερα.

Ο προτεινόμενος αλγόριθμος, τροποποίηση των [139], [146] και του *Piotr Dollar MATLAB Toolbox* [140], για τη δημιουργία παρακολούθησης πεζών σε ακολουθία frames είναι πιο κοντά στο πλαίσιο παρακολούθησης πυρήνα μέσω κεντροειδών, αλλά μια ισχυρή αναπαράσταση αναπτύσσεται με την επιβολή βαρών για εμφάνιση και εξαγωγή χωρικών πληροφοριών. Μία απλοποιημένη διάταξη του αναπτυσσόμενου πλαισίου παρέχεται στην Εικόνα 7.2. Τα πλαίσια εικόνων κλειστού κυκλώματος τηλεόρασης εισάγονται στον αλγόριθμο που ανιχνεύει και εντοπίζει τους μεμονωμένους πεζούς. Χαρακτηριστικά από τις περικομμένες εικόνες των πεζών εξάγονται από το προ-εκπαιδευμένο CNN. Ο αλγόριθμος χρησιμοποιείται για την πραγματοποίηση αντιστοιχιών των εντοπισμών πεζών σε όλα τα πλαίσια και οι αύξοντες αριθμοί κωδικοποίησης των πεζών (ID) εμφανίζονται στους ξεχωριστούς εντοπισμούς. Τα αποτελέσματα παρακολούθησης εμφανίζονται με την επικάλυψη των ID των εντοπισμών στα αντίστοιχα πλαίσια.

Το τελευταίο στρώμα πριν το στρώμα ταξινόμησης του CNN, δημιουργεί ένα διάνυσμα 1000 χαρακτηριστικών για κάθε είσοδο εικόνα (frame). Οι μεμονωμένες ανιχνεύσεις στο κάθε πλαίσιο έχουν τη μορφή ενός πλαισίου οριοθέτησης γύρω από τους πεζούς. Ακολούθως, οι ανιχνεύσεις περικόπτονται και τροφοδοτούνται στο CNN το οποίο δημιουργεί μια μήτρα διανυσμάτων χαρακτηριστικών. Μαθηματικά αυτό μπορεί να αναπαρασταθεί ως,

$$FV_{i(k)} = \begin{bmatrix} A_{(1,1)} & \dots & A_{(1,i(k))} \\ \vdots & \ddots & \vdots \\ A_{(1000,1)} & \dots & A_{(1000,i(k))} \end{bmatrix}$$

Όπου στον πίνακα δηλώνεται η μήτρα των διανυσμάτων χαρακτηριστικών των ανιχνεύσεων i για ένα πλαίσιο k , δηλ. ο πίνακας αυτός είναι το σύνολο $\{A(1, i(k)), \dots, A(1000, i(k))\}$ εντοπισμών πεζών στο πλαίσιο k , και $i(k)$ υποδηλώνει τον αριθμό των εντοπισμών στο πλαίσιο k .

Τα κεντροειδή των εντοπισμών μπορούν να εκφράζονται με την εξίσωση

$$PC_{i(k)} = \begin{bmatrix} x_{(1,k)} & \dots & x_{(i,k)} \\ \vdots & \ddots & \vdots \\ y_{(1,k)} & \dots & y_{(i,k)} \end{bmatrix}$$

Όπου ο ανωτέρω πίνακας PC δηλώνει τη μήτρα των x και y συντεταγμένων των κεντροειδών των εντοπισμών πεζών i στο πλαίσιο k . Η αντιστοίχιση των χαρακτηριστικών των frames δημιουργείται με τον υπολογισμό ενός της απόστασης των χαρακτηριστικών και την απόσταση σε pixels μεταξύ κάθε ζεύγους εντοπισμών σε δύο διαδοχικά πλαίσια. Έστω $FV_i(k)$ και $FV_j(k+1)$ δηλώνουν αντιστοίχως τους φορείς χαρακτηριστικών για τους εντοπισμούς i και j στο πλαίσιο k και το πλαίσιο $k+1$. Η κανονικοποιημένη απόσταση των χαρακτηριστικών μεταξύ των δύο συνόλων εντοπισμών $F_d(i(k), j(k+1))$ εκφράζεται ως

$$F_d(i(k), j(k+1)) = \frac{|FV_{i(k)} - FV_{j(k+1)}|}{|FV_{i(k)}| |FV_{j(k+1)}|}$$

Έστω τα $PC_i(k)$ και $PC_j(k+1)$ υποδηλώνουν τα κεντροειδή για τα σύνολα εντοπισμών i και j στο πλαίσιο k και το πλαίσιο $k+1$ αντίστοιχα. Ομοίως, η κανονικοποιημένη απόσταση pixels μεταξύ των δύο συνόλων εντοπισμών $P_d(i(k), j(k+1))$ εκφράζεται ως

$$P_d(i(k), j(k+1)) = \frac{|PC_{i(k)} - PC_{j(k+1)}|}{|PC_{i(k)}| |PC_{j(k+1)}|}$$

Στη συνέχεια δημιουργείται μια μήτρα απόστασης $F_d(k+1)$ για τα διανύσματα χαρακτηριστικών από τις κανονικοποιημένες αποστάσεις των ζευγών χαρακτηριστικών και αντιπροσωπεύεται λογικά από την σχέση

$$F_d(k+1) = \begin{bmatrix} F_{d(1,1)} & \dots & F_{d(1,j(k+1))} \\ \vdots & \ddots & \vdots \\ F_{d(i(k),1)} & \dots & F_{d(i(k),j(k+1))} \end{bmatrix}$$

Αντίστοιχα δημιουργείται μια μήτρα για τις αποστάσεις των ζευγών των εικονοστοιχείων $P_d(k+1)$ από τις κανονικοποιημένες αποστάσεις των ζευγών pixels και αντιπροσωπεύεται από την εξίσωση

$$P_d(k+1) = \begin{bmatrix} P_{d(1,1)} & \dots & P_{d(1,j(k+1))} \\ \vdots & \ddots & \vdots \\ P_{d(i(k),1)} & \dots & P_{d(i(k),j(k+1))} \end{bmatrix}$$

Οι πίνακες $F_d(k+1)$ και $P_d(k+1)$ συνδυάζονται χρησιμοποιώντας ένα βάρος w ($0 \leq w \leq 1$). Το αποτέλεσμα συνδυασμού ονομάζεται πίνακας παρακολούθησης $T_d(k+1)$ και ορίζεται από την ακόλουθη σχέση.

$$T_{d(k+1)} = (w)P_{d(k+1)} + (1-w)E_{d(k+1)} = \begin{bmatrix} t_{(1,1)} & \dots & t_{(1,j(k+1))} \\ \vdots & \ddots & \vdots \\ t_{(i(k),1)} & \dots & t_{(i(k),j(k+1))} \end{bmatrix}$$

Στο πρώτο πλαίσιο, τα αναγνωριστικά ID των πεζών παράγονται τυχαία και παρακολουθούνται στα επόμενα πλαίσια. Ο αριθμός των παραγόμενων ID στο πρώτο πλαίσιο είναι ίσα με τον αριθμό των εντοπισμών πεζών όπως είναι φυσιολογικό. Για τις ανιχνεύσεις πεζών στα επόμενα πλαίσια, είτε το αναγνωριστικό ID εντοπίζεται από το προηγούμενο πλαίσιο (το οποίο περιλαμβάνει την αντιστοίχιση βάσει κριτηρίων ελάχιστης απόστασης) ή δημιουργείται ένα νέο αναγνωριστικό ID (στην προφανή περίπτωση ενός νέου προσώπου που εισέρχεται στο πλαίσιο).

Έστω το σύνολο $\{t(1, i(k)), \dots, t(i(k), j(k+1))\}$ το σύνολο διανυσμάτων παρακολούθησης στο πλαίσιο k . Η ελάχιστη τιμή του συνόλου $\{t(1, i(k)), \dots, t(i(k), j(k+1))\}$ χρησιμοποιείται για να κάνει αντιστοίχιση της $j(k+1)$ ανίχνευσης στο πλαίσιο $k+1$, στην $1n, \dots, i(k)$ ανίχνευση στο πλαίσιο k , μόνο εάν αυτή η ελάχιστη τιμή είναι κάτω από ένα όριο. Δηλ. επί της ουσίας κάθε ζευγάρι του πίνακα T_d μπαίνει στη διαδικασία προσδιορισμού αντιστοιχιών για την περίπτωση αυτή όπου συγκρίνονται τα ζεύγη των πλαισίων k και $k+1$. Αν τώρα δεν ανιχνευτούν αντιστοιχίες, δηλ. αν η απόσταση $t(i(k), j(k+1))$ δεν είναι μικρότερη ενός προκαθορισμένου επιπέδου, τότε το $j(k+1)$ συγκρίνεται με εντοπισμούς του πλαισίου $k-1$ στο υπό θεώρηση ζευγάρι $t(i(k-1), j(k+1))$ κλπ, γίνεται στο πλαίσιο k , αλλά η ανίχνευση συγκρίνεται με τις ανιχνεύσεις των προηγούμενων πλαισίων z για έναν αγώνα. Αυτή η διαδικασία μπορεί να επαναληφθεί για όλες τους εντοπισμούς από το πλαίσιο k έως το πλαίσιο $k-z$.

Το πλαίσιο $k-z$ και κάθε πλαίσιο μπορούν να περιέχουν διαφορετικό αριθμό εντοπισμό πεζών (a, b, g, h, \dots, i και j). Αν βρεθεί μια αντιστοίχιση, του εντοπισμού $j(k+1)$ με την $i(k-z)$, τότε αυτός ο εντοπισμός πιστώνεται στον αντίστοιχο id του εντοπισμού σε $(k-z)$. Εάν δεν υπάρχει αντιστοιχία μετά τη σύγκριση των προηγούμενων πλαισίων z , ο εντοπισμός θεωρείται ως νέος πεζός που εισέρχεται στο πλαίσιο. Ο νέος πεζός διαθέτει ένα νέο αναγνωριστικό ID και παρακολουθείται στο επόμενο πλαίσιο. Εάν ένας πεζός εγκαταλείψει τη σκηνή ή είναι απομονωμένος στο πλαίσιο $K+1$, η αντίστοιχη ανίχνευση στο πλαίσιο k δεν θα έχει καμία αντιστοιχία στο πλαίσιο $k+1$, αλλά αυτό το ID θα αποθηκευτεί στη βάση δεδομένων για μελλοντική χρήση σε πιθανές νέες αντιστοιχίες επόμενων πλαισίων καθώς μπορεί να επανεμφανιστεί μετά από αρκετά πλαίσια ένας πεζός.

Πολλαπλές αντιστοιχίες από τους εντοπισμούς $j(k+1)$ έως $i(k)$ μπορεί να προκύψουν, εάν $j(k+1) > i(k)$. Τέτοιες καταστάσεις μπορούν να επιλυθούν δημιουργώντας μια αντιστοιχία της $j(k+1)$ ανίχνευσης με το $n(k)$, αν αυτός ο εντοπισμός είχε την μικρότερη τιμή του συνόλου $\{t(1, i(k)), \dots, t(i(k), j(k+1))\}$. Η αντιστοιχία των μη καταναμημένων εντοπισμών γίνεται από το ίδιο σύνολο T_d παρακολούθησης χρησιμοποιώντας την δεύτερη μικρότερη τιμή του συνόλου $T_d = \{t(1, i(k)), \dots, t(i(k), j(k+1))\}$ αν είναι φυσικά κάτω από το όριο. Αν όχι, τότε το oi μη προσδιορισμένοι με ID εντοπισμοί συγκρίνονται με τους εντοπισμούς των προηγούμενων πλαισίων z για μια ακολουθία frames. [145]

7.3 Υλοποίηση Αλγόριθμου Εντοπισμού Πεζών

Περιγραφή Αλγόριθμου

- **Βήμα 0:** Μετατροπή video κλειστού κυκλώματος τηλεόρασης σε .jpg frames, μέσω εφαρμογής **Free Video to JPG Converter**
- **Βήμα 1:** Φόρτωση δεδομένων. Εφαρμόζεται το προεκπαιδευμένο Faster R-CNN για τον εντοπισμό των πεζών. Ο εντοπισμός πραγματοποιείται από τη εξαγωγή χαρακτηριστικών του τελευταίου στρώματος. Το προεκπαιδευμένο Faster R-CNN που εφαρμόζεται, προέρχεται από το toolbox PIOTR DOLLAR του MATLAB.
- **Βήμα 2:** Αρχικοποίηση παραμέτρων (βάρος του κάθε pixel και χαρακτηριστικό διάνυσμα, όριο της απόστασης κάθε χαρακτηριστικού διανύσματος, όριο για την απόσταση των pixel και αριθμός των frame για σύγκριση επανα-ταυτοποίησης).
- **Βήμα 3:** Εκτίμηση των κεντροειδών των πεζών.

$$P\vec{C}_{i(k)} = \begin{bmatrix} x_{(1,k)} & \dots & x_{(i,k)} \\ \vdots & \ddots & \vdots \\ y_{(1,k)} & \dots & y_{(i,k)} \end{bmatrix}$$

- **Βήμα 4:** Ανάγνωση του πρώτου και επόμενου frame, με διευθέτηση των δεδομένων.
- **Βήμα 5:** Εκτίμηση των κανονικοποιημένων αποστάσεων του χαρακτηριστικού διανύσματος, και των αποστάσεων των pixels δύο διαδοχικών frames.

$$F_{d(i(k),j(k+1))} = \frac{|FV_{i(k)} - FV_{j(k+1)}|}{|FV_{i(k)}| |FV_{j(k+1)}|}$$

- **Βήμα 6:** Συνδυασμός των αποστάσεων του χαρακτηριστικού διανύσματος με τις αποστάσεις pixel, και εκ νέου ταυτοποίηση των διαδοχικών frames. Ο πίνακας παρακολούθησης των πεζών, θα έχει τη μορφή:

$$T_{d(k+1)} = (w)P_{d(k+1)} + (1-w)E_{d(k+1)} = \begin{bmatrix} t_{(1,1)} & \dots & t_{(1,j(k+1))} \\ \vdots & \ddots & \vdots \\ t_{(i(k),1)} & \dots & t_{(i(k),j(k+1))} \end{bmatrix}$$

- **Βήμα 7** Ταίριασμα των IDs των πεζών του προηγούμενου frame, στο νέο frame.
- **Βήμα 8** Δημιουργία νέων IDs πεζών στο επόμενο frame, στην περίπτωση όπου κατά το ταίριασμα του βήματος 7, υπήρξαν νεοί και μη ταυτοποιημένοι πεζοί (χωρίς ID).
- **Βήμα 9** Οπτικοποίηση των αποτελεσμάτων στα διάδοχα frame, εφαρμόζοντας στους πεζούς το αντίστοιχο IDs τους.

Υλοποίηση στο περιβάλλον MATLAB

Ο ακόλουθος αλγόριθμος υλοποιεί την μεθοδολογία που παρουσιάστηκε και αναλύθηκε προηγουμένως, στην πλατφόρμα του MATLAB.

```

%% LOADING DATA
clear;
clc;

%FASTER pretrained RCNN WAS USED FOR THE DETECTION OF THE PEDESTRIANS
bboxes.bboxes=load('town_faster_rcnn');

%LAST FEATURE LAYER WAS USED TO EXTRACT THE %FEATURES OF EACH %
DETECTION
load ('town_feature_stack.mat')
load ('town_test_stack.mat')

%INITIALIZATION OF THE PARAMETERS
check_previous_stack{251}=1;
check_next_stack{251}=1;
acti_next_stack_res{251}=1;
acti_stack_res{251}=1;
A = single(zeros(1,1000));
B = single(zeros(1,2));
average_activation=zeros(1000,250);
next_id = length(bboxes.bboxes.ans{1, 1});
for ooo = 1:250
all_index{ooo} = 0;
end

%This is the folder containing the frames of the dataset, consider it
%changing to the folder where frames are located. In case of a video
file
%it should be converted to individual frames first before running the
code
imageName =
dir(fullfile('C:\TownCentreXVID_images16102018\','*.jpg'));
imageName = {imageName.name}';
dataFolder = 'C:\TownCentreXVID_images16102018\';

weight = 0.3; % weight of the pixel and feature vector
thr_fv= 0.8; % threshold of the distance for feature vector
thr_pxl=0.12; % Threshold for the pixel distance
frames_to_compare = 10; % number of frames to compare for
reidentification

figure;

```

```

%%
for ii = 1:length(imageNames)
%CENTROIDS OF THE PREDICTIONS
cen=[(bboxes.bboxes.ans{1,ii}(:,1)+bboxes.bboxes.ans{1,ii}(:,3))/2,
(bboxes.bboxes.ans{1,ii}(:,2)+bboxes.bboxes.ans{1,ii}(:,4))/2];
cen_nf=[(bboxes.bboxes.ans{1,ii+1}(:,1)+bboxes.bboxes.ans{1,ii+1}(:,3)
)/2,
(bboxes.bboxes.ans{1,ii+1}(:,2)+bboxes.bboxes.ans{1,ii+1}(:,4))/2];

%% READING THE FIRST FRAME AND NEXT FRAME
im =
imread(fullfile('C:\TownCentreXVID_images16102018\',imageNames{ii}));
im1 =
imread(fullfile('C:\TownCentreXVID_images16102018\',imageNames{ii+1}));

%REARRANGING THE DATA
box = [bboxes.bboxes.ans{1,ii}(:,1), bboxes.bboxes.ans{1,ii}(:,2),
(bboxes.bboxes.ans{1,ii}(:,3)-bboxes.bboxes.ans{1,ii}(:,1)),
(bboxes.bboxes.ans{1,ii}(:,4)-bboxes.bboxes.ans{1,ii}(:,2))];
box_nf = [bboxes.bboxes.ans{1,ii+1}(:,1),
bboxes.bboxes.ans{1,ii+1}(:,2), (bboxes.bboxes.ans{1,ii+1}(:,3)-
bboxes.bboxes.ans{1,ii+1}(:,1)), (bboxes.bboxes.ans{1,ii+1}(:,4)-
bboxes.bboxes.ans{1,ii+1}(:,2))];

%% NORMALIZED FEATURE VECTOR DISTANCES

D1 = zeros(length(box), length(box_nf));
D2 = zeros(length(box),1);
D3 = zeros(length(box_nf),1);
D4 = zeros(length(box), length(box_nf));

    for imfeat = 1:length(box)
        for timfeat=1:length(box_nf)
            D1(imfeat,timfeat) =
pdist2(acti_stack_image{ii}(:,imfeat)',acti_stack{ii}(:,timfeat)');
        end
    end
    for imfeat1 = 1:length(box)
        D2(imfeat1) = pdist2(A, acti_stack_image{ii}(:,imfeat1)');
    end
    for timfeat1 = 1:length(box_nf)
        D3(timfeat1) = pdist2(A, acti_stack{ii}(:,timfeat1)');
    end
    for imfeat2 = 1:length(box)
        for timfeat2=1:length(box_nf)
            D4(imfeat2,timfeat2) = (D1(imfeat2,timfeat2))/
(sqrt(D2(imfeat2)*D3(timfeat2)));
        end
    end
end

```

```

%% NORMALIZED PIXEL DISTANCES

d1 = zeros(length(box), length(box_nf));
d2 = zeros(length(box), 1);
d3 = zeros(length(box_nf), 1);
d4 = zeros(length(box), length(box_nf));
d = zeros(length(box_nf), 1);
n = zeros(length(box_nf), 1);

    for imfeat = 1:length(box)
        for timfeat=1:length(box_nf)
            d1(imfeat,timfeat) =
pdist2(cen(imfeat,:), cen_nf(timfeat,:));
        end
    end
    for imfeat1 = 1:length(box)
        d2(imfeat1) = pdist2(B, cen(imfeat1,:));
    end
    for timfeat1 = 1:length(box_nf)
        d3(timfeat1) = pdist2(B, cen_nf(timfeat1,:));
    end
    for imfeat2 = 1:length(box)
        for timfeat2=1:length(box_nf)
            d4(imfeat2,timfeat2) = (d1(imfeat2,timfeat2))/
(sqrt(d2(imfeat2)*d3(timfeat2)));
        end
    end

    for norm = 1:length(box_nf)
        [d(norm), n(norm)] = min(d4(:,norm));
    end

%% COMBINING FEATURE VECTOR DISTANCES WITH PIXEL DISTANCES
Dist_min_norm = zeros(length(box_nf), 1);
Index_min_norm = zeros(length(box_nf), 1);

comb = zeros(length(box), length(box_nf));
    for imfeat = 1:length(box)
        for timfeat=1:length(box_nf)
            comb(imfeat,timfeat) = ((1-weight)*D4(imfeat,timfeat) +
weight*(d4(imfeat,timfeat)));
        end
    end
    for norm = 1:length(box_nf)
        [Dist_min_norm(norm), Index_min_norm(norm)] =
min(comb(:,norm));
    end
end

```

```

%% RE-IDENTIFICATION IN SUBSEQUENT FRAMES
%Mapping IDs of previous frame to new frame
if ii>1
    Index_min_norm_new = zeros(length(Index_min_norm),1);
    for ppp = 1:length(Index_min_norm)
        Index_min_norm_new(ppp) =
Index_min_norm_previous(Index_min_norm(ppp));
    end
end

if ii==1
    Index_min_norm_new=Index_min_norm;
end
Index_stack{ii}=Index_min_norm_new;
dist_stack{ii}=Dist_min_norm;

%% CHECKING (FRAMES TO COMPARE) PREVIOUS FRAMES FOR A MATCH
if ii>1
lol=1;
for qqq = 1:length(Index_min_norm_previous)
    if ~(ismember(qqq, Index_min_norm));
        res_acti(lol) = qqq;
        lol=lol+1;
    end
end
end

if exist ('res_acti', 'var');
    check_previous = zeros(1000, length(res_acti));
    acti_stack_res{ii}=res_acti;
for poll = 1:length(res_acti)
    check_previous(:,poll) = acti_stack{ii-1}(:,res_acti(poll));
end
check_previous_stack{ii}=check_previous;
end

toll=1;
for lok = 1:length(Index_min_norm)
    for loki = lok+1:length(Index_min_norm)
        if Index_min_norm(lok)==Index_min_norm(loki)
            if
comb(Index_min_norm(lok),lok)<comb(Index_min_norm(lok),loki)
                res_acti_next(toll) = loki;
                toll=toll+1;
            else
                res_acti_next(toll) = lok;
                toll=toll+1;
            end
        end
    end
end
end

```

```

if exist ('res_acti_next', 'var');
    res_acti_next = unique(res_acti_next);
    acti_next_stack_res{ii} = res_acti_next;
    check_next = zeros(1000, length(res_acti_next));
    for poll = 1:length(res_acti_next)
        check_next(:,poll) = acti_stack{ii}(:,res_acti_next(poll));
    end

check_next_stack{ii}=check_next;
end

if ii>1
    if acti_next_stack_res{ii}
        dist_second = zeros(length(acti_next_stack_res{ii}),1);
        dist_third = zeros(length(acti_next_stack_res{ii}),1);
        Index_second = zeros(length(acti_next_stack_res{ii}),1);
        Index_third = zeros(length(acti_next_stack_res{ii}),1);
        for norm = 1: length(acti_next_stack_res{ii})
            DDD5=sort(comb(:,acti_next_stack_res{ii}(norm)));
            dist_second(norm)=(DDD5(2));
            dist_third(norm)=(DDD5(3));
            Index_second(norm)= find
(comb(:,acti_next_stack_res{ii}(norm))==dist_second(norm));
            Index_third(norm)= find
(comb(:,acti_next_stack_res{ii}(norm))==dist_third(norm));

            if dist_third(norm)< 0.5 && d4(Index_third(norm),
acti_next_stack_res{ii}(norm)) < thr_pxl
                Index_min_norm_new(acti_next_stack_res{ii}(norm))=
Index_min_norm_previous(Index_third(norm));
            end

            if dist_second(norm)< 0.5 && d4(Index_second(norm),
acti_next_stack_res{ii}(norm)) < thr_pxl
                Index_min_norm_new(acti_next_stack_res{ii}(norm))=
Index_min_norm_previous(Index_second(norm));
            end

        end
    end
end

end

min_dist = thr_fv;
for gulla = 1:frames_to_compare

if ii>gulla+1 && ~isempty(check_previous_stack{1,ii-gulla}) &&
~isempty(check_next_stack{1,ii})

Df1 = zeros(length(check_previous_stack{1,ii-gulla}(1,:)),
length(check_next_stack{1,ii}(1,:)));
Df2 = zeros(length(check_previous_stack{1,ii-gulla}(1,:)),1);
Df3 = zeros(length(check_next_stack{1,ii}(1,:)),1);
Df4 = zeros(length(check_previous_stack{1,ii-gulla}(1,:)),
length(check_next_stack{1,ii}(1,:)));

```



```

Dist_frame = zeros(length(check_next_stack{1,ii}(1,:)),1);
Index_frame = zeros(length(check_next_stack{1,ii}(1,:)),1);

    for imfeat = 1:length(check_previous_stack{1,ii-gulla}(1,:))
        for timfeat=1:length(check_next_stack{1,ii}(1,:))
            Df1(imfeat,timfeat) = pdist2(check_previous_stack{1,ii-gulla}(:,imfeat)',check_next_stack{1,ii}(:,timfeat)');
        end
    end

    for imfeat1 = 1:length(check_previous_stack{1,ii-gulla}(1,:))
        Df2(imfeat1) = pdist2(A, single(check_previous_stack{1,ii-gulla}(:,imfeat1)'));
    end

    for timfeat1 = 1:length(check_next_stack{1,ii}(1,:))

        Df3(timfeat1) = pdist2(A,
single(check_next_stack{1,ii}(:,timfeat1)'));
    end

    for imfeat2 = 1:length(check_previous_stack{1,ii-gulla}(1,:))
        for timfeat2=1:length(check_next_stack{1,ii}(1,:))
            Df4(imfeat2,timfeat2) = (Df1(imfeat2,timfeat2))/
(sqrt(Df2(imfeat2)*Df3(timfeat2)));
        end
    end

    for norm = 1:length(check_next_stack{1,ii}(1,:))
        [Dist_frame(norm), Index_frame(norm)] = min(Df4(:,norm));
        if min_dist>Dist_frame(norm)
            min_dist=Dist_frame(norm);
        end
    end

    dist_frame_stack{ii}{gulla}=Dist_frame;
    index_frame_stack{ii}{gulla}=Index_frame;
for poppo = 1:length(acti_next_stack_res{ii})
    for polly = 1:length(dist_frame_stack{ii})
        if index_frame_stack{ii}{polly}
            acti_next_stack_res_dist{ii}{poppo}{polly} =
dist_frame_stack{ii}{polly}(poppo);

acti_next_stack_res_dist_min{ii}(poppo)=min(cell2mat(acti_next_stack_r
es_dist{ii}{poppo}(:)));
        end
    end
end
for ret = 1:length(index_frame_stack{1,ii}{gulla})
    for rete = ret+1:length(index_frame_stack{1,ii}{gulla})

        if index_frame_stack{1,ii}{gulla}(ret) ==
index_frame_stack{1,ii}{gulla}(rete)

```

```

        if dist_frame_stack{1,ii}{gulla}(ret) <
dist_frame_stack{1,ii}{gulla}(rete) &&
dist_frame_stack{1,ii}{gulla}(ret) < thr_fv &&
dist_frame_stack{1,ii}{gulla}(ret) <=
acti_next_stack_res_dist_min{ii}(ret)
            Index_min_norm_new(acti_next_stack_res{1,ii}(ret))
= Index_stack{1,ii-gulla-1}(acti_stack_res{1,ii-
gulla}(index_frame_stack{1,ii}{gulla}(ret)));
        end

        if dist_frame_stack{1,ii}{gulla}(ret) >
dist_frame_stack{1,ii}{gulla}(rete) &&
dist_frame_stack{1,ii}{gulla}(rete) < thr_fv &&
dist_frame_stack{1,ii}{gulla}(rete) <=
acti_next_stack_res_dist_min{ii}(rete)

Index_min_norm_new(acti_next_stack_res{1,ii}(rete)) =
Index_stack{1,ii-gulla-1}(acti_stack_res{1,ii-
gulla}(index_frame_stack{1,ii}{gulla}(rete)));
        end

    else
        if dist_frame_stack{1,ii}{gulla}(ret) < thr_fv &&
dist_frame_stack{1,ii}{gulla}(ret) <=
acti_next_stack_res_dist_min{ii}(ret)
            Index_min_norm_new(acti_next_stack_res{1,ii}(ret))
= Index_stack{1,ii-gulla-1}(acti_stack_res{1,ii-
gulla}(index_frame_stack{1,ii}{gulla}(ret)));
        end
    end
end
end
end

%% GENERATING NEW IDS

for lok = 1:length(Index_min_norm)
    for loki = lok+1:length(Index_min_norm)
        if Index_min_norm_new(lok)==Index_min_norm_new(loki)

            if comb(Index_min_norm(lok),lok) <
comb(Index_min_norm(lok),loki)
                next_id=next_id+1;
                Index_min_norm_new(loki)=next_id;

            else

                next_id=next_id+1;
                Index_min_norm_new(lok)=next_id;

            end

        end
    end
end
end

```

```
Index_norm_stack{ii}=Index_min_norm;

if ii==1
    Index_min_norm_previous=1:14;
end

%% VISUALIZATION
if ii==1
    labell = cellstr( num2str((Index_min_norm_previous')));
else
    labell = cellstr( num2str((Index_min_norm_previous)));
end
Index_min_norm_previous=Index_min_norm_new;
Index_new_stack{ii}=Index_min_norm_new;

%% DISPLAYING RESULTS

imshow(fullfile('C:\TownCentreXVID_images16102018\',imageNames{ii}))
hold on
text((double(cen(:,1))), double(cen(:,2)), labell,
'Color','red','FontSize',20)
pause(0.001)
hold off

clear -regexp ^res
%%
if ii==(length(imageNames)-2)
    break;
end

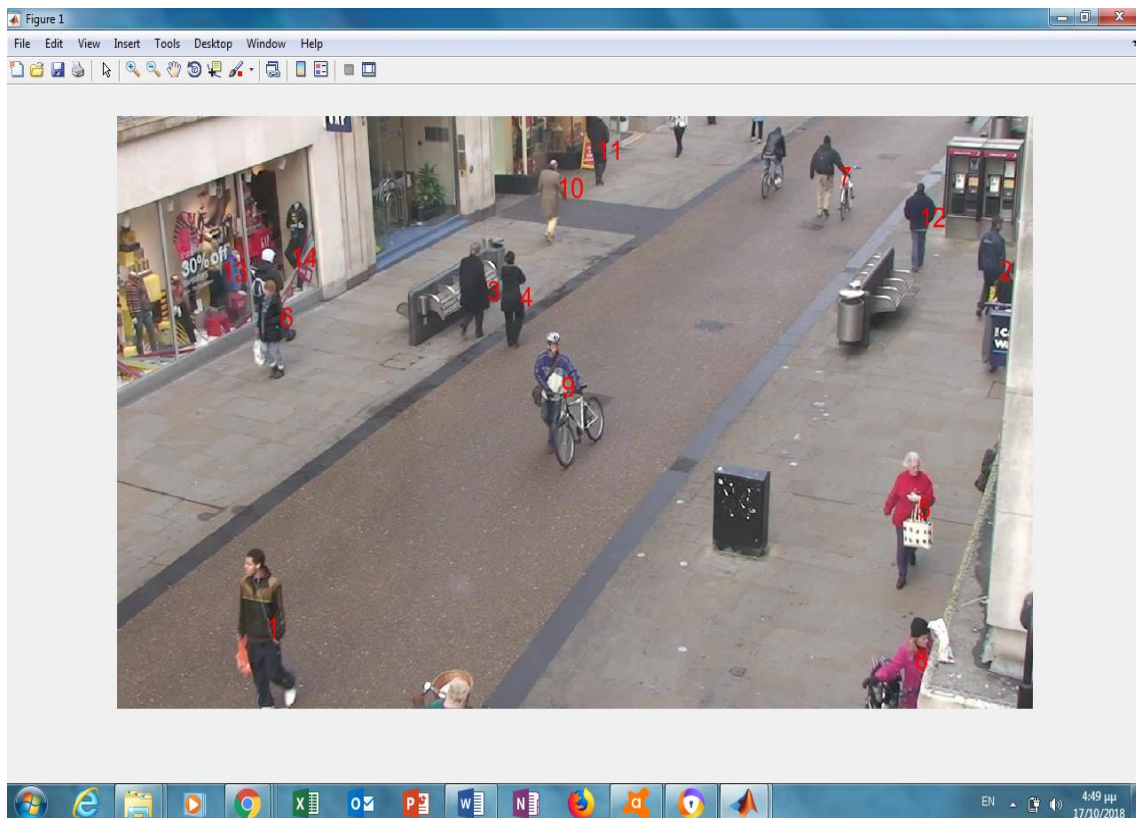
end
```

7.4 Αποτελέσματα

Όπως αναφέρθηκε ανωτέρω, χρησιμοποιήθηκε το video κλειστού κυκλώματος τηλεόρασης αστικού κέντρου, σε μορφή .avi, για αξιολόγηση στη παρούσα μελέτη. Αφού το video .avi μετασχηματίστηκε σε frames (εικόνες jpeg) μέσω της εφαρμογής που επιλέχθηκε, *Free Video to JPG Converter*, επιλέχθηκε αργός ρυθμό λήψης καρέ/frames για τα πρώτα 30-31 δευτερόλεπτα του βίντεο με τιμή ρυθμού 8 καρέ ανά δευτερόλεπτο, συγκεντρώνοντας 250 καρέ συνολικά. Η ανίχνευση και η παρακολούθηση αξιολογούνται ξεχωριστά με τη χρήση πινάκων παρακολούθησης Td όπως αναφέρθηκε ανωτέρω. Τα αποτελέσματα των εντοπισμών και της παρακολούθησης πεζών παρουσιάζονται στα επόμενα σχήματα.

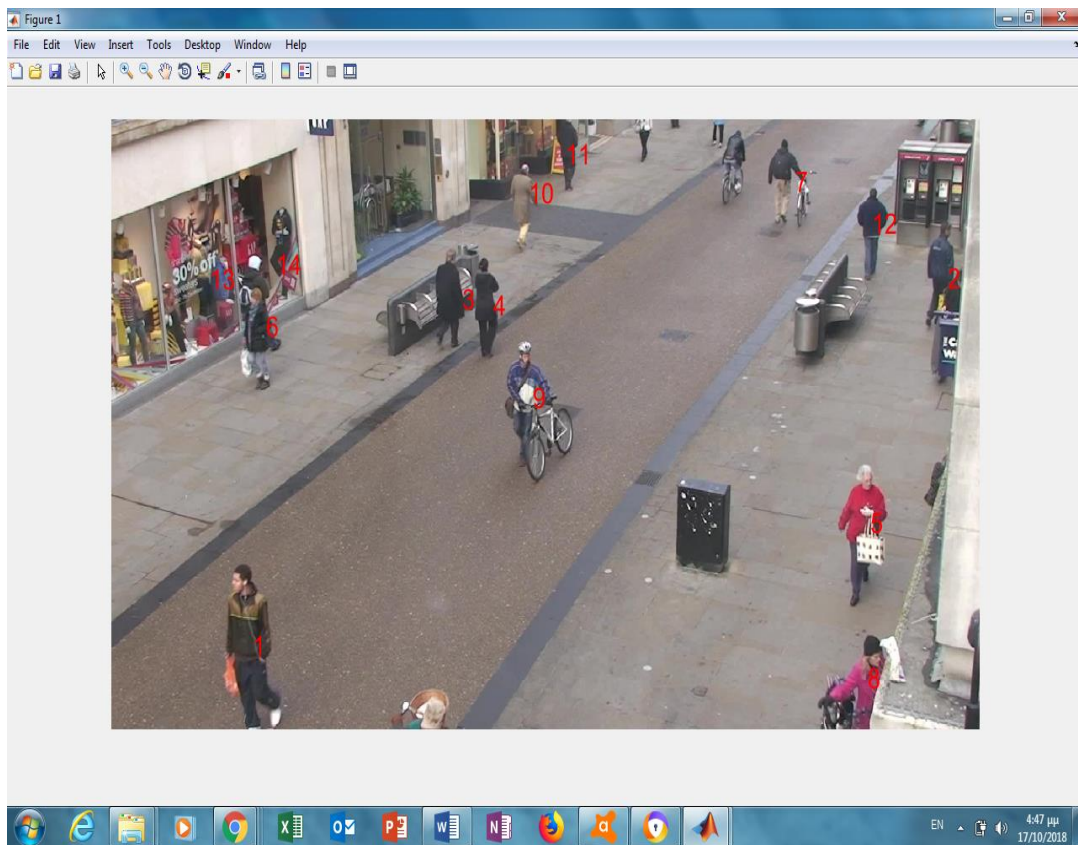
Είναι προφανές ότι στην αρχή το σύστημα ανταποκρίνεται σχετικά καλά αλλά προϊόντος του χρόνου τα αποτελέσματα είναι σχετικά φτωχά. Είναι προφανές ότι χρειάζονται βελτιώσεις σε πολλούς τομείς αυτής της προσπάθειας. Πιθανόν ακόμη και ο μικρός ρυθμός λήψης frames να παίζει ρόλο. Χρειάζεται σαφώς συστηματικότερη μελέτη από την παρούσα, αλλά από ότι διαφαίνεται υπάρχουν πολλές δυνατότητες βελτίωσης ακόμη και με pretrained CNN, όπως φαίνεται στις εντολές MATLAB ,

```
load ('town_feature_stack.mat'), load ('town_test_stack.mat'),
load ('town_faster_rcnn.mat'),
```



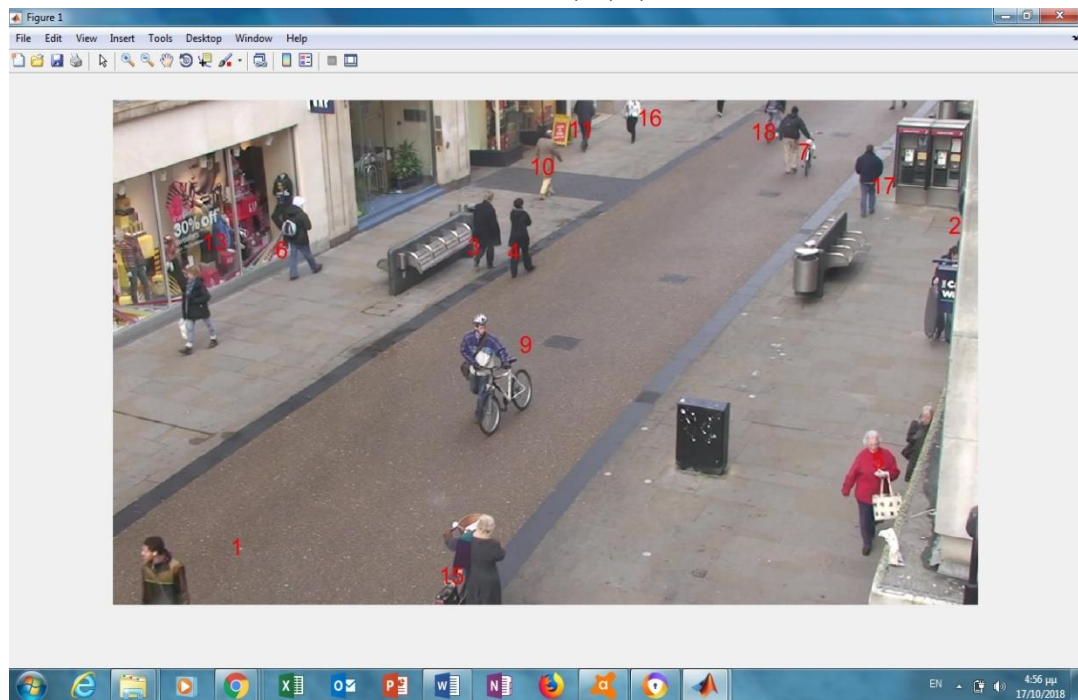
Εικόνα 7.3

Τα αποτελέσματα επιτυχούς παρακολούθησης πεζών μετά το 2^ο frame , όπως φαίνονται μέσω των αντίστοιχων IDs πάνω στους πεζούς.



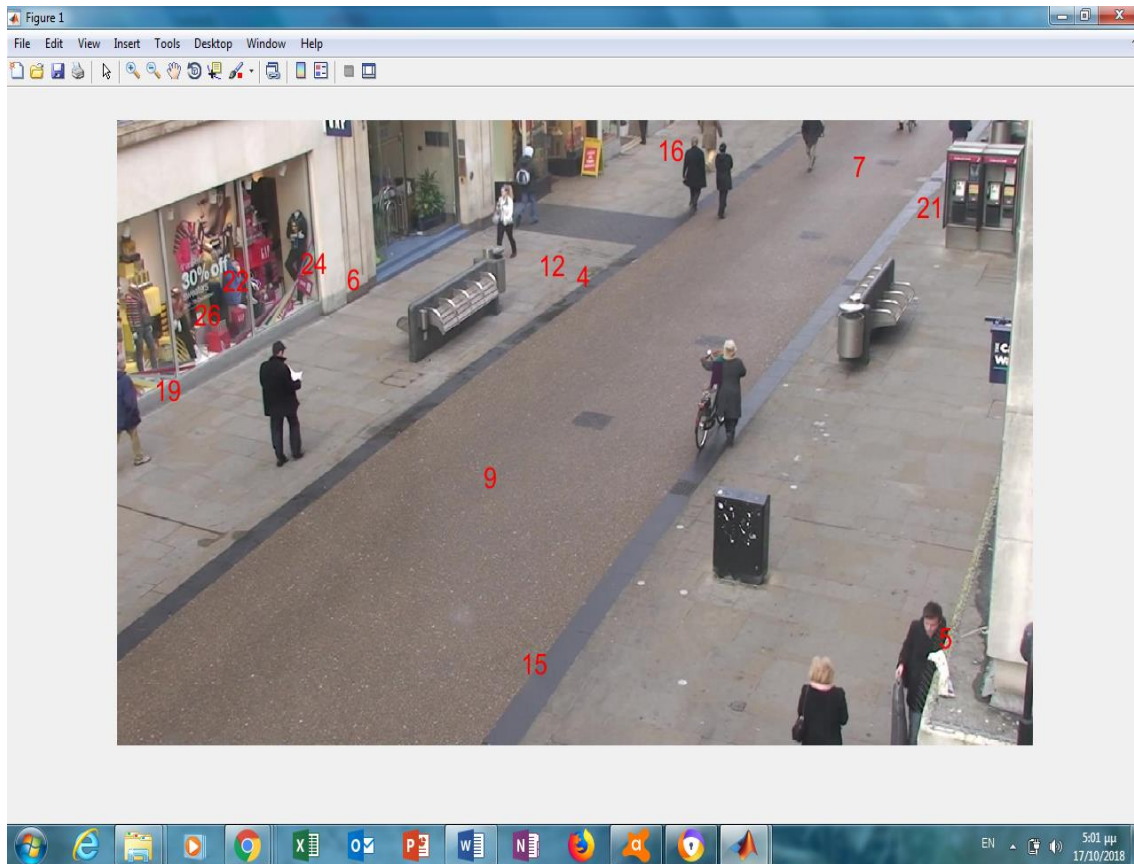
Εικόνα 7.4

Τα αποτελέσματα επιτυχούς παρακολούθησης πεζών μετά το 3^ο frame , όπως φαίνονται μέσω των αντίστοιχων IDs πάνω στους πεζούς.



Εικόνα 7.5

Αναγνώριση πεζών από αυτόνομο ρομποτικό όχημα με χρήση τεχνητής νοημοσύνης και τεχνικών βαθιάς μάθησης



Εικόνα 7.6

Τα αποτελέσματα επιτυχούς παρακολούθησης πεζών μετά το 17^ο frame , όπως φαίνονται μέσω των αντίστοιχων IDs πάνω στους πεζούς.

Όπως φαίνεται στην παραπάνω εικόνα, εμφανίζονται και κάποια IDs που αντιστοιχούν σε πεζούς που έχουν φύγει από το σκηνικό και παραμένουν στη βάση. Όπως και κάποιες αποτυχιές εντοπισμού. Είναι προφανές ότι οι αλγόριθμοι που παρουσιάζονται δεν έχουν επαρκή γενικευτική ικανότητα σε μερικές δύσκολες καταστάσεις στα όρια, όπως επίσης όταν υπάρχουν πεζοί που είναι πολύ κοντά ο ένας στον άλλο. Κατά συνέπεια η μακρόχρονη παρακολούθηση είναι μάλλον φτωχή σε αποτελέσματα.

Προτάσεις Περαιτέρω Βελτίωσης

Η ορθή, αλάνθαστη, ακριβής και σε πραγματικό χρόνο αναγνώριση πεζών από ένα αυτόνομο ρομποτικό όχημα, αποτελεί μία από τις σημαντικότερες προκλήσεις που αντιμετωπίζουν αυτά τα οχήματα. Η αναγνώριση πεζών εν γένει, είναι ίσως ο σπουδαιότερος εντοπισμός όπου θα πρέπει να επιτελέσει ένα αυτόνομο ρομποτικό όχημα, και αν δεν καταστεί δυνατός ο εντοπισμός αυτός σε απόλυτο επίπεδο, δεν θα είναι ποτέ και δυνατή η αφενός αναβάθμιση των μελλοντικών αυτόνομων ρομποτικών οχημάτων στο επίπεδο 5 του *NHTSA*, και αφετέρου δεν θα καταστεί δυνατή η αδειοδότηση κυκλοφορίας τους στα σύγχρονα αστικά και επαρχιακά κέντρα.

Όπως έγινε κατανοητό στα πλαίσια της παρούσας διατριβής, η *Τεχνητή Νοημοσύνη* και συγκεκριμένα η Βαθιά Μάθηση, αποτελεί ένα από τα σημαντικότερα εργαλεία της επιστημονικής κοινότητας ώστε να προβεί στην επίτευξη του σκοπού αυτού. Οι σύγχρονες τεχνικές και μέθοδοι της Βαθιάς Μάθησης, έχοντας σημειώσει ραγδαία εξέλιξη στην ταχύτητα, την ποιότητα και το ποσοστό πιστότητας αναγνώρισης, αποτελούν την βέλτιστη μέθοδο αναγνώρισης προτύπων στις εφαρμογές Υπολογιστικής Όρασης.

Στο κεφάλαιο 6, παρουσιάστηκαν πειραματικά οι δυνατότητες αναγνώρισης διαφόρων αντικειμένων, από κάποιο αυτόνομο όχημα, με χρήση τεχνικών *Βαθιάς Μάθησης*, μέσα από δύο κατάλληλα διαμορφωμένους αλγόριθμους. Όπως προκύπτει από τις δοκιμές που πραγματοποιήθηκαν, οι αλγόριθμοι μπόρεσαν να αναγνωρίσουν επιτυχώς τα προς αναγνώριση πρότυπα, σε πραγματικό χρόνο και με πολύ υψηλά ποσοστά επιτυχίας. Επίσης, φάνηκε πως οι αλγόριθμοι έχουν τη δυνατότητα να προβούν σε πραγματικό χρόνο, σε αναγνώριση πολλαπλών προτύπων ταυτόχρονα.

Παράλληλα, στο κεφάλαιο 7, κατέστη δυνατή η αναγνώριση πολλαπλών πεζών ταυτόχρονα, σε πραγματικό χρόνο. Για την επίτευξη του σκοπού αυτού, εφαρμόστηκε ένας συνδυασμός μερικών εκ των ταχύτερων μεθόδων της Βαθιάς Μάθησης, η μέθοδος των προ-εκπαιδευμένων Συνελκτικών Νευρωνικών Δικτύων, και ο αλγόριθμος των *Faster R-CNN*. Παρότι, τα αποτελέσματα της εφαρμογής αυτής χαρακτηρίζονται από πολύ μεγάλη ταχύτητα, προφανές ότι χρειάζονται βελτιώσεις σε πολλούς τομείς αυτής της προσπάθειας.

Η προσέγγιση βελτίωσης των επιδόσεων της εφαρμογής αυτής, θα πρέπει να περιλαμβάνει βήματα εξέλιξης τόσο του αλγόριθμου αυτού καθαυτού, όσο και των συνθηκών εφαρμογής του. Η αύξηση του ρυθμού λήψης των frames της εξεταζόμενης εικόνας, δύναται να βελτιώσει αισθητά την αποτελεσματικότητα της εφαρμογής.

Ακόμα, η εφαρμογή λόγω της μη επαρκούς γενικευτικής της ικανότητας, δεν δύναται να αναγνωρίσει πεζούς όταν αυτοί βρίσκονται κοντά στα οπτικά όρια ή επικαλύπτονται. Ακόμα, παρατηρείται το γεγονός ορισμένοι πεζοί να μην ανιχνεύονται σε ορισμένα καρέ. Επομένως, γίνεται αντιληπτή η ανάγκη περεταίρω εκπαίδευσης του δικτύου ώστε να παρέχει υψηλότερα επίπεδα αναγνώρισης προτύπων.

Κατά τη λειτουργία της παρακολούθησης ενός πεζού, παρατηρείται το φαινόμενο διατήρησης του αναγνωριστικού ID στη βάση, άρα και στην εικόνα του output του δικτύου, ακόμα και όταν ο πεζός βρίσκεται εκτός εικόνας. Σε αυτή τη περίπτωση απαιτείται η τροποποίηση του αλγόριθμου έτσι ώστε να πραγματοποιείται αυτόματη διαγραφή από την βάση, των IDs των οποίων τα πρότυπα δεν εμφανίζονται πλέον στην εικόνα επί μία σειρά αρκετών frames. Ακόμα, η εφαρμογή της μεθόδου *Mask R-CNN*, θα μπορούσε να συνεισφέρει αποτελεσματικά στην διαδικασία επακριβή εντοπισμού των προτύπων με ακρίβεια επιπέδου pixel.

Τέλος, αξίζει να σημειωθεί ότι ιδιαίτερο ενδιαφέρον παρουσιάζουν και λοιποί νεότεροι αλγόριθμοι *Μηχανικής Μάθησης* οι οποίοι χαρακτηρίζονται από ακόμη υψηλότερη ταχύτητα ενώ παράλληλα οι αλγόριθμοι αυτοί φέρουν διαφορετική αρχιτεκτονική αφού δεν βασίζονται στην φιλοσοφία εξαγωγής προτεινόμενων περιοχών. Κατά τα τελευταία έτη, οι αλγόριθμοι YOLO, Deeplab Xception, SSD και RetinaNet έχουν παρουσιάσει αξιοσημείωτες επιδόσεις στην αναγνώριση προτύπων και ολόένα και περισσότερο χρησιμοποιούνται σε μικτές αρχιτεκτονικές *R-CNN*. Συνεπώς, θα ήταν πολύ ενδιαφέρουσα η προσπάθεια δόμησης ενός δικτύου μικτής αρχιτεκτονικής το οποίο να επιτυγχάνει αναγνώριση προτύπων αλλά και παρακολούθηση αυτών, μέσα από προεκπαιδευμένα – αλλά και συνεχώς εκπαιδευόμενα – *Νευρωνικά Δίκτυα*.

Επίλογος - Συμπεράσματα

Από τη παρούσα εργασία γίνεται κατανοητό ότι η επίλυση του προβλήματος της βελτιστοποίησης αλγορίθμων *Βαθιάς Μάθησης* με σκοπό την εφαρμογή τους στον εντοπισμό πεζών από ένα αυτόνομο ρομποτικό όχημα, αποτελεί ένα πολύ ενδιαφέρον ερευνητικό πεδίο της *Τεχνητής Νοημοσύνης* το οποίο συνδυάζει πολλαπλές μεθόδους και εργαλεία για την επίτευξη του σκοπού αυτού. Συνεπώς, η προαγωγή και βελτιστοποίηση των τεχνολογικών μεθόδων δόμησης αποτελεσματικών και ταχύτατων εργαλείων *Τεχνητής Νοημοσύνης*, αποτελεί αυτοσκοπό για την εκσυγχρόνιση των αυτόνομων αυτοκινήτων και την γενικότερη τους ενσωμάτωση στις σύγχρονες πολιτείες. Η παρούσα διατριβή, πραγματεύτηκε την εφαρμογή σύγχρονων τεχνολογικών εργαλείων με σκοπό την υψίστης αποδοτικότητας και ταχύτητας αναγνώριση πεζών από τα αυτόνομα ρομποτικά οχήματα, με εφαρμογή αλγορίθμων *Βαθιάς Μάθησης* που επιτελούν αυτόνομες και ολοκληρωμένες λειτουργίες *Υπολογιστικής Όρασης*.

Η υιοθέτηση των αυτόνομων ρομποτικών οχημάτων από τις σύγχρονες κοινωνίες, όπως αναλύθηκε και στα προηγούμενα κεφάλαια της παρούσας εργασίας, δύναται να εξαλείψει μία πληθώρα σημαντικών προβλημάτων που αντιμετωπίζουν οι πολίτες σήμερα. Είναι δε αυτονόητο, ότι η αντικατάσταση της συμβατικής οδήγησης από την αυτόνομη οδήγηση, είναι κάτι το τετελεσμένο και αποτελεί ζήτημα μερικών ετών να συμβεί καθολικά σε ολόκληρη την υφήλιο. Η σημαντικότερη πρόκληση που έχει να αντιμετωπίσει η ιδέα αυτή, είναι αυτή της *Υπολογιστικής Όρασης* του αυτόνομου ρομποτικού οχήματος. Πρόκληση που φέρει επισύρει πολλαπλά τεχνολογικά ερωτήματα, και απαιτεί μεγάλο εύρος υψηλής αποδοτικότητας εργαλείων για την αντιμετώπιση της. Έτσι, η διατριβή αυτή υλοποιήθηκε υπό το πρίσμα της γενικότερης συμβολής της, τόσο στην τεχνολογία και την φιλοσοφία των αυτόνομων ρομποτικών οχημάτων που προσπάθησε να αναλύσει επαρκώς, όσο και στο πεδίο της *Υπολογιστικής Όρασης* με χρήση σύγχρονων τεχνολογιών *Βαθιάς Μάθησης* εν γένει, για την όποια περεταίρω εφαρμογή της σε όλα τα τεχνολογικά πεδία.

Οι σύγχρονες τεχνολογίες *Υπολογιστικής Όρασης*, εφαρμόζονται με αξιοποίηση των πλέον αποτελεσματικότερων εργαλείων που παρέχει η *Βαθιά Μάθηση* και των αντίστοιχων αλγορίθμων της, σε εικόνες που λαμβάνονται σε πραγματικό χρόνο, από τα οπτικά αισθητήρια συστήματα που φέρει το αυτόνομο ρομποτικό όχημα. Αξίζει να σημειωθεί ότι τα σύγχρονα εργαλεία *Βαθιάς Μάθησης* που εφαρμόστηκαν στην παρούσα διατριβή, όπως είναι τα *Faster R-CNN*, τα *προεκπαιδευμένα Νευρωνικά Δίκτυα* αλλά και οι μικτές αρχιτεκτονικές δομές, αποτελούν τα *state-of-the-art* μέσα που υπάρχουν στον ευρύτερο χώρο της *Βαθιάς Μάθησης* τα οποία εφαρμόζονται σε προβλήματα *Υπολογιστικής Όρασης* των αυτόνομων ρομποτικών οχημάτων. Το πλήθος των διαφορετικών προτύπων, οι περιορισμοί λόγω μη ιδανικής αποτύπωσης των προτύπων στις εισερχόμενες στο σύστημα εικόνες, η απαίτηση πολύ υψηλής ταχύτητας εξαγωγής αναγνωρισμένων προτύπων, η ενδεχόμενη επικάλυψή τους κα, καθιστούν την εφαρμογή τεχνικών *Βαθιάς Μάθησης* μονόδρομο στην αναγνώριση προτύπων από κάποιο αυτόνομο ρομποτικό όχημα. Για το λόγο αυτό και επιλέχθηκαν τα εργαλεία αυτά στο πλαίσιο υλοποίησης της συγκεκριμένης διατριβής.

Το μείζον ζήτημα της αναγνώρισης πεζού από κάποιο αυτόνομο ρομποτικό όχημα, έχει απασχολήσει πολλές φορές την επιστημονική κοινότητα με σκοπό την ολοένα και ταχύτερη επίλυση του. Για το λόγο αυτό, η προσπάθεια που υλοποιήθηκε στη παρούσα διατριβή, δεν ήταν απλώς η αποτύπωση και εφαρμογή των πλέον ταχύτερων και αξιόπιστων μεθόδων και αλγορίθμων αναγνώρισης προτύπων, αλλά η προσπάθεια συνδυασμού των βέλτιστων τεχνικών που διατίθενται, για την δόμηση ενός πρότυπου αλγορίθμου ο οποίος θα επιλύει το παραπάνω πρόβλημα αντιμετωπίζοντας παράλληλα τις όποιες αστάθειες δημιουργούνταν από αυτό. Η τεχνική που ακολουθήθηκε για τη δόμηση του δικτύου, εστιάζει κυρίως στην συσχέτιση αντιστοιχιών των προηγούμενων με των επόμενων πλαισίων, έτσι ώστε να καταστεί δυνατή αφενός η εξαγωγή των χαρακτηριστικών και αφετέρου η παρακολούθηση των αναγνωρισμένων προτύπων, εντός των *frames*. Με τον τρόπο αυτό, επιτυγχάνεται η δόμηση μίας πρότυπης αρχιτεκτονικής υψηλής ταχύτητας η οποία φέρει ικανοποιητικά – οριακών αποδεκτά επίπεδα ακριβείας. Παράλληλα, εντοπίζονται όλα τα πιθανά προς βελτίωση τμήματα της προσπάθειας αυτής με

απώτερο σκοπό την εξάλειψη τους μέσω διαφοροποίησης στην αρχιτεκτονική και στον τρόπο εφαρμογής της στα δεδομένα εισόδου.

Το μάλλον οριακά αποδεκτό ποσοστό ακριβείας, η αδυναμία εντοπισμού επικαλυπτόμενων πεζών καθώς και πεζών που δεν εμφανίζονται αποκλειστικά εντός της εικόνας, η διατήρηση του αυξαντα αριθμού στην βάση της εφαρμογής ακόμα και μετά την έξοδο του παρακολουθούμενου στόχου από το πλάνο και άλλα πολλά, αποτελούν κάποιες από τις προκλήσεις επίλυσης που φέρει η παρούσα εφαρμογή. Ακόμα, η ραγδαία εξέλιξη που πραγματοποιήθηκε στους νέους αλγορίθμους Βαθιάς Μάθησης, όπως αυτοί αναφέρθηκαν και προηγουμένως, αποτελεί μία νέα πρόκληση επιτάχυνσης της λειτουργίας του μοντέλου, ενώ παράλληλα θα μπορούσαν να εφαρμοστούν και δυαδικές μάσκες με σκοπό τον ακόμη υψηλότερης ακρίβειας εντοπισμό τους παρακολουθούμενου στόχου. Γίνεται λοιπόν αντιληπτό, ότι η βελτίωση της λειτουργικής απόδοσης του αλγορίθμου που παρουσιάστηκε σε συνδυασμό με την εφαρμογή των νέων εργαλείων και τεχνικών που αναφέρθηκαν, μπορούν να αυξήσουν σε πολύ υψηλό επίπεδο τους αποδοτικούς δείκτες της εφαρμογής.

Ωστόσο, είναι ιδιαίτερα σημαντικό να τονιστεί ότι ο πρότυπος αυτός αλγόριθμος, πέρας από το πεδίο της αναγνώρισης πεζών και τις εφαρμογές *Υπολογιστικής Όρασης* των αυτόνομων ρομποτικών οχημάτων εν γένει, αποτελεί μία πολύ ισχυρή βάση και για λοιπές άλλες μελλοντικές εφαρμογές Υπολογιστικής όρασης ποικίλου περιεχομένου. Για παράδειγμα, η αναγνώριση και πρόβλεψη των ανθρώπινων ενεργειών σε περιβάλλοντα πραγματικού κόσμου βρίσκει εφαρμογή μία πληθώρα καταστάσεων που λαμβάνουν χώρα σε περιβάλλοντα επιτηρούμενης συμπεριφοράς. Βέβαια, η τεχνική αυτή βρίσκει εφαρμογή και σε λοιπές άλλες περιπτώσεις εντοπισμού και παρακολούθησης αντικειμένων. Ο συνδυασμός των *Προεκπαιδευμένων Βαθιών Συνελικτικών Νευρωνικών Δικτύων* για την τάχιστα εξαγωγή χαρακτηριστικών, με τις τεχνικές παρακολούθησης αντιστοιχιών εντός ακολουθίας παραγόμενων πλαισίων αποτελεί ένα ισχυρό εργαλείο για την επίτευξη των αντίστοιχων εφαρμογών.

Τέλος, μέσα από την εκπόνηση της διατριβής αυτής, παρουσιάστηκαν οι πραγματικές ανάγκες εφαρμογής των τεχνολογιών *Τεχνητής Νοημοσύνης* στα αυτόνομα ρομποτικά οχήματα, οι περιορισμοί που αντιμετωπίζουν οι κλασικοί αλγόριθμοι της Τεχνητής νοημοσύνης σε εφαρμογές εντοπισμού προτύπων, η γενικότερη αλλά και ειδικότερη φιλοσοφία των τεχνικών *Βαθιάς Μάθησης* αλλά και η εφαρμογή τους στο μείζον ζήτημα της αναγνώρισης πεζού. Έτσι, και κατόπιν των σχετικών εργαστηριακών δοκιμών του αλγορίθμου που παρουσιάστηκε, η προαγωγή των σύγχρονων εργαλείων *Βαθιάς Μάθησης*, αποτελεί επιτακτική ανάγκη για την βελτιστοποίηση των διεργασιών αναγνώρισης προτύπων εξαγόμενων από οπτικά δεδομένα. Η *Υπολογιστική Όραση*, αποτελεί μία πολλά υποσχόμενη επιστήμη που αναμένεται να αλλάξει ριζικά τις τεχνολογικές υποδομές των επόμενων δεκαετιών, προάγοντας σε πολύ υψηλά επίπεδα την βελτίωση των ανθρώπινων αναγκών. Τα αυτόνομα ρομποτικά οχήματα, θα επιφέρουν πραγματική επανάσταση στην καθημερινότητα των πολιτών λύνοντας. Και οι σύγχρονες και ολοένα αποδοτικότερες τεχνικές *Βαθιάς Μάθησης* που εφαρμόζονται καθημερινά, αποτελούν μία «υπόσχεση» βραχυπρόθεσμης επίτευξης του τεχνολογικού θαύματος των αυτόνομων ρομποτικών οχημάτων.

Βιβλιογραφία – Αναφορές

- [1] *A Philosophy for Developing Trust in Self-Driving Cars* ,
Michael Wagner, Philip Koopman ,
Carnegie Mellon University, Robotics Institute & Department of Electrical and Computer Engineering ,
Springer ,
2015
- [2] *Printed / Flexible / Stretchable Sensors: Creative Solutions for Empowering Abundance Realities* ,
Roger H. Grace ,
Roger Grace Associates ,
www.rgrace.com ,
2015
- [3] *End to End Learning for Self-Driving Cars* ,
Mariusz Bojarski, Davide Del Testa, Daniel Dworakowski, Bernhard Firner, Beat Flepp, Praseem Goyal,
Lawrence D. Jackel, Mathew Monfort, Urs Muller, Jiakai Zhang, Xin Zhang, Jake Zhao, Karol Zieba ,
NVIDIA Corporation ,
NVIDIA Press ,
2006
- [4] *Autonomic Vehicles* ,
Kevin Curran, James Lee ,
Intelligent Systems Research Centre, Faculty of Engineering, University of Ulster ,
Computer and Information Science ,
2008
- [5] *Computer Vision and its Application in Self-Driving Cars* ,
Benjamin Cutilli ,
Haverford College | Bryn Mawr College ,
University Degree Thesis ,
2013
- [6] *The Google Car: Driving Toward A Better Future?* ,
Sharon L. Poczter, Luka M. Jankovic ,
Cornell University ,
Journal of Business Case Studies ,
2014
- [7] *An Autonomous Driverless Car: An Idea to Overcome the Urban Road Challenges* ,
Sheetal Ds Rathod ,
Department of Information Technology, JDIET Yavatmal, Amaravati University ,
Journal of Information Engineering and Applications ,
2013
- [8] *An Intelligent Architecture for Legged Robot Terrain Classification Using Proprioceptive and Exteroceptive Data* ,
Mohammed Nour Abdel Gwad Ahmed ,
University of Bremen ,
Doctoral Dissertation ,
2015

- [9] *Self-Driving Cars* ,
James Martin ,
2014
- [10] *Driverless Dreams: Technological Narratives and the Shape of the Automated Car* ,
Erik Lee Stayton ,
Massachusetts Institute of Technology ,
Master's Degree Thesis ,
2015
- [11] *A Seminar Report on "Google Driveless Car"* ,
Hrishikesh Dongare ,
University of Pune ,
Bachelor's Degree Thesis ,
2015
- [12] *Advanced Microsystems for Automotive Applications 2006* ,
J. Valldorf, W. Gessner ,
VDI/VDE Innovation + Technik GmbH ,
Springer ,
2006
- [13] *Research Priorities for Robust and Beneficial Artificial Intelligence* ,
Stuart Russell, Daniel Dewey, Max Tegmark ,
Association for the Advancement of Artificial Intelligence ,
AI Magazine ,
2015
- [14] *Practical Object Recognition in Autonomous Driving and Beyond* ,
Alex Teichman, Sebastian Thrun ,
Computer Science Department, Stanford University ,
IEEE ,
2012
- [15] *Automobile Insurance in the Rea of Autonomus Vehicles* ,
Jerry Albright, Alex Bell, Joe Schneider, Chris Nyce ,
KPMG ,
KPMG Press ,
2015
- [16] *Technology and Computing Requirements for Self-Driving Cars* ,
Intel Automotive R&D Dpt. ,
Intel ,
Intel Press ,
2014
- [17] *Autonomous Vehicles: Handling Over Control: Opportunities and Risks for Insurance* ,
Gillian Yeomans ,
Lloyd's ,
Lloyd's Press ,
2014

- [18] *Autonomous Vehicles: Technical and Legal Challenges: An Overview of the State of the Art in Autonomous Vehicle Technology and Policy* ,
William Covington ,
Autonomous Vehicles Team, Technology Law & Public Policy Clinic ,
Technology Law & Public Policy Clinic Press ,
2013
- [19] *Autonomous Vehicle Implementation Predictions: Implications for Transport Planning* ,
Todd Litman ,
Victoria Transport Policy Institute ,
Victoria Transport Policy Institute Press ,
2015
- [20] *The Costs of Self-Driving Cars: Reconciling Freedom and Privacy with Tort Liability in Autonomous Vehicle Regulation* ,
Jack Boeglin ,
Law School, Yale University ,
Yale University Press ,
2015
- [21] *Control & Electronic News Journal: September 2014* ,
Shaun Howell ,
Ricardo Knowledge ,
Ricardo Knowledge Press ,
2014
- [22] *Learning to Drive: Perception for Autonomous Cars* ,
David Michael Stavens ,
Artificial Intelligence Laboratory, Stanford University ,
Doctoral Dissertation ,
2011
- [23] *Self-Driving Vehicles in Logistics* ,
Matthias Heutger ,
DHL Customer Solutions & Innovation ,
DHL Customer Solutions & Innovation Press ,
2014
- [24] *Petit Self-Driving And Connected Cars Fooling Sensors And Tracking Drivers* ,
Jonathan Petit, Bas Stottelaar, Michael Feiri, Frank Kargl ,
Security Innovation Wilmington, MA, United States | Services, Cybersecurity and Safety, University of Twente, The Netherlands | Institute of Distributed Systems, University of Ulm, Germany ,
Semantic Scholar ,
2015
- [25] *Self-Driving Cars: Diffusion of Radical Innovations and Technology Acceptance* ,
Hannes Enqvist ,
Faculty of Engineering, Lund University ,
Master's Degree Thesis ,
2014

- [26] *Google Self-Driving Car Project* ,
Sarah Hunter ,
Google Corporation ,
Google Corporation Press ,
2016
- [27] *Let the Cars Drive Themselves* ,
Christopher Schultz ,
University of Gothenburg ,
Bachelor's Degree Thesis ,
2015
- [28] *Obstacle Detection for Self-Driving Cars Using Only Monocular Cameras and Wheel Odometry* ,
Christian Hane, Torsten Sattler, Marc Pollefeys ,
ETH University, Zürich ,
IEEE ,
2015
- [29] *How Google's Self-Driving Car Works* ,
Erico Guizzo ,
IEEE Blogs ,
IEEE ,
2011
- [30] *Robust Vehicle Localization in Urban Environments Using Probabilistic Maps* ,
Jesse Levinson, Sebastian Thrun ,
Artificial Intelligence Laboratory, Stanford University ,
IEEE ,
2010
- [31] *Traffic Light Mapping, Localization, and State Detection for Autonomous Vehicles* ,
Jesse Levinson, Jake Askeland, Jennifer Dolson, Sebastian Thrun ,
Artificial Intelligence Laboratory, Stanford University ,
IEEE ,
2011
- [32] *Towards 3D Object Recognition via Classification of Arbitrary Object Tracks* ,
Alex Teichman, Jesse Levinson, Sebastian Thrun ,
Artificial Intelligence Laboratory, Stanford University ,
IEEE ,
2011
- [33] *A Probabilistic Framework for Car Detection in Images using Context and Scale* ,
David Held, Jesse Levinson and Sebastian Thrun ,
Artificial Intelligence Laboratory, Stanford University ,
IEEE ,
2012

- [34] *Precision Tracking with Sparse 3D and Dense Color 2D Data* ,
David Held, Jesse Levinson, and Sebastian Thrun ,
Artificial Intelligence Laboratory, Stanford University ,
IEEE ,
2013
- [35] *Sharing the Road: Autonomous Vehicles Meet Human Drivers* ,
Kurt Dresner, Peter Stone ,
Department of Computer Sciences, The University of Texas at Austin ,
IJCAI ,
2007
- [36] *Exponential Family Sparse Coding with Applications to Self-taught Learning* ,
Honglak Lee, Rajat Raina, Alex Teichman, Andrew Y. Ng ,
Computer Science Department, Stanford University ,
IJCAI ,
2009
- [37] *Going Driveless with Sensors* ,
Geetinderkaur, Sourabh Joshi, Jaspreet Kaur, Ssmreet Kaur ,
Department of Computer Science and Engineering , CT Institute of Technology & Research
IJSET ,
2014
- [38] *Autonomous Vehicle: Obstacle Detection and Decision-Based Navigation* ,
Medha Kalelkar, Anand Kelkar, Shashidhar Pamarthi ,
Computer Engineering School, Vidyalankar Institute of Technology, Mumbai, India ,
IJSRP ,
2013
- [39] *Autonomous Cars and Society* ,
Alex Forrest, Mustafa Konca ,
Department of Social Science and Policy Studies, Worcester Polytechnic Institute, Worcester, MA ,
Bachelor's Degree Thesis ,
2007
- [40] *Group Induction* ,
Alex Teichman, Sebastian Thrun ,
Artificial Intelligence Laboratory, Stanford University ,
Semantic Scholar ,
2014
- [41] *Using Artificial Intelligence to Create a Low Cost Self-Driving Car* ,
Ionut Alexandru Budisteanu ,
University of Bucharest ,
ISEF ,
2013

- [42] *Unsupervised Calibration for Multi-beam Lasers* ,
Jesse Levinson, Sebastian Thrun ,
Artificial Intelligence Laboratory, Stanford University ,
Springer ,
2010
- [43] *Σχεδιασμός Διαδρομής και Έλεγχος Κίνησης για Ρομποτικό Αυτοκίνητο* ,
Εμμανουήλ Βαρδάκης ,
Τμήμα Ηλεκτρονικών Μηχανικών και Μηχανικών Υπολογιστών, Πολυτεχνείο Κρήτης ,
Διπλωματική Εργασία ,
2009
- [44] *Driving the Self-Driving Vehicle: Expanding the Technological Design Horizon* ,
Pascale L. Blyth, Milos N. Mladenovic, Bonnie A. Nardi, Norman M. Su, Hamid R. Ekbia ,
IEEE ,
2015
- [45] *Towards Fully Autonomous Driving: Systems and Algorithms* ,
Jesse Levinson, Jake Askeland, Jan Becker, Jennifer Dolson, David Held, Soeren Kammel, J. Zico Kolter,
Dirk Langer, Oliver Pink, Vaughan Pratt, Michael Sokolsky, Ganymed Stanek, David Stavens, Alex
Teichman, Moritz Werling, and Sebastian Thrun ,
Artificial Intelligence Laboratory, Stanford University ,
IEEE ,
2011
- [46] *Automatic Laser Calibration, Mapping, and Localization for Autonomous Vehicles* ,
Jesse Sol Levinson ,
Artificial Intelligence Laboratory, Stanford University ,
Stanford University Press ,
2011
- [47] *New Challenges For Microsystems-Technology In Automotive Applications* ,
Sven Krueger, Roger Grace ,
MST News ,
MST News ,
2001
- [48] *Micro-Technology for Positioning, Navigation, and Timing Towards PNT Everywhere and Always* ,
Robert Lutwak ,
Microsystems Technology Office, Defense Advanced Research Projects Agency, DARPA ,
DARPA Press ,
2014
- [49] *Χαρτογράφηση Χώρου με Αυτοκινούμενο Ρομπότ* ,
Μανιάς Νίκος ,
Τμήμα Μηχανικών Πληροφορικής, ΑΤΕΙ Κρήτης ,
Πτυχιακή Εργασία ,
2010

- [50] *Mobility & Energy Futures Series: Self-Driving Cars* ,
Zia Wadud, Greg Marsden ,
University of Leeds ,
University of Leeds Press ,
2014
- [51] *Ευφυή Συστήματα Μεταφορών* ,
Μπλέτσος Φώτης ,
Τμήμα Πολιτικών Έργων Υποδομής, ΑΤΕΙ Θεσσαλονίκης ,
Πτυχιακή Εργασία ,
2013
- [52] *Navigation and Control of an Autonomous Vehicle* ,
Ian Schworer ,
Virginia Polytechnic Institute | State University ,
Master's Degree Thesis ,
2005
- [53] *Self-Driving Cars: Mapping Access to a Technology Revolution* ,
Jeff Rosen Chairperson ,
National Council of Disability ,
NCD Press ,
2015
- [54] *Autonomous Cars: Self-Driving the New Auto Industry Paradigm* ,
Ravi Shanker, Adam Jonas, Scott Devitt, Katy Huberty, Simon Flannery, William Greene, Benjamin Swinburne, Gregory Locraft, Adam Wood, Keith Weiss, Joseph Moore, Andrew Schenker, Paresh Jain, Yejay Ying, Shinji Kakiuchi, Ryosuke Hoshino, Andrew Humphrey ,
Morgan Stanley & Co ,
Morgan Stanley & Co Press ,
2013
- [55] *Combining 3D Shape, Color, and Motion for Robust Anytime Tracking* ,
David Held, Jesse Levinson, Sebastian Thrun, Silvio Savarese ,
Artificial Intelligence Laboratory, Stanford University ,
Stanford University Press ,
2014
- [56] *Programming a Self-Driving Car* ,
Onishim Hasdak ,
Department of Computer Science & Engineering, School of Engineering & Computer Science, Brac University ,
Bachelor's Degree Thesis ,
2015
- [57] *Autonomous Vehicle Technology* ,
James M. Anderson, Nidhi Kalra, Karlyn D. Stanley, Paul Sorensen, Constantine Samaras, Oluwatobi A. Oluwatola ,
RAND Corporation ,
RAND Corporation Press ,
2016

- [58] *Google Self-Driving Car Project Monthly Report 5/15* ,
Google Team ,
Google Corporation ,
Google Corporation Press ,
2015
- [59] *Google Self-Driving Car Project Monthly Report 10/15* ,
Google Team ,
Google Corporation ,
Google Corporation Press ,
2015
- [60] *Google Self-Driving Car Project Monthly Report 11/15* ,
Google Team ,
Google Corporation ,
Google Corporation Press ,
2015
- [61] *Map-Based Precision Vehicle Localization in Urban Environments* ,
Jesse Levinson, Michael Montemerlo, Sebastian Thrun ,
Artificial Intelligence Laboratory, Stanford University ,
IEEE ,
2015
- [62] *Tracking-Based Semi-Supervised Learning* ,
Alex Teichman, Sebastian Thrun ,
Department of Computer, Science Stanford University ,
Stanford University Press ,
2011
- [63] *Online, semi-supervised learning for long-term interaction with object recognition systems* ,
Alex Teichman and Sebastian Thrun ,
Department of Computer, Science Stanford University ,
Stanford University Press ,
2012
- [64] *Automatic Online Calibration of Cameras and Lasers* ,
Jesse Levinson, Sebastian Thrun ,
Artificial Intelligence Laboratory, Stanford University ,
Stanford University Press ,
2013
- [65] *Self-Driving Vehicles: Current Status of Autonomous Vehicle Development and Minnesota Policy Implications* ,
Adeel Lari, Frank Douma, Ify Onyiah ,
University of Minnesota ,
University of Minnesota Press ,
2014

[66] *Self Driving Cars* ,
Ratan Hudda, Clint Kelly, Garrett Long, Jun Luo, Atul Pandit, Dave Phillips, Lubab Sheet, Ikhlaz Sidhu ,
College of Engineering, University of California, Berkeley ,
University of California Press ,
2013

[67] *Self-driving cars* ,
Hari Chittilla, Dennis Sun ,
Computer Science Department, University of North Carolina ,
University of North Carolina Press ,
2015

[68] *Self-Driving Cars: Disruptive or Incremental?* ,
Tao Jiang, Srdjan Petrovic, Uma Ayyer, Anand Tolani, Anand Tolani ,
Berkley University ,
Berkley University Press ,
2015

[69] *Self-Driving Cars: A Case Study in Making New Markets* ,
Luis Araujo, Katy Mason and Martin Spring ,
Lancaster University ,
Lancaster University Press ,
2012

[70] *Self-Driving Cars: Are We Ready?* ,
Gary Silberg, Michael Corley, Mitch Manassa, Hugh Fraser, Kevin Everhart, Vivek Sinha, Deepak Subramanian ,
KPMG LLP ,
KPMG LLP Press ,
2013

[71] *Self-Driving Cars: The Next Revolution* ,
Gary Silberg, Jos Plessers, Richard Wallace, Chris Brower, Gary Matuszak, Deepak Subramanian ,
KPMG LLP | CarGroup ,
KPMG LLP Press ,
2012

[72] *Removing Roadblocks to Intelligent Vehicles and Driverless Cars* ,
Adam Thierer, Ryan Hagemann ,
Mercatus Center, George Mason University ,
Mercatus Working Paper ,
2014

[73] *Development of Predictive Vehicle & Drivetrain Operating Strategies Based Upon Advanced Information & Communication Technologies* ,
Stephen Jones, Arno Huss, Emre Kural, Alexander Massoner, Edoardo Morra, Christa Simon, Reinhard Tatschl, Christian Vock ,
AVL List GmbH ,
Transport Research Arena ,
2014

- [74] *Transorming Personal Mobility* ,
Lawrence D. Burns, William C. Jordan, Bonnie A. Scarborough ,
The Earth Institute, Columbia University ,
Columbia University Press ,
2013
- [75] *Revolutionizing Mobility: Connected and Self-Driving Vehicles* ,
Richard Wallace ,
CarGroup ,
Health Effects Institute Annual Conference ,
2014
- [76] *Self Driving Cars: Future Has Already Begun* ,
Muhammad Azmat, Clemens Schuhmayer ,
Vienna University Of Economics and Business ,
Vienna University Press ,
2015
- [77] *Multi-Agent Rendezvous on Street Networks* ,
Malika Meghjani, Gregory Dudek ,
Semantic Scholar ,
IEEE ,
2014
- [78] *Self Driving Car* ,
Sameh Hamzawy ,
University of Wollongong ,
University of Wollongong Press ,
2013
- [79] *Reachability Analysis and its Application to the Safety Assessment of Autonomous Cars* ,
Matthias Althoff ,
Technical University of Munich ,
Doctoral Dissertation ,
2010
- [80] *Autonomous Car Driving using a Low-Cost On-Board Computer* ,
Brian Pfretzschner ,
Intelligent Autonomous Systems Lab, Technical University of Darmstadt ,
Bachelor's Degree Thesis ,
2013
- [81] *Ραντάρ* ,
Ζαχαρίας Τσουκαλάς ,
Ανώτερη Δημόσια Σχολή Εμπορικού Ναυτικού ,
Ίδρυμα Ευγενίδου ,
2006

- [82] *Τεχνητή Νοημοσύνη* ,
Ιωάννης Ραφανίδης ,
Τμήμα Εφαρμοσμένης Πληροφορικής, Πανεπιστήμιο Μακεδονίας ,
Σύγγραμμα Πανεπιστημίου Μακεδονίας ,
2015
- [83] *Study of a Shared Autonomous Vehicles Based Mobility Solution in Stockholm* ,
Pierre-Jean Rigole ,
Royal Institute of Technology ,
Master's Degree Thesis ,
2014
- [84] *Τι είναι η Τεχνητή Νοημοσύνη* ,
Μανόλης Κουμπάρκης ,
Τμήμα Πληροφορικής και Επικοινωνιών, Πανεπιστήμιο Αθήνας ,
Σύγγραμμα Πανεπιστημίου Αθήνας ,
2011
- [85] *LIDAR-based 3D Object Perception* ,
M. Himmelsbach, A. Muller, T. Luttel, H.-J. Wunsche ,
Semantic Scholar ,
Semantic Scholar Press ,
2008
- [86] *Policy-Gradient Algorithms for Partially Observable Markov Decision Processes* ,
Aberdeen, D. ,
Australian National University ,
PhD thesis ,
2003
- [87] *Learning algorithms for Markov decision processes with average cost* ,
Abounadi, J., Bertsekas, D., Borkar, V. S. ,
SIAM ,
Control and Optimization Journal ,
2002
- [88] *Steps toward artificial intelligence* ,
Minsky, M. ,
McGraw-Hill ,
Computers and Thought ,
1963
- [89] *A logical calculus of the ideas immanent in nervous activity* ,
McCulloch, W., Pitts, W. ,
Bulletin of Mathematical Biophysics ,
1943
- [90] *Principles of risk minimization for learning theory* ,
Vapnik, V. ,
Morgan Kaufmann. ,
Neural Information Processing Systems ,
1992

- [91] *Advances in Kernel Methods - Support Vector Learning* ,
Scholkopf, B., Burges, C. J. C., Smola, A. J. ,
MIT Press ,
1998
- [92] *ImageNet Large Scale Visual Recognition Challenge* ,
ILSVRC2015 ,
2015
- [93] *Mitosis Detection in Breast Cancer Histology Images using Deep Neural Networks* ,
Dan Claudiu Ciresan, Alessandro Giusti, Luca Maria Gambardella, Jürgen Schmidhuber ,
MICCAI ,
2013
- [94] *Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images* ,
D. Ciresan, A. Giusti, L. Gambardella, J. Schmidhuber ,
Advances in Neural Information Processing Systems ,
2012
- [95] *Segmentation of neuronal structures in EM stacks* ,
ISBI ,
2012
- [96] *Turing computability with neural nets* ,
Siegelmann, H. T., Sontag, E. D. ,
Applied Mathematics Letters ,1991
- [97] *Dynamic neural nets and the fundamental spatio-temporal credit assignment problem* ,
Schmidhuber, J. ,
Technical University of Munich ,
PhD Dissertation ,
1990
- [98] *PROW: a step toward automatic program writing* ,
Waldinger, R. J., Lee, R. C. T. ,
Morgan Kaufmann ,
International Joint Conference on Artificial Intelligence ,
1969
- [99] *A 15 year perspective on automatic programming* ,
Balzer, R. ,
IEEE Transactions on Software Engineering ,
1985 ,
- [100] *Learning to program = learning to construct mechanisms and explanations* ,
Soloway, E. ,
ACM ,
1986 ,

- [101] *Logic program synthesis* ,
Deville, Y., Lau, K. K. ,
ACM ,
1994
- [102] *Reinforcement learning: a survey* ,
Kaelbling, L. P., Littman, M. L., Moore, A.W. ,
Journal of AI research ,
1996
- [103] *Reinforcement learning: An introduction* ,
Sutton, R., Barto, A. ,
MIT Press ,
1998
- [104] *Reinforcement Learning* ,
Wiering, M., van Otterlo, M. ,
Springer ,
2012
- [105] *Universal Artificial Intelligence: Sequential Decisions based on Algorithmic Probability* ,
Hutter, M., ,
J. Schmidhuber's SNF grant ,
Springer ,
2005
- [106] *Universal sequential search problems* Levin, L. A. ,
Problems of Information Transmission ,
1973
- [107] *The complexity of theorem-proving procedures* ,
Cook, S. A. ,
Annual ACM Symposium on the Theory of Computing ,
1971
- [108] *A survey of computational complexity results in systems and control* ,
Blondel, V. D., Tsitsiklis, J. N. ,
Automatica ,
2000
- [109] *On the undecidability of probabilistic planning and related stochastic optimization problem* ,
Madani, O., Hanks, S., Condon, A. ,
Artificial Intelligence ,
2003
- [110] *On the computational complexity of stochastic controller optimization in POMDPs* ,
Vlassis, N., Littman, M. L., Barber, D. ,
ACM Transactions on Computation Theory ,
2012

- [111] *Neural networks for control and system identification* ,
Werbos, P. J. ,
IEEE ,
1989
- [112] *A dual back-propagation scheme for scalar reinforcement learning* ,
Munro, P.W. ,
Conference of the Cognitive Science Society ,
1987
- [113] *Dynamic reinforcement driven error propagation networks with application to game playing* ,
Robinson, T., Fallside, F. ,
Conference of the Cognitive Science Society ,
1989
- [114] *An on-line algorithm for dynamic reinforcement learning and planning in reactive environments* ,
Schmidhuber, J. ,
International Joint Conference on Neural Networks ,
IEEE
1990
- [115] *Identification and control of dynamical systems using neural networks* ,
Narendra, K. S., Parthasarathy, K. ,
IEEE ,
1990
- [116] *Neural networks for optimization and signal processing* ,
Cochocki, A., Unbehauen, R. ,
John Wiley & Sons, Inc. ,
1993
- [117] *Control of nonlinear dynamical systems using neural networks* ,
Levin, A. U. and Narendra, K. S. ,
IEEE ,
1995
- [118] *System identification* ,
Ljung, L. ,
Springer ,
1998
- [119] *Dynamical neural networks for control* ,
Prokhorov, D., Puskorius, G., Feldkamp, L. ,
A field guide to dynamical recurrent networks ,
IEEE
2001
- [120] *Stable adaptive neural network control* ,
Ge, S., Hang, C. C., Lee, T. H., Zhang, T. ,
Springer ,
2010

- [121] *The truck backer-upper: An example of self learning in neural networks* ,
Nguyen, N., Widrow, B. ,
International Joint Conference on Neural Networks ,
IEEE ,
1989
- [122] *Learning to generate artificial fovea trajectories for target detection* ,
Schmidhuber, J., Huber, R. ,
International Journal of Neural Systems ,
1991
- [123] *Reinforcement Learning for the adaptive control of perception and action* ,
Whitehead, S. ,
University of Rochester ,
PhD thesis ,
1992
- [124] *A signal processing framework based on dynamic neural networks with application to problems in adaptation, filtering, and classification* ,
Feldkamp, L. A., Puskorius, G. V. ,
IEEE ,
1998
- [125] *FU-Fighters Small Size 2004, Team Description* ,
Egorova, A., Glove, A., G`oktekin, C., Liers, A., Luft, M., Rojas, R., Simon, M., Tenchio, O., Wiesel, F. ,
RoboCup Symposium ,
2004
- [126] *RoboCup 2004: Robot Soccer World Cup VIII, Lecture Notes in Artificial Intelligence* ,
Nardi, D., Riedmiller, M., Sammut, C., Santos-Victor, J. ,
Springer ,
2005
- [127] *Reinforcing the driving quality of soccer playing robots by anticipation* ,
Glove, A., Wiesel, F., Tenchio, O., Simon, M. ,
Information Technology ,
2005
- [128] *Prototype resilient, self-modeling robots* ,
Schmidhuber, J. ,
Science ,
2007
- [129] *My first Deep Learning system of 1991 + Deep Learning timeline 1962-2013* ,
Schmidhuber, J. ,
The Swiss AI Lab IDSIA ,
arXiv ,
2013

[130] *POWERPLAY: Training an Increasingly General Problem Solver by Continually Searching for the Simplest Still Unsolvable Problem* ,

Schmidhuber, J. ,

Psychology ,

2013

[131] *Intrinsically motivated learning of real world sensorimotor skills with developmental constraint* ,

Oudeyer, P.-Y., Baranes, A., Kaplan, F. ,

Intrinsically Motivated Learning in Natural and Artificial Systems ,

Springer

2013

[132] *Neuro-dynamic Programming* ,

Bertsekas, D. P., Tsitsiklis, J. N. ,

Athena Scientific ,

1996

[133] *GQ(_): A general gradient algorithm for temporal-difference prediction learning with eligibility traces* ,

Maei, H. R., Sutton, R. S. ,

Conference on Artificial General Intelligence ,

2010

[134] *Reinforcement learning in continuous state and action spaces* ,

Hasselt, H. ,

Reinforcement Learning ,

Springer

2012

[135] *Playing Atari with deep reinforcement learning* ,

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., Riedmiller, M ,

Deepmind Technologies ,

arXiv

2013

[136] *Deep learning for real-time Atari game play using offline Monte-Carlo tree search planning* Guo ,

X., Singh, S., Lee, H., Lewis, R., Wang, X. ,

Neural Information Processing Systems ,

2014

[137] *Multi-Dimensional Deep Memory Atari-Go Players for Parameter Exploring Policy Gradients* ,

Gruttner, M., Sehnke, F., Schaul, T., Schmidhuber, J. ,

International Conference on Artificial Neural Networks ,

Springer

2010

[138] *Evolving large-scale neural networks for vision-based reinforcement learning* ,

Koutník, J., Cuccu, G., Schmidhuber, J., Gomez, F. ,

Genetic and Evolutionary Computation Conference ,

ACM ,

2013

- [139] *Pedestrian detection: An evaluation of the state of the art* ,
Dollar, P., Wojek, C., Schiele, B., Perona, P. ,
Pattern Analysis and Machine Intelligence ,
IEEE
2012
- [140] *Fast feature pyramids for object detection* ,
Dollar, P., Appel, R., Belongie, S., Perona, P. ,
Pattern Analysis and Machine Intelligence ,
IEEE ,
2014
- [141] *Visual tracking: An experimental survey* ,
Smeulders, A. W., D. M. Chu, R. Cucchiara, S. Calderara, A. Dehghan, M. Shah ,
Pattern Analysis and Machine Intelligence ,
IEEE
2014
- [142] *3d convolutional neural networks for human action recognition* ,
Ji, S., W. Xu, M. Yang, K. Yu ,
Pattern Analysis and Machine Intelligence ,
IEEE ,
2013
- [143] *Cnntracker: Online discriminative object tracking via deep convolutional neural network* ,
Chen, Y., X. Yang, B. Zhong, S. Pan, D. Chen, H. Zhang ,
Applied Soft Computing ,
2016
- [144] *Boosting object detection performance in crowded surveillance videos* ,
Feris, R., A. Datta, S. Pankanti, M. T. Sun ,
Workshop on Applications of Computer Vision ,
IEEE ,
2013
- [145] *Faster r-cnn: Towards real-time object detection with region proposal networks* ,
Ren, S., K. He, R. Girshick, J. Sun ,
Advances in Neural Information Processing Systems ,
Curran Associates Inc. ,
2015
- [146] *Tracking with deep neural networks* ,
Jin, J., A. Dundar, J. Bates, C. Farabet, E. Culurciello ,
Annual Conference on Information Sciences and Systems ,
2013
- [147] *Decaf: A deep convolutional activation feature for generic visual recognition* ,
Donahue, J., Y. Jia, O. Vinyals, J. Ho_man, N. Zhang, E. Tzeng, T. Darrell ,
International Conference on Machine Learning ,
2014

- [148] *Imagenet classification with deep convolutional neural networks* ,
Krizhevsky, A., I. Sutskever, G. E. Hinton ,
Advances in Neural Information Processing Systems ,
Curran Associates, Inc. ,
2012
- [149] *Gradient-based learning applied to document recognition* ,
Lecun, Y., L. Bottou, Y. Bengio, P. Haffner ,
IEEE ,
1998
- [150] *Tracking with deep neural networks* ,
Jin, J., A. Dundar, J. Bates, C. Farabet, E. Culurciell ,
Annual Conference on Information Sciences and Systems ,
2013
- [151] *Transferring rich feature hierarchies for robust visual tracking* ,
Wang, N., S. Li, A. Gupta, D. Yeung ,
Computing Research Repository ,
2015
- [152] *Learning a deep compact image representation for visual tracking* ,
Wang, N., D. Y. Yeung ,
Advances in Neural Information Processing Systems ,
Curran Associates, Inc.
2013
- [153] *Learning multi-domain convolutional neural networks for visual tracking* ,
Nam, H., B. Han ,
Computing Research Repository ,
2015
- [154] *Deeptrack: Learning discriminative feature representations online for robust visual tracking* ,
Li, H., Y. Li, F. Porikli ,
Transactions on Image Processing ,
IEEE
2016
- [155] *A minimum description length framework for unsupervised learning* ,
Zemel, R. S. ,
University of Toronto ,
PhD thesis ,
1993
- [156] *Developing population codes by minimizing description length* ,
Zemel, R. S., Hinton, G. E. ,
Morgan Kaufmann Press ,
1994

- [157] *Discrete recurrent neural networks for grammatical inference* ,
Zeng, Z., Goodman, R., Smyth, P. ,
Transactions on Neural Networks ,
IEEE ,
1994
- [158] *Forecasting with recurrent neural networks: 12 tricks* ,
Zimmermann, H.-G., Tietz, C., Grothmann, R. ,
2012
- [159] *Tricks of the Trade (2nd ed.)* ,
Lecture Notes in Computer Science ,
Springer ,
2012
- [160] *A spiking network model of short-term active memory* ,
Zipser, D., Kehoe, B., Littlewort, G., Fuster, J. ,
The Journal of Neuroscience ,
1993
- [161] *Rich feature hierarchies for accurate object detection and semantic segmentation
Tech Report (v5)* ,
Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik ,
UC Berkeley ,
UC Berkeley Press ,
2014
- [162] *Fast R-CNN* ,
Ross Girshick ,
Microsoft Research ,
Microsoft Research Press ,
2015
- [163] *Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks* ,
Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun ,
Microsoft Research ,
Microsoft Research Press ,
2016
- [164] *Is Faster R-CNN Doing Well for Pedestrian Detection?* ,
Liliang Zhang, Liang Lin, Xiaodan Liang, Kaiming He ,
School of Data and Computer Science, Sun Yat-sen University ,
arXiv ,
2016
- [165] *A Unified Multi-scale Deep Convolutional Neural Network for Fast Object Detection* ,
Zhaowei Cai, Quanfu Fan, Rogerio S. Feris, Nuno Vasconcelos ,
SVCL, UC San Diego | IBM T. J. Watson Research ,
arXiv ,
2016

[166] *Fused DNN: A deep neural network fusion approach to fast and robust pedestrian detection* ,
Xianzhi Du, Mostafa El-Khamy, Jungwon Lee, Larry S. Davis ,
Computer Vision Laboratory, UMIACS, University of Maryland | Modern Systems R&D, Samsung
Electronics ,
arXiv ,
2017

[167] *A Closer Look at Faster R-CNN for Vehicle Detection* ,
Quanfu Fan, Lisa Brown, John Smith ,
Intelligent Vehicles Symposium ,
IEEE ,
2016

[168] *Scale-Aware Fast R-CNN for Pedestrian Detection* ,
Jianan Li, Xiaodan Liang, Shengmei Shen, Tingfa Xu, Jiashi Feng, Shuicheng Yan ,
Transactions of Multimedia ,
IEEE ,
2018

[169] *CS231n: Convolutional Neural Networks for Visual Recognition* ,
Fei-Fei Li, Justin Johnson, Serena Yeung ,
Lecture Notes and Slides ,
Stanford University Press ,
2018

[170] *CS131: Computer Vision: Foundations and Applications* ,
Juan Carlos Niebles, Ranjay Krishna ,
Lecture Notes and Slides ,
Stanford University Press ,
2017

[171] *CS231a: Computer Vision, from 3D Reconstruction to Recognition* ,
Silvio Savarese ,
Lecture Notes and Slides ,
Stanford University Press ,
2018

[172] *CS230: Deep Learning* ,
Andrew Ng, Kian Katanforoosh ,
Lecture Notes and Slides ,
Stanford University Press ,
2018

[173] *Feature Pyramid Networks for Object Detection* ,
Tsung-Yi Lin, Piotr Dollar, Ross Girshick, Kaiming He, Bharath Hariharan, Serge Belongie ,
Facebook AI Research | Cornell University | Cornell Tech ,
IEEE ,
2017

- [174] *You Only Look Once: Unified, Real-Time Object Detection* ,
Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi ,
University of Washington | Allen Institute for AI | Facebook AI Research ,
IEEE ,
2016
- [175] *Faster R-CNN Features for Instance Search* ,
Amaia Salvador, Xavier Giro-i-Nieto, Ferran Marques, Shinichi Satoh ,
Universitat Politecnica de Catalunya | National Institute of Informatics ,
IEEE ,
2016
- [176] *End-to-end people detection in crowded scenes* ,
Russell Stewart, Mykhaylo Andriluka, Andrew Y. Ng ,
Stanford University | Max Planck Institute for Informatics ,
IEEE ,
2016
- [177] *How Far are We from Solving Pedestrian Detection?* ,
Shanshan Zhang, Rodrigo Benenson, Mohamed Omran, Jan Hosang, Bernt Schiele ,
Max Planck Institute for Informatics ,
IEEE ,
2016
- [178] *Reinforcement Learning I: Introduction Course Notes* ,
Richard S. Sutton, Andrew G. Barto ,
Stanford University ,
Stanford University Press ,
2014
- [179] *Application of Convoluted Neural Networks for Pedestrian Detection* ,
Anil Variyar ,
Stanford University ,
IEEE ,
2016
- [180] *Deep convolutional neural networks for pedestrian detection* ,
D. Tome, F. Monti, L. Barofio, L. Bondi, M. Tagliasacchi, S. Tubaro ,
Politecnico di Milano ,
arXiv ,
2016
- [181] *Two Parallel Deep Convolutional Neural Networks for Pedestrian Detection* ,
Bo-Yao Lin, Chu-Song Chen ,
Institute of Information Science Academia Sinica | Institute of Information Science & Research Center
for Information Technology Innovation Academia Sinica ,
IEEE ,
2015

- [182] *Real-Time Pedestrian Detection With Deep Network Cascades* ,
Anelia Angelova, Alex Krizhevsky, Vincent Vanhoucke, Abhijit Ogale, Dave Ferguson ,
Google Research ,
Google Research Press ,
2015
- [183] *On the use of Convolutional Neural Networks for Pedestrian Detection* ,
Sergi Canyameres Masip ,
University of Barcelona ,
University of Barcelona Press ,
2015
- [184] *Markov Decision Processes: Concepts and Algorithms* ,
Martijn van Otterlo ,
Computer Science School, Kuleuven University ,
Kuleuven University Press ,
2009
- [185] *Pedestrian Detection Based on Deep Convolutional Neural Network with Ensemble Inference Network* ,
Hiroshi Fukui, Takayoshi Yamashita, Yuji Yamauchi, Hironobu Fujiyoshi, Hiroshi Murase ,
Chubu University | Nagoya University ,
IEEE ,
2015
- [186] *Pedestrian Detection with Deep Convolutional Neural Network* ,
Xiaogang Chen, Pengxu Wei, Wei Ke, Qixiang Ye, Jianbin Jiao ,
School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Science ,
University of Chinese Academy of Science Press ,
2014
- [187] *Learning to Drive: Perception for Autonomous Cars a Dissertation* ,
David Michael Stavens ,
Stanford University ,
Stanford University Press ,
2011
- [188] *Pedestrian Detection with Multichannel Convolutional Neural Networks* ,
David José Lopes, Brito Duarte Ribeiro ,
Electrical and Computer Engineering School, Tecnico Lisboa University ,
Master's Degree Thesis ,
2015
- [189] *The development of autonomous vehicles* ,
Filip Hucko ,
Aalborg University ,
Master's Degree Thesis ,
2017

- [190] *A Survey of Applications of Markov Decision Processes* ,
D. J. White ,
Operational Research Society ,
Palgrave Macmillan Journal ,
1993
- [191] *Multispectral Deep Neural Networks for Pedestrian Detection* ,
Jingjing Liu, Shaoting Zhang, Shu Wang, Dimitris N. Metaxas ,
Department of Computer Science, Rutgers University | Department of Computer Science, UNC
Charlotte ,
dx.doi.org ,
2016
- [192] *Beyond Pedestrian Detection: Deep Neural Networks Level-Up Automotive Safety* ,
Ikuro Sato, Hideki Niihara ,
Denso IT Laboratory Inc. ,
Denso Inc. Press ,
2014
- [193] *Development of Autonomous Driving Using Robot Operating System* ,
Alexandar Zivkovic ,
Polytechnic University of Madrid ,
Master's Degree Thesis ,
2018
- [194] *Object Detection Using Deep Learning - Learning where to search using visual attention* ,
Alina Kloss ,
Eberhard Karls Tübingen University ,
Master's Degree Thesis ,
2015
- [195] *Combining 3D Shape, Color, and Motion for Robust Velocity Estimation* ,
David Held, Jesse Levinson, Sebastian Thrun, Silvio Savarese ,
Stanford University ,
Stanford University Press ,
2015
- [196] *Μηχανική Μάθηση για το Συγκεκριασμό Πολλαπλών, Κατευθυνόμενων Ευφοιών Συστημάτων* ,
Γρηγόριος Τσουμάκας ,
Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, Τμήμα Πληροφορικής ,
Διδακτορική Διατριβή ,
2005
- [197] *Παρουσιάσεις Διαλέξεων Μαθήματος Τεχνητής Νοημοσύνης* ,
Εθνικό Μετσόβιο Πολυτεχνείο, Τμήμα ΗΜΜΥ ,
2003
- [198] *Πιθανοτική Ρομποτική* ,
Sebastian Thrun, Wolfram Burgard, Dieter Fox ,
Εκδόσεις Κλειδάριθμος ,
2012

- [199] *Deep Learning* ,
Ian Goodfellow, Yoshua Bengio, Aaron Courville ,
The MIT Press ,
2017
- [200] *Τεχνητή Νοημοσύνη* ,
Stuart Russell, Peter Norvig ,
Εκδόσεις Κλειδάριθμος ,
2004
- [201] *Τεχνητή Νοημοσύνη* ,
Ιωάννης Βλαχάβας, Νικόλαος Βασιλειάδης, Φώτης Κόκκορας, Ηλίας Σακελλαρίου, Πέτρος Κεφάλας ,
Εκδόσεις Πανεπιστημίου Μακεδονίας ,
2011
- [202] *Piotr Dollar Research* ,
Piotr Dollar ,
GitHub ,
<https://www.pdollar.github.io/research.html> ,
- [203] *What is driverless car* ,
TechTarget ,
<https://www.searchenterpriseai.techtarget.com/definition/driverless-car>
- [204] *History of self-driving cars*
https://www.en.wikipedia.org/wiki/History_of_self-driving_cars
- [205] *MIT 6.S094: Deep Learning for Self-Driving Cars* ,
MIT University ,
<https://www.selfdrivingcars.mit.edu>
- [206] *Machine Learning is Fun! Part 3: Deep Learning and Convolutional Neural Networks* ,
Adam Geitgey ,
Medium.com ,
<https://www.medium.com/@ageitgey/machine-learning-is-fun-part-3-deep-learning-and-convolutional-neural-networks-f40359318721>
- [207] *A Brief History of CNNs in Image Segmentation: From R-CNN to Mask R-CNN* ,
Dhruv Parthasarathy ,
Medium.com , ,
<https://blog.athelas.com/a-brief-history-of-cnns-in-image-segmentation-from-r-cnn-to-mask-r-cnn-34ea83205de4>
- [208] *R-CNN, Fast R-CNN, Faster R-CNN, YOLO — Object Detection Algorithms* ,
Rohith Gandhi ,
Medium.com , ,
<https://towardsdatascience.com/r-cnn-fast-r-cnn-faster-r-cnn-yolo-object-detection-algorithms-36d53571365e>
- [209] *Autonomous Vehicles | Self-driving vehicles enacted Legislation* ,
<http://www.ncsl.org/research/transportation/autonomous-vehicles-self-driving-vehicles-enacted-legislation.aspx>

Ευρετήριο Αρκτικόλεξων

ΔΙΕΘΝΗΣ

ACC	Adaptive Cruise Control
ACV	Amphibious Combat Vehicle
ADAS	Advanced Driver Assistance Systems
ADS	Advanced Driving System
AI	Artificial Intelligence
ALCA	Active Lane Change Assist
BP	Backpropagation Error
CAP	Credit Assignment Path
CMOS	Complimentary Metal Oxide Semiconductor
CNN	Convolutional Neural Network
ConvNet	Convolutional Neural Network
CPU	Central Processing Unit
DARPA	Defense Advanced Projects Agency
Deep NN	Deep Neural Network
DL	Deep Learning
DP	Dynamic Programming
Fast R-CNN	Fast Region-based Convolutional Neural Network
Faster R-CNN	Faster Region-based Convolutional Neural Network
FHWA	Federal Highway Administration
GNSS	Global Navigational Satellite Systems
GPS	Global Positioning System
GPU	Graphics Processing Unit
IoT	Internet of Things
LAN	Local Area Network
LCA	Lane Change Assist
LDWS	Lane Departure Warning Systems
LIDAR	Light Detection And Ranging
LKA	Lane Keeping Assist
Mask R-CNN	Mask Region-based Convolutional Neural Network
MDP	Markov Decision Processes
ML	Machine Learning
MLP	Multilayer Perceptrons
NAHSC	National Automated Highway System Consortium
NASA	National Aeronautics and Space Administration
NHTSA	National Highway Traffic Safety Administration
NN	Neural Networks

PAP	Parallel Assist Parking
RADAR	RAdio Detection And Ranging
RAM	Random Access Memory
RBM	Restricted Boltzmann Machine
R-CNN	Region-based Convolutional Neural Network
R-FCN	Region-based Fully Convolutional Networks
RL	Reinforcement Learning
RNN	Recurrent Neural Networks
RoIAlign	Region of Interest Align
RoIPool	Region of Interest Pooling
ROM	Read only Memory
RPM	Revolutions Per Minute
RPN	Region Proposal Network
RT	Real Time
SAE	Society of Automotive Engineers
SDC	Self Driving Car
SL	Supervised Learning
SPP NET	Spatial Pyramid Pooling in Deep Convolutional Networks
SSD	Single Shot Detector
SVM	Support Vector Machine
UL	Unsupervised Learning
USDOT	US Department of Transportation
V2D	Vehicle to Device
V2G	Vehicle to Grid
V2I	Vehicle to Infrastructure
V2P	Vehicle to Pedestrian
V2V	Vehicle to Vehicle
V2X	Vehicle to Everything
WLAN	Wireless Local Area Network
YOLO	You Only Loon Once

ΕΛΛΗΝΙΚΟΙ

ΑΑΥ	Αλγόριθμος Απαλοιφής Υποψηφίων
ΚΟΚ	Κώδικας Οδικής Κυκλοφορίας
ΜΔΥ	Μηχανές Διανυσμάτων Υποστήριξης
ΜΜΕ	Μέσα Μαζικής Ενημέρωσης
ΝΔ	Νευρωνικά Δίκτυα
ΣΝΔ	Συνελικτικά Νευρωνικά Δίκτυα
ΤΝΔ	Τεχνητά Νευρωνικά Δίκτυα

