

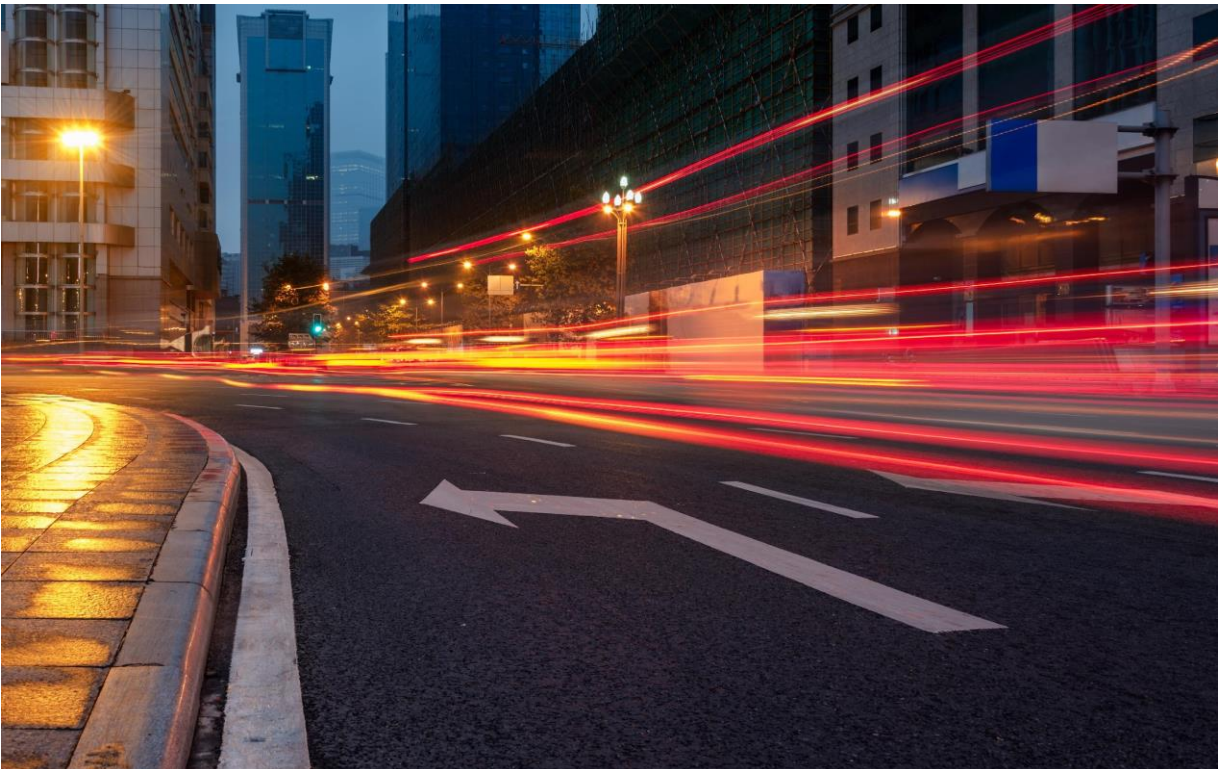


Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Πολιτικών Μηχανικών

Τομέας Μεταφορών & Συγκοινωνιακής Υποδομής

Ανάλυση ακατάλληλης ταχύτητας οχημάτων σε διαβάσεις πεζών στην Αθήνα με χρήση αλγορίθμων οπτικής αναγνώρισης



Διπλωματική Εργασία

Χατζηθεοδοσίου Ιωάννης

Αθήνα, Μάρτιος 2026

Επιβλέπων: Γιαννής Γεώργιος, Καθηγητής ΕΜΠ

Ευχαριστίες

Με την ολοκλήρωση της διπλωματικής εργασίας ευχαριστώ ειλικρινά τον επιβλέποντα καθηγητή, κ. Γεώργιο Γιαννή, για την εμπιστοσύνη που μου παρείχε αναθέτοντάς μου το συγκεκριμένο ερευνητικό αντικείμενο και για την καθοδήγησή του.

Αισθάνομαι την ανάγκη να ευχαριστήσω επίσης, την ερευνήτρια, κα Στυλιανή Ρούσσου για την αδιάκοπη στήριξη, την βοήθεια και την συμβολή της στην υλοποίηση της εργασίας.

Τέλος, ευχαριστώ την οικογένειά μου και τους φίλους μου για την ηθική υποστήριξη, την υπομονή και την ενθάρρυνση που μου προσέφεραν καθ' όλη την διάρκεια των σπουδών μου.

Αθήνα, Μάρτιος 2026

Χατζηθεοδοσίου Ιωάννης

Ανάλυση ακατάλληλης ταχύτητας οχημάτων σε διαβάσεις πεζών στην Αθήνα με χρήση αλγορίθμων οπτικής αναγνώρισης

Χατζηθεοδοσίου Ιωάννης

Επιβλέπων: Γιαννής Γεώργιος, Καθηγητής ΕΜΠ

Σύνοψη

Η οδική ασφάλεια στις αστικές περιοχές και ιδιαίτερα στις διαβάσεις πεζών αποτελεί σημαντικό πεδίο έρευνας, καθώς οι χώροι αυτοί χαρακτηρίζονται από έντονη αλληλεπίδραση μεταξύ οχημάτων και ευάλωτων χρηστών της οδού. Η παρούσα διπλωματική εργασία εξετάζει την ακατάλληλη ταχύτητα οχημάτων σε σηματοδοτούμενες διαβάσεις πεζών στην Αθήνα, με αξιοποίηση δεδομένων βιντεοσκόπησης και τεχνικών υπολογιστικής όρασης. Τα δεδομένα συλλέχθηκαν σε διασταύρωση υψηλής κυκλοφορίας στην Λεωφόρο Βασιλέως Κωνσταντίνου, μέσω δύο παράλληλων μεθόδων: επιτόπιας χειροκίνητης παρατήρησης και ενός συστήματος υπολογιστικής όρασης που συνδυάζει τους αλγόριθμους YOLOv8, ResNet-50, ομογραφικό μετασχηματισμό και φιλτράρισμα Kalman. Μέσω της επεξεργασίας των δεδομένων εξάγονται μεταβλητές κίνησης και δείκτες οδικής ασφάλειας, όπως ο χρόνος μέχρι σύγκρουση (Time to Collision – TTC) και το Post-Encroachment Time (PET), ώστε να διερευνηθεί η σχέση τους με τη συμπεριφορά των οδηγών. Στη συνέχεια εφαρμόζονται στατιστικές μέθοδοι και αναλύσεις γραμμικής παλινδρόμησης για τον εντοπισμό των παραγόντων που σχετίζονται με την εμφάνιση ακατάλληλης ταχύτητας. Στόχος της εργασίας είναι η καλύτερη κατανόηση της οδηγικής συμπεριφοράς κοντά σε διαβάσεις πεζών και η συμβολή στην αξιολόγηση της οδικής ασφάλειας μέσω σύγχρονων μεθόδων ανάλυσης. Τα αποτελέσματα μπορούν να υποστηρίξουν τον σχεδιασμό παρεμβάσεων για τη βελτίωση της ασφάλειας των πεζών στο αστικό περιβάλλον.

Λέξεις-κλειδιά: Παράνομες διελεύσεις οχημάτων, Υπολογιστική Όραση, YOLO-v8, Λογιστική Παλινδρόμηση, Γενικευμένα Γραμμικά Μοντέλα

Analyzing speeding vehicle behavior at crosswalks in Athens using computer vision

Chatzitheodosiou Ioannis

Supervisor: George Yannis, Professor NTUA

Abstract

Road safety in urban areas, and especially at pedestrian crossings, constitutes an important field of research, as these spaces are characterized by intense interaction between vehicles and vulnerable road users. This diploma thesis examines inappropriate vehicle speed at signalized pedestrian crossings in Athens, using video data and computer vision techniques. The data were collected at a high-traffic intersection in Leoforos Vassileos Konstantinou through two parallel methods: on-site manual observation and a computer vision system combining YOLOv8, ResNet-50, homographic transformation, and Kalman filtering. Through data processing, motion variables and road safety indicators are extracted, such as Time to Collision (TTC) and Post-Encroachment Time (PET), in order to investigate their relationship with driver behavior. Statistical methods and regression models are then applied to identify the factors associated with the occurrence of inappropriate speed. The aim of the thesis is to achieve a better understanding of driver behavior near pedestrian crossings and to contribute to road safety assessment through modern analytical methods. The results may support the design of interventions to improve pedestrian safety in the urban environment.

Keywords: Illegal vehicle crossings, Computer Vision, YOLO-v8, Logistic Regression, Generalized Linear Models

Περίληψη

Η παρούσα διπλωματική εργασία έχει ως κύριο στόχο τη διερεύνηση της ακατάλληλης ταχύτητας οχημάτων σε σηματοδοτούμενες διαβάσεις πεζών στην Αθήνα και την εκτίμηση του τρόπου με τον οποίο η συμπεριφορά αυτή συνδέεται με κρίσιμες κυκλοφοριακές αλληλεπιδράσεις. Για την επίτευξη του στόχου αυτού πραγματοποιήθηκε αρχικά βιβλιογραφική ανασκόπηση σχετικά με τη διαχείριση ταχύτητας, τη συμπεριφορά των οδηγών σε αστικά περιβάλλοντα, τους surrogate safety indicators και τις εφαρμογές της υπολογιστικής όρασης στην ανάλυση της κυκλοφορίας. Στη συνέχεια συλλέχθηκαν και επεξεργάστηκαν δεδομένα βίντεο από σηματοδοτούμενη διάβαση πεζών, από τα οποία εξήχθησαν τροχιές οχημάτων, μεταβλητές κίνησης και δείκτες ασφάλειας όπως ο TTC και ο PET.

Ακολούθως, αναπτύχθηκε βάση δεδομένων υψηλής χρονικής ανάλυσης και εφαρμόστηκαν στατιστικές μέθοδοι για τον έλεγχο συσχέτισης, τον εντοπισμό πιθανής πολυσυγγραμμικότητας και την ανάπτυξη παλινδρομικών υποδειγμάτων. Ειδικότερα, εκτιμήθηκαν υποδείγματα διωνυμικής λογιστικής παλινδρόμησης για δυαδικές εκβάσεις σχετικές με το speeding, το PET και το TTC, καθώς και Γενικευμένα Γραμμικά Μοντέλα για τη διερεύνηση της επίδρασης των ανεξάρτητων μεταβλητών σε συνεχείς δείκτες επικινδυνότητας. Τα αποτελέσματα αξιολογήθηκαν με βάση δείκτες καλής προσαρμογής, στατιστικής σημαντικότητας και ερμηνευσιμότητας των συντελεστών.

Η συνολική ανάλυση ανέδειξε ότι η ακατάλληλη ταχύτητα κοντά στις διαβάσεις πεζών δεν αποτελεί τυχαίο φαινόμενο, αλλά συνδέεται με τη χωρική θέση του οχήματος, την παρουσία του στην περιοχή της διάβασης, το διαθέσιμο διαμήκες κενό, καθώς και με τους χρονικούς δείκτες αλληλεπίδρασης με τα λοιπά οχήματα. Παράλληλα, διαπιστώθηκε ότι οι δείκτες PET και TTC μπορούν να χρησιμοποιηθούν αποτελεσματικά για την προληπτική εκτίμηση της επικινδυνότητας, χωρίς να απαιτείται η εκ των υστέρων καταγραφή πραγματικών ατυχημάτων.

Συμπεράσματα

Παρουσίαση αποτελεσμάτων λογιστικών παλινδρομήσεων						
	Binomial Regression PET_MinPredicted_leq_3_0		Binomial Regression Speeding_during_Green_50_100		Binomial Regression TTC_Leader_leq_1_5	
Dep. Variable:	PET_MinPredicted_leq_3_0		Speeding_during_Green_50_100		TTC_Leader_leq_1_5	
AIC:	10051.584		13896.930		6906.863	
AUC:	0.831		0.794		0.928	
BIC:	10131.281		13962.137		6964.825	
Pseudo R-squ.:	0.281		0.290		0.515	
Variable	Coef.	P> z	Coef.	P> z	Coef.	P> z
Intercept	30.765	<0.001	-54.335	<0.001	0.078889597	0.527
GroundPlaneCentroid_1			0.392	<0.001		
GroundPlaneCentroid_2	0.013	<0.001				
IsInCrossingArea	0.398	<0.001	-0.404	<0.001	-0.185241545	0.029
magnitude	0.000	0.908			0.063092021	<0.001
VehicleStatus	0.216	0.119				
Speeding_during_Orange_50_100	0.341	0.390			-0.154656305	0.796
GapAlongPath_m	-0.002	0.003	0.004	<0.001	-0.209145706	<0.001
TTC_Leader_s	0.000	0.822	-0.001	0.087		
LateralOffsetToLeader_m	0.021	0.467	0.036	0.132	-0.410401177	<0.001
PET_MinPredicted_s			-0.004	0.767	0.027662049	0.116
PET_TimeToConflict_Self_s	-0.554	<0.001	-0.030	0.005	-0.093582771	<0.001

Πίνακας 0.1 Αποτελέσματα δυαδικής λογιστικής παλινδρόμησης

GLM Regression Scenarios					
	Generalized Linear Model για PET		Generalized Linear Model για Speeding_during_Green		
Dep. Variable:	PET_MinPredicted_s		Speeding_during_Green_50_100		
AIC:	40167.181		12319.175		
AUC:	—		0.793		
BIC:	40326.576		12427.854		
Pseudo R-squ.:	0.619		0.177		
Variable	Coef.	P> z	Coef.	P> z	
Intercept	-21.631	<0.001	-36.830	<0.001	
C(IsInCrossingArea)[T.1]	-0.295	<0.001	0.165	0.005	
C(TrafficLightStatus)[T.orange]	0.689	0.038			
C(TrafficLightStatus)[T.red]	0.642	0.001			
C(ProcessedTrafficLightStatus)[T.orange]	-0.299	0.093			
C(ProcessedTrafficLightStatus)[T.red]	-0.040	0.505			
C(VehicleStatus)[T.legal]	0.510	<0.001			
C(Vehicle_Classification)[T.moto]	-0.017	0.826			
Confidence	0.047	0.787	1.042	<0.001	
GroundPlaneCentroid_1	0.159	<0.001	0.247	<0.001	
vx			-0.181	<0.001	
vy	-0.019	<0.001			
Speeding_during_Orange_50_100	-0.824	0.041			
Speeding_during_Green_50_100	-0.066	0.082			
speed_calc	-0.007	0.099	-0.015	0.007	
GapAlongPath_m	0.002	<0.001	0.000	0.708	
LateralOffsetToLeader_m	-0.061	0.004	0.045	0.106	
TTC_Leader_s	0.000	0.819	0.000	0.294	
PET_MinPredicted_s			-0.023	0.081	
PET_TimeToConflict_Self_s	0.604	<0.001	0.028	0.016	
PET_TimeToConflict_Other_s	0.033	<0.001	-0.021	0.049	

Πίνακας 0.2 Αποτελέσματα Γενικευμένων Γραμμικών Μοντέλων

Από τα αποτελέσματα των μοντέλων προκύπτει ότι η **ακατάλληλη ταχύτητα** και γενικότερα η **επικινδυνότητα των αλληλεπιδράσεων** στη σηματοδοτούμενη διάβαση δεν αποτελούν τυχαία φαινόμενα, αλλά επηρεάζονται από συγκεκριμένα **χωρικά, κινηματικά** και **χρονικά χαρακτηριστικά** της κυκλοφορίας. Ειδικότερα, στα υποδείγματα για το speeding κατά το πράσινο και το πορτοκαλί διαπιστώθηκε ότι η συμπεριφορά των οδηγών σχετίζεται με τη **θέση του οχήματος** στο πεδίο της διασταύρωσης, με το αν το όχημα βρίσκεται **εντός της περιοχής διάβασης**, καθώς και με το διαθέσιμο **διαμήκες κενό** μπροστά του. Τα ευρήματα δείχνουν ότι όταν ο οδηγός διαθέτει περισσότερο ελεύθερο χώρο ή μεγαλύτερο αντιλαμβανόμενο χρονικό περιθώριο, είναι πιθανότερο να διατηρήσει ή να αναπτύξει υψηλότερη ταχύτητα, ενώ η είσοδος στην κρίσιμη περιοχή της διάβασης σε ορισμένες περιπτώσεις λειτουργεί ανασταλτικά, πιθανότατα λόγω αυξημένης προσοχής. Παράλληλα, το speeding κατά το πορτοκαλί εμφανίστηκε γενικά πιο δύσκολα προβλέψιμο από ό,τι το speeding κατά το πράσινο, γεγονός που υποδηλώνει ότι στη συγκεκριμένη συμπεριφορά υπεισέρχονται και παράγοντες που δεν αποτυπώνονται πλήρως στα διαθέσιμα δεδομένα, όπως η στιγμιαία απόφαση του οδηγού και η υποκειμενική αντίληψη κινδύνου. Περαιτέρω, η χαμηλότερη ερμηνευτική επίδοση του υποδείγματος είναι πιθανό να οφείλεται και στο **σχετικά μικρό μέγεθος του δείγματος** που αφορά τη συγκεκριμένη διάβαση, το οποίο περιορίζει τη δυνατότητα ανάδειξης στατιστικά σταθερών σχέσεων μεταξύ των μεταβλητών.

Ακόμη πιο ουσιαστικά είναι τα ευρήματα των υποδειγμάτων για τους δείκτες **PET** και **TTC**, καθώς εκεί αναδεικνύονται με μεγαλύτερη σαφήνεια οι παράγοντες που συνδέονται με τις **κρίσιμες συνθήκες αλληλεπίδρασης**. Τα μοντέλα έδειξαν ότι χαμηλές τιμές PET και TTC σχετίζονται κυρίως με μικρότερες **χωρικές αποστάσεις**, χαμηλότερο **χρονικό περιθώριο μέχρι πιθανή σύγκρουση** και, σε ορισμένες περιπτώσεις, με υψηλότερη ταχύτητα του οχήματος. Ιδιαίτερα το μοντέλο για **TTC \leq 1,5 sec** παρουσίασε πολύ ισχυρή επίδοση, γεγονός που υποδηλώνει ότι οι κρίσιμες συνθήκες ακολουθίας ερμηνεύονται αποτελεσματικά από μεταβλητές όπως η ταχύτητα, το διαμήκες και πλευρικό κενό και οι δείκτες χρονικής αλληλεπίδρασης. Συνολικά, τα μοντέλα της διπλωματικής δείχνουν ότι η επικινδυνότητα κοντά στη διάβαση διαμορφώνεται από τον συνδυασμό της **γεωμετρίας της κίνησης**, της **χρονικής εγγύτητας** μεταξύ οχημάτων και της **οδηγικής συμπεριφοράς**, γεγονός που ενισχύει την αξία των surrogate safety indicators για την κατανόηση και αξιολόγηση της οδικής ασφάλειας σε αστικό περιβάλλον.

Περιεχόμενα

Σύνοψη	3
Abstract	4
Περίληψη	5
Συμπεράσματα	6
Περιεχόμενα.....	8
Κεφάλαιο 1: Εισαγωγή	11
1.1 Γενική ανασκόπηση	11
1.2 Στόχος της διπλωματικής εργασίας	13
1.3 Ερευνητικά ερωτήματα.....	14
1.4 Μεθοδολογική προσέγγιση	14
1.5 Δομή της διπλωματικής εργασίας.....	15
Κεφάλαιο 2: Βιβλιογραφική Ανασκόπηση.....	16
2.1 Εισαγωγή.....	16
2.2 Μοτίβα συμπεριφοράς οδηγών και υπέρβαση ταχύτητας σε αστικές περιοχές.....	16
2.2.1 Εισαγωγή.....	16
2.2.2 Προσαρμογή ταχύτητας και συμπεριφορά πέδησης σε διαβάσεις.....	17
2.2.3 Αλληλεπίδραση με πεζούς και παραχώρηση προτεραιότητας.....	18
2.3 Παράγοντες επιρροής της ταχύτητας και της παραβατικότητας.....	18
2.3.1 Υποδομή, οπτικό περιβάλλον και γεωμετρικά χαρακτηριστικά	18
2.3.2 Περιβαλλοντικές συνθήκες και τύποι περιοχών	19
2.3.3 Ανθρώπινοι παράγοντες και αντίληψη κινδύνου.....	20
2.4 Αυτόματη εκτίμηση οδικής ασφάλειας με χρήση Υπολογιστικής Όρασης	20
2.4.1 Εισαγωγή στις εφαρμογές Computer Vision στις Μεταφορές	20
2.4.2 Αλγόριθμοι ανίχνευσης και παρακολούθησης οχημάτων (Detection & Tracking).....	21
2.4.3 Αυτόματη εκτίμηση ταχύτητας (Speed Estimation)	22
2.4.4 Ανάλυση Εμπλοκών και δεικτών ασφαλείας (Surrogate measures of Safety).....	22
2.5 Σύνοψη βιβλιογραφίας	23
Κεφάλαιο 3: Θεωρητικό Υπόβαθρο	24
3.1 Εισαγωγή.....	24

3.2 Βασικές έννοιες στατιστικής	24
3.3 Βασικές μέθοδοι στατιστικής ανάλυσης.....	25
3.3.1 Γραμμική παλινδρόμηση	26
3.3.2 Probit ανάλυση.....	26
3.3.3 Διαχωριστική ανάλυση	26
3.3.4 Λογιστική παλινδρόμηση.....	27
3.4 Διωνυμική λογιστική παλινδρόμηση	27
3.5 Γενικευμένα Γραμμικά Μοντέλα (Generalized Linear Models)	28
3.6 Κριτήρια αποδοχής και αξιολόγησης μοντέλου.....	28
3.6.1 Συντελεστές εξίσωσης.....	29
3.6.2 Στατιστική σημαντικότητα	29
3.6.3 Συσχέτιση ανεξάρτητων μεταβλητών.....	29
3.6.4 Δείκτες καλής προσαρμογής	29
3.6.5 Μέγιστη πιθανοφάνεια και λόγος πιθανοφάνειας.....	30
3.6.6 Κριτήρια πληροφοριών AIC και BIC	30
3.7 Θεωρία στοχαστικής χρησιμότητας	30
3.8 Σύνδεση του θεωρητικού πλαισίου με την παρούσα εργασία	31
Κεφάλαιο 4: Συλλογή και Επεξεργασία Στοιχείων	32
4.1 Εισαγωγή.....	32
4.2 Συλλογή δεδομένων	32
4.3 Προεπεξεργασία και ενοποίηση των αρχικών δεδομένων	36
4.4 Προσδιορισμός προπορευόμενου οχήματος και γεωμετρικών σχέσεων κίνησης	38
4.5 Αναλυτική περιγραφή των μεταβλητών του τελικού πίνακα	45
4.6 Στατιστική αποτύπωση δεδομένων - Συγκεντρωτικά Στοιχεία	48
Κεφάλαιο 5: Στατιστική Ανάλυση.....	56
5.1 Εισαγωγή.....	56
5.2 Ανάπτυξη Μοντέλων Διωνυμικής Παλινδρόμησης.....	56
5.2.1 Εισαγωγή δεδομένων, προεπεξεργασία και έλεγχος συσχέτισης	57
5.3 Αποτελέσματα Μοντέλων Διωνυμικής Παλινδρόμησης	64
5.3.1 Παρουσίαση αποτελεσμάτων λογιστικής παλινδρόμησης για το PET	64
5.3.2 Παρουσίαση αποτελεσμάτων λογιστικής παλινδρόμησης για Speeding_during_Green67	

5.3.3 Παρουσίαση αποτελεσμάτων λογιστικής παλινδρόμησης για το TTC_Leader.....	69
5.4 Αποτελέσματα Γενικευμένων Γραμμικών Μοντέλων.....	71
5.4.1 Αποτελέσματα GLM για το PET	71
5.4.2 Αποτελέσματα GLM για το Speeding_during_Green_.....	73
Κεφάλαιο 6: Συμπεράσματα.....	75
6.1 Σύνοψη.....	75
6.2 Συμπεράσματα.....	76
6.2 Προτάσεις για την Αξιοποίηση των Αποτελεσμάτων.....	78
6.3 Προτάσεις για Περαιτέρω Έρευνα	79
Βιβλιογραφία.....	80

Κεφάλαιο 1: Εισαγωγή

1.1 Γενική ανασκόπηση

Η οδική ασφάλεια στις αστικές περιοχές αποτελεί διαχρονικά έναν από τους κρισιμότερους τομείς του συγκοινωνιακού σχεδιασμού, καθώς στους ίδιους περιορισμένους χώρους συνυπάρχουν οχήματα, πεζοί, ποδηλάτες και χρήστες δημόσιων συγκοινωνιών. Ιδιαίτερα στις διαβάσεις πεζών, όπου η αλληλεπίδραση μεταξύ μηχανοκίνητης και μη μηχανοκίνητης κυκλοφορίας είναι άμεση, η ανάπτυξη ακατάλληλης ή υπερβολικής ταχύτητας από τα οχήματα αυξάνει ουσιαστικά την πιθανότητα σύγκρουσης αλλά και τη σοβαρότητα των συνεπειών της. Τόσο μέσω ελληνικής όσο και της διεθνούς βιβλιογραφίας, έχει αναδειχθεί ότι η **ταχύτητα αποτελεί έναν από τους βασικότερους παράγοντες κινδύνου** στο οδικό σύστημα, ιδίως μέσα στον αστικό ιστό, όπου η έκθεση ευάλωτων χρηστών είναι αυξημένη (Yannis & Michelaraki, 2024).

Στην Ελλάδα, τα οδικά τροχαία ατυχήματα εξακολουθούν να συνιστούν σημαντικό κοινωνικό και δημόσιο πρόβλημα. Τα πλέον πρόσφατα οριστικά ετήσια στοιχεία της ΕΛΣΤΑΤ καταγράφουν 10.553 οδικά τροχαία ατυχήματα με παθόντες για το έτος 2023, γεγονός που επιβεβαιώνει ότι, **παρά τη βελτίωση ορισμένων δεικτών σε βάθος χρόνου, το ζήτημα παραμένει ιδιαίτερα επίκαιρο** (ΕΛΣΤΑΤ, 2025). Στο αστικό περιβάλλον, η προστασία των πεζών αποκτά ακόμη μεγαλύτερη βαρύτητα, επειδή η καθημερινή τους μετακίνηση περιέχει επαναλαμβανόμενες διασταυρώσεις με την κυκλοφορία των οχημάτων σε σημεία όπου η ασφαλής συμπεριφορά όλων των χρηστών είναι απαραίτητη (Παπαδημητρίου, 2010).

Οι **διαβάσεις πεζών** δεν αποτελούν απλώς στοιχεία υποδομής, αλλά κρίσιμους κόμβους οργάνωσης της αστικής κινητικότητας και της βιώσιμης μεταφοράς. Η αποτελεσματικότητά τους εξαρτάται από τη γεωμετρία της οδού, την ορατότητα, τη διαγράμμιση, τη σηματοδότηση, τον κυκλοφοριακό φόρτο, αλλά και από τη συμπεριφορά των οδηγών κατά την προσέγγιση και διέλευσή τους (Σιδόπουλος, 2025). Όταν οι οδηγοί δεν προσαρμόζουν έγκαιρα την ταχύτητά τους στις επικρατούσες συνθήκες ή παραβιάζουν το πλαίσιο προτεραιότητας των πεζών, **ο χώρος της διάβασης μετατρέπεται από μέσο προστασίας σε πεδίο αυξημένου κινδύνου** (Αρβανιτάκη, 2022). Η σύγχρονη προσέγγιση της οδικής ασφάλειας αποδίδει, επομένως, ιδιαίτερη σημασία όχι μόνο στις τελικές συγκρούσεις, αλλά και στις επικίνδυνες αλληλεπιδράσεις που προηγούνται αυτών (Αλβέρτη, 2025).

Στο πλαίσιο της παρούσας εργασίας, η έννοια της ακατάλληλης ταχύτητας δεν ταυτίζεται αποκλειστικά με την υπέρβαση του εκάστοτε θεσμοθετημένου ορίου ταχύτητας, αλλά προσεγγίζεται υπό ένα ευρύτερο λειτουργικό και κυκλοφοριακό πρίσμα. Ειδικότερα, ως **ακατάλληλη νοείται κάθε ταχύτητα η οποία δεν είναι προσαρμοσμένη στις επικρατούσες συνθήκες του οδικού περιβάλλοντος**, όπως

η παρουσία και η πιθανή κίνηση πεζών, τα γεωμετρικά χαρακτηριστικά της θέσης, η λειτουργία της φωτεινής σηματοδότησης και οι εκάστοτε συνθήκες κυκλοφορίας (WHO, 2018; ΚΟΚ, Άρθρο 20).

Η σχετική ελληνική βιβλιογραφία αναδεικνύει ότι η ασφαλής διαχείριση της ταχύτητας δεν εξαντλείται στη συμμόρφωση προς το ισχύον κανονιστικό πλαίσιο, **καθώς ακόμη και ταχύτητες χαμηλότερες του ορίου ενδέχεται να κρίνονται μη κατάλληλες**, όταν δεν ανταποκρίνονται στο πραγματικό επίπεδο κινδύνου που χαρακτηρίζει μια συγκεκριμένη θέση, ιδίως σε περιοχές όπου παρατηρείται αυξημένη αλληλεπίδραση μεταξύ διαφορετικών κατηγοριών χρηστών της οδού (European Transport Safety Council (ETSC), 2010). Υπό το πρίσμα αυτό, η αξιολόγηση της ταχύτητας στις διαβάσεις πεζών οφείλει να στηρίζεται όχι μόνο στην εξέταση της τυπικής νομιμότητας της κίνησης, αλλά και στην εκτίμηση του βαθμού λειτουργικής ασφάλειας που αυτή διασφαλίζει.

Τα τελευταία χρόνια, η αξιοποίηση δεδομένων υψηλής ανάλυσης και τεχνικών υπολογιστικής όρασης (computer vision) έχει μετασχηματίσει τις δυνατότητες παρακολούθησης της κυκλοφορίας. Σε αντίθεση με τις παραδοσιακές μεθόδους συλλογής στοιχείων, όπως οι επιτόπιες παρατηρήσεις ή οι χειροκίνητες καταγραφές από βίντεο, τα σύγχρονα συστήματα **computer vision επιτρέπουν την αυτοματοποιημένη ανίχνευση**, ταυτοποίηση και παρακολούθηση οχημάτων και πεζών ανά χρονική στιγμή. Μέσω αυτών μπορούν να εξαχθούν τροχιές, ταχύτητες, αποστάσεις, χρόνοι αλληλεπίδρασης και δείκτες σύγκρουσης, παρέχοντας ένα ιδιαίτερα πλούσιο πληροφοριακό υπόβαθρο για την κατανόηση της κυκλοφοριακής συμπεριφοράς (Μυρογιάννη, 2025).

Ιδιαίτερο ενδιαφέρον παρουσιάζει η δυνατότητα εξαγωγής έμμεσων δεικτών ασφάλειας, όπως ο χρόνος μέχρι σύγκρουση (**Time to Collision - TTC**) και το **Post-Encroachment Time (PET)**, οι οποίοι επιτρέπουν την ποσοτικοποίηση της επικινδυνότητας χωρίς να απαιτείται η ύπαρξη πραγματικού ατυχήματος. Οι δείκτες αυτοί είναι ιδιαίτερος χρήσιμοι στις αστικές διαβάσεις πεζών, όπου τα ατυχήματα αποτελούν σχετικά σπάνια αλλά υψηλής σοβαρότητας γεγονότα, ενώ οι επικίνδυνες αλληλεπιδράσεις εμφανίζονται συχνότερα και μπορούν να λειτουργήσουν ως προειδοποιητικά σήματα για την ανάγκη παρέμβασης. Η χρήση τους συμβάλλει στην υιοθέτηση μιας πιο προληπτικής προσέγγισης για την οδική ασφάλεια, βασισμένης στην ανάλυση δεδομένων (Ventura et al., 2025).

Παράλληλα, η σύγχρονη βιβλιογραφία και πολιτική πρακτική συγκλίνουν στην **ανάγκη μείωσης των ταχυτήτων στις πόλεις** και ιδιαίτερα στους χώρους όπου συνυπάρχουν οχήματα και ευάλωτοι χρήστες. Πρόσφατες ανασκοπήσεις και αναλύσεις κόστους-οφέλους με επίκεντρο την ελληνική πραγματικότητα καταδεικνύουν ότι τα χαμηλότερα όρια ταχύτητας και η συστηματική επιτήρηση μπορούν να επιφέρουν σημαντικές

βελτιώσεις στην ασφάλεια, χωρίς ουσιαστική επιβάρυνση στη λειτουργία του δικτύου (Yannis & Michelaraki, 2024 ; Roussou et al., 2024). Ωστόσο, για να σχεδιαστούν αποτελεσματικά μέτρα πολιτικής απαιτείται λεπτομερής γνώση του πού, πότε και υπό ποιες συνθήκες εκδηλώνεται η ακατάλληλη ταχύτητα.

Η παρούσα διπλωματική εργασία εντάσσεται σε αυτό ακριβώς το πεδίο, εξετάζοντας τη συμπεριφορά οχημάτων σε **σηματοδοτούμενες διαβάσεις πεζών στην Αθήνα** με χρήση αλγορίθμων οπτικής αναγνώρισης. Το αντικείμενο επικεντρώνεται στη διερεύνηση της επίδρασης που έχει στην οδική ασφάλεια η ανάπτυξη ακατάλληλης ταχύτητας από οχήματα στις διαβάσεις πεζών, με αξιοποίηση βιντεοληπτικού υλικού και ανάπτυξη στατιστικών και υπολογιστικών μοντέλων για την ερμηνεία της συμπεριφοράς των παραγόντων που έχουν επίδραση στην υπέρβαση του ορίου ταχύτητας των οδηγών. Η προσέγγιση αυτή εδράζεται στη συνδυαστική αξιοποίηση αρχών της συγκοινωνιακής μηχανικής, της ανάλυσης οδικής ασφάλειας και μεθόδων υπολογιστικής όρασης, συγκροτώντας ένα ενιαίο ερευνητικό πλαίσιο.

1.2 Στόχος της διπλωματικής εργασίας

Βασικός στόχος της παρούσας διπλωματικής εργασίας είναι η **ανάλυση της ακατάλληλης ταχύτητας** οχημάτων σε σηματοδοτούμενες διαβάσεις πεζών στην Αθήνα και η διερεύνηση του τρόπου με τον οποίο αυτή συνδέεται με την εμφάνιση επικίνδυνων κυκλοφοριακών συνθηκών. Η εργασία επιδιώκει να αναδείξει, με βάση πραγματικά δεδομένα πεδίου, **ποια χαρακτηριστικά του οδικού περιβάλλοντος, της κυκλοφορίας και της αλληλεπίδρασης οχημάτων-πεζών επηρεάζουν** περισσότερο την πιθανότητα εμφάνισης μη ασφαλούς οδηγικής συμπεριφοράς κοντά στις διαβάσεις.

Ειδικότερα, η παρούσα έρευνα στοχεύει στην εφαρμογή τεχνικών υπολογιστικής όρασης για την εξαγωγή και συστηματική αποτύπωση τροχιών, ταχυτήτων και κρίσιμων δεικτών αλληλεπίδρασης, στην ποσοτική αποτίμηση της σχέσης μεταξύ ακατάλληλης ταχύτητας και δεικτών επικινδυνότητας, όπως ο TTC και ο PET, καθώς και στη διερεύνηση της επίδρασης των χαρακτηριστικών της υποδομής και της λειτουργίας της σηματοδότησης στη συμπεριφορά των οδηγών. Σε ευρύτερο επίπεδο, η εργασία φιλοδοξεί να συμβάλει στην ανάδειξη της αξίας των αυτοματοποιημένων συστημάτων παρακολούθησης ως εργαλείων λήψης αποφάσεων για φορείς σχεδιασμού, διαχείρισης και ελέγχου της κυκλοφορίας.

1.3 Ερευνητικά ερωτήματα

Η παρούσα διπλωματική εργασία επιδιώκει να διερευνήσει κρίσιμες πτυχές της οδηγικής συμπεριφοράς και της επικινδυνότητας στις διαβάσεις πεζών, με έμφαση στην αξιοποίηση τεχνικών υπολογιστικής όρασης και κυκλοφοριακών δεικτών. Στο πλαίσιο αυτό, εξετάζεται αρχικά **κατά πόσο οι μέθοδοι computer vision μπορούν να χρησιμοποιηθούν αξιόπιστα για τον εντοπισμό περιπτώσεων ακατάλληλης ταχύτητας**, αλλά και για την αναγνώριση επικίνδυνων αλληλεπιδράσεων μεταξύ οχημάτων και πεζών στον χώρο της διάβασης.

Παράλληλα, διερευνάται η επίδραση παραμέτρων που συνδέονται με τη λειτουργία της οδού, τη φωτεινή σηματοδότηση, καθώς και τις τοπικές κυκλοφοριακές συνθήκες, προκειμένου να αποσαφηνιστεί **ποιοι παράγοντες επηρεάζουν εντονότερα τη συμπεριφορά των οδηγών** κατά την προσέγγιση των διαβάσεων πεζών. Ιδιαίτερη έμφαση δίνεται επίσης στην αξιολόγηση της δυνατότητας χρήσης δεικτών σύγκρουσης και μεταβλητών κυκλοφοριακής συμπεριφοράς ως εργαλείων πρόβλεψης ή έγκαιρης αναγνώρισης συνθηκών αυξημένου κινδύνου. Τέλος, διερευνάται η προοπτική ενσωμάτωσης των παραπάνω μεθόδων σε εφαρμογές έξυπνης πόλης και σε συστήματα διαχείρισης οδικής ασφάλειας, ώστε να καταστεί δυνατή η πιο έγκαιρη αναγνώριση επικίνδυνων συνθηκών και η αποτελεσματικότερη ιεράρχηση των απαιτούμενων παρεμβάσεων.

1.4 Μεθοδολογική προσέγγιση

Για την επίτευξη των παραπάνω στόχων ακολουθείται μια μεθοδολογική προσέγγιση που συνδυάζει **βιβλιογραφική ανασκόπηση, επεξεργασία βιντεοληπτικού υλικού, παραγωγή βάσης δεδομένων υψηλής χρονικής ανάλυσης και στατιστική/model-based ανάλυση**. Αρχικά, πραγματοποιείται ανασκόπηση της σχετικής βιβλιογραφίας γύρω από την ταχύτητα, την ασφάλεια πεζών, τη λειτουργία των διαβάσεων και τις εφαρμογές υπολογιστικής όρασης στην ανάλυση της κυκλοφορίας. Στη συνέχεια αξιοποιούνται δεδομένα από κάμερες σε επιλεγμένες διαβάσεις πεζών της Αθήνας, από τα οποία εξάγονται τροχιές οχημάτων και συναφείς μεταβλητές κίνησης.

Ακολούθως, υπολογίζονται δείκτες που περιγράφουν τόσο τη συμπεριφορά των οχημάτων κατά την πράσινη και πορτοκαλί ένδειξη του φαναριού όσο και το επίπεδο δυνητικού κινδύνου, όπως στιγμιαίες και μέσες ταχύτητες, σχετικές αποστάσεις, TTC και PET. Τα δεδομένα αυτά οργανώνονται σε κατάλληλη βάση ανάλυσης και χρησιμοποιούνται για την ανάπτυξη στατιστικών μοντέλων, με σκοπό τον εντοπισμό των σημαντικότερων παραγόντων που συνδέονται με την ακατάλληλη ταχύτητα και τις

κρίσιμες αλληλεπιδράσεις. Η ανάλυση εστιάζει τόσο στην ερμηνεία των σχέσεων όσο και στη συγκριτική αξιολόγηση των επιμέρους θέσεων μελέτης.

Η επιλογή αυτής της μεθοδολογίας είναι συμβατή με τις σύγχρονες τάσεις στην ανάλυση οδικής ασφάλειας, όπου η χρήση surrogate safety measures και αλγοριθμικής επεξεργασίας βίντεο επιτρέπει τη διερεύνηση κινδύνων με μεγαλύτερη ανάλυση, ακρίβεια και προληπτικό χαρακτήρα σε σχέση με την αποκλειστική χρήση δεδομένων ατυχημάτων.



Γράφημα 1 : Μεθοδολογική προσέγγιση

1.5 Δομή της διπλωματικής εργασίας

Η εργασία οργανώνεται σε διακριτά κεφάλαια, ώστε να παρουσιάζεται με συστηματικό τρόπο τόσο το θεωρητικό πλαίσιο όσο και η εμπειρική διερεύνηση. Μετά το εισαγωγικό κεφάλαιο ακολουθεί βιβλιογραφική ανασκόπηση, στην οποία εξετάζονται ζητήματα οδικής ασφάλειας πεζών, διαχείρισης ταχυτήτων, λειτουργίας διαβάσεων και εφαρμογών υπολογιστικής όρασης στην κυκλοφοριακή ανάλυση. Στη συνέχεια παρουσιάζονται τα δεδομένα, οι θέσεις μελέτης και η μεθοδολογία επεξεργασίας του βιντεοληπτικού υλικού.

Στα επόμενα κεφάλαια αναπτύσσονται τα στατιστικά και υπολογιστικά μοντέλα, παρουσιάζονται τα αποτελέσματα της ανάλυσης και σχολιάζονται τα βασικά ευρήματα ως προς τους παράγοντες που επηρεάζουν την ακατάλληλη ταχύτητα στις διαβάσεις πεζών. Η εργασία ολοκληρώνεται με τα κύρια συμπεράσματα, τις προτάσεις πολιτικής και τις κατευθύνσεις για περαιτέρω έρευνα.

Κεφάλαιο 2: Βιβλιογραφική Ανασκόπηση

2.1 Εισαγωγή

Η οδική ασφάλεια σε αστικά περιβάλλοντα αποτελεί αντικείμενο εκτενούς έρευνας παγκοσμίως, με ιδιαίτερη έμφαση στα σημεία αλληλεπίδρασης ευάλωτων χρηστών (Vulnerable Road Users - VRUs) και μηχανοκίνητων οχημάτων, όπως οι διαβάσεις πεζών (Klanjic et al, 2022). **Η ακατάλληλη ταχύτητα** και η επιθετική συμπεριφορά των οδηγών αναγνωρίζονται ως κρίσιμοι παράγοντες πρόκλησης ατυχημάτων. Ενώ παραδοσιακά η αξιολόγηση της οδικής ασφάλειας βασιζόταν σε ιστορικά δεδομένα ατυχημάτων, η σύγχρονη έρευνα στρέφεται στην προληπτική ανάλυση της οδηγικής συμπεριφοράς και των κυκλοφοριακών εμπλοκών (traffic conflicts).

Η ανάγκη για μείωση των τροχαίων ατυχημάτων και την ενίσχυση της οδικής ασφάλειας, ειδικά σε αστικό περιβάλλον, έχει οδηγήσει στην ανάπτυξη αυτόματων συστημάτων παρακολούθησης της κυκλοφορίας. Με την πρόοδο στην **Τεχνητή Όραση (Computer Vision, CV)** και στις μεθόδους βαθιάς μάθησης (Deep Learning, DL), έχει καταστεί δυνατή η **αυτοματοποιημένη ανίχνευση οχημάτων**, πεζών, η εκτίμηση ταχύτητας, και η ανάλυση συμπεριφοράς οδηγών και πεζών. Τα συστήματα αυτά υπόσχονται αποτελεσματική, οικονομική και αξιόπιστη εναλλακτική ή συμπληρωματική λύση σε σχέση με παραδοσιακούς αισθητήρες (ραντάρ, lidar, βρόχους, ...) (Dilek & Dener, 2023).

Στο πλαίσιο αυτό, η βιβλιογραφική ανασκόπηση που ακολουθεί διαρθρώνεται σε τρεις κεντρικούς άξονες. Αρχικά, εξετάζονται τα **μοτίβα συμπεριφοράς των οδηγών** κατά την προσέγγιση σε διαβάσεις, με έμφαση στην προσαρμογή της ταχύτητας και την παραχώρηση προτεραιότητας. Στη συνέχεια, αναλύονται οι **παράγοντες** - περιβαλλοντικοί, γεωμετρικοί και ανθρωπinoi-που **επηρεάζουν την επιλογή ταχύτητας**. Τέλος, παρουσιάζεται η ραγδαία εξελισσόμενη τεχνολογία της Τεχνητής Όρασης (**Computer Vision**) ως **μεθοδολογικό εργαλείο** για την αυτόματη ανίχνευση, ιχνηλάτηση και εκτίμηση της ταχύτητας οχημάτων, προσφέροντας νέες δυνατότητες για την ακριβή και μαζική συλλογή δεδομένων.

2.2 Μοτίβα συμπεριφοράς οδηγών και υπέρβαση ταχύτητας σε αστικές περιοχές

2.2.1 Εισαγωγή

Η κατανόηση της συμπεριφοράς του οδηγού αποτελεί τον ακρογωνιαίο λίθο για την ανάπτυξη στρατηγικών μείωσης των ατυχημάτων. Η σύγχρονη βιβλιογραφία υποδεικνύει ότι η **αξιολόγηση της οδικής ασφάλειας δεν μπορεί** να περιορίζεται αποκλειστικά **στην εκ των υστέρων καταγραφή συμβάντων**, αλλά πρέπει να

επεκτείνεται στη μελέτη των φυσικών (naturalistic) δεδομένων οδήγησης. Μέσα από ανασκοπήσεις των μεθόδων αξιολόγησης, τεκμηριώνεται ότι τα δεδομένα αυτά είναι απαραίτητα για να αποκαλυφθούν λανθάνουσες επικίνδυνες συμπεριφορές, όπως η υπερβολική ταχύτητα και η καθυστερημένη αντίδραση, οι οποίες συχνά δεν καταγράφονται στις στατιστικές ατυχημάτων (Zijun et al, 2023).

Σε αυτό το πλαίσιο, η ανάλυση των συμπεριφορών μέσω βίντεο και αλγορίθμων Τεχνητής Όρασης (Computer Vision) αναδεικνύεται σε κρίσιμο εργαλείο. Η τεχνολογία αυτή επιτρέπει την **ακριβή ταυτοποίηση μοτίβων που οδηγούν σε συγκρούσεις**, ιδιαίτερα σε σύνθετα αστικά περιβάλλοντα, όπου η πολυπλοκότητα της αλληλεπίδρασης και η γεωμετρία της οδού απαιτούν εξειδικευμένη παρακολούθηση για τον εντοπισμό επισφαλών ενεργειών (St-Aubin, 2016).

2.2.2 Προσαρμογή ταχύτητας και συμπεριφορά πέδησης σε διαβάσεις

Ένα κρίσιμο στοιχείο της ασφάλειας στις διαβάσεις είναι ο τρόπος με τον οποίο οι οδηγοί διαχειρίζονται την ταχύτητά τους πλησιάζοντας το σημείο διεπαφής με τους πεζούς. Η ανάγκη για αντικειμενική μέτρηση αυτής της συμπεριφοράς έχει οδηγήσει στην ανάπτυξη μεθοδολογιών ασφάλειας βασισμένων σε βίντεο .

Οι μεθοδολογίες αυτές παρέχουν ένα δομημένο πλαίσιο, μετατρέποντας την απλή οπτική καταγραφή σε ποσοτικούς δείκτες κινδύνου. Για παράδειγμα, τέτοιες προσεγγίσεις έχουν εφαρμοστεί επιτυχώς σε πανεπιστημιακούς χώρους για την αξιολόγηση της συμμόρφωσης των οδηγών (Hnooohom et al., 2024). Παράλληλα, διευρύνοντας το πεδίο εφαρμογής τους, αντίστοιχα μεθοδολογικά πλαίσια χρησιμοποιούνται πλέον σε μη σηματοδοτούμενες διασταυρώσεις για τον υπολογισμό σε πραγματικό χρόνο δεικτών όπως ο **Χρόνος Μετά την Εμπλοκή (Post-Encroachment Time - PET)**, προσφέροντας μια δυναμική και άμεση αξιολόγηση του δυνητικού κινδύνου για τους πεζούς (Tengfeng et al, 2024).

Διερευνώντας τη δυναμική της απόκρισης των οδηγών, έρευνες που αξιοποιούν δεδομένα βίντεο έχουν αναλύσει τη συμπεριφορά πέδησης (braking behavior) σε μεσοδιαβάσεις (mid-block crossings). Χρησιμοποιώντας προηγμένα στατιστικά εργαλεία, όπως **μοντέλα mixed logit με ετερογένεια στους μέσους όρους**, οι μελετητές έχουν καταφέρει να ερμηνεύσουν τις σημαντικές διαφοροποιήσεις στις αντιδράσεις των οδηγών, παρέχοντας μια πιο λεπτομερή εικόνα της διαδικασίας επιβράδυνσης (Ahsan et al, 2025).

Παράλληλα, η συνολική αξιολόγηση της απόδοσης των οδηγών δεν εξαρτάται μόνο από την κινηματική του οχήματος αλλά και από ψυχοφυσιολογικούς παράγοντες. Η **ανάλυση συνδυαστικών δεδομένων ταχύτητας και αντίληψης (perceptive data)** κατά τη διάρκεια δοκιμών πεδίου, έχει καταδείξει ότι, αν και η έγκαιρη μείωση της ταχύτητας είναι καθοριστική για την αποφυγή ατυχημάτων, η απόφαση αυτή συχνά

επηρεάζεται και αλλοιώνεται από την υποκειμενική αντίληψη του κινδύνου που έχει ο οδηγός τη δεδομένη στιγμή (Meocci et al, 2024).

2.2.3 Αλληλεπίδραση με πεζούς και παραχώρηση προτεραιότητας

Η απόφαση ενός οδηγού να παραχωρήσει προτεραιότητα (yielding behavior) σε έναν πεζό είναι μια σύνθετη διαδικασία που εξαρτάται άμεσα και καθοριστικά από την ταχύτητα προσέγγισης. Η **υψηλή ταχύτητα μειώνει δραματικά τον διαθέσιμο χρόνο αντίδρασης**, καθιστώντας επιτακτική την ανάγκη για έγκαιρη πρόβλεψη. Σύγχρονα πλαίσια, όπως το VRUCrossSafe, στοχεύουν ακριβώς στην πρόβλεψη της πρόθεσης διάσχισης (crossing intention prediction) των ευάλωτων χρηστών, δίνοντας στον οδηγό τα απαραίτητα δευτερόλεπτα για να μειώσει ταχύτητα προληπτικά και να αποφύγει την απότομη πέδηση (Abdelrahman et al, 2025).

Ωστόσο, η συμμόρφωση και η επιλογή ταχύτητας δεν είναι ενιαίες. Μελέτες σε ημι-ελεγχόμενες διαβάσεις έχουν δείξει ότι η επιθετικότητα της οδήγησης και η τάση για **υπερβολική ταχύτητα ποικίλλουν** σημαντικά ανάλογα **με τον τύπο του οχήματος (ετερογένεια)**. Ιδιαίτερα για τα βαρέα οχήματα, η τήρηση των ορίων ταχύτητας είναι κρίσιμη, καθώς η μεγάλη μάζα αυξάνει την απαιτούμενη απόσταση φρεναρίσματος, καθιστώντας κάθε υπέρβαση ταχύτητας δυνητικά μοιραία σε περίπτωση αλληλεπίδρασης με πεζό (Wang et al, 2021).

Για την αντικειμενική καταγραφή αυτών των φαινομένων, η ανάπτυξη αλγορίθμων Υπολογιστικής Όρασης (Computer Vision) επιτρέπει πλέον την αυτόματη ανίχνευση του εάν ένα όχημα επιβραδύνει και σταματά για τους πεζούς, παρέχοντας ποσοτικά δεδομένα για την επικινδυνότητα που απορρέει από την ακατάλληλη ταχύτητα διέλευσης (Wan et al, 2023). Τέλος, η ασφάλεια συναρτάται και από την κατάσταση του οδηγού· συστήματα ταυτόχρονης παρακολούθησης (simultaneous monitoring) έχουν αναδείξει ότι η **απόσπαση προσοχής (π.χ. από χρήση κινητού τηλεφώνου κλπ.)** οδηγεί συχνά σε καθυστερημένη αντίληψη των οδικών κινδύνων και, κατ' επέκταση, σε αδυναμία έγκαιρης προσαρμογής της ταχύτητας στα επιτρεπτά όρια (Rezaei, 2016).

2.3 Παράγοντες επιρροής της ταχύτητας και της παραβατικότητας

2.3.1 Υποδομή, οπτικό περιβάλλον και γεωμετρικά χαρακτηριστικά

Το **φυσικό περιβάλλον και τα γεωμετρικά χαρακτηριστικά της οδού** διαδραματίζουν πρωτεύοντα ρόλο στη διαμόρφωση της επιλογής ταχύτητας. Πρόσφατες έρευνες έχουν διερευνήσει διεξοδικά την επίδραση του οπτικού περιβάλλοντος (visual surroundings) στην τάση για υπερβολική ταχύτητα, υποδεικνύοντας ότι στοιχεία όπως η πυκνότητα της δόμησης και η ανοιχτότητα του πεδίου ορατότητας δεν είναι αμελητέα, αλλά μπορούν να ενθαρρύνουν ή να αποτρέψουν την επιθετική οδήγηση (Abdel-Aty et al, 2024).

Για την αξιολόγηση όμως της συμπεριφοράς σε συγκεκριμένα σημεία, προαπαιτούμενο αποτελεί ο **ακριβής εντοπισμός της υποδομής**. Προς αυτή την κατεύθυνση, έχουν αναπτυχθεί ημι-αυτόματες τεχνικές που συνδυάζουν Τεχνική Όραση και Βαθιά Μάθηση (Deep Learning) για την επίλυση του προβλήματος ανίχνευσης διαβάσεων (Crosswalk Detection Problem), επιτρέποντας την ταχεία και ακριβή χαρτογράφηση των σημείων ενδιαφέροντος (Rúbio et al., 2020).

Αφού εντοπιστεί η διάβαση, τα ειδικά χαρακτηριστικά της, όπως η ύπαρξη κεντρικής νησίδας, η ποιότητα της διαγράμμισης και η σήμανση, φαίνεται να επηρεάζουν καθοριστικά τη συμμόρφωση των οδηγών (yielding compliance) ως προς την παραχώρηση προτεραιότητας, ειδικά σε σηματοδοτούμενες μεσοδιαβάσεις (Kutela, 2021). Τέλος, σε ευρωπαϊκό επίπεδο, η **συστηματική ταυτοποίηση των αστικών χαρακτηριστικών** που σχετίζονται με την ασφάλεια των ευάλωτων χρηστών έχει αναδείξει την ανάγκη σχεδιασμού «συγχωρητικών» υποδομών που λαμβάνουν υπόψη την ανθρώπινη φύση και τα περιβαλλοντικά δεδομένα (Klanjcic et al, 2022).

2.3.2 Περιβαλλοντικές συνθήκες και τύποι περιοχών

Η συμπεριφορά των οδηγών και η επιλογή ταχύτητας διαφοροποιούνται σημαντικά ανάλογα με τη λειτουργική κατηγορία και τις ειδικές συνθήκες της περιοχής. Στις ευαίσθητες **σχολικές ζώνες**, μελέτες που αξιοποιούν τεχνικές Deep Learning (όπως στην Osan City της Νότιας Κορέας) έχουν δείξει ότι η εφαρμογή ενισχυμένων πολιτικών βελτιώνει την ασφάλεια, αν και η πλήρης συμμόρφωση στα όρια ταχύτητας παραμένει ένα διαρκές ζητούμενο (Zhixiong et al, 2024).

Αντίστοιχα, οι **ζώνες εκτέλεσης έργων** (work zones) παρουσιάζουν μοναδικές προκλήσεις, καθώς το απρόβλεπτο οδικό περιβάλλον διαταράσσει την ομαλή συμπεριφορά των οδηγών. Για την ακριβή εκτίμηση της επικινδυνότητας σε αυτά τα σημεία, οι ερευνητές χρησιμοποιούν **εξειδικευμένα στατιστικά μοντέλα, όπως οι προσεγγίσεις διακριτής διάρκειας** (discretized duration approach). Με αυτή τη μέθοδο, το εξεταζόμενο σημείο ενδιαφέροντος χωρίζεται σε μικρότερα τμήματα, επιτρέποντας την πρόβλεψη του ακριβούς σημείου ή της χρονικής στιγμής όπου ένας οδηγός είναι πιθανότερο να αναπτύξει υπερβολική ταχύτητα (Thapa et al, 2023).

Σε ευρύτερο πλαίσιο, η σύγκριση μεταξύ αστικού και αγροτικού περιβάλλοντος αποκαλύπτει διαφορετικά μοτίβα. Στις **αγροτικές περιοχές**, πολυμεταβλητές προσεγγίσεις έχουν εντοπίσει ότι παράγοντες όπως η γεωμετρία της οδού και η ροή της κυκλοφορίας ωθούν τους οδηγούς σε παραβάσεις ταχύτητας με διαφορετικό τρόπο σε σχέση με τον αστικό ιστό (Ferko et al, 2024). Στον αντίποδα, εντός των **πόλεων** (π.χ. Βερολίνο), η σοβαρότητα των ατυχημάτων με πεζούς συνδέεται στενά με την πολυπλοκότητα του δικτύου και την πυκνότητα των αλληλεπιδράσεων (Kopsacheilis et al, 2024). Τέλος, για την ολιστική αξιολόγηση της ασφάλειας στις διαβάσεις ανεξαρτήτως περιοχής, η χρήση Συνελικτικών Νευρωνικών Δικτύων (CNNs) που

εκπαιδεύονται σε ζεύγη εναέριων και επίγειων εικόνων προσφέρει μια νέα προοπτική στην κατανόηση του πώς η προοπτική θέασης επηρεάζει την εκτίμηση της ασφάλειας (Russon et al, 2025).

2.3.3 Ανθρώπινοι παράγοντες και αντίληψη κινδύνου

Πέραν των εξωτερικών και περιβαλλοντικών παραγόντων, η ψυχολογία του οδηγού και η ικανότητά του να αντιλαμβάνεται υποκειμενικά τον κίνδυνο (perceived risk) καθορίζουν σε μεγάλο βαθμό την οδηγική συμπεριφορά. Η ανάπτυξη μοντέλων τεχνικής όρασης για την πρόβλεψη του αντιληπτού κινδύνου σε κυκλοφοριακές σκηνές προσφέρει πλέον τη δυνατότητα **να κατανοήσουμε πώς οι οδηγοί «διαβάζουν» το δρόμο** και προσαρμόζουν (ή αποτυγχάνουν να προσαρμόσουν) την ταχύτητά τους βάσει των οπτικών ερεθισμάτων που λαμβάνουν (Winter et al, 2023).

Ωστόσο, η έρευνα δείχνει ότι η απόκλιση μεταξύ του πραγματικού και του αντιληπτού κινδύνου είναι συχνά η γενεσιουργός αιτία της ακατάλληλης ταχύτητας σε κρίσιμα σημεία όπως οι διαβάσεις. Προς αυτή την **κατεύθυνση, ποσοτική μελέτη των αλληλεπιδράσεων οχήματος-πεζού έχει αναδείξει την ανάγκη για σαφέστερη επικοινωνία προθέσεων** (π.χ. μέσω προσαρμοστικών λειτουργιών φωτισμού σε μελλοντικά οχήματα), ώστε να γεφυρωθεί το χάσμα αντίληψης και να διασφαλιστεί η έγκαιρη προσαρμογή της ταχύτητας (Zang et al, 2021).

2.4 Αυτόματη εκτίμηση οδικής ασφάλειας με χρήση Υπολογιστικής Όρασης

2.4.1 Εισαγωγή στις εφαρμογές Computer Vision στις Μεταφορές

Η τεχνολογία της Υπολογιστικής Όρασης (Computer Vision - CV) έχει φέρει επανάσταση στον τομέα των Ευφυών Συστημάτων Μεταφορών (ITS), αποτελώντας πλέον ένα από τα πλέον αξιόπιστα και οικονομικά μέσα για την αναλυτική κατανόηση της κυκλοφορίας σε πραγματικό χρόνο.

Εργασίες επισκόπησης επιβεβαιώνουν ότι οι αλγόριθμοι CV επιτρέπουν την αυτοματοποιημένη ανάλυση τεράστιου όγκου δεδομένων βίντεο, καλύπτοντας το φάσμα από την απλή παρακολούθηση έως τον πλήρη έλεγχο της κυκλοφορίας (Dilek & Dener, 2023). Ιδιαίτερη βαρύτητα δίνεται στη χρήση Deep Learning, συνελκτικών νευρωνικών δικτύων (CNNs), **ανιχνευτών τύπου YOLO** και συστημάτων tracking όπως SORT/DeepSORT, τα οποία αποτελούν τη βάση και για την παρούσα διπλωματική. Παράλληλα, σύγχρονες μελέτες παρουσιάζουν τις μεθοδολογίες αιχμής που χρησιμοποιούνται για την αναγνώριση οχημάτων και την εκτίμηση τροχιών, θέτοντας τη βάση για την ανάπτυξη πιο έξυπνων και προσαρμοστικών συστημάτων ελέγχου (Abiamamela et al, 2024).

Σε επίπεδο επιβολής κανόνων, η εφαρμογή συστημάτων ανίχνευσης παραβάσεων που **βασίζονται στην αλληλεπίδραση ανθρώπου-υπολογιστή και στο CV προσφέρει ακρίβεια** και αντικειμενικότητα που δεν μπορούν να επιτευχθούν με τις παραδοσιακές μεθόδους, επιτρέποντας τον εντοπισμό σύνθετων παραβατικών συμπεριφορών (Ren, 2024).

Πέρα όμως από την παρατήρηση, η τεχνολογία εξελίσσεται και προς την «Ενεργητική Ασφάλεια» (Active Safety). Με στόχο λοιπόν την πρόληψη και αποτροπή των εμπλοκών, πρωτότυπα συστήματα έχουν ήδη δοκιμαστεί για την ενσωμάτωση του CV με τεχνολογίες επικοινωνίας ITS-G5, όπου **η κάμερα ανιχνεύει τον κίνδυνο και ενημερώνει** ασύρματα τα οχήματα, αυξάνοντας δραστικά την ασφάλεια στις διαβάσεις (Gaspar et al, 2020).

Τέλος, απαντώντας στις ανησυχίες περί προσωπικών δεδομένων, καινοτόμες προσεγγίσεις όπως το σύστημα VTPM (Video-to-Text Pedestrian Monitoring) συνδυάζουν το CV με **Μεγάλα Γλωσσικά Μοντέλα (LLMs)**, μετατρέποντας την οπτική πληροφορία σε κειμενική περιγραφή δραστηριότητας, διασφαλίζοντας έτσι την προστασία της ιδιωτικότητας κατά την παρακολούθηση σε διασταυρώσεις (Abdelrahman et al, 2024).

2.4.2 Αλγόριθμοι ανίχνευσης και παρακολούθησης οχημάτων (Detection & Tracking)

Η ακριβής εκτίμηση της ταχύτητας προϋποθέτει ως θεμέλιο λίθο τον αλάνθαστο εντοπισμό και την ιχνηλάτηση των οχημάτων. Σύγχρονα υπολογιστικά πλαίσια (frameworks) αξιοποιούν προηγμένους **αλγόριθμους βαθιάς μάθησης (Deep Learning)** για να επιτύχουν υψηλή ακρίβεια στον εντοπισμό και την παρακολούθηση οχημάτων απευθείας από βίντεο οδικής επιτήρησης, αντικαθιστώντας τις παλαιότερες, λιγότερο αξιόπιστες μεθόδους (Yadav et al, 2023).

Ωστόσο, το αστικό περιβάλλον παρουσιάζει ιδιαίτερες προκλήσεις λόγω της οπτικής πολυπλοκότητας. Για τον λόγο αυτό, εξειδικευμένοι αλγόριθμοι, όπως ο **VRU-YOLO**, έχουν αναπτυχθεί συγκεκριμένα για να αντιμετωπίσουν το πρόβλημα του εντοπισμού μικρών αντικειμένων και ευάλωτων χρηστών (VRUs) σε πολύπλοκα σκηνικά, βελτιώνοντας σημαντικά την ικανότητα του συστήματος να διακρίνει πεζούς και δίκυκλα μέσα στην πυκνή κυκλοφορία (Yunxiang Liu, 2025).

Για την επίτευξη λειτουργίας σε πραγματικό χρόνο, τεχνικές όπως το «Tracking-by-Detection» έχουν αποδειχθεί εξαιρετικά αποτελεσματικές, επιτρέποντας την **παρακολούθηση υψηλής ταχύτητας (High-Speed Tracking)**. Η μέθοδος αυτή βελτιστοποιεί την απόδοση του συστήματος, καθώς ελαχιστοποιεί την ανάγκη για βαριά επεξεργασία εικόνας σε κάθε καρέ, εστιάζοντας στη γεωμετρική αντιστοίχιση των ανιχνεύσεων (Bochinski et al, 2017). Τέλος, η εφαρμογή αυτών των τεχνολογιών **σε βίντεο ληφθέντα από το επίπεδο του δρόμου (street-level videos)** επιτρέπει πλέον

τη λεπτομερή ανάλυση όχι μόνο της κίνησης, αλλά και της συμπεριφοράς και αλληλεπίδρασης πεζών-οχημάτων, παρέχοντας αξιόπιστα δεδομένα υψηλής εγκυρότητας για την έρευνα οδικής ασφάλειας (Ventura et al, 2025).

2.4.3 Αυτόματη εκτίμηση ταχύτητας (Speed Estimation)

Η εκτίμηση της ταχύτητας από μονοσκοπικάβίντεο (βίντεο που έχουν καταγραφεί από μία και μόνο κάμερα) αποτελεί μια **σύνθετη τεχνική πρόκληση**, η οποία ωστόσο αντιμετωπίζεται πλέον αποτελεσματικά μέσω προηγμένων μεθόδων γεωμετρικής αντιστοίχισης. Συγκεκριμένα, η βιβλιογραφία προτείνει τη χρήση εξελικτικών αλγορίθμων για την αυτόματη βαθμονόμηση της κάμερας (evolutionary camera calibration), μια καινοτόμος προσέγγιση που επιτρέπει τον υπολογισμό της ταχύτητας των οχημάτων με υψηλή ακρίβεια, ελαχιστοποιώντας τα σφάλματα μέτρησης (Mejia et al, 2021).

Πέραν της απλής μέτρησης, συνδυαστικές προσεγγίσεις που ενσωματώνουν Deep Learning τεχνικές αξιοποιούνται για την αυτόματη ανίχνευση σοβαρών τροχαίων παραβάσεων. Τα συστήματα αυτά είναι σε θέση να εντοπίζουν και να ταξινομούν περιστατικά **που οφείλονται σε υπερβολική ταχύτητα** ή κίνηση στο αντίθετο ρεύμα, παρέχοντας αξιόπιστα εργαλεία τόσο για την επιβολή του νόμου όσο και για την ερευνητική ανάλυση της οδικής ασφάλειας (Shailendra et al, 2025).

Τέλος, μια εναλλακτική και ιδιαίτερα αποδοτική προοπτική για την εξαγωγή κινηματικών δεδομένων προσφέρει η **χρήση μη επανδρωμένων αεροσκαφών (UAV/drones)**. Μελέτες έχουν καταδείξει ότι η εναέρια βίντεο-καταγραφή επιτρέπει την λεπτομερή διερεύνηση της ροής και των παραβάσεων, προσφέροντας δεδομένα και οπτικές γωνίες που συχνά δεν είναι προσβάσιμα από τις συμβατικές επίγειες κάμερες (Bhavsar et al, 2023).

2.4.4 Ανάλυση Εμπλοκών και δεικτών ασφαλείας (Surrogate measures of Safety)

Πέρα από την απλή μέτρηση της ταχύτητας, η Τεχνική Όραση επιτρέπει την **εκτίμηση έμμεσων δεικτών ασφαλείας** (Surrogate Measures of Safety- SMOs) για την πρόβλεψη ατυχημάτων. Η σύγχρονη έρευνα στο πεδίο της Τεχνητής Νοημοσύνης επιτρέπει τη μετάβαση από τις ετερογενείς πηγές δεδομένων και τα video analytics στην ακριβή εκτίμηση μέτρων για τους πεζούς και την τελική πρόβλεψη της συχνότητας των συγκρούσεων (Lian et al, 2024).

Για την ποσοτικοποίηση της επικινδυνότητας, δείκτες όπως ο Χρόνος Μετά την Εμπλοκή (**Post-Encroachment Time - PET**) υπολογίζονται πλέον σε πραγματικό χρόνο. Νέα πλαίσια αξιολόγησης χρησιμοποιούν τον προβλεπόμενο PET για να εκτιμήσουν το δυνητικό ρίσκο των πεζών, ιδιαίτερα σε μη σηματοδοτούμενες διασταυρώσεις όπου η αβεβαιότητα είναι αυξημένη (Lin et al, 2024). Αντίστοιχα, για τον

δείκτη “Χρόνος μέχρι την Σύγκρουση” (**Time-to-Collision - TTC**), έχουν αναπτυχθεί στιβαρά πλαίσια υπολογισμού βασισμένα στην αλληλεπίδραση πεζών με αυτόνομα οχήματα (shuttles), τα οποία προσφέρουν μεθοδολογίες εφαρμόσιμες και στα συμβατικά οχήματα (Avignone et al, 2024).

Η ανάλυση επεκτείνεται και σε ειδικές κατηγορίες οχημάτων, όπως τα λεωφορεία, όπου συστήματα στερεοσκοπικής όρασης εντός του οχήματος (in-vehicle stereo vision) χρησιμοποιούνται για τον εντοπισμό κυκλοφοριακών συγκρούσεων με πεζούς, προσφέροντας ακριβέστερη αντίληψη βάθους (Cafiso et al, 2017). Παράλληλα, μελέτες πεδίου **σε σηματοδοτούμενες διασταυρώσεις αξιοποιούν αυτά τα εργαλεία** για τη συνολική αξιολόγηση της ασφάλειας των αλληλεπιδράσεων (Gagliardi et al, 2024). Τέλος, η προγνωστική ικανότητα των συστημάτων έχει εξελιχθεί σε σημείο που, χρησιμοποιώντας τη δυναμική του οχήματος και αλγορίθμους όπως το LightGBM, καθίσταται δυνατή η πρόβλεψη ακόμη και παράνομων διελεύσεων (illegal crossing) από πεζούς, επιτρέποντας την έγκαιρη αντίδραση του συστήματος (Roussou et al, 2025).

2.5 Σύνοψη βιβλιογραφίας

Η ανασκόπηση της βιβλιογραφίας καταδεικνύει ότι **η ακατάλληλη ταχύτητα σε διαβάσεις πεζών αποτελεί ένα πολυδιάστατο πρόβλημα** που επηρεάζεται από την υποδομή, το περιβάλλον και την ανθρώπινη αντίληψη. Ενώ υπάρχουν εκτενείς μελέτες για τη συμπεριφορά των οδηγών, παρατηρείται έλλειψη ερευνών που να συνδυάζουν προηγμένους αλγόριθμους τεχνικής όρασης (όπως YOLOv8 και αυτοματοποιημένη βαθμονόμηση (Ventura et al., 2025)) για την ειδική ανάλυση της ταχύτητας σε διαβάσεις της Αθήνας.

Επιπλέον, η βιβλιογραφία αναδεικνύει **την ανάγκη μετάβασης** από την παραδοσιακή, αντιδραστική ανάλυση που βασίζεται σε σπάνια ιστορικά δεδομένα ατυχημάτων, σε προληπτικές προσεγγίσεις. Η χρήση **«έμμεσων δεικτών ασφαλείας» (Surrogate Measures of Safety)** επιτρέπει την αξιολόγηση της επικινδυνότητας σε περιβάλλοντα με έντονη ετερογένεια κυκλοφορίας (όπως η Αθήνα), όπου τα συμβατικά μοντέλα συχνά αδυνατούν να αποδώσουν την πολυπλοκότητα των αλληλεπιδράσεων. **Η παρούσα διπλωματική εργασία** έρχεται να καλύψει αυτό το κενό, αξιοποιώντας τις μεθοδολογίες που αναλύθηκαν για την ακριβή, αυτόματη και μαζική αξιολόγηση της επικινδυνότητας σε πραγματικές συνθήκες.

Η σημασία αυτής της προσέγγισης, ωστόσο, εκτείνεται πέραν της στατιστικής καταγραφής, καθώς αποτελεί τον προθάλαμο για την επόμενη γενιά συστημάτων ασφαλείας. Όταν ενσωματώνεται η Υπολογιστική Όραση με καινοτόμες τεχνολογίες καθίσταται πλέον δυνατή η ακριβής ανίχνευση του κινδύνου σε πραγματικό χρόνο. Είναι η απαραίτητη προϋπόθεση για τη μετάβαση από την παθητική παρατήρηση στην ενεργητική παρέμβαση και την προειδοποίηση των οδηγών (Gaspar et al, 2020).

Κεφάλαιο 3: Θεωρητικό Υπόβαθρο

3.1 Εισαγωγή

Στο παρόν κεφάλαιο παρουσιάζεται το **θεωρητικό υπόβαθρο** στο οποίο στηρίχθηκε η **στατιστική ανάλυση** της παρούσας διπλωματικής εργασίας. Αρχικά, παρατίθενται βασικές έννοιες της στατιστικής, οι οποίες είναι απαραίτητες για την κατανόηση της φύσης των δεδομένων, της διάκρισης των μεταβλητών και των βασικών δεικτών αξιοπιστίας. Στη συνέχεια εξετάζονται οι κυριότερες μέθοδοι παλινδρόμησης που χρησιμοποιούνται στη συγκοινωνιακή έρευνα, με έμφαση στη διωνυμική λογιστική παλινδρόμηση και στα γενικευμένα γραμμικά μοντέλα, τα οποία εφαρμόζονται ευρέως σε αναλύσεις συμπεριφοράς και οδικής ασφάλειας.

Παράλληλα παρουσιάζονται τα βασικά κριτήρια αποδοχής και αξιολόγησης ενός στατιστικού υποδείγματος, όπως η **στατιστική σημαντικότητα** των συντελεστών, η καλή προσαρμογή, η πιθανοφάνεια και τα **κριτήρια πληροφοριών**. Η ανάπτυξη του θεωρητικού αυτού πλαισίου είναι αναγκαία, ώστε η μεθοδολογική προσέγγιση και η ερμηνεία των αποτελεσμάτων που ακολουθούν στα επόμενα κεφάλαια να είναι συνεπείς, τεκμηριωμένες και επιστημονικά ορθές.

3.2 Βασικές έννοιες στατιστικής

Στο πλαίσιο μιας στατιστικής έρευνας, ως **πληθυσμός** ορίζεται το σύνολο των μονάδων που αποτελούν το αντικείμενο μελέτης ως προς ένα ή περισσότερα χαρακτηριστικά. Επειδή στην πράξη η εξέταση του συνόλου του πληθυσμού είναι συνήθως δύσκολη ή και ανέφικτη, η ανάλυση βασίζεται σε ένα **δείγμα**, δηλαδή σε ένα υποσύνολο του πληθυσμού, το οποίο επιλέγεται με κατάλληλη διαδικασία δειγματοληψίας. Η εγκυρότητα των συμπερασμάτων εξαρτάται σε σημαντικό βαθμό από το κατά πόσο το δείγμα είναι αντιπροσωπευτικό του πληθυσμού από τον οποίο προέρχεται.

Τα χαρακτηριστικά που παρατηρούνται και καταγράφονται για κάθε μονάδα του δείγματος ονομάζονται **μεταβλητές**. Οι μεταβλητές διακρίνονται σε **ποιοτικές και ποσοτικές**. Οι ποιοτικές μεταβλητές μπορούν να περιγράψουν κατηγορίες ή ιδιότητες, όπως για παράδειγμα την ένδειξη του φωτεινού σηματοδότη ή ακόμη και την παρουσία οχήματος στην διάβαση πεζών, ενώ οι ποσοτικές μεταβλητές λαμβάνουν αριθμητικές τιμές και δύνανται να μετρηθούν. Οι ποσοτικές μεταβλητές διακρίνονται περαιτέρω σε διακριτές, όταν λαμβάνουν μεμονωμένες αριθμητικές τιμές, και σε συνεχείς, όταν μπορούν να λάβουν οποιαδήποτε τιμή εντός ενός διαστήματος.

Για την αποτίμηση της αξιοπιστίας των στατιστικών εκτιμήσεων χρησιμοποιούνται συνήθως το **επίπεδο εμπιστοσύνης** και το **επίπεδο σημαντικότητας**. Το επίπεδο εμπιστοσύνης εκφράζει τον βαθμό βεβαιότητας με τον οποίο ένα διάστημα εμπιστοσύνης περιέχει την πραγματική τιμή μιας παραμέτρου του πληθυσμού. Στις περισσότερες εφαρμογές υιοθετείται επίπεδο εμπιστοσύνης 95%, το οποίο αντιστοιχεί σε επίπεδο σημαντικότητας 5%. Το επίπεδο σημαντικότητας, που συμβολίζεται συνήθως με α , εκφράζει την πιθανότητα να απορριφθεί εσφαλμένα μία αληθής μηδενική υπόθεση.

Ιδιαίτερη σημασία έχει επίσης η έννοια της **συσχέτισης** μεταξύ μεταβλητών. Ο συντελεστής συσχέτισης Pearson (r) λαμβάνει τιμές στο διάστημα $[-1, 1]$ και εκφράζει τον βαθμό και την κατεύθυνση της γραμμικής σχέσης μεταξύ δύο ποσοτικών μεταβλητών. Τιμές κοντά στο 1 υποδηλώνουν ισχυρή θετική συσχέτιση, τιμές κοντά στο -1 ισχυρή αρνητική συσχέτιση, ενώ τιμές κοντά στο 0 υποδηλώνουν απουσία ισχυρής γραμμικής σχέσης. Στα υποδείγματα παλινδρόμησης, η υψηλή συσχέτιση μεταξύ ανεξάρτητων μεταβλητών είναι δυνατόν να οδηγήσει σε προβλήματα **πολυσυστασιαστικότητας** και να δυσχεράνει την αξιόπιστη ερμηνεία των εκτιμώμενων συντελεστών.

3.3 Βασικές μέθοδοι στατιστικής ανάλυσης

Η **ανάλυση παλινδρόμησης** αποτελεί έναν από τους σημαντικότερους κλάδους της εφαρμοσμένης στατιστικής και χρησιμοποιείται για τη διερεύνηση της σχέσης μεταξύ μιας **εξαρτημένης μεταβλητής** και ενός συνόλου ανεξάρτητων μεταβλητών. Βασικός σκοπός της είναι είτε η ερμηνεία του τρόπου με τον οποίο οι ανεξάρτητες μεταβλητές επηρεάζουν την εξαρτημένη, είτε η πρόβλεψη της τιμής ή της πιθανότητας εμφάνισής της.

Η **μορφή της εξαρτημένης μεταβλητής**, δηλαδή αν είναι συνεχής, δυαδική, ονομαστική ή διατεταγμένη, καθορίζει σε μεγάλο βαθμό και τη **μεθοδολογική προσέγγιση** που κρίνεται κατάλληλη. Στη συγκοινωνιακή έρευνα έχουν χρησιμοποιηθεί διάφορες μορφές παλινδρόμησης, όπως η γραμμική παλινδρόμηση, τα λογιστικά υποδείγματα, τα probit υποδείγματα και άλλες μέθοδοι διακριτής επιλογής. Η επιλογή της καταλληλότερης μεθόδου εξαρτάται από τη φύση των δεδομένων, τη θεωρητική δομή του προβλήματος και τους στόχους της έρευνας.

3.3.1 Γραμμική παλινδρόμηση

Η **γραμμική παλινδρόμηση** χρησιμοποιείται όταν η εξαρτημένη μεταβλητή είναι **συνεχής** και η σχέση της με τις ανεξάρτητες μεταβλητές μπορεί να προσεγγιστεί ικανοποιητικά μέσω γραμμικής συνάρτησης. Στο απλούστερο σχήμα της, η αναμενόμενη τιμή της εξαρτημένης μεταβλητής εκφράζεται ως γραμμικός συνδυασμός των ανεξάρτητων μεταβλητών και ενός όρου σφάλματος.

Η εκτίμηση των παραμέτρων πραγματοποιείται συνήθως με τη **μέθοδο των ελαχίστων τετραγώνων**, ώστε να ελαχιστοποιείται το άθροισμα των τετραγώνων των αποκλίσεων μεταξύ παρατηρούμενων και εκτιμημένων τιμών. Στις περιπτώσεις όπου η εξαρτημένη μεταβλητή είναι κατηγορική, και ειδικότερα **διχοτομική**, η γραμμική παλινδρόμηση δεν αποτελεί την πλέον κατάλληλη μεθοδολογική επιλογή, επειδή μπορεί να αποδώσει προβλεπόμενες τιμές πέραν του διαστήματος $[0,1]$, ενώ παράλληλα ενδέχεται να παραβιάζονται θεμελιώδεις υποθέσεις στις οποίες στηρίζεται το γραμμικό υπόδειγμα.

3.3.2 Probit ανάλυση

Τα **probit υποδείγματα** χρησιμοποιούνται για την ανάλυση διακριτών επιλογών, όταν η πιθανότητα πραγματοποίησης ενός γεγονότος συνδέεται με τις ανεξάρτητες μεταβλητές μέσω της αθροιστικής συνάρτησης της **κανονικής κατανομής**. Η λογική τους είναι παρεμφερής με εκείνη της λογιστικής παλινδρόμησης, διαφοροποιούνται όμως ως προς τη μορφή της συνάρτησης σύνδεσης.

Τα probit μοντέλα είναι ιδιαίτερα χρήσιμα σε θεωρητικές εφαρμογές διακριτής επιλογής, όπου η υπόθεση κανονικότητας του σφάλματος κρίνεται καταλληλότερη. Παρά ταύτα, σε πολλές εφαρμογές η **λογιστική παλινδρόμηση** προτιμάται, λόγω της ευκολότερης ερμηνείας των συντελεστών μέσω των **odds ratios** και της ευρύτερης χρήσης της στη διεθνή και ελληνική βιβλιογραφία.

3.3.3 Διαχωριστική ανάλυση

Η **διαχωριστική ανάλυση** αποσκοπεί στην **ταξινόμηση παρατηρήσεων** σε προκαθορισμένες κατηγορίες με βάση ένα σύνολο επεξηγηματικών μεταβλητών. Η μέθοδος αυτή έχει χρησιμοποιηθεί σε αρκετές ερευνητικές περιοχές, προϋποθέτει όμως ισχυρότερες στατιστικές υποθέσεις ως προς την κατανομή των δεδομένων και την ομοιογένεια των διακυμάνσεων.

Για ερωτήματα που αφορούν κυρίως την **πιθανότητα εμφάνισης** ενός γεγονότος ή την επίδραση συγκεκριμένων χαρακτηριστικών σε μια δυαδική επιλογή, τα **λογιστικά υποδείγματα** θεωρούνται συνήθως καταλληλότερα και περισσότερο ευέλικτα ως προς την εφαρμογή και την ερμηνεία τους.

3.3.4 Λογιστική παλινδρόμηση

Η **λογιστική παλινδρόμηση** χρησιμοποιείται όταν η εξαρτημένη μεταβλητή είναι κατηγορική. Στην απλούστερη μορφή της, η **διωνυμική λογιστική παλινδρόμηση** εφαρμόζεται όταν η εξαρτημένη μεταβλητή λαμβάνει αποκλειστικά δύο τιμές, όπως «ναι/όχι», «συμβαίνει/δεν συμβαίνει» ή «νόμιμη/παράνομη επιλογή».

Η μέθοδος αυτή επιτρέπει την εκτίμηση της **πιθανότητας εμφάνισης** ενός γεγονότος ως συνάρτηση ενός συνόλου ανεξάρτητων μεταβλητών και χρησιμοποιείται εκτενώς σε αναλύσεις συμπεριφοράς μετακίνησης, οδικής ασφάλειας και διακριτής επιλογής. Το βασικό της πλεονέκτημα είναι ότι οι προβλεπόμενες τιμές περιορίζονται στο **διάστημα [0,1]**, ενώ παράλληλα παρέχεται η δυνατότητα ερμηνείας της κατεύθυνσης και της έντασης επίδρασης κάθε μεταβλητής.

3.4 Διωνυμική λογιστική παλινδρόμηση

Η **διωνυμική λογιστική παλινδρόμηση** περιγράφει τη σχέση μεταξύ μιας δυαδικής εξαρτημένης μεταβλητής και ενός συνόλου ανεξάρτητων μεταβλητών μέσω της **συνάρτησης logit**. Αν P_i είναι η πιθανότητα πραγματοποίησης του γεγονότος i , τότε η λογαριθμισμένη αναλογία πιθανοτήτων δίνεται από την ακόλουθη σχέση:

$$\log(P_i / (1 - P_i)) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k \quad (Eq. 1)$$

όπου β_0 είναι ο **σταθερός όρος** και $\beta_1, \beta_2, \dots, \beta_k$ οι **συντελεστές** των ανεξάρτητων μεταβλητών.

Από την παραπάνω σχέση προκύπτει η **πιθανότητα πραγματοποίησης** του γεγονότος:

$$P_i = e^{(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)} / [1 + e^{(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)}] \quad (Eq. 2)$$

Η εκτίμηση των παραμέτρων δεν πραγματοποιείται με τη μέθοδο των ελαχίστων τετραγώνων αλλά με τη μέθοδο της **μέγιστης πιθανοφάνειας**, η οποία αναζητά το σύνολο των παραμέτρων που καθιστά τα παρατηρούμενα δεδομένα περισσότερο

πιθανά. Οι εκτιμώμενοι συντελεστές ερμηνεύονται ως μεταβολές στο log-odds του γεγονότος, ενώ με κατάλληλο εκθετικό μετασχηματισμό προκύπτουν τα **odds ratios**, τα οποία είναι ιδιαίτερα χρήσιμα στην πρακτική ερμηνεία των αποτελεσμάτων.

Η λογιστική παλινδρόμηση μπορεί να επεκταθεί τόσο σε **πολυωνυμικά** όσο και σε **διατεταγμένα υποδείγματα**, όταν η εξαρτημένη μεταβλητή έχει περισσότερες από δύο κατηγορίες. Στις περιπτώσεις αυτές η μεθοδολογική βάση παραμένει κοινή, διαφοροποιείται όμως η μορφή της συνάρτησης σύνδεσης και η ερμηνεία των αντίστοιχων πιθανοτήτων.

3.5 Γενικευμένα Γραμμικά Μοντέλα (Generalized Linear Models)

Τα **Γενικευμένα Γραμμικά Μοντέλα** (Generalized Linear Models – GLM) αποτελούν γενίκευση της κλασικής γραμμικής παλινδρόμησης και επιτρέπουν τη μοντελοποίηση εξαρτημένων μεταβλητών που δεν ακολουθούν κατ' ανάγκη κανονική κατανομή. Το πλαίσιο των GLM βασίζεται σε τρία βασικά στοιχεία: την τυχαία συνιστώσα, δηλαδή την κατανομή της εξαρτημένης μεταβλητής, τη συστηματική συνιστώσα, δηλαδή τον γραμμικό προβλεπτή των ανεξάρτητων μεταβλητών, και τη **συνάρτηση σύνδεσης**, η οποία συνδέει τη μέση τιμή της εξαρτημένης μεταβλητής με τον γραμμικό προβλεπτή.

Συνεπώς, τα λογιστικά υποδείγματα δεν βρίσκονται εκτός του πλαισίου των GLM, αλλά αποτελούν **ειδική περίπτωση** τους, όταν η εξαρτημένη μεταβλητή ακολουθεί **διωνυμική κατανομή** και χρησιμοποιείται ως συνάρτηση σύνδεσης το logit. Αντίστοιχα, για διατεταγμένες κατηγορικές εξαρτημένες μεταβλητές μπορεί να χρησιμοποιηθεί η ordinal logistic regression, η οποία είναι κατάλληλη όταν οι κατηγορίες διαθέτουν φυσική σειρά.

Τα **GLM** χρησιμοποιούνται ευρέως σε εφαρμογές όπου η μορφή της απόκρισης αποκλίνει από τις προϋποθέσεις της γραμμικής παλινδρόμησης, όπως σε δυαδικά, μετρητικά ή **διατεταγμένα δεδομένα**. Στη συγκοινωνιακή έρευνα προσφέρουν ιδιαίτερα χρήσιμο εργαλείο για την ανάλυση επιλογών, προτιμήσεων, επιπέδων επικινδυνότητας και συμπεριφορών χρηστών του οδικού δικτύου.

3.6 Κριτήρια αποδοχής και αξιολόγησης μοντέλου

Η αξιολόγηση ενός στατιστικού μοντέλου δεν περιορίζεται μόνο στη μαθηματική του διατύπωση, αλλά επεκτείνεται στην **επιστημονική ερμηνευσιμότητα** και στη **στατιστική επάρκεια** των αποτελεσμάτων. Ένα αποδεκτό μοντέλο οφείλει να παρουσιάζει συντελεστές με πρόσημα συμβατά με τη θεωρία και τη λογική του

προβλήματος, ικανοποιητική στατιστική σημαντικότητα και επαρκή καλή προσαρμογή στα δεδομένα.

3.6.1 Συντελεστές εξίσωσης

Το **πρόσημο** κάθε εκτιμώμενου συντελεστή υποδηλώνει την κατεύθυνση της σχέσης μεταξύ της αντίστοιχης ανεξάρτητης μεταβλητής και της εξαρτημένης. Θετικός συντελεστής σημαίνει ότι η αύξηση της μεταβλητής συνδέεται με αύξηση της πιθανότητας ή του επιπέδου του εξεταζόμενου γεγονότος, ενώ αρνητικός συντελεστής υποδηλώνει την αντίθετη τάση. Στα λογιστικά μοντέλα, η ερμηνεία αφορά πρωτίστως τη μεταβολή των **log-odds** και δευτερευόντως, μέσω του e^{β} , τη μεταβολή στον λόγο πιθανοτήτων.

3.6.2 Στατιστική σημαντικότητα

Η **στατιστική σημαντικότητα** των παραμέτρων εξετάζεται συνήθως μέσω του ελέγχου Wald, του αντίστοιχου z- ή t-στατιστικού, καθώς και μέσω του **p-value**. Σε επίπεδο σημαντικότητας 5%, μία μεταβλητή θεωρείται στατιστικά σημαντική όταν το p-value είναι μικρότερο από 0,05. Η ύπαρξη στατιστικά σημαντικών συντελεστών ενισχύει την αξιοπιστία του μοντέλου, υπό την προϋπόθεση ότι η ερμηνεία των αποτελεσμάτων παραμένει θεωρητικά συνεπής.

3.6.3 Συσχέτιση ανεξάρτητων μεταβλητών

Η ύπαρξη **υψηλής συσχέτισης** μεταξύ ανεξάρτητων μεταβλητών μπορεί να προκαλέσει προβλήματα **πολυσυγγραμμικότητας** (μεταβλητές να συσχετίζονται σε μεγάλο βαθμό) , γεγονός που οδηγεί σε ασταθείς εκτιμήσεις συντελεστών και δυσχεραίνει την αξιόπιστη αποτίμηση της επίδρασης κάθε μεταβλητής χωριστά. Για τον λόγο αυτό προηγείται συνήθως έλεγχος συσχέτισης, καθώς και συμπληρωματικοί διαγνωστικοί έλεγχοι όπου απαιτείται.

3.6.4 Δείκτες καλής προσαρμογής

Στα γραμμικά μοντέλα, η ποιότητα προσαρμογής εκτιμάται συνήθως μέσω του **συντελεστή προσδιορισμού R^2** . Στα λογιστικά μοντέλα, επειδή η εξαρτημένη μεταβλητή είναι κατηγορική, χρησιμοποιούνται συχνότερα οι **pseudo- R^2** δείκτες, όπως οι Cox–Snell και Nagelkerke, καθώς και έλεγχοι καλής προσαρμογής όπως το Hosmer–Lemeshow test. Οι δείκτες αυτοί δεν ερμηνεύονται με τον ίδιο ακριβώς τρόπο όπως το κλασικό R^2 , παρέχουν όμως χρήσιμη εικόνα για τη σχετική ποιότητα προσαρμογής του υποδείγματος.

3.6.5 Μέγιστη πιθανοφάνεια και λόγος πιθανοφάνειας

Η εκτίμηση των λογιστικών και, γενικότερα, των GLM υποδειγμάτων βασίζεται στη **μέγιστη πιθανοφάνεια**. Για τη σύγκριση εναλλακτικών μοντέλων χρησιμοποιείται συχνά ο έλεγχος λόγου πιθανοφάνειας (Likelihood Ratio Test), ο οποίος αξιολογεί κατά πόσο ένα πιο σύνθετο μοντέλο υπερέχει στατιστικά έναντι ενός απλούστερου. Όσο καλύτερη είναι η προσαρμογή του μοντέλου στα δεδομένα, τόσο ευνοϊκότερη είναι η τιμή της λογαριθμικής πιθανοφάνειας.

3.6.6 Κριτήρια πληροφοριών AIC και BIC

Τα κριτήρια **AIC** και **BIC** χρησιμοποιούνται για τη σύγκριση μοντέλων, λαμβάνοντας υπόψη όχι μόνο την καλή προσαρμογή αλλά και την πολυπλοκότητά τους. Χαμηλότερες τιμές AIC και BIC υποδηλώνουν γενικά προτιμητέο μοντέλο, με το BIC να επιβάλλει αυστηρότερη ποινή για την εισαγωγή μεγάλου αριθμού παραμέτρων. Τα κριτήρια αυτά είναι ιδιαίτερα χρήσιμα όταν εξετάζονται περισσότερες από μία εναλλακτικές προδιαγραφές υποδειγμάτων.

3.7 Θεωρία στοχαστικής χρησιμότητας

Η **Θεωρία Στοχαστικής Χρησιμότητας** αποτελεί τη βασική θεωρητική βάση των **μοντέλων διακριτής επιλογής**. Σύμφωνα με αυτή, κάθε άτομο αποδίδει σε κάθε διαθέσιμη εναλλακτική μία συνολική χρησιμότητα, η οποία αποτελείται από ένα συστηματικό και ένα τυχαίο μέρος. Η χρησιμότητα της εναλλακτικής i για το άτομο n εκφράζεται ως εξής:

$$U_{in} = V_{in} + \varepsilon_{in} \quad (\text{Eq. 3})$$

όπου V_{in} είναι το **προσδιοριστικό μέρος** της χρησιμότητας, το οποίο εξαρτάται από παρατηρούμενα χαρακτηριστικά, και ε_{in} ο **στοχαστικός όρος**, ο οποίος αποτυπώνει μη παρατηρούμενους ή μη μετρήσιμους παράγοντες.

Το άτομο θεωρείται ότι επιλέγει την εναλλακτική με τη **μέγιστη χρησιμότητα**. Από τη μορφή της κατανομής του στοχαστικού όρου προκύπτουν διαφορετικά μοντέλα διακριτής επιλογής, όπως τα logit και τα probit. Η θεωρία αυτή είναι ιδιαίτερα χρήσιμη σε έρευνες μεταφορικής συμπεριφοράς, καθώς επιτρέπει τη σύνδεση των ατομικών επιλογών με ποσοτικοποιήσιμα χαρακτηριστικά του περιβάλλοντος, του χρήστη ή του σεναρίου απόφασης.

3.8 Σύνδεση του θεωρητικού πλαισίου με την παρούσα εργασία

Με βάση τα παραπάνω, η επιλογή της **διωνυμικής λογιστικής παλινδρόμησης** και των **γενικευμένων γραμμικών μοντέλων** είναι θεωρητικά τεκμηριωμένη για προβλήματα όπου η εξαρτημένη μεταβλητή εκφράζει κατηγορικές ή διακριτές αποκρίσεις. Στο πλαίσιο της παρούσας διπλωματικής, τα υποδείγματα αυτά παρέχουν τη δυνατότητα να διερευνηθεί η επίδραση επιμέρους χαρακτηριστικών στην πιθανότητα εκδήλωσης μιας συγκεκριμένης συμπεριφοράς ή στη μεταβολή κρίσιμων δεικτών που συνδέονται με την κυκλοφοριακή συμπεριφορά και την οδική ασφάλεια.

Η **θεωρητική θεμελίωση** που προηγήθηκε υποστηρίζει τη μεθοδολογία που ακολουθεί στα επόμενα κεφάλαια και επιτρέπει τη συνεπή ερμηνεία των ευρημάτων. Κατά συνέπεια, το παρόν κεφάλαιο λειτουργεί ως γέφυρα μεταξύ της βιβλιογραφικής ανασκόπησης και της **εμπειρικής ανάλυσης** που εφαρμόζεται στην παρούσα εργασία.

Κεφάλαιο 4: Συλλογή και Επεξεργασία Στοιχείων

4.1 Εισαγωγή

Στο παρόν κεφάλαιο παρουσιάζονται τα δεδομένα που αξιοποιήθηκαν για να διερευνηθούν οι **παράγοντες παράνομης διέλευσης** των οχημάτων σε διάβαση πεζών και συγκεκριμένα οι **μεταβλητές** που σχετίζονται με την **παραβατική συμπεριφορά των οδηγών**.

Συγκεκριμένα, περιγράφεται η **διαδικασία συλλογής και επεξεργασίας των δεδομένων**, με την οποία διαμορφώθηκε το **τελικό αρχείο** που χρησιμοποιήθηκε στην ανάλυση. ώστε να δοθεί πλήρης εικόνα της **αξιοπιστίας και της ποιότητάς τους** και να αποσαφηνίζεται το ακριβές περιεχόμενο κάθε μεταβλητής του τελικού πίνακα.

Αρχικά, παρουσιάζεται η **μεθοδολογία καταγραφής** των στοιχείων που χρησιμοποιήθηκαν, ενώ στη συνέχεια δίνεται έμφαση στον προσδιορισμό της **ακρίβειας των ενδείξεων** του **φωτεινού σηματοδότη** ανά κύκλο εναλλαγής, καθώς και της ακρίβειας στον εντοπισμό **νόμιμων και παράνομων περασμάτων**.

Τα δεδομένα αφορούν τη διάβαση στην **Λεωφόρο Βασιλέως Κωνσταντίνου**, στο ύψος του **Παναθηναϊκού Σταδίου**, και περιλαμβάνουν:

1. **Δεδομένα αλγόριθμοι** για την **αυτόματη καταγραφή παράνομων περασμάτων** οχημάτων και **διαβάσεων πεζών**,
2. **Χειροκίνητες καταγραφές** από **μετρήσεις πεδίου**,
3. Την **ταχύτητα** και τον **χρόνο αντίδρασης πεζών και οχημάτων**.

Επιπλέον, παρουσιάζονται στους παρακάτω πίνακες τα **περιγραφικά στατιστικά στοιχεία** της **βάσης δεδομένων**, με σκοπό την κατανόηση της **κατανομής**, της **μεταβλητότητας** και των βασικών χαρακτηριστικών των υπό ανάλυση μεταβλητών.

4.2 Συλλογή δεδομένων

Τα δεδομένα που χρησιμοποιήθηκαν στα πλαίσια της συγκεκριμένης **Διπλωματικής Εργασίας** συλλέχθηκαν από **καταγραφή βίντεο** με χρήση **κινητού τηλεφώνου** τοποθετημένο σε τρίποδο διάρκειας μιας ώρας **πρωινής αιχμής**. Η τοποθεσία που επιλέχθηκε είναι στο κέντρο της Αθήνας και συγκεκριμένα στη Λεωφόρο Βασιλέως Κωνσταντίνου, στο ύψος του Παναθηναϊκού Σταδίου.

Το βίντεο που καταγράφηκε εισήχθη στον **αλγόριθμο** που έχει υλοποιηθεί από την ερευνητική ομάδα του εργαστηρίου Κυκλοφοριακής Τεχνικής (Ventura et al., 2025) βασισμένος στην **υπολογιστική όραση** (computer vision) και χρησιμοποιείται για την παρακολούθηση πεζών και οχημάτων, την ανάλυση της συμπεριφοράς τους, την

εκτίμηση του Χρόνου έως τη Σύγκρουση (**Time-to-Collision - TTC**) και του δείκτη **Post Encroachment Type (PET)**. Ο αλγόριθμος αποτελείται από τα παρακάτω:

1. Χρήση των μοντέλων **YOLOv8** για **ανίχνευση αντικειμένων** και **ResNet-50** για **εξαγωγή χαρακτηριστικών**.
2. Εφαρμογή δικτύου **Re-Identification (Re-ID)** για **αναγνώριση και διατήρηση των ταυτοτήτων** των αντικειμένων.
3. Ορισμός **ομογραφίας (homography)** για **μετατροπή των συντεταγμένων** των αντικειμένων από το επίπεδο του βίντεο στο **πραγματικό επίπεδο του δρόμου**.
4. Πρόβλεψη κίνησης με **Kalman Filter** για να μειώσει το **θόρυβο** και να διατηρήσει **σταθερότητα στην παρακολούθηση**.
5. **Αντιστοίχιση αντικειμένων** μεταξύ **διαδοχικών καρτέ** με χρήση του **Hungarian Algorithm**.
6. **Φιλτράρισμα θέσεων** σε **χαμένα αντικείμενα** με **Savitzky-Golay Filter**.
7. Ανίχνευση της κατάστασης του **φωτεινού σηματοδότη** μέσω **ανάλυσης χρωμάτων** σε συγκεκριμένες **περιοχές ενδιαφέροντος** (Region of Interest).
8. **Ανάλυση συμπεριφοράς** πεζών και οχημάτων, εντοπίζοντας **παραβάσεις σηματοδότησης** (π.χ. παράνομες διαβάσεις).
9. **Καταμέτρηση πεζών και οχημάτων** σε διάφορα **σημεία διέλευσης**.
10. **Υπολογισμός ταχύτητας** αντικειμένων και εκτίμηση **Χρόνου έως Σύγκρουση (TTC)**.

Η ανάλυση του βίντεο πραγματοποιήθηκε έχοντας ορίσει ως ανώτατο όριο **29.74** (σχεδόν 30) καρτέ (**Frames per Second - fps**), ώστε να αποφευχθεί η **υπερφόρτωση της μνήμης**.

Η εικόνα (Εικόνα 1) δείχνει τα **σημεία αναφοράς** στην **προοπτική του βίντεο**, ενώ η δεύτερη εικόνα (Εικόνα 2) δείχνει τα αντίστοιχα σημεία στο **επίπεδο του χάρτη**. Αυτά τα σημεία είναι απαραίτητα για την προσαρμογή των συντεταγμένων μεταξύ των δύο επιπέδων, επιτρέποντας την ακριβή αντιστοίχιση των δεδομένων μεταξύ βίντεο και **πραγματικού κόσμου**. Ουσιαστικά, χρησιμοποιούνται για την ακριβή μετατροπή και χαρτογράφηση της θέσης αντικειμένων στο βίντεο σε πραγματικές συντεταγμένες.



Εικόνα 4.1: Σημεία αναφοράς στην προοπτική του βίντεο



Εικόνα 4.2: Σημεία αναφοράς στο επίπεδο του χάρτη



Εικόνα 4.3: Φωτεινοί Σηματοδότες Οχημάτων

Στην ανίχνευση της κατάστασης **φωτεινών σηματοδοτών οχημάτων**, καθορίστηκαν περιοχές ενδιαφέροντος (**Region Of Interest -ROIs**) για τα **κόκκινα, πορτοκαλί και πράσινα** φώτα του σηματοδότη (Εικόνα 3). Παρόμοια, στην ανάλυση της συμπεριφοράς πεζών και οχημάτων, καθορίστηκε η περιοχή ενδιαφέροντος της **διάβασης πεζών** (Εικόνα 1).

Για την καταμέτρηση πεζών και οχημάτων, ορίστηκαν **έξι «είσοδοι» ανίχνευσης** για κάθε σημείο: **τρεις για πεζούς** και **τρεις για οχήματα**. Οι είσοδοι τοποθετήθηκαν στη μέση της διάβασης και στις άκρες της, ενώ οι είσοδοι των οχημάτων τοποθετήθηκαν στη μέση, ανάντη και κατάντη της διάβασης (Εικόνα 1). Ο στόχος ήταν η **διασταύρωση των μετρήσεων** για αξιόπιστα αποτελέσματα και η αποφυγή υπερ- ή υπο-καταμέτρησης.

Οι **μέγιστες επιτρεπτές ταχύτητες** ορίστηκαν σε **10 m/s** για πεζούς (ταχύτητα τρεξίματος) και **40 m/s** για οχήματα, ώστε να ληφθούν υπόψη σενάρια υπερβολικής ταχύτητας αλλά να αποφευχθούν ακραίες, **μη ρεαλιστικές τιμές**. Αυτές οι ρυθμίσεις διασφαλίζουν ότι το σύστημα είναι σταθερό, αποδοτικό και προσαρμόζεται στις διαφορετικές συνθήκες του περιβάλλοντος.

Η **συλλογή υλικού μέσω βιντεοσκόπησης** της εξεταζόμενης διάβασης αποτέλεσε τη βάση για την ανάπτυξη μιας ολοκληρωμένης μεθοδολογίας ανάλυσης δεδομένων. Μέσω προηγμένων τεχνικών επεξεργασίας στο λογισμικό **MATLAB**, πραγματοποιήθηκε ανάλυση και επεξεργασία των καταγραφών, οδηγώντας στη δημιουργία της **βάσης δεδομένων** της αυτοματοποιημένης μεθόδου. Παράλληλα, το

ίδιο υλικό παρακολουθήθηκε λεπτομερώς, καταγράφοντας τα δεδομένα χειροκίνητα. Με αυτόν τον τρόπο, διαμορφώθηκε τη βάση δεδομένων της χειροκίνητης μεθόδου.

Όλα όσα προαναφέρθηκαν αξιοποιήθηκαν κατά τρόπο που θα αναλυθεί παρακάτω με τη χρήση της **γλώσσας προγραμματισμού Python**, εφαρμόζοντας κατάλληλες **βιβλιοθήκες και εργαλεία** για την επεξεργασία, ανάλυση και **οπτικοποίηση των δεδομένων**.

4.3 Προεπεξεργασία και ενοποίηση των αρχικών δεδομένων

Αρχικά, εξάγεται αρχείο **vehicle_data.csv** στο οποίο πραγματοποιήθηκε μετατροπή των τιμών της ταχύτητας που περιλαμβάνονται στη **στήλη magnitude**, ώστε οι τιμές να εκφράζονται στη σωστή μονάδα μέτρησης, δηλαδή σε **km/h** αντί για m/sec. Η παρέμβαση αυτή ήταν απαραίτητη, ώστε όλες οι επόμενες αναλύσεις να βασίζονται σε **ενιαία και άμεσα ερμηνεύσιμη μονάδα ταχύτητας**.

```
import csv
import math

input_file = "vehicle_data.csv"
output_file = "vehicle_data_kmh.csv"

with open(input_file, newline="", encoding="utf-8") as f:
    reader = csv.DictReader(f)
    fieldnames = reader.fieldnames
    rows = list(reader)

    for row in rows:
        value = row.get("magnitude", "")
        try:
            value = float(value)
            if not math.isnan(value):
                row["magnitude"] = value * 3.6
        except ValueError:
            pass

    with open(output_file, "w", newline="", encoding="utf-8") as f:
        writer = csv.DictWriter(f, fieldnames=fieldnames)
        writer.writeheader()
        writer.writerows(rows)

print(f"Converted 'magnitude' from m/s to km/h and saved to {output_file}")
```

Εικόνα 4.4. Ενδεικτική απεικόνιση της εντολής μετατροπής της ταχύτητας στη στήλη magnitude.

Στη συνέχεια, διαπιστώθηκε ότι ο αλγόριθμος ανίχνευσης αναγνώριζε ορισμένες **μοτοσυκλέτες ως πεζούς**. Για τον λόγο αυτό, από το αρχείο **pedestrian_data.csv**

απομονώθηκαν χειροκίνητα οι εγγραφές που αντιστοιχούσαν σε μηχανές, έγινε **επαναταξινόμησή τους** με νέα vehicle IDs και κατόπιν οι εγγραφές αυτές ενσωματώθηκαν στο **βασικό αρχείο** μέσω χειροκίνητης αντιγραφής και επικόλλησης. Με τον τρόπο αυτό εξασφαλίστηκε ότι τα δίκυκλα δεν θα λογίζονται εσφαλμένα ως πεζοί κατά την επόμενη ανάλυση.

Παράλληλα, για τις μοτοσικλές διορθώθηκε κατάλληλα και η μεταβλητή **TrafficLightStatus**, ώστε ως **illegal** να καταγράφονται οι περιπτώσεις στις οποίες η διέλευση πραγματοποιήθηκε υπό κόκκινη ένδειξη σηματοδότη. Επιπλέον, δημιουργήθηκε η νέα στήλη **Vehicle_Classification**, στην οποία αποδόθηκε η τιμή car για τα τετράτροχα οχήματα και η τιμή moto για τις μοτοσικλές. Η ίδια διόρθωση μονάδων ταχύτητας εφαρμόστηκε και στις μοτοσικλές, ώστε όλο το σύνολο δεδομένων να είναι **συνεπές** ως προς την αποτύπωση της ταχύτητας.

```
import csv
from decimal import Decimal, InvalidOperation, getcontext

# Increase precision for decimal calculations
getcontext().prec = 28

input_file = "vehicle_data_kmh.csv"
output_file = "vehicle_data_kmh_converted.csv"

def format_decimal(value: Decimal) -> str:
    """
    Convert Decimal to plain string without scientific notation
    and remove unnecessary trailing zeros.
    """
    s = format(value, "f")
    if "." in s:
        s = s.rstrip("0").rstrip(".")
    return s

with open(input_file, "r", encoding="utf-8-sig", newline="") as infile, \
    open(output_file, "w", encoding="utf-8", newline="") as outfile:

    reader = csv.reader(infile)
    writer = csv.writer(outfile)

    header = next(reader)
    writer.writerow(header)

    # Find the column positions by header name
    magnitude_idx = header.index("magnitude") # K column
    vehicle_class_idx = header.index("Vehicle Classification") # P column

    for row in reader:
        # Check if the row is long enough
        if len(row) <= max(magnitude_idx, vehicle_class_idx):
            writer.writerow(row)
            continue

        # Only modify rows where Vehicle Classification = "moto"
        if row[vehicle_class_idx].strip().lower() == "moto":
            value = row[magnitude_idx].strip()

            if value != "":
                try:
                    new_value = Decimal(value) * Decimal("3.6")
                    row[magnitude_idx] = format_decimal(new_value)
                except InvalidOperation:
                    # Leave the value unchanged if it is not numeric
                    pass

            writer.writerow(row)

    print(f"Done. New file saved as: {output_file}")
```

Εικόνα 4.5. Ενδεικτική απεικόνιση του σταδίου ενοποίησης και επαναταξινόμησης των μοτοσικλετών στο τελικό αρχείο

4.4 Προσδιορισμός προπορευόμενου οχήματος και γεωμετρικών σχέσεων κίνησης

Μετά την ολοκλήρωση της βασικής προεπεξεργασίας, ακολούθησε η δημιουργία μεταβλητών που περιγράφουν τη σχέση κάθε οχήματος με το **προπορευόμενό του**. Η διαδικασία βασίστηκε στη χωρική θέση των οχημάτων ανά frame, στη διεύθυνση κίνησης και στη σχετική εγκάρσια απόστασή τους, ώστε να εκτιμηθεί αν δύο οχήματα ανήκουν στην **ίδια ροή ή λωρίδα κυκλοφορίας**. Από τη διαδικασία αυτή προέκυψαν οι στήλες **LeadingVehicleID**, **GapAlongPath_m**, **LateralOffsetToLeader_m** και **HasLeaderSameLane**.

```
# =====
# Υπολογισμός προπορευόμενου οχήματος και gap
# κατά μήκος πορείας, ανά frame
# =====

# Εισαγωγή των βασικών βιβλιοθηκών.
# Η pandas χρησιμοποιείται για ανάγνωση/διαχείριση του CSV,
# ενώ η numpy για αριθμητικούς υπολογισμούς και πράξεις πινάκων.
import pandas as pd
import numpy as np

# -----
# 1. Ορισμός αρχείων εισόδου και εξόδου
# -----

# Όνομα του αρχικού αρχείου CSV που περιέχει τα δεδομένα των οχημάτων.
input_file = "vehicle_data_kmh_moto_status_reversed.csv"

# Όνομα του νέου αρχείου CSV που θα δημιουργηθεί
# και θα περιέχει τις νέες στήλες με το προπορευόμενο όχημα και το gap.
output_file = "vehicle_data_with_leading_gaps.csv"

# -----
# 2. Ανάγνωση του αρχείου CSV
# -----

# Το low_memory=False βοηθά στην ασφαλέστερη ανάγνωση αρχείων
# με πολλές στήλες ή μεικτούς τύπους δεδομένων.
df = pd.read_csv(input_file, low_memory=False)
```

```
# -----
# 3. Μετατροπή των βασικών στηλών σε αριθμητική μορφή
# -----

# Οι παρακάτω στήλες πρέπει να είναι αριθμητικές,
# διότι θα χρησιμοποιηθούν σε γεωμετρικούς και κινηματικούς υπολογισμούς.
numeric_cols = [
    "VehicleID",
    "Frame",
    "GroundPlaneCentroid_1",
    "GroundPlaneCentroid_2",
    "vx",
    "vy",
    "magnitude"
]

# Για κάθε στήλη της λίστας γίνεται μετατροπή σε numeric τύπο.
```

```
# -----
# 4. Αφαίρεση γραμμών με κρίσιμες ελλείπουσες τιμές
# -----

# Για να είναι δυνατός ο υπολογισμός της σχετικής θέσης μεταξύ οχημάτων,
# απαιτείται κάθε γραμμή να έχει:
# - VehicleID
# - Frame
# - x συντεταγμένη
# - y συντεταγμένη
#
# Όσες γραμμές δεν διαθέτουν αυτά τα βασικά στοιχεία αφαιρούνται.
df = df.dropna(subset=[
    "VehicleID",
    "Frame",
    "GroundPlaneCentroid_1",
    "GroundPlaneCentroid_2"
]).copy()

# -----
# 5. Ταξινόμηση των δεδομένων
# -----

# Η ταξινόμηση γίνεται πρώτα κατά VehicleID και μετά κατά Frame,
# ώστε τα διαδοχικά στιγμιότυπα του ίδιου οχήματος να είναι σε σωστή σειρά.
# Αυτό είναι χρήσιμο για τη συμπλήρωση της διεύθυνσης κίνησης στα επόμενα βήματα.
df = df.sort_values(["VehicleID", "Frame"]).reset_index(drop=True)
```

```

# -----
# 6. Υπολογισμός της στιγμιαίας ταχύτητας
# -----

# Η στιγμιαία ταχύτητα υπολογίζεται από τις συνιστώσες vx και vy:
# speed = sqrt(vx^2 + vy^2)
#
# Χρησιμοποιείται fillna(0), ώστε τυχόν κενές τιμές στις vx, vy
# να μη δημιουργήσουν σφάλμα στον υπολογισμό.
speed = np.sqrt(df["vx"].fillna(0)**2 + df["vy"].fillna(0)**2)

# Η υπολογισμένη ταχύτητα αποθηκεύεται προαιρετικά σε νέα στήλη,
# για λόγους ελέγχου ή μεταγενέστερης αξιοποίησης.
df["speed_calc"] = speed

# -----
# 7. Υπολογισμός μοναδιαίου διανύσματος κατεύθυνσης
# -----

# Για να εντοπιστεί ποιο όχημα βρίσκεται "μπροστά",
# δεν αρκεί μόνο η θέση του οχήματος. Χρειάζεται και η κατεύθυνση κίνησης.
#
# Η κατεύθυνση προκύπτει από το μοναδιαίο διάνυσμα:
# ux_raw = vx / speed
# uy_raw = vy / speed
#
# Ωστόσο, όταν η ταχύτητα είναι πολύ μικρή, η διεύθυνση δεν είναι αξιόπιστη.
# Για τον λόγο αυτό χρησιμοποιείται ένα κατώφλι ταχύτητας.
speed_threshold = 0.5 # m/s

# Αν η ταχύτητα είναι μεγαλύτερη από το κατώφλι,
# υπολογίζεται το μοναδιαίο διάνυσμα κατεύθυνσης.
# Διαφορετικά, αποδίδεται NaN.
df["ux_raw"] = np.where(speed > speed_threshold, df["vx"] / speed, np.nan)
df["uy_raw"] = np.where(speed > speed_threshold, df["vy"] / speed, np.nan)

```

```

# -----
# 8. Συμπλήρωση ελλειπουσών διευθύνσεων ανά όχημα
# -----

# Σε ορισμένα frames ένα όχημα μπορεί να εμφανίζεται σχεδόν ακινητοποιημένο,
# άρα η κατεύθυνσή του να μην είναι αξιόπιστη και να έχει τιμή NaN.
#
# Για να αντιμετωπιστεί αυτό, γίνεται συμπλήρωση των κενών τιμών
# με χρήση της προηγούμενης ή/και της επόμενης διαθέσιμης έγκυρης κατεύθυνσης
# του ίδιου οχήματος.
#
# ffill() = forward fill -> γεμίζει από προηγούμενη τιμή
# bfill() = backward fill -> γεμίζει από επόμενη τιμή
df["ux"] = df.groupby("VehicleID")["ux_raw"].transform(lambda s: s.ffill().bfill())
df["uy"] = df.groupby("VehicleID")["uy_raw"].transform(lambda s: s.ffill().bfill())

# Αν κάποιο όχημα δεν είχε ποτέ έγκυρη διεύθυνση σε κανένα frame,
# τότε οι τιμές που απομένουν ως NaN αντικαθίστανται με 0.
df["ux"] = df["ux"].fillna(0.0)
df["uy"] = df["uy"].fillna(0.0)

# -----
# 9. Παράμετροι για τον προσδιορισμό ίδιας λωρίδας / ίδιας ροής
# -----

# lane_thresh:
# Μέγιστη επιτρεπτή εγκάρσια απόσταση (σε μέτρα) μεταξύ δύο οχημάτων,
# ώστε να θεωρηθούν περίπου στην ίδια λωρίδα.
lane_thresh = 4.0

# cos_thresh:
# Ελάχιστη επιτρεπτή ομοιότητα διεύθυνσης.
# Η τιμή αυτή βασίζεται στο εσωτερικό γινόμενο δύο μοναδιαίων διανυσμάτων.
# Τιμή κοντά στο 1 σημαίνει πολύ παρόμοια κατεύθυνση.
cos_thresh = 0.92

```

```

# -----
# 10. Προετοιμασία πινάκων για τα αποτελέσματα
# -----

# Συνολικός αριθμός γραμμών του dataset.
n = len(df)

# Δημιουργία πινάκων για αποθήκευση:
# - του ID του προπορευόμενου οχήματος,
# - της διαμήκουσ απόστασης gap,
# - της εγκάρσιας απόστασης από το προπορευόμενο όχημα.
#
# Αρχικά όλοι οι πίνακες γεμίζουν με NaN.
leading_vehicle_id = np.full(n, np.nan)
gap_along_path = np.full(n, np.nan)
lateral_offset_to_leader = np.full(n, np.nan)

# -----
# 11. Ομαδοποίηση ανά frame
# -----

# Ο υπολογισμός του προπορευόμενου οχήματος γίνεται ανεξάρτητα για κάθε frame,
# δηλαδή συγκρίνονται μόνο οχήματα που υπάρχουν την ίδια χρονική στιγμή.
frame_indices = df.groupby("Frame").indices

```

```

# -----
# 12. Επεξεργασία κάθε frame ξεχωριστά
# -----

for frame, idx in frame_indices.items():

    # Μετατροπή των δεικτών των γραμμών σε numpy array,
    # ώστε να διευκολυνθούν οι υπολογισμοί.
    idx = np.array(idx)

    # Δημιουργία υποπίνακα μόνο για το συγκεκριμένο frame.
    sub = df.iloc[idx]

    # Ανάκτηση των συντεταγμένων x και y των οχημάτων του frame.
    x = sub["GroundPlaneCentroid_1"].to_numpy()
    y = sub["GroundPlaneCentroid_2"].to_numpy()

    # Ανάκτηση των τελικών μοναδιαίων διανυσμάτων κατεύθυνσης.
    ux = sub["ux"].to_numpy()
    uy = sub["uy"].to_numpy()

    # Ανάκτηση των VehicleID.
    vid = sub["VehicleID"].to_numpy()

    # Πλήθος οχημάτων στο συγκεκριμένο frame.
    m = len(sub)

    # -----
    # 12a. Υπολογισμός σχετικών θέσεων ανά ζεύγος οχημάτων
    # -----

    # Δημιουργούνται πίνακες διαφορών για όλα τα ζεύγη οχημάτων (i, j).
    #
    # dx[i, j] = x_j - x_i
    # dy[i, j] = y_j - y_i
    #
    # Δηλαδή περιγράφουν τη θέση του οχήματος j
    # ως προς το όχημα i.
    dx = x[None, :] - x[:, None]
    dy = y[None, :] - y[:, None]

    # -----
    # 12b. Υπολογισμός διαμήκου απόστασης
    # -----

    # Η διαμήκης απόσταση (longitudinal) προκύπτει
    # από την προβολή του διανύσματος θέσης (dx, dy)
    # πάνω στη διεύθυνση κίνησης του οχήματος i.
    #
    # longitudinal > 0 σημαίνει ότι το όχημα j
    # βρίσκεται μπροστά από το όχημα i.
    longitudinal = dx * ux[:, None] + dy * uy[:, None]

```

```

# -----
# 12c. Υπολογισμός εγκάρσιας απόστασης
# -----

# Η εγκάρσια απόσταση (lateral) εκφράζει πόσο "εκτός πορείας"
# βρίσκεται το όχημα j σε σχέση με την κατεύθυνση του i.
#
# Όσο μικρότερη είναι η lateral απόσταση,
# τόσο πιθανότερο είναι τα δύο οχήματα να βρίσκονται στην ίδια λωρίδα.
lateral = np.abs(dx * uy[:, None] - dy * ux[:, None])

# -----
# 12d. Υπολογισμός ομοιότητας κατεύθυνσης
# -----

# Η ομοιότητα κατεύθυνσης υπολογίζεται μέσω εσωτερικού γινομένου
# των μοναδιαίων διανυσμάτων των οχημάτων i και j.
#
# cos_sim κοντά στο 1 => πολύ παρόμοια κατεύθυνση
# cos_sim κοντά στο 0 => περίπου κάθετη κίνηση
# cos_sim αρνητικό => αντίθετη κατεύθυνση
cos_sim = ux[:, None] * ux[None, :] + uy[:, None] * uy[None, :]

# -----
# 12e. Επιλογή έγκυρων υποψηφίων leader
# -----

# Ένα όχημα j θεωρείται έγκυρος υποψήφιος leader για το i όταν:
# 1) βρίσκεται μπροστά (longitudinal > 0)
# 2) η εγκάρσια απόσταση είναι μικρή (lateral <= lane_thresh)
# 3) η κατεύθυνσή του είναι παρόμοια (cos_sim >= cos_thresh)
valid = (
    (longitudinal > 0) &
    (lateral <= lane_thresh) &
    (cos_sim >= cos_thresh)
)

# -----
# 12f. Επιλογή του κοντινότερου προπορευόμενου
# -----

# Για τα μη έγκυρα ζεύγη αποδίδεται άπειρη τιμή,
# ώστε να αγνοηθούν στην αναζήτηση του ελάχιστου.
masked_longitudinal = np.where(valid, longitudinal, np.inf)

# Για κάθε όχημα i εντοπίζεται ο δείκτης j του οχήματος
# με τη μικρότερη θετική διαμήκη απόσταση.
j_best = masked_longitudinal.argmin(axis=1)

# Η ελάχιστη αυτή απόσταση αποτελεί το gap προς τον leader.
best_dist = masked_longitudinal[np.arange(m), j_best]

# Η συνθήκη ok δείχνει σε ποια οχήματα βρέθηκε πράγματι leader.
ok = np.isfinite(best_dist)

# Αποθήκευση του vehicleID του leader
# στις αντίστοιχες θέσεις του αρχικού dataframe.
leading_vehicle_id[idx[ok]] = vid[j_best[ok]]

# Αποθήκευση της διαμήκου απόστασης gap.
gap_along_path[idx[ok]] = best_dist[ok]

# Αποθήκευση της εγκάρσιας απόστασης από τον leader.
lateral_offset_to_leader[idx[ok]] = lateral[np.arange(m)[ok], j_best[ok]]

```

```

# -----
# 13. Δημιουργία νέων στηλών στο dataframe
# -----

# Το ID του προπορευόμενου οχήματος αποθηκεύεται ως nullable integer,
# ώστε να επιτρέπονται και κενές τιμές.
df["LeadingVehicleID"] = pd.Series(leading_vehicle_id).astype("Int64")

# Η διαμήκης απόσταση gap σε μέτρα.
df["GapAlongPath_m"] = gap_along_path

# Η εγκάρσια απόσταση από το προπορευόμενο όχημα σε μέτρα.
df["LateralOffsetToLeader_m"] = lateral_offset_to_leader

# Λογική μεταβλητή που δείχνει αν βρέθηκε leader περίπου στην ίδια λωρίδα.
df["HasLeaderSameLane"] = ~pd.isna(df["GapAlongPath_m"])

# -----
# 14. Προαιρετική τελική ταξινόμηση
# -----

# Για λόγους αναγνωσιμότητας το τελικό αρχείο ταξινομείται
# πρώτα κατά Frame και έπειτα κατά VehicleID.
df_out = df.sort_values(["Frame", "VehicleID"]).copy()

# -----
# 15. Αποθήκευση σε νέο CSV
# -----

# Το τελικό dataframe αποθηκεύεται σε νέο αρχείο CSV
# χωρίς να γραφτεί η αρίθμηση γραμμών του pandas index.
df_out.to_csv(output_file, index=False)

# -----
# 16. Βασικά μηνύματα ελέγχου
# -----

# Εκτύπωση συνοπτικών πληροφοριών για έλεγχο της διαδικασίας.
print(f"Done. File saved as: {output_file}")
print(f"Rows: {len(df_out)}")
print(f"Frames: {df_out['Frame'].nunique()}")
print(f"Vehicles: {df_out['VehicleID'].nunique()}")
print(f"Rows with leader found: {df_out['HasLeaderSameLane'].mean() * 100:.2f}%")
print(f"Median gap: {df_out['GapAlongPath_m'].median():.2f} m")

```

Εικόνα 4.6. Script για δημιουργία *LeadingVehicleID*, *GapAlongPath_m*, *LateralOffsetToLeader_m*, *HasLeaderSameLane*

Η μεταβλητή **LeadingVehicleID** αποθηκεύει το αναγνωριστικό του προπορευόμενου οχήματος που θεωρήθηκε ως το καταλληλότερο σημείο αναφοράς για το εκάστοτε όχημα. Η μεταβλητή **GapAlongPath_m** εκφράζει τη διαμήκη απόσταση, σε μέτρα, κατά μήκος της πορείας κίνησης μεταξύ του οχήματος και του leader, δηλαδή το πραγματικό longitudinal gap που χρησιμοποιείται σε συγκοινωνιακές αναλύσεις ακολουθίας. Η μεταβλητή **LateralOffsetToLeader_m** αποτυπώνει την εγκάρσια απόσταση μεταξύ των δύο οχημάτων, ώστε να ελέγχεται αν η μεταξύ τους σχέση είναι

συμβατή με κοινή λωρίδα κίνησης. Τέλος, η δυαδική μεταβλητή **HasLeaderSameLane** λαμβάνει την τιμή True όταν έχει εντοπιστεί προπορευόμενο όχημα στην ίδια ροή/λωρίδα και την τιμή False όταν τέτοιο όχημα δεν βρέθηκε ή όταν η χωρική διάταξη δεν δικαιολογεί συσχέτιση leader–follower.

Σε επόμενο στάδιο δημιουργήθηκαν μεταβλητές που ποσοτικοποιούν την αλληλεπίδραση μεταξύ οχημάτων και τον δυνητικό κίνδυνο σύγκρουσης. Αρχικά υπολογίσθηκε η σχετική διαμήκης ταχύτητα ως προς το προπορευόμενο όχημα μέσω της μεταβλητής **LeaderLongitudinalRelativeSpeed_mps**. Με βάση τη διαμήκη απόσταση και τη σχετική ταχύτητα υπολογίσθηκε στη συνέχεια ο δείκτης Time to Collision (**TTC_Leader_s**), ο οποίος εκφράζει σε δευτερόλεπτα τον **θεωρητικό χρόνο** μέχρι **πιθανή σύγκρουση** με το προπορευόμενο όχημα, υπό την παραδοχή διατήρησης της τρέχουσας κινηματικής κατάστασης.

Παράλληλα, δημιουργήθηκαν και οι μεταβλητές που σχετίζονται με τον δείκτη **Post-Encroachment Time (PET)**. Για κάθε όχημα εξετάστηκαν οι **πιθανές αλληλεπιδράσεις** με τα υπόλοιπα οχήματα του ίδιου frame, θεωρώντας σταθερές ταχύτητες και κατευθύνσεις. Από το σύνολο των πιθανών αλληλεπιδράσεων επιλέχθηκε εκείνη που οδηγεί στη **μικρότερη προβλεπόμενη τιμή PET** και αποθηκεύτηκε στις **αντίστοιχες στήλες** του πίνακα.

```

import numpy as np
import pandas as pd

INPUT_CSV = "vehicle_data_kmh_moto_status_reversed.csv"
OUTPUT_CSV = "vehicle_data_with_TTC_PET.csv"

LANE_THRESHOLD_M = 4.0
COSINE_THRESHOLD = 0.92
MIN_SPEED_FOR_HEADING_MPS = 0.5
MAX_PET_HORIZON_S = 10.0
MIN_PET_SPEED_MPS = 0.5
POINT_MERGE_TOLERANCE_M = 2.0

df = pd.read_csv(INPUT_CSV, low_memory=False)

numeric_columns = [
    "VehicleID",
    "Frame",
    "GroundPlaneCentroid_1",
    "GroundPlaneCentroid_2",
    "vx",
    "vy",
    "magnitude",
]

for col in numeric_columns:
    df[col] = pd.to_numeric(df[col], errors="coerce")

df = df.dropna(
    subset=["VehicleID", "Frame", "GroundPlaneCentroid_1", "GroundPlaneCentroid_2"]
).copy()

df["_orig_order"] = np.arange(len(df))

speed = np.sqrt(df["vx"].fillna(0.0) ** 2 + df["vy"].fillna(0.0) ** 2)
df["speed_calc_mps"] = speed

df["ux_raw"] = np.where(
    speed > MIN_SPEED_FOR_HEADING_MPS,
    df["vx"] / speed,
    np.nan,
)

df["uy_raw"] = np.where(
    speed > MIN_SPEED_FOR_HEADING_MPS,
    df["vy"] / speed,
    np.nan,
)

df = df.sort_values(["VehicleID", "Frame"]).reset_index(drop=True)

df["ux"] = df.groupby("VehicleID")["ux_raw"].transform(lambda s: s.ffill().bfill())
df["uy"] = df.groupby("VehicleID")["uy_raw"].transform(lambda s: s.ffill().bfill())

df["ux"] = df["ux"].fillna(0.0)
df["uy"] = df["uy"].fillna(0.0)

n = len(df)

leading_vehicle_id = np.full(n, np.nan)
gap_along_path_m = np.full(n, np.nan)
lateral_offset_to_leader_m = np.full(n, np.nan)

frame_to_indices = df.groupby("Frame").indices

```

```

for _, idx in frame_to_indices.items():
    idx = np.array(idx)
    sub = df.iloc[idx]

    x = sub["GroundPlaneCentroid_1"].to_numpy()
    y = sub["GroundPlaneCentroid_2"].to_numpy()
    ux = sub["ux"].to_numpy()
    uy = sub["uy"].to_numpy()
    vid = sub["VehicleID"].to_numpy()

    dx = x[None, :] - x[:, None]
    dy = y[None, :] - y[:, None]

    longitudinal = dx * ux[:, None] + dy * uy[:, None]
    lateral = np.abs(dx * uy[:, None] - dy * ux[:, None])
    cosine_similarity = ux[:, None] * ux[None, :] + uy[:, None] * uy[None, :]

    valid_pairs = (
        (longitudinal > 0.0)
        & (lateral <= LANE_THRESHOLD_M)
        & (cosine_similarity >= COSINE_THRESHOLD)
    )

    candidate_gaps = np.where(valid_pairs, longitudinal, np.inf)

    best_j = candidate_gaps.argmax(axis=1)
    best_gap = candidate_gaps[np.arange(len(idx)), best_j]

    found_leader = np.isfinite(best_gap)

    leading_vehicle_id[idx[found_leader]] = vid[best_j[found_leader]]
    gap_along_path_m[idx[found_leader]] = best_gap[found_leader]
    lateral_offset_to_leader_m[idx[found_leader]] = lateral[
        np.arange(len(idx))[found_leader],
        best_j[found_leader],
    ]

df["LeadingVehicleID"] = pd.Series(leading_vehicle_id).astype("Int64")
df["GapAlongPath_m"] = gap_along_path_m
df["LateralOffsetToLeader_m"] = lateral_offset_to_leader_m
df["HasLeaderSameLane"] = ~pd.isna(df["GapAlongPath_m"])

leader_relative_speed_mps = np.full(n, np.nan)

lookup = df.set_index(["Frame", "VehicleID"])

for i, row in df.iterrows():
    leader_id = row["LeadingVehicleID"]

    if pd.isna(leader_id):
        continue

    key = (row["Frame"], int(leader_id))

    if key not in lookup.index:
        continue

    leader_row = lookup.loc[key]

    if isinstance(leader_row, pd.DataFrame):
        leader_row = leader_row.iloc[0]

    dvx = row["vx"] - leader_row["vx"]
    dvy = row["vy"] - leader_row["vy"]

    leader_relative_speed_mps[i] = dvx * row["ux"] + dvy * row["uy"]

```

```

df["LeaderLongitudinalRelativeSpeed_mps"] = leader_relative_speed_mps

ttc_leader_s = np.full(n, np.nan)

valid_ttc = (
    df["GapAlongPath_m"].notna()
    & df["LeaderLongitudinalRelativeSpeed_mps"].notna()
    & (df["LeaderLongitudinalRelativeSpeed_mps"] > 0.0)
)

ttc_leader_s[valid_ttc.to_numpy()] = (
    df.loc[valid_ttc, "GapAlongPath_m"].to_numpy()
    / df.loc[valid_ttc, "LeaderLongitudinalRelativeSpeed_mps"].to_numpy()
)

df["TTC_Leader_s"] = ttc_leader_s

def estimate_pet_for_pair(x1, y1, vx1, vy1, x2, y2, vx2, vy2):
    """
    Εκτιμά το PET για ένα ζεύγος οχημάτων υπό σταθερή ταχύτητα / σταθερή κατεύθυνση.

    Θεωρούνται οι προβλεπόμενες τροχιές:
    P1(t) = (x1, y1) + t * (vx1, vy1)
    P2(s) = (x2, y2) + s * (vx2, vy2)
    """

    speed1 = np.hypot(vx1, vy1)
    speed2 = np.hypot(vx2, vy2)

    if speed1 < MIN_PET_SPEED_MPS or speed2 < MIN_PET_SPEED_MPS:
        return (np.nan, np.nan, np.nan, np.nan, np.nan)

    A = np.array([[vx1, -vx2], [vy1, -vy2]], dtype=float)
    b = np.array([x2 - x1, y2 - y1], dtype=float)

    det = np.linalg.det(A)
    if abs(det) < 1e-9:
        return (np.nan, np.nan, np.nan, np.nan, np.nan)

    t1, t2 = np.linalg.solve(A, b)

    if t1 < 0.0 or t2 < 0.0:
        return (np.nan, np.nan, np.nan, np.nan, np.nan)

    if t1 > MAX_PET_HORIZON_S or t2 > MAX_PET_HORIZON_S:
        return (np.nan, np.nan, np.nan, np.nan, np.nan)

    cx1 = x1 + vx1 * t1
    cy1 = y1 + vy1 * t1
    cx2 = x2 + vx2 * t2
    cy2 = y2 + vy2 * t2

    if np.hypot(cx1 - cx2, cy1 - cy2) > POINT_MERGE_TOLERANCE_M:
        return (np.nan, np.nan, np.nan, np.nan, np.nan)

    pet_value = abs(t1 - t2)

    return (pet_value, cx1, cy1, t1, t2)

```

```

pet_min_predicted_s = np.full(n, np.nan)
pet_conflict_vehicle_id = np.full(n, np.nan)
pet_conflict_x_m = np.full(n, np.nan)
pet_conflict_y_m = np.full(n, np.nan)
pet_time_to_conflict_self_s = np.full(n, np.nan)
pet_time_to_conflict_other_s = np.full(n, np.nan)

for _, idx in frame_to_indices.items():
    idx = np.array(idx)
    sub = df.iloc[idx]

    x = sub["GroundPlaneCentroid_1"].to_numpy()
    y = sub["GroundPlaneCentroid_2"].to_numpy()
    vx = sub["vx"].fillna(0.0).to_numpy()
    vy = sub["vy"].fillna(0.0).to_numpy()
    vid = sub["VehicleID"].to_numpy()

    m = len(idx)

    for a in range(m):
        best_pet = np.nan
        best_vid = np.nan
        best_cx = np.nan
        best_cy = np.nan
        best_t_self = np.nan
        best_t_other = np.nan

        for b in range(m):
            if a == b:
                continue

            pet_value, cx, cy, t_self, t_other = estimate_pet_for_pair(
                x[a], y[a], vx[a], vy[a],
                x[b], y[b], vx[b], vy[b],
            )

            if np.isnan(pet_value):
                continue

            if np.isnan(best_pet) or pet_value < best_pet:
                best_pet = pet_value
                best_vid = vid[b]
                best_cx = cx
                best_cy = cy
                best_t_self = t_self
                best_t_other = t_other

        pet_min_predicted_s[idx[a]] = best_pet
        pet_conflict_vehicle_id[idx[a]] = best_vid
        pet_conflict_x_m[idx[a]] = best_cx
        pet_conflict_y_m[idx[a]] = best_cy
        pet_time_to_conflict_self_s[idx[a]] = best_t_self
        pet_time_to_conflict_other_s[idx[a]] = best_t_other

df["PET_MinPredicted_s"] = pet_min_predicted_s
df["PET_ConflictVehicleID"] = pd.Series(pet_conflict_vehicle_id).astype("Int64")
df["PET_ConflictPoint_X_m"] = pet_conflict_x_m
df["PET_ConflictPoint_Y_m"] = pet_conflict_y_m
df["PET_TimeToConflict_Self_s"] = pet_time_to_conflict_self_s
df["PET_TimeToConflict_Other_s"] = pet_time_to_conflict_other_s

df = df.sort_values("orig_order").drop(columns=["orig_order"])

df.to_csv(OUTPUT_CSV, index=False)

print(f"H επεξεργασία ολοκληρώθηκε. Το αρχείο αποθηκεύτηκε ως: {OUTPUT_CSV}")

```

Εικόνα 4.7. Script για δημιουργία *TTC_Leader_s*, *PET_MinPredicted_s*, *PET_ConflictVehicleID*, *PET_ConflictPoint_X_m*, *PET_ConflictPoint_Y_m*, *PET_TimeToConflict_Self_s*, *PET_TimeToConflict_Other_s*,

Ειδικότερα, η μεταβλητή **PET_MinPredicted_s** αντιστοιχεί στη μικρότερη προβλεπόμενη τιμή Post-Encroachment Time, δηλαδή στη μικρότερη χρονική διαφορά με την οποία δύο οχήματα αναμένεται να διέλθουν από το ίδιο conflict point. Η μεταβλητή **PET_ConflictVehicleID** καταγράφει το vehicle ID του οχήματος με το οποίο προκύπτει αυτή η κρίσιμη αλληλεπίδραση. Οι μεταβλητές **PET_ConflictPoint_X_m** και **PET_ConflictPoint_Y_m** αποθηκεύουν τις συντεταγμένες του προβλεπόμενου σημείου σύγκρουσης στο σύστημα αναφοράς του ground plane. Τέλος, οι μεταβλητές **PET_TimeToConflict_Self_s** και **PET_TimeToConflict_Other_s** εκφράζουν, αντίστοιχα, τον χρόνο που απαιτείται για να φθάσει στο conflict point το εξεταζόμενο όχημα και το άλλο όχημα που συμμετέχει στην κρίσιμη αλληλεπίδραση.

Το τελικό αρχείο δεδομένων έχει αυτήν την μορφή:

The image shows a detailed view of a data table with columns labeled A through Z. The data includes various numerical values, some with scientific notation, and some cells containing text like 'green', 'red', 'yellow', 'orange', 'blue', 'purple', 'brown', 'pink', 'grey', 'white', 'black', 'red', 'green', 'blue', 'yellow', 'orange', 'purple', 'brown', 'pink', 'grey', 'white', 'black'. The table appears to be a continuation of the data from the previous image.

Εικόνα 4.8. Μέρος του Master Table

4.5 Αναλυτική περιγραφή των μεταβλητών του τελικού πίνακα

Για λόγους πληρότητας, στον Πίνακα που ακολουθεί συνοψίζεται η σημασία των **βασικών στηλών** του **τελικού αρχείου δεδομένων**. Οι αρχικές μεταβλητές διατηρήθηκαν χωρίς αλλοίωση των καταγεγραμμένων τιμών, ενώ οι **νέες στήλες** προστέθηκαν αποκλειστικά για τις ανάγκες της επεξεργασίας και της επακόλουθης **στατιστικής ανάλυσης**.

Μεταβλητή	Περιγραφή
VehicleID	Μοναδικό αναγνωριστικό κάθε οχήματος στο σύνολο δεδομένων.
Frame	Αριθμός καρέ του βίντεο στο οποίο καταγράφεται η συγκεκριμένη παρατήρηση.
Confidence	Βαθμός εμπιστοσύνης της ανίχνευσης/ταυτοποίησης του αντικειμένου από τον αλγόριθμο.

GroundPlaneCentroid_1	Συντεταγμένη x του κέντρου βάρους του οχήματος στο επίπεδο αναφοράς του οδοστρώματος.
GroundPlaneCentroid_2	Συντεταγμένη y του κέντρου βάρους του οχήματος στο επίπεδο αναφοράς του οδοστρώματος.
IsInCrossingArea	Διαδική ένδειξη για το αν το όχημα βρίσκεται εντός της καθορισμένης περιοχής ενδιαφέροντος/διάβασης.
TrafficLightStatus	Ένδειξη της φωτεινής σηματοδότησης κατά τη χρονική στιγμή της παρατήρησης.
ProcessedTrafficLightStatus	Διορθωμένη ένδειξη της φωτεινής σηματοδότησης κατά τη χρονική στιγμή της παρατήρησης.
VehicleStatus	Χαρακτηρισμός της διέλευσης ως legal ή illegal, ανάλογα με τη συμμόρφωση προς τη σηματοδότηση.
vx	Συνιστώσα ταχύτητας κατά τον άξονα x, σε m/s.
vy	Συνιστώσα ταχύτητας κατά τον άξονα y, σε m/s.
magnitude	Μέτρο της ταχύτητας του οχήματος. Στο τελικό αρχείο αποδίδεται σε km/h.
Speeding_during_Orange_50_100	Διαδική/δεικτική μεταβλητή που καταγράφει εμφάνιση υπέρβασης ταχύτητας κατά τη διάρκεια πορτοκαλί ένδειξης.
Speeding_during_Green_50_100	Διαδική/δεικτική μεταβλητή που καταγράφει εμφάνιση υπέρβασης ταχύτητας κατά τη διάρκεια πράσινης ένδειξης.
Is in cr Area	Ένδειξη που χρησιμοποιείται για να αποτυπώσει υπέρβαση ταχύτητας κατά τη διάρκεια πράσινης ένδειξης ενώ το όχημα βρίσκεται στην περιοχή της διάβασης πεζών.
Vehicle_Classification	Κατηγοριοποίηση του οχήματος ως car ή moto.

LeadingVehicleID	Αναγνωριστικό του προπορευόμενου οχήματος που αντιστοιχίσθηκε στο εξεταζόμενο όχημα.
GapAlongPath_m	Διαμήκης απόσταση, σε μέτρα, μεταξύ οχήματος και προπορευόμενου κατά μήκος της πορείας.
LateralOffsetToLeader_m	Εγκάρσια απόσταση, σε μέτρα, από το προπορευόμενο όχημα.
HasLeaderSameLane	Λογική μεταβλητή True/False που δηλώνει αν βρέθηκε προπορευόμενο όχημα στην ίδια λωρίδα ή ροή.
LeaderLongitudinalRelativeSpeed_mps	Σχετική διαμήκης ταχύτητα μεταξύ οχήματος και προπορευόμενου, σε m/s.
TTC_Leader_s	Χρόνος μέχρι πιθανή σύγκρουση με το προπορευόμενο όχημα, σε δευτερόλεπτα.
PET_MinPredicted_s	Ελάχιστη προβλεπόμενη τιμή Post-Encroachment Time, σε δευτερόλεπτα.
PET_ConflictVehicleID	Αναγνωριστικό του οχήματος με το οποίο προκύπτει το ελάχιστο Post-Encroachment Time (PET).
PET_ConflictPoint_X_m	Συντεταγμένη x του προβλεπόμενου conflict point, σε μέτρα.
PET_ConflictPoint_Y_m	Συντεταγμένη y του προβλεπόμενου conflict point, σε μέτρα.
PET_TimeToConflict_Self_s	Χρόνος του εξεταζόμενου οχήματος μέχρι το conflict point, σε δευτερόλεπτα.
PET_TimeToConflict_Other_s	Χρόνος του άλλου οχήματος μέχρι το ίδιο conflict point, σε δευτερόλεπτα.

Συνολικά, η παραπάνω διαδικασία οδήγησε στη δημιουργία ενός **συνεκτικού και τεκμηριωμένου αρχείου δεδομένων**, στο οποίο έχουν ενσωματωθεί τόσο οι

πρωτογενείς πληροφορίες της ανίχνευσης όσο και οι **παράγωγες μεταβλητές** που απαιτούνται για την ανάλυση της συμπεριφοράς, της αλληλεπίδρασης και της **επικινδυνότητας των οχημάτων** στη σηματοδοτούμενη διασταύρωση.

4.6 Στατιστική αποτύπωση δεδομένων - Συγκεντρωτικά Στοιχεία

Στο πλαίσιο της παρούσας εργασίας πραγματοποιήθηκε **στατιστική αποτύπωση**

των μεταβλητών που περιλαμβάνονται στις τρεις επιμέρους βάσεις δεδομένων:

πεζών, οχημάτων και δεδομένων ελάχιστου χρονικού περιθωρίου πριν από πιθανή σύγκρουση. Για τις **αριθμητικές μεταβλητές** υπολογίστηκαν βασικά μεγέθη, όπως η **μέση τιμή** (ενδεικτικό του κεντρικού επιπέδου των τιμών), η **τυπική απόκλιση** (ένδειξη της διασποράς τους), τα **ποσοστημόρια** (για την κατανοήση της κατανομής)

και οι ακραίες τιμές (ελάχιστο και μέγιστο). Αντίστοιχα, για τις **κατηγορικές και**

δυσιαδικές μεταβλητές υπολογίστηκαν οι απόλυτες και σχετικές συχνότητες

εμφάνισης των κατηγοριών τους, περιλαμβανομένων και των περιπτώσεων

αβεβαιότητας (π.χ. **τιμές unknown**), ώστε να αξιολογηθεί η **πληρότητα των δεδομένων**.

Η παρουσίαση των αποτελεσμάτων οργανώθηκε σε **πίνακες και διαγράμματα**, με στόχο την παροχή μιας καθαρής εικόνας της **δομής**, της **ποιότητας** και της ποικιλίας των **διαθέσιμων δεδομένων** πριν την εφαρμογή περαιτέρω αναλυτικών μεθόδων.

Στατιστικά αριθμητικών μεταβλητών

Μεταβλητή	Πλήθος (count)	Μέση Τιμή (mean)	Τυπ. Απόκλιση (std)	Ελάχιστο (min)	25%	Διάμεσος 50%	75%	Μέγιστο (max)
Confidence	429,126	0.77	0.131	0.33	0.678	0.791	0.878	0.977
GroundPlaneCentroid_1	449,533	145.395	325.619	-9,248.30	85.42	112.86	146.119	39,029.80
GroundPlaneCentroid_2	449,533	79.527	2,921.09	-8,110.80	37.76	56.73	82.063	505,271.90
vx	449,533	0.053	3.275	-39.13	-0.844	-0.01	0.129	35.722
vy	449,533	-2.01	4.937	-39.401	-4.614	-2.259	0.556	39.892
magnitude	449,533	26.624	18.482	0	9.79	27.064	44.137	143.998
speed_calc	449,533	4.966	4.675	0	1.405	4.644	6.73	39.999
GapAlongPath_m	239,592	31.408	102.17	0	6.165	13.847	28.576	8,815.48
LateralOffsetToLeader_m	239,592	0.973	0.915	0	0.317	0.772	1.46	4
TTC_Leader_s	112,951	152.397	5,615.88	0	3.225	5.496	11.536	803,277.80
PET_MinPredicted_s	29,103	2.974	2.303	0	1.001	2.525	4.494	9.989
PET_TimeToConflict_Self_s	29,103	3.128	2.376	0	1.036	2.685	4.958	9.999
PET_TimeToConflict_Other_s	29,103	2.857	2.167	0	0.984	2.418	4.52	9.999

Η **περιγραφική στατιστική** των αριθμητικών μεταβλητών δείχνει αρχικά ότι το σύνολο των παρατηρήσεων είναι **πολύ μεγάλο**, γεγονός που ενισχύει τη **στατιστική αξιοπιστία** της ανάλυσης. Παράλληλα, το πλήθος διαφέρει μεταξύ των μεταβλητών, κάτι που σημαίνει ότι ορισμένες στήλες, όπως οι δείκτες αλληλεπίδρασης **TTC και PET**,

δεν είναι διαθέσιμες για όλα τα οχήματα, αλλά μόνο για εκείνες τις περιπτώσεις όπου μπορούσε να υπολογιστεί σχετική σύγκρουση ή αλληλεπίδραση με άλλο όχημα.

Η μεταβλητή **Confidence** εμφανίζει **μέση τιμή 0.770** και σχετικά μικρή τυπική απόκλιση. Αυτό υποδηλώνει ότι, κατά μέσο όρο, η **αξιοπιστία της ανίχνευσης** των οχημάτων από το σύστημα είναι ικανοποιητική και δεν παρουσιάζει ακραία διασπορά. Συνεπώς, τα δεδομένα παρακολούθησης μπορούν να θεωρηθούν γενικά **αξιόπιστα** ως βάση για περαιτέρω ανάλυση.

Οι μεταβλητές **GroundPlaneCentroid_1** και **GroundPlaneCentroid_2** παρουσιάζουν **πολύ μεγάλες διακυμάνσεις** και ακραίες μέγιστες τιμές. Αυτό δείχνει ότι οι συντεταγμένες των οχημάτων καλύπτουν μεγάλο εύρος στο επίπεδο αναφοράς, πιθανόν λόγω διαφορετικών θέσεων μέσα στο χώρο, αλλά και πιθανών ακραίων παρατηρήσεων ή σφαλμάτων μετασχηματισμού/ανίχνευσης. Επομένως, οι μεταβλητές θέσης είναι χρήσιμες για χωρική ανάλυση, αλλά απαιτούν προσοχή ως προς τα **outliers**.

Για τις συνιστώσες ταχύτητας **v_x** και **v_y** , η μέση τιμή του v_x είναι πολύ κοντά στο μηδέν, ενώ η v_y είναι αρνητική κατά μέσο όρο. Αυτό μπορεί να σημαίνει ότι η **κύρια κατεύθυνση κίνησης** των οχημάτων στη συγκεκριμένη θέση μελέτης αναπτύσσεται κυρίως κατά τον άξονα y και μάλιστα προς **αρνητική φορά** του συστήματος συντεταγμένων. Με άλλα λόγια, η γεωμετρία της κάμερας και του μετασχηματισμού φαίνεται να έχει αποτυπώσει μια **κυρίαρχη κατεύθυνση ροής**.

Η μεταβλητή **magnitude** εμφανίζει **μέση τιμή περίπου 26.6**, με αρκετά μεγάλη τυπική απόκλιση. Αυτό δείχνει **σημαντική ετερογένεια** στις ταχύτητες των οχημάτων, δηλαδή συνυπάρχουν οχήματα με χαμηλές αλλά και πολύ υψηλότερες ταχύτητες. Η ύπαρξη μέγιστης τιμής σχεδόν 144 υποδηλώνει είτε πολύ γρήγορες κινήσεις είτε πιθανές **ακραίες τιμές** που πρέπει να ελεγχθούν πριν από μοντελοποίηση, ειδικά αν η μονάδα είναι km/h. Αντίστοιχα, η **speed_calc** έχει χαμηλότερη μέση τιμή, γεγονός που δείχνει ότι πιθανόν πρόκειται για διαφορετικό τρόπο υπολογισμού της ταχύτητας ή για πιο «συντηρητική» εκτίμηση.

Η μεταβλητή **GapAlongPath_m** παρουσιάζει **μεγάλη διασπορά**. Η διάμεσος είναι αρκετά χαμηλότερη από τη μέση τιμή, κάτι που δείχνει **δεξιά ασυμμετρία**: οι περισσότερες αποστάσεις μεταξύ διαδοχικών οχημάτων είναι σχετικά μικρές, αλλά υπάρχουν λίγες πολύ μεγάλες τιμές που ανεβάζουν τον μέσο όρο. Αυτό είναι αναμενόμενο σε κυκλοφοριακά δεδομένα, επειδή σε κάποιες χρονικές στιγμές τα οχήματα κινούνται κοντά μεταξύ τους, ενώ σε άλλες υπάρχουν **μεγάλα κενά στη ροή**.

Η **LateralOffsetToLeader_m** έχει μικρή σχετικά μέση τιμή, γεγονός που υποδηλώνει ότι στις περισσότερες περιπτώσεις το προπορευόμενο όχημα βρίσκεται κοντά στην **ίδια λωρίδα** ή σε παρόμοια εγκάρσια θέση με το εξεταζόμενο όχημα. Αυτό ενισχύει την ορθότητα του υπολογισμού ηγούμενου οχήματος, ειδικά αν η μεταβλητή αυτή χρησιμοποιείται για να επιβεβαιώνει σχέση **car-following**.

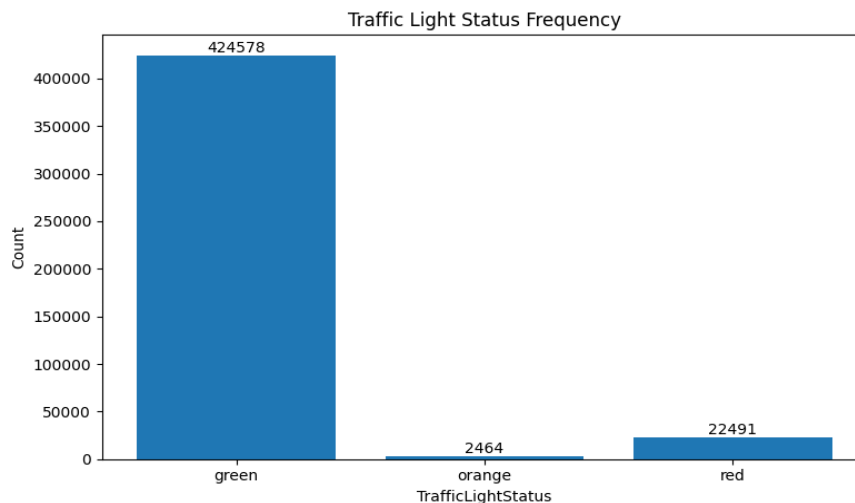
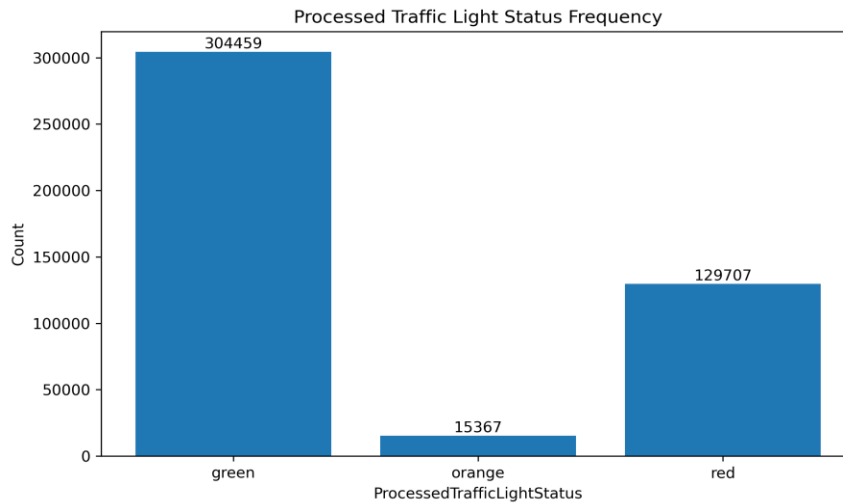
Η μεταβλητή **TTC_Leader_s** εμφανίζει πολύ μεγάλη μέση τιμή και εξαιρετικά μεγάλη τυπική απόκλιση, ενώ η διάμεσος είναι πολύ μικρή συγκριτικά. Αυτό σημαίνει ότι η κατανομή είναι **έντονα ασύμμετρη** και περιλαμβάνει πολλές **ακραίες τιμές**. Στην πράξη, αυτό δείχνει ότι στις περισσότερες περιπτώσεις ο κίνδυνος άμεσης σύγκρουσης με το προπορευόμενο όχημα είναι μάλλον μικρός, αλλά υπάρχουν λίγες περιπτώσεις με πολύ μεγάλες τιμές TTC που διογκώνουν τον μέσο όρο. Για τον λόγο αυτό, στη σχετική ερμηνεία είναι πιο κατάλληλο να δίνεται έμφαση στη **διάμεσο και στα τεταρτημόρια** παρά μόνο στη μέση τιμή.

Οι μεταβλητές **PET_MinPredicted_s**, **PET_TimeToConflict_Self_s** και **PET_TimeToConflict_Other_s** έχουν μέσες τιμές περίπου από 2.9 έως 3.1 s. Αυτές οι τιμές υποδηλώνουν ότι, στις περιπτώσεις όπου ανιχνεύθηκε πιθανή αλληλεπίδραση, οι χρονικές αποστάσεις μεταξύ οχημάτων δεν είναι αμελητέες αλλά ούτε και πολύ υψηλές. Δηλαδή, υπάρχουν καταστάσεις με **δυναμική σύγκρουση** ή εγγύτητα, χωρίς όμως να σημαίνει απαραίτητα ότι πρόκειται πάντα για κρίσιμα συμβάντα. Εφόσον στη διεθνή βιβλιογραφία μικρές τιμές TTC και PET συνδέονται συνήθως με **αυξημένη επικινδυνότητα**, οι χαμηλές τιμές των κατώτερων τεταρτημορίων δείχνουν ότι μέσα στο δείγμα υπάρχουν και αρκετές περιπτώσεις που αξίζουν ειδικότερη διερεύνηση.

Σε πιο ουσιαστικό επίπεδο, τα αποτελέσματα φαίνεται να δείχνουν ότι το δείγμα περιλαμβάνει **μεγάλη ποικιλία κυκλοφοριακών συνθηκών**: από **ελεύθερη ροή** και μεγάλες αποστάσεις μέχρι περιπτώσεις πιο **στενής αλληλεπίδρασης**, με μέτριες ή και αυξημένες ταχύτητες. Αυτό είναι ιδιαίτερα χρήσιμο για τη συνέχεια της εργασίας, επειδή σημαίνει ότι το dataset είναι κατάλληλο για διερεύνηση σχέσεων μεταξύ ταχύτητας, σηματοδότησης και **δεικτών επικινδυνότητας**.

4.6.1 Ένδειξη φωτεινού σηματοδότη

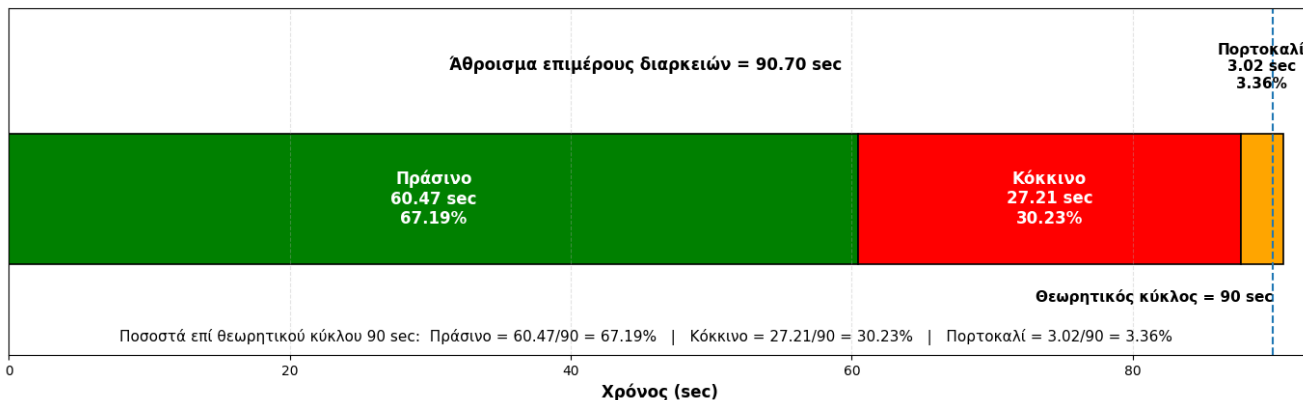
Μεταβλητή	Κατάσταση Φωτεινού Σηματοδότη		
	Green	Orange	Red
Πλήθος	304,459	15,367	129,707
Ποσοστό	67.73%	3.42%	28.85%



Για τη διασφάλιση της **αξιοπιστίας και της συνέπειας** των δεδομένων, εφαρμόστηκε διαδικασία καθαρισμού και εξομάλυνσης των τιμών στη στήλη **Traffic Light Status**. Με τις απαραίτητες διορθώσεις σε τιμές που ήταν κενές ή όπου εμφανώς προέκυψαν λάθη, καθορίστηκαν με ακρίβεια οι **κύκλοι λειτουργίας** του φωτεινού σηματοδότη και έγινε αντιστοίχιση κάθε χρονικής στιγμής, κάθε καρέ με την σωστή ένδειξη του σηματοδότη. Η ανάλυση της ακρίβειας των ενδείξεων του φωτεινού σηματοδότη βασίστηκε στη σύγκριση μεταξύ της αυτοματοποιημένης και της χειροκίνητης μεθόδου. Για την

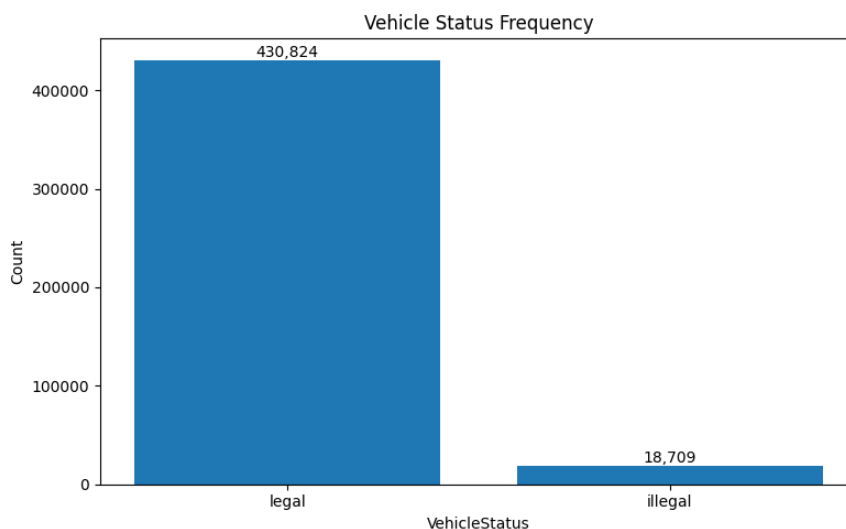
εξασφάλιση της εγκυρότητας των αποτελεσμάτων, κρίθηκε απαραίτητη η προεπεξεργασία των δεδομένων της αυτοματοποιημένης μεθόδου, ώστε να εναρμονιστούν οι διαφορετικές μορφές καταγραφής και να διευκολυνθεί η σύγκριση. Για την ανάλυση των δεδομένων, ήταν απαραίτητο να καθοριστούν οι κύκλοι λειτουργίας του φωτεινού σηματοδότη και να αντιστοιχιστεί η χρονική στιγμή κάθε εγγραφής με τον σωστό κύκλο και τη φάση στην οποία ανήκει. Κατά την ανάλυση των “raw” δεδομένων έγινε διασταύρωση των τιμών Traffic Light Status και έπειτα από διορθώσεις σφαλμάτων ή ελλείψεων διαπιστώθηκε πως κατά μέσο όρο παρεμβάλλονται μεταξύ πράσινης, πορτοκαλί και κόκκινης ένδειξης του φωτεινού σηματοδότη: 1798.47 frames, 809.27 frames & 89.83 frames. Αυτό μεταφράζεται στις εξής χρονικές διάρκειες των ενδείξεων του φωτεινού σηματοδότη: **60.47 sec διάρκεια Πρασίνου, 27.21 sec διάρκεια Κόκκινου και 3.02 sec διάρκεια Πορτοκαλί**. Έτσι, γίνεται η σύνθεση των κύκλων λειτουργίας του φωτεινού σηματοδότη, με περίοδο (προκύπτει και αθροιστικά – περίπου) **90 seconds**.

Διάρκεια ενδείξεων φωτεινού σηματοδότη και περίοδος κύκλου



4.6.2 Τύπος Διέλευσης Οχημάτων

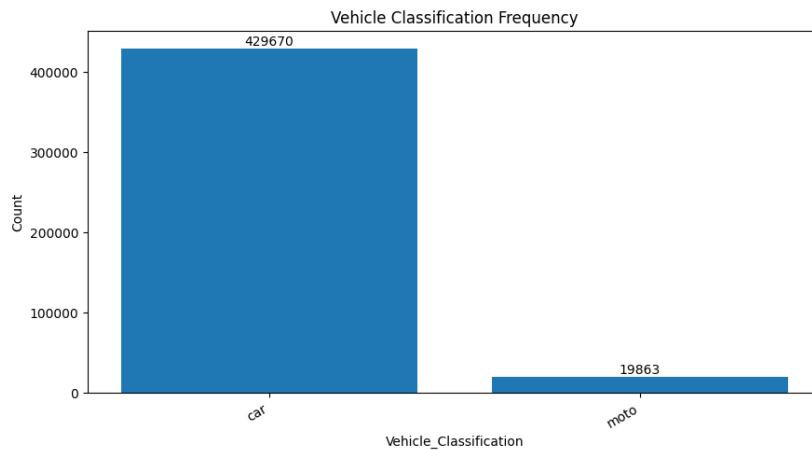
Μεταβλητή	Τύπος Διέλευσης	
	Legal	Illegal
Πλήθος	430,824	18,709
Ποσοστό	95.84%	4.16%



Η κατανομή της μεταβλητής **VehicleStatus** καταδεικνύει ότι το μεγαλύτερο μέρος των παρατηρήσεων αντιστοιχεί σε **νόμιμη συμπεριφορά (legal)**, με ποσοστό **95,84%**, ενώ οι **παράνομες ή παραβατικές περιπτώσεις (illegal)** ανέρχονται μόλις στο 4,16% του συνολικού δείγματος. Το εύρημα αυτό υποδηλώνει ότι οι παραβάσεις αποτελούν μειοψηφικό αλλά ιδιαίτερα κρίσιμο τμήμα του δείγματος, καθώς συνδέονται με τις πλέον επικίνδυνες εκδηλώσεις οδηγικής συμπεριφοράς. Παράλληλα, η έντονη ανισοκατανομή μεταξύ των δύο κατηγοριών θα πρέπει να ληφθεί υπόψη κατά την εφαρμογή στατιστικών μοντέλων, ώστε να αποφευχθούν στρεβλώσεις στην ερμηνεία των αποτελεσμάτων.

4.6.3 Κατηγορία Οχήματος

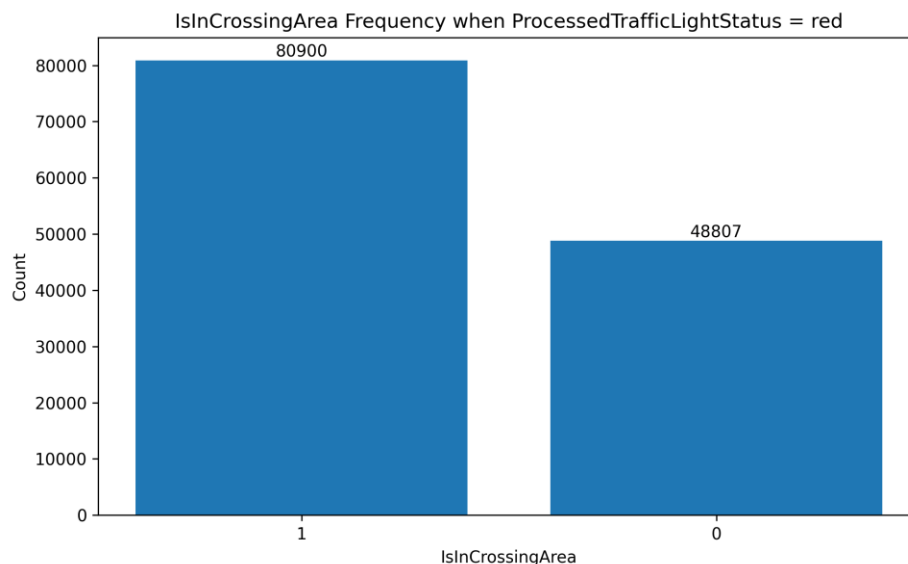
Μεταβλητή	Κατηγορία Οχήματος	
	Car	Moto
Πλήθος	429,670	19,863
Ποσοστό	95.58%	4.42%



Η κατανομή της μεταβλητής **Vehicle_Classification** δείχνει ότι το δείγμα αποτελείται κυρίως από **επιβατικά οχήματα**, ενώ οι **μοτοσυκλές** αντιπροσωπεύουν μικρότερο ποσοστό του συνόλου. Η παρατήρηση αυτή είναι ιδιαίτερα σημαντική, καθώς υποδηλώνει ότι τα συνολικά αποτελέσματα της ανάλυσης αντανακλούν κυρίως τη συμπεριφορά των επιβατικών οχημάτων. Επιπλέον, η περιορισμένη συμμετοχή των μοτοσυκλετών στο δείγμα απαιτεί προσοχή κατά την ερμηνεία τυχόν **συγκρίσεων μεταξύ κατηγοριών οχημάτων**.

4.6.4 Παρουσία Οχήματος στην Διάβαση Πεζών ενώ το φανάρι είναι κόκκινο

Μεταβλητή	Όχημα να βρίσκεται στην περιοχή Διάβασης Πεζών ενώ ο σηματοδότης δείχνει κόκκινο	
	ΝΑΙ (1)	ΌΧΙ (0)
Πλήθος	80,900	48,807
Ποσοστό	62.37%	37.63%



Παρατηρείται ότι για τις περιπτώσεις όπου ο φωτεινός σηματοδότης λαμβάνει την **ένδειξη red**, το **62,37%** των καταγραφών αντιστοιχεί σε οχήματα που βρίσκονται **εντός της περιοχής διέλευσης** (IsInCrossingArea = 1). Το εύρημα αυτό υποδηλώνει ότι σημαντικό μέρος της κυκλοφορίας εξακολουθεί να καταλαμβάνει τη ζώνη της διασταύρωσης κατά την κόκκινη ένδειξη, γεγονός που μπορεί να συνδέεται είτε με οχήματα που εισήλθαν οριακά πριν τη λήξη της επιτρεπόμενης φάσης είτε με καθυστερημένη εκκένωση της διασταύρωσης. Ως εκ τούτου, το συγκεκριμένο ποσοστό αποτελεί ένδειξη αυξημένης έκθεσης σε **δυναμικά συγκρουσιακές συνθήκες** και παρουσιάζει ιδιαίτερο ενδιαφέρον για την ανάλυση της οδικής ασφάλειας.

Κεφάλαιο 5: Στατιστική Ανάλυση

5.1 Εισαγωγή

Σε συνέχεια της συλλογής και επεξεργασίας των στοιχείων που παρουσιάστηκε στο Κεφάλαιο 4, ακολούθησε η στατιστική τους ανάλυση, με στόχο τη διερεύνηση βασικών ερευνητικών ερωτημάτων που σχετίζονται με τη συμπεριφορά των οχημάτων στη σηματοδοτημένη διάβαση και με τους δείκτες επικινδυνότητας που υπολογίστηκαν από τα δεδομένα τροχιάς. Για τον σκοπό αυτό εφαρμόστηκαν δύο συμπληρωματικές μεθοδολογικές προσεγγίσεις, η Διωνυμική Λογιστική Παλινδρόμηση (**Binomial Logistic Regression**) και τα Γενικευμένα Γραμμικά Μοντέλα (Generalized Linear Models – **GLM**).

Η πρώτη μέθοδος χρησιμοποιήθηκε στις περιπτώσεις όπου η εξαρτημένη μεταβλητή ήταν δυαδική, δηλαδή λάμβανε τιμές 0 και 1, όπως για παράδειγμα στην περίπτωση εκδήλωσης speeding ή στην περίπτωση δυαδικοποίησης των δεικτών **PET** και **TTC** με κατάλληλα κατώφλια (thresholds). Αντίστοιχα, τα **Γενικευμένα Γραμμικά Μοντέλα** αξιοποιήθηκαν τόσο για δυαδικές όσο και για συνεχείς εξαρτημένες μεταβλητές, με σκοπό την εκτίμηση της κατεύθυνσης και της έντασης της επίδρασης των ανεξάρτητων μεταβλητών.

Η στατιστική ανάλυση υλοποιήθηκε στη γλώσσα προγραμματισμού **Python**, σε περιβάλλον Visual Studio Code, με χρήση βιβλιοθηκών ανοικτού κώδικα για την προεπεξεργασία των δεδομένων, την εκτίμηση των υποδειγμάτων και την εξαγωγή συγκριτικών δεικτών προσαρμογής.

5.2 Ανάπτυξη Μοντέλων Διωνυμικής Παλινδρόμησης

Η ανάπτυξη των μοντέλων **διωνυμικής λογιστικής παλινδρόμησης** προϋπέθεσε τον σαφή καθορισμό της εξαρτημένης μεταβλητής κάθε υποδείγματος, καθώς και την επιλογή κατάλληλου συνόλου ανεξάρτητων μεταβλητών. Η εξαρτημένη μεταβλητή στις περιπτώσεις αυτές όφειλε να έχει δυαδική μορφή, έτσι ώστε το μοντέλο να εκτιμά την πιθανότητα εμφάνισης ενός συγκεκριμένου συμβάντος.

Στο πλαίσιο της παρούσας εργασίας αναπτύχθηκαν υποδείγματα για τρεις βασικές κατηγορίες εξαρτημένων μεταβλητών: α) το **PET** σε διαφορετικά κατώφλια δυαδικοποίησης, β) το **Speeding_during_Green_50_100**, γ) το **TTC_Leader** σε κατώφλι 1,5 s. Για κάθε περίπτωση δοκιμάστηκαν εναλλακτικοί συνδυασμοί ανεξάρτητων μεταβλητών, με στόχο να εντοπιστεί η προδιαγραφή που παρουσιάζει την καλύτερη ισορροπία μεταξύ στατιστικής σημαντικότητας, ερμηνευσιμότητας και συνολικής προσαρμογής.

Η αξιολόγηση των υποδειγμάτων βασίστηκε κυρίως στην τιμή των συντελεστών παλινδρόμησης, στα πρόσημά τους, στη στατιστική σημαντικότητα των μεταβλητών, καθώς και στους δείκτες AIC, BIC, **Pseudo R²** και **AUC**, όπου αυτό ήταν διαθέσιμο.

5.2.1 Εισαγωγή δεδομένων, προεπεξεργασία και έλεγχος συσχέτισης

Πριν από την εκτίμηση των μοντέλων πραγματοποιήθηκε προεπεξεργασία του συνόλου δεδομένων, μετατροπή των κατηγορικών μεταβλητών σε αριθμητική μορφή, επιλογή του **complete-case δείγματος** και υπολογισμός πίνακα συσχετίσεων Pearson. Παράλληλα υπολογίστηκε και ο δείκτης **πολυσυγγραμμικότητας** VIF, ώστε να εντοπιστούν ζεύγη μεταβλητών που ενδέχεται να μεταφέρουν κοινή πληροφορία.

Στο παρόν στάδιο της ανάλυσης δημιουργήθηκε ειδικό script σε Python, μέσω του οποίου πραγματοποιήθηκαν η φόρτωση του αρχείου δεδομένων, η αφαίρεση των μη χρήσιμων αναγνωριστικών στηλών, η αντιστοίχιση των κατηγορικών μεταβλητών σε αριθμητικές τιμές και η επιλογή των τελικών μεταβλητών που συμμετείχαν στον πίνακα συσχετίσεων. Η διαδικασία αυτή κρίθηκε αναγκαία, ώστε οι μεταβλητές να αποκτήσουν ενιαία μορφή και να καταστεί εφικτός ο υπολογισμός τόσο του πίνακα **συσχέτισης** Pearson όσο και του δείκτη **VIF**.

Ιδιαίτερη έμφαση δόθηκε στον έλεγχο της **πολυσυγγραμμικότητας** πριν από την ανάπτυξη των παλινδρομικών υποδειγμάτων. Για τον σκοπό αυτό χρησιμοποιήθηκε ως σημείο αναφοράς η τιμή $|r| < 0,70$ για τον αρχικό έλεγχο **συσχέτισης**, ενώ συμπληρωματικά εξετάστηκε και ο δείκτης Variance Inflation Factor (**VIF**), ώστε να επιβεβαιωθεί αν τυχόν έντονα συσχετισμένες μεταβλητές θα μπορούσαν να δημιουργήσουν προβλήματα σταθερότητας κατά την εκτίμηση των συντελεστών.

Στα Σχήματα 5.1–5.6 παρουσιάζονται ενδεικτικά αποσπάσματα του κώδικα που χρησιμοποιήθηκε για την προεπεξεργασία των δεδομένων και τη δημιουργία του διαγράμματος θερμότητας, καθώς και το τελικό **correlation matrix** του **complete-case δείγματος**.

```

# Εισαγωγή των απαραίτητων βιβλιοθηκών
import pandas as pd
import matplotlib.pyplot as plt
from statsmodels.stats.outliers_influence import variance_inflation_factor

# =====
# 1. Φόρτωση του αρχείου δεδομένων
# =====
# Το αρχείο csv διαβάζεται σε μορφή πίνακα δεδομένων (DataFrame).
# Η παράμετρος low_memory=False χρησιμοποιείται ώστε να γίνει
# πιο σταθερή αναγνώριση των τύπων δεδομένων ανά στήλη.
df = pd.read_csv("MasterTable_Chatzitheodosiou_v2.csv", low_memory=False)

# =====
# 2. Αφαίρεση στηλών που δεν χρησιμοποιούνται στην ανάλυση
# =====
# Οι παρακάτω στήλες αγνοούνται είτε επειδή αποτελούν αναγνωριστικά,
# είτε επειδή δεν είναι χρήσιμες για την παρούσα ανάλυση συσχέτισης.
ignore = [
    "Frame",                # Αριθμός frame βίντεο
    "_ + Is in cr Area",    # Βοηθητική / παράγωγη στήλη
    "speed_calc",          # Ενδιάμεση υπολογιστική στήλη
    "PET_ConflictVehicleID", # Αναγνωριστικό άλλου οχήματος
    "PET_TimeToConflict_Other_s", # Χρόνος σύγκρουσης του άλλου οχήματος
    "VehicleID",           # Αναγνωριστικό οχήματος
    "LeadingVehicleID"     # Αναγνωριστικό προπορευόμενου οχήματος
]

# Διαγράφονται μόνο όσες από τις παραπάνω στήλες υπάρχουν πράγματι στο αρχείο
df = df.drop(columns=[c for c in ignore if c in df.columns], errors="ignore")

# =====

```

Εικόνα 5.1: Απόσπασμα κώδικα Python για την εισαγωγή βιβλιοθηκών, φόρτωση δεδομένων και αφαίρεση μη χρήσιμων στηλών.

```

# =====
# 3. Μετατροπή κατηγορικών μεταβλητών σε αριθμητική μορφή
# =====
# Για να είναι δυνατός ο υπολογισμός πίνακα συσχέτισης και VIF,
# οι κατηγορικές / δυαδικές μεταβλητές πρέπει να μετατραπούν
# σε αριθμητικές τιμές.
map_dicts = {
    "IsInCrossingArea": {
        "0": 0, "1": 1, 0: 0, 1: 1,
        False: 0, True: 1, "False": 0, "True": 1
    },
    "VehicleStatus": {
        "legal": 0,
        "illegal": 1
    },
    "HasLeaderSameLane": {
        "False": 0, "True": 1,
        False: 0, True: 1
    },
    "Speeding_during_Orange_50_100": {
        "0": 0, "1": 1, 0: 0, 1: 1
    },
    "Speeding_during_Green_50_100": {
        "0": 0, "1": 1, 0: 0, 1: 1
    },
    "TrafficLightStatus": {
        "green": 0,
        "orange": 1,
        "red": 2
    },
    "ProcessedTrafficLightStatus": {
        "green": 0,
        "orange": 1,
        "red": 2
    },
    "Vehicle_Classification": {
        "car": 0,
        "moto": 1
    },
}

# Εφαρμογή της αντιστοίχισης σε κάθε στήλη που υπάρχει στο DataFrame
for col, mapping in map_dicts.items():
    if col in df.columns:
        df[col] = df[col].map(mapping)

```

Εικόνα 5.2: Απόσπασμα κώδικα Python για τη μετατροπή κατηγορικών μεταβλητών σε αριθμητική μορφή.

```

# =====
# 4. Μετατροπή όλων των στηλών σε αριθμητική μορφή
# =====
# Κάθε στήλη μετατρέπεται σε numeric.
# Τιμές που δεν μπορούν να μετατραπούν γίνονται NaN.
for col in df.columns:
    df[col] = pd.to_numeric(df[col], errors="coerce")

# =====
# 5. Επιλογή των μεταβλητών που θα συμμετέχουν στην ανάλυση
# =====
# Ορίζονται οι μεταβλητές που θα χρησιμοποιηθούν τόσο
# στον πίνακα συσχετίσεων όσο και στον έλεγχο VIF.
sel = [
    "Confidence",
    "GroundPlaneCentroid_1",
    "GroundPlaneCentroid_2",
    "IsInCrossingArea",
    "TrafficLightStatus",
    "ProcessedTrafficLightStatus",
    "VehicleStatus",
    "vx",
    "vy",
    "magnitude",
    "Speeding_during_Orange_50_100",
    "Speeding_during_Green_50_100",
    "Vehicle_Classification",
    "GapAlongPath_m",
    "LateralOffsetToLeader_m",
    "HasLeaderSameLane",
    "TTC_Leader_s",
    "PET_MinPredicted_s",
    "PET_TimeToConflict_Self_s"
]

# =====
# 6. Δημιουργία complete-case δείγματος
# =====
# Διατηρούνται μόνο οι γραμμές για τις οποίες υπάρχουν διαθέσιμες
# τιμές σε όλες τις επιλεγμένες μεταβλητές.
# Με αυτόν τον τρόπο αποφεύγεται η επίδραση ελλειπών παρατηρήσεων
# στους υπολογισμούς.
cc = df[sel].dropna().copy()

```

Εικόνα 5.3: Απόσπασμα κώδικα Python για τη μετατροπή μεταβλητών σε numeric μορφή και τον ορισμό του complete-case δείγματος.

```

# -----
# 7. Υπολογισμός πίνακα συσχέτισης
# -----
# Υπολογίζεται ο πίνακας συσχέτισης Pearson μεταξύ των μεταβλητών
# του complete-case δείγματος.
corr = cc.corr()

# Ενδεικτική εκτύπωση της συσχέτισης μεταξύ vx και vy
print("corr(vx, vy) =", round(corr.loc["vx", "vy"], 4))

# -----
# 8. Υπολογισμός δείκτη VIF
# -----
# Ο δείκτης VIF (Variance Inflation Factor) χρησιμοποιείται για
# τον έλεγχο πολυσυγγραμμικότητας μεταξύ των ανεξάρτητων μεταβλητών.
# Αφαιρούνται πρώτα οι στήλες που έχουν μία μόνο μοναδική τιμή,
# επειδή δεν μπορούν να δώσουν έγκυρο VIF.
X = cc.loc[:, [c for c in cc.columns if cc[c].nunique() > 1]].astype(float)

# Δημιουργία πίνακα με τη μεταβλητή και την αντίστοιχη τιμή VIF
vif = pd.DataFrame({
    "Variable": X.columns,
    "VIF": [variance_inflation_factor(X.values, i) for i in range(X.shape[1])]
}).sort_values("VIF", ascending=False)

# Εκτύπωση των αποτελεσμάτων VIF
print("\nΠίνακας VIF:")
print(vif)

# -----
# 9. Επιλογή μεταβλητών για το διάγραμμα θερμότητας
# -----
# Από τον συνολικό πίνακα συσχέτισης επιλέγονται συγκεκριμένες
# μεταβλητές, ώστε το διάγραμμα να είναι πιο ευανάγνωστο.
plot_cols = [
    "PET_MinPredicted_s",
    "PET_TimeToConflict_Self_s",
    "TTC_Leader_s",
    "GapAlongPath_m",
    "GroundPlaneCentroid_1",
    "GroundPlaneCentroid_2",
    "vx",
    "vy",
    "magnitude",
    "TrafficLightStatus",
    "ProcessedTrafficLightStatus",
    "VehicleStatus",
    "IsInCrossingArea"
]

# Διατηρούνται μόνο οι στήλες που υπάρχουν πράγματι στον πίνακα συσχέτισης
plot_cols = [c for c in plot_cols if c in corr.index]

# Δημιουργία υποπίνακα συσχέτισης μόνο για τις επιλεγμένες μεταβλητές
subcorr = corr.loc[plot_cols, plot_cols]

# -----

```

Εικόνα 5.4: Απόσπασμα κώδικα Python για τον υπολογισμό του πίνακα συσχέτισης και του δείκτη VIF.

```

# =====
# 10. Δημιουργία heatmap πίνακα συσχέτισης
# =====
# Κατασκευάζεται διάγραμμα θερμότητας για την οπτική απεικόνιση
# των συσχετίσεων μεταξύ των μεταβλητών.
fig, ax = plt.subplots(figsize=(12, 10))

# Προβολή του πίνακα συσχέτισης με χρωματική κλίμακα από -1 έως 1
im = ax.imshow(subcorr.values, cmap="coolwarm", vmin=-1, vmax=1)

# Ορισμός των ετικετών του οριζόντιου άξονα
ax.set_xticks(range(len(subcorr.columns)))
ax.set_xticklabels(subcorr.columns, rotation=90, fontsize=9)

# Ορισμός των ετικετών του κάθετου άξονα
ax.set_yticks(range(len(subcorr.index)))
ax.set_yticklabels(subcorr.index, fontsize=9)

# Τίτλος του διαγράμματος
ax.set_title("Correlation Matrix (Complete-case sample)", fontsize=15)

# =====
# 11. Εμφάνιση των τιμών συσχέτισης μέσα στα κελιά
# =====
# Σε κάθε κελί του heatmap εμφανίζεται και η αριθμητική τιμή
# της αντίστοιχης συσχέτισης.
for i in range(subcorr.shape[0]):
    for j in range(subcorr.shape[1]):
        val = float(subcorr.iloc[i, j])
        ax.text(
            j, i, f"{val:.2f}",
            ha="center",
            va="center",
            color="white" if abs(val) >= 0.5 else "black",
            fontsize=8
        )

# Προσθήκη χρωματικής μπάρας
fig.colorbar(im, ax=ax, fraction=0.046, pad=0.04)

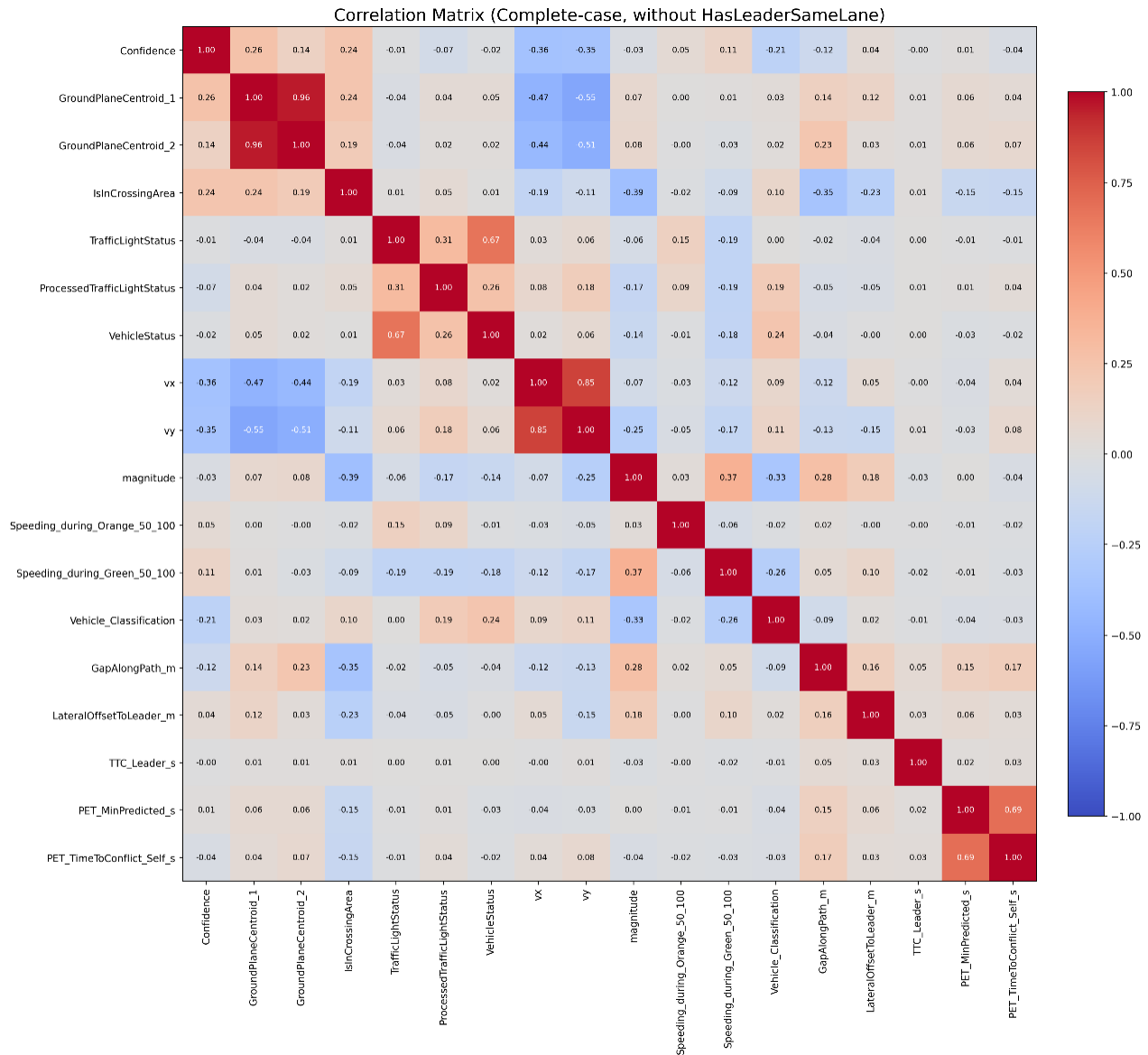
# Προσαρμογή διάταξης για αποφυγή επικάλυψης στοιχείων
plt.tight_layout()

# Αποθήκευση του διαγράμματος σε αρχείο εικόνας
plt.savefig("complete_case_correlation_diagram.png", dpi=220, bbox_inches="tight")

# Εμφάνιση του διαγράμματος στην οθόνη
plt.show()

```

Εικόνα 5.5: Απόσπασμα κώδικα Python για τη δημιουργία του **heatmap συσχέτισης** και την αποθήκευσή του σε αρχείο εικόνας.



Σχήμα 5.6: Correlation matrix του complete-case δείγματος.

Όπως προκύπτει από τα παραπάνω, ο έλεγχος **συσχέτισης** δεν αποτέλεσε απλώς ένα ενδιάμεσο τεχνικό βήμα, αλλά βασικό στάδιο επιλογής μεταβλητών πριν από τη διαμόρφωση των τελικών υποδειγμάτων. Με αυτόν τον τρόπο εξασφαλίστηκε ότι τα μοντέλα που παρουσιάζονται στις επόμενες ενότητες στηρίζονται σε ένα σύνολο μεταβλητών με όσο το δυνατόν μεγαλύτερη ερμηνευτική αυτοτέλεια.

Από τον πίνακα συσχετίσεων προέκυψε ότι οι περισσότερες μεταβλητές εμφανίζουν χαμηλό έως μέτριο βαθμό γραμμικής σχέσης, γεγονός που θεωρείται θετικό για την ανάπτυξη πολυμεταβλητών υποδειγμάτων. Ωστόσο, εντοπίστηκαν δύο σαφή ζεύγη με

ιδιαίτερα υψηλή θετική συσχέτιση: οι μεταβλητές **GroundPlaneCentroid_1** και **GroundPlaneCentroid_2** ($r = 0,9571$), καθώς και οι μεταβλητές v_x και v_y ($r = 0,8542$). Τα αποτελέσματα αυτά υποδηλώνουν ότι μέρος της χωρικής και κινηματικής πληροφορίας αλληλεπικαλύπτεται και, συνεπώς, η ταυτόχρονη εισαγωγή των αντίστοιχων μεταβλητών στα ίδια υποδείγματα απαιτεί προσοχή.

Αντίθετα, οι μεταβλητές **GapAlongPath_m**, **LateralOffsetToLeader_m**, **TTC_Leader_s**, **PET_MinPredicted_s** και **PET_TimeToConflict_Self_s** εμφάνισαν γενικά χαμηλές συσχετίσεις με τις περισσότερες λοιπές μεταβλητές, στοιχείο που ενισχύει την ερμηνευτική τους αυτοτέλεια. Επιπλέον, η μεταβλητή **HasLeaderSameLane** δεν εμφάνισε χρήσιμη διακύμανση στο complete-case δείγμα και, για τον λόγο αυτό, αποκλείστηκε από μέρος των τελικών προδιαγραφών.

5.3 Αποτελέσματα Μοντέλων Διωνυμικής Παλινδρόμησης

5.3.1 Παρουσίαση αποτελεσμάτων λογιστικής παλινδρόμησης για το PET

Πραγματοποιήθηκαν **τρία υποδείγματα διωνυμικής λογιστικής παλινδρόμησης** με εξαρτημένη μεταβλητή τον δείκτη **Post-Encroachment Time (PET)**, ο οποίος δυαδικοποιήθηκε σε τρία διαφορετικά κατώφλια, ήτοι **1,5 sec**, **3,0 sec** και **5,0 sec**, προκειμένου να εξεταστεί η ευαισθησία των αποτελεσμάτων ως προς τον ορισμό της κρίσιμης χρονικής εγγύτητας.

Η μεταβλητή **IsInCrossingArea** εμφάνισε **θετική και στατιστικά σημαντική επίδραση**. Το αποτέλεσμα αυτό υποδηλώνει ότι, όταν το όχημα βρίσκεται εντός της περιοχής διάβασης ή διασταύρωσης, η πιθανότητα εμφάνισης χαμηλού **PET** αυξάνεται σε σύγκριση με τις περιπτώσεις όπου το όχημα βρίσκεται εκτός της εν λόγω περιοχής. Η διαπίστωση αυτή είναι κυκλοφοριακά εύλογη, καθώς εντός της περιοχής σύγκρουσης οι αλληλεπιδράσεις μεταξύ οχημάτων τείνουν να είναι εντονότερες και τα διαθέσιμα χρονικά περιθώρια μικρότερα.

Αντίστοιχα, η μεταβλητή **GapAlongPath_m** παρουσίασε **αρνητική και στατιστικά σημαντική σχέση** με την πιθανότητα εμφάνισης χαμηλού **PET**. Το εύρημα αυτό σημαίνει ότι όσο αυξάνεται το **διαμήκες κενό** από το προπορευόμενο όχημα, τόσο μειώνεται η πιθανότητα παρατήρησης ιδιαίτερα χαμηλών τιμών PET. Η σχέση αυτή

θεωρείται αναμενόμενη, δεδομένου ότι η μεγαλύτερη διαμήκης απόσταση συνεπάγεται αυξημένο περιθώριο ασφάλειας και, κατά συνέπεια, μικρότερη πιθανότητα εμφάνισης κρίσιμης χρονικής εγγύτητας.

Ιδιαίτερα ισχυρή επίδραση καταγράφηκε για τη μεταβλητή **PET_TimeToConflict_Self_s**, η οποία εμφάνισε **αρνητική και εξαιρετικά στατιστικά σημαντική σχέση** με την πιθανότητα εμφάνισης χαμηλού **PET**. Το αποτέλεσμα αυτό δείχνει ότι όσο αυξάνεται ο χρόνος μέχρι το σημείο σύγκρουσης για το ίδιο το όχημα, τόσο μειώνεται ουσιαστικά η πιθανότητα εμφάνισης πολύ χαμηλού PET. Πρόκειται για εύρημα απολύτως συνεπές με τη φυσική σημασία των δεικτών επικινδυνότητας, καθώς ο μεγαλύτερος διαθέσιμος χρόνος αντανάκλα λιγότερο κρίσιμες συνθήκες αλληλεπίδρασης.

Συνολικά, τα αποτελέσματα της λογιστικής παλινδρόμησης για το **PET_MinPredicted_** δείχνουν ότι η πιθανότητα εμφάνισης ιδιαίτερα χαμηλών τιμών PET επηρεάζεται κυρίως από τη **θέση του οχήματος εντός της περιοχής διασταύρωσης**, από το **διαμήκες κενό** σε σχέση με το προπορευόμενο όχημα, καθώς και από τον **χρόνο μέχρι το σημείο σύγκρουσης** για το ίδιο το όχημα. Επομένως, οι μεταβλητές αυτές αναδεικνύονται ως βασικοί προσδιοριστικοί παράγοντες της επικινδυνότητας των αλληλεπιδράσεων στο εξεταζόμενο αστικό περιβάλλον.

Παρουσίαση αποτελεσμάτων λογιστικής παλινδρόμησης για το PET_MinPredicted_leq_1_5					
Binomial Regression PET_MinPredicted_leq_1_5					
Dep. Variable:	PET_MinPredicted_leq_1_5	No. Observ	10,355		
Model:	Logit	Df Residual	10,344	AIC:	10640.599
Method:	MLE	Df Model:	10	BIC:	10720.297
Date:	—	Pseudo R ² :	0.220	AUC:	0.808
converged:	TRUE	Log-Likelihood	-5309.300	LLR p-value:	0
Covariance Type:	—	LL-Null:	-6804.413		

Variable	Coef.	Std.Err.	z	P> z	[0.025	0.975]	OddsRatio
Intercept	35.031	4.308	8.131	<0.001	26.587	43.475	1635611232137740.000
GroundPlaneCentroid_2	0.017	0.003	6.765	<0.001	0.012	0.022	1.017
IsInCrossingArea	0.428	0.066	6.488	<0.001	0.299	0.557	1.534
magnitude	0.002	0.001	2.183	0.029	0.000	0.004	1.002
VehicleStatus	0.529	0.127	4.175	<0.001	0.281	0.777	1.697
Speeding_during_Orange_50_100	-0.150	0.374	-0.401	0.689	-0.884	0.584	0.861
GapAlongPath_m	-0.003	0.001	-3.594	<0.001	-0.005	-0.001	0.997
TTC_Leader_s	0.000	0.000	0.926	0.354	0.000	8.70E-04	1.000
LateralOffsetToLeader_m	0.023	0.028	0.822	0.411	-0.032	0.079	1.024
PET_TimeToConflict_Self_s	-0.525	0.013	-40.587	<0.001	-0.551	-0.500	0.591

Παρουσίαση αποτελεσμάτων λογιστικής παλινδρόμησης για το PET_MinPredicted_leq_3_0						
Binomial Regression PET_MinPredicted_leq_3_0						
Dep. Variable:	PET_MinPredicted_leq_3_0	No. Observations:	10,355			
Model:	Logit	Df Residuals:	10,344	AIC:	10051.584	
Method:	MLE	Df Model:	10	BIC:	10131.281	
Date:	—	Pseudo R-squ.:	0.281	AUC:	0.831	
converged:	TRUE	Log-Likelihood:	-5014.792	LLR p-value:	0	
Covariance Type:	—	LL-Null:	-6970.654			

Variable	Coef.	Std.Err.	z	P> z	[0.025	0.975]	OddsRatio
Intercept	30.765	4.238	7.259	<0.001	22.458	39.072	22964647117123.800
GroundPlaneCentroid_2	0.013	0.002	5.555	<0.001	0.009	0.018	1.013
IsInCrossingArea	0.398	0.066	6.072	<0.001	0.270	0.527	1.489
magnitude	0.000	0.001	0.115	0.908	-0.002	0.002	1.000
VehicleStatus	0.216	0.139	1.557	0.119	-0.056	0.488	1.241
Speeding_during_Orange_50_100	0.341	0.397	0.859	0.390	-0.436	1.118	1.406
GapAlongPath_m	-0.002	0.001	-3.013	0.003	-0.004	-7.96E-04	0.998
TTC_Leader_s	0.000	0.000	-0.224	0.822	-0.001	4.07E-04	1.000
LateralOffsetToLeader_m	0.021	0.029	0.727	0.467	-0.036	0.078	1.021
PET_TimeToConflict_Self_s	-0.554	0.011	-48.220	<0.001	-0.576	-0.531	0.575

Παρουσίαση αποτελεσμάτων λογιστικής παλινδρόμησης για το Speeding_during_Green_50_100						
# Binomial Regression Speeding_during_Green_50_100						
Dep. Variable:	Speeding_during_Green_50_100	No. Observations:	10355			
Model:	Logit	Df Residuals:	10346	AIC:	13896.930	
Method:	MLE	Df Model:	8	BIC:	13962.137	
Date:	—	Pseudo R-squ.:	0.299	AUC:	0.794	
converged:	TRUE	Log-Likelihood:	-6939.465	LLR p-value:	<0.001	
Covariance Type:	—	LL-Null:	-7150.282			

Variable	Coef.	Std.Err.	z	P> z	[0.025	0.975]
Intercept	-54.335	3.684	-14.749	<0.001	-61.555	-47.115
GroundPlaneCentroid_1	0.392	0.027	14.741	<0.001	0.340	0.444
IsInCrossingArea	-0.404	0.051	-7.859	<0.001	-0.505	-0.303
GapAlongPath_m	0.004	0.001	6.126	<0.001	0.003	0.006
TTC_Leader_s	-0.001	0.000	-1.709	0.087	-0.001	0.000
LateralOffsetToLeader_m	0.036	0.024	1.507	0.132	-0.011	0.082
PET_MinPredicted_s	-0.004	0.012	-0.297	0.767	-0.027	0.020
PET_TimeToConflict_Self_s	0.010	0.011	-2.810	0.005	-0.050	-0.009

Παρουσίαση αποτελεσμάτων λογιστικής παλινδρόμησης για το PET_MinPredicted_leq_5_0						
Binomial Regression PET_MinPredicted_leq_5_0						
Dep. Variable:	PET_MinPredicted_leq_5_0	No. Observations:	10,355			
Model:	Logit	Df Residuals:	10,344	AIC:	6837.834	
Method:	MLE	Df Model:	10	BIC:	6917.531	
Date:	—	Pseudo R-squ.:	0.340	AUC:	0.858	
converged:	TRUE	Log-Likelihood:	-3407.917	LLR p-value:	<0.001	
Covariance Type:	—	LL-Null:	-5163.594			

Variable	Coef.	Std.Err.	z	P> z	[0.025	0.975]	OddsRatio
Intercept	37.164	5.043	7.369	<0.001	27.280	47.049	1.38E+16
GroundPlaneCentroid_2	0.017	0.003	6.099	<0.001	0.012	0.023	1.017
IsInCrossingArea	0.199	0.079	2.520	0.012	0.044	0.353	1.220
magnitude	-0.002	0.001	-1.953	0.051	-0.005	8.55E-06	0.998
VehicleStatus	0.214	0.178	1.205	0.228	-0.134	0.563	1.239
Speeding_during_Orange_50_100	0.667	0.668	0.999	0.318	-0.642	1.977	1.949
GapAlongPath_m	-0.001	0.001	-1.202	0.229	-0.003	6.56E-04	0.999
TTC_Leader_s	0.000	0.000	-0.712	0.476	-0.001	3.18E-04	1.000
LateralOffsetToLeader_m	0.034	0.035	0.970	0.332	-0.035	0.104	1.035
PET_TimeToConflict_Self_s	-0.626	0.014	-46.110	<0.001	-0.653	-0.600	0.535

Πίνακας 5.1 Λογιστική Παλινδρόμηση για Post-Encroachment Time με τρία διαφορετικά thresholds

5.3.2 Παρουσίαση αποτελεσμάτων λογιστικής παλινδρόμησης για το Speeding_during_Green_

Παρουσίαση αποτελεσμάτων λογιστικής παλινδρόμησης για το Speeding_during_Green_50_100					
# Binomial Regression Speeding_during_Green_50_100					
Dep. Variable:	Speeding_during_Green_50_100	No. Observations:	10355		
Model:	Logit	Df Residuals:	10346	AIC:	13896.930
Method:	MLE	Df Model:	8	BIC:	13962.137
Date:	—	Pseudo R-squ.:	0.299	AUC:	0.794
converged:	TRUE	Log-Likelihood:	-6939.465	LLR p-value:	<0.001
Covariance Type:	—	LL-Null:	-7150.282		

Variable	Coef.	Std.Err.	z	P> z	[0.025	0.975]
Intercept	-54.335	3.684	-14.749	<0.001	-61.555	-47.115
GroundPlaneCentroid_1	0.392	0.027	14.741	<0.001	0.340	0.444
IsInCrossingArea	-0.404	0.051	-7.859	<0.001	-0.505	-0.303
GapAlongPath_m	0.004	0.001	6.126	<0.001	0.003	0.006
TTC_Leader_s	-0.001	0.000	-1.709	0.087	-0.001	0.000
LateralOffsetToLeader_m	0.036	0.024	1.507	0.132	-0.011	0.082
PET_MinPredicted_s	-0.004	0.012	-0.297	0.767	-0.027	0.020
PET_TimeToConflict_Self_s	0.010	0.011	-2.810	0.005	-0.050	-0.009

Πίνακας 5.2 Λογιστική Παλινδρόμηση για Speeding_during_Green

Η εξαρτημένη μεταβλητή του υποδείγματος ορίστηκε ως δυαδική, λαμβάνοντας την τιμή 1 όταν το όχημα εμφανίζει speeding κατά τη διάρκεια του πράσινου σηματοδότη και την τιμή 0 όταν δεν εμφανίζει. Το λογιστικό υπόδειγμα παρουσίασε Pseudo R² = 0.299 και AUC = 0.794, γεγονός που υποδηλώνει ότι διαθέτει ικανοποιητική έως καλή ερμηνευτική και διακριτική ικανότητα. Με άλλα λόγια, το σύνολο των επιλεγμένων ανεξάρτητων μεταβλητών αποτυπώνει σε σημαντικό βαθμό τη συμπεριφορά speeding, χωρίς ωστόσο να εξηγεί πλήρως τη διακύμανσή της.

Ως προς την ερμηνεία των στατιστικά σημαντικών μεταβλητών, η μεταβλητή **IsInCrossingArea** παρουσίασε αρνητική και στατιστικά σημαντική σχέση με την πιθανότητα εμφάνισης speeding. Το αποτέλεσμα αυτό σημαίνει ότι όταν το όχημα βρίσκεται εντός της περιοχής διάβασης ή διασταύρωσης, η πιθανότητα να εμφανίσει speeding μειώνεται. Η διαπίστωση αυτή είναι κυκλοφοριακά εύλογη, καθώς εντός της κρίσιμης περιοχής οι οδηγοί τείνουν να επιδεικνύουν μεγαλύτερη προσοχή, είτε λόγω αυξημένης αντιληπτής επικινδυνότητας είτε λόγω της πιθανής παρουσίας πεζών και άλλων συγκρουόμενων κινήσεων. Επομένως, η είσοδος του οχήματος στην περιοχή διάβασης φαίνεται να λειτουργεί ανασταλτικά ως προς την εκδήλωση πιο επιθετικής κινηματικής συμπεριφοράς.

Η μεταβλητή **GapAlongPath_m** εμφάνισε επίσης θετική και στατιστικά σημαντική επίδραση στην πιθανότητα speeding κατά το πράσινο. Ο θετικός συντελεστής

υποδηλώνει ότι όσο αυξάνεται το διαμήκες κενό μπροστά από το όχημα, τόσο αυξάνεται και η πιθανότητα εμφάνισης speeding. Από κυκλοφοριακή σκοπιά, το αποτέλεσμα αυτό δείχνει ότι όταν ο οδηγός διαθέτει περισσότερο ελεύθερο χώρο κίνησης μπροστά του και αισθάνεται λιγότερο περιορισμένος από προπορευόμενο όχημα, είναι πιθανότερο να αναπτύξει υψηλότερη ταχύτητα. Κατά συνέπεια, το διαθέσιμο διαμήκες περιθώριο φαίνεται να αποτελεί σημαντικό προσδιοριστικό παράγοντα της εκδήλωσης speeding.

Τέλος, η μεταβλητή **PET_TimeToConflict_Self_s** εμφάνισε θετική σχέση με την πιθανότητα εκδήλωσης speeding, γεγονός που σημαίνει ότι όσο αυξάνεται ο χρόνος μέχρι το σημείο σύγκρουσης για το ίδιο το όχημα, τόσο αυξάνεται η πιθανότητα ο οδηγός να αναπτύξει υπερβολική ταχύτητα. Η σχέση αυτή θα μπορούσε να ερμηνευθεί με βάση το ότι το μεγαλύτερο διαθέσιμο χρονικό περιθώριο ενισχύει το αίσθημα άνεσης ή ασφάλειας του οδηγού, οδηγώντας τον σε λιγότερο συγκρατημένη συμπεριφορά ως προς την ταχύτητα.

Συνολικά, τα αποτελέσματα του υποδείγματος δείχνουν ότι η εκδήλωση speeding κατά τη διάρκεια του πράσινου σηματοδότη επηρεάζεται κυρίως από τη χωρική θέση του οχήματος, από το αν αυτό βρίσκεται εντός της κρίσιμης περιοχής της διάβασης, από το διαθέσιμο διαμήκες κενό μπροστά του, καθώς και από τη χρονική εγγύτητα της αλληλεπίδρασης. Τα ευρήματα αυτά ενισχύουν την άποψη ότι η ακατάλληλη ταχύτητα δεν αποτελεί τυχαίο φαινόμενο, αλλά συνδέεται με συγκεκριμένα χαρακτηριστικά του κυκλοφοριακού περιβάλλοντος και της δυναμικής κίνησης του οχήματος.

5.3.3 Παρουσίαση αποτελεσμάτων λογιστικής παλινδρόμησης για το TTC_Leader

Παρουσίαση αποτελεσμάτων λογιστικής παλινδρόμησης για το TTC_Leader_leq_1_5					
# Binomial Regression TTC_Leader_leq_1_5					
Dep. Variable:	TTC_Leader_leq_1_5	No. Obser	10,355		
Model:	Logit	Df Residual	10,347	AIC:	6906.863
Method:	MLE	Df Model:	7	BIC:	6964.825
Date:	—	Pseudo R-	0.515	AUC:	0.928
converged:	TRUE	Log-Likelihood	-3445.432	LLR p-value	0
Covariance Type:	—	LL-Null:	-7101.057		

Variable	Coef.	Std.Err.	z	P> z	[0.025	0.975]	OddsRatio
Intercept	0.079	0.125	0.633	0.527	-0.166	0.323	1.082
IsInCrossingArea	-0.185	0.085	-2.179	0.029	-0.352	-0.019	0.831
magnitude	0.063	0.002	38.626	<0.001	0.060	0.066	1.065
Speeding_during_Orange_50_100	-0.155	0.599	-0.258	0.796	-1.329	1.019	0.857
GapAlongPath_m	-0.209	0.005	-41.833	<0.001	-0.219	-0.199	0.811
LateralOffsetToLeader_m	-0.410	0.037	-10.954	<0.001	-0.484	-0.337	0.663
PET_MinPredicted_s	0.028	0.018	1.571	0.116	-0.007	0.062	1.028
PET_TimeToConflict_Self_s	-0.094	0.016	-6.032	<0.001	-0.124	-0.063	0.911

Πίνακας 5.3 Λογιστική Παλινδρόμηση για Time to Collision

Η εξαρτημένη μεταβλητή ορίστηκε ως δυαδική, λαμβάνοντας την τιμή 1 όταν το $TTC_Leader_s \leq 1,5$ sec και την τιμή 0 όταν το $TTC_Leader_s > 1,5$ sec. Κατά συνέπεια, το υπόδειγμα εξετάζει ποιοι παράγοντες επηρεάζουν την πιθανότητα εμφάνισης ιδιαίτερα χαμηλών τιμών TTC, δηλαδή συνθηκών αυξημένης χρονικής εγγύτητας με το προπορευόμενο όχημα.

Το λογιστικό υπόδειγμα βασίστηκε σε 10.355 παρατηρήσεις και παρουσίασε $Pseudo R^2 = 0.519$ και $AUC = 0.928$, γεγονός που δείχνει ότι διαθέτει πολύ υψηλή διακριτική ικανότητα και πολύ ικανοποιητική συνολική προσαρμογή. Επομένως, οι μεταβλητές που περιλήφθηκαν στο μοντέλο ερμηνεύουν σε σημαντικό βαθμό την πιθανότητα εμφάνισης πολύ χαμηλού TTC.

Ως προς τις επιμέρους επιδράσεις, η μεταβλητή *IsInCrossingArea* εμφάνισε αρνητική και στατιστικά σημαντική σχέση με την πιθανότητα εμφάνισης $TTC_Leader_s \leq 1,5$ sec. Το αποτέλεσμα αυτό δείχνει ότι, όταν το όχημα βρίσκεται εντός της περιοχής διάβασης ή διασταύρωσης, η πιθανότητα εμφάνισης πολύ χαμηλού TTC μειώνεται. Κυκλοφοριακά, αυτό μπορεί να ερμηνευθεί ως ένδειξη ότι εντός της κρίσιμης περιοχής οι οδηγοί τείνουν να κινούνται με μεγαλύτερη προσοχή ή να προσαρμόζουν τη συμπεριφορά τους με τρόπο που μειώνει τις εξαιρετικά κρίσιμες συνθήκες ακολουθίας.

Η μεταβλητή *magnitude* εμφάνισε θετική και έντονα στατιστικά σημαντική επίδραση. Συνεπώς, όσο αυξάνεται η ταχύτητα του οχήματος, τόσο αυξάνεται και η πιθανότητα εμφάνισης $TTC_Leader_s \leq 1,5$ sec. Το εύρημα αυτό είναι πλήρως συμβατό με τη φυσική σημασία του δείκτη TTC, καθώς η υψηλότερη ταχύτητα μειώνει το διαθέσιμο

χρονικό περιθώριο αντίδρασης και αυξάνει την πιθανότητα δημιουργίας συνθηκών επικίνδυνης εγγύτητας με το προπορευόμενο όχημα.

Η μεταβλητή **GapAlongPath_m** παρουσίασε αρνητική και στατιστικά σημαντική επίδραση. Αυτό σημαίνει ότι όσο αυξάνεται το διαμήκες κενό από το προπορευόμενο όχημα, τόσο μειώνεται η πιθανότητα εμφάνισης ιδιαίτερα χαμηλού TTC. Από κυκλοφοριακή άποψη, το αποτέλεσμα αυτό θεωρείται απολύτως αναμενόμενο, καθώς η μεγαλύτερη διαμήκης απόσταση αυξάνει το περιθώριο ασφαλείας και περιορίζει την πιθανότητα δημιουργίας κρίσιμων χρονικών αποστάσεων.

Αντίστοιχα, η μεταβλητή **LateralOffsetToLeader_m** εμφάνισε επίσης αρνητική και στατιστικά σημαντική σχέση. Το αποτέλεσμα αυτό υποδηλώνει ότι όσο αυξάνεται η πλευρική απόσταση από το προπορευόμενο όχημα, τόσο μειώνεται η πιθανότητα εμφάνισης πολύ χαμηλού TTC. Η ερμηνεία αυτή είναι εύλογη, καθώς **η χωρική απομάκρυνση μεταξύ των οχημάτων**, όχι μόνο κατά μήκος αλλά και εγκάρσια, συνδέεται με **λιγότερο κρίσιμες συνθήκες αλληλεπίδρασης**.

Ιδιαίτερα σημαντική είναι και η επίδραση της μεταβλητής **PET_TimeToConflict_Self_s**, η οποία εμφάνισε αρνητική και στατιστικά σημαντική σχέση. Το αποτέλεσμα αυτό σημαίνει ότι **όσο αυξάνεται ο χρόνος μέχρι το σημείο σύγκρουσης** για το ίδιο το όχημα, **τόσο μειώνεται η πιθανότητα εμφάνισης χαμηλού $TTC_{Leader_s} \leq 1,5 \text{ sec}$** . Η σχέση αυτή είναι απολύτως συνεπής με τη σημασία των χρονικών δεικτών ασφάλειας, καθώς μεγαλύτερος διαθέσιμος χρόνος συνεπάγεται λιγότερο κρίσιμες συνθήκες αλληλεπίδρασης.

Συνολικά, τα αποτελέσματα του υποδείγματος δείχνουν ότι η πιθανότητα εμφάνισης χαμηλού TTC_{Leader_s} επηρεάζεται κυρίως από την ταχύτητα του οχήματος, το διαμήκες και πλευρικό κενό από το προπορευόμενο όχημα, καθώς και από τον χρόνο μέχρι το σημείο σύγκρουσης για το ίδιο το όχημα. Τα ευρήματα αυτά επιβεβαιώνουν ότι οι ιδιαίτερα κρίσιμες τιμές TTC_{Leader_s} δεν αποτελούν τυχαίο φαινόμενο, αλλά συνδέονται συστηματικά με τη χωρική και χρονική δομή της αλληλεπίδρασης μεταξύ οχημάτων.

5.4 Αποτελέσματα Γενικευμένων Γραμμικών Μοντέλων

5.4.1 Αποτελέσματα GLM για το PET

GLM Regression Scenarios						
Dep. Variable:	PET_MinPredicted_s	No. Observations:	10355			
Model:	GLM	Df Residuals:	10333			
Model Family:	Gaussian	Df Model:	21	AIC:	40167.181	
Link Function:	Identity	Scale:	1			
Method:	IRLS	Log-Likelihood:	-20061.591	BIC:	40326.576	
Date:	—	Deviance:	29205.482			
No. Iterations:	3	Pearson chi2:	29205.482	AUC:	—	
Covariance Type:	nonrobust	Pseudo R-squ. (CS):	0.619			
Variable	Coef.	Std.Err.	z	P> z	[0.025	0.975]
Intercept	-21.631	3.261	-6.633	<0.001	-28.023	-15.240
C(IsInCrossingArea)[T.1]	-0.295	0.045	-6.599	<0.001	-0.383	-0.207
C(TrafficLightStatus)[T.orange]	0.689	0.332	2.074	0.038	0.038	1.341
C(TrafficLightStatus)[T.red]	0.642	0.199	3.225	0.001	0.252	1.032
C(ProcessedTrafficLightStatus)[T.orange]	-0.299	0.178	-1.680	0.093	-0.649	0.050
C(ProcessedTrafficLightStatus)[T.red]	-0.040	0.059	-0.667	0.505	-0.156	0.077
C(VehicleStatus)[T.legal]	0.510	0.131	3.904	<0.001	0.254	0.766
C(Vehicle_Classification)[T.moto]	-0.017	0.076	-0.219	0.826	-0.166	0.132
Confidence	0.047	0.174	0.270	0.787	-0.294	0.388
GroundPlaneCentroid_1	0.159	0.024	6.716	<0.001	0.113	0.206
vy	-0.019	0.003	-6.909	<0.001	-0.024	-0.013
Speeding_during_Orange_50_100	-0.824	0.403	-2.048	0.041	-1.614	-0.035
Speeding_during_Green_50_100	-0.066	0.038	-1.741	0.082	-0.140	0.008
speed_calc	-0.007	0.004	-1.649	0.099	-0.015	0.001
GapAlongPath_m	0.002	5.16E-04	3.833	<0.001	0.001	0.003
LateralOffsetToLeader_m	-0.061	0.021	-2.903	0.004	-0.102	-0.020
TTC_Leader_s	3.38E-05	1.48E-04	0.229	0.819	0.000	0.000
PET_TimeToConflict_Self_s	0.604	0.006	93.463	<0.001	0.591	0.617
PET_TimeToConflict_Other_s	0.033	0.008	4.000	<0.001	0.017	0.049

Πίνακας 5.4 GLM για Post-Encroachment Time

Η εξαρτημένη μεταβλητή του υποδείγματος είναι η συνεχής μεταβλητή PET_MinPredicted_s και, για τον λόγο αυτό, το μοντέλο εκτιμήθηκε ως Generalized Linear Model (GLM) με οικογένεια Gaussian και συνάρτηση σύνδεσης identity. Το υπόδειγμα βασίστηκε σε 10.355 παρατηρήσεις και παρουσίασε ιδιαίτερα ικανοποιητική συνολική προσαρμογή, γεγονός που δείχνει ότι εξηγεί σημαντικό μέρος της διακύμανσης της εξαρτημένης μεταβλητής. Κατά συνέπεια, το συγκεκριμένο μοντέλο μπορεί να θεωρηθεί κατάλληλο για την ερμηνεία των παραγόντων που επηρεάζουν το επίπεδο του Post-Encroachment Time (PET).

Σε όρους ερμηνείας, θετικός συντελεστής σημαίνει ότι η αντίστοιχη μεταβλητή συνδέεται με υψηλότερες τιμές PET, δηλαδή με μεγαλύτερο χρονικό περιθώριο ασφάλειας, ενώ αρνητικός συντελεστής σημαίνει ότι συνδέεται με χαμηλότερες τιμές PET, άρα με περισσότερο κρίσιμες συνθήκες αλληλεπίδρασης.

Η μεταβλητή **IsInCrossingArea** εμφάνισε αρνητική και στατιστικά σημαντική επίδραση στο επίπεδο του PET. Το αποτέλεσμα αυτό υποδηλώνει ότι όταν το όχημα βρίσκεται εντός της περιοχής διάβασης ή διασταύρωσης, το PET τείνει να λαμβάνει χαμηλότερες τιμές. Συνεπώς, η παρουσία του οχήματος στην κρίσιμη περιοχή συνδέεται με

μικρότερο χρονικό περιθώριο μεταξύ αλληλεπιδρώντων κινήσεων και, κατά συνέπεια, με αυξημένη επικινδυνότητα.

Η μεταβλητή **Speeding_during_Orange_50_100**, η οποία συνδέθηκε αρνητικά και στατιστικά σημαντικά με το PET. Το αποτέλεσμα αυτό δείχνει ότι η εκδήλωση speeding κατά το πορτοκαλί συνδέεται με χαμηλότερες τιμές PET, δηλαδή με περισσότερο κρίσιμες χρονικές συνθήκες. Η διαπίστωση αυτή είναι κυκλοφοριακά εύλογη, καθώς η ανάπτυξη υψηλότερης ταχύτητας κοντά στη μετάβαση του σηματοδότη περιορίζει τα διαθέσιμα περιθώρια ασφάλειας.

Αντίθετα, για τη μεταβλητή **Speeding_during_Green_50_100** η οποία συνδέθηκε επίσης αρνητικά, δεν προέκυψε στατιστικά σημαντική ανεξάρτητη (ενδεχομένως οριακά) επίδραση στο επίπεδο του PET, γεγονός που υποδηλώνει ότι η επίδρασή της δεν είναι εξίσου σαφής όταν συνεκτιμώνται οι υπόλοιποι παράγοντες του υποδείγματος.

Η μεταβλητή **GapAlongPath_m** εμφάνισε θετική και στατιστικά σημαντική σχέση με το PET. Αυτό σημαίνει ότι όσο αυξάνεται η διαμήκης απόσταση από το προπορευόμενο όχημα, τόσο αυξάνεται και το επίπεδο του PET. Το εύρημα αυτό είναι απολύτως αναμενόμενο, καθώς μεγαλύτερο διαμήκης κενό μεταφράζεται σε μεγαλύτερο χρονικό περιθώριο ασφάλειας και, συνεπώς, σε λιγότερο κρίσιμες συνθήκες αλληλεπίδρασης.

Τέλος, ιδιαίτερα ισχυρή ήταν η επίδραση της μεταβλητής **PET_TimeToConflict_Self_s**, η οποία συνδέθηκε θετικά και στατιστικά σημαντικά με το PET. Το αποτέλεσμα αυτό υποδηλώνει ότι όσο αυξάνεται ο χρόνος μέχρι το σημείο σύγκρουσης για το ίδιο το όχημα, τόσο αυξάνεται και το PET. Πρόκειται για εύρημα απολύτως συμβατό με τη φυσική σημασία των δεικτών οδικής ασφάλειας, καθώς και οι δύο μεταβλητές αποτυπώνουν διαφορετικές όψεις της χρονικής εγγύτητας της σύγκρουσης. Αντίστοιχα, και ο δείκτης που αφορά το άλλο εμπλεκόμενο όχημα κινείται προς την ίδια κατεύθυνση, ενισχύοντας την ερμηνεία ότι η χρονική δομή της αλληλεπίδρασης αποτελεί καθοριστικό προσδιοριστικό παράγοντα του PET.

Συνολικά, τα αποτελέσματα του υποδείγματος δείχνουν ότι το **PET_MinPredicted_s** διαμορφώνεται από συνδυασμό χωρικών, λειτουργικών και χρονικών παραμέτρων. Ειδικότερα, η παρουσία του οχήματος εντός της περιοχής διασταύρωσης, η φάση της σηματοδότησης, η νομιμότητα της κίνησης, το διαθέσιμο διαμήκης κενό και οι χρονικοί δείκτες σύγκρουσης αναδεικνύονται ως οι σημαντικότεροι προσδιοριστικοί παράγοντες του δείκτη. Επομένως, το PET δεν αποτελεί μια απομονωμένη χρονική μέτρηση, αλλά έναν σύνθετο δείκτη επικινδυνότητας που αντανακλά τη συνολική δομή της αλληλεπίδρασης μεταξύ των οχημάτων στο εξεταζόμενο αστικό περιβάλλον.

5.4.2 Αποτελέσματα GLM για το Speeding_during_Green_

GLM Regression Scenarios						
Dep. Variable:	Speeding_during_Green_50_100	No. Observations:	10355			
Model:	GLM	Df Residuals:	10340			
Model Family:	Binomial	Df Model:	14	AIC:	12319.175	
Link Function:	Logit	Scale:	1			
Method:	IRLS	Log-Likelihood:	-6144.588	BIC:	12427.854	
Date:	—	Deviance:	12289.175			
No. Iterations:	4	Pearson chi2:	12018.405	AUC:	0.793	
Covariance Type:	nonrobust	Pseudo R-squ. (CS):	0.177			
Variable	Coef.	Std.Err.	z	P> z	[0.025	0.975]
Intercept	-36.830	4.206	-8.756	<0.001	-45.074	-28.586
C(IsInCrossingArea)[T.1]	0.165	0.059	2.821	0.005	0.050	0.280
Confidence	1.042	0.225	4.629	<0.001	0.601	1.483
GroundPlaneCentroid_1	0.247	0.031	8.024	<0.001	0.186	0.307
vx	-0.181	0.034	-5.395	<0.001	-0.247	-0.115
GapAlongPath_m	0.000	0.001	-0.374	0.708	-0.002	0.001
LateralOffsetToLeader_m	0.045	0.028	1.616	0.106	-0.010	0.101
TTC_Leader_s	0.000	0.000	-1.050	0.294	-0.001	0.000
PET_MinPredicted_s	-0.023	0.013	-1.747	0.081	-0.049	0.003
PET_TimeToConflict_Self_s	0.028	0.012	2.403	0.016	0.005	0.051
PET_TimeToConflict_Other_s	-0.021	0.011	-1.965	0.049	-0.042	0.000

Πίνακας 5.5 GLM για Speeding_during_Green

Η εξαρτημένη μεταβλητή του υποδείγματος είναι η δυαδική μεταβλητή **Speeding_during_Green_50_100**, η οποία λαμβάνει την τιμή **1** όταν το όχημα εμφανίζει speeding κατά τη διάρκεια της πράσινης ένδειξης και την τιμή **0** όταν δεν εμφανίζει. Για τον λόγο αυτό, το υπόδειγμα εκτιμήθηκε ως **Generalized Linear Model (GLM)** με οικογένεια **Binomial** και συνάρτηση σύνδεσης **logit**. Το μοντέλο βασίστηκε σε **10.355 παρατηρήσεις** και παρουσίασε **Pseudo R² (CS) = 0.177** και **AUC = 0.793**, γεγονός που δείχνει ότι διαθέτει **μέτρια προς καλή διακριτική ικανότητα** και ικανοποιητική συνολική ερμηνευτική ισχύ.

Η μεταβλητή **IsInCrossingArea** εμφάνισε θετική και στατιστικά σημαντική επίδραση στην πιθανότητα εμφάνισης Speeding_during_Green_50_100. Το αποτέλεσμα αυτό υποδηλώνει ότι όταν το όχημα βρίσκεται εντός της κρίσιμης περιοχής της διασταύρωσης ή της διάβασης, αυξάνεται η πιθανότητα να αναπτύξει speeding κατά τη διάρκεια της πράσινης ένδειξης. Κυκλοφοριακά, η σχέση αυτή δείχνει ότι η χωρική ένταξη του οχήματος στην περιοχή σύγκρουσης συνδέεται με πιο επιθετική ή πιο αδιάκοπη κινηματική συμπεριφορά, ενδεχομένως επειδή ο οδηγός επιλέγει να διατηρήσει τη ροή της κίνησής του και να μην επιβραδύνει καθώς διέρχεται από την περιοχή της διασταύρωσης.

Αντίστοιχα, η μεταβλητή **PET_TimeToConflict_Self_s** εμφάνισε επίσης θετική και στατιστικά σημαντική επίδραση στην πιθανότητα εμφάνισης speeding κατά το πράσινο. Το εύρημα αυτό σημαίνει ότι όσο αυξάνεται ο χρόνος μέχρι το σημείο σύγκρουσης για το ίδιο το όχημα, τόσο αυξάνεται και η πιθανότητα εμφάνισης speeding. Η ερμηνεία αυτού του αποτελέσματος είναι ότι όταν ο οδηγός αντιλαμβάνεται πως διαθέτει μεγαλύτερο χρονικό περιθώριο ασφαλείας στην αλληλεπίδραση με τα υπόλοιπα

οχήματα ή χρήστες της οδού, αισθάνεται περισσότερο «άνετα» να διατηρήσει ή και να αυξήσει την ταχύτητά του. Συνεπώς, το μεγαλύτερο αντιλαμβανόμενο χρονικό περιθώριο δεν λειτουργεί περιοριστικά, αλλά αντιθέτως φαίνεται να ευνοεί την εκδήλωση speeding κατά την πράσινη φάση.

Συνολικά, τα δύο αυτά αποτελέσματα υποδηλώνουν ότι η εκδήλωση speeding κατά το πράσινο δεν σχετίζεται μόνο με τη γενική ταχύτητα του οχήματος, αλλά και με τη θέση του στο χώρο καθώς και με τη χρονική εγγύτητα της αλληλεπίδρασής του με τα υπόλοιπα στοιχεία της κυκλοφορίας. Με άλλα λόγια, όταν το όχημα κινείται μέσα στην περιοχή της διασταύρωσης και ταυτόχρονα διαθέτει μεγαλύτερο χρονικό περιθώριο σε σχέση με πιθανή σύγκρουση, αυξάνεται η πιθανότητα ο οδηγός να υιοθετήσει πιο επιθετική κινηματική συμπεριφορά.

Κεφάλαιο 6: Συμπεράσματα

6.1 Σύνοψη

Η παρούσα διπλωματική εργασία είχε ως βασικό αντικείμενο τη διερεύνηση της **ακατάλληλης ταχύτητας** οχημάτων σε **σηματοδοτούμενες διαβάσεις** πεζών στην **Αθήνα** και την αποτίμηση του τρόπου με τον οποίο η συμπεριφορά αυτή συνδέεται με κρίσιμες συνθήκες αλληλεπίδρασης οχημάτων και πεζών. Η ανάλυση βασίστηκε σε πραγματικά δεδομένα βίντεο και σε μεθόδους **υπολογιστικής όρασης** και στατιστικής μοντελοποίησης, με στόχο όχι μόνο την καταγραφή του φαινομένου αλλά και την ερμηνεία των παραγόντων που το επηρεάζουν.

Ως προς την πορεία δημιουργίας της έρευνας, αρχικά καθορίστηκε το **ερευνητικό πρόβλημα**, με κεντρικό ερώτημα το πότε και υπό ποιες συνθήκες τα οχήματα αναπτύσσουν ακατάλληλη ταχύτητα κοντά στη διάβαση πεζών και πώς η συμπεριφορά αυτή σχετίζεται με την επικινδυνότητα της αλληλεπίδρασης. Στη συνέχεια πραγματοποιήθηκε **βιβλιογραφική ανασκόπηση**, κατά την οποία αναζητήθηκαν και αξιολογήθηκαν σύγχρονες ερευνητικές προσεγγίσεις σχετικά με τη διαχείριση της ταχύτητας, τη συμπεριφορά των οδηγών σε αστικό περιβάλλον, τη χρήση surrogate safety indicators και τις εφαρμογές της υπολογιστικής όρασης στην ανάλυση της κυκλοφορίας. Ακολούθως συγκεντρώθηκε και επεξεργάστηκε το βιντεοληπτικό υλικό από σηματοδοτούμενη διάβαση πεζών, ώστε να εντοπιστούν τα οχήματα, να παρακολουθηθούν οι τροχιές τους και να εξαχθούν οι βασικές μεταβλητές κίνησης.

Με βάση τα δεδομένα αυτά υπολογίστηκαν κρίσιμοι **δείκτες ασφάλειας**, όπως η ταχύτητα, οι χωρικές αποστάσεις, τα διαθέσιμα κενά, ο χρόνος μέχρι σύγκρουση (TTC) και ο χρόνος μετά την εμπλοκή (PET), και στη συνέχεια δημιουργήθηκε μια ενιαία βάση δεδομένων υψηλής χρονικής ανάλυσης. Έπειτα πραγματοποιήθηκε στατιστικός έλεγχος και προετοιμασία των μεταβλητών, με έλεγχο συσχέτισης, εντοπισμό πιθανής πολυσυγγραμμικότητας και επιλογή των καταλληλότερων ανεξάρτητων μεταβλητών πριν από την ανάπτυξη των υποδειγμάτων.

Σε επόμενο στάδιο αναπτύχθηκαν τα **στατιστικά μοντέλα**, και συγκεκριμένα μοντέλα **δυναμικής λογιστικής παλινδρόμησης** για τις δυναμικές εκβάσεις που σχετίζονται με το speeding, το PET και το TTC, καθώς και **Γενικευμένα Γραμμικά Μοντέλα (GLM)** για τη διερεύνηση της επίδρασης των παραγόντων σε συνεχείς δείκτες επικινδυνότητας. Τέλος, τα αποτελέσματα ερμηνεύτηκαν και συντέθηκαν με βάση τη στατιστική σημαντικότητα, την κατεύθυνση και το μέγεθος των συντελεστών, αλλά και την ικανότητα προσαρμογής των μοντέλων, ώστε να εξαχθούν πρακτικά συμπεράσματα για την οδική ασφάλεια στις διαβάσεις πεζών.

6.2 Συμπεράσματα

Η συνολική ανάλυση έδειξε ότι η **ακατάλληλη ταχύτητα** κοντά στις διαβάσεις πεζών δεν αποτελεί τυχαίο φαινόμενο, αλλά επηρεάζεται από συγκεκριμένα χωρικά, κινηματικά και χρονικά χαρακτηριστικά της κυκλοφορίας. Τα μοντέλα για το speeding κατά το πράσινο και το πορτοκαλί ανέδειξαν ότι η συμπεριφορά των οδηγών σχετίζεται κυρίως με τη θέση του οχήματος στο πεδίο της διασταύρωσης, με το αν το όχημα βρίσκεται εντός της περιοχής της διάβασης και με το διαθέσιμο **διαμήκες κενό** μπροστά του. Με απλά λόγια, όταν ο οδηγός διαθέτει μεγαλύτερο ελεύθερο χώρο και μεγαλύτερο αντιλαμβανόμενο χρονικό περιθώριο, είναι πιθανότερο να διατηρήσει ή και να αναπτύξει υψηλότερη ταχύτητα. Αντίθετα, η είσοδος στην κρίσιμη περιοχή της διάβασης φαίνεται σε αρκετές περιπτώσεις να λειτουργεί ανασταλτικά, πιθανότατα λόγω αυξημένης προσοχής.

Ιδιαίτερο ενδιαφέρον παρουσιάζουν τα αποτελέσματα για το **speeding κατά το πορτοκαλί**, καθώς η συμπεριφορά αυτή αποδείχθηκε γενικά δυσκολότερο να προβλεφθεί σε σχέση με το speeding κατά το πράσινο. Η εικόνα αυτή υποδηλώνει ότι στην απόφαση του οδηγού να επιταχύνει ή να συνεχίσει με ακατάλληλη ταχύτητα όταν το σήμα αλλάζει εμπλέκονται και παράγοντες που δεν αποτυπώνονται πλήρως στα διαθέσιμα δεδομένα, όπως η στιγμιαία **εκτίμηση κινδύνου**, η πίεση χρόνου ή η υποκειμενική κρίση του οδηγού. Επιπλέον, το σχετικά περιορισμένο μέγεθος δείγματος για ορισμένες υποπεριπτώσεις της διάβασης είναι πιθανό να περιορίσει την ικανότητα ανάδειξης ακόμη πιο σταθερών στατιστικών σχέσεων.

Ακόμη πιο καθαρά και ουσιαστικά ήταν τα ευρήματα για τους δείκτες **PET και TTC**, όπου αναδείχθηκαν με μεγαλύτερη σαφήνεια οι παράγοντες που συνδέονται με κρίσιμες συνθήκες αλληλεπίδρασης. Τα μοντέλα έδειξαν ότι χαμηλές τιμές PET και TTC συνδέονται κυρίως με μικρότερες χωρικές αποστάσεις, με χαμηλότερο διαθέσιμο χρονικό περιθώριο μέχρι πιθανή σύγκρουση και, σε ορισμένες περιπτώσεις, με υψηλότερη ταχύτητα του οχήματος. Ειδικά το υπόδειγμα για **TTC \leq 1,5 sec** παρουσίασε την ισχυρότερη επίδοση από τα εξεταζόμενα μοντέλα, γεγονός που δείχνει ότι οι πιο κρίσιμες συνθήκες ακολουθίας και εγγύτητας ερμηνεύονται αποτελεσματικά από μεταβλητές όπως η ταχύτητα, το διαμήκες και το πλευρικό κενό και οι χρονικοί δείκτες αλληλεπίδρασης.

Παράλληλα, τα αποτελέσματα των **μοντέλων PET** επιβεβαίωσαν ότι η επικινδυνότητα δεν εξαρτάται από έναν μόνο παράγοντα, αλλά από τον συνδυασμό γεωμετρίας κίνησης, θέσης του οχήματος και χρονικής εγγύτητας με τα υπόλοιπα κινούμενα στοιχεία του συστήματος. Η συστηματική εμφάνιση στατιστικά σημαντικής σχέσης μεταξύ των χρονικών δεικτών αλληλεπίδρασης και των εκβάσεων PET και TTC ενισχύει την αξία των **surrogate safety indicators** ως εργαλείων προληπτικής εκτίμησης της οδικής ασφάλειας, χωρίς να απαιτείται εκ των υστέρων καταγραφή πραγματικών ατυχημάτων.

Συνολικά, η παρούσα εργασία καταλήγει στο ότι η **επικινδυνότητα** κοντά στη διάβαση πεζών διαμορφώνεται από τον συνδυασμό της γεωμετρίας της κίνησης, της χρονικής εγγύτητας μεταξύ οχημάτων και της οδικής συμπεριφοράς. Κατά συνέπεια, η ακατάλληλη ταχύτητα πρέπει να αντιμετωπίζεται όχι μόνο ως υπέρβαση ενός τυπικού ορίου, αλλά ως συμπεριφορά που εξαρτάται από το πραγματικό λειτουργικό πλαίσιο της διάβασης. Η συμβολή της εργασίας έγκειται ακριβώς στο ότι παρουσιάζει, με συστηματικό και μετρήσιμο τρόπο, πώς δεδομένα βίντεο, αλγόριθμοι **υπολογιστικής όρασης** και στατιστικά υποδείγματα μπορούν να συνδυαστούν για την κατανόηση και την πρόληψη επικινδυνών αλληλεπιδράσεων σε αστικό περιβάλλον.

Παρακάτω δίνεται ο συγκεντρωτικός πίνακας των αποτελεσμάτων των μοντέλων για τις μεθόδους που χρησιμοποιήθηκαν:

<i>Παρουσίαση αποτελεσμάτων λογιστικών παλινδρομήσεων</i>						
	Binomial Regression PET_MinPredicted_Leq_3_0		Binomial Regression Speeding_during_Green_50_100		Binomial Regression TTC_Leader_Leq_1_5	
Dep. Variable:	PET_MinPredicted_Leq_3_0		Speeding_during_Green_50_100		TTC_Leader_Leq_1_5	
AIC:	10051.584		13896.930		6906.863	
AUC:	0.831		0.794		0.928	
BIC:	10131.281		13962.137		6964.825	
Pseudo R-squ.:	0.281		0.290		0.515	
Variable	Coef.	P> z	Coef.	P> z	Coef.	P> z
Intercept	30.765	<0.001	-54.335	<0.001	0.078889597	0.527
GroundPlaneCentroid_1			0.392	<0.001		
GroundPlaneCentroid_2	0.013	<0.001				
IsInCrossingArea	0.398	<0.001	-0.404	<0.001	-0.185241545	0.029
magnitude	0.000	0.908			0.063092021	<0.001
VehicleStatus	0.216	0.119				
Speeding_during_Orange_50_100	0.341	0.390			-0.154656305	0.796
GapAlongPath_m	-0.002	0.003	0.004	<0.001	-0.209145706	<0.001
TTC_Leader_s	0.000	0.822	-0.001	0.087		
LateralOffsetToLeader_m	0.021	0.467	0.036	0.132	-0.410401177	<0.001
PET_MinPredicted_s			-0.004	0.767	0.027662049	0.116
PET_TimeToConflict_Self_s	-0.554	<0.001	-0.030	0.005	-0.093582771	<0.001

Πίνακας 6.1 Αποτελέσματα μοντέλων δυαδικής λογιστικής παλινδρόμησης

GLM Regression Scenarios				
	Generalized Linear Model για PET		Generalized Linear Model για Speeding_during_Green	
Dep. Variable:	PET_MinPredicted_s		Speeding_during_Green_50_100	
AIC:	40167.181		12319.175	
AUC:	—		0.793	
BIC:	40326.576		12427.854	
Pseudo R-squ.:	0.619		0.177	
Variable	Coef.	P> z	Coef.	P> z
Intercept	-21.631	<0.001	-36.830	<0.001
C(IsInCrossingArea)[T.1]	-0.295	<0.001	0.165	0.005
C(TrafficLightStatus)[T.orange]	0.689	0.038		
C(TrafficLightStatus)[T.red]	0.642	0.001		
C(ProcessedTrafficLightStatus)[T.orange]	-0.299	0.093		
C(ProcessedTrafficLightStatus)[T.red]	-0.040	0.505		
C(VehicleStatus)[T.legal]	0.510	<0.001		
C(Vehicle_Classification)[T.moto]	-0.017	0.826		
Confidence	0.047	0.787	1.042	<0.001
GroundPlaneCentroid_1	0.159	<0.001	0.247	<0.001
vx			-0.181	<0.001
vy	-0.019	<0.001		
Speeding_during_Orange_50_100	-0.824	0.041		
Speeding_during_Green_50_100	-0.066	0.082		
speed_calc	-0.007	0.099	-0.015	0.007
GapAlongPath_m	0.002	<0.001	0.000	0.708
LateralOffsetToLeader_m	-0.061	0.004	0.045	0.106
TTC_Leader_s	0.000	0.819	0.000	0.294
PET_MinPredicted_s			-0.023	0.081
PET_TimeToConflict_Self_s	0.604	<0.001	0.028	0.016
PET_TimeToConflict_Other_s	0.033	<0.001	-0.021	0.049

Πίνακας 6.2 Αποτελέσματα Γενικευμένων Γραμμικών Μοντέλων

6.3 Προτάσεις για την Αξιοποίηση των Αποτελεσμάτων

Τα αποτελέσματα της παρούσας εργασίας μπορούν να αξιοποιηθούν για τη **στοχευμένη ιεράρχηση** παρεμβάσεων σε **σηματοδοτούμενες διαβάσεις πεζών** με αυξημένη πιθανότητα εμφάνισης speeding ή κρίσιμων αλληλεπιδράσεων.

Η χρήση **αυτοματοποιημένων συστημάτων παρακολούθησης** με βάση την **υπολογιστική όραση** μπορεί να υποστηρίξει τους φορείς διαχείρισης κυκλοφορίας στην τακτική παρακολούθηση της συμπεριφοράς των οδηγών και στον εντοπισμό σημείων αυξημένου κινδύνου.

Τα ευρήματα σχετικά με τον ρόλο της σηματοδότησης και του διαθέσιμου **διαμήκους κενού** μπορούν να αξιοποιηθούν **για την επανεξέταση παραμέτρων λειτουργίας της φωτεινής σηματοδότησης**, της διαγράμμισης και της οργάνωσης του χώρου προσέγγισης της διάβασης.

Η **ενσωμάτωση SMoS**, όπως ο **PET** και ο **TTC**, σε συστήματα αξιολόγησης της ασφάλειας μπορεί να ενισχύσει τη μετάβαση από την αντιδραστική αντιμετώπιση ατυχημάτων σε περισσότερο προληπτικές πολιτικές οδικής ασφάλειας.

6.4 Προτάσεις για Περαιτέρω Έρευνα

Η παρούσα ανάλυση θα μπορούσε να επεκταθεί σε **περισσότερες θέσεις μελέτης** στην Αθήνα, αλλά και σε άλλες ελληνικές πόλεις, ώστε να διερευνηθεί η επίδραση διαφορετικών γεωμετρικών, λειτουργικών και κυκλοφοριακών χαρακτηριστικών στη συμπεριφορά των οδηγών. Ιδιαίτερο ενδιαφέρον θα παρουσίαζε η εφαρμογή της μεθοδολογίας σε διαβάσεις του κέντρου της Αθήνας που λειτουργούν υπό διαφορετικές συνθήκες κυκλοφορίας και παρουσιάζουν αυξημένη πολυπλοκότητα, προκειμένου να αξιολογηθεί η **δυνατότητα γενίκευσης** των αποτελεσμάτων σε πιο απαιτητικά αστικά περιβάλλοντα.

Χρήσιμη θα ήταν επίσης η αξιολόγηση της συμπεριφοράς των οδηγών σε διαφορετικές χρονικές περιόδους και υπό διαφορετικές κυκλοφοριακές και **περιβαλλοντικές συνθήκες**, συμπεριλαμβανομένων νυχτερινών περιόδων, δυσμενών καιρικών συνθηκών και εναλλαγών στον φόρτο κυκλοφορίας. Παράλληλα, η διερεύνηση της επίδρασης της ημέρας και της ώρας στην πιθανότητα εκδήλωσης **speeding** ή κρίσιμων αλληλεπιδράσεων θα μπορούσε να ενισχύσει σημαντικά την ερμηνεία των σχετικών υποδειγμάτων.

Ιδιαίτερο ενδιαφέρον παρουσιάζει ακόμη η μελλοντική ενσωμάτωση δεδομένων που αφορούν τους πεζούς, ώστε να αναπτυχθούν κοινά υποδείγματα αλληλεπίδρασης **οχήματος–πεζού** και να διερευνηθεί συστηματικότερα η επικινδυνότητα για όλους τους χρήστες της οδού. Προς την ίδια κατεύθυνση, θα ήταν ιδιαίτερα χρήσιμη η ενσωμάτωση **δημογραφικών** ή άλλων συμπληρωματικών μεταβλητών, καθώς και η σύγκριση των αποτελεσμάτων με δεδομένα τηλεματικής ή άλλες εξωτερικές πηγές πληροφόρησης, ώστε να διαμορφωθούν πιο ολοκληρωμένες και κοινά αξιοποιήσιμες μετρικές.

Τέλος, θα μπορούσαν να εξεταστούν πιο προχωρημένες μεθοδολογικές προσεγγίσεις, όπως μη γραμμικά ή υβριδικά υποδείγματα **μηχανικής μάθησης**, με στόχο τη βελτίωση της πρόβλεψης συμπεριφορών speeding και κρίσιμων αλληλεπιδράσεων. Επιπλέον, η σύγκριση των παραγόμενων **surrogate safety indicators** με πραγματικά δεδομένα οδικών ατυχημάτων θα μπορούσε να συμβάλει ουσιαστικά στην προγνωστική αξιολόγηση του κινδύνου και στη μελλοντική αξιοποίηση των αποτελεσμάτων για τον σχεδιασμό στοχευμένων πολιτικών οδικής ασφάλειας.

Βιβλιογραφία

1. Abiamamela, O.-O. &. (2024). *Computer Vision for Intelligent Traffic Monitoring and Control*.
2. Ahmed S. Abdelrahman, M. A.-A. (2024). *Video-to-Text Pedestrian Monitoring (VTPM): Leveraging Computer Vision and Large Language Models for Privacy-Preserve Pedestrian Activity Monitoring at Intersections*.
3. Ahmed S. Abdelrahman, Z. I.-A. (2025). *VRUCrossSafe for crossing intention prediction of vulnerable road users for improving safe crossing at intersections. Communications Engineering*.
4. Αλβέρτη, Μ. (2025). *Διερεύνηση των παράνομων διελεύσεων των πεζών σε σηματοδοτούμενες διαβάσεις στο κέντρο της Αθήνας αξιοποιώντας τεχνική υπολογιστικής όρασης*. Διπλωματική Εργασία, Εθνικό Μετσόβιο Πολυτεχνείο.
5. Ahmed S. Abdelrahman, Z. I.-A. (2025). *VRUCrossSafe for crossing intention prediction of vulnerable road users for improving safe crossing at intersections. Communications Engineering*.
6. Αρβανιτάκη, Δ. (2022). *Διερεύνηση των παραμέτρων επιρροής της παραβατικότητας των πεζών*. Διπλωματική Εργασία, Εθνικό Μετσόβιο Πολυτεχνείο.
7. Aristomenis Kopsacheillis, I. P. (2024). *Exploring factors influencing pedestrian accident severity: a multi-source study in the city of Berlin. European Transport Research Review*.
8. Avignone, B. B. (2024). *A robust time-to-collision computation framework for pedestrian-autonomous shuttle interaction*.
9. Bonela, P. and Kadali, B. (2022). *A review of surrogate safety measures in road safety analysis*.
10. Diwas Thapa, S. M. (2023). *Assessing Driver Behavior in Work Zones: A Discretized Duration Approach to Predict Speeding. Work Zone Safety Analysis*.
11. Dilek E, D. M. (2023). *Computer Vision Applications in Intelligent Transportation Systems: A Survey. Sensors*.
12. Dylan Russon, A. G.-T. (2025). *Evaluating pedestrian crossing safety: Implementing and evaluating a convolutional neural network model trained on paired aerial and subjective perspective images. Heliyon*.
13. Erik Bochinski, V. E. (2017, August). *High-Speed Tracking-by-Detection Without Using Image Information. IWOT4S Challenge Winner*.
14. Ελληνική Στατιστική Αρχή - ΕΛΣΤΑΤ. (2025). *Οδικά Τροχαία Ατυχήματα: Έτος 2023*. Πειραιάς.
15. Ferko M, P. A. (2024). *Exploring Factors Influencing Speeding on Rural Roads: A Multivariable Approach. Infrastructures*.
16. Gagliardi, V. F. (2024, February). *Safety assessment of pedestrian-vehicle interaction at signalized intersections: An observational study. Journal of Transportation Safety & Security*.

17. Gaspar, F. V. (2020). *Prototype to Increase Crosswalk Safety by Integrating Computer Vision with ITS-G5 Technologies*. Information.
18. Guoqin Zang, S. A. (2021). *Quantitative study of vehicle-pedestrian interactions: Towards pedestrian-adapted lighting communication functions for autonomous vehicles*. *Proceedings of IEEE*.
19. Hosmer, D., Lemeshow, S. and Sturdivant, R. (2013). *Applied Logistic Regression* (3rd ed.). New York: Wiley.
20. Ismail, K., Sayed, T., Saunier, N. and Lim, C. (2013). *Automated analysis of pedestrian–vehicle conflicts using video data*.
21. Jamil Ahsan, M. A.-A. (2025, May). *Evaluating drivers' braking behavior at mid-block pedestrian crosswalks using video data and a mixed logit model with heterogeneity in the means*.
22. Joost de Winter, J. H. (2023). *Predicting perceived risk of traffic scenes using computer vision*. *Transportation Research Part F: Traffic Psychology and Behaviour*.
23. Γιαννής, Γ., & Μιχαλαράκη, Ε. (2024). *Review of City-Wide 30 km/h Speed Limit Benefits in Europe*. *Sustainability*, 16(11), 4382.
24. Γιαννής, Γ., Λαΐου, Α., & Ευγενικός, Π. (2013). *Διαχείριση ταχυτήτων*. Εθνικό Μετσόβιο Πολυτεχνείο.
25. Καλαϊτζή, Μ. (2016). *Αυτοματοποίηση Συλλογής Δεδομένων Πεζών με χρήση Τεχνικών Computer Vision*. Διπλωματική Εργασία, Εθνικό Μετσόβιο Πολυτεχνείο.
26. Κατσόγιαννος, Σ. και Καμβύσης, Μ. (2009). *Ανάλυση σε βάθος της στάσης Ελλήνων και άλλων Ευρωπαίων οδηγών απέναντι στην ταχύτητα*. Διπλωματική Εργασία, Εθνικό Μετσόβιο Πολυτεχνείο.
27. Kutela, B. (2021). *The role of crosswalk-related features on drivers' spatial yielding compliance at signalized midblock crosswalks*. *Journal of Traffic and Transportation Engineering (English Edition)*.
28. Marina Klanjic, L. G. (2022). *Identifying urban features for vulnerable road user safety in Europe*. *EPJ Data Science*.
29. Mejia, H. (2021). *Vehicle Speed Estimation Using Computer Vision And Evolutionary Camera Calibration*. *NeurIPS 2021*.
30. Mohamed Abdel-Aty, J. U. (2024). *Exploring the influence of drivers' visual surroundings on speeding behavior*. *Accident Analysis & Prevention*.
31. Monica Meocci, A. T. (2024, April). *Drivers' performance assessment approaching pedestrian crossings through the analysis of the speed and perceptive data recorded during on-field tests*.
32. Μυρογιάννη, Μ.-Ν. (2025). *Διερεύνηση μη συμμορφούμενων διελεύσεων πεζών σε σηματοδοτούμενες διασταυρώσεις αξιοποιώντας τεχνική υπολογιστικής όρασης*. Διπλωματική Εργασία, Εθνικό Μετσόβιο Πολυτεχνείο.

33. Narit Hnoohom, P. C. (2024, March). *The video-based safety methodology for pedestrian crosswalk safety measured: The case of Thammasat University, Thailand.*
34. Ν. 5209/2025. *Κώδικας Οδικής Κυκλοφορίας και άλλες διατάξεις.* ΦΕΚ 100/Α/13.06.2025.
35. Νικολάου, Δ. (2024). *Machine-learning based road crash risk assessment fusing infrastructure, traffic and driver behaviour data.* Διδακτορική Διατριβή, Εθνικό Μετσόβιο Πολυτεχνείο.
36. Παπαδημητρίου, Ε. (2010). *Πρότυπα συμπεριφοράς και ασφάλειας πεζών σε αστικά οδικά δίκτυα.* Διδακτορική Διατριβή, Εθνικό Μετσόβιο Πολυτεχνείο.
37. Rezaei, M. (2016). *Computer Vision for Road Safety: A System for Simultaneous Monitoring of Driver Behaviour and Road Hazards.*
38. Ren, Y. (2024, February). *Intelligent Vehicle Violation Detection System Under Human–Computer Interaction and Computer Vision.* *International Journal of Computational Intelligence Systems.*
39. Roussou, S., Πετράκη, Β., Δεληάλη, Κ., Κοντάξη, Α., & Γιαννής, Γ. (2024). *Cost benefit analysis of reducing speed limits in Athens to 30 Km/h.* *Case Studies on Transport Policy*, 18, 101289.
40. Roussou, S., Z. A. (2025). *Predicting pedestrian illegal crossing using vehicle dynamics with LightGBM.* *IRCOBI Conference 2025.*
41. Salvatore Cafiso, A. D. (2017, February). *In-vehicle stereo vision system for identification of traffic conflicts between bus and pedestrian.* *Journal of Traffic and Transportation Engineering (English Edition).*
42. Saunier, N. and Sayed, T. (2008). *Probabilistic framework for automated analysis of exposure to road collisions.*
43. Shailendra, A. S. (2025, February). *Computer Vision and Deep Learning based Approach for Traffic Violations due to Over-speeding and Wrong Direction Detection.* *International Journal of Computer Applications.*
44. St-Aubin, P. (2016). *Driver Behaviour and Road Safety Analysis Using Computer Vision and Applications in Roundabout Safety.*
45. Τάκκα, Μ. (2024). *Διερεύνηση επιρροής της υπηρεσίας αυτόνομων μικρών λεωφορείων στην ασφάλεια με προσομοίωση κυκλοφορίας δικτύου.* Διπλωματική Εργασία, Εθνικό Μετσόβιο Πολυτεχνείο.
46. Tengfeng Lin, Z. J. (2024, April). *A Real-time Evaluation Framework for Pedestrian’s Potential Risk at Non-Signalized Intersections Based on Predicted Post-Encroachment Time.*
47. Thiago R. P. M. Rúbio, J. A. (2020). *A Semi-automatic Object Identification Technique Combining Computer Vision and Deep Learning for the Crosswalk Detection Problem.* In *Intelligent Data Engineering and Automated Learning – IDEAL 2020.*

48. Ting Lian, B. P. (2024, January). *Advances in estimating pedestrian measures through artificial intelligence: From data sources, computer vision, video analytics to the prediction of crash frequency. Computers, Environment and Urban Systems.*
49. Τζουτζούλης, Δ. Μ. (2020). *Κρίσιμοι Παράγοντες Επιρροής της Υπέρβασης των Ορίων Ταχύτητας*. Διπλωματική Εργασία, Εθνικό Μετσόβιο Πολυτεχνείο.
50. Ventura, R., Roussou, S., Ziakopoulos, A., Barabino, B., & Yannis, G. (2025). *Using computer vision and street-level videos for pedestrian-vehicle tracking and behaviour analysis. Transportation Research Interdisciplinary Perspectives, 30,* 101366.
51. Wan, Y. X. (2023). *A Computer Vision-Based Algorithm for Detecting Vehicle Yielding to Pedestrians. Sustainability.*
52. Yadav, V. &. (2023). *A COMPUTER VISION FRAMEWORK FOR TRACKING & DETECT VEHICLES FROM ROAD SURVEILLANCE VIDEO.*
53. Yagnik M. Bhavsar, M. S. (2023, December). *Vision-based investigation of road traffic and violations at urban roundabout in India using UAV video: A case study. Transportation Engineering.*
54. Yongjie Wang, Q. S. (2021). *Investigating yielding behavior of heterogeneous vehicles at a semi-controlled crosswalk. Accident Analysis & Prevention.*
55. Yunxiang Liu. (2025). *VRU-YOLO: A Small Object Detection Algorithm for Vulnerable Road Users in Complex Scenes.*
56. Zhixiong Jin, W. N. (2024). *Do enhanced school zone policies improve pedestrians' safety? A deep learning-based case study of Osan City, South Korea. Cities.*
57. Zijun Du, M. D. (2023, October). *A review of road safety evaluation methods based on driving behavior.*
58. Φραντζεσκάκης, Ι., Γκόλιας, Ι. και Πιτσιάβα-Λατινοπούλου, Μ. (2009). *Κυκλοφοριακή Τεχνική*. Αθήνα: Παπασωτηρίου.