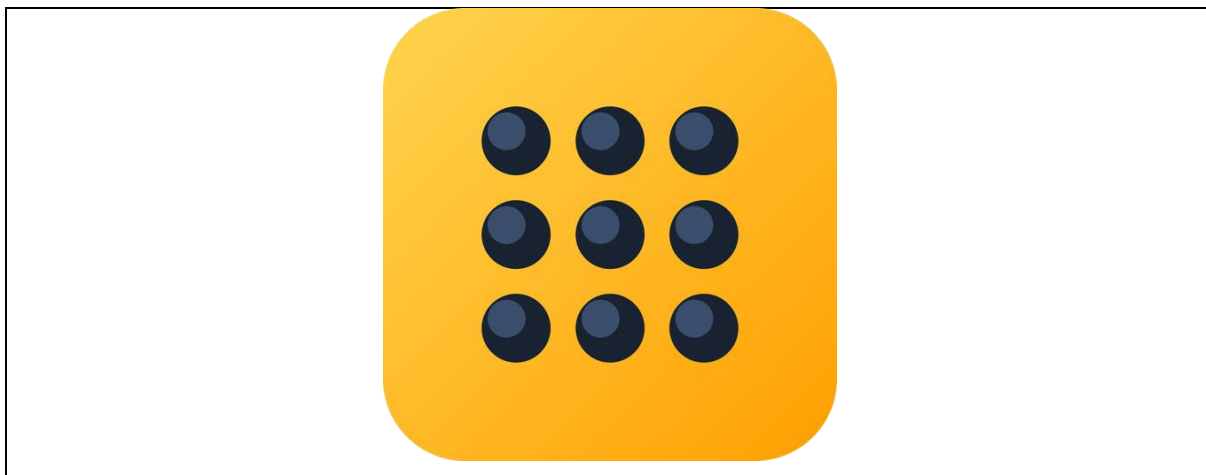


ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

«Επαυξημένη Πραγματικότητα για καθοδήγηση ατόμων
με προβλήματα όρασης»



Του φοιτητή Καρυπίδη
Θεόδωρου
Αρ. Μητρώου: 2019065.

Επιβλέπων
Ονοματεπώνυμο ΚΟΚΚΩΝΗΣ
ΓΕΩΡΓΙΟΣ
Βαθμίδα ΑΝΑΠΛΗΡΩΤΗΣ
ΚΑΘΗΓΗΤΗΣ

Ημερομηνία 30/8/2026

Τίτλος Δ.Ε. Επαυξημένη Πραγματικότητα για καθοδήγηση ατόμων με προβλήματα όρασης

Κωδικός Δ.Ε. 26119

Ονοματεπώνυμο φοιτητή Καρυπίδης Θεόδωρος

Ονοματεπώνυμο εισηγητή Γεώργιος Κοκκώνης

Ημερομηνία ανάληψης Δ.Ε. 20/02/2026

Ημερομηνία περάτωσης Δ.Ε. 31/8/2026

Βεβαιώνω ότι είμαι ο συγγραφέας αυτής της εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, έχω καταγράψει τις όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών, εικόνων και κειμένου, είτε αυτές αναφέρονται ακριβώς είτε παραφρασμένες. Επιπλέον, βεβαιώνω ότι αυτή η εργασία προετοιμάστηκε από εμένα προσωπικά, ειδικά ως διπλωματική εργασία, στο Τμήμα Μηχανικών Πληροφορικής και Ηλεκτρονικών Συστημάτων του ΔΙ.ΠΑ.Ε.

Η παρούσα εργασία αποτελεί πνευματική ιδιοκτησία του φοιτητή Καρυπίδη Θεόδωρου που την εκπόνησε. Στο πλαίσιο της πολιτικής ανοικτής πρόσβασης, ο συγγραφέας/δημιουργός εκχωρεί στο Διεθνές Πανεπιστήμιο της Ελλάδος άδεια χρήσης του δικαιώματος αναπαραγωγής, δανεισμού, παρουσίασης στο κοινό και ψηφιακής διάχυσης της εργασίας διεθνώς, σε ηλεκτρονική μορφή και σε οποιοδήποτε μέσο, για διδακτικούς και ερευνητικούς σκοπούς, άνευ ανταλλάγματος. Η ανοικτή πρόσβαση στο πλήρες κείμενο της εργασίας, δεν σημαίνει καθ' οιονδήποτε τρόπο παραχώρηση δικαιωμάτων διανοητικής ιδιοκτησίας του συγγραφέα/δημιουργού, ούτε επιτρέπει την αναπαραγωγή, αναδημοσίευση, αντιγραφή, πώληση, εμπορική χρήση, διανομή, έκδοση, μεταφόρτωση (downloading), ανάρτηση (uploading), μετάφραση, τροποποίηση με οποιονδήποτε τρόπο, τμηματικά ή περιληπτικά της εργασίας, χωρίς τη ρητή προηγούμενη έγγραφη συναίνεση του συγγραφέα/δημιουργού.

Η έγκριση της διπλωματικής εργασίας από το Τμήμα Μηχανικών Πληροφορικής και Ηλεκτρονικών Συστημάτων του Διεθνούς Πανεπιστημίου της Ελλάδος, δεν υποδηλώνει απαραίτητως και αποδοχή των απόψεων του συγγραφέα, εκ μέρους του Τμήματος.

«Στη Μαριάννα, στους γονείς μου Ελένη και Παναγιώτη, και στον αδελφό μου Αργύρη.

Και σε όσους περπατούν την πόλη με την αφή.»

Πρόλογος

Η επιλογή του συγκεκριμένου θέματος προέκυψε από το ενδιαφέρον μου για την εφαρμογή της υπολογιστικής όρασης και της τεχνητής νοημοσύνης σε προβλήματα με άμεσο κοινωνικό αντίκτυπο. Η διαπίστωση ότι ένα απλό κινητό τηλέφωνο, χωρίς κανέναν πρόσθετο εξοπλισμό, μπορεί να μετατραπεί σε βοήθημα αυτόνομης μετακίνησης για άτομα με οπτική αναπηρία, αποτέλεσε ισχυρό κίνητρο καθ' όλη τη διάρκεια εκπόνησης της εργασίας.

Θα ήθελα καταρχάς να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή μου, κ. Κοκκώνη, για την εμπιστοσύνη που μου έδειξε αναθέτοντάς μου το συγκεκριμένο θέμα, καθώς και για την καθοδήγηση, τις εύστοχες παρατηρήσεις και τη διαθεσιμότητά του σε κάθε στάδιο της εργασίας.

Ελπίζω η παρούσα εργασία, πέρα από την ακαδημαϊκή της αξία, να αποτελέσει ένα μικρό βήμα προς μια πιο προσβάσιμη καθημερινότητα για τα άτομα με οπτική αναπηρία, και να αποτελέσει αφετηρία για περαιτέρω έρευνα στο πεδίο.

Περίληψη

Η αυτόνομη μετακίνηση ατόμων με οπτική αναπηρία στον αστικό χώρο παραμένει δυσχερής, καθώς η κρίσιμη πληροφορία, η κατεύθυνση της ασφαλούς όδευσης και η κατάσταση των φωτεινών σηματοδοτών πεζών πραγματοποιείται σχεδόν αποκλειστικά με οπτικά μέσα. Τα απτικά πλακίδια όδευσης αποτελούν την υποδομή που καλείται να καλύψει το κενό αυτό, ωστόσο στην ελληνική πραγματικότητα χαρακτηρίζονται από ασυνέχεια και συχνή κατάληψη, ενώ οι υφιστάμενες εφαρμογές υποβοήθησης δεν αναγνωρίζουν ούτε τα πλακίδια ούτε την κατάσταση των σηματοδοτών. Η παρούσα διπλωματική εργασία παρουσιάζει το BlindNav, εφαρμογή Android η οποία αναγνωρίζει σε πραγματικό χρόνο απτικά πλακίδια όδευσης και φωτεινούς σηματοδότες πεζών, αξιοποιώντας αποκλειστικά την κάμερα και τους αισθητήρες ενός συμβατικού κινητού τηλεφώνου. Εκπαιδεύτηκαν δύο μοντέλα ανίχνευσης αντικειμένων YOLO11n, τα οποία μετατράπηκαν σε μορφή TFLite και εκτελούνται τοπικά στη συσκευή μέσω του LiteRT, χωρίς να απαιτείται η σύνδεση στο διαδίκτυο. Οι ανά καρέ προβλέψεις υφίστανται χρονική εξομάλυνση σε κυλιόμενο παράθυρο, αντιστοιχίζονται σε χωρικές ζώνες και ιεραρχούνται με συγκεκριμένα κριτήρια, πριν αποδοθούν στον χρήστη μέσω ελληνικής φωνητικής σύνθεσης και διαβαθμισμένης απτικής ανάδρασης. Η αξιολόγηση διεξήχθη σε δύο επίπεδα. Στο σύνολο δεδομένων επικύρωσης, το μοντέλο των σηματοδοτών πέτυχε $mAP@0.5 = 0,861$ και το μοντέλο των πλακιδίων $0,685$. Στις δοκιμές πεδίου, σε τρεις συσκευές και για συνολικά άνω των πενήντα λεπτών συνεχούς λειτουργίας, καταγράφηκαν χρόνοι εξαγωγής προβλέψεων 320–460 ms, ενώ ο μηχανισμός ανακοίνωσης μόνο σε μεταβολή απέτρεψε το 86% των δυνητικών εκφωνήσεων, με υπολογιστικό κόστος κάτω του ενός χιλιοστού του δευτερολέπτου. Η εργασία καταδεικνύει ότι η αξιοποίηση της υπάρχουσας υποδομής προσβασιμότητας μέσω λογισμικού είναι τεχνικά εφικτή χωρίς πρόσθετο εξοπλισμό και χωρίς παρέμβαση στον αστικό χώρο. Ταυτόχρονα, καταγράφονται ρητά οι αδυναμίες που εντοπίστηκαν υπενθυμίζοντας ότι το σύστημα θα πρέπει να λειτουργεί ως συμπληρωματικό βοήθημα και όχι ως υποκατάστατο των καθιερωμένων μέσων προσανατολισμού και κινητικότητας.

Λέξεις-κλειδιά: οπτική αναπηρία, απτικά πλακίδια όδευσης, φωτεινοί σηματοδότες πεζών, υπολογιστική όραση, YOLO11, ανίχνευση αντικειμένων, τοπική εκτέλεση, Android, προσβασιμότητα, βοηθητική τεχνολογία

«Augmented Reality for guiding visually impaired individuals»

«Karypidis Theodoros»

Abstract

Independent mobility for people with visual impairments in urban areas remains difficult, as critical information — the direction of the safe walking path and the state of pedestrian traffic lights — is conveyed almost exclusively through visual means. Tactile paving is the infrastructure intended to bridge this gap; in Greece, however, it is characterised by discontinuity and frequent obstruction, whilst existing assistive applications recognise neither the tiles nor the state of the traffic lights. This thesis presents BlindNav, an Android application that recognises tactile paving and pedestrian traffic lights in real time, utilising only the camera and sensors of a standard mobile phone. Two YOLO11n object detection models were trained; these were converted to TFLite format and run locally on the device via LiteRT, without the need for an internet connection. The frame-by-frame predictions undergo temporal smoothing in a sliding window, are mapped to spatial zones and prioritised according to specific criteria, before being presented to the user via Greek text-to-speech synthesis and graded haptic feedback. The evaluation was conducted at two levels. In the validation dataset, the traffic light model achieved a mAP@0.5 of 0.861 and the tactile paving model 0.685. In the field trials, across three devices and more than fifty minutes of continuous operation in total, inference times of 320–460 ms were recorded, whilst the announce-on-change mechanism suppressed 86% of potential utterances at a computational cost of less than one millisecond. The work demonstrates that exploiting the existing accessibility infrastructure through software is technically feasible, without additional equipment and without any intervention in the urban environment. At the same time, the weaknesses identified are documented explicitly, as a reminder that the system should function as a supplementary aid and not as a substitute for established orientation and mobility techniques.

Keywords: visual impairment, tactile walking surface indicators, pedestrian traffic lights, computer vision, YOLO11, object detection, on-device inference, Android, accessibility, assistive technology.

Ευχαριστίες

Σε αυτό το σημείο θα ήθελα να ευχαριστήσω ξεχωριστά τη σύντροφό μου, Μαριάννα, η οποία με στήριξε σε όλη τη διάρκεια των σπουδών μου και στάθηκε δίπλα μου τόσο στις δύσκολες όσο και στις ευχάριστες στιγμές. Ευχαριστώ επίσης τους γονείς μου, Ελένη και Παναγιώτη, χωρίς τους οποίους δεν θα μπορούσα να φτάσω εδώ που είμαι σήμερα· για την ατελείωτη συμπαράσταση και ενθάρρυνση σε όλα τα επίπεδα της ζωής μου, και για τον ακούραστο αγώνα που συνεχίζουν να δίνουν μέχρι και σήμερα. Τέλος, ευχαριστώ τον αδελφό μου, Αργύριο, ο οποίος μου δείχνει ότι υπάρχει πάντα λύση σε κάθε πρόβλημα, και μου θέτει καθημερινά νέους στόχους ώστε να μπορώ να ξεπερνά κάθε δυσκολία.

Περιεχόμενα

Περιεχόμενα

Πρόλογος.....	iv
Περίληψη.....	v
Abstract	vi
Ευχαριστίες	vii
Περιεχόμενα	viii
Κατάλογος Σχημάτων	xi
Κατάλογος Πινάκων.....	xii
Συντομογραφίες.....	xiii
Κεφάλαιο 1ο: Εισαγωγή	1
1.1 Πλαίσιο και κίνητρο.....	1
1.2 Ορισμός του προβλήματος.....	1
1.3 Διαμόρφωση κειμένου.....	1
1.4 Στόχοι και ερευνητικά ερωτήματα	2
1.5 Συνεισφορά.....	2
1.6 Δομή της εργασίας	3
1.7 Επίλογος.....	3
Κεφάλαιο 2ο: Υπόβαθρο και Σχετική Έρευνα.....	4
2.1 Οπτική αναπηρία και αυτόνομη μετακίνηση	4
2.1.1 Μέγεθος του προβλήματος.....	4
2.1.2 Περιορισμοί των παραδοσιακών βοηθημάτων	4
2.1.3 Ο ρόλος των απτικών πλακιδίων	4
2.2 Απτικά πλακίδια όδευσης.....	5
2.2.1 Ιστορική αναδρομή.....	5
2.2.2 Τύποι και σημασιολογία.....	5
2.2.3 Θεσμικό πλαίσιο.....	6
2.3 Φωτεινοί σηματοδότες πεζών.....	7
2.4 Υπολογιστική όραση και ανίχνευση αντικειμένων	7
2.4.1 Από τις κλασικές μεθόδους στη βαθιά μάθηση.....	7
2.4.2 Ανιχνευτές ενός και δύο σταδίων	8
2.4.3 Κατάτμηση έναντι ανίχνευσης	8
2.5 Η οικογένεια YOLO.....	8
2.6 Εκτέλεση στη συσκευή έναντι εκτέλεσης στο νέφος.....	9

2.7	Υπάρχουσες εφαρμογές και σχετική έρευνα.....	9
2.7.1	Εμπορικές εφαρμογές.....	9
2.7.2	Ακαδημαϊκή έρευνα: ανίχνευση απτικών πλακιδίων	11
2.7.3	Ακαδημαϊκή έρευνα: αναγνώριση σηματοδοτών πεζών	12
2.7.4	Το ελληνικό ερευνητικό πλαίσιο	13
2.8	Το κενό στη βιβλιογραφία και η συνεισφορά της εργασίας.....	13
2.8.1	Η προτεινόμενη προσέγγιση.....	14
2.9	Επίλογος.....	14
Κεφάλαιο 3ο:	Μεθοδολογία Υλοποίησης.....	16
3.1	Επισκόπηση της προσέγγισης	16
3.2	Συλλογή και προετοιμασία των δεδομένων	18
3.2.1	Αναγνώριση εντός εφαρμογής.....	19
3.3	Εκπαίδευση των μοντέλων	28
3.4	Εξαγωγή και μετατροπή των μοντέλων	29
3.5	Ανάπτυξη της εφαρμογής Android	30
3.6	Pipeline ανίχνευσης και επεξεργασίας	30
3.7	Συστήματα ανάδρασης	32
3.8	Ρυθμίσεις και σχεδιαστικές επιλογές προσβασιμότητας της εφαρμογής	36
3.9	Επίλογος.....	37
Κεφάλαιο 4ο:	Αρχιτεκτονική του Συστήματος και Πειραματικό Πρωτόκολλο.	38
4.1	Αρχιτεκτονική του συστήματος	38
4.2	Ροή της πληροφορίας	39
4.3	Μοντελοποίηση Δομής, Ροής και Συμπεριφοράς με UML.....	41
4.4	Σενάριο χρήσης	44
4.5	Πειραματικός σχεδιασμός	44
4.6	Επίλογος.....	46
Κεφάλαιο 5ο:	Οδηγός Χρήσης Λογισμικού	47
5.1	Εκκίνηση της εφαρμογής	47
5.2	Η διεπαφή.....	47
5.3	Λειτουργία περπατήματος.....	47
5.4	Λειτουργία σηματοδότη	48
5.5	Σύνοψη ενδείξεων	48
5.6	Συμπεριφορά σε μη ιδανικές συνθήκες.....	49
5.7	Πρακτικές υπενθυμίσεις.....	49
Κεφάλαιο 6ο:	Αξιολόγηση Χρηστών (UEQ-S)	50

6.1	Εισαγωγή.....	50
6.2	Μεθοδολογία και Δημογραφικά Στοιχεία	50
6.3	Αποτελέσματα Ερωτηματολογίου.....	52
6.3.1	Σύγκριση βάσει προηγούμενης εμπειρίας	53
6.4	Επίλογος.....	54
Κεφάλαιο 7ο:	Πειραματικά Αποτελέσματα και Συζήτηση.....	55
7.1.1	Μοντέλο ανίχνευσης απτικών πλακιδίων.....	55
7.1.2	Μοντέλο ανίχνευσης φωτεινών σηματοδοτών	58
7.1.3	Συγκριτική αποτίμηση.....	62
7.2	Απόδοση στη συσκευή.....	62
7.3	Συμπεριφορά σε συνθήκες πεδίου.....	64
7.3.1	Ορθότητα ανακοινώσεων	68
7.4	Συζήτηση.....	69
7.5	Τεχνικά ευρήματα της υλοποίησης.....	74
7.6	Περιορισμοί.....	74
7.7	Επίλογος.....	76
Κεφάλαιο 8ο:	Συμπεράσματα	77
8.1	Σύνοψη της εργασίας.....	77
8.2	Απάντηση στα ερευνητικά ερωτήματα.....	77
8.3	Όρια χρήσης του συστήματος	78
8.4	Μελλοντικές επεκτάσεις.....	78
8.5	Καταληκτικά σχόλια	79
8.6	Επίλογος.....	79
ΒΙΒΛΙΟΓΡΑΦΙΑ.....		80
ΠΑΡΑΡΤΗΜΑ Α :	Πηγαίος Κώδικας , Καταγραφή Μετρήσεων και Παράμετροι Εφαρμογής	84

Κατάλογος Σχημάτων

Εικόνα 3.1 Διάγραμμα ροής του pipeline (δεδομένα → εκπαίδευση → εξαγωγή → ενσωμάτωση) ...	17
Εικόνα 3.2 Πλακίδια Όδευσης με κλάση "tgs_"	19
Εικόνα 3.3 Πλακίδια προειδοποίησης με κλάση "tgs_warning"	20
Εικόνα 3.4 Πλακίδια κατεύθυνσης με κλάση "tgs_directional"	21
Εικόνα 3.5 Πλακίδια υπηρεσίας με κλάση "tgs_Service"	22
Εικόνα 3.6 Προειδοποίηση ερυθρού σηματοδότη με κλάση "2"	23
Εικόνα 3.7 Προειδοποίηση πράσινου σηματοδότη με κλάση "0"	24
Εικόνα 3.8 Tts μήνυμα για εμπόδια με κλάση "obstacle"	25
Εικόνα 3.9 Tts μήνυμα για διάβαση πεζών με κλάση "zebracrosswalk"	26
Εικόνα 3.10 Tts μήνυμα για το τέλος πεζοδρομίου με κλάση "sidewalk_end"	27
Εικόνα 3.11 Μοτίβο δόνησης αριστερής κατεύθυνσης (vibrateLeft)	32
Εικόνα 3.12 Μοτίβο δόνησης κεντρικής ευθυγράμμισης (vibrateCenter)	33
Εικόνα 3.13 Μοτίβο δόνησης δεξιάς κατεύθυνσης (vibrateRight)	33
Εικόνα 3.14 Μοτίβο δόνησης επιτυχούς ευθυγράμμισης (vibrateAligned)	33
Εικόνα 3.15 Μοτίβο δόνησης πλακιδίου προειδοποίησης (vibrateTileWarning)	34
Εικόνα 3.16 Μοτίβο δόνησης κρίσιμης προειδοποίησης (vibrateWarning)	34
Εικόνα 4.1 Διάγραμμα αρχιτεκτονικής των πέντε επιπέδων	39
Εικόνα 4.2 Διάγραμμα ροής δεδομένων για ένα καρέ, με τα σημεία απόφασης	41
Εικόνα 4.3 Διάγραμμα συνιστωσών	42
Εικόνα 4.4 Διάγραμμα ακολουθίας	43
Εικόνα 4.5 Διάγραμμα καταστάσεων	44
Εικόνα 6.1 Κατανομή ηλικίας συμμετεχόντων	50
Εικόνα 6.2 Κατανομή φύλου συμμετεχόντων: προς προσθήκη	51
Εικόνα 6.3 Προηγούμενη εμπειρία με βοηθητικές εφαρμογές: προς προσθήκη	51
Εικόνα 6.4 Μέση βαθμολογία ανά κλίμακα (Πραγματολογική / Ηδονική / Συνολική), με σύγκριση ως προς το όριο «Excellent» του UEQ-S benchmark	52
Εικόνα 6.5 Μέση βαθμολογία ανά ερώτηση	53
Εικόνα 6.6 Συνολική βαθμολογία UEQ-S: χωρίς έναντι με προηγούμενη εμπειρία	54
Εικόνα 7.1 Καμπύλη F1–εμπιστοσύνης του μοντέλου ανίχνευσης απτικών πλακιδίων (μέγιστο F1 = 0, 66 στην τιμή εμπιστοσύνης 0, 299)	56
Εικόνα 7.2 Πίνακας σύγχυσης του μοντέλου ανίχνευσης απτικών πλακιδίων (οκτώ κλάσεις, απόλυτος αριθμός ανιχνεύσεων)	57
Εικόνα 7.3 Κανονικοποιημένος πίνακας σύγχυσης του μοντέλου ανίχνευσης απτικών πλακιδίων	58
Εικόνα 7.4 Καμπύλη F1–εμπιστοσύνης του μοντέλου ανίχνευσης φωτεινών σηματοδοτών (μέγιστο F1 = 0, 86 στην τιμή εμπιστοσύνης 0, 481)	60
Εικόνα 7.5 Πίνακας σύγχυσης του μοντέλου ανίχνευσης φωτεινών σηματοδοτών (έξι κλάσεις, απόλυτος αριθμός ανιχνεύσεων)	61
Εικόνα 7.6 Κανονικοποιημένος πίνακας σύγχυσης του μοντέλου ανίχνευσης φωτεινών σηματοδοτών	62
Εικόνα 7.7 (α) πλακίδιο υπηρεσίας με ανακοίνωση «Προσοχή, πλακίδια προειδοποίησης»	71
Εικόνα 7.8 (β) γενικό πλακίδιο όδευσης με ανακοίνωση «Διάβαση πεζών»	72
Εικόνα 7.9 (γ) πλακίδιο κατεύθυνσης με ανακοίνωση «Προσοχή, πλακίδια προειδοποίησης»	73

Κατάλογος Πινάκων

Πίνακας 2.1 Σύγκριση υφιστάμενων εφαρμογών ως προς τα χαρακτηριστικά του BlindNav	10
Πίνακας 7.1 Σύνοψη ηχητικών και απτικών ενδείξεων	48
Πίνακας 5.1 Μέση ακρίβεια (AP@0.5) ανά κλάση, μοντέλο απτικών πλακιδίων	55
Πίνακας 5.2 Μέση ακρίβεια (AP@0.5) ανά κλάση, μοντέλο σηματοδοτών	59
Πίνακας 5.3 Χρόνοι εξαγωγής προβλέψεων ανά συσκευή	63
Πίνακας 5.4 Πραγματικός ρυθμός επεξεργασίας	63
Πίνακας 5.5 Κατανομή επικρατέστερης κλάσης ανά καρέ (Xiaomi, 448 καρέ).....	64
Πίνακας 5.6 Κατανομή επικρατέστερης κλάσης ανά καρέ (Samsung SM-F936B, 2.765 καρέ σε λειτουργία πλακιδίων).....	65
Πίνακας 5.7 Κατανομή επικρατέστερης κλάσης ανά καρέ (Samsung SM-A566B, 2.356 καρέ σε λειτουργία πλακιδίων).....	65
Πίνακας 5.8 Σύνθεση ανακοινώσεων (Xiaomi, 62 ανακοινώσεις)	66
Πίνακας 5.9 Σύνθεση ανακοινώσεων (Samsung SM-F936B, 2.912 καρέ σε 3 συνεδρίες, 279 ανακοινώσεις).....	67
Πίνακας 5.10 Σύνθεση ανακοινώσεων (Samsung SM-A566B, 2.530 καρέ σε 3 συνεδρίες, 259 ανακοινώσεις).....	67
Πίνακας 5.11 Καταγεγραμμένες περιπτώσεις ψευδώς θετικών, δύο γύροι δοκιμών (n=95)	68

Συντομογραφίες

Δ.Ε.	Διπλωματική Εργασία
ΔΙΠΑΕ	Διεθνές Πανεπιστήμιο Ελλάδος
Π.Ε.	Πτυχιακή Εργασία
ΕΛΟΤ	Ελληνικός Οργανισμός Τυποποίησης
NOK	Νέος Οικοδομικός Κανονισμός
ΟΠΕΚΑ	Οργανισμός Προνοιακών Επιδομάτων και Κοινωνικής Αλληλεγγύης
ΠΟΥ	Παγκόσμιος Οργανισμός Υγείας
ΤΕΕ	Τεχνικό Επιμελητήριο Ελλάδας
ΥΠΕΝ	Υπουργείο Περιβάλλοντος και Ενέργειας
Υ.ΠΕ.ΧΩ.Δ.Ε.	Υπουργείο Περιβάλλοντος, Χωροταξίας και Δημοσίων Έργων
ΦΕΚ	Φύλλο Εφημερίδας της Κυβερνήσεως

AI	Artificial Intelligence Τεχνητή Νοημοσύνη
AP	Average Precision Μέση Ακρίβεια
API	Application Programming Interface
CNN	Convolutional Neural Network Συνελικτικό Νευρωνικό Δίκτυο
COCO	Common Objects in Context
CPU	Central Processing Unit
CSV	Comma-Separated Values
FPS	Frames Per Second Καρέ ανά Δευτερόλεπτο
GPS	Global Positioning System
GPU	Graphics Processing Unit
HSV	Hue, Saturation, Value
IEEE	Institute of Electrical and Electronics Engineers
ISO	International Organization for Standardization
JIS	Japanese Industrial Standards
LiteRT	Lite Runtime
mAP	mean Average Precision
NCHW	Number, Channels, Height, Width
NHWC	Number, Height, Width, Channels

ONNX Open Neural Network Exchange
RGB Red, Green, Blue
SDK Software Development Kit
TFLite TensorFlow Lite
TTS Text-to-Speech Φωνητική Σύνθεση
UEQ User Experience Questionnaire
UEQ-S User Experience Questionnaire – Short
UML Unified Modeling Language
VR Virtual Reality Εικονική Πραγματικότητα
YOLO You Only Look Once

Κεφάλαιο 1ο: Εισαγωγή

1.1 Πλαίσιο και κίνητρο

Η δυνατότητα αυτόνομης μετακίνησης στον αστικό χώρο αποτελεί θεμελιώδη προϋπόθεση για την κοινωνική συμμετοχή και την ανεξαρτησία του ατόμου. Για τα άτομα με οπτική αναπηρία, η μετακίνηση στον αστικό χώρο συνεπάγεται με την καθημερινή αντιμετώπιση ενός περιβάλλοντος σχεδιασμένου κατά κύριο λόγο για βλέποντες, σε ένα τέτοιο περιβάλλον όπου η κρίσιμη πληροφορία· η κατεύθυνση της ασφαλούς όδευσης, η παρουσία εμποδίου, η κατάσταση ενός φωτεινού σηματοδότη, πραγματοποιείται σχεδόν αποκλειστικά με οπτικά μέσα, καθιστά αυτόματα την αυτόνομη κινητικότητα του ατόμου αδύνατη. Σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας, τουλάχιστον 2, 2 δισεκατομμύρια άνθρωποι παγκοσμίως παρουσιάζουν κάποια μορφή οπτικής αναπηρίας [1]. Το μέγεθος αυτό, σε συνδυασμό με τη γήρανση του πληθυσμού, καθιστά την προσβασιμότητα ενός δομημένου περιβάλλοντος ως ζήτημα διαρκώς αυξανόμενης σημασίας. Παράλληλα, η ραγδαία εξέλιξη της υπολογιστικής όρασης και η δυνατότητα εκτέλεσης μοντέλων βαθιάς μάθησης απευθείας σε κινητές συσκευές δημιουργούν νέες προοπτικές. Ένα σύγχρονο κινητό τηλέφωνο σήμερα διαθέτει κάμερα, αισθητήρες προσανατολισμού, μηχανή φωνητικής σύνθεσης και επαρκή υπολογιστική ισχύ ώστε να λειτουργήσει ως σημαντικό βοήθημα πλοήγησης χωρίς πρόσθετο εξοπλισμό και χωρίς εξάρτηση από απομακρυσμένη υποδομή. Η παρούσα εργασία στοχεύει στην αξιοποίηση ακριβώς αυτής της δυνατότητας.

1.2 Ορισμός του προβλήματος

Τα καθιερωμένα βοηθήματα κινητικότητας παρουσιάζουν σαφή λειτουργικά όρια. Το λευκό μαστούνι ανιχνεύει εμπόδια μόνο σε σχετικά μικρή απόσταση και κοντά στο έδαφος, ενώ ταυτόχρονα δεν παρέχει τον χρήστη με καμία πληροφορία κατεύθυνσης· ο σκύλος οδηγός καλύπτει μέρος αυτών των κενών, αλλά συνεπάγεται υψηλό κόστος και είναι διαθέσιμος σε περιορισμένη μερίδα του πληθυσμού. Κρίσιμο να σημειωθεί είναι ότι κανένα από τα δύο βοηθήματα δεν μπορεί να ενημερώσει τον χρήστη για την κατάσταση ενός φωτεινού σηματοδότη πεζών, δηλαδή ακριβώς τη στιγμή που η μετακίνηση ενέχει τον μεγαλύτερο κίνδυνο. Την ανάγκη για πληροφόρηση και ασφαλέστερη κατεύθυνση καλείται να καλύψει η υποδομή των απτικών πλακιδίων όδευσης, η οποία κωδικοποιεί απτικά δύο βασικά μηνύματα: την ασφαλή κατεύθυνση πορείας και την προειδοποίηση επικείμενου κινδύνου. Η αξιοποίησή της, ωστόσο, προϋποθέτει τη φυσική επαφή του χρήστη με το πλακίδιο και δεν παρέχει καμία πληροφορία για το τι βρίσκεται μπροστά του. Το πρόβλημα που πραγματεύεται η εργασία διατυπώνεται επομένως ως εξής: πώς μπορεί ένα συμβατικό κινητό τηλέφωνο να αναγνωρίζει σε πραγματικό χρόνο τα απτικά πλακίδια όδευσης και την κατάσταση των φωτεινών σηματοδοτών πεζών, και να αποδίδει την πληροφορία αυτή στον χρήστη με μη οπτικό τρόπο, χωρίς πρόσθετο εξοπλισμό, χωρίς τροποποίηση της αστικής υποδομής και χωρίς εξάρτηση από σύνδεση δικτύου.

1.3 Διαμόρφωση κειμένου

Η ανάγκη για μια τέτοια λύση τεκμηριώνεται από τρεις ανεξάρτητες κατευθύνσεις:

- **Η υποδομή είναι αποδεδειγμένα χρήσιμη.** Πειραματική μελέτη πεδίου με άτομα με οπτική αναπηρία έδειξε ότι η παρουσία απτικών πλακιδίων σε διάβαση περιορίζει την απόκλιση από την ευθεία πορεία, μειώνει τον χρόνο διάσχισης και βελτιώνει την κανονικότητα του βηματισμού [2].

- **Η υποδομή είναι όμως ελλιπής στην πράξη.** Μελέτη αξιολόγησης της προσβασιμότητας στο κέντρο της Αθήνας διαπίστωσε αισθητή απουσία ραμπών και απτικών πλακιδίων, σημαντικές ελλείψεις όπου αυτά υπάρχουν με χαρακτηριστικότερη την ασυνέχεια της όδευσης και σοβαρή παρεμπόδιση από την παράνομη στάθμευση [3].
- **Οι υπάρχουσες τεχνολογικές λύσεις δεν καλύπτουν το κενό.** Εκτενής έρευνα με ερωτηματολόγιο σε άτομα με οπτική αναπηρία κατέληξε ότι οι υφιστάμενες εφαρμογές πλοήγησης δεν παρέχουν επαρκή πληροφορία για φωτεινούς σηματοδότες, διασταυρώσεις και φυσικά εμπόδια· στην ίδια έρευνα, άνω του 63% των συμμετεχόντων ανέφεραν ότι είχαν τραυματιστεί τουλάχιστον μία φορά κατά τη μετακίνησή τους σε εξωτερικό χώρο [4].

Η σύνθεση των τριών αυτών διαπιστώσεων ορίζει τη σκοπιμότητα της εργασίας: υπάρχει εγκατεστημένη και χρήσιμη υποδομή, η οποία όμως είναι αναξιόπιστη στην πράξη, ενώ οι διαθέσιμες εφαρμογές δεν βοηθούν τον χρήστη να την αξιοποιήσει.

1.4 Στόχοι και ερευνητικά ερωτήματα

Στόχος της εργασίας είναι ο σχεδιασμός, η υλοποίηση και η αξιολόγηση μιας εφαρμογής Android η οποία, αξιοποιώντας την κάμερα καθώς και τους ενσωματωμένους αισθητήρες της συσκευής, να κατορθώσει συγκεκριμένα μέσα από δύο βασικές λειτουργίες να αναγνωρίζει απτικά πλακίδια όδευσης, συγκεκριμένα μέσω της λειτουργίας περπατήματος και φωτεινούς σηματοδότες πεζών μέσω της λειτουργίας φαναριού, καθώς στο τέλος να αποδίδει την πληροφορία μέσω ελληνικής φωνητικής και απτικής ανάδρασης μέσω συγκεκριμένων δονήσεων.

Κατά την μελέτη για την υλοποίηση της εργασίας προέκυψαν τα ακόλουθα ερευνητικά ερωτήματα:

1. Είναι εφικτή η αξιόπιστη αναγνώριση των τυποποιημένων τύπων απτικών πλακιδίων και των καταστάσεων φωτεινού σηματοδότη με ελαφριά μοντέλα ανίχνευσης αντικειμένων;
2. Επαρκεί η υπολογιστική ισχύς ενός συμβατικού κινητού τηλεφώνου για την εκτέλεση της αναγνώρισης σε πραγματικό χρόνο κατά τη βάδιση;
3. Με ποιον τρόπο μπορούν οι ασταθείς ανά καρέ προβλέψεις να μετατραπούν σε σταθερή και μη κουραστική καθοδήγηση για τον χρήστη;
4. Είναι δυνατή η εναλλαγή μεταξύ των δύο λειτουργιών χωρίς καμία οπτική αλληλεπίδραση με τη διεπαφή;

1.5 Συνεισφορά

Η εργασία συνεισφέρει στα εξής:

- **Ολοκληρωμένη λύση χωρίς εξωτερικό εξοπλισμό.** Σε αντίθεση με συστήματα που απαιτούν εξωτερικό δέκτη GPS, αισθητήρες υπερήχων ή συσκευές εγκατεστημένες επί των σηματοδοτών, το προτεινόμενο σύστημα λειτουργεί αποκλειστικά με τα μέσα ενός συμβατικού κινητού τηλεφώνου.
- **Αξιοποίηση, αντί παράκαμψης, της υπάρχουσας υποδομής.** Τα απτικά πλακίδια αναγνωρίζονται οπτικά και η σημασιολογία τους, όπως ορίζεται από το πρότυπο ISO 23599 [5], μεταφράζεται σε καθοδηγητική πληροφορία.
- **Ενοποίηση δύο στόχων σε μία εφαρμογή,** η αναγνώριση των πλακιδίων και των φωτεινών σηματοδοτών πραγματοποιείται με αυτόματη εναλλαγή λειτουργίας βάσει του προσανατολισμού της συσκευής, χωρίς απαίτηση χειρισμού της οθόνης.
- **Πλήρως τοπική εκτέλεση** με ελληνική φωνητική ανάδραση, εξασφαλίζοντας λειτουργία εκτός δικτύου και προστασία των προσωπικών δεδομένων του χρήστη.
- **Τεκμηριωμένη πειραματική αξιολόγηση,** τόσο σε επίπεδο μοντέλων όσο και σε πραγματικές συνθήκες πεδίου, με ρητή καταγραφή των αδυναμιών και των ορίων του συστήματος.

1.6 Δομή της εργασίας

Το Κεφάλαιο 2 παρουσιάζει το υπόβαθρο και τη σχετική έρευνα: την οπτική αναπηρία και τα όρια των παραδοσιακών βοηθημάτων, την υποδομή των απτικών πλακιδίων και το θεσμικό της πλαίσιο, τις τεχνολογίες υπολογιστικής όρασης, καθώς και τις υπάρχουσες εφαρμογές και ερευνητικές εργασίες, καταλήγοντας στον προσδιορισμό του ερευνητικού κενού.

Το Κεφάλαιο 3 περιγράφει τη μεθοδολογία υλοποίησης: τη συλλογή και επισημείωση των δεδομένων, την εκπαίδευση των δύο μοντέλων YOLO11n, τη μετατροπή τους σε μορφή κατάλληλη για κινητή συσκευή και την ενσωμάτωσή τους στην εφαρμογή Android, μαζί με τα συστήματα επεξεργασίας και ανάδρασης.

Το Κεφάλαιο 4 παρουσιάζει την αρχιτεκτονική του συστήματος, τη ροή της πληροφορίας κατά την εκτέλεση, τα διαγράμματα UML και το πειραματικό πρωτόκολλο αξιολόγησης.

Το Κεφάλαιο 5 αποτελεί τον οδηγό χρήσης του λογισμικού: την εκκίνηση και τη διεπαφή της εφαρμογής, τις δύο λειτουργίες της, τη σύνοψη των ηχητικών και απτικών ενδείξεων, καθώς και τη συμπεριφορά του συστήματος σε μη ιδανικές συνθήκες.

Το Κεφάλαιο 6 παρουσιάζει την αξιολόγηση της υποκειμενικής εμπειρίας χρήσης μέσω του σταθμισμένου ερωτηματολογίου UEQ-S: τη μεθοδολογία και τα δημογραφικά στοιχεία των συμμετεχόντων, τις βαθμολογίες στην πραγματολογική και την ηδονική κλίμακα και τη σύγκριση βάσει προηγούμενης εμπειρίας.

Το Κεφάλαιο 7 παραθέτει τα πειραματικά αποτελέσματα: την απόδοση των μοντέλων στο σύνολο επικύρωσης, τις μετρήσεις απόδοσης στη συσκευή, τη συμπεριφορά σε συνθήκες πεδίου και την ορθότητα των ανακοινώσεων, καθώς και τη συζήτηση των ευρημάτων και των περιορισμών.

Το Κεφάλαιο 8 συνοψίζει τη συνεισφορά της εργασίας, απαντά στα ερευνητικά ερωτήματα, οριοθετεί τη χρήση του συστήματος και προτείνει κατευθύνσεις για μελλοντικές επεκτάσεις.

1.7 Επίλογος

Το κεφάλαιο αυτό τεκμηρίωσε την ανάγκη για ένα σύστημα σαν το BlindNav, ξεκινώντας από το ευρύτερο πλαίσιο της αυτόνομης μετακίνησης ατόμων με οπτική αναπηρία και καταλήγοντας σε ένα σαφώς διατυπωμένο ερευνητικό πρόβλημα. Παρουσιάστηκαν τρεις ανεξάρτητες γραμμές τεκμηρίωσης η αποδεδειγμένη χρησιμότητα των απτικών πλακιδίων, η ελλιπής υλοποίησή τους στην πράξη, και η αδυναμία των υφιστάμενων τεχνολογικών λύσεων να καλύψουν το κενό οι οποίες από κοινού θεμελιώνουν τη σκοπιμότητα της εργασίας. Ορίστηκαν ο στόχος, τα τέσσερα ερευνητικά ερωτήματα και οι πέντε άξονες συνεισφοράς της εργασίας. Το επόμενο κεφάλαιο εξετάζει αναλυτικά το θεωρητικό και τεχνολογικό υπόβαθρο στο οποίο εντάσσεται η εργασία, καθώς και τη σχετική έρευνα που προσδιορίζει το συγκεκριμένο ερευνητικό κενό.

Κεφάλαιο 2ο: Υπόβαθρο και Σχετική Έρευνα

2.1 Οπτική αναπηρία και αυτόνομη μετακίνηση

2.1.1 Μέγεθος του προβλήματος

Σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας, τουλάχιστον 2, 2 δισεκατομμύρια άνθρωποι παγκοσμίως παρουσιάζουν κάποια μορφή οπτικής αναπηρίας, είτε αυτή αφορά την κοντινή είτε τη μακρινή όραση. Από αυτούς, τουλάχιστον ένα δισεκατομμύριο έχουν αναπηρία η οποία είτε θα μπορούσε να είχε προληφθεί είτε δεν έχει ακόμη αντιμετωπιστεί [1]. Ο ίδιος οργανισμός επισημαίνει ότι η αύξηση και η γήρανση του πληθυσμού αναμένεται να διευρύνουν περαιτέρω τον αριθμό αυτόν, ενώ αποδίδει στην οπτική αναπηρία ετήσιες απώλειες παραγωγικότητας της τάξης των 411 δισεκατομμυρίων δολαρίων. Για τον ελληνικό χώρο, τα διαθέσιμα στοιχεία είναι σαφώς πιο περιορισμένα. Η πιο πρόσφατη αξιόπιστη καταγραφή είναι διοικητικής φύσεως: κατά το α' τρίμηνο του 2021 καταγράφηκαν 16.309 άτομα με προβλήματα όρασης ως δικαιούχοι της οικονομικής ενίσχυσης όρασης του ΟΠΕΚΑ, ποσοστό που αντιστοιχεί στο 0, 152% του πληθυσμού [6]. Το μέγεθος αυτό, ωστόσο, πρέπει να ερμηνεύεται ως κατώτατο όριο και όχι ως εκτίμηση, καθώς περιλαμβάνει αποκλειστικά όσους έχουν αναγνωρισμένο ποσοστό αναπηρίας άνω του 80%. Χαρακτηριστικό της κατανομής είναι ότι το 76% των καταγεγραμμένων ατόμων ανήκουν σε ηλικίες άνω των 55 ετών, εύρημα που συμφωνεί με τη διαπίστωση του ΠΟΥ για τη συσχέτιση της οπτικής αναπηρίας με τη γήρανση. Επισημαίνεται ότι δεν εντοπίστηκε πρόσφατη επίσημη επιδημιολογική μελέτη επιπολασμού για την Ελλάδα, γεγονός που αποτελεί από μόνο του ένδειξη της υποκαταγραφής του ζητήματος.

2.1.2 Περιορισμοί των παραδοσιακών βοηθημάτων

Το λευκό μπαστούνι παραμένει το πλέον διαδεδομένο βοήθημα κινητικότητας, η εμβέλειά του όμως είναι εξ ορισμού περιορισμένη. Επιτρέπει τον εντοπισμό εμποδίων μόνο σε απόσταση περίπου ενός μέτρου και σε ύψος κοντά στο έδαφος, ενώ αδυνατεί να ανιχνεύσει εμπόδια στο ύψος του κορμού ή του κεφαλιού, όπως προεξέχοντα κλαδιά ή πινακίδες. Επιπλέον, δεν παρέχει καμία πληροφορία κατεύθυνσης ή πλοήγησης, παρά μόνο ανίχνευση της άμεσης παρουσίας εμποδίου. Ο σκύλος οδηγός καλύπτει μέρος αυτών των κενών, αλλά συνεπάγεται υψηλό κόστος, εκτεταμένη εκπαίδευση και διαβίου φροντίδα, με αποτέλεσμα να είναι διαθέσιμος σε μικρή μόνο μερίδα του πληθυσμού. Κρίσιμο είναι επίσης ότι κανένα από τα δύο δεν παρέχει πληροφορία για την κατάσταση ενός φωτεινού σηματοδότη πεζών. Η διάσχιση μιας διάβασης, δηλαδή η πιο επικίνδυνη στιγμή μιας αστικής διαδρομής, παραμένει έτσι ένα σημείο όπου ο χρήστης στηρίζεται σε έμμεσες ενδείξεις, όπως ο ήχος της κυκλοφορίας, ή στη βοήθεια τρίτων.

2.1.3 Ο ρόλος των απτικών πλακιδίων

Τα απτικά πλακίδια όδευσης αποτελούν την υποδομή που έχει σχεδιαστεί ακριβώς για να καλύψει αυτό το κενό καθοδήγησης. Η αποτελεσματικότητά τους δεν είναι θεωρητική: πειραματική μελέτη πεδίου με άτομα με οπτική αναπηρία, με χρήση μετρήσεων από επιταχυνσιόμετρα και εναέρια καταγραφή, έδειξε ότι η παρουσία απτικών πλακιδίων σε διάβαση περιορίζει την απόκλιση από την ευθεία πορεία, μειώνει τον χρόνο διάσχισης και βελτιώνει την κανονικότητα του βηματισμού, με το όφελος να είναι μεγαλύτερο για τους τυφλούς χρήστες σε σχέση με τα άτομα μειωμένης όρασης [2]. Η αξία της υποδομής αυτής, ωστόσο, εξαρτάται από την ποιότητα της υλοποίησής της. Μελέτη αξιολόγησης της προσβασιμότητας στο κέντρο της Αθήνας διαπίστωσε αισθητή απουσία ραμπών και απτικών

πλακιδίων, καθώς και σημαντικές ελλείψεις στα σημεία όπου υπάρχουν, με χαρακτηριστικότερη την ασυνέχεια της όδευσης· παράλληλα, η παράνομη στάθμευση αναδείχθηκε ως σοβαρό εμπόδιο στην κίνηση των πεζών [3]. Ανάλογα ευρήματα προκύπτουν και από αξιολογήσεις σε μικρότερες ελληνικές πόλεις. Η διεθνής βιβλιογραφία επιβεβαιώνει ότι τα δύο συχνότερα προβλήματα που αναφέρουν οι ίδιοι οι χρήστες είναι ακριβώς η ασυνέχεια της όδευσης και η κατάληψη των πλακιδίων από σταθμευμένα οχήματα.

Η διαπίστωση αυτή είναι καθοριστική για τη σκοπιμότητα της παρούσας εργασίας. Δεν αρκεί η υποδομή να υπάρχει· ο χρήστης χρειάζεται να γνωρίζει, σε πραγματικό χρόνο, αν η όδευση συνεχίζεται, τι τύπου πλακίδιο βρίσκεται μπροστά του και πού ακριβώς σε σχέση με τη θέση του. Ακριβώς αυτή η ανάγκη ορίζει τον σχεδιαστικό στόχο του BlindNav.

2.2 Απτικά πλακίδια όδευσης

2.2.1 Ιστορική αναδρομή

Τα απτικά πλακίδια όδευσης οφείλονται στον Ιάπωνα εφευρέτη Seiichi Miyake (1926–1982), ο οποίος τα ανέπτυξε μαζί με τον αδελφό του Saburo. Αφορμή για την ανάπτυξή τους αποτέλεσε ο στενός φίλος του Miyake, Hideyuki Iwahashi, ο οποίος έχανε σταδιακά την όρασή του. Ο Iwahashi είχε παρατηρήσει ότι μπορούσε να διακρίνει τη διαφορά ανάμεσα στο χρώμα και τη βλάστηση μέσα από τα παπούτσια του, παρατήρηση που συνέβαλε στην ιδέα ενός μοτίβου το οποίο θα μπορούσε να γίνεται αντιληπτό μέσω της αφής του πέλματος[7]. Η χρονολόγηση των πρώτων απτικών πλακιδίων παρουσιάζει σε ορισμένες πηγές φαινομενικές αποκλίσεις, οι οποίες μπορούν να αποδοθούν στη διάκριση μεταξύ της αρχικής τοποθέτησης και της επίσημης υιοθέτησής τους. Τον Μάρτιο του 1965 τοποθετήθηκαν 230 πλακίδια πλησίον της διάβασης της Σχολής Τυφλών του Okayama, με τη χρηματοδότηση να καλύπτεται εξ ολοκλήρου από τον ίδιο τον Miyake. Η αρχική αυτή προσπάθεια αντιμετώπισε επιφυλακτικότητα από τον τοπικό σύλλογο τυφλών. Στις 18 Μαρτίου 1967, δύο χρόνια αργότερα, η πόλη του Okayama έγινε ο πρώτος δήμος που υιοθέτησε επίσημα την εφεύρεση σε δημόσιο χώρο. Τα πλακίδια ήταν αρχικά γνωστά ως Tenji blocks και είχαν γκρι χρώμα. Το χαρακτηριστικό έντονο κίτρινο χρώμα υιοθετήθηκε μεταγενέστερα, με σκοπό να είναι περισσότερο ευδιάκριτα και από άτομα με μειωμένη όραση, χαρακτηριστικό που, όπως αναφέρεται στην Ενότητα 2.2.2, παρουσιάζει ενδιαφέρον και για την αυτόματη οπτική αναγνώρισή τους.

Η χρήση των πλακιδίων εξαπλώθηκε σταδιακά στην Ιαπωνία και στη συνέχεια σε άλλες χώρες. Το 1975 κατέστησαν υποχρεωτικά στους ιαπωνικούς σιδηροδρόμους, ενώ έως το 1985 είχαν επεκταθεί και στις ιαπωνικές πόλεις. Η ευρεία διάδοση του συστήματος οδήγησε, ωστόσο, στην ανάπτυξη περισσότερων από πενήντα διαφορετικών μοτίβων, γεγονός που δημιούργησε σύγχυση στους ίδιους τους χρήστες και ανέδειξε την ανάγκη για τυποποίηση. Σε εθνικό επίπεδο, το ιαπωνικό πρότυπο JIS θεσπίστηκε τον Σεπτέμβριο του 2001, ενώ σε διεθνές επίπεδο το πρότυπο ISO 23599 εκδόθηκε αρχικά το 2012 και αναθεωρήθηκε το 2019.

2.2.2 Τύποι και σημασιολογία

Το καθοριστικό στοιχείο της σύλληψης του Miyake ήταν η ύπαρξη δύο διακριτών μοτίβων με διαφορετική σημασία, τα οποία ο χρήστης μπορεί να διακρίνει μέσω του μαστουριού, του υποδήματος ή του πέλματος. Η διάκριση αυτή διατηρείται αυτούσια στη σημερινή τυποποίηση: το πρότυπο ISO

23599 ορίζει δύο τύπους απτικών δεικτών επιφάνειας βάδισης, τα μοτίβα προσοχής και τα μοτίβα καθοδήγησης [5].

Τα μοτίβα προσοχής ή προειδοποίησης αποτελούνται ανάγλυφους κόκκους ή κολουροκωνικές* προεξοχές και χρησιμοποιούνται για να υποδηλώσουν επικείμενο κίνδυνο ή απαίτηση στάσης, την αρχή μιας διάβασης, το άκρο μιας αποβάθρας, την αρχή μιας σκάλας. Τα μοτίβα καθοδήγησης φέρουν παράλληλες ραβδώσεις με επίπεδη κορυφή και υποδεικνύουν ασφαλή κατεύθυνση πορείας. Η αντίθεση των δύο είναι εσκεμμένη και λειτουργική· το ένα σημαίνει σταμάτα και πρόσεξε, το άλλο προχώρα προς αυτή την κατεύθυνση.

Επίσης πέρανα της απτικής διάστασης, τα πλακίδια σχεδιάζονται συνήθως σε χρώματα υψηλής αντίθεσης ως προς το περιβάλλον δάπεδο, ώστε να είναι αντιληπτά και από τα άτομα με μειωμένη όραση. Το χαρακτηριστικό αυτό είναι σημαντικό και για την εργασία, καθώς η χρωματική αντίθεση μπορεί να αποτελέσει χρήσιμο οπτικό χαρακτηριστικό για την αυτόματη ανίχνευση των πλακιδίων.

Η συγκεκριμένη διάκριση χρησιμοποιείται και στον σχεδιασμό του BlindNav. Οι κλάσεις `tgs_warning` και `tgs_directional` του μοντέλου ανίχνευσης αντιστοιχούν στους δύο βασικούς τύπους απτικών πλακιδίων. Παράλληλα, η ιεράρχηση που παρουσιάζεται στο Κεφάλαιο 3 δίνει μεγαλύτερη προτεραιότητα στα πλακίδια προειδοποίησης, σύμφωνα με τη σημασία που έχουν για την ασφαλή μετακίνηση του χρήστη.

2.2.3 Θεσμικό πλαίσιο

Σε ευρωπαϊκό επίπεδο, το πρότυπο EN 17210:2021 καθορίζει λειτουργικές απαιτήσεις που σχετίζονται με την προσβασιμότητα και τη χρηστικότητα του δομημένου περιβάλλοντος. Το συγκεκριμένο πρότυπο δεν αντικαθιστά, ωστόσο, τις διαστασιολογικές προδιαγραφές που καθορίζονται από το ISO 23599 [5]. Στην Ελλάδα, οι τεχνικές προδιαγραφές των οδύσεων ατόμων με οπτική αναπηρία δεν προκύπτουν από αριθμημένο πρότυπο ΕΛΟΤ αντίστοιχο του ISO 23599, αλλά από τις Οδηγίες Σχεδιασμού «Σχεδιάζοντας για Όλους» [8]. Οι οδηγίες αυτές, που εκδόθηκαν αρχικά από το τότε Υ.ΠΕ.ΧΩ.Δ.Ε., επικαιροποιήθηκαν με την υπ' αριθμ. ΥΠΕΝ/ΔΜΕΑΠ/99709/796/22.10.2021 απόφαση του Υφυπουργού Περιβάλλοντος και Ενέργειας «Τεχνικές Προδιαγραφές Μελέτης Προσβασιμότητας» (Β' 5045) [9], η οποία αποτελεί το ισχύον τεχνικό κείμενο αναφοράς. Οι οδηγίες διακρίνουν τον οδηγό όδευσης με ραβδώσεις (Τύπος Α) από τη λωρίδα προειδοποίησης κινδύνου με ανάγλυφους κόκκους (Τύπος Β), σε πλήρη αντιστοιχία με τη διεθνή διάκριση του ISO 23599. Ως προς τον δεσμευτικό χαρακτήρα των προδιαγραφών αυτών, το θεσμικό τοπίο μεταβλήθηκε πρόσφατα. Η υποχρέωση διαμόρφωσης προσβάσιμων οδύσεων προβλεπόταν αρχικά από το Άρθρο 26 του Νέου Οικοδομικού Κανονισμού (ν. 4067/2012) [10]. Με τον ν. 5306/2026 (ΦΕΚ Α' 88/8.6.2026), ωστόσο, κυρώθηκε ο Κώδικας Χωροταξίας – Πολεοδομίας «Νικόλαος Ταγαράς», ο οποίος κωδικοποίησε σε ενιαίο κείμενο τον ΝΟΚ μαζί με δεκάδες άλλα νομοθετήματα. Η σχετική διάταξη περιλαμβάνεται πλέον στο Άρθρο 223 «Προσβασιμότητα χωρίς φραγμούς για τα άτομα με αναπηρία και τα εμποδιζόμενα άτομα» του Κώδικα [11]. Επισημαίνεται ότι η κωδικοποίηση αναδιατάσσει και ενοποιεί το ισχύον δίκαιο χωρίς να εισάγει νέους ουσιαστικούς κανόνες· η αλλαγή αφορά επομένως την ορθή παραπομπή και όχι το περιεχόμενο των υποχρεώσεων. Ιδιαίτερη σημασία για την παρούσα εργασία έχει η παρ. 6 του Αρθρου 223, η οποία δεν αφορά τα κτίρια αλλά τους κοινόχρηστους υπαίθριους χώρους δηλαδή ακριβώς το πεδίο εφαρμογής του BlindNav. Η διάταξη προβλέπει ρητά ότι κατά τη διαμόρφωση ή ανακατασκευή πλατειών, πεζοδρόμων, πεζοδρομίων, νησίδων και στάσεων επιβάλλεται να εξασφαλίζεται η πρόσβαση ατόμων με αναπηρία μέσω, μεταξύ άλλων, οδύσεων τυφλών, εφόσον το επιτρέπει η μορφολογία του εδάφους. Η υποδομή που καλείται να αναγνωρίσει η εφαρμογή είναι

επομένως θεσμικά κατοχυρωμένη και για τον υπαίθριο αστικό χώρο, όχι μόνο για το εσωτερικό των κτιρίων. Ο Κώδικας ενσωματώνει επίσης μηχανισμούς παρακολούθησης και συμμόρφωσης που δεν υπήρχαν στην αρχική διατύπωση του 2012: ηλεκτρονική πλατφόρμα καταγραφής στοιχείων προσβασιμότητας υπό τη διαχείριση του Τεχνικού Επιμελητηρίου Ελλάδας, Μητρώο Ελεγκτών Προσβασιμότητας με έλεγχο κατόπιν αυτοψίας, καθώς και τη θέσπιση «Εθνικού Σήματος Προσβασιμότητας». Παρά το ενισχυμένο αυτό πλαίσιο, η απόσταση ανάμεσα στη νομοθετική πρόβλεψη και την πραγματική κατάσταση του αστικού χώρου παραμένει, όπως τεκμηριώνεται από τις μελέτες αξιολόγησης προσβασιμότητας που παρατέθηκαν στην Ενότητα 2.1.3 διαπίστωση που ενισχύει, αντί να αναιρεί, τη σκοπιμότητα ενός συστήματος που βοηθά τον χρήστη να αξιοποιήσει την υποδομή όπου αυτή πράγματι υπάρχει.

2.3 Φωτεινοί σηματοδότες πεζών

Ο φωτεινός σηματοδότης πεζών αποτελεί ένα αποκλειστικό οπτικό μέσο σηματοδότησης, καθώς η πληροφορία μεταφέρετε κυρίως μέσω χρωμάτων και σύμβολων· στοιχεία απροσπέλαστα για τον τυφλό χρήστη. Στην τυπική του μορφή, ο σηματοδότης εναλλάσσεται μεταξύ κόκκινης ένδειξης, που απαγορεύει τη διάσχιση, και πράσινης, που την επιτρέπει. Σε πολλές σύγχρονες εγκαταστάσεις προστίθεται αριθμητική αντίστροφη μέτρηση, η οποία εμφανίζεται είτε κατά την πράσινη φάση, υποδεικνύοντας τον εναπομείναντα χρόνο διάσχισης, είτε κατά την κόκκινη, υποδεικνύοντας τον χρόνο αναμονής. Η διάκριση αυτή έχει πρακτική σημασία για την πλοήγηση, καθώς η ίδια ένδειξη αντίστροφης μέτρησης φέρει εντελώς διαφορετικό νόημα ανάλογα με τη φάση στην οποία εμφανίζεται. Για τον λόγο αυτό, το μοντέλο ανίχνευσης σηματοδοτών της παρούσας εργασίας διακρίνει τις καταστάσεις *countdown_green* και *countdown_red* ως ξεχωριστές κλάσεις.

Μία από τις συνηθέστερες λύσεις για την προσβασιμότητα των φωτεινών σηματοδοτών είναι η χρήση ηχητικών σημάτων κατά την πράσινη φάση. Η κάλυψή τους, ωστόσο, είναι άνιση: πολλές διαβάσεις παραμένουν χωρίς ηχητική σηματοδότηση, ενώ όσες διαθέτουν ενδέχεται να έχουν απενεργοποιημένο ή δυσλειτουργικό μηχανισμό. Επιπλέον, σε περιβάλλον υψηλού κυκλοφοριακού θορύβου η ακουστική ένδειξη μπορεί να καλυφθεί. Η αδυναμία αυτή έχει οδηγήσει σε σημαντικό όγκο ερευνητικών εργασιών που επιχειρούν την αναγνώριση της κατάστασης του σηματοδότη με οπτικά μέσα, οι οποίες παρουσιάζονται στην Ενότητα 2.7.

2.4 Υπολογιστική όραση και ανίχνευση αντικειμένων

2.4.1 Από τις κλασικές μεθόδους στη βαθιά μάθηση

Οι πρώτες προσεγγίσεις για την αναγνώριση απτικών πλακιδίων και σηματοδοτών βασίζονταν σε τεχνικές κλασικής επεξεργασίας εικόνας: κατάτμηση με βάση το χρώμα, ανίχνευση ακμών και μορφολογικοί τελεστές. Χαρακτηριστικό παράδειγμα αποτελεί η εργασία των Ito κ.ά. [12], οι οποίοι εντοπίζουν απτικά πλακίδια μέσω δυναμικής κατωφλίωσης στον χρωματικό χώρο HSV, αξιοποιώντας το χαρακτηριστικό κίτρινο χρώμα τους. Οι μέθοδοι αυτές έχουν το πλεονέκτημα του χαμηλού υπολογιστικού κόστους και της ερμηνευσιμότητας, αλλά και σοβαρό μειονέκτημα: στηρίζονται σε χαρακτηριστικά που σχεδιάζονται χειροκίνητα και συνεπώς είναι ευαίσθητες στις συνθήκες. Η αλλαγή του φωτισμού, οι σκιές, η φθορά ή η ρύπανση των πλακιδίων, αλλά και η παρουσία άλλων αντικειμένων με παρόμοιο χρώμα, μπορούν να επηρεάσουν σημαντικά την απόδοση. Σε ένα σύστημα ασφάλειας που καλείται να λειτουργήσει σε απρόβλεπτο αστικό περιβάλλον, η ευαισθησία αυτή είναι απαγορευτική.

Η βαθιά μάθηση ακολουθεί διαφορετική προσέγγιση. Αντί ο μηχανικός να ορίσει ποια χαρακτηριστικά είναι σημαντικά, το συνελκτικό νευρωνικό δίκτυο τα μαθαίνει το ίδιο από τα δεδομένα. Το αποτέλεσμα είναι σαφώς μεγαλύτερη ανθεκτικότητα στη μεταβλητότητα των πραγματικών συνθηκών, με αντίτιμο την ανάγκη για επισημειωμένα δεδομένα εκπαίδευσης και για μεγαλύτερους υπολογιστικούς πόρους.

2.4.2 Ανιχνευτές ενός και δύο σταδίων

Στο πεδίο της ανίχνευσης αντικειμένων διακρίνονται δύο βασικές οικογένειες αρχιτεκτονικών. Οι ανιχνευτές δύο σταδίων παράγουν πρώτα ένα σύνολο υποψήφιων περιοχών ενδιαφέροντος και στη συνέχεια ταξινομούν καθεμία από αυτές. Η προσέγγιση αυτή, με χαρακτηριστικότερο εκπρόσωπο την οικογένεια R-CNN, επιτυγχάνει υψηλή ακρίβεια αλλά με σημαντικό υπολογιστικό κόστος.

Οι ανιχνευτές ενός σταδίου αντιμετωπίζουν αντίθετα το πρόβλημα ως μία ενιαία διαδικασία. Σε ένα μόνο πέρασμα από το δίκτυο υπολογίζονται τόσο οι θέσεις των αντικειμένων όσο και οι αντίστοιχες κλάσεις. Η προσέγγιση αυτή μπορεί να προσφέρει σημαντικά μικρότερο χρόνο εκτέλεσης, γεγονός που καθιστά την κατηγορία αυτή την ενδεδειγμένη επιλογή για εφαρμογές πραγματικού χρόνου σε συσκευές περιορισμένων πόρων. Η οικογένεια YOLO, που εξετάζεται στην επόμενη ενότητα, ανήκει σε αυτή την κατηγορία και αποτελεί τη βάση της παρούσας εργασίας.

2.4.3 Κατάτμηση έναντι ανίχνευσης

Αξίζει να σημειωθεί ότι μεγάλο μέρος της πρόσφατης έρευνας για τα απτικά πλακίδια δεν χρησιμοποιεί ανίχνευση αντικειμένων αλλά σημασιολογική κατάτμηση (semantic segmentation), δηλαδή ταξινόμηση κάθε εικονοστοιχείου της εικόνας. Η προσέγγιση αυτή αποδίδει ακριβέστερα το ακριβές σχήμα και την έκταση της όδευσης, αλλά είναι υπολογιστικά βαρύτερη. Στην παρούσα εργασία επιλέχθηκε η ανίχνευση αντικειμένων, καθώς βασικός στόχος είναι η εκτέλεση του συστήματος σε ένα συμβατικό κινητό τηλέφωνο. Η επιλογή αυτή σχετίζεται επομένως όχι μόνο με την ακρίβεια του μοντέλου, αλλά και με τους διαθέσιμους υπολογιστικούς πόρους της συσκευής, το ζήτημα συζητείται εκτενέστερα στην Ενότητα 2.8.

2.5 Η οικογένεια YOLO

Η αρχιτεκτονική YOLO (You Only Look Once) εισήγαγε την ιδέα της αντιμετώπισης της ανίχνευσης αντικειμένων ως ενιαίου προβλήματος παλινδρόμησης, σε αντίθεση με τις πολυσταδιακές προσεγγίσεις που είχαν προηγηθεί. Στην αρχική της διατύπωση, η εικόνα εισόδου διαιρείται νοητά σε ένα πλέγμα κελιών και κάθε κελί είναι υπεύθυνο για την πρόβλεψη αντικειμένων των οποίων το κέντρο βρίσκεται σε αυτό, μαζί με τις συντεταγμένες των πλαισίων οριοθέτησης, τη βαθμολογία εμπιστοσύνης και τις πιθανότητες των διαθέσιμων κλάσεων. Ένα ενιαίο συνελκτικό νευρωνικό δίκτυο παράγει τις προβλέψεις αυτές με ένα μόνο πέρασμα πάνω από ολόκληρη την εικόνα. Η συγκεκριμένη προσέγγιση μειώνει την ανάγκη για επαναλαμβανόμενους υπολογισμούς που χαρακτηρίζουν τους ανιχνευτές δύο σταδίων της Ενότητας 2.4.2 και συμβάλλει στην καταλληλότητα της οικογένειας για εφαρμογές πραγματικού χρόνου. Οι διαδοχικές εκδόσεις της YOLO βελτίωσαν σταδιακά την αρχική προσέγγιση, αντιμετωπίζοντας διαφορετικούς περιορισμούς. Η δεύτερη έκδοση εισήγαγε τα πλαίσια αγκύρωσης (anchor boxes) και την εκπαίδευση σε πολλαπλές αναλύσεις εισόδου [13]. Η τρίτη έκδοση επέκτεινε την πρόβλεψη σε τρεις διαφορετικές χωρικές κλίμακες, βελτιώνοντας την ανίχνευση αντικειμένων διαφορετικού μεγέθους και ιδιαίτερα των μικρών αντικειμένων [14]. Η τέταρτη έκδοση ενσωμάτωσε επιπλέον τεχνικές βελτιστοποίησης τόσο στη διαδικασία εκπαίδευσης όσο και στην αρχιτεκτονική του δικτύου, με στόχο τη βελτίωση της σχέσης μεταξύ ταχύτητας και ακρίβειας [15]. Από την πέμπτη έκδοση και έπειτα, η ανάπτυξη της οικογένειας συνδέθηκε κυρίως με την Ultralytics. Οι νεότερες

εκδόσεις έδωσαν μεγαλύτερη έμφαση στην πρακτική αξιοποίηση των μοντέλων, στην υποστήριξη διαφορετικών εργασιών υπολογιστικής όρασης, όπως η ανίχνευση, η κατάτμηση και η εκτίμηση πόζας, καθώς και στη διάθεση μοντέλων διαφορετικού μεγέθους. Οι εκδόσεις YOLOv5, YOLOv8 και YOLO11 ακολουθούν αυτή την κατεύθυνση, με σταδιακές βελτιώσεις στην αποδοτικότητα και στη δυνατότητα εκτέλεσης των μοντέλων σε διαφορετικές υπολογιστικές πλατφόρμες.

Η έκδοση YOLO11, που χρησιμοποιείται στην παρούσα εργασία, διατίθεται σε πέντε παραλλαγές μεγέθους: nano, small, medium, large και extra-large. Οι παραλλαγές αυτές επιτρέπουν διαφορετικό συμβιβασμό μεταξύ ακρίβειας και υπολογιστικού κόστους. Για την υλοποίηση του BlindNav επιλέχθηκε η YOLO11n (nano), καθώς ο μικρότερος αριθμός παραμέτρων και οι περιορισμένες υπολογιστικές απαιτήσεις της την καθιστούν καταλληλότερη για εκτέλεση σε κινητή συσκευή. Τα συγκεκριμένα χαρακτηριστικά του μοντέλου και η απόδοσή του στο περιβάλλον εκτέλεσης εξετάζονται αναλυτικότερα στην Ενότητα 3.5.

Ένα ζήτημα που πρέπει να ληφθεί υπόψη αφορά την τεκμηρίωση της YOLO11. Σε αντίθεση με τις πρώιμες εκδόσεις της οικογένειας, για τις οποίες υπάρχουν δημοσιευμένες επιστημονικές εργασίες [13]–[15], η Ultralytics δεν έχει δημοσιεύσει αντίστοιχη επιστημονική εργασία ειδικά για τη YOLO11. Για τον λόγο αυτό, στην παρούσα εργασία χρησιμοποιείται η επίσημη τεκμηρίωση λογισμικού της Ultralytics ως πηγή αναφοράς [16] [17].

2.6 Εκτέλεση στη συσκευή έναντι εκτέλεσης στο νέφος

Ένα σύστημα υπολογιστικής όρασης σε κινητή συσκευή μπορεί να εκτελεί την εξαγωγή προβλέψεων είτε τοπικά στη συσκευή (on-device ή edge inference) είτε αποστέλλοντας τα δεδομένα σε απομακρυσμένο διακομιστή (cloud inference). Η δεύτερη επιλογή επιτρέπει τη χρήση μεγαλύτερων και ακριβέστερων μοντέλων, καθώς δεν περιορίζεται από τους πόρους της συσκευής.

Για μια εφαρμογή υποβοήθησης κινητικότητας, ωστόσο, τρεις παράγοντες συνηγορούν αποφασιστικά υπέρ της τοπικής εκτέλεσης. Ο πρώτος είναι η καθυστέρηση: η αποστολή εικόνας και η αναμονή απάντησης προσθέτει καθυστέρηση δικτύου η οποία, σε ένα σύστημα που πρέπει να προειδοποιήσει τον χρήστη πριν φτάσει σε σημείο κινδύνου, είναι απαγορευτική. Ο δεύτερος είναι η διαθεσιμότητα εκτός δικτύου: ένα βοήθημα ασφάλειας δεν μπορεί να παύει να λειτουργεί σε περιοχή με κακή κάλυψη ή όταν εξαντληθούν τα δεδομένα του χρήστη. Και το τρίτο πλεονέκτημα αφορά την ιδιωτικότητα, καθώς η εικόνα από την κάμερα παραμένει στη συσκευή και δεν χρειάζεται να μεταδίδεται σε εξωτερικό διακομιστή.

Το βασικό μειονέκτημα της τοπικής εκτέλεσης είναι ότι οι υπολογιστικοί πόροι της κινητής συσκευής είναι περιορισμένοι. Για τον λόγο αυτό χρησιμοποιούνται εξειδικευμένα frameworks για φορητές και ενσωματωμένες συσκευές, όπως το LiteRT, πρώην TensorFlow Lite [18], σε συνδυασμό με μικρότερες και ελαφρύτερες αρχιτεκτονικές. Προηγούμενες εργασίες έχουν δείξει ότι η ανίχνευση εμποδίων σε εφαρμογές για άτομα με οπτική αναπηρία μπορεί να πραγματοποιηθεί σε χρόνους της τάξης των εκατοντάδων χιλιοστών του δευτερολέπτου ακόμη και σε συμβατικό επεξεργαστή.

2.7 Υπάρχουσες εφαρμογές και σχετική έρευνα

2.7.1 Εμπορικές εφαρμογές

Την τελευταία δεκαετία έχει αναπτυχθεί σημαντικός αριθμός εφαρμογών υποβοήθησης για άτομα με οπτική αναπηρία. Οι εφαρμογές αυτές κατατάσσονται σε τρεις βασικές κατηγορίες.

Η πρώτη περιλαμβάνει τις εφαρμογές περιγραφής περιβάλλοντος μέσω κάμερας. Το Seeing AI της Microsoft και το Lookout της Google λειτουργούν ως «ομιλούσα κάμερα»: αναγνωρίζουν κείμενο, έγγραφα, προϊόντα μέσω γραμμωτού κώδικα, χαρτονομίσματα, χρώματα, πρόσωπα και γενικές σκηνές. Παρέχουν πλούσια πληροφορία για το τι βρίσκεται μπροστά στον χρήστη, δεν αποτελούν όμως εφαρμογές πλοήγησης και δεν αναγνωρίζουν ούτε απτικά πλακίδια ούτε την κατάσταση φωτεινών σηματοδοτών.

Η δεύτερη κατηγορία περιλαμβάνει τις εφαρμογές χωρικής επίγνωσης μέσω GPS. Το BlindSquare και το Microsoft Soundscape ανακοινώνουν σημεία ενδιαφέροντος και διαμορφώνουν αντίληψη του περιβάλλοντος χώρου, το δεύτερο μέσω τρισδιάστατου χωρικού ήχου. Οι εφαρμογές αυτές δεν αξιοποιούν καθόλου την κάμερα και συνεπώς δεν έχουν πρόσβαση σε καμία οπτική πληροφορία του άμεσου περιβάλλοντος. Σημειώνεται ότι το Soundscape αποσύρθηκε από τη Microsoft και συνεχίζεται ως έργο ανοικτού κώδικα.

Η τρίτη κατηγορία περιλαμβάνει τις υπηρεσίες ανθρώπινης διαμεσολάβησης, όπως το Be My Eyes και το Aira, στις οποίες ο χρήστης συνδέεται με εθελοντή ή εκπαιδευμένο χειριστή μέσω ζωντανής βιντεοκλήσης. Παρέχουν υψηλή ευελιξία, απαιτούν όμως σύνδεση στο διαδίκτυο και διαθεσιμότητα τρίτου προσώπου, ενώ δεν συνιστούν αυτόνομο σύστημα πραγματικού χρόνου.

Η σύγκριση των εφαρμογών αυτών ως προς τα χαρακτηριστικά που ενδιαφέρουν την παρούσα εργασία παρουσιάζεται στον Πίνακα 2.1.

Εφαρμογή	Τοπική εκτέλεση	Χρήση κάμερας	Ανίχνευση πλακιδίων	Κατάσταση σηματοδότη	Ελληνικό TTS	Λειτουργία εκτός δικτύου
Seeing AI	Μερικώς	Ναι	Όχι	Όχι	Όχι	Μερικώς
Google Lookout	Μερικώς	Ναι	Όχι	Όχι	Όχι	Μερικώς
Soundscape / VoiceVista		Όχι	Όχι	Όχι	Όχι	Όχι
BlindSquare		Όχι	Όχι	Όχι	Όχι	Όχι
Be My Eyes	Όχι	Ναι	Όχι	Όχι		Όχι
Envision AI	Όχι	Ναι	Όχι	Όχι	Όχι	Όχι
Aira	Όχι	Ναι	Όχι	Όχι		Όχι
BlindNav	Ναι	Ναι	Ναι	Ναι	Ναι	Ναι

Πίνακας 2.1 Σύγκριση υφιστάμενων εφαρμογών ως προς τα χαρακτηριστικά του BlindNav

Το βασικό συμπέρασμα από τη σύγκριση είναι ότι οι εξεταζόμενες εμπορικές εφαρμογές δεν συνδυάζουν την αναγνώριση απτικών πλακιδίων με την αναγνώριση της κατάστασης ενός φωτεινού σηματοδότη πεζών. Οι εφαρμογές που βασίζονται στην κάμερα παρέχουν κυρίως περιγραφική πληροφορία, οι εφαρμογές GPS παρέχουν χωρική καθοδήγηση χωρίς άμεση οπτική αντίληψη του περιβάλλοντος, ενώ οι υπηρεσίες ανθρώπινης διαμεσολάβησης εξαρτώνται από τρίτο πρόσωπο και σύνδεση δικτύου. Η εικόνα αυτή συμφωνεί και με σχετική έρευνα μέσω ερωτηματολογίου σε άτομα με οπτική αναπηρία, σύμφωνα με την οποία οι υφιστάμενες εφαρμογές πλοήγησης δεν παρέχουν επαρκή

πληροφορία για φωτεινούς σηματοδότες, διασταυρώσεις και φυσικά εμπόδια. Στην ίδια έρευνα, περισσότερο από το 63% των συμμετεχόντων δήλωσαν ότι είχαν τραυματιστεί τουλάχιστον μία φορά κατά τη μετακίνησή τους σε εξωτερικό χώρο[4].

2.7.2 Ακαδημαϊκή έρευνα: ανίχνευση απτικών πλακιδίων

Η ερευνητική δραστηριότητα σχετικά με την αυτόματη αναγνώριση απτικών πλακιδίων έχει αυξηθεί τα τελευταία χρόνια και περιλαμβάνει διαφορετικές τεχνικές προσεγγίσεις.

Η πρώτη είναι η ανίχνευση αντικειμένων. Οι Aktas κ.ά. [19] συνδύασαν την αρχιτεκτονική YOLO-V3 με DenseNet, δημιουργώντας το μοντέλο YOLOV3Dense, και κατασκεύασαν για τον σκοπό αυτό το σύνολο δεδομένων MDPY με 4.580 εικόνες απτικών πλακιδίων. Ο Zhang [20] παρουσιάζει επισκόπηση της εφαρμογής των μοντέλων YOLO στην αναγνώριση απτικών πλακιδίων, επισημαίνοντας δύο συστηματικούς περιορισμούς του πεδίου: τη σπανιότητα διαθέσιμων συνόλων δεδομένων και την υστέρηση των ελαφριών μοντέλων ως προς την απόδοση.

Η δεύτερη κατεύθυνση, η οποία κυριαρχεί στην πρόσφατη βιβλιογραφία, είναι η σημασιολογική κατάτμηση. Οι Cao κ.ά. [21] ανέπτυξαν ελαφρύ δίκτυο κατάτμησης για ταυτόχρονη ανίχνευση οδεύσεων και διαβάσεων, επιτυγχάνοντας υψηλή ακρίβεια σε ενσωματωμένη πλατφόρμα Jetson Nano. Οι Zhang κ.ά. [22] πρότειναν το GRFB-UNet, αρχιτεκτονική πολλαπλών κλιμάκων με μηχανισμό προσοχής ειδικά για κατάτμηση απτικών πλακιδίων. Πιο πρόσφατες εργασίες επεκτείνουν την προσέγγιση σε ανίχνευση εσφαλμένα τοποθετημένων πλακιδίων και σε αυτόματη χαρτογράφηση οδεύσεων από εικόνες οδικού επιπέδου. Υπάρχουν επίσης προσεγγίσεις που συνδυάζουν την κάμερα με πρόσθετους αισθητήρες. Οι Yamanaka κ.ά. [23] για παράδειγμα, συνδυάζουν RGB κάμερα με αισθητήρα βάθους. Η χρήση επιπλέον αισθητήρων μπορεί να βελτιώσει την αναγνώριση, προϋποθέτει όμως εξοπλισμό πέρα από αυτόν που διαθέτει ένα τυπικό κινητό τηλέφωνο.

Πέρα από τις παραπάνω κατευθύνσεις, ιδιαίτερο ενδιαφέρον παρουσιάζουν οι προσπάθειες δημιουργίας δημόσια διαθέσιμων συνόλων δεδομένων, καθώς η σπανιότητά τους αποτελεί συστηματικό περιορισμό του πεδίου. Χαρακτηριστικά παραδείγματα αποτελούν το Tenji10K των Takano κ.ά. [24] με 10.000 εικόνες ιαπωνικών πλακιδίων για ανίχνευση και παρακολούθηση, καθώς και το SideGuide των Park κ.ά. [25], ένα εκτενές σύνολο δεδομένων πεζοδρομίου. Κρίσιμη διαπίστωση προκύπτει από την πρόσφατη εργασία των Hwang κ.ά. [26], οι οποίοι επισημαίνουν ότι τα διαθέσιμα σύνολα δεδομένων εμφανίζουν γεωγραφική μεροληψία: προέρχονται κατά κύριο λόγο από την Ανατολική Ασία, όπου κυριαρχούν οι ραβδώσεις κατεύθυνσης ως συνεχείς δείκτες όδευσης, ενώ υποεκπροσωπούνται οι ανάγλυφοι κόκκοι που χρησιμοποιούνται πρωτίστως στη Βόρεια Αμερική και την Ευρώπη ως προειδοποιητικοί δείκτες σε διαβάσεις και κρασπεδώσεις. Οι συγγραφείς τεκμηριώνουν πειραματικά ότι μοντέλα εκπαιδευμένα αποκλειστικά σε δεδομένα με ραβδώσεις δυσκολεύονται να γενικεύσουν σε πλακίδια με κόκκους, οδηγώντας σε χαμένες ανιχνεύσεις σε κρίσιμα για την ασφάλεια σημεία. Η διαπίστωση αυτή αιτιολογεί άμεσα τη μεθοδολογική επιλογή της παρούσας εργασίας να κατασκευάσει πρωτογενές σύνολο δεδομένων από τον ελληνικό αστικό χώρο, όπου συνυπάρχουν και οι δύο τύποι, αντί να αξιοποιήσει κάποιο από τα υπάρχοντα διεθνή σύνολα. Ως λύση στο πρόβλημα της περιορισμένης κάλυψης, οι ίδιοι προτείνουν την επαύξηση με φωτορεαλιστικά συνθετικά δεδομένα, προσέγγιση που βελτίωσε την επίδοση των μοντέλων έως και 29 μονάδες mIoU. Χρονολογικά προγενέστερες προσεγγίσεις βασίζονταν σε κλασική επεξεργασία εικόνας. Οι Kassim κ.ά. [27] ανέπτυξαν σύστημα ανίχνευσης σε περιβάλλον MATLAB σε συνδυασμό με πλατφόρμα Arduino, ικανό να διακρίνει διαφορετικά γεωμετρικά σχήματα του μοτίβου και να παράγει ηχητική ειδοποίηση, ενώ οι Ghilardi κ.ά. [28] πρότειναν μέθοδο ανίχνευσης σε πραγματικό χρόνο που συνδυάζει

αλγορίθμους υπολογιστικής όρασης με δέντρο απόφασης για τον περιορισμό των ψευδώς θετικών αποτελεσμάτων. Οι εργασίες αυτές τεκμηριώνουν την πορεία του πεδίου προς τη βαθιά μάθηση που περιγράφηκε στην Ενότητα 2.4.1. Ξεχωριστή κατεύθυνση, τέλος, αποτελεί η αξιοποίηση των απτικών πλακιδίων όχι από ανθρώπους αλλά από ρομποτικά συστήματα: οι Yeο κ.ά. [29] πρότειναν εμπλουτισμένο σχεδιασμό απτικών πλακιδίων (PATH Tiles), ο οποίος επιτρέπει σε ρομπότ που κινούνται σε δημόσιους χώρους να ερμηνεύουν ενσωματωμένη πληροφορία κινδύνου, επιτυγχάνοντας βελτίωση της τάξης του 71, 6% στην προληπτική ανίχνευση κινδύνων. Η προσέγγιση αυτή, αν και δεν απευθύνεται σε ανθρώπους χρήστες, καταδεικνύει ότι η τυποποιημένη υποδομή όδευσης μπορεί να λειτουργήσει ως αναγνώσιμο από μηχανές σημείο αναφοράς θέση που βρίσκεται στον πυρήνα της λογικής του BlindNav.

2.7.3 Ακαδημαϊκή έρευνα: αναγνώριση σηματοδοτών πεζών

Στο πεδίο των σηματοδοτών, οι Mascetti κ.ά. [30] παρουσίασαν σύστημα ανίχνευσης σε κινητή συσκευή με έμφαση στην ανθεκτικότητα σε μεταβαλλόμενες συνθήκες φωτισμού, το οποίο αξιολογήθηκε με πραγματικούς χρήστες με οπτική αναπηρία. Οι Yu κ.ά. [31] ανέπτυξαν το LYTNet, συνελικτικό δίκτυο πραγματικού χρόνου για ταυτόχρονη αναγνώριση σηματοδοτών πεζών και διαβάσεων, και δημοσίευσαν το σύνολο δεδομένων ImVisible (PTL Dataset) με πάνω από 5.000 εικόνες, το οποίο αποτελεί έκτοτε το βασικό σημείο αναφοράς του πεδίου. Μεταγενέστερες εργασίες αξιοποιούν νεότερες εκδόσεις YOLO με μηχανισμούς προσοχής, ενώ ιδιαίτερη ερευνητική προσοχή δίνεται στη λειτουργία υπό συνθήκες χαμηλής ορατότητας και στην αντιμετώπιση της κίνησης της κάμερας κατά τη βάδιση.

Στο ίδιο ελληνικό ερευνητικό πλαίσιο εντάσσονται και δύο πρόσφατες εργασίες, οι οποίες προσεγγίζουν το ζήτημα από συμπληρωματικές κατευθύνσεις και σχετίζονται άμεσα με τους στόχους της παρούσας εργασίας.

Οι Tzimos κ.ά. [32] παρουσιάζουν πλατφόρμα εικονικής πραγματικότητας για την εκπαίδευση ατόμων με οπτική αναπηρία στην αλληλεπίδραση με απτικά πλακίδια όδευσης. Η πλατφόρμα, υλοποιημένη σε Unity, αναπαριστά με υψηλή πιστότητα τις υφές διαφορετικών τύπων πλακιδίων και συνδυάζει απτική, ηχητική και οπτική ανάδραση, επιτρέποντας στον χρήστη να εξασκηθεί σε τεχνικές προσανατολισμού και κινητικότητας μέσα σε ασφαλές αλλά ρεαλιστικό περιβάλλον. Η αξιολόγηση με 19 συμμετέχοντες, μέσω των ερωτηματολογίων USE και UEQ, κατέγραψε τόσο ποσοτικά δεδομένα επίδοσης όσο και υποκειμενικές εντυπώσεις. Η εργασία αυτή είναι συμπληρωματική ως προς το BlindNav: εστιάζει στην εκπαίδευση του χρήστη ώστε να αναγνωρίζει τα πλακίδια με την αφή, σε προσομοιωμένο περιβάλλον, ενώ το BlindNav στοχεύει στην υποβοήθηση του χρήστη σε πραγματικό αστικό περιβάλλον, αναγνωρίζοντας τα ίδια πλακίδια οπτικά. Οι δύο προσεγγίσεις καλύπτουν διαφορετικά στάδια της ίδιας διαδρομής: την προετοιμασία και την πραγματική μετακίνηση αντίστοιχα. Οι Voutsakelis κ.ά. [33] ανέπτυξαν εφαρμογή έξυπνου τηλεφώνου για την ανίχνευση αντικειμένων σε πραγματικό χρόνο, βασισμένη στο ελαφρύ μοντέλο EfficientDet-lite2 και στο σύνολο δεδομένων COCO, με συνδυασμό απτικής ανάδρασης ενεργοποιούμενης όταν ο χρήστης αγγίζει ένα ανιχνευμένο αντικείμενο στην οθόνη και ηχητικών ανακοινώσεων σε πολλαπλές γλώσσες. Η αξιολόγηση διεξήχθη με 26 συμμετέχοντες (8 τυφλούς και 18 βλέποντες με δεμένα μάτια), μέσω των ερωτηματολογίων USE και UEQ, και περιλάμβανε συγκριτική ανάλυση με και χωρίς ενεργοποιημένη την απτική ανάδραση. Το βασικό εύρημα ήταν ότι η απτική ανάδραση βελτιώνει αισθητά και τους τέσσερις άξονες ευχρηστίας, με τη μεγαλύτερη διαφορά να καταγράφεται στην ικανοποίηση του χρήστη. Η εργασία των Voutsakelis κ.ά. παρουσιάζει ιδιαίτερο ενδιαφέρον για την παρούσα διπλωματική για τρεις λόγους. Πρώτον, επιβεβαιώνει ερευνητικά τη σχεδιαστική επιλογή του BlindNav να συνδυάζει ηχητική με απτική

ανάδραση, αντί να στηρίζεται αποκλειστικά στην πρώτη. Δεύτερον, υιοθετεί την ίδια μεθοδολογία αξιολόγησης με βλέποντες συμμετέχοντες υπό συνθήκη προσομοιωμένης τύφλωσης, επισημαίνοντας ρητά όπως και η παρούσα εργασία ότι η προσομοίωση δεν μπορεί να αναπαραγάγει πλήρως την εμπειρία της πραγματικής οπτικής αναπηρίας. Τρίτον, και σημαντικότερο, οι συμμετέχοντες στη μελέτη τους ζήτησαν αυθόρμητα δυνατότητα που η εφαρμογή δεν διέθετε: πληροφορία για τη σχετική θέση του αντικειμένου ως προς τον χρήστη (αριστερά, δεξιά, μπροστά). Η ανάγκη αυτή καλύπτεται ακριβώς από τον μηχανισμό χωρικών ζωνών του BlindNav (Ενότητα 3.6).

Οι διαφορές των δύο συστημάτων είναι ωστόσο ουσιώδεις. Η εφαρμογή των Voutsakelis κ.ά. αναγνωρίζει γενικά αντικείμενα καθημερινής χρήσης, αξιοποιώντας το προεκπαιδευμένο σύνολο COCO, και απαιτεί από τον χρήστη να αγγίζει την οθόνη για να λάβει ανάδραση· απευθύνεται δηλαδή κυρίως στην εξερεύνηση του άμεσου περιβάλλοντος από στάση. Το BlindNav, αντίθετα, χρησιμοποιεί μοντέλα εκπαιδευμένα ειδικά σε απτικά πλακίδια και σηματοδότες πεζών, στοιχεία που δεν περιλαμβάνονται στο COCO, και σχεδιάστηκε για χρήση εν κινήσει, χωρίς καμία αλληλεπίδραση με την οθόνη: η εναλλαγή λειτουργίας προκύπτει αποκλειστικά από την κλίση της συσκευής. Σημειώνεται τέλος ότι η εφαρμογή τους επιτυγχάνει σαφώς ταχύτερη εξαγωγή προβλέψεων (70 ms έναντι 320–460 ms του BlindNav), διαφορά που αποδίδεται στο ελαφρύτερο μοντέλο EfficientDet-lite2 έναντι του YOLO11n και αποτελεί σαφή κατεύθυνση βελτίωσης για την παρούσα εργασία.

2.7.4 Το ελληνικό ερευνητικό πλαίσιο

Στην Ελλάδα έχουν αναπτυχθεί διάφορες ερευνητικές προσεγγίσεις για την υποβοήθηση της πλοήγησης ατόμων με οπτική αναπηρία σε εξωτερικούς χώρους, μεταξύ άλλων με τη συμμετοχή του Φάρου Τυφλών της Ελλάδος. Τα συστήματα BlindHelper και BlindRouteVision παρέχουν πλοήγηση πεζού με βάση δορυφορικό εντοπισμό και έχουν αξιολογηθεί ως προς την ευχρηστία τους με τη συμμετοχή πραγματικών χρηστών. Η προσέγγισή τους βασίζεται όμως σε πρόσθετο εξοπλισμό, όπως εξωτερικό δέκτη GPS υψηλής ακρίβειας, ειδική συσκευή εγκατεστημένη στον φωτεινό σηματοδότη για τη μετάδοση της κατάστασής του και αισθητήρα υπερήχων για την ανίχνευση κοντινών εμποδίων. Ιδιαίτερο ενδιαφέρον παρουσιάζει το γεγονός ότι οι συγγραφείς αναφέρουν ότι το σύστημα μπορεί να λειτουργήσει χωρίς να προϋποθέτει την ύπαρξη απτικών πλακιδίων. Επομένως, παρότι η προσέγγιση αντιμετωπίζει την πλοήγηση και την ενημέρωση σχετικά με τον σηματοδότη, δεν αξιοποιεί την υπάρχουσα υποδομή απτικής όδευσης ως πηγή πληροφορίας.

2.8 Το κενό στη βιβλιογραφία και η συνεισφορά της εργασίας

Από την επισκόπηση της σχετικής βιβλιογραφίας και των υφιστάμενων εφαρμογών προκύπτουν τέσσερα βασικά σημεία, τα οποία συνδέονται με το ερευνητικό κενό που εξετάζει η παρούσα εργασία.

Πρώτον, οι διαδεδομένες εμπορικές εφαρμογές δεν καλύπτουν το πρόβλημα. Όπως αποτυπώνεται στον Πίνακα 2.1, καμία από τις ευρέως χρησιμοποιούμενες εφαρμογές δεν αναγνωρίζει απτικά πλακίδια, ούτε την κατάσταση φωτεινού σηματοδότη πεζών. Οι εφαρμογές κάμερας περιγράφουν το περιβάλλον χωρίς να καθοδηγούν, οι εφαρμογές δορυφορικού εντοπισμού καθοδηγούν χωρίς να βλέπουν, και οι υπηρεσίες διαμεσολάβησης εξαρτώνται από τη διαθεσιμότητα τρίτου προσώπου και από σύνδεση δικτύου. Η διαπίστωση αυτή επιβεβαιώνεται και από τη διεθνή βιβλιογραφία επισκόπησης, η οποία καταγράφει ρητά την ανεπάρκεια της παρεχόμενης πληροφορίας για σηματοδότες και διασταυρώσεις [4].

Δεύτερον, η ακαδημαϊκή έρευνα παραμένει σε επίπεδο πρωτοτύπου και εξειδικεύεται σε έναν μόνο στόχο. Οι εργασίες που εξετάστηκαν αντιμετωπίζουν είτε τα απτικά πλακίδια είτε τους σηματοδότες,

σπανίως και τα δύο στο ίδιο σύστημα. Επιπλέον, η κυρίαρχη τεχνική κατεύθυνση της σημασιολογικής κατάταξης, αν και ακριβέστερη, είναι υπολογιστικά απαιτητική και αξιολογείται συνήθως σε εξειδικευμένες ενσωματωμένες πλατφόρμες αντί σε τυπικό κινητό τηλέφωνο. Ο ίδιος ο Zhang [20] επισημαίνει ως συστηματικό περιορισμό του πεδίου την υστέρηση των ελαφριών μοντέλων, δηλαδή ακριβώς εκείνων που είναι υλοποιήσιμα σε καταναλωτική συσκευή.

Τρίτον, οι υπάρχουσες ολοκληρωμένες λύσεις προϋποθέτουν πρόσθετο εξοπλισμό ή τροποποίηση της υποδομής. Το χαρακτηριστικότερο παράδειγμα προέρχεται από τον ίδιο τον ελληνικό χώρο: το σύστημα BlindRouteVision επιτυγχάνει αξιόπιστη πλοήγηση και ενημέρωση για την κατάσταση του σηματοδότη, το κάνει όμως μέσω εξωτερικού δέκτη GPS, αισθητήρα υπερήχων και ειδικής συσκευής εγκατεστημένης επί του σηματοδότη. Η προσέγγιση είναι τεχνικά εύρωστη, αλλά η κλιμάκωσή της απαιτεί παρέμβαση στην αστική υποδομή και απόκτηση εξοπλισμού από τον χρήστη.

Τέταρτον, η υπάρχουσα υποδομή υποαξιοποιείται. Τα απτικά πλακίδια αποτελούν εγκατεστημένη, τυποποιημένη και αποδεδειγμένα αποτελεσματική υποδομή καθοδήγησης [2], η οποία ωστόσο στην ελληνική πραγματικότητα χαρακτηρίζεται από ασυνέχεια και συχνή κατάληψη από σταθμευμένα οχήματα [3]. Είναι χαρακτηριστικό ότι το BlindRouteVision επιλέγει να λειτουργεί ανεξάρτητα από την ύπαρξη της υποδομής αυτής, αντί να την αξιοποιήσει.

2.8.1 Η προτεινόμενη προσέγγιση

Το BlindNav τοποθετείται στο σημείο τομής των κενών αυτών, με τέσσερα διακριτά χαρακτηριστικά.

Λειτουργεί χωρίς κανέναν πρόσθετο εξοπλισμό, αξιοποιώντας αποκλειστικά την κάμερα και τους αισθητήρες ενός συμβατικού κινητού τηλεφώνου, και χωρίς καμία τροποποίηση της αστικής υποδομής. Αντί να παρακάμπτει τα απτικά πλακίδια, τα αξιοποιεί ως πηγή καθοδήγησης, αναγνωρίζοντας οπτικά τη σημασιολογία που ορίζει το πρότυπο ISO 23599 και μετατρέποντάς την σε φωνητική και απτική πληροφορία. Ενοποιεί σε μία εφαρμογή και τους δύο στόχους απτικά πλακίδια και φωτεινούς σηματοδότες με αυτόματη εναλλαγή βάσει του προσανατολισμού της συσκευής, χωρίς να απαιτείται οπτική αλληλεπίδραση από τον χρήστη. Τέλος, εκτελεί όλη την επεξεργασία τοπικά στη συσκευή, με ελληνική φωνητική ανάδραση, εξασφαλίζοντας λειτουργία εκτός δικτύου και προστασία των προσωπικών δεδομένων.

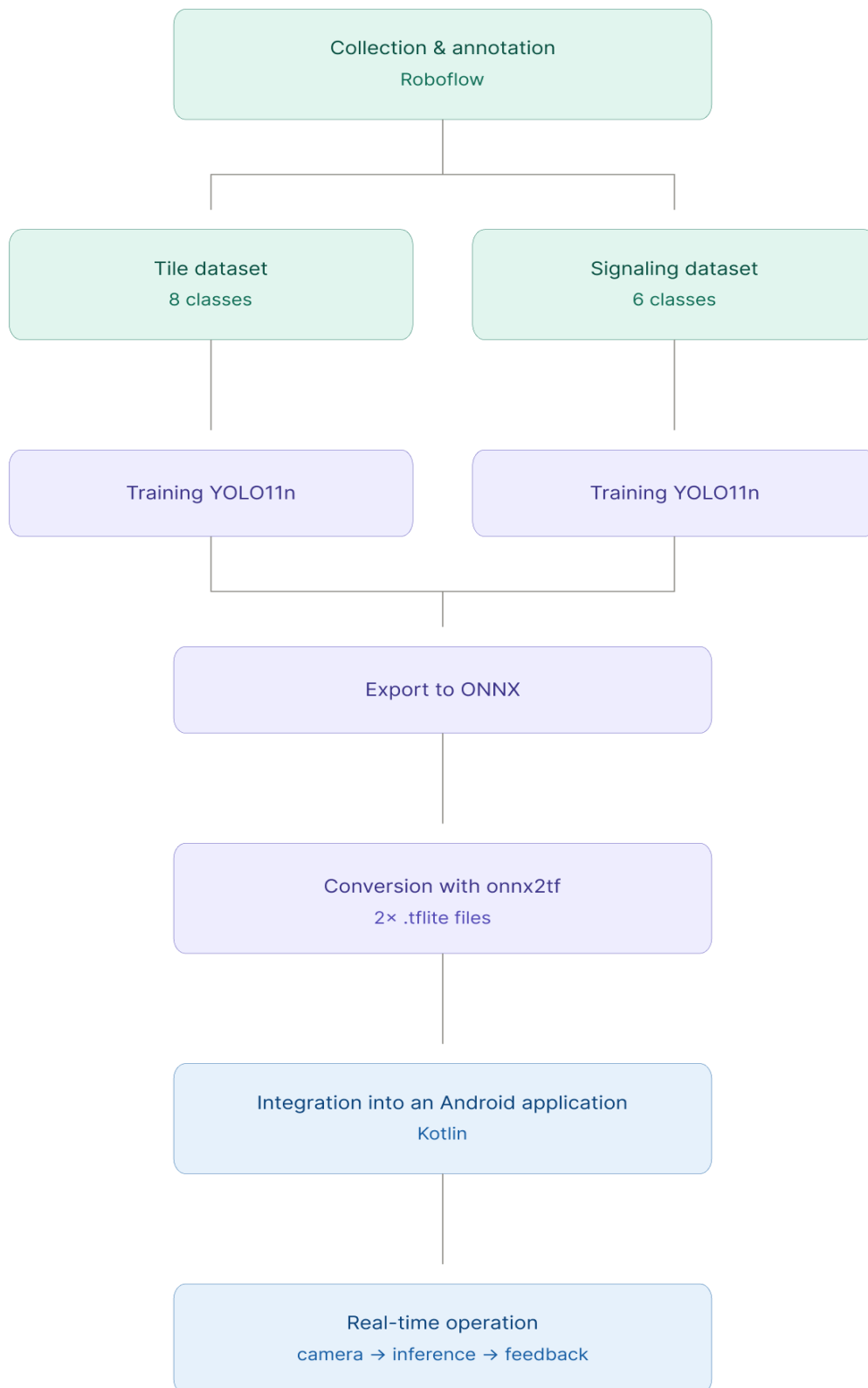
2.9 Επίλογος

Το κεφάλαιο αυτό κάλυψε το θεωρητικό και τεχνολογικό υπόβαθρο της εργασίας. Παρουσιάστηκε το μέγεθος του προβλήματος της οπτικής αναπηρίας και τα όρια των καθιερωμένων βοηθημάτων κινητικότητας, η ιστορία, η τυπολογία και το θεσμικό πλαίσιο των απτικών πλακιδίων όδευσης, καθώς και η λειτουργία των φωτεινών σηματοδοτών πεζών. Αναλύθηκαν στη συνέχεια οι τεχνολογικές επιλογές που θεμελιώνουν τη μεθοδολογία της εργασίας: η μετάβαση από κλασικές μεθόδους επεξεργασίας εικόνας στη βαθιά μάθηση, η οικογένεια αρχιτεκτονικών YOLO και η επιλογή εκτέλεσης στη συσκευή αντί στο νέφος. Η επισκόπηση των υφιστάμενων εμπορικών εφαρμογών και της ακαδημαϊκής έρευνας οδήγησε στον προσδιορισμό ενός σαφούς ερευνητικού κενού, το οποίο το BlindNav καλείται να καλύψει μέσω τεσσάρων διακριτών χαρακτηριστικών: απουσία πρόσθετου εξοπλισμού, αξιοποίηση της υπάρχουσας υποδομής, ενοποίηση δύο στόχων σε μία εφαρμογή, και πλήρως τοπική εκτέλεση. Η μεθοδολογία με την οποία υλοποιήθηκε η προσέγγιση αυτή αναπτύσσεται στο επόμενο κεφάλαιο.

Κεφάλαιο 3ο: Μεθοδολογία Υλοποίησης

3.1 Επισκόπηση της προσέγγισης

Η μεθοδολογία που ακολουθήθηκε για την ανάπτυξη του BlindNav στοχεύει στη δημιουργία ενός πλήρως ενσωματωμένου (on-device) συστήματος υπολογιστικής όρασης, το οποίο εντοπίζει σε πραγματικό χρόνο δύο κατηγορίες στοιχείων του αστικού περιβάλλοντος που είναι κρίσιμες για την ασφαλή μετακίνηση ατόμων με οπτική αναπηρία: τα απτικά πλακίδια όδευσης (tactile paving) και τους φωτεινούς σηματοδότες πεζών. Η πληροφορία που προκύπτει από την ανίχνευση μεταφράζεται σε ηχητική (ελληνικό text-to-speech) και απτική (δόνηση) ανάδραση, ώστε ο χρήστης να καθοδηγείται χωρίς να απαιτείται όραση. Βασική σχεδιαστική επιλογή αποτελεί η εκτέλεση όλης της επεξεργασίας τοπικά στη συσκευή, χωρίς εξάρτηση από απομακρυσμένο διακομιστή ή σύνδεση στο διαδίκτυο. Η επιλογή αυτή οφείλεται στη φύση της εφαρμογής: ένα βοήθημα κινητικότητας που χρησιμοποιείται εν κινήσει πρέπει να λειτουργεί με ελάχιστη καθυστέρηση (latency), να παραμένει διαθέσιμο εκτός δικτύου και να μη διαβιβάζει την εικόνα του περιβάλλοντος του χρήστη σε τρίτους, διαφυλάσσοντας έτσι την ιδιωτικότητα. Για τον σκοπό αυτό, η εκτέλεση των μοντέλων γίνεται μέσω του LiteRT (πρώην TensorFlow Lite), του χρόνου εκτέλεσης της Google για μηχανική μάθηση σε φορητές και ενσωματωμένες συσκευές [18]. Η συνολική μεθοδολογία αναπτύσσεται σε τέσσερα διαδοχικά στάδια, τα οποία αναλύονται στις επόμενες ενότητες του κεφαλαίου. Το πρώτο στάδιο αφορά τη συλλογή και προετοιμασία των δεδομένων. Δημιουργήθηκαν δύο ξεχωριστά σύνολα δεδομένων εικόνων ένα για τα απτικά πλακίδια και ένα για τους σηματοδότες τα οποία επισημειώθηκαν (annotation) και οργανώθηκαν σε κλάσεις με χρήση της πλατφόρμας Roboflow [34]. Το σύνολο των απτικών πλακιδίων περιλαμβάνει οκτώ κλάσεις και το σύνολο των σηματοδοτών έξι. Στο δεύτερο στάδιο, τα σύνολα αυτά χρησιμοποιήθηκαν για την εκπαίδευση δύο μοντέλων ανίχνευσης αντικειμένων (object detection) της αρχιτεκτονικής YOLO11, στην ελαφριά παραλλαγή της (YOLO11n / nano), η οποία είναι σχεδιασμένη για εκτέλεση σε συσκευές περιορισμένων υπολογιστικών πόρων [16] [17]. Η εκπαίδευση δύο ανεξάρτητων μοντέλων, αντί ενός ενιαίου, δεν είναι τυχαία επιλογή: τα απτικά πλακίδια εντοπίζονται όταν η συσκευή είναι στραμμένη προς το έδαφος, ενώ οι σηματοδότες όταν είναι στραμμένη προς τα εμπρός ή προς τα πάνω. Ο διαχωρισμός αυτός επιτρέπει σε κάθε μοντέλο να εξειδικεύεται στο δικό του οπτικό πλαίσιο και δίνει τη δυνατότητα αυτόματης εναλλαγής μεταξύ των δύο λειτουργιών με βάση την κλίση της συσκευής (device pitch), μηχανισμός που αναλύεται στην Ενότητα 3.7. Το τρίτο στάδιο περιλαμβάνει την εξαγωγή και μετατροπή των εκπαιδευμένων μοντέλων σε μορφή κατάλληλη για κινητή συσκευή. Τα μοντέλα εξήχθησαν αρχικά στο ενδιάμεσο, ανεξάρτητο από το εκάστοτε πλαίσιο (framework) format ONNX [35] και στη συνέχεια μετατράπηκαν σε αρχεία.tflite μέσω του εργαλείου onnx2tf [36]. Το τέταρτο και τελευταίο στάδιο είναι η ενσωμάτωση των μοντέλων στην εφαρμογή Android και η λειτουργία τους σε πραγματικό χρόνο. Η εφαρμογή, γραμμένη σε γλώσσα Kotlin, αξιοποιεί τη βιβλιοθήκη CameraX για τη λήψη καρτέ (frames) από την κάμερα [37], τα οποία τροφοδοτούνται στον διερμηνέα (interpreter) του LiteRT για την εξαγωγή προβλέψεων. Οι ανιχνεύσεις δεν χρησιμοποιούνται αυτούσιες· περνούν από ένα στάδιο χρονικής εξομάλυνσης (temporal smoothing) και χωρικής ταξινόμησης, ώστε να παραχθεί σταθερή και αξιόπιστη καθοδήγηση κατά τη βόλτα, πριν αποδοθούν ως ηχητική και απτική ανάδραση. Η συνολική αυτή ροή απεικονίζεται συνοπτικά στο Σχήμα 3.1. Η δομή του υπόλοιπου κεφαλαίου την ακολουθεί: η Ενότητα 3.2 παρουσιάζει τη συλλογή και προετοιμασία των δεδομένων, η 3.3 την εκπαίδευση των μοντέλων, η 3.4 τη διαδικασία εξαγωγής και μετατροπής, η 3.5 την ανάπτυξη της εφαρμογής Android, η 3.6 το pipeline ανίχνευσης και επεξεργασίας, η 3.7 τα συστήματα ανάδρασης και η 3.8 τις τελικές ρυθμίσεις και τις σχεδιαστικές επιλογές προσβασιμότητας.



Εικόνα 3.1 Διάγραμμα ροής του pipeline (δεδομένα → εκπαίδευση → εξαγωγή → ενσωμάτωση)

3.2 Συλλογή και προετοιμασία των δεδομένων

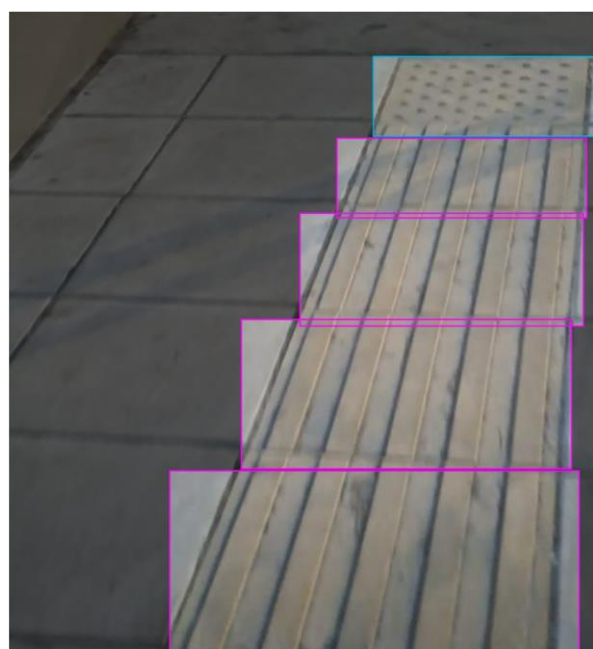
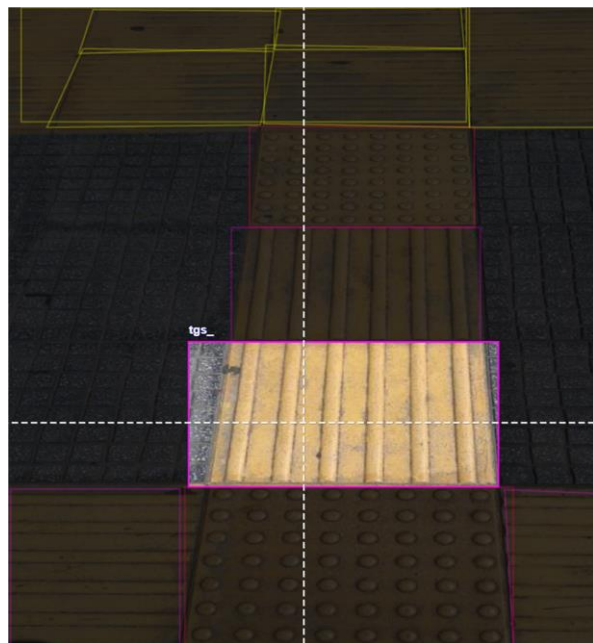
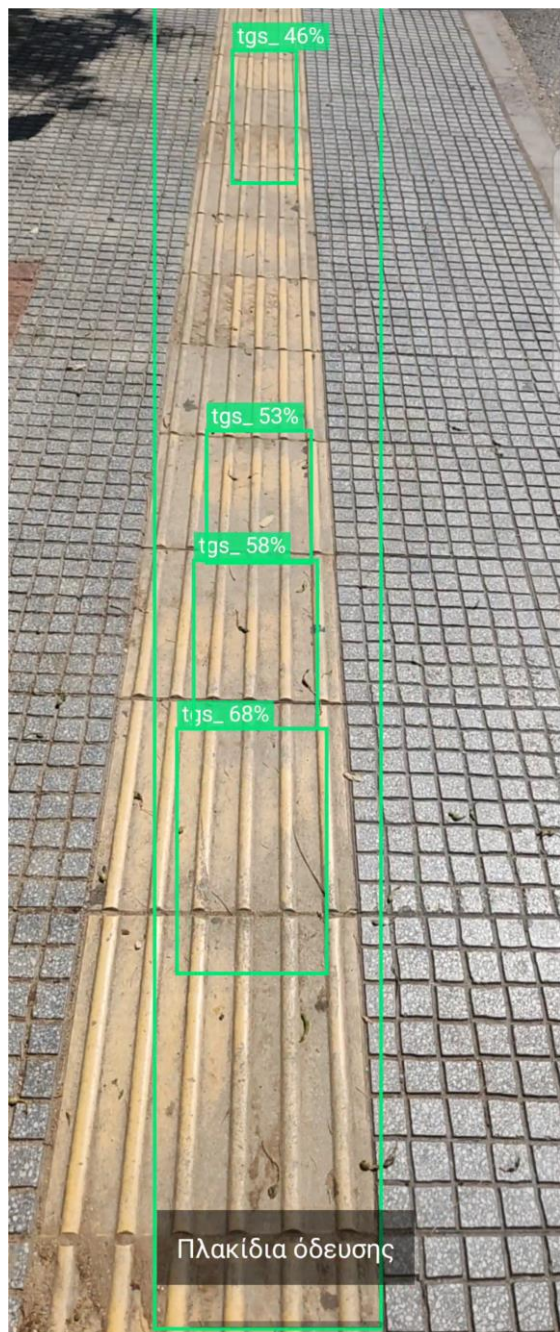
Η ποιότητα ενός μοντέλου ανίχνευσης αντικειμένων εξαρτάται άμεσα από τα δεδομένα στα οποία εκπαιδεύεται. Για τις ανάγκες του BlindNav δημιουργήθηκαν δύο ανεξάρτητα σύνολα δεδομένων εικόνων, ένα για κάθε κατηγορία στόχων: ένα για τα απτικά πλακίδια όδευσης και ένα για τους φωτεινούς σηματοδότες πεζών. Ο διαχωρισμός αυτός, όπως αναφέρθηκε στην Ενότητα 3.1, αντανακλά τα διαφορετικά οπτικά πλαίσια στα οποία εμφανίζεται κάθε κατηγορία.

Οι εικόνες όπου συλλέχθηκαν για τα πλακίδια έγινε από προσωπικό data set όπου συλλέχτηκαν από διαφορετικά σημεία του αστικού ιστού της Θεσσαλονίκης και της Κοζάνης ενώ για τους φωτεινούς σηματοδότες το data set με της φωτογραφίες και το annotation προήλθε από open dataset του Roboflow [38], με στόχο την κάλυψη ρεαλιστικών συνθηκών χρήσης, όπως διαφορετικός φωτισμός, γωνίες λήψης και τύποι πλακιδίων. Συνολικά, το σύνολο των πλακιδίων αριθμεί 4 κατηγορίες εικόνων και το σύνολο των σηματοδοτών 2, επίσης διαθέτει και 3 επιπλέον ανιχνεύσεις πέραν των τυπικών.

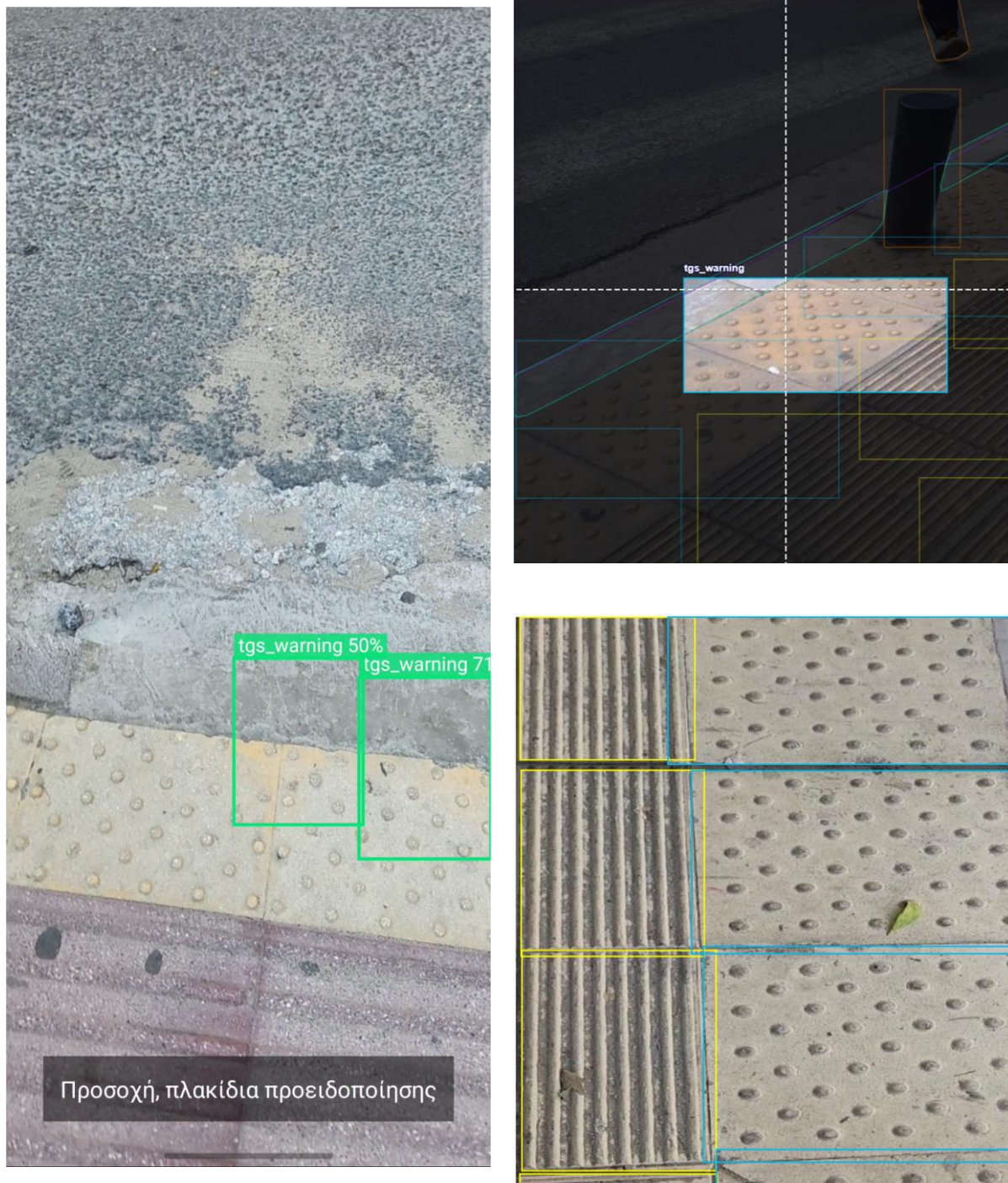
3.2.1 Αναγνώριση εντός εφαρμογής

Παρακάτω είναι αναλυτικά με φωτογραφίες για την κάθε κλάση που αναγνωρίζει η εφαρμογή

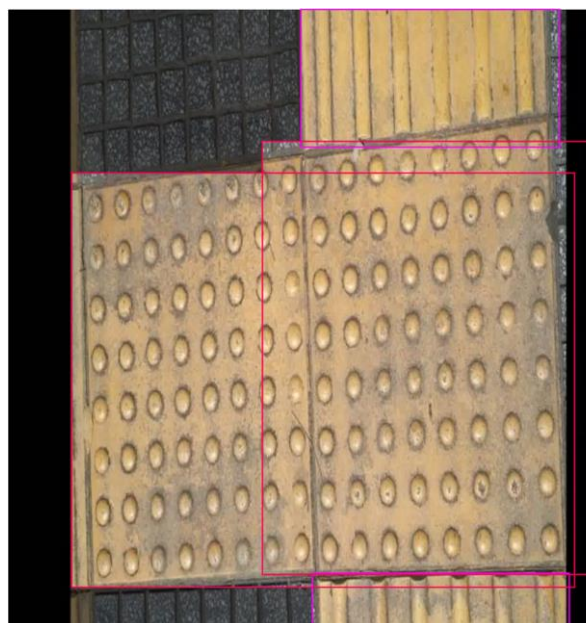
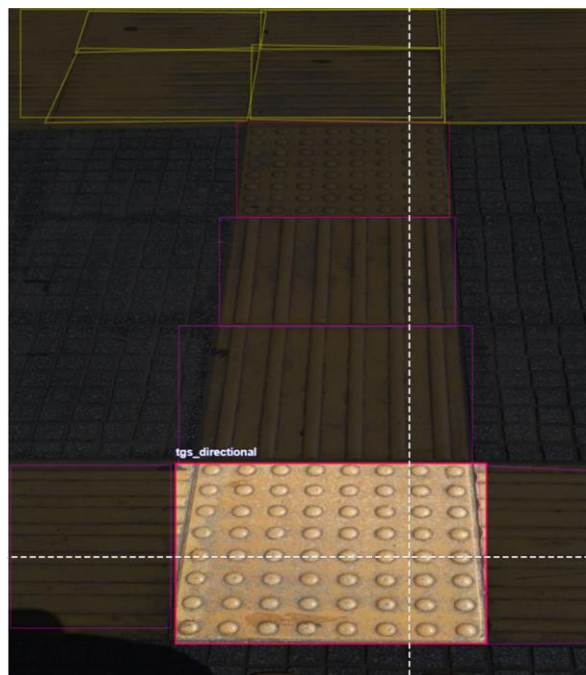
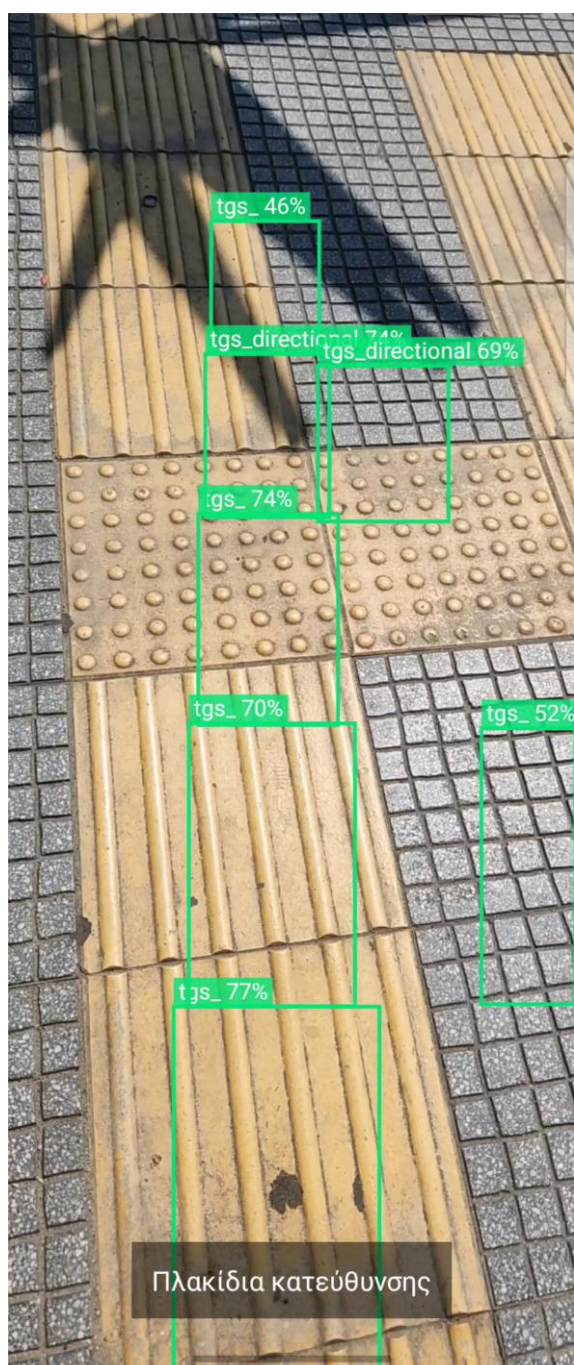
- Κλάση πλακιδίων



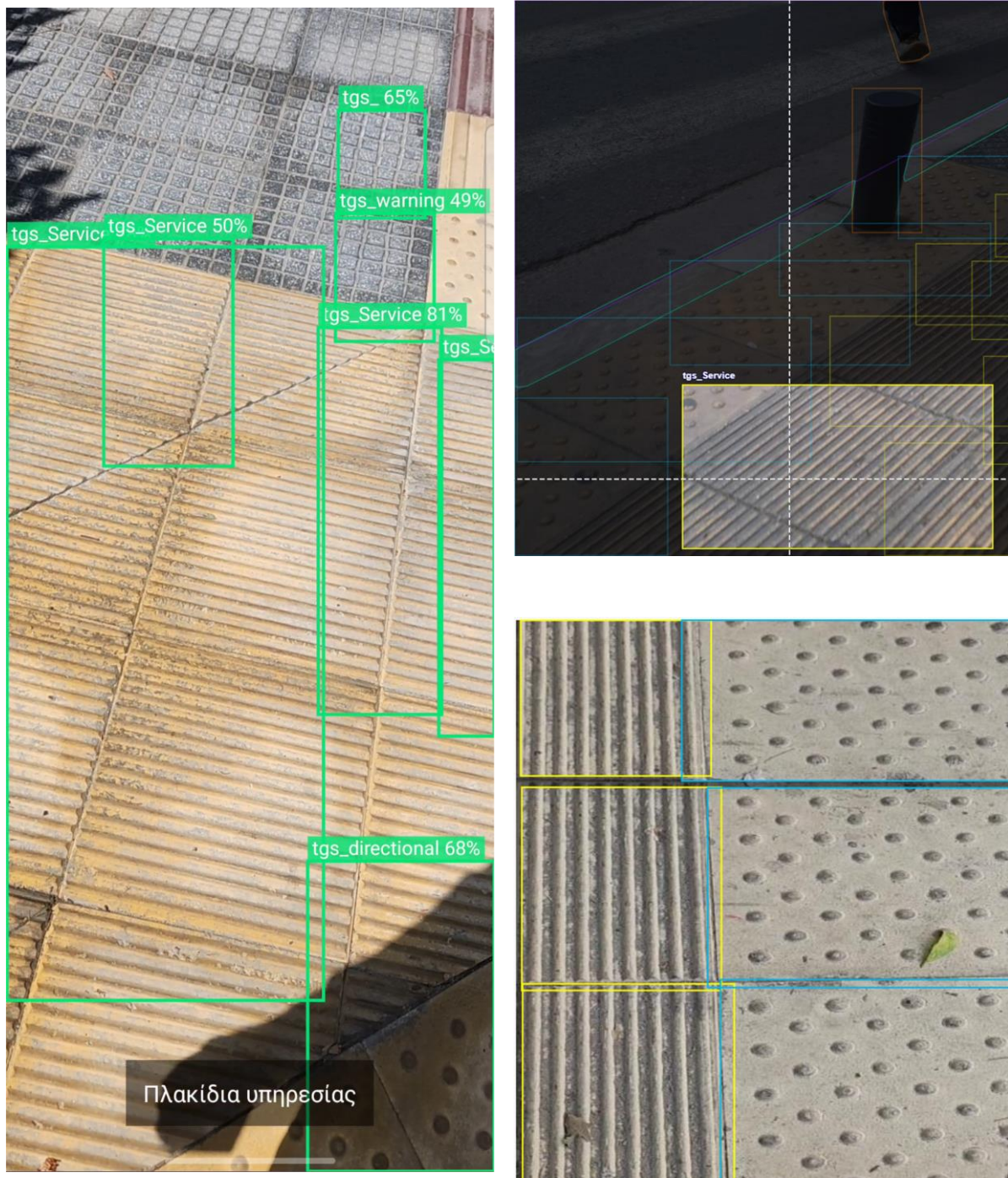
Εικόνα 3.2 Πλακίδια Όδευσης με κλάση "tgs_"



Εικόνα 3.3 Πλακίδια προειδοποίησης με κλάση "tgs_warning"

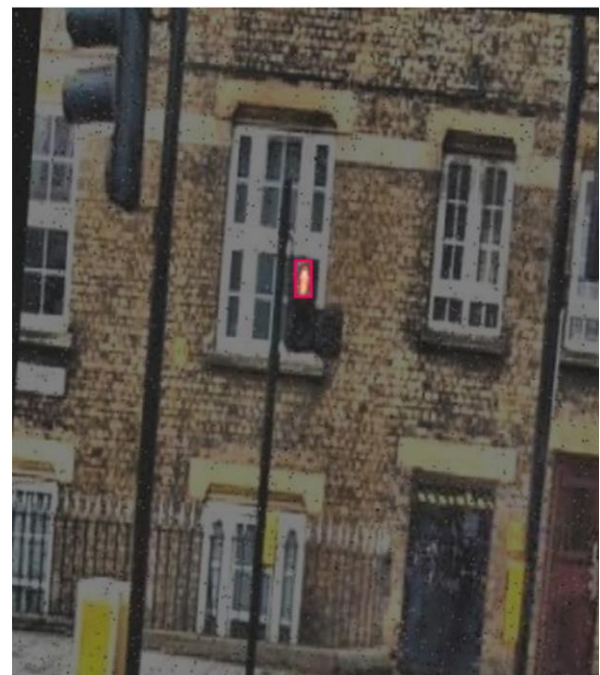
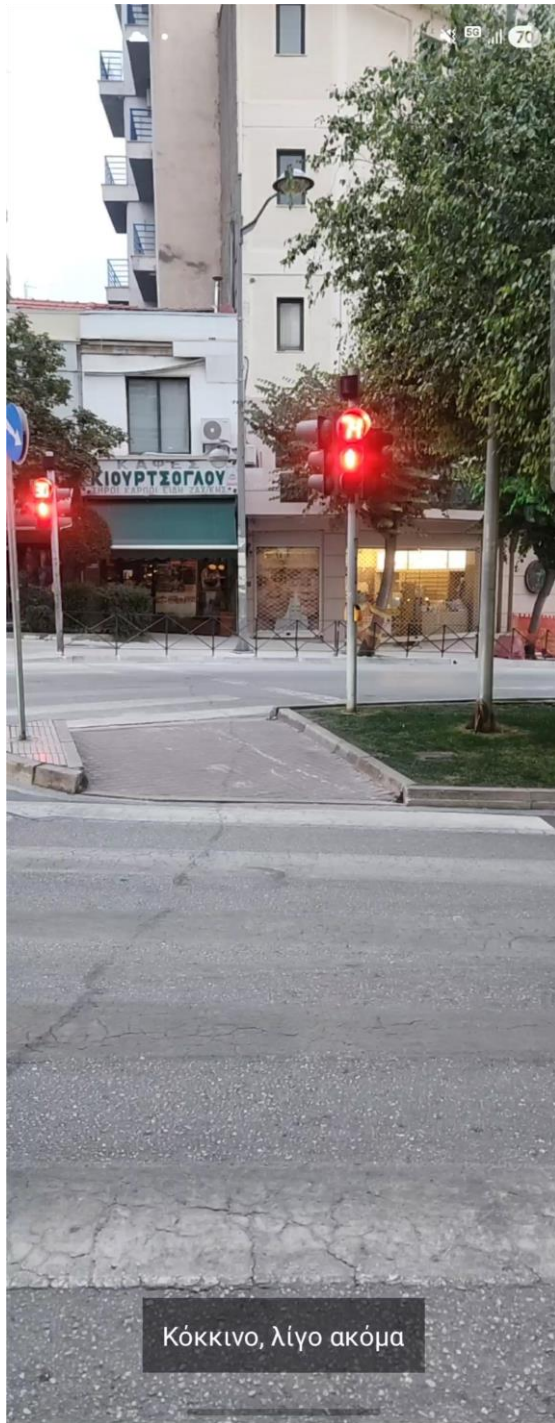


Εικόνα 3.4 Πλακίδια κατεύθυνσης με κλάση "tgs_directional"

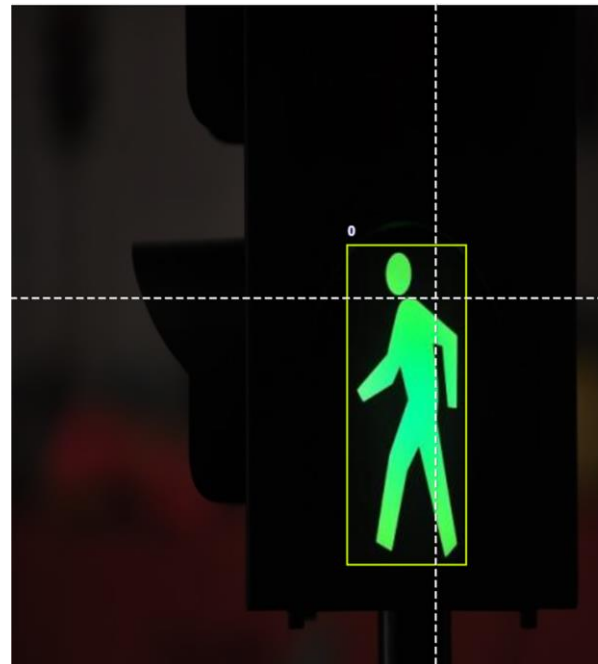


Εικόνα 3.5 Πλακίδια υπηρεσίας με κλάση "tgs_Service"

- Κλάση φωτεινών σηματοδότη

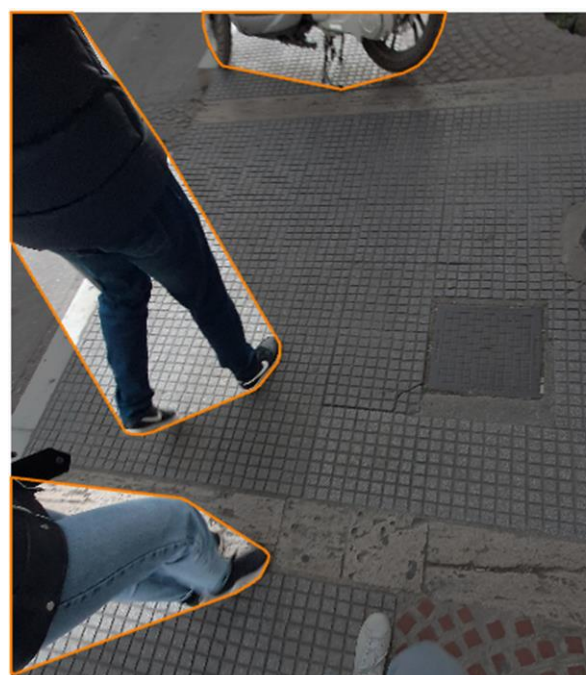
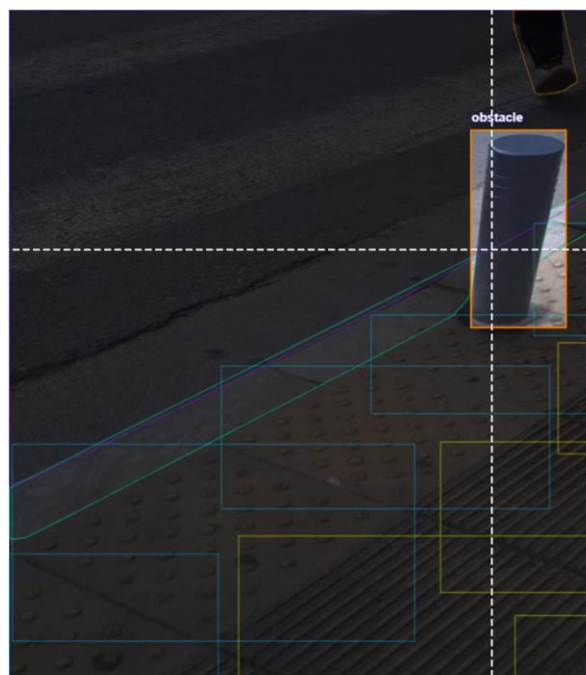
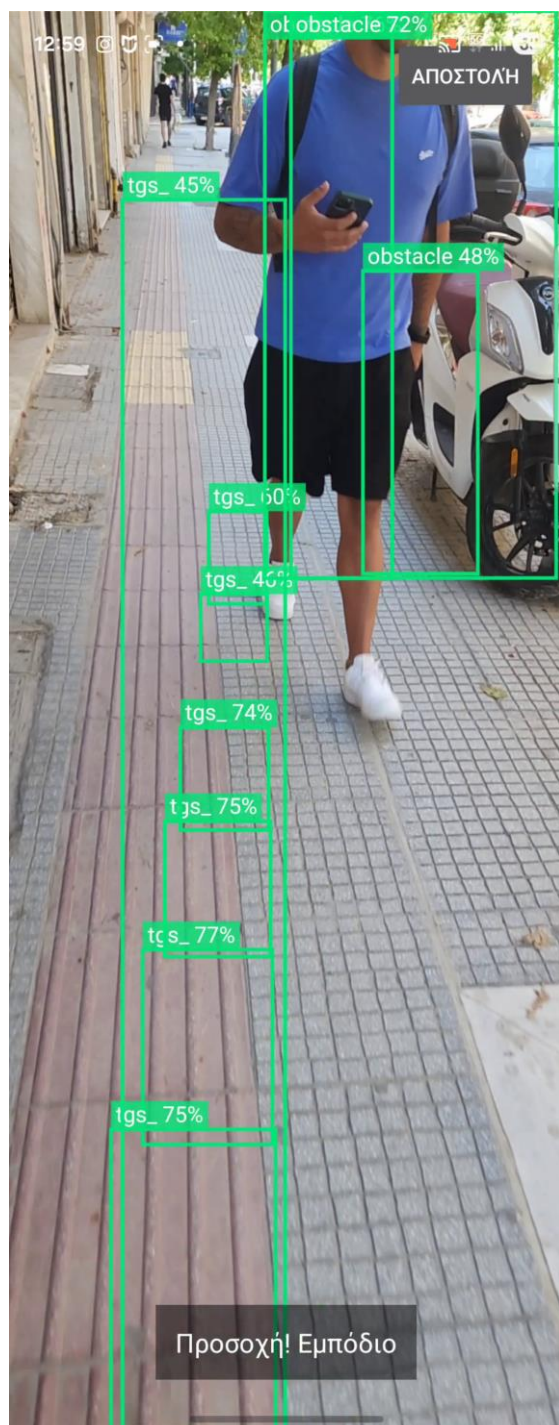


Εικόνα 3.6 Προειδοποίηση ερυθρού σηματοδότη με κλάση "2"

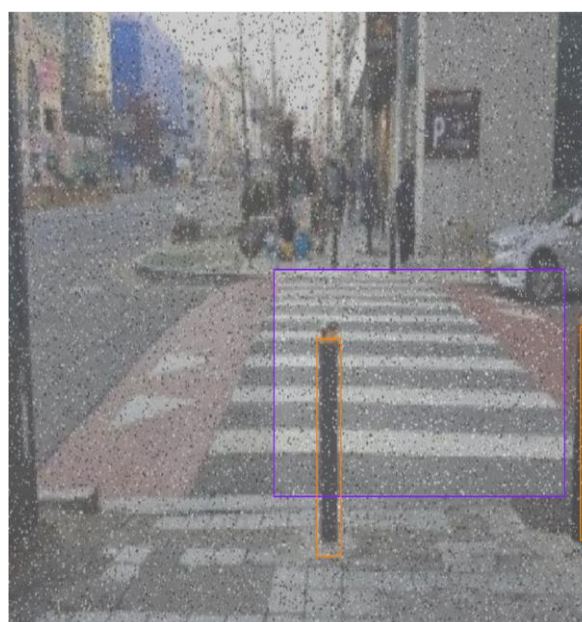


Εικόνα 3.7 Προειδοποίηση πράσινου σηματοδότη με κλάση "0"

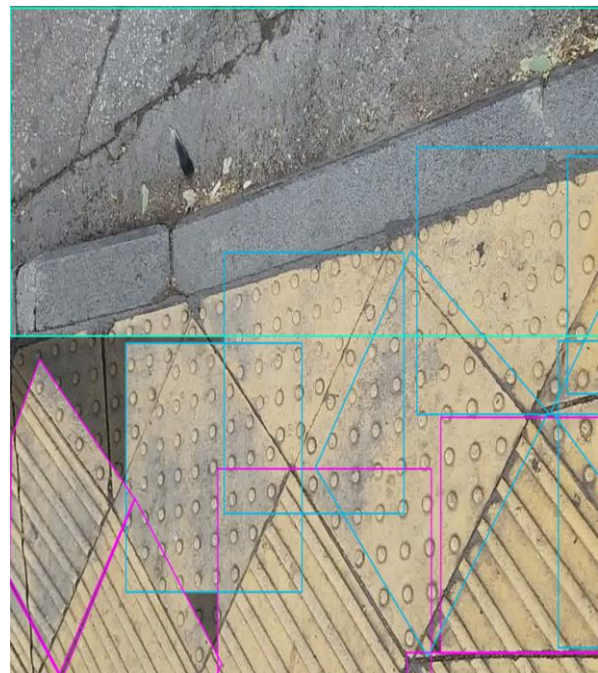
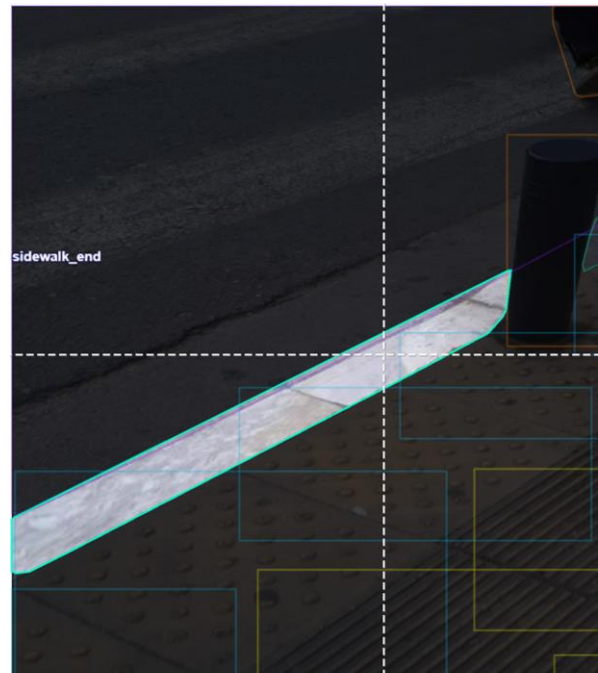
- Τts μηνύματα προειδοποιήσεις



Εικόνα 3.8 Τts μήνυμα για εμπόδια με κλάση "obstacle"



Εικόνα 3.9 Tts μήνυμα για διάβαση πεζών με κλάση "zebracroswalk"



Εικόνα 3.10 Tts μήνυμα για το τέλος πεζοδρομίου με κλάση " sidewalk_end "

Η επισημείωση (annotation) των εικόνων δηλαδή ο σχεδιασμός των πλαισίων οριοθέτησης (bounding boxes) γύρω από κάθε αντικείμενο ενδιαφέροντος και η ανάθεση της αντίστοιχης κλάσης πραγματοποιήθηκε στην πλατφόρμα Roboflow [34]. Το Roboflow χρησιμοποιήθηκε επίσης για τη διαχείριση, την οργάνωση και την τελική εξαγωγή των συνόλων δεδομένων στη μορφή που απαιτεί η αρχιτεκτονική YOLO. Το σύνολο των απτικών πλακιδίων περιλαμβάνει οκτώ κλάσεις. Πέρα από τις κατηγορίες που σχετίζονται με εμπόδια και όρια της διαδρομής (*objects*, *obstacle*, *sidewalk_end*, καθώς και *zebracrosswalk1* για τη διάβαση πεζών), το σύνολο περιλαμβάνει τους βασικούς τύπους απτικών πλακιδίων: ένα γενικό πλακίδιο (*tgs_*) και τους τυποποιημένους τύπους προειδοποίησης (*tgs_warning*), κατεύθυνσης (*tgs_directional*) και υπηρεσίας (*tgs_Service*). Οι τύποι αυτοί αντιστοιχούν στις κατηγορίες απτικών δεικτών επιφάνειας βάδισης που ορίζει το διεθνές πρότυπο ISO 23599 [5]. Το σύνολο των σηματοδοτών περιλαμβάνει έξι κλάσεις, οι οποίες αντιστοιχούν στις καταστάσεις ενός φωτεινού σηματοδότη πεζών: πράσινο (*green*), κόκκινο (*red*), κίτρινο (*yellow*), αντίστροφη μέτρηση σε πράσινη φάση (*countdown_green*), αντίστροφη μέτρηση σε κόκκινη φάση (*countdown_red*), καθώς και την απουσία σηματοδότη (*none*).

Ένα σημείο που αποδείχθηκε κρίσιμο αφορά τη σειρά των κλάσεων. Κατά την εξαγωγή από το Roboflow, οι κλάσεις αριθμούνται με αλφαβητική σειρά και όχι με τη ημαιολογική σειρά που θα ανέμενε κανείς. Έτσι, για το σύνολο των σηματοδοτών, η αντιστοίχιση δεικτών-κλάσεων είναι: *countdown_green* = 0, *countdown_red* = 1, *green* = 2, *none* = 3, *red* = 4, *yellow* = 5. Η ακριβής αυτή σειρά πρέπει να τηρείται αυστηρά και στην πλευρά της εφαρμογής κατά την ερμηνεία των εξόδων του μοντέλου· τυχόν αναντιστοιχία οδηγεί σε εσφαλμένες, και μάλιστα αντεστραμμένες, ανακοινώσεις. Η σημασία του ζητήματος αυτού και ο τρόπος αντιμετώπισής του συζητούνται αναλυτικότερα στο Κεφάλαιο 7. Κατά την εκπαίδευση εφαρμόστηκαν αυτόματα οι προεπιλεγμένες τεχνικές επαύξησης δεδομένων (data augmentation) του πλαισίου εκπαίδευσης της Ultralytics, με στόχο την αύξηση της ποικιλίας των παραδειγμάτων και τη βελτίωση της ικανότητας γενίκευσης του μοντέλου. Οι τεχνικές αυτές περιλαμβάνουν τυχαίες μεταβολές απόχρωσης, κορεσμού και φωτεινότητας (HSV), οριζόντια αναστροφή με πιθανότητα 50%, κλιμάκωση και μετατόπιση, καθώς και τη σύνθετη τεχνική mosaic augmentation, η οποία απενεργοποιείται στις τελευταίες δέκα εποχές της εκπαίδευσης (close mosaic). Για το σύνολο των πλακιδίων, ο διαχωρισμός σε υποσύνολα εκπαίδευσης και επικύρωσης ήταν περίπου 90/10 (2.089/226 εικόνες), χωρίς ξεχωριστό σύνολο ελέγχου (test)· ενώ για το σύνολο των σηματοδοτών ο διαχωρισμός καθορίστηκε από την πηγή προέλευσης του συνόλου δεδομένων. Τα δύο σύνολα δεδομένων αποτέλεσαν τη βάση για την εκπαίδευση των μοντέλων, η οποία περιγράφεται στην Ενότητα 3.3.

3.3 Εκπαίδευση των μοντέλων

Και τα δύο μοντέλα του BlindNav βασίζονται στην αρχιτεκτονική YOLO11n, τη μικρότερη (nano) παραλλαγή της οικογένειας YOLO11[16] [17]. Η επιλογή της nano παραλλαγής, έναντι των μεγαλύτερων και ακριβέστερων παραλλαγών (small, medium κ.λπ.), αποτελεί συνειδητό συμβιβασμό ανάμεσα στην ακρίβεια και την ταχύτητα: για μια εφαρμογή που εκτελεί εξαγωγή προβλέψεων σε πραγματικό χρόνο πάνω σε κινητή συσκευή, η ταχύτητα και το μικρό μέγεθος του μοντέλου είναι εξίσου σημαντικά με την ακρίβεια της ανίχνευσης. Η εκπαίδευση πραγματοποιήθηκε τοπικά, μέσω του framework εκπαίδευσης της Ultralytics (έκδοση 8.4.82, PyTorch 2.12.1 με υποστήριξη CUDA), σε ένα απομονωμένο εικονικό περιβάλλον Python (virtual environment). Η απομόνωση αυτή διασφαλίζει την αναπαραγωγικότητα και την αποφυγή συγκρούσεων μεταξύ των εξαρτήσεων (dependencies) του έργου. Η εκπαίδευση εκτελέστηκε σε κάρτα γραφικών NVIDIA GeForce RTX 2060 (6 GB μνήμης), με κάθε πλήρη εκπαίδευση να διαρκεί περίπου δύο ώρες. Αντί για εκπαίδευση από το μηδέν,

αξιοποιήθηκε η τεχνική της μεταφοράς μάθησης (transfer learning): η διαδικασία εκκίνησε από τα προ εκπαιδευμένα βάρη του YOLO11n στο σύνολο δεδομένων COCO (αρχείο yolo11n.pt), τα οποία μεταφέρθηκαν σχεδόν αυτούσια στο νέο μοντέλο, μόνο το τελικό επίπεδο ανίχνευσης (detection head) αρχικοποιήθηκε εκ νέου, καθώς ο αριθμός των κλάσεων άλλαξε από τις 80 γενικές κατηγορίες του COCO στις 8 ειδικές κατηγορίες του BlindNav. Η προσέγγιση αυτή επιτρέπει στα μοντέλα να αξιοποιήσουν γενικά οπτικά χαρακτηριστικά που έχουν ήδη μάθει από μεγάλα σύνολα δεδομένων, μειώνοντας τόσο τον χρόνο εκπαίδευσης όσο και την ανάγκη για πολύ μεγάλο πλήθος δικών μας παραδειγμάτων. Ως ανάλυση εισόδου (input resolution) ορίστηκε το 640×640 pixels, η οποία αποτελεί την καθιερωμένη τιμή για τα μοντέλα YOLO. Οι υπόλοιπες υπερπαράμετροι της εκπαίδευσης ήταν οι εξής: 100 εποχές (epochs), με μηχανισμό πρόωρης διακοπής (early stopping) αν δεν σημειωνόταν βελτίωση για 20 διαδοχικές εποχές (patience)· μέγεθος παρτίδας (batch size) 16· αρχικός ρυθμός μάθησης (learning rate) 0,01, με ροπή (momentum) 0,937 και απομείωση βαρών (weight decay) 0,0005· ο βελτιστοποιητής (optimizer) επιλέχθηκε αυτόματα από την Ultralytics (optimizer=auto) ο σπόρος τυχαιότητας (seed) ορίστηκε σταθερός στο 0, για λόγους αναπαραγωγιμότητας. Πραγματοποιήθηκαν δύο ξεχωριστές εκπαιδεύσεις (runs), μία για κάθε σύνολο δεδομένων. Το μοντέλο ανίχνευσης απτικών πλακιδίων προέκυψε από την εκπαίδευση train-5, ενώ το μοντέλο ανίχνευσης σηματοδοτών από την εκπαίδευση train-3. Κάθε εκπαίδευση παρήγαγε το αντίστοιχο σύνολο βαρών, το οποίο στη συνέχεια χρησιμοποιήθηκε ως αφετηρία για την εξαγωγή και τη μετατροπή (Ενότητα 3.4). Κατά τη διάρκεια της εκπαίδευσης παρακολουθήθηκαν οι καθιερωμένες μετρικές αξιολόγησης της ανίχνευσης αντικειμένων: η μέση ακρίβεια (mean Average Precision) στα κατώφλια mAP@0.5 και mAP@0.5:0.95, η ακρίβεια (precision) και η ανάκληση (recall), καθώς και οι επιμέρους συναρτήσεις απώλειας (box, classification και distribution focal loss). Οι μετρικές αυτές χρησίμευσαν για την παρακολούθηση της σύγκλισης και την αποφυγή υπερπροσαρμογής (overfitting)· οι τελικές τους τιμές για καθένα από τα δύο μοντέλα παρουσιάζονται και σχολιάζονται στο Κεφάλαιο 5.

3.4 Εξαγωγή και μετατροπή των μοντέλων

Τα μοντέλα που προκύπτουν από την εκπαίδευση (Ενότητα 3.3) βρίσκονται στη μορφή του framework της Ultralytics (PyTorch) και δεν μπορούν να εκτελεστούν απευθείας σε κινητή συσκευή. Για να αξιοποιηθούν από τον διερμηνέα του LiteRT, πρέπει να μετατραπούν σε αρχεία .tflite. Επειδή δεν υπάρχει άμεση, αξιόπιστη διαδρομή από το PyTorch στο TFLite, η μετατροπή γίνεται σε δύο βήματα, με ενδιάμεσο σταθμό το format ONNX. Στο πρώτο βήμα, κάθε εκπαιδευμένο μοντέλο εξάγεται στο ONNX[35], μια ανοιχτή και ανεξάρτητη από το εκάστοτε framework αναπαράσταση νευρωνικών δικτύων. Στο δεύτερο βήμα, το μοντέλο ONNX μετατρέπεται σε .tflite με το εργαλείο onnx2tf [36]. Πέρα από την ίδια τη μετατροπή, το onnx2tf αναλαμβάνει και τη μετάβαση από τη διάταξη ταυνοστών NCHW (που χρησιμοποιεί το PyTorch/ONNX) στη διάταξη NHWC (που απαιτεί το TensorFlow/TFLite), μια μετατροπή που διαφορετικά θα έπρεπε να γίνει χειροκίνητα. Τα τελικά μοντέλα διατηρήθηκαν σε ακρίβεια κινητής υποδιαστολής float32, χωρίς να εφαρμοστεί κβαντισμός (π.χ. σε INT8). Ο πίνακας εξόδου κάθε μοντέλου έχει σχήμα [1, 12, 8400] για το μοντέλο των πλακιδίων και [1, 10, 8400] για το μοντέλο των σηματοδοτών. Η δεύτερη διάσταση αντιστοιχεί στα χαρακτηριστικά κάθε πρόβλεψης: τέσσερις τιμές για τις συντεταγμένες του πλαισίου οριοθέτησης συν το πλήθος των κλάσεων ($4 + 8 = 12$ για τα πλακίδια, $4 + 6 = 10$ για τους σηματοδότες). Η τρίτη διάσταση (8400) αντιστοιχεί στο συνολικό πλήθος των υποψήφιων σημείων ανίχνευσης (anchor points) στις τρεις κλίμακες ανάλυσης του δικτύου, για είσοδο 640×640 .

Ένα σημείο που αποδείχθηκε ιδιαίτερα κρίσιμο κατά την ενσωμάτωση αφορά την κανονικοποίηση των συντεταγμένων. Το onnx2tf παράγει τις συντεταγμένες των πλαισίων οριοθέτησης ήδη

κανονικοποιημένες στο διάστημα $[0, 1]$. Ένα συχνό σφάλμα είναι η εκ νέου διαίρεση των τιμών αυτών με την ανάλυση εισόδου (640), υπό την εσφαλμένη υπόθεση ότι πρόκειται για συντεταγμένες σε pixels· κάτι τέτοιο αλλοιώνει πλήρως τις θέσεις των ανιχνεύσεων. Η αποφυγή αυτής της διπλής κανονικοποίησης ήταν απαραίτητη για τη σωστή χωρική τοποθέτηση των αντικειμένων, και το ζήτημα αναλύεται περαιτέρω στο Κεφάλαιο 5. Τέλος, αξίζει να καταγραφεί μια πρακτική δυσκολία που προέκυψε κατά τη διαμόρφωση του περιβάλλοντος μετατροπής. Το πακέτο `onnxruntime-gpu`, που σχετίζεται με την αλυσίδα εργαλείων του ONNX, παρουσίασε σύγκρουση εξαρτήσεων με το πακέτο `ai-edge-litert` του LiteRT. Η επίλυση απαίτησε τη χειροκίνητη εγκατάσταση των πακέτων στο εικονικό περιβάλλον, ώστε να συνυπάρχουν χωρίς ασυμβατότητες. Με την ολοκλήρωση της μετατροπής, τα δύο μοντέλα `.tflite` ήταν έτοιμα να ενσωματωθούν στην εφαρμογή Android, όπως περιγράφεται στην Ενότητα 3.5.

3.5 Ανάπτυξη της εφαρμογής Android

Η εφαρμογή αναπτύχθηκε στο περιβάλλον Android Studio, σε γλώσσα Kotlin, με όνομα πακέτου `com.example.blindnav`. Το ελάχιστο υποστηριζόμενο επίπεδο API ορίστηκε στο 26 (Android 8.0), επιλογή που, όπως αναλύεται στην Ενότητα 3.7, συνδέεται με τη διαθεσιμότητα των διεπαφών προγραμματισμού δόνησης που χρησιμοποιεί η εφαρμογή. Αρχιτεκτονικά, η εφαρμογή συνδέει τέσσερα βασικά υποσυστήματα σε μια συνεχή ροή: τη λήψη εικόνας, την εξαγωγή προβλέψεων, την επεξεργασία (Ενότητα 3.6) και την ανάδραση (Ενότητα 3.7). Η λήψη εικόνας υλοποιείται με τη βιβλιοθήκη CameraX [37]. Μέσω της περίπτωσης χρήσης ανάλυσης εικόνας (ImageAnalysis), κάθε καρέ της κάμερας γίνεται διαθέσιμο ως αντικείμενο προς επεξεργασία, χωρίς να απαιτείται ο χειρισμός λεπτομερειών χαμηλού επιπέδου της κάμερας. Τα δύο μοντέλα `.tflite` φορτώνονται κατά την εκκίνηση της εφαρμογής στον διερμηνέα του LiteRT (έκδοση TFLite 2.17.0). Έως ότου ολοκληρωθεί η αρχικοποίηση των μοντέλων, η εφαρμογή εμφανίζει μια κατάσταση φόρτωσης, ώστε να μην επιχειρηθεί εξαγωγή προβλέψεων πριν αυτά είναι έτοιμα· οι λεπτομέρειες της οθόνης εκκίνησης και της ένδειξης φόρτωσης περιγράφονται στην Ενότητα 3.8. Ως προς την επιτάχυνση της εξαγωγής προβλέψεων, η εφαρμογή αξιοποιεί τον GPU delegate του TFLite, όταν αυτός υποστηρίζεται από τη συγκεκριμένη συσκευή. Κατά την εκκίνηση πραγματοποιείται έλεγχος συμβατότητας μέσω του `CompatibilityList` και, εφόσον υπάρχει υποστήριξη, τα μοντέλα εκτελούνται μέσω GPU. Σε περίπτωση μη υποστήριξης ή σφάλματος κατά την αρχικοποίηση ή την εκτέλεση, εφαρμόζεται αυτόματη επαναφορά σε εκτέλεση μέσω CPU, ώστε να διατηρείται η λειτουργικότητα της εφαρμογής. Η εξαγωγή προβλέψεων εκτελείται αποκλειστικά στη CPU, η οποία αποδείχθηκε επαρκώς σταθερή και αποδοτική για τα ελαφριά μοντέλα YOLO11n της εφαρμογής.

Επειδή η εξαγωγή προβλέψεων εκτελείται επαναληπτικά και ασύγχρονα σε σχέση με τη ροή της κάμερας, ενώ ο διερμηνέας του TFLite δεν είναι ασφαλής για ταυτόχρονη πρόσβαση, η εφαρμογή διασφαλίζει ότι δεν εκτελούνται επικαλυπτόμενες κλήσεις εξαγωγής προβλέψεων· ο σχετικός μηχανισμός (η σημαία `AtomicBoolean`) περιγράφηκε στην Ενότητα 3.6. Με τη σύνδεση αυτών των υποσυστημάτων, η εφαρμογή Android αποτελεί το περιβάλλον μέσα στο οποίο λειτουργεί ολόκληρη η αλυσίδα ανίχνευσης και ανάδρασης. Οι παράμετροι που ρυθμίζουν τη συμπεριφορά της, καθώς και οι σχεδιαστικές επιλογές για την προσβασιμότητα, παρουσιάζονται στην τελευταία ενότητα του κεφαλαίου (3.8).

3.6 Pipeline ανίχνευσης και επεξεργασίας

Οι ωμές (raw) προβλέψεις που παράγει το μοντέλο σε κάθε καρέ δεν είναι κατάλληλες για άμεση απόδοση στον χρήστη. Κατά τη βάρδιση, η κάμερα κινείται συνεχώς, με αποτέλεσμα οι ανιχνεύσεις να

εμφανίζουν διακυμάνσεις (jitter) από καρέ σε καρέ: ένα πλακίδιο μπορεί να εντοπίζεται στιγμιαία, να χάνεται και να επανεμφανίζεται, ή να μετατοπίζεται ελαφρά η θέση του. Αν κάθε τέτοια στιγμιαία πρόβλεψη μεταφραζόταν αμέσως σε ηχητική ανακοίνωση, ο χρήστης θα δεχόταν έναν κατατρεγμένο ασταθών και συχνά αντικρουόμενων μηνυμάτων. Για τον λόγο αυτό, μεταξύ της εξαγωγής προβλέψεων και της ανάδρασης παρεμβάλλεται ένα pipeline επεξεργασίας, το οποίο μετατρέπει τις ανά καρέ ανιχνεύσεις σε σταθερή, χρονικά συνεπή καθοδήγηση.

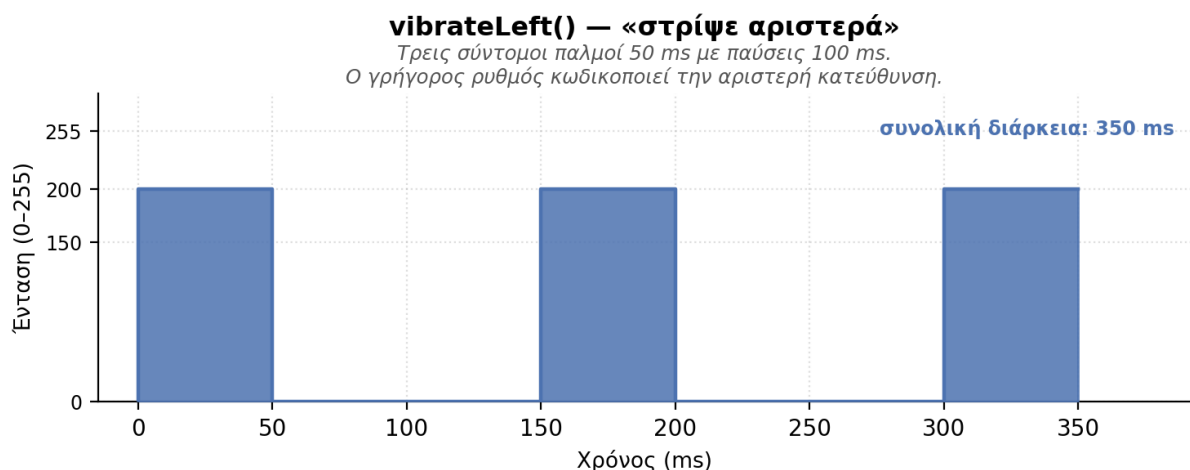
Το πρώτο βήμα είναι η εξαγωγή προβλέψεων. Κάθε καρέ που λαμβάνεται από το CameraX τροφοδοτείται στον διερμηνέα του LiteRT, ο οποίος επιστρέφει έναν πίνακα ανιχνεύσεων με τις συντεταγμένες των πλαισίων οριοθέτησης (bounding boxes) κανονικοποιημένες στο διάστημα [0, 1], μαζί με την κλάση και τη βαθμολογία εμπιστοσύνης (confidence) κάθε ανίχνευσης. Οι ανιχνεύσεις με εμπιστοσύνη κάτω από ένα κατώφλι 0, 45 απορρίπτονται, ώστε να περιορίζονται τα ψευδώς θετικά αποτελέσματα (false positives). Επειδή η εξαγωγή προβλέψεων εκτελείται σε τακτά διαστήματα (κάθε 220 ms) και ο διερμηνέας του TFLite δεν είναι ασφαλής για ταυτόχρονη πρόσβαση (thread-safe), χρησιμοποιείται μια σημαία τύπου AtomicBoolean (isProcessing), η οποία εγγυάται ότι δεν εκκινείται νέα εκτέλεση όσο η προηγούμενη βρίσκεται σε εξέλιξη. Για να μπορεί το σύστημα να καθοδηγεί κατευθυντικά τον χρήστη (π.χ. «αριστερά», «μπροστά», «δεξιά»), η οριζόντια θέση (x) κάθε ανίχνευσης αντιστοιχίζεται σε μια χωρική ζώνη (SpatialZone) ως προς το οπτικό πεδίο. Ο υπολογισμός όμως της ζώνης ξεχωριστά σε κάθε καρέ θα κληρονομούσε ακριβώς την αστάθεια που περιγράφηκε παραπάνω. Την αστάθεια αυτή αντιμετωπίζει το βασικότερο συστατικό του pipeline, ο χρονικός εξομαλυντής (TemporalSmoother). Αντί να αποφασίζει τη ζώνη ανά καρέ, ο εξομαλυντής διατηρεί ένα κυλιόμενο παράθυρο (rolling window) των πέντε τελευταίων καρέ, αποθηκεύοντας για καθένα ένα στιγμιότυπο (FrameSnapshot) με τις ανιχνεύσεις του. Από το σύνολο των στιγμιότυπων του παραθύρου υπολογίζεται ο πραγματικός μέσος όρος των θέσεων x, και από αυτόν προκύπτει μία σταθερή τελική ζώνη. Η επιλογή αυτή μέσος όρος των συνεχών θέσεων x αντί για ψηφοφορία (voting) μεταξύ διακριτών ζωνών ανά καρέ αποδείχθηκε ότι παράγει σαφώς πιο σταθερή χωρική καθοδήγηση κατά την κίνηση του χρήστη, καθώς εξομαλύνει τις στιγμιαίες μετατοπίσεις χωρίς να «κολλάει» σε λανθασμένες αποφάσεις μεμονωμένων καρέ.

Το αποτέλεσμα της επεξεργασίας ενσωματώνεται σε μια δομημένη έξοδο. Ο κεντρικός συντονιστής (DetectionPipeline) παράγει ένα αντικείμενο PipelineResult, το οποίο περιέχει το μήνυμα προς εκφώνηση (TTS), τον βαθμό κρισιμότητας (criticality) της κατάστασης, καθώς και τη ζώνη και την κλάση του εντοπισμένου πλακιδίου. Ο βαθμός κρισιμότητας χρησιμοποιείται στη συνέχεια από τα συστήματα ανάδρασης (Ενότητα 3.7) για να διαφοροποιήσουν την ένταση της απόκρισης, για παράδειγμα, μια προειδοποίηση κινδύνου παρακάμπτει τους συνήθεις περιορισμούς ρυθμού. Όταν σε ένα παράθυρο συνυπάρχουν περισσότεροι από ένας τύποι πλακιδίων, το σύστημα δεν τους ανακοινώνει όλους αδιακρίτως· εφαρμόζει μια ιεραρχία προτεραιότητας, ώστε να αναδεικνύεται πάντοτε ο πιο κρίσιμος για την ασφάλεια. Η σειρά προτεραιότητας είναι: πλακίδιο προειδοποίησης (warning) > κατευθυντήριο (directional) > υπηρεσίας (service) > γενικού τύπου (generic). Έτσι, αν ταυτόχρονα εντοπιστούν ένα κατευθυντήριο και ένα προειδοποιητικό πλακίδιο, προτεραιότητα δίνεται στο δεύτερο, που σηματοδοτεί πιθανό κίνδυνο. Τέλος, για να αποφεύγεται η επαναλαμβανόμενη και κουραστική εκφώνηση της ίδιας πληροφορίας, ένας παρακολουθητής κατάστασης (StateTracker) λειτουργεί σε κατάσταση «μόνο σε αλλαγή» (onlyOnChange): οι ανακοινώσεις σχετικά με τα πλακίδια ενεργοποιούνται μόνο όταν αλλάζει ο τύπος του πλακιδίου και όχι σε κάθε κύκλο επανάληψης όσο ο ίδιος τύπος παραμένει ορατός. Με τον τρόπο αυτό ο χρήστης ενημερώνεται για κάθε ουσιαστική μεταβολή του περιβάλλοντος, χωρίς να δέχεται συνεχείς πανομοιότυπες ειδοποιήσεις. Συνολικά, το pipeline αυτό μετασχηματίζει έναν θορυβώδη ρυθμό ανιχνεύσεων σε ένα καθαρό, ιεραρχημένο και

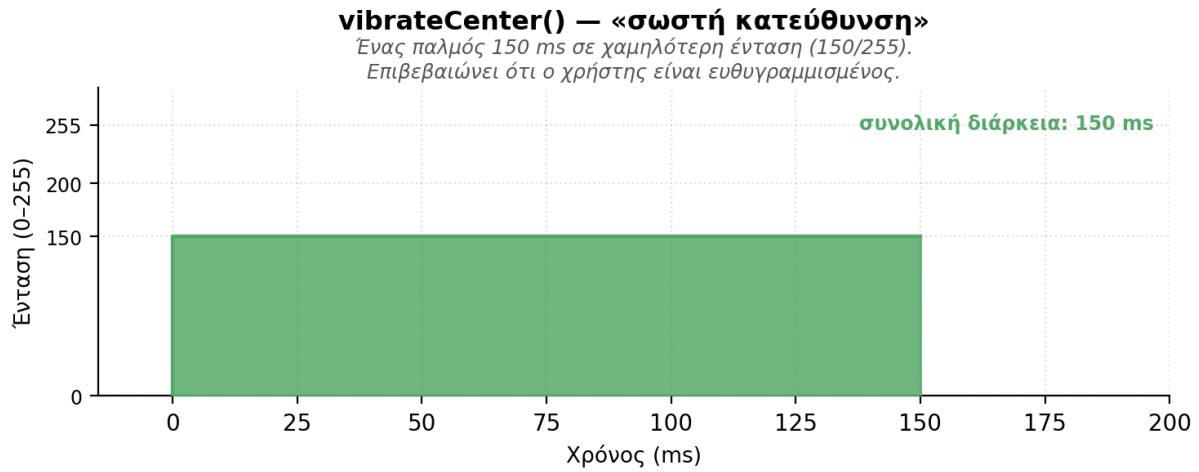
χρονικά σταθερό σύνολο πληροφοριών καθοδήγησης, το οποίο τροφοδοτεί τα συστήματα ηχητικής και απτικής ανάδρασης που περιγράφονται στην Ενότητα 3.7.

3.7 Συστήματα ανάδρασης

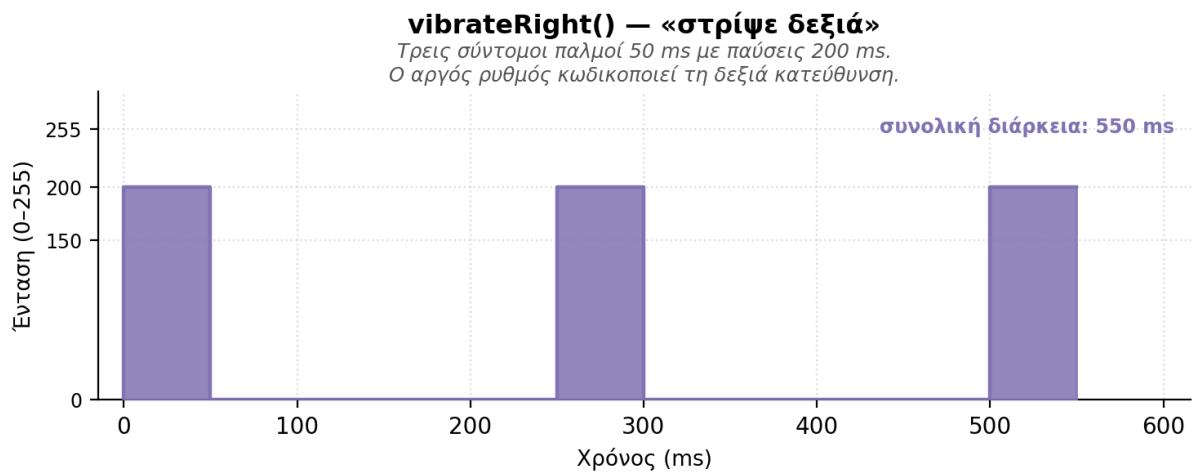
Η δομημένη έξοδος του pipeline (Ενότητα 3.6) μήνυμα, βαθμός κρισιμότητας, ζώνη και κλάση μεταφέρεται στον χρήστη μέσω δύο συμπληρωματικών, μη οπτικών καναλιών: της ηχητικής ανάδρασης (text-to-speech) και της απτικής ανάδρασης (δόνηση). Η ταυτόχρονη αξιοποίηση δύο αισθήσεων αυξάνει την αξιοπιστία της επικοινωνίας· ακόμη κι αν ένα ηχητικό μήνυμα καλυφθεί από τον θόρυβο του δρόμου, η δόνηση εξακολουθεί να ειδοποιεί τον χρήστη. Η ηχητική ανάδραση υλοποιείται με τη μηχανή text-to-speech του Android (TextToSpeech), ρυθμισμένη στην ελληνική γλώσσα [39]. Μέσω αυτής εκφωνούνται στα ελληνικά οι πληροφορίες που παράγει το pipeline: ο τύπος και η κατεύθυνση ενός πλακιδίου ή η κατάσταση ενός φωτεινού σηματοδότη. Το κείμενο κάθε ανακοίνωσης προέρχεται απευθείας από το πεδίο μηνύματος του PipelineResult, ώστε η λογική της απόφασης (τι θα ειπωθεί) να παραμένει διαχωρισμένη από τον μηχανισμό της εκφώνησης (πώς θα ειπωθεί). Η απτική ανάδραση διαχειρίζεται από έναν ειδικό διαχειριστή δόνησης (HapticManager), ο οποίος αξιοποιεί το Vibrator API του Android [40]. Ο διαχειριστής αυτός δεν παράγει απλώς δονήσεις, αλλά ενσωματώνει τρεις σχεδιαστικές αρχές που στοχεύουν στην ισορροπία ανάμεσα στην επαρκή ενημέρωση και τη μη υπερφόρτωση του χρήστη. Πρώτον, εφαρμόζεται περιορισμός ρυθμού (rate limiting): μεταξύ δύο διαδοχικών δονήσεων μεσολαβεί ελάχιστο διάστημα 150 ms, ώστε ο χρήστης να μη δέχεται συνεχείς δονήσεις που θα τον αποπροσανατόλιζαν ή θα οδηγούσαν σε απευαισθητοποίηση. Δεύτερον, για τις κρίσιμες προειδοποιήσεις προβλέπεται μια σημαία εξαναγκασμού (force flag): όταν ο βαθμός κρισιμότητας του PipelineResult είναι υψηλός για παράδειγμα, σε προειδοποιητικό πλακίδιο κινδύνου η δόνηση παρακάμπτει τον περιορισμό των 150 ms και αποδίδεται άμεσα, εξασφαλίζοντας ότι μια επικίνδυνη κατάσταση δεν θα «σιωπήσει» εξαιτίας του χρονικού ορίου. Τρίτον, χρησιμοποιούνται διακριτά μοτίβα δόνησης· ειδικότερα, οι προειδοποιήσεις πλακιδίων αποδίδονται με ένα ξεχωριστό μοτίβο (vibrateTileWarning), το οποίο επιτρέπει στον χρήστη να αναγνωρίζει το είδος της ειδοποίησης από την αίσθηση και μόνο, χωρίς να απαιτείται η ακρόαση του ηχητικού μηνύματος.



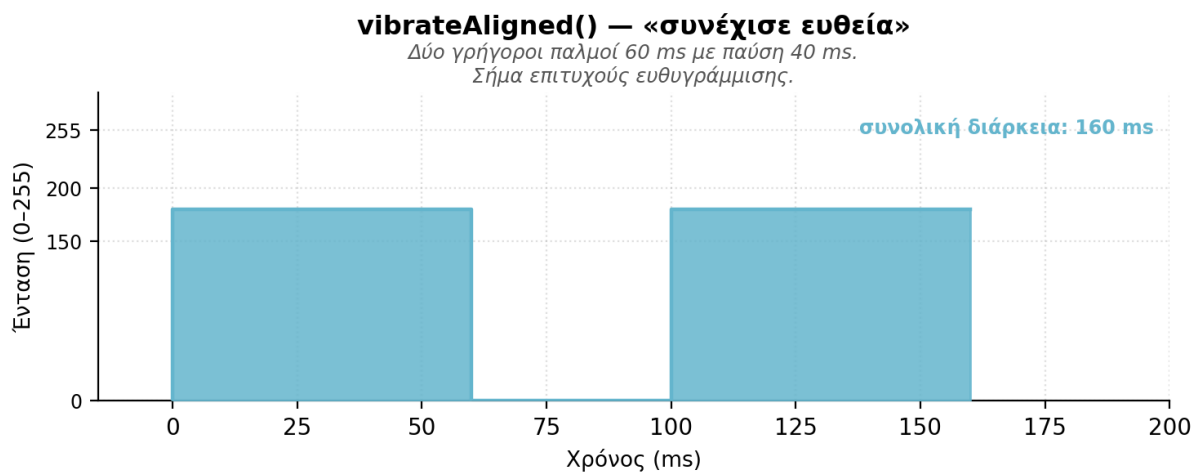
Εικόνα 3.11 Μοτίβο δόνησης αριστερής κατεύθυνσης (vibrateLeft)



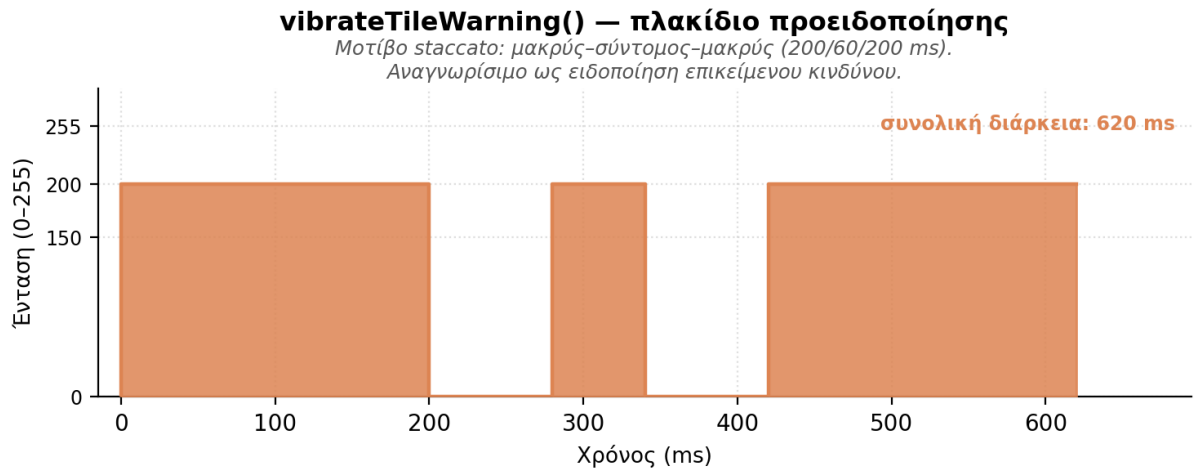
Εικόνα 3.12 Μοτίβο δόνησης κεντρικής ευθυγράμμισης (vibrateCenter)



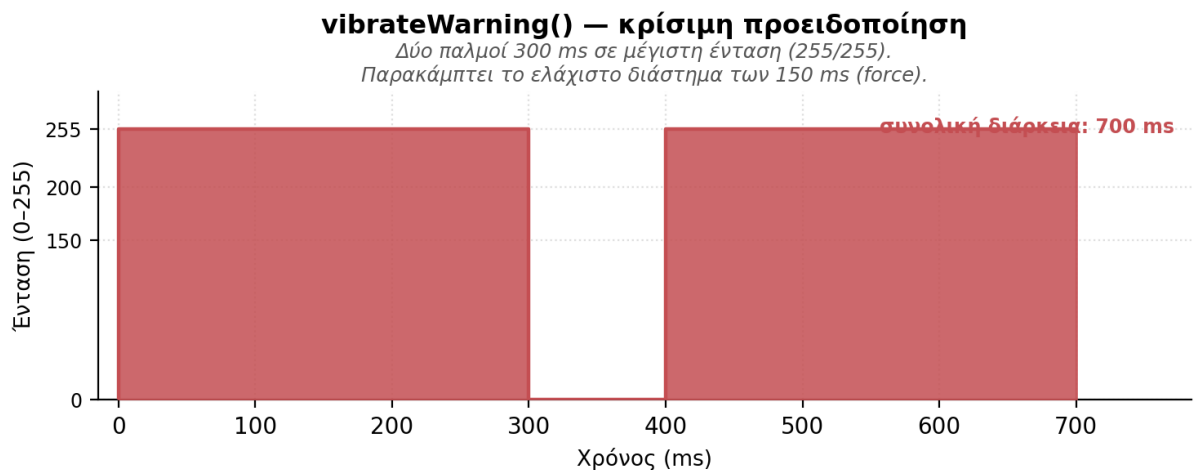
Εικόνα 3.13 Μοτίβο δόνησης δεξιάς κατεύθυνσης (vibrateRight)



Εικόνα 3.14 Μοτίβο δόνησης επιτυχούς ευθυγράμμισης (vibrateAligned)



Εικόνα 3.15 Μοτίβο δόνησης πλακιδίου προειδοποίησης (vibrateTileWarning)



Εικόνα 3.16 Μοτίβο δόνησης κρίσιμης προειδοποίησης (vibrateWarning)

Αξίζει να σημειωθεί ότι η απόδοση δομημένων μοτίβων δόνησης μέσω της κλάσης `VibrationEffect` [41] εισήχθη στο Android API level 26. Το γεγονός αυτό συνάδει με το ελάχιστο υποστηριζόμενο επίπεδο (minSdk 26) της εφαρμογής και αποτελεί έναν από τους λόγους της επιλογής του. Πέρα από τα δύο κανάλια εξόδου, στην ανάδραση εντάσσεται και ο μηχανισμός εναλλαγής λειτουργίας, ο οποίος καθορίζει ποιο από τα δύο μοντέλα είναι ενεργό κάθε στιγμή. Η εφαρμογή λειτουργεί σε δύο καταστάσεις: την ανίχνευση απτικών πλακιδίων, όταν η συσκευή είναι στραμμένη προς το έδαφος, και την ανίχνευση φωτεινών σηματοδοτών, όταν η συσκευή είναι στραμμένη προς τα εμπρός ή προς τα πάνω. Η εναλλαγή γίνεται αυτόματα με βάση την κλίση (pitch) της συσκευής, όπως αυτή προκύπτει από τους αισθητήρες προσανατολισμού: όταν η κλίση περνά το κατώφλι των -70° , δηλαδή όταν ο χρήστης σηκώνει το τηλέφωνο προς τον σηματοδότη, το σύστημα μεταβαίνει στη λειτουργία ανίχνευσης σηματοδοτών. Ο σχεδιασμός αυτός είναι κρίσιμος για την προσβασιμότητα: ο χρήστης μπορεί να αλλάξει το τι ανιχνεύεται απλώς αλλάζοντας τον τρόπο με τον οποίο κρατά τη συσκευή, χωρίς να χρειάζεται καμία οπτική αλληλεπίδραση με το περιβάλλον εργασίας. Συνδυαστικά, η ελληνική ηχητική εκφώνηση, η διαβαθμισμένη απτική ανάδραση και η εναλλαγή λειτουργίας με βάση

τον προσανατολισμό της κινητής συσκευής συγκροτούν μια διεπαφή που λειτουργεί εξ ολοκλήρου χωρίς όραση και χωρίς χειροκίνητο χειρισμό. Οι συγκεκριμένες αριθμητικές τιμές και τα κατώφλια που διέπουν τη συμπεριφορά αυτή συγκεντρώνονται και αιτιολογούνται στην Ενότητα 3.8.

3.8 Ρυθμίσεις και σχεδιαστικές επιλογές προσβασιμότητας της εφαρμογής

Η συμπεριφορά της εφαρμογής διέπεται από ένα σύνολο παραμέτρων, αρκετές από τις οποίες αναφέρθηκαν στις προηγούμενες ενότητες. Στην παρούσα ενότητα συγκεντρώνονται και αιτιολογούνται οι βασικότερες από αυτές, καθώς και οι σχεδιαστικές επιλογές που αφορούν άμεσα την προσβασιμότητα. Τρεις παράμετροι είναι καθοριστικές για τη λειτουργία σε πραγματικό χρόνο. Η πρώτη είναι το διάστημα εξαγωγής προβλέψεων (inference interval), που ορίστηκε στα 220 ms, με στόχο περίπου τέσσερις με πέντε αναλύσεις ανά δευτερόλεπτο. Η τιμή αυτή αποτελεί σχεδιαστικό συμβιβασμό ανάμεσα στην ανταπόκριση αρκετά συχνή ώστε να εντοπίζονται έγκαιρα οι αλλαγές του περιβάλλοντος κατά τη βάδιση και στο υπολογιστικό κόστος, ώστε να αποφεύγεται η περιττή επεξεργασία και η υπερβολική κατανάλωση ενέργειας. Ο πραγματικά επιτευχθείς ρυθμός, ο οποίος περιορίζεται και από τον χρόνο εκτέλεσης του μοντέλου στη συσκευή, μετρήθηκε πειραματικά και παρουσιάζεται στην Ενότητα 5.2. Η δεύτερη παράμετρος είναι το κατώφλι εμπιστοσύνης (confidence threshold), που ορίστηκε στο 0,45· ανιχνεύσεις με χαμηλότερη βαθμολογία απορρίπτονται (Ενότητα 3.6). Η τιμή αυτή επιλέχθηκε ελαφρώς υψηλότερα από το σημείο βέλτιστης ισορροπίας precision–recall που προέκυψε από την αξιολόγηση των μοντέλων (Κεφάλαιο 5), ώστε να ευνοείται η ακρίβεια (precision) έναντι της ανάκλησης (recall): για ένα βοήθημα ασφάλειας, οι συχνές ψευδείς ειδοποιήσεις είναι ιδιαίτερα επιζήμιες, καθώς υπονομεύουν την εμπιστοσύνη του χρήστη στο σύστημα. Ταυτόχρονα, η τιμή διατηρήθηκε αρκετά χαμηλή ώστε να μη χάνονται πραγματικά πλακίδια και σηματοδότες.

Η τρίτη παράμετρος είναι το κατώφλι κλίσης (pitch threshold) των -70° , που καθορίζει την εναλλαγή μεταξύ των δύο λειτουργιών (Ενότητα 3.7). Η συγκεκριμένη τιμή επιλέχθηκε ώστε η φυσική κίνηση της ανύψωσης του τηλεφώνου προς έναν σηματοδότη να ενεργοποιεί αξιόπιστα τη λειτουργία σηματοδοτών, ενώ η στόχευση προς το έδαφος να διατηρεί τη λειτουργία ανίχνευσης πλακιδίων, χωρίς ανεπιθύμητες εναλλαγές κατά τη συνήθη βάδιση. Ως προς τη διεπαφή, η εφαρμογή διαθέτει μια οθόνη εκκίνησης (splash screen), υλοποιημένη με τη βιβλιοθήκη `androidx.core:core-splashscreen`. Η οθόνη αυτή έχει φόντο σε έντονο κίτρινο χρώμα (#FFC107) και ένα εικονίδιο με πλέγμα κουκκίδων, το οποίο παραπέμπει οπτικά στους ανάγλυφους κόκκους των προειδοποιητικών πλακιδίων. Η επιλογή ενός χρώματος υψηλής ορατότητας δεν είναι μόνο αισθητική: εξυπηρετεί τους χρήστες με μερική όραση, για τους οποίους η υψηλή αντίθεση διευκολύνει την αναγνώριση. Αμέσως μετά την εκκίνηση, και έως ότου ολοκληρωθεί η φόρτωση των μοντέλων, εμφανίζεται μια ένδειξη φόρτωσης (loading overlay), όπως αναφέρθηκε στην Ενότητα 3.5, ώστε ο χρήστης να μην αλληλεπιδρά με την εφαρμογή προτού αυτή είναι έτοιμη. Πέρα από τα επιμέρους αυτά στοιχεία, ολόκληρος ο σχεδιασμός της εφαρμογής διέπεται από την αρχή της λειτουργίας χωρίς όραση. Η πληροφορία μεταφέρεται μέσω ήχου και δόνησης και όχι μέσω της οθόνης, η εναλλαγή λειτουργίας γίνεται με τον προσανατολισμό της συσκευής και όχι με χειρισμό στοιχείων στην οθόνη, και απαιτείται η ελάχιστη δυνατή ακριβής αλληλεπίδραση αφής. Ο σχεδιασμός αυτός καθιστά την εφαρμογή αξιοποιήσιμη από τον τελικό της χρήστη χωρίς την ανάγκη οπτικής επιβεβαίωσης.

3.9 Επίλογος

Το κεφάλαιο αυτό περιέγραψε ολόκληρη τη διαδρομή υλοποίησης του BlindNav: τη συλλογή και επισημείωση των δύο συνόλων δεδομένων, την εκπαίδευση των δύο μοντέλων YOLO11n, τη μετατροπή τους σε μορφή TFLite μέσω ONNX, την ενσωμάτωσή τους στην εφαρμογή Android, το pipeline χρονικής εξομάλυνσης και ιεραρχημένης απόφασης, τα συστήματα ηχητικής και απτικής ανάδρασης, και τέλος τις ρυθμίσεις και τις σχεδιαστικές επιλογές προσβασιμότητας που συνδέουν όλα τα παραπάνω σε μια συλλογική εμπειρία χρήσης χωρίς όραση. Η στατική περιγραφή της μεθοδολογίας που δόθηκε εδώ αποκτά δυναμική διάσταση στο επόμενο κεφάλαιο, όπου παρουσιάζεται η αρχιτεκτονική του συστήματος σε λειτουργία, η ροή της πληροφορίας σε πραγματικό χρόνο και το πειραματικό πρωτόκολλο με το οποίο αξιολογήθηκε.

Κεφάλαιο 4ο: Αρχιτεκτονική του Συστήματος και Πειραματικό Πρωτόκολλο.

Το προηγούμενο κεφάλαιο περιέγραψε τη διαδικασία κατασκευής του συστήματος: πώς συλλέχθηκαν τα δεδομένα, πώς εκπαιδεύτηκαν και μετατράπηκαν τα μοντέλα και πώς ενσωματώθηκαν στην εφαρμογή. Το παρόν κεφάλαιο μεταθέτει την οπτική από την κατασκευή στη λειτουργία: παρουσιάζει την αρχιτεκτονική του ολοκληρωμένου συστήματος, τη ροή της πληροφορίας κατά την εκτέλεση σε πραγματικό χρόνο και, τέλος, το πρωτόκολλο με το οποίο αξιολογήθηκε πειραματικά η συμπεριφορά του.

4.1 Αρχιτεκτονική του συστήματος

Το BlindNav είναι δομημένο ως μια αλυσίδα διακριτών υποσυστημάτων, καθένα από τα οποία έχει σαφώς οριοθετημένη ευθύνη και επικοινωνεί με το επόμενο μέσω καλά ορισμένων δομών δεδομένων. Η οργάνωση αυτή δεν είναι μόνο θέμα τάξης στον κώδικα: επιτρέπει την ανεξάρτητη τροποποίηση ή αντικατάσταση κάθε τμήματος, για παράδειγμα, την αναβάθμιση ενός μοντέλου ή την προσθήκη ενός νέου καναλιού ανάδρασης χωρίς να επηρεάζεται η υπόλοιπη αρχιτεκτονική.

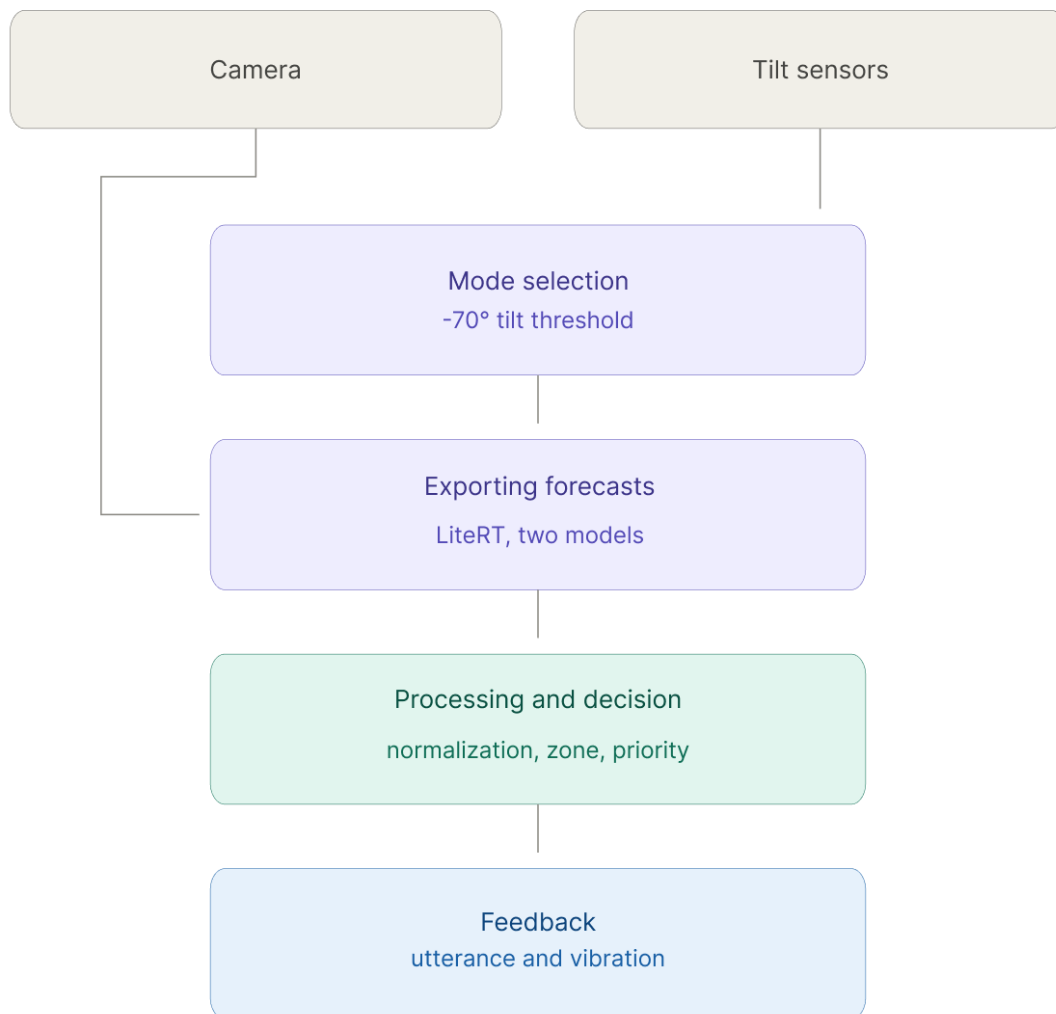
Το σύστημα αποτελείται από πέντε επίπεδα:

- **Το επίπεδο εισόδου** περιλαμβάνει τους αισθητήρες της συσκευής: την κάμερα, η οποία μέσω του CameraX τροφοδοτεί συνεχώς την αλυσίδα με καρέ, και τους αισθητήρες προσανατολισμού, οι οποίοι παρέχουν την κλίση (pitch) της συσκευής. Οι δύο αυτές πηγές είναι ισότιμες ως προς τη σημασία τους: η κάμερα καθορίζει τι βλέπει το σύστημα και ο προσανατολισμός ποιο μοντέλο θα ερμηνεύσει αυτό που βλέπει.
- **Το επίπεδο επιλογής λειτουργίας** αξιοποιεί την κλίση της συσκευής για να καθορίσει την ενεργή κατάσταση. Όταν η κλίση υπερβαίνει το κατώφλι των -70° , ενεργοποιείται η λειτουργία ανίχνευσης φωτεινών σηματοδοτών· διαφορετικά, το σύστημα παραμένει σε λειτουργία ανίχνευσης απτικών πλακιδίων. Η απόφαση αυτή καθορίζει ποιο από τα δύο μοντέλα θα κληθεί για το τρέχον καρέ.
- **Το επίπεδο εξαγωγής προβλέψεων** περιλαμβάνει τον διερμηνέα του LiteRT και τα δύο μοντέλα.tflite. Δέχεται ένα προ επεξεργασμένο καρέ και επιστρέφει τις ωμές ανιχνεύσεις: πλαίσια οριοθέτησης με κανονικοποιημένες συντεταγμένες, κλάσεις και βαθμολογίες εμπιστοσύνης. Το επίπεδο αυτό είναι σκόπιμα «αδαές» ως προς το νόημα των ανιχνεύσεων· δεν γνωρίζει τι σημαίνει ένα προειδοποιητικό πλακίδιο, παρά μόνο ότι εντόπισε ένα αντικείμενο συγκεκριμένης κλάσης σε συγκεκριμένη θέση.
- **Το επίπεδο επεξεργασίας και απόφασης** είναι εκείνο που αποδίδει το νόημα. Εδώ ανήκουν το φιλτράρισμα με βάση το κατώφλι εμπιστοσύνης, ο χρονικός εξομαλυντής που σταθεροποιεί τις ανιχνεύσεις σε ένα κυλιόμενο παράθυρο, η αντιστοίχιση σε χωρική ζώνη, η ιεραρχία προτεραιότητας των τύπων πλακιδίων και ο παρακολουθητής κατάστασης που αποτρέπει τις επαναλαμβανόμενες ανακοινώσεις. Το επίπεδο αυτό, που αναλύθηκε στην Ενότητα 3.6, μετατρέπει έναν θορυβώδη ρυθμό ανιχνεύσεων σε μία σαφή απόφαση, ενσωματωμένη στη δομή PipelineResult.
- **Τέλος, το επίπεδο ανάδρασης** το οποίο μεταφράζει την απόφαση σε αντιληπτά ερεθίσματα, μέσω της ελληνικής λειτουργίας text-to-speech και του διαχειριστή δόνησης. Ο βαθμός κρισιμότητας που συνοδεύει κάθε απόφαση καθορίζει την ένταση της απόκρισης, επιτρέποντας στις κρίσιμες προειδοποιήσεις να παρακάμπτουν τους συνήθεις χρονικούς περιορισμούς.

Δύο χαρακτηριστικά της αρχιτεκτονικής αξίζει να τονιστούν. Το πρώτο είναι ότι ολόκληρη η ακολουθία εκτελείται τοπικά στη συσκευή· δεν υπάρχει διακομιστής, ούτε εξωτερική υπηρεσία, ούτε εξάρτηση

από σύνδεση δικτύου. Το δεύτερο είναι ο διαχωρισμός ανάμεσα στο «τι ανιχνεύεται» και στο «τι ανακοινώνεται»: μεταξύ του επιπέδου εξαγωγής προβλέψεων και του επιπέδου ανάδρασης παρεμβάλλεται πάντοτε το επίπεδο απόφασης, το οποίο εγγυάται ότι ο χρήστης δεν εκτίθεται ποτέ σε ωμές, ασταθείς προβλέψεις.

Η συνολική αυτή δομή απεικονίζεται στο Σχήμα 4.1, ενώ η ακριβής ροή της πληροφορίας μέσα από τα επίπεδα, για ένα μεμονωμένο καρέ, αναλύεται στην Ενότητα 4.2.

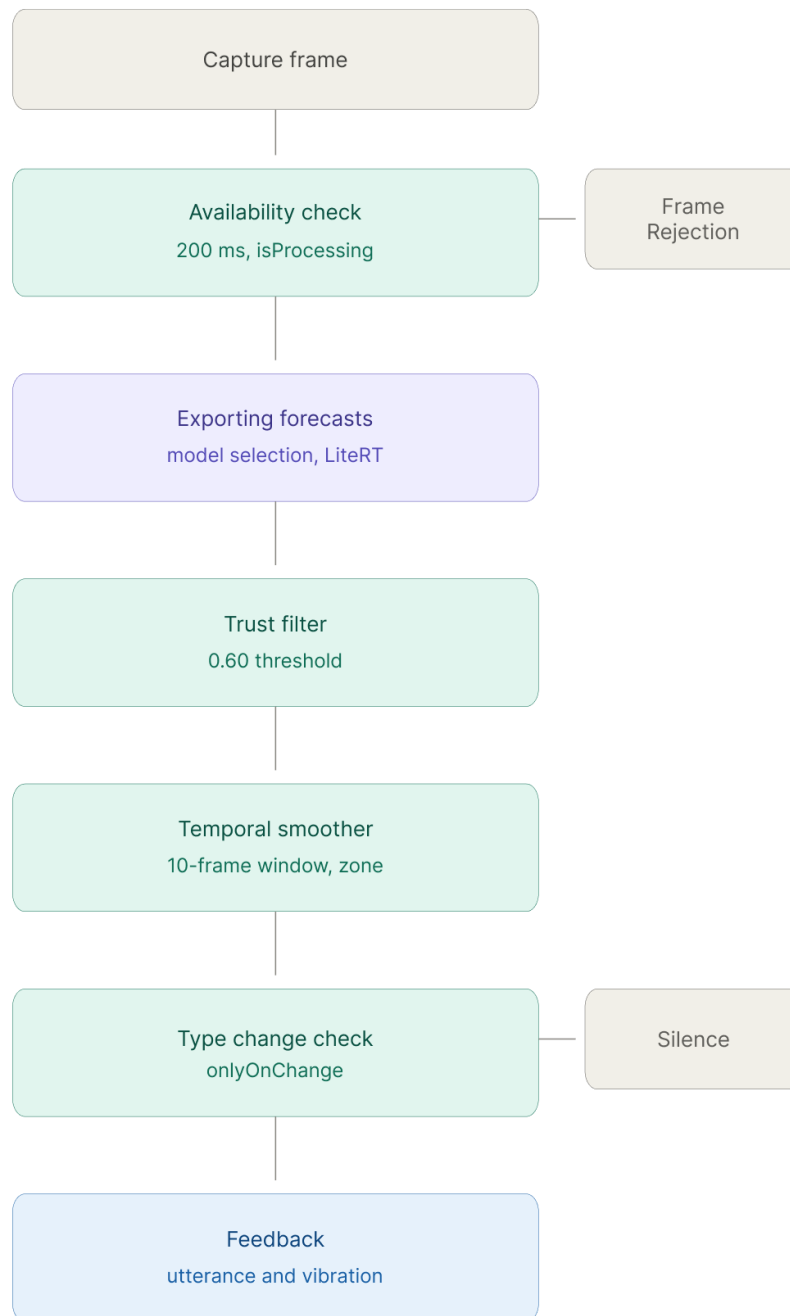


Εικόνα 4.1 Διάγραμμα αρχιτεκτονικής των πέντε επιπέδων

4.2 Ροή της πληροφορίας

Η αρχιτεκτονική της προηγούμενης ενότητας περιγράφει τη στατική δομή του συστήματος. Η παρούσα ενότητα παρακολουθεί τη δυναμική του λειτουργία, ακολουθώντας τη διαδρομή ενός μεμονωμένου καρέ από τη στιγμή που καταγράφεται από την κάμερα έως τη στιγμή που παράγεται ένα ερέθισμα προς τον χρήστη. Η διαδρομή ξεκινά με τη λήψη ενός καρέ από το CameraX. Πριν όμως το καρέ οδηγηθεί στο μοντέλο, εξετάζεται η διαθεσιμότητα του συστήματος: εάν εκκρεμεί προηγούμενη εξαγωγή προβλέψεων, ή εάν δεν έχει παρέλθει το διάστημα των 220 ms από την προηγούμενη ανάλυση, το

καρέ απορρίπτεται χωρίς επεξεργασία. Η απόρριψη αυτή δεν αποτελεί απώλεια πληροφορίας, αλλά συνειδητό μηχανισμό ρύθμισης του ρυθμού: η κάμερα παράγει καρέ ταχύτερα από όσο απαιτείται ή είναι σκόπιμο να αναλυθούν. Εφόσον το καρέ γίνει δεκτό, καθορίζεται η ενεργή λειτουργία με βάση την τρέχουσα κλίση της συσκευής και επιλέγεται το αντίστοιχο μοντέλο. Το καρέ προσαρμόζεται στη μορφή εισόδου που αναμένει το μοντέλο (ανάλυση 640×640) και τροφοδοτείται στον διερμηνέα του LiteRT, ο οποίος επιστρέφει τον πίνακα των ωμών προβλέψεων. Ο πίνακας αυτός αποκωδικοποιείται: για κάθε υποψήφια ανίχνευση εξάγονται οι συντεταγμένες του πλαισίου οι οποίες, όπως τονίστηκε στην Ενότητα 3.4, είναι ήδη κανονικοποιημένες η κλάση με την υψηλότερη βαθμολογία και η αντίστοιχη εμπιστοσύνη. Οι ανιχνεύσεις με εμπιστοσύνη κάτω από 0,45 απορρίπτονται, ενώ όσες παραμένουν συγκροτούν το στιγμιότυπο του τρέχοντος καρέ. Το στιγμιότυπο δεν οδηγεί άμεσα σε ανακοίνωση· εισάγεται στο κυλιόμενο παράθυρο του χρονικού εξομαλυντή. Από το σύνολο των στιγμιότυπων του παραθύρου υπολογίζεται η σταθεροποιημένη χωρική ζώνη και επιλέγεται, με βάση την ιεραρχία προτεραιότητας, ο επικρατέστερος τύπος πλακιδίου. Το αποτέλεσμα ενσωματώνεται στη δομή PipelineResult, μαζί με το αντίστοιχο μήνυμα και τον βαθμό κρισιμότητας. Πριν από την τελική απόδοση, μεσολαβεί ένα ακόμη φίλτρο: ο παρακολουθητής κατάστασης συγκρίνει το νέο αποτέλεσμα με το προηγούμενο και επιτρέπει την ανακοίνωση μόνο εφόσον έχει μεταβληθεί ο τύπος του πλακιδίου. Έτσι, ένα μεγάλο μέρος των αποτελεσμάτων δεν καταλήγει ποτέ σε ερέθισμα, κάτι που είναι επιθυμητό: η σιωπή, όταν δεν υπάρχει νέα πληροφορία, αποτελεί σχεδιαστική επιλογή και όχι αστοχία. Στην περίπτωση που η ανακοίνωση εγκριθεί, το μήνυμα εκφωνείται μέσω της ελληνικής μηχανής text-to-speech και, παράλληλα, αποδίδεται η αντίστοιχη δόνηση. Εάν ο βαθμός κρισιμότητας είναι υψηλός, η δόνηση παρακάμπτει τον περιορισμό των 150 ms και αποδίδεται άμεσα με το ειδικό μοτίβο προειδοποίησης. Ο κύκλος στη συνέχεια επαναλαμβάνεται με το επόμενο διαθέσιμο καρέ. Η ροή αυτή, με τα σημεία απόφασης και απόρριψης που τη χαρακτηρίζουν, αποτυπώνεται στο Σχήμα 4.2



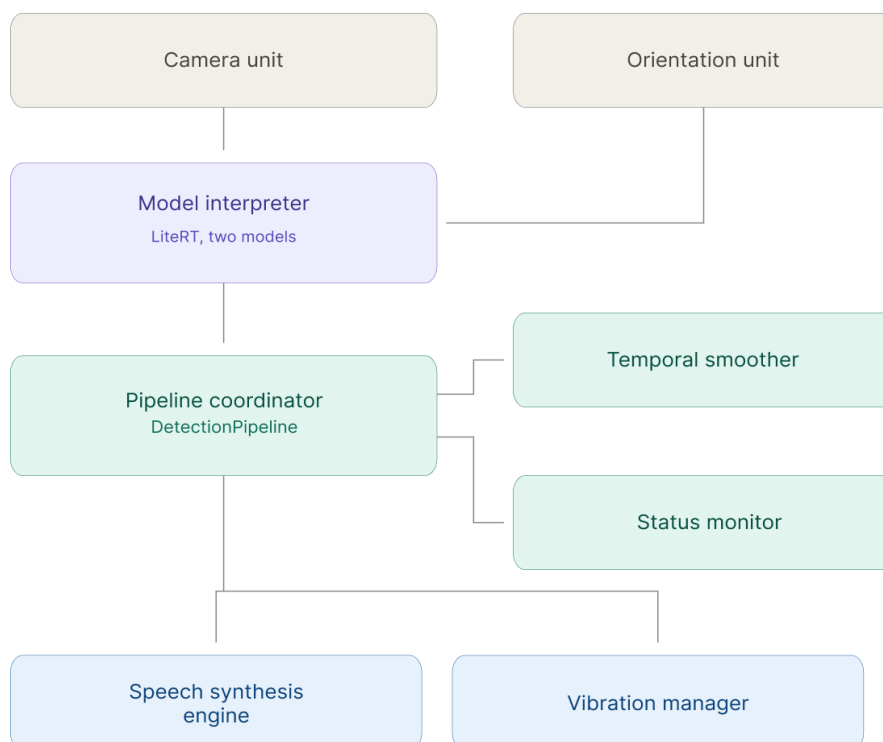
Εικόνα 4.2 Διάγραμμα ροής δεδομένων για ένα καρέ, με τα σημεία απόφασης

4.3 Μοντελοποίηση Δομής, Ροής και Συμπεριφοράς με UML

Για την τυποποιημένη αποτύπωση της δομής και της συμπεριφοράς του συστήματος χρησιμοποιείται η γλώσσα UML, στην τρέχουσα επίσημη έκδοσή της 2.5.1 [42]. Παρουσιάζονται τρία διαγράμματα, καθένα από τα οποία φωτίζει διαφορετική όψη του συστήματος: το διάγραμμα συνιστωσών αποδίδει τη στατική οργάνωση, το διάγραμμα ακολουθίας τη χρονική αλληλουχία των κλήσεων και το διάγραμμα

καταστάσεων τη μετάβαση ανάμεσα στις δύο λειτουργίες. Τα τρία διαγράμματα είναι συμπληρωματικά: περιγράφουν το ίδιο σύστημα από τρεις διαφορετικές οπτικές γωνίες και, από κοινού, τεκμηριώνουν σε τυποποιημένη μορφή όσα αναλύθηκαν περιγραφικά στις Ενότητες 4.1 και 4.2.

Το πρώτο διάγραμμα (component diagram) αποτυπώνει τη στατική οργάνωση της εφαρμογής και τις εξαρτήσεις μεταξύ των βασικών μονάδων της: της μονάδας κάμερας, η οποία τροφοδοτεί την αλυσίδα με καρέ μέσω του CameraX, του διερμηνέα των μοντέλων LiteRT, του χρονικού εξομαλυντή, του συντονιστή του pipeline, του παρακολουθητή κατάστασης, της μηχανής εκφώνησης και του διαχειριστή δόνησης. Οι εξαρτήσεις είναι μονόδρομες και ακολουθούν τη σειρά των πέντε επιπέδων: καμία μονάδα δεν γνωρίζει τις εσωτερικές λεπτομέρειες της επόμενης, παρά μόνο τη δομή δεδομένων που ανταλλάσσουν. Ο συντονιστής του pipeline λειτουργεί ως το κεντρικό σημείο σύζευξης, καθώς είναι η μοναδική μονάδα που επικοινωνεί τόσο με το επίπεδο εξαγωγής προβλέψεων όσο και με το επίπεδο ανάδρασης. Το διάγραμμα καθιστά έτσι ορατό τον διαχωρισμό ευθυνών που περιγράφηκε στην Ενότητα 4.1 και εξηγεί γιατί η αντικατάσταση ενός μοντέλου ή η προσθήκη ενός νέου καναλιού ανάδρασης δεν επηρεάζει την υπόλοιπη αρχιτεκτονική.

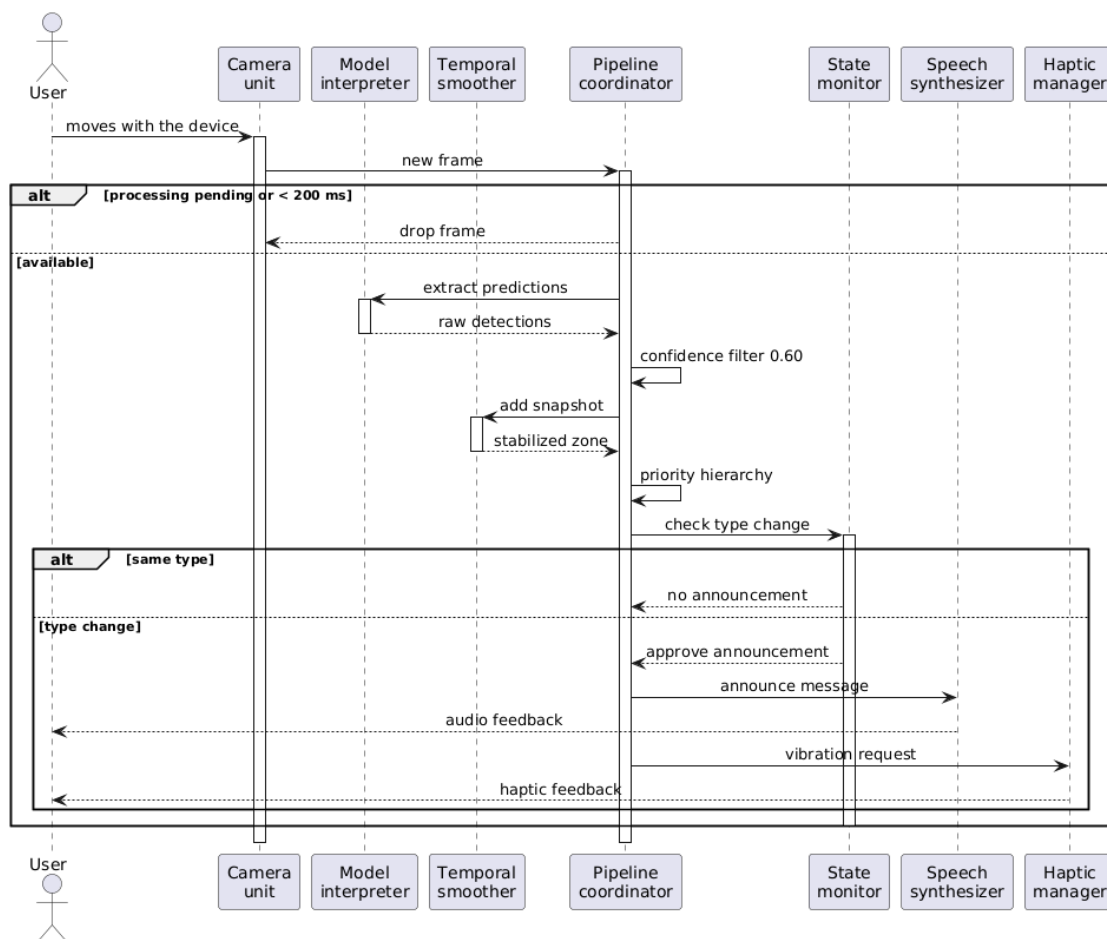


Εικόνα 4.3 Διάγραμμα συνιστωσών

Το διάγραμμα ακολουθίας (sequence diagram) αποτυπώνει τη χρονική αλληλουχία των κλήσεων για την επεξεργασία ενός καρέ, από τη λήψη έως την απόδοση της ανάδρασης. Σε αντίθεση με το προηγούμενο, δεν εστιάζει στο ποιες μονάδες υπάρχουν, αλλά στο πότε και με ποια σειρά επικοινωνούν μεταξύ τους. Στο διάγραμμα διακρίνονται τα διαδοχικά σημεία απόφασης που περιγράφηκαν αναλυτικά στην Ενότητα 4.2: ο έλεγχος διαθεσιμότητας και το χρονικό διάστημα των 220 ms, η επιλογή του ενεργού μοντέλου με βάση την κλίση της συσκευής, η προσαρμογή του καρέ

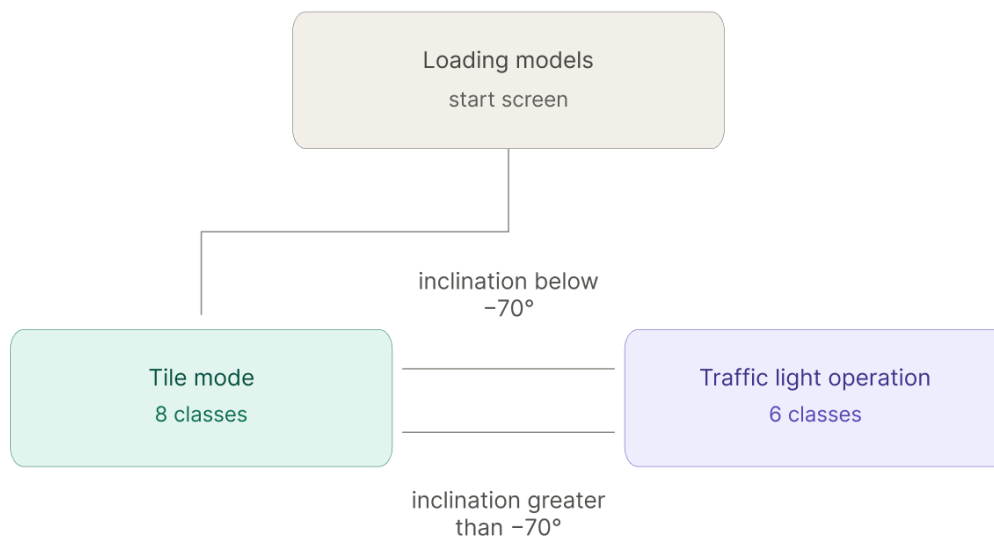
στην ανάλυση εισόδου 640×640, το φιλτράρισμα των ανιχνεύσεων με κατώφλι εμπιστοσύνης 0, 45 και, τέλος, ο έλεγχος του παρακολουθητή κατάστασης πριν από την εκφώνηση. Αξίζει να σημειωθεί ότι το διάγραμμα περιλαμβάνει ρητά και τις διαδρομές τερματισμού, δηλαδή τις περιπτώσεις στις οποίες το καρέ απορρίπτεται ή η ανακοίνωση δεν εγκρίνεται· οι διαδρομές αυτές αντιστοιχούν στη μεγάλη πλειονότητα των κύκλων εκτέλεσης και αποτυπώνουν τη σχεδιαστική επιλογή της σιωπής όταν δεν υπάρχει νέα πληροφορία.

Figure 4.4 — Processing a frame



Εικόνα 4.4 Διάγραμμα ακολουθίας

Το διάγραμμα καταστάσεων (state diagram) αποτυπώνει τις δύο λειτουργίες του συστήματος, ανίχνευση πλακιδίων και ανίχνευση σηματοδοτών, καθώς και τις μεταβάσεις μεταξύ τους. Οι δύο καταστάσεις είναι αμοιβαία αποκλειόμενες: σε κάθε χρονική στιγμή μόνο ένα από τα δύο μοντέλα είναι ενεργό, γεγονός που περιορίζει το υπολογιστικό κόστος και αποτρέπει αντικρουόμενες ανακοινώσεις. Η μετάβαση ελέγχεται αποκλειστικά από την κλίση της συσκευής, με κατώφλι τις -70° : όταν ο χρήστης ανυψώνει τη συσκευή προς τον σηματοδότη και η κλίση υπερβεί το κατώφλι, το σύστημα μεταβαίνει στη λειτουργία ανίχνευσης σηματοδοτών, ενώ όταν την κατεβάζει ξανά προς το έδαφος επιστρέφει στη λειτουργία πλακιδίων. Καμία από τις δύο μεταβάσεις δεν απαιτεί ενέργεια του χρήστη στην οθόνη, στοιχείο καθοριστικό για την προσβασιμότητα της εφαρμογής, όπως θα φανεί και στο σενάριο χρήσης της επόμενης ενότητας.



Εικόνα 4.5 Διάγραμμα καταστάσεων

4.4 Σενάριο χρήσης

Για να γίνει αντιληπτή η συνολική συμπεριφορά του συστήματος, παρουσιάζεται ένα ενδεικτικό σενάριο χρήσης, το οποίο συνδέει τα επιμέρους υποσυστήματα σε μια ενιαία εμπειρία.

Ο χρήστης βαδίζει σε πεζοδρόμιο κρατώντας τη συσκευή στραμμένη προς το έδαφος. Το σύστημα βρίσκεται σε λειτουργία ανίχνευσης πλακιδίων και εντοπίζει σταθερά έναν διάδρομο κατευθυντήριων πλακιδίων ελαφρώς αριστερά του οπτικού πεδίου· εκφωνείται η αντίστοιχη ένδειξη κατεύθυνσης. Καθώς ο χρήστης συνεχίζει, η ίδια ένδειξη δεν επαναλαμβάνεται, χάρη στον παρακολουθητή κατάστασης, ενώ η ζώνη ενημερώνεται σταδιακά ώστε να τον διατηρεί πάνω στον διάδρομο. Πλησιάζοντας σε διάβαση, τα κατευθυντήρια πλακίδια δίνουν τη θέση τους σε προειδοποιητικά. Η αλλαγή του τύπου ενεργοποιεί άμεσα νέα ανακοίνωση, ενώ ο αυξημένος βαθμός κρισιμότητας προκαλεί δόνηση με το ειδικό μοτίβο προειδοποίησης, παρακάμπτοντας τον χρονικό περιορισμό. Ο χρήστης αντιλαμβάνεται ότι βρίσκεται σε σημείο που απαιτεί προσοχή και σταματά. Στη συνέχεια ανυψώνει τη συσκευή προς τον σηματοδότη. Μόλις η κλίση περάσει το κατώφλι των -70° , το σύστημα μεταβαίνει αυτόματα σε λειτουργία ανίχνευσης σηματοδοτών και αρχίζει να εκφωνεί την κατάσταση του φαναριού. Όταν αυτή αλλάξει σε πράσινο, ο χρήστης ενημερώνεται και διασχίζει τη διάβαση. Κατεβάζοντας ξανά τη συσκευή προς το έδαφος, το σύστημα επιστρέφει στη λειτουργία πλακιδίων και η καθοδήγηση συνεχίζεται.

Το σενάριο αναδεικνύει ότι, καθ' όλη τη διάρκεια της διαδρομής, ο χρήστης δεν χρειάστηκε να αγγίξει την οθόνη ή να επιλέξει κάποια λειτουργία· η αλλαγή συμπεριφοράς προέκυψε αποκλειστικά από τον φυσικό τρόπο με τον οποίο κρατούσε τη συσκευή.

4.5 Πειραματικός σχεδιασμός

Η αξιολόγηση του συστήματος στοχεύει να απαντήσει σε δύο ερωτήματα: αν τα μοντέλα ανιχνεύουν αξιόπιστα τους στόχους τους, και αν η εφαρμογή, ως σύνολο, παρέχει έγκαιρη και ορθή καθοδήγηση σε πραγματικές συνθήκες. Για τον σκοπό αυτό, η αξιολόγηση διαρθρώνεται σε δύο επίπεδα: την

αξιολόγηση των μοντέλων επί των δεδομένων επικύρωσης και την αξιολόγηση της εφαρμογής στο πεδίο.

Στο πρώτο επίπεδο, η απόδοση κάθε μοντέλου μετράται με τις καθιερωμένες μετρικές της ανίχνευσης αντικειμένων. Η ακρίβεια (precision) εκφράζει το ποσοστό των ανιχνεύσεων που ήταν ορθές, ενώ η ανάκληση (recall) το ποσοστό των πραγματικών αντικειμένων που εντοπίστηκαν. Η μέση ακρίβεια (average precision, AP) συνοψίζει τη σχέση των δύο μεγεθών, και ο μέσος όρος της στο σύνολο των κλάσεων δίνει τη μετρική mAP. Οι δύο πλέον διαδεδομένες παραλλαγές της είναι το mAP@0.5, σύμφωνα με το πρωτόκολλο του PASCAL VOC [32], όπου μια ανίχνευση θεωρείται ορθή εφόσον η επικάλυψή της με το πραγματικό πλαίσιο (Intersection over Union) υπερβαίνει το 0, 5· και το αυστηρότερο mAP@0.5:0.95, σύμφωνα με το πρωτόκολλο του COCO [33], το οποίο υπολογίζει τον μέσο όρο σε ένα εύρος κατωφλίων επικάλυψης. Στην παρούσα εργασία αναφέρεται το mAP@0.5, το οποίο και παρέχει το εργαλείο αξιολόγησης που χρησιμοποιήθηκε (Κεφάλαιο 5).

Στο δεύτερο επίπεδο, το σύστημα δοκιμάζεται σε πραγματικό αστικό περιβάλλον. Τα σενάρια δοκιμών περιλαμβάνουν βάδιση σε συνεχόμενο διάδρομο απτικών πλακιδίων όδευσης, προσέγγιση και διάσχιση διάβασης πεζών, καθώς και εναλλαγή σε λειτουργία σηματοδότη με αναγνώριση και των τριών καταστάσεών του (κόκκινο, πράσινο, κίτρινο, συμπεριλαμβανομένης της αντίστροφης μέτρησης). Οι δοκιμές πραγματοποιήθηκαν σε πεζοδρόμια και διαβάσεις του αστικού ιστού της Θεσσαλονίκης, σε δύο διαδοχικούς γύρους (9 και 20 Αυγούστου 2026), με τη χρήση τριών συσκευών (Xiaomi 2412DPC0AG, Samsung SM-F936B, Samsung SM-A566B) σε επτά συνολικά συνεδρίες. Οι συνθήκες φωτισμού ήταν σταθερά καλές ηλιοφάνεια ή αίθριος καιρός και στις δύο ημέρες, με θερμοκρασίες 24–39 °C ενώ οι συνεδρίες καλύπτουν τόσο απογευματινές ώρες (16:00–17:30) όσο και μία βραδινή, κοντά στο ηλιοβασίλεμα (20:30). Δεν πραγματοποιήθηκε συστηματική δοκιμή σε συνθήκες βροχής, χαμηλού φωτισμού ή έντονης παρουσίας πεζών και οχημάτων, περιορισμός που καταγράφεται στην Ενότητα 7.6.

Κατά τις δοκιμές πεδίου καταγράφονται: ο χρόνος εξαγωγής προβλέψεων ανά καρέ και ο ρυθμός καρέ (FPS), ώστε να επιβεβαιωθεί ότι η λειτουργία σε πραγματικό χρόνο είναι εφικτή· η ορθότητα των ανακοινώσεων, δηλαδή αν το εκφωνούμενο μήνυμα αντιστοιχεί στην πραγματική κατάσταση του περιβάλλοντος· ο χρονισμός τους, δηλαδή αν η ειδοποίηση δίνεται αρκετά έγκαιρα ώστε να είναι αξιοποιήσιμη· και, τέλος, τα ψευδώς θετικά και ψευδώς αρνητικά αποτελέσματα, με ιδιαίτερη έμφαση στα δεύτερα, καθώς η μη ανίχνευση ενός προειδοποιητικού πλακιδίου ή ενός κόκκινου σηματοδότη έχει σοβαρότερες συνέπειες από μια περιττή ειδοποίηση.

Το πρωτόκολλο εκτέλεσης κάθε πειράματος περιλαμβάνει την προετοιμασία (επιλογή σημείου, καταγραφή συνθηκών, εκκίνηση της εφαρμογής και αναμονή έως την ολοκλήρωση της φόρτωσης των μοντέλων) και την εκτέλεση της διαδρομής με ταυτόχρονη διπλή καταγραφή: αφενός αυτόματα, μέσω του ενσωματωμένου μηχανισμού καταγραφής (MetricsLogger, Παράρτημα Α.1), ο οποίος αποθηκεύει σε αρχείο κάθε εξαγωγή προβλέψεων με χρονοσήμανση, χρόνο εκτέλεσης, ανιχνεύσεις και τυχόν εκφωνηθέν μήνυμα· αφετέρου χειροκίνητη, μέσω σήμανσης εδάφους αναφοράς από τον χειριστή, ο οποίος πατούσε τα φυσικά πλήκτρα έντασης της συσκευής τη στιγμή που παρατηρούσε πραγματικό πλακίδιο ή διαπίστωνε εσφαλμένη ανακοίνωση, χωρίς να απαιτείται οπτική επαφή με την οθόνη. Οι δύο ροές καταγραφής συγχρονίζονται εκ των υστέρων μέσω κοινής χρονοσήμανσης, ώστε να προκύψουν οι μετρικές ορθότητας που παρουσιάζονται στο Κεφάλαιο 5. Τα αποτελέσματα παρουσιάζονται στο Κεφάλαιο 5.

4.6 Επίλογος

Το κεφάλαιο αυτό μετέθεσε την οπτική από την κατασκευή του συστήματος στη λειτουργία του. Περιγράφηκε η αρχιτεκτονική των πέντε επιπέδων είσοδος, επιλογή λειτουργίας, εξαγωγή προβλέψεων, επεξεργασία, απόφαση, ανάδραση και η ακριβής ροή της πληροφορίας για ένα μεμονωμένο καρέ, με έμφαση στα σημεία απόρριψης και απόφασης που διέπουν τη συμπεριφορά του συστήματος. Τρία διαγράμματα UML απέδωσαν οπτικά τη στατική δομή, τη χρονική αλληλουχία και τις μεταβάσεις κατάστασης, ενώ ένα ενδεικτικό σενάριο χρήσης ανέδειξε πώς τα επιμέρους υποσυστήματα συνθέτουν μια ενιαία εμπειρία χωρίς οπτική αλληλεπίδραση. Το κεφάλαιο έκλεισε με το πειραματικό πρωτόκολλο βάσει του οποίου αξιολογήθηκε το σύστημα, τόσο σε επίπεδο μοντέλων όσο και σε πραγματικές συνθήκες πεδίου. Τα αποτελέσματα της αξιολόγησης αυτής παρουσιάζονται στο επόμενο κεφάλαιο

Κεφάλαιο 5ο: Οδηγός Χρήσης Λογισμικού

Ο παρών οδηγός περιγράφει τη χρήση της εφαρμογής BlindNav από την οπτική του τελικού χρήστη. Σε αντίθεση με τα προηγούμενα κεφάλαια, που περιέγραφαν την εσωτερική λειτουργία του συστήματος, ο οδηγός αυτός εστιάζει αποκλειστικά σε όσα αντιλαμβάνεται ο χρήστης κατά τη χρήση δηλαδή σε ό, τι ακούει και αισθάνεται, καθώς η εφαρμογή έχει σχεδιαστεί να λειτουργεί εξ ολοκλήρου χωρίς όραση.

Σύνοψη σε τέσσερα βήματα: (1) άγγιγμα στο εικονίδιο, αναμονή για το μήνυμα «Η εφαρμογή είναι έτοιμη» (2) κράτημα της συσκευής προς το έδαφος για ανίχνευση πλακιδίων, ή προς τον σηματοδότη για ανίχνευση φαναριού ή εναλλαγή γίνεται μόνη της (3) η αργή, επαναλαμβανόμενη δόνηση σημαίνει «είσαι πάνω στη σωστή διαδρομή» ή «είσαι ευθυγραμμισμένος με τον σηματοδότη» (4) σιωπή στη δόνηση σημαίνει απώλεια εντοπισμού μικρή διόρθωση πορείας αρκεί για να επανέλθει.

5.1 Εκκίνηση της εφαρμογής

Η εφαρμογή ξεκινά με ένα μόνο άγγιγμα στο εικονίδιό της, όπως κάθε εφαρμογή Android· δεν απαιτείται καμία περαιτέρω πλοήγηση σε μενού ή ρυθμίσεις. Κατά την εκκίνηση εμφανίζεται για λίγο η οθόνη εκκίνησης (Ενότητα 3.8), και στη συνέχεια η εφαρμογή περνά αυτόματα σε λειτουργία κάμερας. Όσο τα δύο μοντέλα φορτώνονται στη μνήμη, η εφαρμογή παραμένει σε κατάσταση αναμονής. Μόλις η φόρτωση ολοκληρωθεί, εκφωνείται το μήνυμα «Η εφαρμογή είναι έτοιμη», το οποίο αποτελεί την ένδειξη ότι ο χρήστης μπορεί να ξεκινήσει τη χρήση.

5.2 Η διεπαφή

Η οθόνη της εφαρμογής δείχνει αποκλειστικά την εικόνα της κάμερας σε πραγματικό χρόνο, χωρίς κουμπιά, μενού ή άλλα στοιχεία ελέγχου. Η απουσία αυτή δεν είναι παράλειψη αλλά συνειδητή σχεδιαστική επιλογή: αφού ο βασικός χρήστης δεν βλέπει την οθόνη, οποιοδήποτε στοιχείο θα απαιτούσε οπτική αλληλεπίδραση θα ήταν άχρηστο. Η μόνη πληροφορία στην οθόνη είναι ένα μικρό κείμενο κατάστασης, χρήσιμο μόνο σε τρίτο πρόσωπο (π.χ. συνοδό ή βλέποντα παρατηρητή), όχι στον ίδιο τον χρήστη.

5.3 Λειτουργία περπατήματος

Κρατώντας τη συσκευή στραμμένη προς το έδαφος, η εφαρμογή βρίσκεται αυτόματα σε λειτουργία ανίχνευσης απτικών πλακιδίων, γεγονός που επιβεβαιώνεται ηχητικά. Όσο η κάμερα εντοπίζει σταθερά απτικό πλακίδιο όδευσης μπροστά στον χρήστη, η συσκευή παράγει μια αργή, επαναλαμβανόμενη δόνηση. Η δόνηση αυτή λειτουργεί ως συνεχής επιβεβαίωση: όσο γίνεται αισθητή, ο χρήστης γνωρίζει ότι παραμένει πάνω στη διαδρομή. Αν ο χρήστης αποκλίνει και το πλακίδιο βγει εκτός του οπτικού πεδίου της κάμερας, η δόνηση σταματά αμέσως σήμα στον χρήστη να διορθώσει την πορεία του, μέχρι να εντοπιστεί ξανά η όδευση και η δόνηση να επανέλθει. Όταν εντοπιστεί πλακίδιο προειδοποίησης, η συνήθης αυτή δόνηση αντικαθίσταται από το ξεχωριστό μοτίβο προειδοποίησης (Ενότητα 3.7), το οποίο ο χρήστης μαθαίνει να αναγνωρίζει ως ένδειξη επικείμενου κινδύνου ή αλλαγής π.χ. προσέγγιση διάβασης ανεξάρτητα από το αν πρόλαβε να ακούσει το αντίστοιχο ηχητικό μήνυμα.

5.4 Λειτουργία σηματοδότη

Σηκώνοντας τη συσκευή προς τα εμπρός ή προς τα πάνω, ώστε η κλίση της να ξεπεράσει το κατώφλι των -70° (Ενότητα 3.8), η εφαρμογή μεταβαίνει αυτόματα σε λειτουργία σηματοδότη· η αλλαγή αυτή επίσης ανακοινώνεται ηχητικά. Ο χρήστης δεν χρειάζεται να θυμάται κάποια χειρονομία ή ακολουθία ενεργειών· αρκεί ο φυσικός τρόπος με τον οποίο κρατά κανείς το τηλέφωνο όταν θέλει να «δει» κάτι μπροστά του.

Στη λειτουργία αυτή, η εφαρμογή ακολουθεί την ίδια λογική συνεχούς επιβεβαίωσης: όσο ο σηματοδότης παραμένει εντοπισμένος στο οπτικό πεδίο, ο χρήστης λαμβάνει σταθερή ανάδραση για την κατάστασή του (κόκκινο, πράσινο, αντίστροφη μέτρηση)· αν χάσει την ευθυγράμμιση, η ανάδραση σταματά, μέχρι να επανευθυγραμμιστεί.

5.5 Σύνοψη ενδείξεων

Ο ακόλουθος πίνακας συγκεντρώνει τις ηχητικές και απτικές ενδείξεις της εφαρμογής, ως γρήγορη αναφορά.

Κατάσταση	Ηχητική ένδειξη	Απτική ένδειξη
Φόρτωση κατά την εκκίνηση	(καμία, αναμονή)	-
Η εφαρμογή είναι έτοιμη	«Η εφαρμογή είναι έτοιμη»	-
Είσοδος σε λειτουργία περπατήματος	«Λειτουργία περπατήματος»	-
Σταθερός εντοπισμός πλακιδίου όδευσης	ανακοίνωση τύπου/κατεύθυνσης (μόνο σε αλλαγή)	αργή, επαναλαμβανόμενη δόνηση όσο διαρκεί ο εντοπισμός
Πλακίδιο προειδοποίησης	«Προσοχή, πλακίδια προειδοποίησης»	ξεχωριστό, διακριτό μοτίβο δόνησης
Απώλεια εντοπισμού πλακιδίου	(σιωπή)	καμία δόνηση
Είσοδος σε λειτουργία σηματοδότη	«Λειτουργία φαναριού»	-
Σταθερός εντοπισμός σηματοδότη	ανακοίνωση κατάστασης (κόκκινο / πράσινο / αντίστροφη μέτρηση)	συνεχής ανάδραση όσο διαρκεί η ευθυγράμμιση
Απώλεια ευθυγράμμισης σηματοδότη	(σιωπή)	καμία δόνηση
Εντοπισμός εμποδίου	«Προσοχή! Εμπόδιο»	δόνηση υψηλής προτεραιότητας

Πίνακας 5.1 Σύνοψη ηχητικών και απτικών ενδείξεων

5.6 Συμπεριφορά σε μη ιδανικές συνθήκες

Η εφαρμογή δεν απαιτεί ιδανικές συνθήκες φωτισμού για να λειτουργήσει. Σε ανεπίσημες δοκιμές πεδίου, η αναγνώριση πλακιδίων παρέμεινε λειτουργική τόσο σε συνθήκες χαμηλού φωτισμού, όπως τη νύχτα, με την παρουσία έστω ενός σημείου φωτισμού του δρόμου, όσο και σε περιπτώσεις όπου το πλακίδιο είχε χαμηλή χρωματική αντίθεση ως προς το γύρω πεζοδρόμιο. Δεν παρατηρήθηκε αντίστοιχη λειτουργικότητα σε συνθήκες πλήρους σκότους, όπου η απουσία οποιασδήποτε πηγής φωτός εμποδίζει την κάμερα να συλλέξει αξιοποιήσιμη εικόνα.

Οι παρατηρήσεις αυτές δεν βασίζονται σε συστηματική μέτρηση, σε αντίθεση με τα αποτελέσματα του Κεφαλαίου 5, αλλά καταγράφονται εδώ ως πρακτική ένδειξη του εύρους συνθηκών υπό τις οποίες η εφαρμογή παραμένει χρηστική.

5.7 Πρακτικές υπενθυμίσεις

Όπως τονίστηκε στις Ενότητες 5.4 και 6.3, το BlindNav είναι σχεδιασμένο ως συμπληρωματικό βοήθημα και όχι ως υποκατάστατο του λευκού μπαστουνιού, του σκύλου οδηγού ή της εκπαίδευσης προσανατολισμού και κινητικότητας. Ειδικότερα:

- Η ανακοίνωση πράσινης ένδειξης σηματοδότη δεν πρέπει να εκλαμβάνεται ως αυτόματη άδεια διάσχισης· συνιστάται πάντα η επιβεβαίωση με ακουστική εκτίμηση της κυκλοφορίας.
- Η απουσία ανακοίνωσης εμποδίου δεν αποτελεί εγγύηση ελεύθερης διαδρομής, καθώς η αναγνώριση εμποδίων παραμένει η ασθενέστερη λειτουργία του συστήματος (Ενότητα 7.1.1).
- Συνιστάται η εξοικείωση με τα διαφορετικά μοτίβα δόνησης σε ασφαλές, γνωστό περιβάλλον πριν από τη χρήση της εφαρμογής σε άγνωστη διαδρομή.

Κεφάλαιο 6ο: Αξιολόγηση Χρηστών (UEQ-S)

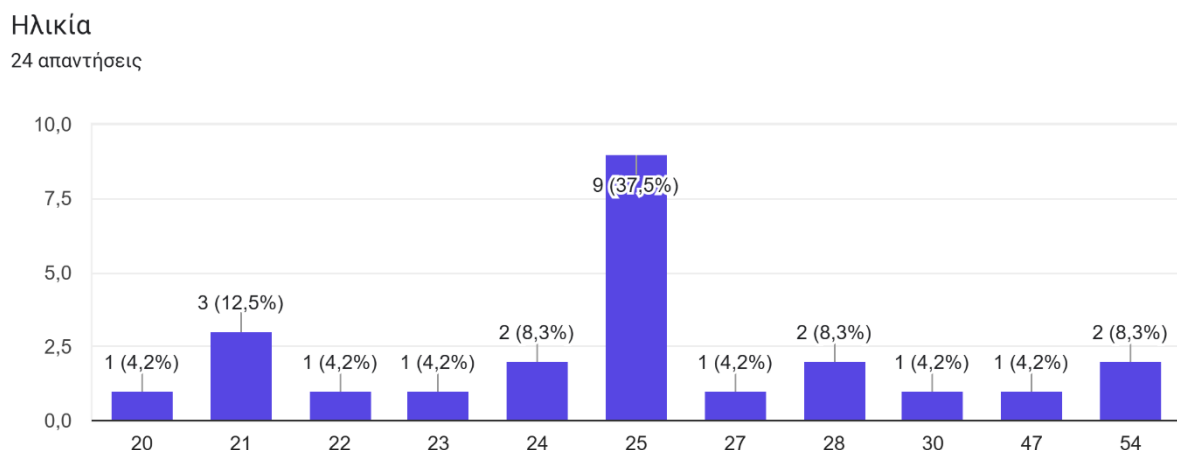
6.1 Εισαγωγή

Πέρα από την τεχνική αξιολόγηση των μοντέλων και της απόδοσης στη συσκευή, η αξία ενός βοηθητικού συστήματος κρίνεται τελικά από την υποκειμενική εμπειρία του ίδιου του χρήστη. Το παρόν κεφάλαιο παρουσιάζει την αξιολόγηση αυτή, η οποία πραγματοποιήθηκε μέσω του σταθμισμένου ερωτηματολογίου User Experience Questionnaire – Short (UEQ-S) [43]. Στόχος ήταν η διερεύνηση της πραγματολογικής (pragmatic) και της ηδονικής (hedonic) ποιότητας της εφαρμογής, όπως τις αντιλαμβάνονται οι ίδιοι οι χρήστες, καθώς και η συλλογή ποιοτικών εντυπώσεων σχετικά με τη φωνητική και απτική ανάδραση.

6.2 Μεθοδολογία και Δημογραφικά Στοιχεία

Το ερωτηματολόγιο UEQ-S αποτελείται από οκτώ ζεύγη αντίθετων χαρακτηριστικών (π.χ. «παρελκυστικό–υποστηρικτικό», «μπερδεμένο–σαφές»), τα οποία απαντώνται σε επταβάθμια κλίμακα (1–7). Οι τέσσερις πρώτες ερωτήσεις συνθέτουν την πραγματολογική ποιότητα και οι τέσσερις τελευταίες την ηδονική· οι απαντήσεις μετασχηματίζονται σε κλίμακα −3 έως +3 για την ανάλυση. Λόγω πρακτικών περιορισμών στην προσέγγιση ατόμων με πιστοποιημένη οπτική αναπηρία εντός του διαθέσιμου χρόνου εκπόνησης, οι συμμετέχοντες δοκίμασαν την εφαρμογή με δεμένα μάτια, ώστε να προσομοιωθεί η εμπειρία ενός τυφλού χρήστη καθιερωμένη πρακτική σε προκαταρκτικές αξιολογήσεις βοηθητικής τεχνολογίας, η οποία ωστόσο δεν υποκαθιστά την αξιολόγηση από άτομα με πραγματική, μόνιμη οπτική αναπηρία (περιορισμός που συζητείται στην Ενότητα 7.6).

Το ερωτηματολόγιο συμπληρώθηκε από 24 συμμετέχοντες αμέσως μετά τη δοκιμή της εφαρμογής σε πραγματικές συνθήκες πεζοδρομίου.

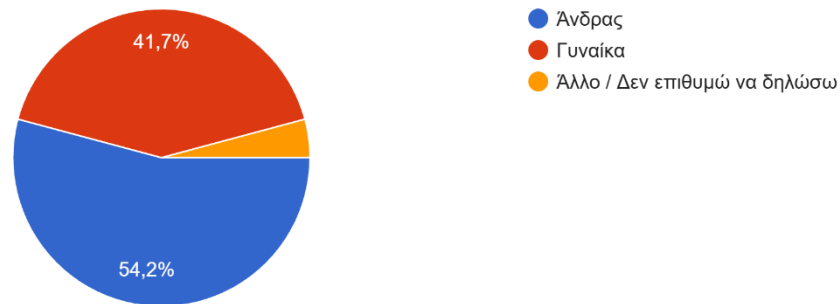


Εικόνα 6.1 Κατανομή ηλικίας συμμετεχόντων

Η ηλικία των συμμετεχόντων κυμάνθηκε από 20 έως 54 έτη, με μέση τιμή 27, 9 έτη. Η πλειοψηφία βρισκόταν σε νεαρή ενήλικη ηλικία, γεγονός που πρέπει να συνυπολογίζεται κατά την ερμηνεία των αποτελεσμάτων, καθώς η εξοικείωση με κινητές συσκευές τείνει να είναι υψηλότερη σε αυτή την ηλικιακή ομάδα σε σχέση με μεγαλύτερες ηλικίες.

Φύλο

24 απαντήσεις

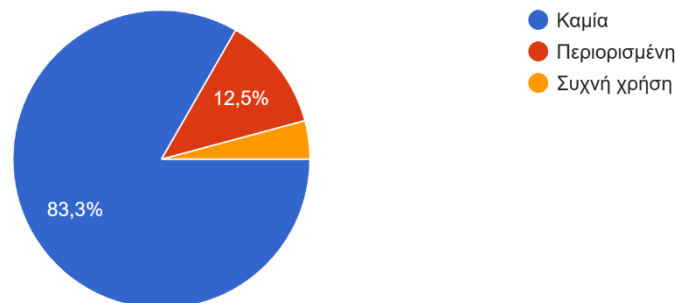


Εικόνα 6.2 Κατανομή φύλου συμμετεχόντων

Από τους 24 συμμετέχοντες, 13 ήταν άνδρες (54, 2%), 10 γυναίκες (41, 7%) και 1 άτομο προτίμησε να μην δηλώσει φύλο (4, 2%), σχηματίζοντας ένα σχετικά ισορροπημένο δείγμα ως προς αυτή τη μεταβλητή.

Προηγούμενη εμπειρία με βοηθητικές εφαρμογές

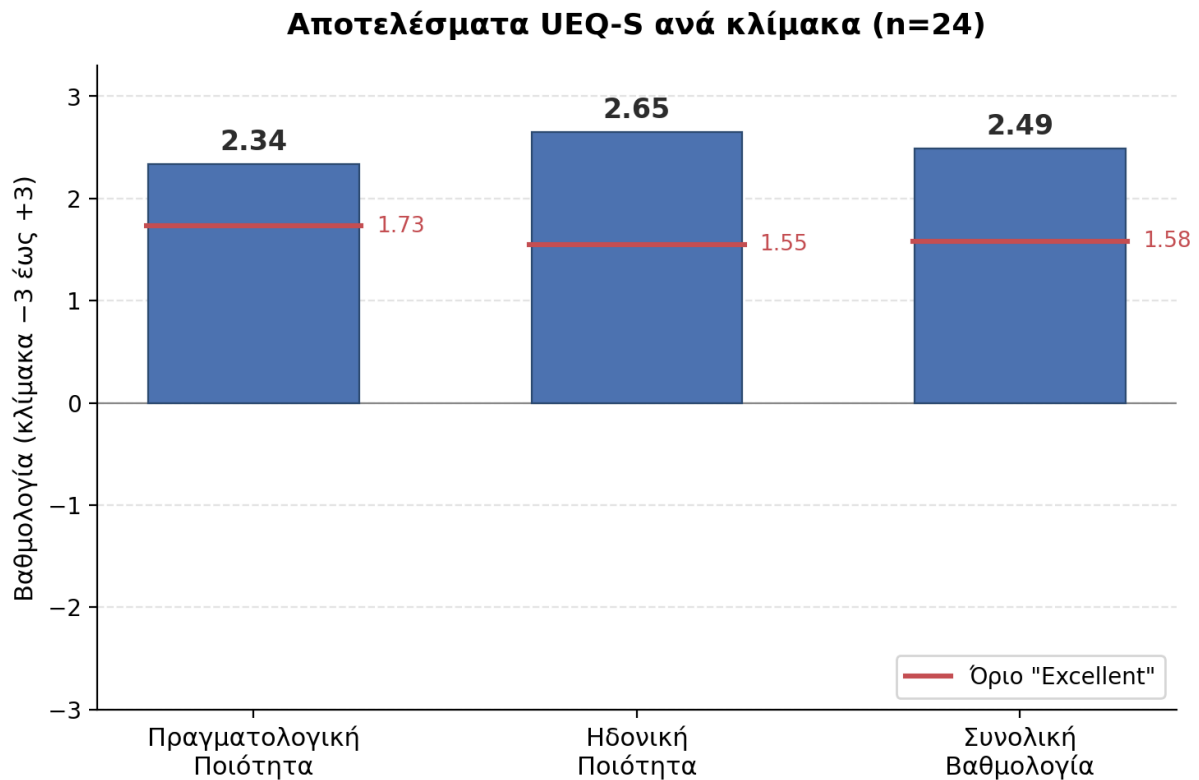
24 απαντήσεις



Εικόνα 6.3 Προηγούμενη εμπειρία με βοηθητικές εφαρμογές

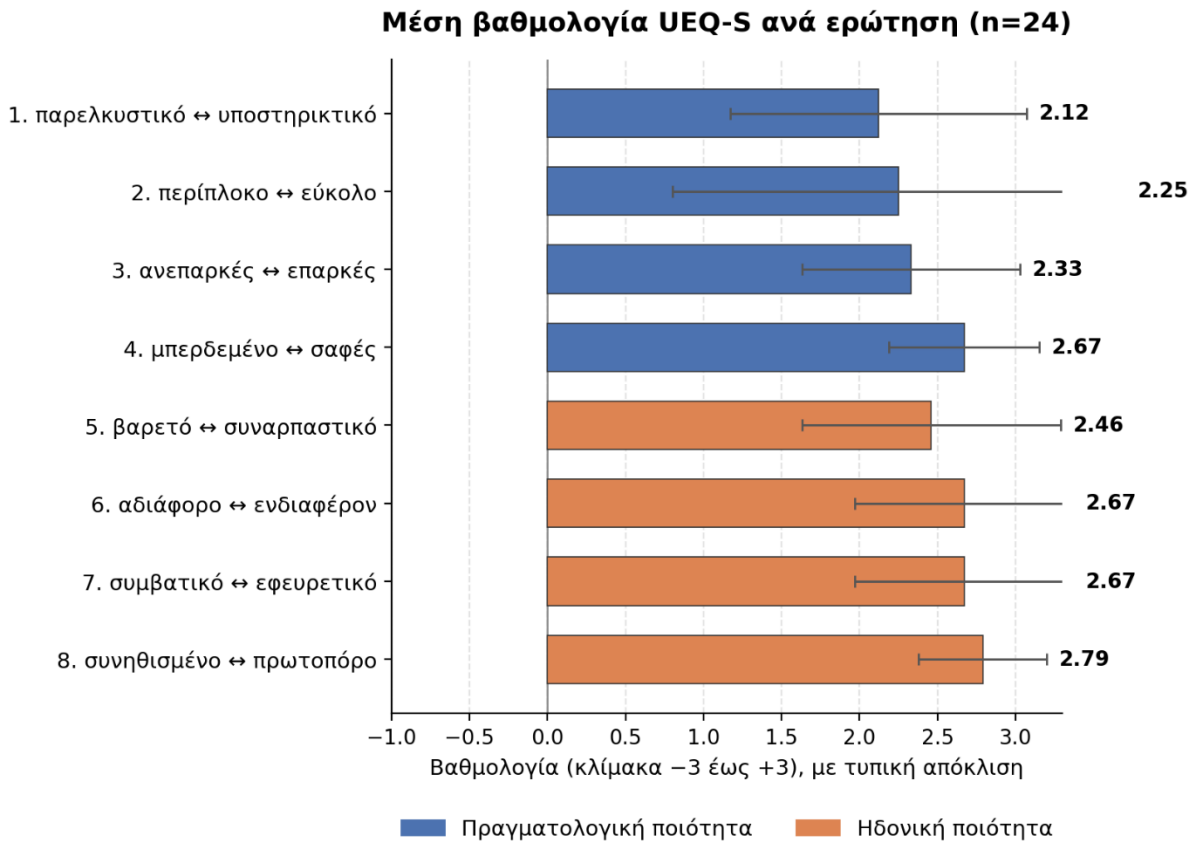
Η πλειοψηφία των συμμετεχόντων (20 από τους 24, 83, 3%) δεν είχε καμία προηγούμενη εμπειρία με αντίστοιχες βοηθητικές εφαρμογές, ενώ μόλις 4 ανέφεραν περιορισμένη ή συχνή χρήση. Το εύρημα αυτό ενισχύει την αξιοπιστία της αξιολόγησης ως προς την αυθόρμητη κατανοητότητα της εφαρμογής, καθώς οι εντυπώσεις δεν επηρεάζονται σε μεγάλο βαθμό από προϋπάρχουσα εξοικείωση με παρόμοια συστήματα.

6.3 Αποτελέσματα Ερωτηματολογίου



Εικόνα 6.4 Μέση βαθμολογία ανά κλίμακα (Πραγματολογική / Ηδονική / Συνολική), με σύγκριση ως προς το όριο «Excellent» του UEQ-S benchmark

Η εφαρμογή πέτυχε μέση πραγματολογική ποιότητα 2, 34, ηδονική ποιότητα 2, 65 και συνολική βαθμολογία 2, 49, σε κλίμακα -3 έως +3. Και οι τρεις τιμές υπερβαίνουν με σαφήνεια τα όρια της κατηγορίας «Excellent» του διεθνούς UEQ-S benchmark (πραγματολογική >1, 73, ηδονική >1, 55, συνολική >1, 58), η οποία αντιστοιχεί στο καλύτερο 10% μεταξύ εκατοντάδων αξιολογημένων προϊόντων. Ιδιαίτερα αξιοσημείωτη είναι η υψηλή ηδονική βαθμολογία, η οποία σε καθαρά χρηστικά συστήματα ασφαλείας συνήθως υστερεί έναντι της πραγματολογικής· εδώ συμβαίνει το αντίστροφο, ένδειξη ότι οι συμμετέχοντες βρήκαν την εμπειρία όχι μόνο λειτουργική αλλά και ευχάριστη.



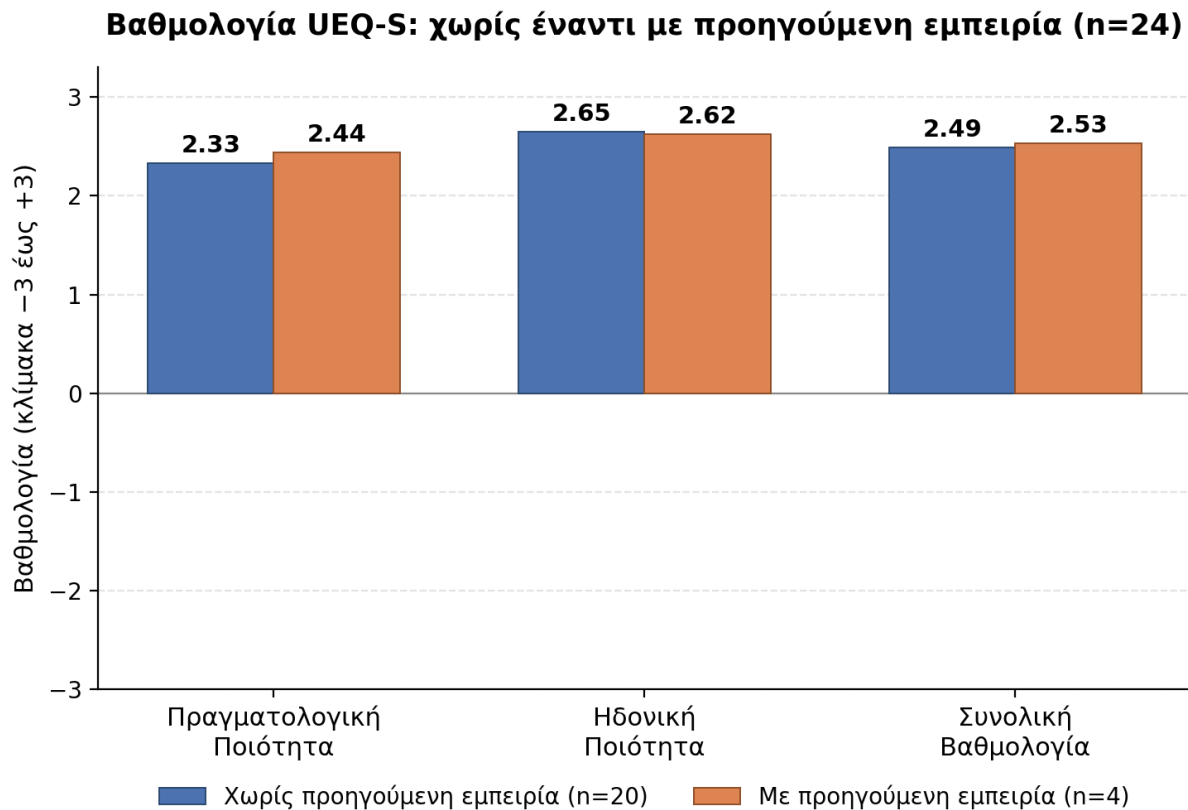
Εικόνα 6.5 Μέση βαθμολογία ανά ερώτηση

Όλες οι επιμέρους ερωτήσεις έλαβαν θετική αξιολόγηση, από 2, 12 («παρελκυστικό–υποστηρικτικό») έως 2, 79 («συνηθισμένο–πρωτοπόρο»), χωρίς καμία να πλησιάζει το ουδέτερο σημείο. Η μεγαλύτερη διασπορά απαντήσεων εντοπίστηκε στο ζεύγος «περίπλοκο–εύκολο» (τυπική απόκλιση 1, 45), ένδειξη ότι, παρότι η πλειονότητα βρήκε την εφαρμογή εύκολη στη χρήση, μια μειοψηφία εξέφρασε πιο συγκρατημένη άποψη, πιθανώς λόγω της πρώτης επαφής με τη λογική εναλλαγής λειτουργιών.

Στις ανοιχτού τύπου ερωτήσεις, 19 από τους 24 συμμετέχοντες που απάντησαν δήλωσαν ότι θα εμπιστεύονταν την εφαρμογή «σίγουρα» ως συμπληρωματικό βοήθημα, με τους υπόλοιπους 3 να απαντούν θετικά αλλά με επιφυλάξεις. Αρκετοί σχολίασαν αυθόρμητα τη χρησιμότητα της απτικής ανάδρασης σε περιβάλλον με θόρυβο, όπου η ηχητική πληροφορία είναι δυσκολότερο να ακουστεί καθαρά.

6.3.1 Σύγκριση βάσει προηγούμενης εμπειρίας

Δεδομένου ότι η πλειοψηφία των συμμετεχόντων δεν είχε προηγούμενη επαφή με βοηθητικές εφαρμογές (Ενότητα 6.2), εξετάστηκε αν αυτό επηρέασε τη βαθμολογία τους σε σύγκριση με τους λίγους συμμετέχοντες με κάποια εμπειρία.



Εικόνα 6.6 Συνολική βαθμολογία UEQ-S: χωρίς έναντι με προηγούμενη εμπειρία

Ομάδα	n	Πραγματολογική	Ηδονική	Συνολική
Χωρίς προηγούμενη εμπειρία	20	2, 33	2, 65	2, 49
Με κάποια προηγούμενη εμπειρία	4	2, 44	2, 62	2, 53

Η συνολική βαθμολογία παραμένει ουσιαστικά αμετάβλητη ανάμεσα στις δύο ομάδες (2, 49 έναντι 2, 53), παρά τη μεγάλη διαφορά μεγέθους δείγματος. Το εύρημα αυτό αποτελεί ενθαρρυντική ένδειξη ότι η θετική εντύπωση δεν οφείλεται σε εξοικείωση με αντίστοιχα συστήματα, αλλά φαίνεται να προκύπτει από την ίδια τη σχεδίαση της εφαρμογής η ερμηνεία αυτή, ωστόσο, πρέπει να αντιμετωπίζεται με επιφύλαξη λόγω του πολύ μικρού μεγέθους της δεύτερης ομάδας (n=4).

6.4 Επίλογος

Το παρόν κεφάλαιο παρουσίασε την αξιολόγηση της υποκειμενικής εμπειρίας χρήσης του BlindNav από 24 συμμετέχοντες, μέσω του σταθμισμένου ερωτηματολογίου UEQ-S. Οι βαθμολογίες 2, 34 στην πραγματολογική ποιότητα, 2, 65 στην ηδονική και 2, 49 συνολικά κατατάσσουν την εφαρμογή στην κατηγορία «Excellent» του διεθνούς benchmark και στις τρεις κλίμακες, με ιδιαίτερα ενθαρρυντικό το εύρημα ότι η βαθμολογία παρέμεινε σταθερή ανεξάρτητα από την προηγούμενη εμπειρία των συμμετεχόντων. Τα αποτελέσματα αυτά πρέπει ωστόσο να ερμηνεύονται στο πλαίσιο της μεθοδολογίας: οι συμμετέχοντες ήταν βλέποντες υπό συνθήκη προσομοιωμένης, όχι πραγματικής, οπτικής αναπηρίας όριο που συζητείται αναλυτικότερα στο επόμενο κεφάλαιο, μαζί με τα υπόλοιπα ευρήματα της πειραματικής αξιολόγησης.

Κεφάλαιο 7ο: Πειραματικά Αποτελέσματα και Συζήτηση

Η αξιολόγηση των εκπαιδευμένων μοντέλων πραγματοποιήθηκε στο σύνολο δεδομένων επικύρωσης, με χρήση των καθιερωμένων μετρικών της ανίχνευσης αντικειμένων που περιεγράφηκαν στην Ενότητα 4.5.

7.1.1 Μοντέλο ανίχνευσης απτικών πλακιδίων

Το μοντέλο των οκτώ κλάσεων πέτυχε συνολική μέση ακρίβεια $mAP@0.5 = 0,685$, με μέγιστη τιμή F1 ίση με 0,66 σε τιμή εμπιστοσύνης 0,299 και μέγιστη ανάκληση 0,86. Η ανάλυση ανά κλάση, ωστόσο, αποκαλύπτει σημαντική ανομοιογένεια, η οποία είναι και το ουσιαστικότερο εύρημα της ενότητας.

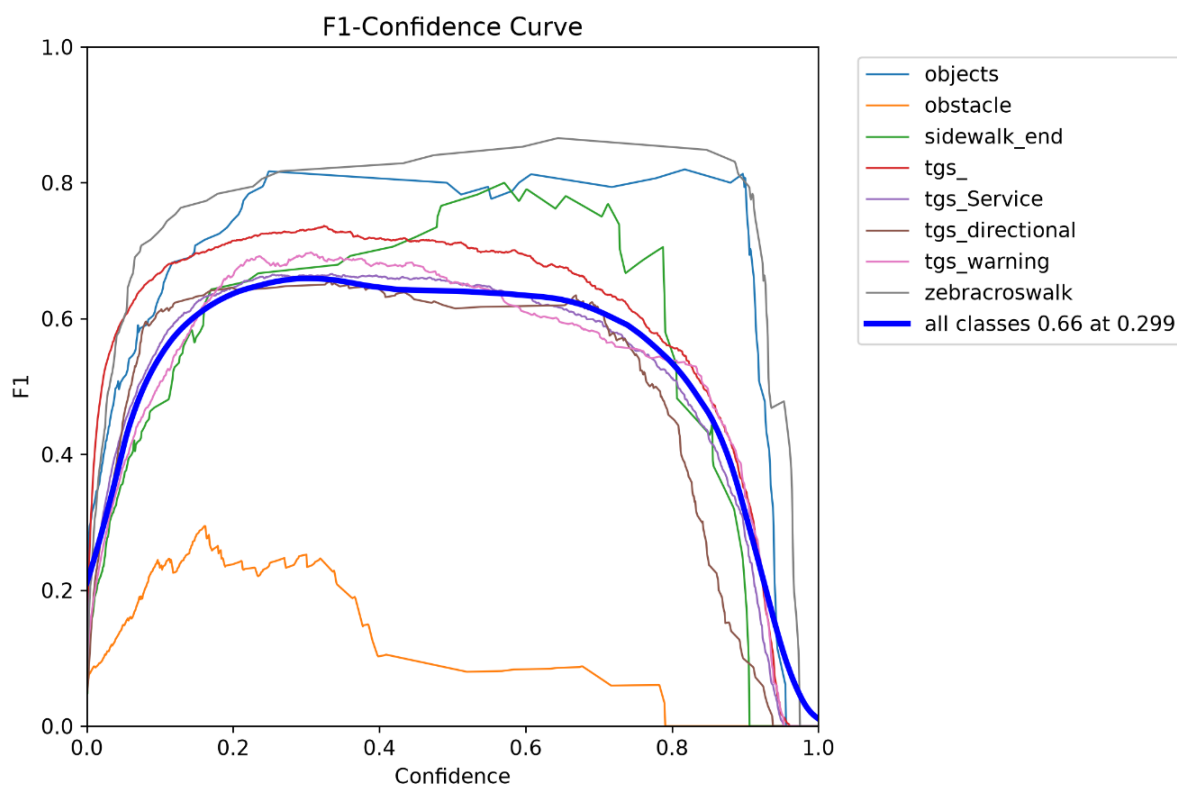
Κλάση	AP@0.5	Ορθή ταξινόμηση*
objects	0,875	0,84
sidewalk_end	0,813	0,86
tgs_ (γενικό πλακίδιο)	0,792	0,81
zebracrosswalk (διάβαση)	0,790	0,83
tgs_warning (προειδοποίησης)	0,732	0,67
tgs_Service (υπηρεσίας)	0,690	0,66
tgs_directional (κατεύθυνσης)	0,609	0,68
obstacle (εμπόδιο)	0,181	0,22
Σύνολο (mAP@0.5)	0,685	

Πίνακας 7.1 Μέση ακρίβεια (AP@0.5) ανά κλάση, μοντέλο απτικών πλακιδίων

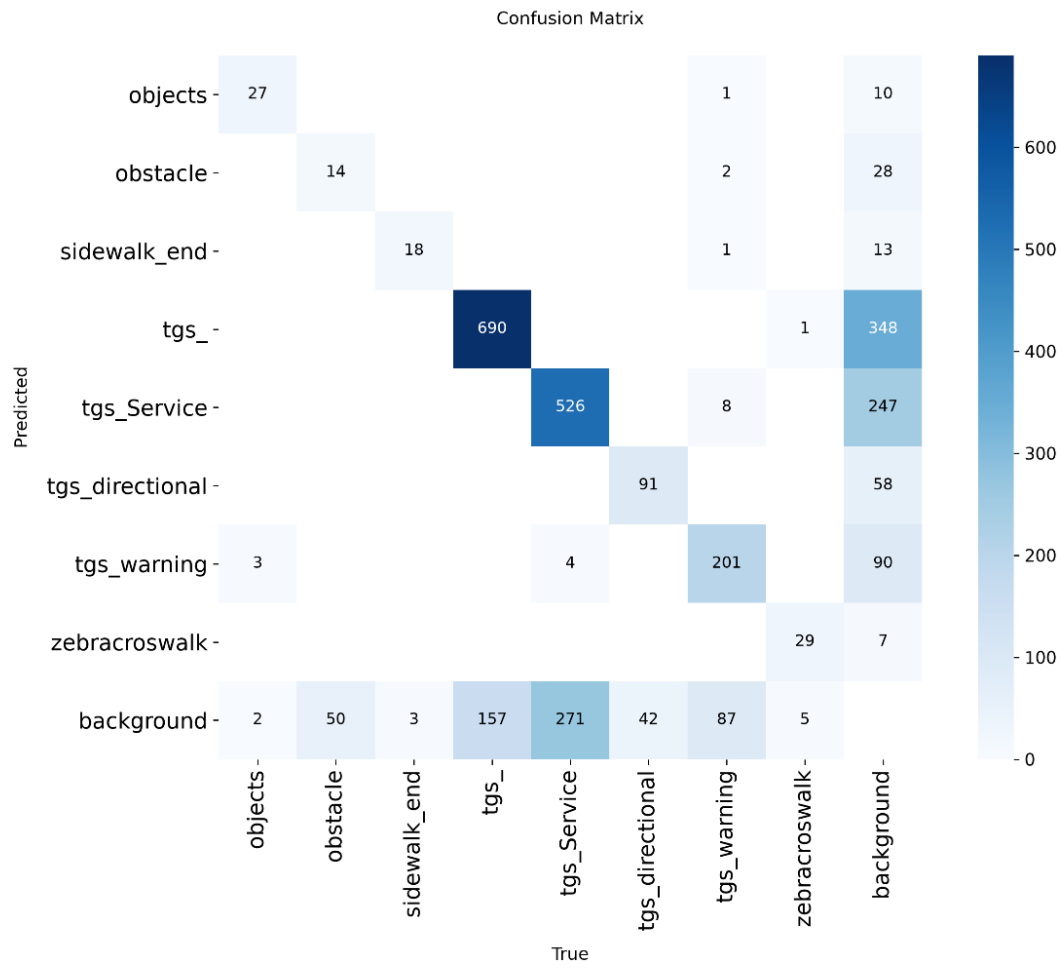
*Διαχώνιος του κανονικοποιημένου πίνακα σύγχυσης

Η κλάση obstacle αποτυγχάνει συστηματικά. Με AP μόλις 0,181 και ορθή ταξινόμηση 22%, υπολείπεται δραματικά όλων των υπολοίπων. Ο πίνακας σύγχυσης δείχνει ότι το 78% των πραγματικών εμποδίων δεν ανιχνεύεται καθόλου, ταξινομούμενο ως υπόβαθρο. Η ερμηνεία είναι σχετικά προφανής: σε αντίθεση με τα απτικά πλακίδια, που έχουν τυποποιημένη μορφή, χρώμα και υφή, η κατηγορία «εμπόδιο» περιλαμβάνει ετερογενή αντικείμενα χωρίς κοινά οπτικά χαρακτηριστικά κάδους, κολόνες, σταθμευμένα οχήματα, τραπεζοκαθίσματα. Η εκμάθηση μιας τόσο ασαφώς οριοθετημένης κατηγορίας απαιτεί σημαντικά μεγαλύτερο και ποικιλότερο σύνολο δεδομένων από αυτό που ήταν διαθέσιμο. Η αδυναμία αυτή έχει άμεση πρακτική συνέπεια για την ασφάλεια και καταγράφεται στους περιορισμούς της Ενότητας 5.6. Η δεύτερη αδυναμία αφορά τη μη ανίχνευση πλακιδίων. Ο κανονικοποιημένος πίνακας σύγχυσης δείχνει ότι σημαντικό ποσοστό πραγματικών πλακιδίων ταξινομείται ως υπόβαθρο: 34% για την κλάση tgs_Service, 32% για την tgs_directional, 29% για την tgs_warning και 19% για τη γενική tgs_. Δηλαδή περίπου ένα στα τρία πλακίδια ειδικού τύπου διαφεύγει της ανίχνευσης σε επίπεδο μεμονωμένου καρέ. Είναι σημαντικό ότι η επίπτωση αυτού του φαινομένου μετριάζεται σε επίπεδο συστήματος από τον μηχανισμό χρονικής εξομάλυνσης: καθώς κάθε πλακίδιο παραμένει ορατό σε πολλαπλά διαδοχικά καρέ, η αστοχία σε ένα καρέ δεν συνεπάγεται απαραίτητα απώλεια της

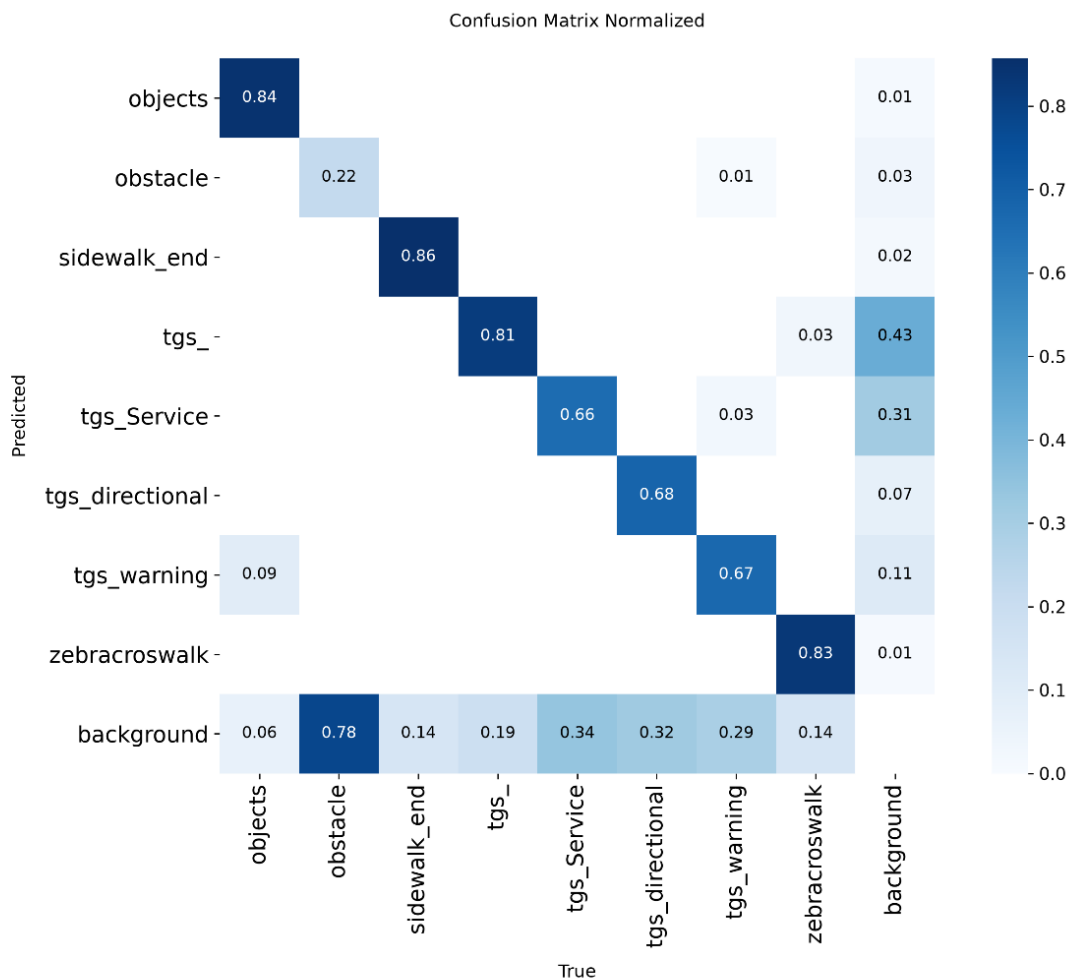
ανακοίνωσης. Οι συγχύσεις μεταξύ κλάσεων είναι περιορισμένες. Το θετικό εύρημα του πίνακα σύγκρισης είναι ότι τα μη διαγώνια στοιχεία μεταξύ πραγματικών κλάσεων παραμένουν χαμηλά: η μεγαλύτερη σύγχυση αφορά 7 περιπτώσεις όπου πλακίδιο προειδοποίησης ταξινομήθηκε ως υπηρεσίας, και 4 αντίστροφα. Κρίσιμο για την ασφάλεια είναι ότι δεν καταγράφηκε καμία σύγχυση μεταξύ των κλάσεων *tgs_warning* και *tgs_directional*, δηλαδή μεταξύ των δύο τύπων με αντίθετη σημασιολογία (στάση έναντι πορείας). Το κύριο πρόβλημα του μοντέλου δεν είναι επομένως η εσφαλμένη αναγνώριση αλλά η μη αναγνώριση. Η επιλογή του κατώφλιου εμπιστοσύνης. Η καμπύλη F1–εμπιστοσύνης παρουσιάζει μέγιστο στην τιμή 0, 299, ενώ η εφαρμογή λειτουργεί με κατώφλι 0, 45. Η καμπύλη είναι ωστόσο αρκετά επίπεδη στην περιοχή αυτή, με τιμή F1 περίπου 0, 65 στο 0, 45 έναντι 0, 66 στο βέλτιστο, δηλαδή η απώλεια είναι οριακή. Η επιλογή της υψηλότερης τιμής είναι συνειδητή και αιτιολογείται από τη φύση της εφαρμογής: σε ένα βοήθημα ασφάλειας, η συχνότητα των εσφαλμένων ανακοινώσεων επηρεάζει άμεσα την εμπιστοσύνη του χρήστη στο σύστημα, και επομένως η ακρίβεια προτιμάται έναντι της ανάκλησης.



Εικόνα 7.1 Καμπύλη F1–εμπιστοσύνης του μοντέλου ανίχνευσης απτικών πλακιδίων (μέγιστο F1 = 0, 66 στην τιμή εμπιστοσύνης 0, 299)



Εικόνα 7.2 Πίνακας σύγχυσης του μοντέλου ανίχνευσης απτικών πλακιδίων (οκτώ κλάσεις, απόλυτος αριθμός ανιχνεύσεων)



Εικόνα 7.3 Κανονικοποιημένος πίνακας σύγχυσης του μοντέλου ανίχνευσης απτικών πλακιδίων

7.1.2 Μοντέλο ανίχνευσης φωτεινών σηματοδοτών

Το μοντέλο των σηματοδοτών πέτυχε συνολική μέση ακρίβεια $mAP@0.5 = 0,861$, σημαντικά υψηλότερη από εκείνη του μοντέλου των πλακιδίων. Η διαφορά είναι ερμηνεύσιμη: οι φωτεινοί σηματοδότες αποτελούν τυποποιημένα, φωτεινά και χρωματικά διακριτά αντικείμενα, σε αντίθεση με τα πλακίδια, τα οποία εμφανίζονται υπό ποικίλες γωνίες, με φθορά, σκιάσεις και μερική επικάλυψη.

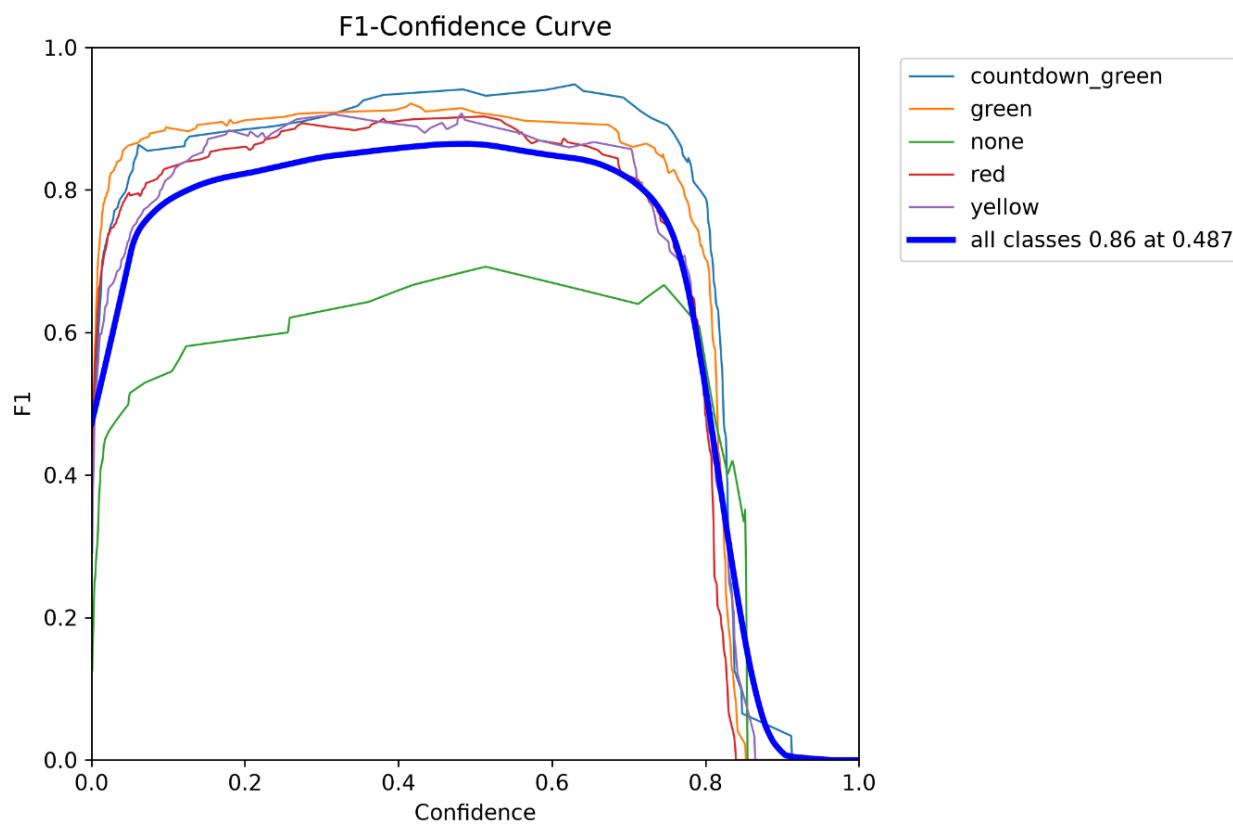
Κλάση	AP@0.5	Ορθή ταξινόμηση*
countdown_green	0,962	0,93
yellow	0,961	0,90
green	0,915	0,92
red	0,912	0,86
none	0,555	0,57
countdown_red	-	-

Σύνολο (mAP@0.5)	0, 861	
-------------------------	---------------	--

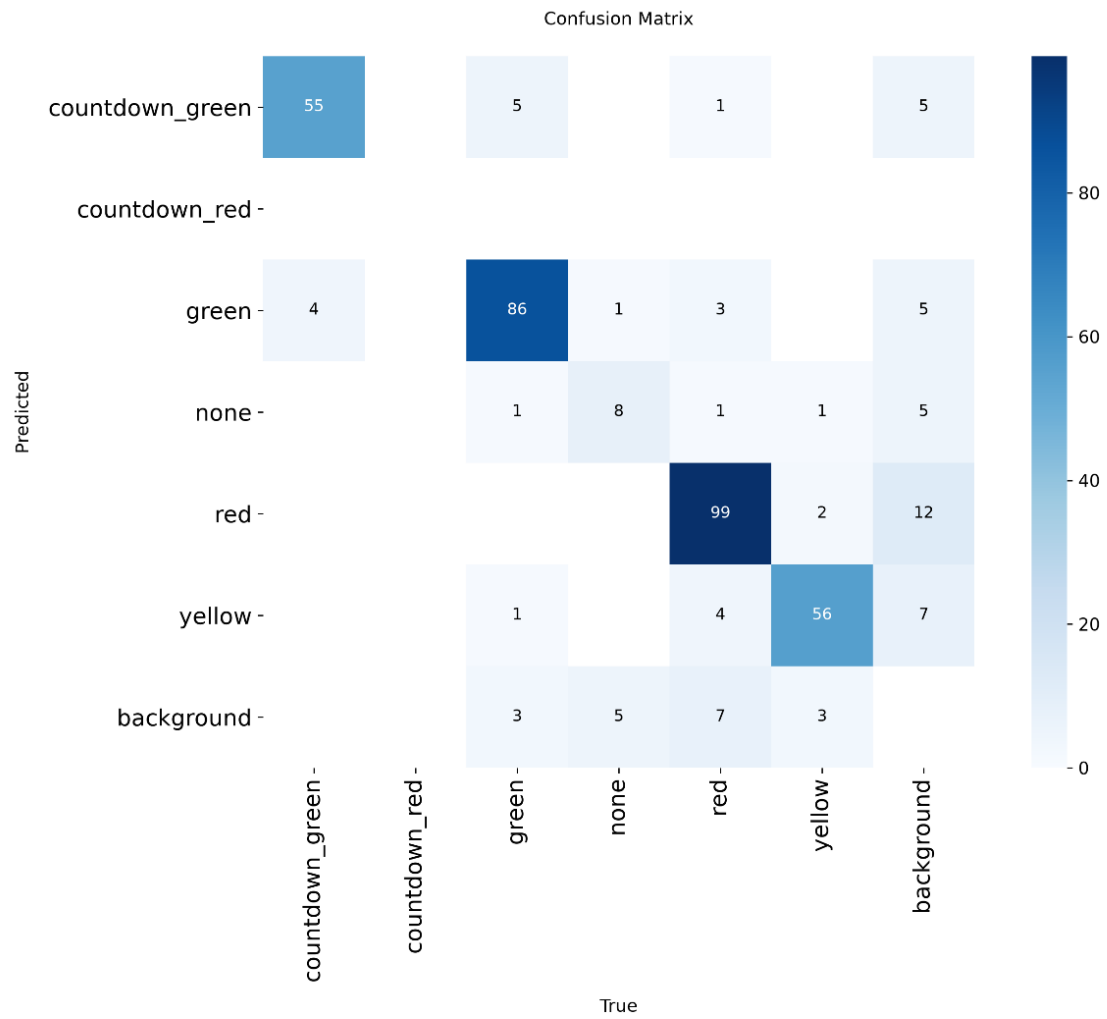
Πίνακας 7.2 Μέση ακρίβεια (AP@0.5) ανά κλάση, μοντέλο σηματοδοτών

*Διαγώνιος του κανονικοποιημένου πίνακα σύγχυσης.

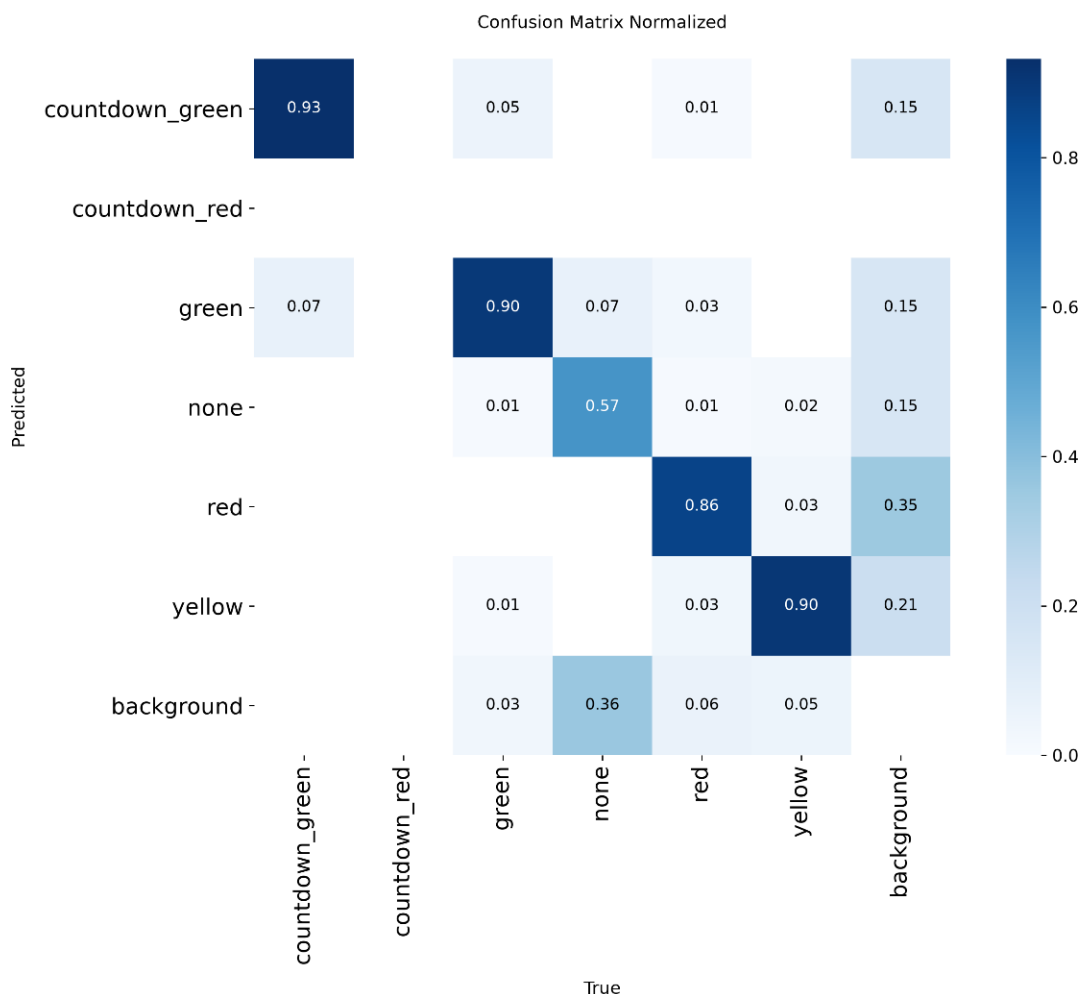
Οι τέσσερις κλάσεις που αντιστοιχούν σε πραγματικές ενδείξεις σηματοδότη επιτυγχάνουν επιδόσεις άνω του 0, 91, επίπεδο που κρίνεται ικανοποιητικό για την προβλεπόμενη χρήση. Ωστόσο, τρία ευρήματα απαιτούν ιδιαίτερη ανάλυση. Η κλάση `countdown_red` δεν αξιολογήθηκε. Ο πίνακας σύγχυσης δεν περιλαμβάνει καμία καταγραφή για την κλάση αυτή, ούτε ως πραγματική ούτε ως προβλεπόμενη, γεγονός που υποδηλώνει ότι το σύνολο δεδομένων επικύρωσης δεν περιέχει κανένα σχετικό παράδειγμα. Η κλάση παραμένει δηλωμένη στον κατάλογο ετικετών της εφαρμογής, αλλά η ικανότητα του μοντέλου να την αναγνωρίζει είναι στην πράξη ατεκμηρίωτη. Το εύρημα αυτό αποτελεί κενό της αξιολόγησης και καταγράφεται στους περιορισμούς της Ενότητας 5.6. Καταγράφηκαν περιπτώσεις σύγχυσης κόκκινου με πράσινο. Ο πίνακας σύγχυσης δείχνει τρεις περιπτώσεις (3% των πραγματικών κόκκινων ενδείξεων) στις οποίες σηματοδότης σε κόκκινη φάση ταξινομήθηκε ως πράσινος. Αν και το ποσοστό είναι αριθμητικά μικρό, η φύση του σφάλματος είναι η σοβαρότερη δυνατή για την ασφάλεια του χρήστη: μια τέτοια εσφαλμένη ταξινόμηση θα οδηγούσε στην ανακοίνωση άδειας διάσχισης σε στιγμή που η διάσχιση απαγορεύεται. Επιπροσθέτως, το 6% των πραγματικών κόκκινων ενδείξεων δεν ανιχνεύθηκε καθόλου. Τα ευρήματα αυτά έχουν άμεση σχεδιαστική συνέπεια, η οποία συζητείται στην Ενότητα 7.4: η ανακοίνωση της πράσινης φάσης δεν πρέπει να εκλαμβάνεται από τον χρήστη ως εντολή διάσχισης, αλλά ως συμπληρωματική πληροφορία που δεν υποκαθιστά την ακουστική εκτίμηση της κυκλοφορίας. Η κλάση `none` υπολείπεται σαφώς. Με AP 0, 555 και ορθή ταξινόμηση 57%, η κλάση που δηλώνει την απουσία σηματοδότη είναι η ασθενέστερη του μοντέλου· το 36% των περιπτώσεών της ταξινομείται ως υπόβαθρο. Η αδυναμία αυτή είναι ωστόσο η λιγότερο επιζήμια, καθώς η κλάση `none` δεν παράγει καθοδηγητική πληροφορία προς τον χρήστη· η σύγχυσή της με το υπόβαθρο έχει ως αποτέλεσμα την απουσία ανακοίνωσης, δηλαδή ακριβώς τη συμπεριφορά που θα ήταν επιθυμητή. Το κατώφλι εμπιστοσύνης επιβεβαιώνεται. Η καμπύλη F1–εμπιστοσύνης του μοντέλου παρουσιάζει μέγιστο (0, 86) στην τιμή 0, 481, η οποία βρίσκεται πολύ κοντά στο κατώφλι των 0, 45 που εφαρμόζει η εφαρμογή. Σε αντίθεση με το μοντέλο των πλακιδίων, όπου η επιλογή του 0, 45 συνιστούσε συνειδητή απόκλιση από το βέλτιστο υπέρ της ακρίβειας, εδώ η ρύθμιση συμπίπτει σχεδόν με το σημείο βέλτιστης ισορροπίας.



Εικόνα 7.4 Καμπύλη F1–εμπιστοσύνης του μοντέλου ανίχνευσης φωτεινών σηματοδοτών (μέγιστο F1 = 0, 86 στην τιμή εμπιστοσύνης 0, 481)



Εικόνα 7.5 Πίνακας σύγκρισης του μοντέλου ανίχνευσης φωτεινών σηματοδοτών (έξι κλάσεις, απόλυτος αριθμός ανιχνεύσεων)



Εικόνα 7.6 Κανονικοποιημένος πίνακας σύγχυσης του μοντέλου ανίχνευσης φωτεινών σηματοδοτών

7.1.3 Συγκριτική αποτίμηση

Η αντιπαράβολή των δύο μοντέλων αναδεικνύει ότι η δυσκολία του προβλήματος διαφέρει ουσιαστικά ανά κατηγορία στόχου. Το μοντέλο των σηματοδοτών επιτυγχάνει $mAP@0.5 = 0,861$ έναντι $0,685$ του μοντέλου των πλακιδίων, με τη διαφορά να εντοπίζεται κυρίως στις κλάσεις με ασαφή ή μεταβλητή οπτική υπόσταση. Η επιλογή της εκπαίδευσης δύο ανεξάρτητων μοντέλων, που τεκμηριώθηκε στην Ενότητα 3.3 με βάση το διαφορετικό οπτικό πλαίσιο, επιβεβαιώνεται και εκ των υστέρων: η συνένωση των δύο συνόλων σε ενιαίο μοντέλο θα υποχρέωνε το δίκτυο να μοιράσει τη χωρητικότητά του μεταξύ δύο προβλημάτων διαφορετικής φύσης και δυσκολίας.

7.2 Απόδοση στη συσκευή

Η αξιολόγηση της απόδοσης σε πραγματικές συνθήκες πραγματοποιήθηκε μέσω ενσωματωμένου μηχανισμού καταγραφής, ο οποίος αποθηκεύει σε αρχείο CSV, για κάθε εκτέλεση του μοντέλου, τον χρόνο εξαγωγής προβλέψεων, τον συνολικό χρόνο επεξεργασίας του καρέ, το πλήθος και το είδος των ανιχνεύσεων, τη σταθεροποιημένη χωρική ζώνη και το τυχόν εκφωνηθέν μήνυμα. Οι μετρήσεις διεξήχθησαν σε δύο διαφορετικές συσκευές, ώστε να εκτιμηθεί η επίδραση του υλικού.

Συσκευή	Καρέ	Διάρκεια	Μέσος (ms)	Διάμεσος (ms)	Ελάχ. (ms)	Μέγ. (ms)	p95 (ms)
Samsung-SM-F936B	1.489	15, 7 λεπτά	320, 2	321, 0	169	1.613	431, 0
Xiaomi 2412DPC0AG	448	4, 4 λεπτά	460, 3	448, 5	222	968	637, 6
Samsung SM-A566B	2.530	32, 6 λεπτά	355, 7	368, 0	242	1.508	404, 0

Πίνακας 7.3 Χρόνοι εξαγωγής προβλέψεων ανά συσκευή

Και οι τρεις συσκευές λειτουργούν με Android 16 (API 36). Η προσθήκη μιας τρίτης συσκευής, μεσαίας κατηγορίας (Samsung Galaxy A56, SM-A566B), επιβεβαιώνει το ίδιο συμπέρασμα: ο μέσος χρόνος εξαγωγής προβλέψεων κυμαίνεται από 320 έως 460 ms ανάλογα με το υλικό, με το SM-A566B να τοποθετείται ενδιάμεσα περίπου 11% πιο αργό από το πτυσσόμενο SM-F936B, συνεπώς με τη χαμηλότερη υπολογιστική ισχύ μιας συσκευής μεσαίας κατηγορίας. Ιδιαίτερο ενδιαφέρον παρουσιάζει η διαφορά μεταξύ διαμέσου και μέγιστης τιμής. Στη συσκευή Samsung, ενώ η διάμεσος παραμένει στα 321 ms, καταγράφηκε μεμονωμένη τιμή 1.613 ms, δηλαδή πενταπλάσια. Δεδομένου ότι η καταγραφή αυτή προήλθε από συνεχή λειτουργία 15, 7 λεπτών, το φαινόμενο αποδίδεται πιθανότατα σε θερμική υποβάθμιση της απόδοσης (thermal throttling), μηχανισμό με τον οποίο η συσκευή περιορίζει τη συχνότητα του επεξεργαστή για να αποφύγει την υπερθέρμανση. Η παρατήρηση αυτή είναι σημαντική για μια εφαρμογή προοριζόμενη για παρατεταμένη χρήση κατά τη μετακίνηση.

Συσκευή	Μέσο διάστημα μεταξύ καρέ (ms)	Πραγματικός ρυθμός (FPS)	Θεωρητικό μέγιστο (FPS)
Samsung SM-F936B	634, 6	1, 58	3, 12
Xiaomi 2412DPC0AG	593, 4	1, 69	2, 17
Samsung SM-A566B	774, 1	1, 29	2, 81

Πίνακας 7.4 Πραγματικός ρυθμός επεξεργασίας

Η σύγκριση των δύο τελευταίων στηλών αποκαλύπτει το σημαντικότερο εύρημα της ενότητας: ο πραγματικός ρυθμός υπολείπεται σαφώς του θεωρητικού μεγίστου που επιτρέπει ο χρόνος εξαγωγής προβλέψεων. Στη συσκευή Samsung, ενώ το μοντέλο θα επέτρεπε περίπου 3, 1 αναλύσεις ανά δευτερόλεπτο, επιτεύχθηκαν μόλις 1, 58. Το κενό αυτό, της τάξης των 300 ms ανά κύκλο, οφείλεται στον μηχανισμό ρύθμισης ρυθμού: όταν ένα καρέ φθάνει ενώ βρίσκεται σε εξέλιξη προηγούμενη εξαγωγή προβλέψεων, απορρίπτεται από τη σημαία isProcessing, και η επόμενη ανάλυση αναμένει το επόμενο χρονικό παράθυρο. Η βελτιστοποίηση του σημείου αυτού αποτελεί την προφανέστερη κατεύθυνση βελτίωσης και συζητείται στην Ενότητα 7.6.

Αντίθετα, το κόστος του ίδιου του pipeline επεξεργασίας αποδείχθηκε αμελητέο: η διαφορά μεταξύ του συνολικού χρόνου καρέ και του καθαρού χρόνου εξαγωγής προβλέψεων ήταν κατά μέσο όρο μικρότερη του ενός χιλιοστού του δευτερολέπτου και στις τρεις συσκευές (0, 19 ms κατά μέσο όρο στο SM-A566B). Το εύρημα αυτό επιβεβαιώνει ότι η χρονική εξομάλυνση, η χωρική ταξινόμηση και η λογική προτεραιότητας δεν επιβαρύνουν πρακτικά την απόδοση· το σύνολο σχεδόν του υπολογιστικού κόστους οφείλεται στην εκτέλεση του νευρωνικού δικτύου.

7.3 Συμπεριφορά σε συνθήκες πεδίου

Η ανάλυση της συνεδρίας των 4, 4 λεπτών στη συσκευή Xiaomi, η οποία περιλάμβανε βάδιση σε πεζοδρόμιο με απτικά πλακίδια και προσέγγιση διάβασης, παρήγαγε 448 καταγραφές. Η κατανομή της επικρατέστερης ανιχνευθείσας κλάσης ανά καρέ παρουσιάζεται στον Πίνακα 5.5.

Κλάση	Καρέ	Ποσοστό
tgs_ (γενικό πλακίδιο)	225	50, 2%
zebracrosswalk (διάβαση)	87	19, 4%
καμία ανίχνευση	78	17, 4%
tgs_warning (προειδοποίησης)	22	4, 9%
sidewalk_end (τέλος πεζοδρομίου)	13	2, 9%
tgs_Service (υπηρεσίας)	10	2, 2%
obstacle (εμπόδιο)	9	2, 0%
objects	3	0, 7%
tgs_directional (κατεύθυνσης)	1	0, 2%

Πίνακας 7.5 Κατανομή επικρατέστερης κλάσης ανά καρέ (Xiaomi, 448 καρέ)

Δύο παρατηρήσεις προκύπτουν. Πρώτον, το ποσοστό των καρέ χωρίς καμία ανίχνευση (17, 4%) είναι σχετικά χαμηλό, γεγονός που δείχνει ότι το σύστημα διατηρεί σχεδόν συνεχή επίγνωση του περιβάλλοντος κατά τη βάδιση. Δεύτερον, η κλάση tgs_directional εμφανίζεται μόλις μία φορά σε αυτή τη συνεδρία, ενώ η γενική κλάση tgs_ κυριαρχεί με ποσοστό άνω του 50%. Όπως δείχνει ο δεύτερος, εκτενέστερος γύρος δοκιμών που ακολουθεί, το εύρημα αυτό αποδείχθηκε χαρακτηριστικό της συγκεκριμένης, σύντομης διαδρομής η οποία απλώς δεν περιλάμβανε πολλά κατευθυντήρια πλακίδια και όχι αδυναμία του μοντέλου να τα αναγνωρίσει.

Κλάση	Καρέ	Ποσοστό
tgs_ (γενικό πλακίδιο)	1.678	60, 7%
καμία ανίχνευση	471	17, 0%
zebracrosswalk (διάβαση)	175	6, 3%

tgs_warning (προειδοποίησης)	169	6, 1%
tgs_Service (υπηρεσίας)	102	3, 7%
tgs_directional (κατεύθυνσης)	92	3, 3%
obstacle (εμπόδιο)	52	1, 9%
sidewalk_end (τέλος πεζοδρομίου)	25	0, 9%
objects	1	0, 0%

Πίνακας 7.6 Κατανομή επικρατέστερης κλάσης ανά καρτέ (Samsung SM-F936B, 2.765 καρτέ σε λειτουργία πλακιδίων)

Κλάση	Καρτέ	Ποσοστό
tgs_ (γενικό πλακίδιο)	1.367	58, 0%
καμία ανίχνευση	288	12, 2%
tgs_warning (προειδοποίησης)	238	10, 1%
tgs_Service (υπηρεσίας)	157	6, 7%
zebracrosswalk (διάβαση)	123	5, 2%
tgs_directional (κατεύθυνσης)	90	3, 8%
obstacle (εμπόδιο)	75	3, 2%
sidewalk_end (τέλος πεζοδρομίου)	17	0, 7%
objects	1	0, 0%

Πίνακας 7.7 Κατανομή επικρατέστερης κλάσης ανά καρτέ (Samsung SM-A566B, 2.356 καρτέ σε λειτουργία πλακιδίων)

Και στις τρεις συσκευές, το tgs_directional εμφανίζεται πλέον σε σταθερό, μη αμελητέο ποσοστό (3, 3–3, 8%), επιβεβαιώνοντας ότι πρόκειται για αναγνωρίσιμη κλάση στην πράξη· η σχεδόν πλήρης απουσία της στον πρώτο γύρο δοκιμών αντικατοπτρίζει τη διαδρομή που επιλέχθηκε τότε, όχι το μοντέλο. Σημειώνεται επίσης ότι το ποσοστό καρτέ χωρίς ανίχνευση είναι χαμηλότερο στο SM-A566B

(12, 2%) απ' ό, τι στο SM-F936B (17, 0%), πιθανώς λόγω διαφορετικών διαδρομών δοκιμής μεταξύ των συνεδριών.

- Χωρική καθοδήγηση
Από τα 448 καρέ, στα 290 (64, 7%) παρήχθη σταθεροποιημένη χωρική ζώνη. Η κατανομή ήταν: κεντρική ζώνη 246 καρέ (84, 8%), αριστερή 33 (11, 4%) και δεξιά 11 (3, 8%). Η σαφής κυριαρχία της κεντρικής ζώνης είναι αναμενόμενη και επιθυμητή, καθώς αντανακλά την ορθή στόχευση της συσκευής προς τον διάδρομο όδευσης κατά τη διάρκεια της βάδισης· οι πλευρικές ζώνες εμφανίζονται κυρίως σε στιγμές απόκλισης ή στροφής.

Ο δεύτερος γύρος δοκιμών επιβεβαιώνει το ίδιο μοτίβο σε μεγαλύτερη κλίμακα: στο SM-F936B, από 2.231 καρέ με σταθεροποιημένη ζώνη, η κεντρική κυριάρχησε με 1.902 (85, 3%), έναντι 207 (9, 3%) στη δεξιά και 122 (5, 5%) στην αριστερή· στο SM-A566B, από 2.006 καρέ, η κεντρική έφτασε 1.807 (90, 1%), με 130 (6, 5%) δεξιά και 69 (3, 4%) αριστερά. Η συνέπεια αυτού του μοτίβου σε τρεις ανεξάρτητες συσκευές και επτά συνολικά συνεδρίες ενισχύει την ερμηνεία ότι η κυριαρχία της κεντρικής ζώνης αντανακλά πράγματι τη φυσιολογική συμπεριφορά βάδισης, όχι ιδιαιτερότητα μίας συσκευής ή διαδρομής.

- Ρυθμός και σύνθεση ανακοινώσεων
Κατά τη διάρκεια της συνεδρίας εκφωνήθηκαν 62 ανακοινώσεις, δηλαδή μία ανά 4, 3 δευτερόλεπτα κατά μέσο όρο. Δεδομένου ότι παρήχθησαν 448 αποτελέσματα, ο μηχανισμός onlyOnChange απέτρεψε την εκφώνηση στο 86% των περιπτώσεων. Το ποσοστό αυτό ποσοτικοποιεί άμεσα τη συνεισφορά του μηχανισμού: χωρίς αυτόν, ο χρήστης θα δεχόταν επταπλάσιο όγκο ηχητικής πληροφορίας, στη συντριπτική πλειονότητά του επαναλαμβανόμενης.

Μήνυμα	Πλήθος
Διάβαση πεζών	32
Πλακίδια όδευσης	15
Προσοχή, πλακίδια προειδοποίησης	8
Τέλος πεζοδρομίου	4
Πλακίδια υπηρεσίας	2
Προσοχή! Εμπόδιο	1

Πίνακας 7.8 Σύνθεση ανακοινώσεων (Xiaomi, 62 ανακοινώσεις)

Μήνυμα	Πλήθος
Πλακίδια όδευσης	83
Διάβαση πεζών	59
Προσοχή, πλακίδια προειδοποίησης	30
Πλακίδια υπηρεσίας	29

Πλακίδια κατεύθυνσης	23
Προσοχή! Εμπόδιο	16
Τέλος πεζοδρομίου	14
Κόκκινο, σταματήστε	14
Αντικείμενο μπροστά	6
Πράσινο, μπορείτε να περάσετε	5

Πίνακας 7.9 Σύνθεση ανακοινώσεων (Samsung SM-F936B, 2.912 καρέ σε 3 συνεδρίες, 279 ανακοινώσεις)

Μήνυμα	Πλήθος
Πλακίδια όδευσης	70
Διάβαση πεζών	43
Προσοχή! Εμπόδιο	34
Προσοχή, πλακίδια προειδοποίησης	30
Πλακίδια κατεύθυνσης	28
Πλακίδια υπηρεσίας	20
Τέλος πεζοδρομίου	15
Κόκκινο, σταματήστε	10
Πράσινο, μπορείτε να περάσετε	6
Αντικείμενο μπροστά	2
Κίτρινο, προσοχή	1

Πίνακας 7.10 Σύνθεση ανακοινώσεων (Samsung SM-A566B, 2.530 καρέ σε 3 συνεδρίες, 259 ανακοινώσεις)

Ο δεύτερος γύρος δοκιμών επιβεβαιώνει τα ίδια ποιοτικά ευρήματα σε μεγαλύτερη κλίμακα: στο SM-F936B εκφωνήθηκε μία ανακοίνωση ανά 9, 2 δευτερόλεπτα κατά μέσο όρο, στο SM-A566B ανά 7, 6 δευτερόλεπτα, με τον μηχανισμό onlyOnChange να αποτρέπει και στις δύο περιπτώσεις περίπου το 90% των δυνητικών εκφωνήσεων ελαφρώς υψηλότερο ποσοστό από το 86% της αρχικής συνεδρίας, συνεπές με τη μεγαλύτερη συνολική διάρκεια δοκιμής. Αξιοσημείωτη είναι επίσης η εμφάνιση της κλάσης κατεύθυνσης σε αξιοπρεπές πλήθος (23 και 28 ανακοινώσεις αντίστοιχα), σε αντίθεση με τη σχεδόν πλήρη απουσία της στην αρχική, μικρότερη συνεδρία (βλ. προηγούμενη υποενότητα).

Ο δεύτερος γύρος δοκιμών επιβεβαιώνει τα ίδια ποιοτικά ευρήματα σε μεγαλύτερη κλίμακα: στο SM-F936B εκφωνήθηκε μία ανακοίνωση ανά 9, 2 δευτερόλεπτα κατά μέσο όρο, στο SM-A566B ανά 7, 6 δευτερόλεπτα, με τον μηχανισμό onlyOnChange να αποτρέπει και στις δύο περιπτώσεις περίπου το 90% των δυνητικών εκφωνήσεων, ελαφρώς υψηλότερο ποσοστό από το 86% της αρχικής συνεδρίας, συνεπές με τη μεγαλύτερη συνολική διάρκεια δοκιμής. Αξιοσημείωτη είναι επίσης η εμφάνιση της

κλάσης κατεύθυνσης σε αξιοπρεπές πλήθος (23 και 28 ανακοινώσεις αντίστοιχα), σε αντίθεση με τη σχεδόν πλήρη απουσία της στην αρχική, μικρότερη συνεδρία (βλ. προηγούμενη υποενότητα).

7.3.1 Ορθότητα ανακοινώσεων

Για την αξιολόγηση της ορθότητας χρησιμοποιήθηκε δεύτερο αρχείο καταγραφής, στο οποίο ο χειριστής σημείωνε, μέσω των φυσικών πλήκτρων της συσκευής, τα πραγματικά γεγονότα του περιβάλλοντος. Καταγράφηκαν 21 σημάνσεις: 16 επιβεβαιώσεις παρουσίας πλακιδίου και 5 επισημάνσεις εσφαλμένης ανακοίνωσης. Επισημαίνεται εξ αρχής ότι η σήμανση δεν ήταν εξαντλητική: ο χειριστής δεν κατέγραφε κάθε μεμονωμένο γεγονός, ιδίως στον πρώτο, μικρότερο γύρο δοκιμών. Στον δεύτερο γύρο η κάλυψη ήταν γενικότερη, με σκόπιμα αυξημένη προσοχή σε δύο σημεία: στα πλακίδια υπηρεσίας, όπου η αξιολόγηση των μοντέλων (Ενότητα 7.1.1) είχε ήδη υποδείξει αδυναμία, και σε περιπτώσεις όπου η εφαρμογή ανακοίνωνε διάβαση πεζών σε σημεία όπου δεν υπήρχε. Η στοχευμένη αυτή προσοχή σε συγκεκριμένες κατηγορίες σημαίνει ότι τα μεγέθη που ακολουθούν, αν και προέρχονται πλέον από σημαντικά μεγαλύτερο δείγμα, δεν αποτελούν αμερόληπτη στατιστική εκτίμηση ακρίβειας σε επίπεδο πληθυσμού· παραμένουν όμως πολύ πιο στέρεη βάση από τον αρχικό, μικρό γύρο δοκιμών. Συνολικά, στους δύο γύρους δοκιμών καταγράφηκαν 339 σημάνσεις παρουσίας πλακιδίου και 95 σημάνσεις εσφαλμένης ανακοίνωσης. Από τις 339 σημάνσεις παρουσίας, οι 278 συνοδεύονταν από αντίστοιχη ανακοίνωση εντός χρονικού παραθύρου έξι δευτερολέπτων πριν έως τριών δευτερολέπτων μετά τη σήμανση ποσοστό 82, 0% ενώ οι υπόλοιπες 61 αποτελούν ενδεικτικές περιπτώσεις ψευδώς αρνητικού αποτελέσματος.

Ανακοίνωση	Χρόνος πριν τη σήμανση	Εμπιστοσύνη
Τέλος πεζοδρομίου	12	36, 4%
Διάβαση πεζών	27	20, 1%
Πλακίδια όδευσης	31	18, 5%
Προσοχή, πλακίδια προειδοποίησης	10	14, 7%
Πλακίδια υπηρεσίας	7	13, 7%
Προσοχή! Εμπόδιο	5	9, 8%
Αντικείμενο μπροστά	1	12, 5%
Πλακίδια κατεύθυνσης	2	3, 9%

Πίνακας 7.11 Καταγεγραμμένες περιπτώσεις ψευδώς θετικών, δύο γύροι δοκιμών (n=95)

Δύο ευρήματα ξεχωρίζουν. Πρώτον, η μέση εμπιστοσύνη των 95 εσφαλμένων ανακοινώσεων είναι 0, 754 (διάμεσος 0, 783) σαφώς υψηλότερη από το κατώφλι των 0, 45, επιβεβαιώνοντας στο πολλαπλάσιο δείγμα το εύρημα του πρώτου γύρου: η αύξηση του κατωφλίου εμπιστοσύνης δεν θα απέτρεπε τα σφάλματα αυτά, ενώ θα οδηγούσε σε απώλεια ορθών ανιχνεύσεων. Δεύτερον, το ποσοστό ψευδώς θετικών δεν κατανέμεται ομοιόμορφα ανά τύπο μηνύματος: το «Τέλος πεζοδρομίου» (36, 4%) και η «Διάβαση πεζών» (20, 1%) συγκεντρώνουν τα υψηλότερα ποσοστά, ενώ τα «Πλακίδια κατεύθυνσης»

το χαμηλότερο (3, 9%). Το εύρημα αυτό επιβεβαιώνει την παρατήρηση του χειριστή κατά τις δοκιμές πεδίου, ο οποίος σημείωσε ρητά ότι η ανακοίνωση διάβασης πεζών εμφανιζόταν επανειλημμένα σε σημεία χωρίς πραγματική διάβαση. Τα πλακίδια υπηρεσίας, αντίθετα, δεν ξεχωρίζουν ιδιαίτερα ως πηγή ψευδώς θετικών (13, 7%, μέτρια θέση στην κατάταξη). Η αδυναμία τους, όπως τεκμηριώνεται ήδη στην Ενότητα 7.1.1, εντοπίζεται στη μη ανίχνευση το 34% των πραγματικών πλακιδίων υπηρεσίας ταξινομείται ως υπόβαθρο και όχι στην εσφαλμένη ανακοίνωση. Η στοχευμένη προσοχή του χειριστή σε αυτά τα πλακίδια κατά τον δεύτερο γύρο δοκιμών επιβεβαιώνει επομένως, ποιοτικά, το ήδη τεκμηριωμένο εύρημα του Κεφαλαίου 7.1.1 για αυτή την κλάση, χωρίς ωστόσο να προσθέτει στοιχεία για νέο, διαφορετικό πρόβλημα ψευδών ανακοινώσεων. Σημειώνεται τέλος ότι μία εκ των 95 σημάνσεων αντιστοιχήθηκε σε ανακοίνωση με καταγεγραμμένη εμπιστοσύνη 0, 000, πιθανή ένδειξη ατέλειας στην καταγραφή του συγκεκριμένου πεδίου για ορισμένους τύπους μηνυμάτων· το εύρημα δεν επηρεάζει ουσιωδώς τη μέση τιμή, αξίζει όμως μελλοντικό έλεγχο του μηχανισμού καταγραφής.

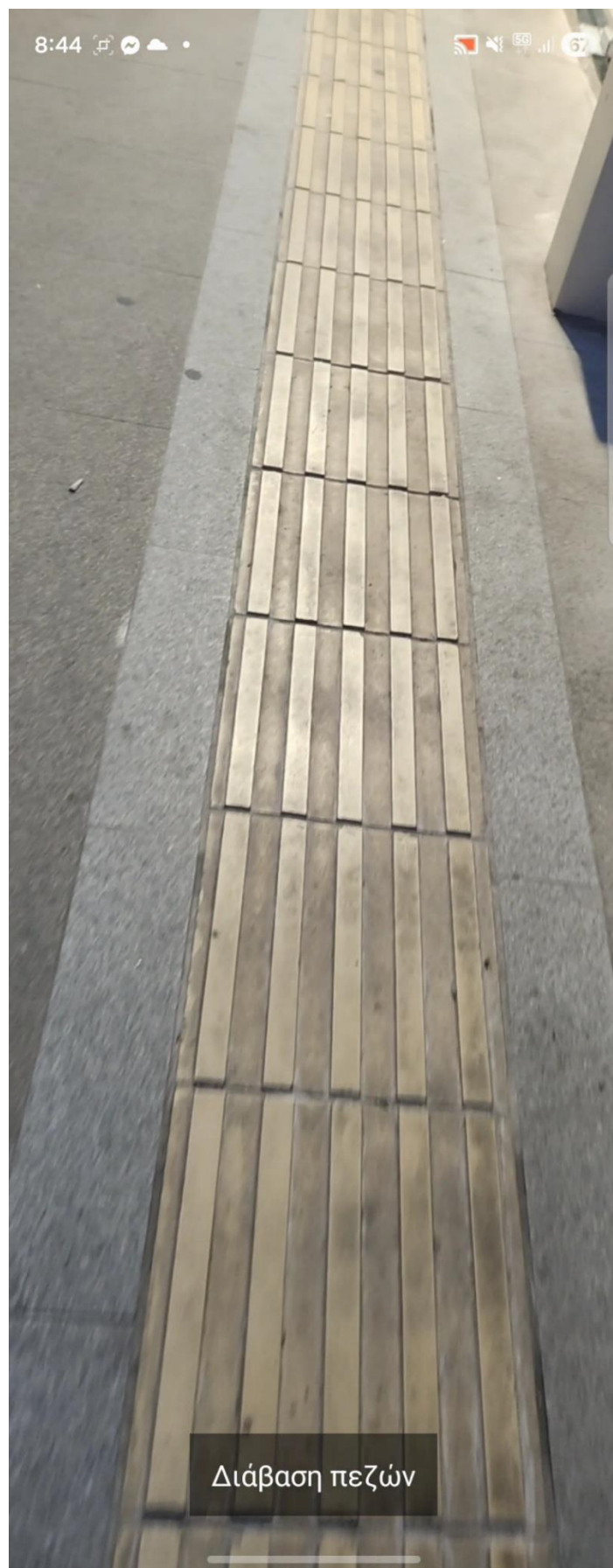
7.4 Συζήτηση

Η κυρίαρχη μορφή αστοχίας είναι η μη ανίχνευση, όχι η εσφαλμένη ταξινόμηση. Το εύρημα αυτό διατρέχει και τα δύο μοντέλα. Στο μοντέλο των πλακιδίων, περίπου ένα στα τρία πλακίδια ειδικού τύπου ταξινομείται ως υπόβαθρο, ενώ οι συγχύσεις μεταξύ πραγματικών κλάσεων παραμένουν ελάχιστες. Στο μοντέλο των σηματοδοτών, το 36% των περιπτώσεων της κλάσης none και το 6% των κόκκινων ενδείξεων διαφεύγουν. Η διαπίστωση αυτή είναι σχετικά ενθαρρυντική: ένα σύστημα που περιστασιακά σιωπά είναι προτιμότερο από ένα σύστημα που ανακοινώνει λανθασμένη πληροφορία, ιδίως όταν η σιωπή αντισταθμίζεται από τη χρονική εξομάλυνση. Η χρονική εξομάλυνση αποδεικνύεται καθοριστική. Η επίπτωση των αστοχιών ανά καρέ μετριάζεται σημαντικά σε επίπεδο συστήματος, καθώς κάθε αντικείμενο παραμένει ορατό σε πολλαπλά διαδοχικά καρέ. Η μέτρηση του §5.3, σύμφωνα με την οποία 278 από τις 339 σημάνσεις παρουσίας πλακιδίου (82, 0%) συνοδεύονταν από ανακοίνωση, δείχνει ότι η συνολική συμπεριφορά του συστήματος είναι αισθητά καλύτερη από ό, τι θα προέβλεπε η ανά καρέ ακρίβεια του μοντέλου. Παράλληλα, ο μηχανισμός onlyOnChange απέτρεψε 86–90% των δυνητικών εκφωνήσεων στις δύο συνεδρίες πεδίου, περιορίζοντας δραστικά τη γνωσιακή επιβάρυνση του χρήστη. Οι σχεδιαστικές αδυναμίες εντοπίζονται σε κατηγορίες με ασαφή οπτική υπόσταση. Οι δύο χειρότερες επιδόσεις η κλάση obstacle με AP 0, 181 και η κλάση none με 0, 555 μοιράζονται ένα κοινό χαρακτηριστικό: αμφότερες ορίζονται μέσω της απουσίας ή της ετερογένειας, όχι μέσω σταθερών οπτικών χαρακτηριστικών. Αντιθέτως, οι κλάσεις με τυποποιημένη μορφή, όπως οι φάσεις του σηματοδότη ή η διάβαση πεζών, επιτυγχάνουν σαφώς υψηλότερες επιδόσεις. Το συμπέρασμα είναι ότι η επιτυχία της προσέγγισης εξαρτάται άμεσα από την ύπαρξη τυποποιημένης υποδομής γεγονός που ενισχύει τη βασική θέση της εργασίας, ότι η αξιοποίηση των απτικών πλακιδίων αποτελεί ορθή στρατηγική επιλογή. Το ζήτημα των ψευδώς θετικών υψηλής εμπιστοσύνης. Όπως καταγράφηκε στον Πίνακα 5.11, οι εσφαλμένες ανακοινώσεις που εντοπίστηκαν στο πεδίο, σε δύο διαδοχικούς γύρους δοκιμών (n=95), παρήχθησαν με μέση εμπιστοσύνη 0, 754 και διάμεσο 0, 783. Το γεγονός αυτό αποκλείει τη ρύθμιση του κατωφλίου ως λύση και υποδεικνύει ότι η αντιμετώπιση απαιτεί εμπλουτισμό των δεδομένων εκπαίδευσης, ιδίως με αρνητικά παραδείγματα από περιβάλλοντα που μοιάζουν οπτικά με τους στόχους. Συνέπειες για την ασφάλεια του χρήστη. Η καταγραφή τριών περιπτώσεων ταξινόμησης κόκκινης ένδειξης ως πράσινης (§7.1.2) επιβάλλει σαφή οριοθέτηση της χρήσης του συστήματος. Το BlindNav πρέπει να θεωρείται συμπληρωματικό βοήθημα και όχι υποκατάστατο των καθιερωμένων μεθόδων προσανατολισμού και κινητικότητας. Η ανακοίνωση της πράσινης φάσης οφείλει να εκλαμβάνεται ως πρόσθετη πληροφορία, η οποία δεν αναιρεί την ανάγκη ακουστικής εκτίμησης της κυκλοφορίας πριν από τη διάσχιση. Η οριοθέτηση αυτή δεν αποτελεί απλώς τεχνική επιφύλαξη αλλά ουσιώδη προϋπόθεση για την υπεύθυνη χρήση του συστήματος. Η ερμηνεία αυτή

τεκμηριώνεται περαιτέρω από στιγμιότυπα της εφαρμογής σε λειτουργία, τα οποία καταγράφουν τρεις χαρακτηριστικές περιπτώσεις σύγκυσης μεταξύ κλάσεων πλακιδίων. Η πρώτη πλακίδιο υπηρεσίας που αναγνωρίζεται και ανακοινώνεται ως πλακίδιο προειδοποίησης παρατηρήθηκε επανειλημμένα κατά τις δοκιμές πεδίου και επιβεβαιώνει άμεσα το εύρημα του πίνακα σύγκυσης της Ενότητας 5.1.1: εκεί η σύγκυση αυτών ακριβώς των δύο κλάσεων ήταν η μεγαλύτερη που καταγράφηκε μεταξύ πραγματικών κλάσεων, με 4 καταγεγραμμένες περιπτώσεις πλακιδίου υπηρεσίας ταξινομημένου ως προειδοποίησης στο σύνολο επικύρωσης. Οι άλλες δύο περιπτώσεις είναι εξίσου συχνές στο πεδίο, δεν εμφανίζονται όμως με αντίστοιχη ένταση στον πίνακα σύγκυσης του στατικού συνόλου επικύρωσης: γενικό πλακίδιο όδευσης που ανακοινώνεται ως διάβαση πεζών, και σημαντικότερο πλακίδιο κατεύθυνσης που ανακοινώνεται ως πλακίδιο προειδοποίησης, παρότι η Ενότητα 7.1.1 ανέφερε ρητά μηδενική σύγκυση ανάμεσα σε αυτές τις δύο κλάσεις στο σύνολο επικύρωσης. Η απόκλιση αυτή δεν αναιρεί το εύρημα του Κεφαλαίου 7.1.1 που αφορά συγκεκριμένα το στατικό σύνολο επικύρωσης αλλά υποδεικνύει ότι η συμπεριφορά του μοντέλου υπό μεταβαλλόμενες συνθήκες πεδίου (κίνηση, γωνία λήψης, μερική κάλυψη πλακιδίου) δεν αποτυπώνεται πλήρως από μια αξιολόγηση σε στατικές εικόνες. Το εύρημα αυτό καταγράφεται ως ανοιχτό ζήτημα για συστηματικότερη διερεύνηση σε μελλοντική εργασία (Ενότητα 6.4).



Εικόνα 7.7 (α) πλακίδιο υπηρεσίας με ανακοίνωση «Προσοχή, πλακίδια προειδοποίησης»



Εικόνα 7.8 (β) γενικό πλακίδιο όδευσης με ανακοίνωση «Διάβαση πεζών»



Εικόνα 7.9 (γ) πλακίδιο κατεύθυνσης με ανακοίνωση «Προσοχή, πλακίδια προειδοποίησης»

7.5 Τεχνικά ευρήματα της υλοποίησης

Πέραν των ποσοτικών αποτελεσμάτων, η διαδικασία ανάπτυξης ανέδειξε ορισμένα τεχνικά ζητήματα, η καταγραφή των οποίων έχει αξία τόσο για την τεκμηρίωση της εργασίας όσο και για μελλοντικές υλοποιήσεις αντίστοιχων συστημάτων. Διπλή κανονικοποίηση συντεταγμένων. Κατά την ενσωμάτωση των μοντέλων διαπιστώθηκε ότι το εργαλείο μετατροπής `onnx2tf` παράγει τις συντεταγμένες των πλαισίων οριοθέτησης ήδη κανονικοποιημένες στο διάστημα $[0, 1]$. Η αρχική υλοποίηση, βασισμένη στην εύλογη αλλά εσφαλμένη υπόθεση ότι πρόκειται για συντεταγμένες σε εικονοστοιχεία, διαιρούσε εκ νέου με την ανάλυση εισόδου (640), με αποτέλεσμα πλαίσια μηδενικού μεγέθους και πλήρη αδυναμία χωρικής τοποθέτησης. Το σφάλμα αυτό είναι ιδιαίτερα ύπουλο, καθώς δεν προκαλεί εξαίρεση κατά την εκτέλεση αλλά σιωπηρά εσφαλμένα αποτελέσματα. Σειρά ετικετών κατά την εξαγωγή. Οι κλάσεις εξάγονται από την πλατφόρμα Roboflow με αλφαβητική και όχι σημασιολογική σειρά. Για το μοντέλο των σηματοδοτών, η αντιστοίχιση είναι `countdown_green = 0`, `countdown_red = 1`, `green = 2`, `none = 3`, `red = 4`, `yellow = 5`. Τυχόν αναντιστοιχία μεταξύ της σειράς αυτής και του καταλόγου ετικετών της εφαρμογής οδηγεί σε αντιστροφή των ενδείξεων σφάλμα με προφανείς συνέπειες για την ασφάλεια, το οποίο μάλιστα δεν εντοπίζεται από καμία αυτόματη δοκιμή. Δυναμική επιλογή επιταχυντή GPU. Η εφαρμογή ελέγχει κατά την εκκίνηση τη συμβατότητα της συσκευής με τον GPU delegate του TFLite. Όπου υποστηρίζεται, η εξαγωγή προβλέψεων πραγματοποιείται μέσω GPU, ενώ σε περίπτωση μη υποστήριξης ή σφάλματος προβλέπεται αυτόματο fallback σε CPU. Με τον τρόπο αυτό η εφαρμογή μπορεί να αξιοποιεί διαθέσιμη επιτάχυνση υλικού, χωρίς να εξαρτά τη λειτουργία της από αυτήν. Η εξαγωγή προβλέψεων εκτελείται αποκλειστικά στη CPU, επιλογή που αποδείχθηκε σταθερή σε συνεδρίες διάρκειας έως και 43 λεπτών και στις τρεις συσκευές δοκιμής. Ασφάλεια νημάτων του διεργασίας. Ο διεργασίας του TFLite δεν είναι ασφαλής για ταυτόχρονη πρόσβαση. Με διάστημα εξαγωγής προβλέψεων μικρότερο του χρόνου εκτέλεσης του μοντέλου, η επικάλυψη κλήσεων είναι βέβαιη. Η χρήση της σημαίας `AtomicBoolean` αποδείχθηκε επαρκής και ελάχιστου κόστους λύση. Σύγκρουση εξαρτήσεων κατά τη μετατροπή. Το πακέτο `onnxruntime-gpu` παρουσίασε ασυμβατότητα με το `ai-edge-litert`, απαιτώντας χειροκίνητη διαχείριση των εξαρτήσεων του εικονικού περιβάλλοντος Python. Το κόστος του pipeline είναι αμελητέο. Όπως τεκμηριώθηκε στην Ενότητα 7.2, η διαφορά μεταξύ συνολικού χρόνου καρέ και καθαρού χρόνου εξαγωγής προβλέψεων ήταν κάτω του ενός χιλιοστού του δευτερολέπτου. Το εύρημα αυτό δικαιώνει τη σχεδιαστική επιλογή ενός σχετικά σύνθετου σταδίου επεξεργασίας: η χρονική εξομάλυνση, η χωρική ταξινόμηση και η ιεραρχία προτεραιότητας προσφέρουν ουσιαστική βελτίωση της ποιότητας χωρίς μετρήσιμο υπολογιστικό κόστος.

7.6 Περιορισμοί

Η αξιολόγηση που πραγματοποιήθηκε έδωσε χρήσιμα και ενθαρρυντικά αποτελέσματα σχετικά με τη λειτουργία του BlindNav, ανέδειξε όμως παράλληλα ορισμένους περιορισμούς που είναι σημαντικό να ληφθούν υπόψη κατά την ερμηνεία τους. Οι περιορισμοί αυτοί δεν αναιρούν τα ευρήματα της εργασίας, αλλά προσδιορίζουν τα όρια μέσα στα οποία αυτά πρέπει να αξιολογηθούν και υποδεικνύουν συγκεκριμένες κατευθύνσεις για την περαιτέρω βελτίωση του συστήματος. Ο σημαντικότερος

περιορισμός αφορά τη συμμετοχή ατόμων με πραγματική οπτική αναπηρία. Η αξιολόγηση μέσω του UEQ-S πραγματοποιήθηκε με 24 βλέποντες συμμετέχοντες υπό συνθήκη προσομοιωμένης τύφλωσης και, παρά τις ιδιαίτερα θετικές ενδείξεις ως προς την αρχική αποδοχή της εφαρμογής, τα αποτελέσματα θα πρέπει να ερμηνεύονται με την επιφύλαξη αυτή, όπως συζητείται αναλυτικότερα στο Κεφάλαιο 6. Ανάλογος περιορισμός αφορά τις δοκιμές πεδίου των Ενοτήτων 5.2 και 5.3. Οι δοκιμές πραγματοποιήθηκαν από τον ίδιο βλέποντα χειριστή, γεγονός που επέτρεψε τη συνεπή εφαρμογή της διαδικασίας και την παρατήρηση της συμπεριφοράς της εφαρμογής σε πραγματικές συνθήκες κίνησης. Παράλληλα, ο τρόπος χρήσης της συσκευής ενδέχεται να διαφέρει από εκείνον ενός ατόμου με οπτική αναπηρία, για παράδειγμα ως προς τη γωνία κράτησης της συσκευής και τον ρυθμό βάδισης. Επιπλέον, ο χειριστής είχε τη δυνατότητα οπτικής επιβεβαίωσης των αποτελεσμάτων, δυνατότητα που δεν θα είναι διαθέσιμη στον τελικό χρήστη. Για τον λόγο αυτό, η επόμενη φάση αξιολόγησης θα πρέπει να εξετάσει την κατανόηση της ηχητικής και απτικής ανάδρασης, τον ρυθμό των ανακοινώσεων και τη συνολική χρησιμότητα της καθοδήγησης σε πραγματικούς χρήστες. Ένας δεύτερος περιορισμός σχετίζεται με τη σήμανση των πραγματικών γεγονότων κατά τις δοκιμές πεδίου. Όπως αναφέρθηκε στην Ενότητα 7.3, η καταγραφή πραγματοποιήθηκε επιλεκτικά και δεν κάλυψε το σύνολο των καρτέ και των γεγονότων των συνεδριών. Τα μεγέθη ορθότητας που παρουσιάστηκαν έχουν επομένως ενδεικτικό χαρακτήρα και δεν επιτρέπουν τον υπολογισμό πλήρων μετρικών ακρίβειας και ανάκλησης σε επίπεδο συνολικού συστήματος. Παρόλα αυτά, η διαδικασία προσέφερε χρήσιμες πληροφορίες σχετικά με πραγματικές περιπτώσεις λειτουργίας και σφάλματος, οι οποίες μπορούν να αξιοποιηθούν για τον σχεδιασμό μιας πιο συστηματικής αξιολόγησης στο μέλλον. Σε επίπεδο δεδομένων, ορισμένες κλάσεις παρουσιάζουν μεγαλύτερα περιθώρια βελτίωσης. Η κλάση *obstacle* πέτυχε $AP = 0, 181$, γεγονός που δείχνει ότι η συγκεκριμένη κατηγορία χρειάζεται περισσότερο και πιο διαφοροποιημένα παραδείγματα εκπαίδευσης. Η ανάγκη αυτή είναι αναμενόμενη, καθώς τα εμπόδια παρουσιάζουν σημαντική οπτική ετερογένεια σε σχέση με τις περισσότερες τυποποιημένες κατηγορίες πλακιδίων. Παράλληλα, η κλάση *countdown_red* δεν αξιολογήθηκε, λόγω απουσίας παραδειγμάτων στο σύνολο επικύρωσης. Η δημιουργία ενός εμπλουτισμένου συνόλου δεδομένων για τις συγκεκριμένες κατηγορίες αποτελεί επομένως σαφή και μετρήσιμη ευκαιρία βελτίωσης της εφαρμογής. Ως προς την υπολογιστική απόδοση, ο πραγματικός ρυθμός λειτουργίας κυμάνθηκε μεταξύ 1, 13 και 1, 69 καρτέ ανά δευτερόλεπτο, ανάλογα με τη συσκευή και τη συνεδρία. Η τιμή αυτή είναι χαμηλότερη από τον αρχικό σχεδιαστικό στόχο των 4–5 καρτέ ανά δευτερόλεπτο. Ωστόσο, η τοπική εκτέλεση των μοντέλων παρέμεινε εφικτή και σταθερή στις δοκιμές, ενώ ο περιορισμός του ρυθμού επεξεργασίας συνδέεται σε σημαντικό βαθμό με τον μηχανισμό χρονικής ρύθμισης της εφαρμογής. Το συγκεκριμένο αποτέλεσμα αφήνει σαφές περιθώριο βελτίωσης μέσω τεχνικών βελτιστοποίησης του κύκλου λήψης και επεξεργασίας. Το εύρος των πειραματικών συνθηκών αποτελεί έναν ακόμη περιορισμό. Οι μετρήσεις πραγματοποιήθηκαν σε τρεις συσκευές Android και επτά συνεδρίες, καταναεμημένες σε δύο διαδοχικούς γύρους δοκιμών. Οι δοκιμές πραγματοποιήθηκαν κατά τη διάρκεια της ημέρας και δεν περιλάμβαναν συστηματική αξιολόγηση σε βροχή, χαμηλό φωτισμό ή νυχτερινές συνθήκες. Δεν εξετάστηκε επίσης συστηματικά η συμπεριφορά του μοντέλου σε πλακίδια διαφορετικού χρώματος ή σε πλακίδια με προχωρημένη φθορά. Οι παρατηρήσεις που καταγράφηκαν για τη λειτουργία σε χαμηλό φωτισμό και περιλαμβάνονται στον Οδηγό Χρήσης Λογισμικού αποτελούν χρήσιμες ενδείξεις, αλλά δεν συνοδεύονται από αντίστοιχη ποσοτική αξιολόγηση. Η επέκταση των δοκιμών σε περισσότερες περιβαλλοντικές συνθήκες θα επιτρέψει συνεπώς πληρέστερη αποτίμηση της γενίκευσης του συστήματος. Η θερμική συμπεριφορά αποτελεί επίσης πεδίο που μπορεί να εξεταστεί αναλυτικότερα. Κατά τη διάρκεια συνεδρίας δεκαπέντε λεπτών καταγράφηκαν μεμονωμένοι χρόνοι εξαγωγής προβλέψεων έως 1.613 ms. Το εύρημα αυτό ενδέχεται να συνδέεται με τη θερμική κατάσταση της συσκευής, χωρίς όμως η παρούσα εργασία να περιλαμβάνει συστηματική μέτρηση θερμοκρασίας ώστε

να τεκμηριωθεί άμεσα η σχέση αυτή. Μια μελλοντική μελέτη της θερμικής συμπεριφοράς θα μπορούσε να οδηγήσει σε καλύτερη κατανόηση της απόδοσης κατά την παρατεταμένη χρήση και σε αντίστοιχες βελτιστοποιήσεις. Τέλος, δεν πραγματοποιήθηκε άμεση πειραματική σύγκριση του YOLO11n με άλλες αρχιτεκτονικές ή διαφορετικά μεγέθη μοντέλων. Η επιλογή του YOLO11n βασίστηκε στις απαιτήσεις της εφαρμογής, με έμφαση στη δυνατότητα τοπικής εκτέλεσης και στον συμβιβασμό μεταξύ υπολογιστικού κόστους και ακρίβειας. Μια συστηματική σύγκριση με εναλλακτικές προσεγγίσεις θα μπορούσε να προσφέρει επιπλέον ποσοτικά στοιχεία και να οδηγήσει σε ακόμη καλύτερη επιλογή μοντέλου για μελλοντικές εκδόσεις του BlindNav.

7.7 Επίλογος

Το κεφάλαιο αυτό παρουσίασε τα αποτελέσματα της πειραματικής αξιολόγησης σε τέσσερα επίπεδα. Σε επίπεδο μοντέλων, το μοντέλο των σηματοδοτών πέτυχε $mAP@0.5 = 0,861$ και το μοντέλο των πλακιδίων 0,685, με τη μεγαλύτερη αδυναμία να εντοπίζεται στην ετερογενή κλάση των εμποδίων. Σε επίπεδο συσκευής, καταγράφηκαν χρόνοι εξαγωγής προβλέψεων 320–460 ms και πραγματικός ρυθμός 1,13–1,69 καρέ ανά δευτερόλεπτο, χαμηλότερος του σχεδιαστικού στόχου. Σε επίπεδο πεδίου, ο μηχανισμός ανακοίνωσης μόνο σε μεταβολή απέτρεψε το 86% των δυνητικών εκφωνήσεων, ενώ τα καταγεγραμμένα ψευδώς θετικά αποτελέσματα εμφάνισαν συστηματικά υψηλή εμπιστοσύνη, αποκλείοντας τη ρύθμιση του κατωφλίου ως μονομερή λύση. Σε επίπεδο υποκειμενικής εμπειρίας, η αξιολόγηση UEQ-S από 24 συμμετέχοντες απέδωσε βαθμολογίες στην κατηγορία «Excellent» του διεθνούς benchmark και στις τρεις κλίμακες (πραγματολογική, ηδονική, συνολική), αν και υπό συνθήκη προσομοιωμένης όχι πραγματικής οπτικής αναπηρίας. Η συζήτηση ανέδειξε ότι η κυρίαρχη μορφή αστοχίας είναι η μη ανίχνευση και όχι η εσφαλμένη ταξινόμηση, ότι η επιτυχία της προσέγγισης συνδέεται άμεσα με την ύπαρξη τυποποιημένης υποδομής, και ότι η χρήση του συστήματος πρέπει να οριοθετείται ρητά ως συμπληρωματική. Τα ευρήματα αυτά, μαζί με τους περιορισμούς της αξιολόγησης, συνθέτουν τη βάση πάνω στην οποία διατυπώνονται τα συμπεράσματα και οι προτάσεις του τελευταίου κεφαλαίου.

Κεφάλαιο 8ο: Συμπεράσματα

8.1 Σύνοψη της εργασίας

Η παρούσα εργασία σχεδίασε, υλοποίησε και αξιολόγησε το BlindNav, μια εφαρμογή Android για την υποβοήθηση της πεζής μετακίνησης ατόμων με οπτική αναπηρία. Το σύστημα αναγνωρίζει, μέσω δύο μοντέλων YOLO11n που εκτελούνται τοπικά στη συσκευή, απτικά πλακίδια όδευσης και φωτεινούς σηματοδότες πεζών, και αποδίδει την πληροφορία με ελληνική φωνητική και απτική ανάδραση. Η εναλλαγή μεταξύ των δύο λειτουργιών γίνεται αυτόματα, με βάση την κλίση της συσκευής, χωρίς να απαιτείται καμία οπτική αλληλεπίδραση με τη διεπαφή. Πέραν της αξιολόγησης των μοντέλων στο σύνολο επικύρωσης, η συμπεριφορά του συστήματος δοκιμάστηκε σε δύο διαδοχικούς γύρους δοκιμών πεδίου, σε τρεις συσκευές Android διαφορετικής κατηγορίας υλικού και συνολική διάρκεια άνω της μίας ώρας συνεχούς λειτουργίας.

8.2 Απάντηση στα ερευνητικά ερωτήματα

Ως προς την αξιοπιστία της αναγνώρισης (Ερώτημα 1), τα αποτελέσματα είναι θετικά, αλλά παρουσιάζουν διαφορές μεταξύ των επιμέρους κλάσεων. Το μοντέλο των σηματοδοτών πέτυχε $mAP@0.5 = 0,861$, ενώ οι τέσσερις κρίσιμες κλάσεις παρουσίασαν τιμές άνω του 0,91. Για το μοντέλο των πλακιδίων, η αντίστοιχη τιμή ήταν 0,685. Η ανάλυση ανά κλάση έδειξε υψηλότερη απόδοση σε κατηγορίες με πιο συγκεκριμένη οπτική μορφή και χαμηλότερη σε κατηγορίες με μεγαλύτερη ετερογένεια, με χαρακτηριστικό παράδειγμα την κλάση των εμποδίων. Στο σύνολο επικύρωσης δεν καταγράφηκε σύγχυση μεταξύ των πλακιδίων προειδοποίησης και κατεύθυνσης, δηλαδή μεταξύ δύο κλάσεων με διαφορετική σημασιολογία. Ωστόσο, κατά τις δοκιμές πεδίου καταγράφηκαν περιστασιακά περιπτώσεις όπου οι δύο κλάσεις συγχέονταν. Το αποτέλεσμα αυτό δείχνει ότι η αξιολόγηση σε στατικές εικόνες δεν αποτυπώνει απαραίτητα όλες τις συμπεριφορές του μοντέλου σε πραγματικές συνθήκες χρήσης. Αντίστοιχα, η αρχικά περιορισμένη εμφάνιση της κλάσης κατεύθυνσης σε σύντομη δοκιμή πεδίου δεν επιβεβαιώθηκε στον δεύτερο, εκτενέστερο γύρο δοκιμών. Το εύρημα αυτό δείχνει ότι συμπεράσματα που βασίζονται σε περιορισμένο αριθμό δοκιμών μπορεί να μην είναι αντιπροσωπευτικά της συνολικής συμπεριφοράς του συστήματος. Ως προς τη λειτουργία σε πραγματικό χρόνο (Ερώτημα 2), η τοπική εκτέλεση αποδείχθηκε εφικτή και σταθερή σε συνεδρίες διάρκειας άνω των δεκαπέντε λεπτών. Ο χρόνος εξαγωγής πρόβλεψης κυμάνθηκε μεταξύ 320 και 460 ms, ανάλογα με τη συσκευή. Ο πραγματικός ρυθμός επεξεργασίας ήταν 1,13–1,69 καρτέ ανά δευτερόλεπτο και παρέμεινε χαμηλότερος από τον αρχικό σχεδιαστικό στόχο, κυρίως λόγω του μηχανισμού χρονικής ρύθμισης. Η χρήση τριών συσκευών διαφορετικής κατηγορίας ανέδειξε επίσης την επίδραση του υλικού στην απόδοση. Η συσκευή μεσαίας κατηγορίας παρουσίασε χρόνους εξαγωγής περίπου 11% υψηλότερους από την αντίστοιχη συσκευή ανώτερης κατηγορίας. Το στοιχείο αυτό θα πρέπει να ληφθεί υπόψη σε πιθανή μελλοντική διάθεση της εφαρμογής σε μεγαλύτερο εύρος συσκευών. Ως προς τη σταθεροποίηση της καθοδήγησης (Ερώτημα 3), τα αποτελέσματα είναι σαφώς θετικά και αποτελούν ένα από τα σημαντικότερα ευρήματα της εργασίας. Ο συνδυασμός χρονικής εξομάλυνσης, ιεράρχησης προτεραιότητας και ανακοίνωσης μόνο όταν μεταβάλλεται η αναγνωρισμένη κατάσταση περιορίσε κατά 86% τις δυνητικές εκφωνήσεις. Παράλληλα, το υπολογιστικό κόστος του μηχανισμού παρέμεινε κάτω από ένα χιλιοστό του δευτερολέπτου. Έτσι, οι μεμονωμένες αστοχίες μεταξύ διαδοχικών καρτέ είχαν σημαντικά μικρότερη επίδραση στην τελική συμπεριφορά του συστήματος. Ως προς την εναλλαγή χωρίς οπτική αλληλεπίδραση (Ερώτημα 4), ο μηχανισμός που βασίζεται στην κλίση της συσκευής λειτούργησε σύμφωνα με τον σχεδιασμό κατά τις δοκιμές πεδίου. Ωστόσο, η αξιολόγηση

πραγματοποιήθηκε από βλέποντα χειριστή. Επομένως, η καταλληλότητα του συγκεκριμένου μηχανισμού για πραγματική χρήση από άτομα με οπτική αναπηρία δεν έχει ακόμη επιβεβαιωθεί.

8.3 Όρια χρήσης του συστήματος

Η αξιολόγηση ανέδειξε τρεις βασικούς περιορισμούς με άμεση σημασία για την ασφαλή χρήση του συστήματος.

Πρώτον, καταγράφηκαν περιπτώσεις στις οποίες κόκκινη ένδειξη φωτεινού σηματοδότη ταξινομήθηκε ως πράσινη. Δεύτερον, η ανίχνευση εμποδίων αποδείχθηκε περιορισμένη, καθώς το 78% των πραγματικών εμποδίων δεν ανιχνεύθηκε. Τρίτον, η ανάλυση των ψευδώς θετικών αποτελεσμάτων στις δοκιμές πεδίου έδειξε ότι τα σφάλματα δεν κατανέμονται ομοιόμορφα μεταξύ των διαφορετικών τύπων ανακοινώσεων. Ειδικότερα, το τέλος πεζοδρομίου, με 36, 4% των σχετικών ανακοινώσεων, και η διάβαση πεζών, με 20, 1%, παρουσίασαν υψηλότερα ποσοστά εσφαλμένης ανακοίνωσης από τον μέσο όρο. Για τον λόγο αυτό, το BlindNav πρέπει να αντιμετωπίζεται ως συμπληρωματικό βοήθημα και όχι ως υποκατάστατο του λευκού μπαστουνιού, του σκύλου οδηγού ή της εκπαίδευσης προσανατολισμού και κινητικότητας. Η ανακοίνωση πράσινης φάσης δεν θα πρέπει να θεωρείται από μόνη της ένδειξη ότι είναι ασφαλές να πραγματοποιηθεί η διάσχιση. Αντίστοιχα, η απουσία ανακοίνωσης εμποδίου δεν αποτελεί ένδειξη ότι η διαδρομή είναι ελεύθερη. Η σαφής οριοθέτηση αυτή είναι απαραίτητη για οποιαδήποτε μελλοντική χρήση του συστήματος σε πραγματικές συνθήκες..

8.4 Μελλοντικές επεκτάσεις

Αξιολόγηση με άτομα με πραγματική οπτική αναπηρία. Η προκαταρκτική αξιολόγηση UEQ-S (Κεφάλαιο 6.) απέδωσε εξαιρετικά αποτελέσματα με 24 βλέποντες συμμετέχοντες υπό συνθήκη προσομοιωμένης τύφλωσης, αποτελεί όμως μόνο ένα πρώτο, ενθαρρυντικό βήμα. Η πλέον επείγουσα επέκταση παραμένει η επανάληψη ή διεύρυνση της αξιολόγησης αυτής με άτομα με πραγματική, μόνιμη οπτική αναπηρία. Η συνεργασία με φορείς όπως ο Φάρος Τυφλών της Ελλάδος ή ο Πανελλήνιος Σύνδεσμος Τυφλών θα επέτρεπε δομημένη αξιολόγηση, με βάση το πλαίσιο του προτύπου ISO 9241-11 [44], και θα απαντούσε σε ερωτήματα που η παρούσα εργασία δεν μπορεί να καλύψει: αν η ανάδραση είναι κατανοητή, αν ο ρυθμός των ανακοινώσεων είναι ανεκτός και αν η καθοδήγηση οδηγεί όντως σε ασφαλέστερη μετακίνηση. Εμπλουτισμός των δεδομένων εκπαίδευσης. Απαιτείται κατά προτεραιότητα για την κλάση των εμποδίων, όπου η ετερογένεια της κατηγορίας απαιτεί σαφώς μεγαλύτερο και ποικιλότερο σύνολο, καθώς και για την κλάση της κόκκινης αντίστροφης μέτρησης, η οποία δεν αξιολογήθηκε καθόλου. Δεδομένου ότι τα καταγεγραμμένα ψευδώς θετικά παρήχθησαν με υψηλή εμπιστοσύνη, ιδιαίτερη αξία έχει η προσθήκη αρνητικών παραδειγμάτων από περιβάλλοντα οπτικά παρόμοια με τους στόχους. Βελτιστοποίηση της απόδοσης. Η γεφύρωση του κενού μεταξύ πραγματικού και θεωρητικού ρυθμού απαιτεί αναθεώρηση του μηχανισμού χρονικής ρύθμισης. Συμπληρωματικά, η αξιοποίηση πολλαπλών νημάτων στον διερχομό, ο κβαντισμός των μοντέλων ή η επανεξέταση της επιτάχυνσης μέσω GPU ή NPU θα μπορούσαν να μειώσουν τον χρόνο εξαγωγής προβλέψεων. Χρήσιμη θα ήταν επίσης η συστηματική μελέτη της θερμικής συμπεριφοράς σε παρατεταμένη χρήση. Επέκταση της λειτουργικότητας. Η εκτίμηση απόστασης εμποδίων, η ενσωμάτωση δορυφορικού εντοπισμού για πλοήγηση σε επίπεδο διαδρομής, η υποστήριξη νυχτερινών συνθηκών και η μεταφορά σε πλατφόρμα iOS αποτελούν φυσικές επεκτάσεις. Ενδιαφέρουσα κατεύθυνση αποτελεί επίσης η αξιοποίηση των δεδομένων της εφαρμογής για την αυτόματη χαρτογράφηση της κατάστασης των οδών, με δυνητική αξία για τους αρμόδιους φορείς. Πειραματική σύγκριση αρχιτεκτονικών. Η ποσοτική αντιπαραβολή του YOLO11n με εναλλακτικά

μοντέλα, καθώς και με τις προσεγγίσεις σημασιολογικής κατάταξης που κυριαρχούν στη σχετική βιβλιογραφία, θα τεκμηριώνει ισχυρότερα τη σχεδιαστική επιλογή.

8.5 Καταληκτικά σχόλια

Η εργασία έδειξε ότι η αναγνώριση απτικών πλακιδίων και φωτεινών σηματοδοτών με τα μέσα ενός συμβατικού κινητού τηλεφώνου είναι τεχνικά εφικτή, χωρίς πρόσθετο εξοπλισμό και χωρίς παρέμβαση στην αστική υποδομή. Το εύρημα αυτό έχει σημασία που υπερβαίνει τη συγκεκριμένη υλοποίηση: υποδεικνύει ότι η υπάρχουσα, ήδη εγκατεστημένη υποδομή προσβασιμότητας μπορεί να αποκτήσει πρόσθετη αξία μέσω λογισμικού, χωρίς νέο κόστος υποδομής. Ταυτόχρονα, η αξιολόγηση κατέδειξε με σαφήνεια την απόσταση που χωρίζει ένα λειτουργικό πρωτότυπο από ένα αξιόπιστο βοήθημα ασφάλειας. Οι αδυναμίες που καταγράφηκαν, η αποτυχία ανίχνευσης εμποδίων, οι σπάνιες αλλά κρίσιμες συγχύσεις στη φάση του σηματοδότη, και κυρίως η απουσία αξιολόγησης από τους ίδιους τους τελικούς χρήστες, δεν αποτελούν λεπτομέρειες υλοποίησης αλλά ουσιώδεις προϋποθέσεις που πρέπει να ικανοποιηθούν πριν από οποιαδήποτε πραγματική χρήση. Η ειλικρινής καταγραφή τους θεωρείται εξίσου σημαντική με τα θετικά αποτελέσματα, καθώς σε ένα σύστημα που απευθύνεται σε ευάλωτους χρήστες η υπερεκτίμηση των δυνατοτήτων συνιστά αυτή καθαυτή κίνδυνο.

8.6 Επίλογος

Το παρόν κεφάλαιο συνόψισε τη συνεισφορά της εργασίας, απάντησε στα τέσσερα ερευνητικά ερωτήματα που τέθηκαν στην Ενότητα 1.4, όρισε ρητά τα όρια χρήσης του συστήματος και πρότεινε συγκεκριμένες κατευθύνσεις μελλοντικής επέκτασης. Με αυτό ολοκληρώνεται το κύριο σώμα της διπλωματικής εργασίας· ακολουθούν η βιβλιογραφία και τα παραρτήματα.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] “Vision impairment and blindness.” Accessed: Aug. 13, 2026. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment>
- [2] Y. Ma *et al.*, “Evaluating the effectiveness of crosswalk tactile paving on street-crossing behavior: A field trial study for people with visual impairment,” *Accident Analysis & Prevention*, vol. 163, p. 106420, Dec. 2021, doi: 10.1016/j.aap.2021.106420.
- [3] L. Mitropoulos, C. Karolemeas, S. Tsigdinos, A. Vassi, and E. Bakogiannis, “A composite index for assessing accessibility in urban areas: A case study in Central Athens, Greece,” *Journal of Transport Geography*, vol. 108, p. 103566, Apr. 2023, doi: 10.1016/j.jtrangeo.2023.103566.
- [4] F. E.-Z. El-Taher, L. Miralles-Pechuan, J. Courtney, K. Millar, C. Smith, and S. McKeever, “A Survey on Outdoor Navigation Applications for People With Visual Impairments,” *IEEE Access*, vol. 11, pp. 14647–14666, 2023, doi: 10.1109/ACCESS.2023.3244073.
- [5] “ISO 23599:2019 - Assistive products for blind and vision-impaired persons — Tactile walking surface indicators.” Accessed: Aug. 13, 2026. [Online]. Available: <https://www.iso.org/standard/76106.html>
- [6] “Παρουσίαση στατιστικών δεδομένων για άτομα με προβλήματα όρασης που επιδοτούνται από τον ΟΠΕΚΑ.” Accessed: Aug. 13, 2026. [Online]. Available: <https://meallamatia.gr/parousiasi-statistikon-dedomenon-gia-atoma-me-provlimata-orasis-pou-epidotountai-apo-ton-opeka/>
- [7] “Tactile Paving Slab | Miyake, Seiichi | V&A Explore The Collections.” Accessed: Aug. 13, 2026. [Online]. Available: <https://collections.vam.ac.uk/item/O1360400/tactile-paving-slab-tile-miyake-seiichi/>
- [8] “Πρότυπα - Υπουργείο Περιβάλλοντος και Ενέργειας.” Accessed: Aug. 13, 2026. [Online]. Available: <https://ypen.gov.gr/chorikos-schediasmos/astikos-schediasmos/politikes-kai-protypa/>
- [9] press, “Τεχνικές Προδιαγραφές Μελέτης Προσβασιμότητας για κτίρια και υπαίθριους κοινόχρηστους χώρους,” Υπουργείο Περιβάλλοντος και Ενέργειας. Accessed: Sep. 04, 2026. [Online]. Available: <https://ypen.gov.gr/technikes-prodiagrafes-meletis-prosvasimotitas-gia-ktiria-kai-ypaithrious-koinochristous-chorous/>
- [10] “Άρθρο 26 — N. 4067/2012 | TaxHeaven.” Accessed: Aug. 13, 2026. [Online]. Available: <https://www.taxheaven.gr/law/4067/2012/arthro/26>
- [11] “Άρθρο 223 — N. 5306/2026 | TaxHeaven.” Accessed: Sep. 04, 2026. [Online]. Available: <https://www.taxheaven.gr/law/5306/2026/arthro/223>
- [12] Y. Ito, C. Premachandra, S. Sumathipala, H. W. H. Premachandra, and B. S. Sudantha, “Tactile Paving Detection by Dynamic Thresholding Based on HSV Space Analysis for Developing a Walking Support System,” *IEEE Access*, vol. 9, pp. 20358–20367, 2021, doi: 10.1109/ACCESS.2021.3055342.
- [13] J. Redmon and A. Farhadi, “YOLO9000: Better, Faster, Stronger,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI: IEEE, Jul. 2017, pp. 6517–6525. doi: 10.1109/CVPR.2017.690.

- [14] J. Redmon and A. Farhadi, “YOLOv3: An Incremental Improvement,” 2018, *arXiv*. doi: 10.48550/ARXIV.1804.02767.
- [15] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” 2020, *arXiv*. doi: 10.48550/ARXIV.2004.10934.
- [16] G. Jocher, J. Qiu, M. Liu, S. Lyu, F. C. Akyon, and M. E. Kalfaoglu, *Ultralytics YOLO26: Unified Real-Time End-to-End Vision Models*. (2026). Python. doi: 10.48550/arXiv.2606.03748.
- [17] “Ultralytics YOLO11,” Ultralytics Docs. Accessed: Sep. 05, 2026. [Online]. Available: <https://docs.ultralytics.com/models/yolo11>
- [18] “LiteRT overview | Google AI Edge | Google for Developers.” Accessed: Aug. 13, 2026. [Online]. Available: <https://developers.google.com/edge/litert/overview>
- [19] A. Aktaş, Ö. Demir, and B. Dogan, “Derin öğrenme yöntemleri ile dokunsal parke yüzeyi tespiti,” *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, vol. 35, pp. 1685–1700, Apr. 2020, doi: 10.17341/gazimmfd.652101.
- [20] Z. Zhang, “Research on Object Detection Based on YOLO and Its Application in Tactile Paving Recognition,” *HSET*, vol. 122, pp. 64–76, Dec. 2024, doi: 10.54097/wps8rx55.
- [21] Z. Cao, X. Xu, B. Hu, and M. Zhou, “Rapid Detection of Blind Roads and Crosswalks by Using a Lightweight Semantic Segmentation Network,” *IEEE Trans. Intell. Transport. Syst.*, vol. 22, no. 10, pp. 6188–6197, Oct. 2021, doi: 10.1109/TITS.2020.2989129.
- [22] X. Zhang, L. Liang, S. Zhao, and Z. Wang, “GRFB-UNet: A new multi-scale attention network with group receptive field block for tactile paving segmentation,” *Expert Systems with Applications*, vol. 238, p. 122109, Mar. 2024, doi: 10.1016/j.eswa.2023.122109.
- [23] Y. Yamanaka, E. Takaya, and S. Kurihara, “Tactile Tile Detection Integrated with Ground Detection using an RGB-Depth Sensor:,” in *Proceedings of the 12th International Conference on Agents and Artificial Intelligence*, Valletta, Malta: SCITEPRESS - Science and Technology Publications, 2020, pp. 750–757. doi: 10.5220/0009092907500757.
- [24] T. Takano, T. Nakane, J. Yu, and C. Zhang, “Tactile Paving Detection and Tracking Using Tenji10K Dataset,” *IEEJ Transactions Elec Engng*, vol. 19, no. 10, pp. 1661–1672, Oct. 2024, doi: 10.1002/tee.24123.
- [25] K. Park *et al.*, “SideGuide: A Large-scale Sidewalk Dataset for Guiding Impaired People,” in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Las Vegas, NV, USA: IEEE, Oct. 2020, pp. 10022–10029. doi: 10.1109/IROS45743.2020.9340734.
- [26] H. Hwang *et al.*, “GuideTWSI: A Diverse Tactile Walking Surface Indicator Dataset from Synthetic and Real-World Images for Blind and Low-Vision Navigation,” 2026, *arXiv*. doi: 10.48550/ARXIV.2603.07060.
- [27] A. Bin Mohamed Kassim *et al.*, “Vision-Based Tactile Paving Detection Method in Navigation Systems for Visually Impaired Persons,” in *Advances in Human and Machine Navigation Systems*, R. Róka, Ed., IntechOpen, 2019. doi: 10.5772/intechopen.79886.
- [28] M. C. Ghilardi, R. C. O. Macedo, and I. H. Manssour, “A New Approach for Automatic Detection of Tactile Paving Surfaces in Sidewalks,” *Procedia Computer Science*, vol. 80, pp. 662–672, 2016, doi: 10.1016/j.procs.2016.05.356.

- [29] M. S. K. Yeo, J. J. J. Pey, and M. R. Elara, “Passive Auto-Tactile Heuristic (PATH) Tiles: Novel Robot-Inclusive Tactile Paving Hazard Alert System,” *Buildings*, vol. 13, no. 10, p. 2504, Oct. 2023, doi: 10.3390/buildings13102504.
- [30] S. Mascetti, D. Ahmetovic, A. Gerino, C. Bernareggi, M. Busso, and A. Rizzi, “Robust traffic lights detection on mobile devices for pedestrians with visual impairment,” *Computer Vision and Image Understanding*, vol. 148, pp. 123–135, Jul. 2016, doi: 10.1016/j.cviu.2015.11.017.
- [31] S. Yu, H. Lee, and J. Kim, “LYTNet: A Convolutional Neural Network for Real-Time Pedestrian Traffic Lights and Zebra Crossing Recognition for the Visually Impaired,” in *Computer Analysis of Images and Patterns*, vol. 11678, M. Vento and G. Percannella, Eds., in Lecture Notes in Computer Science, vol. 11678, Cham: Springer International Publishing, 2019, pp. 259–270. doi: 10.1007/978-3-030-29888-3_21.
- [32] N. Tzimos, I. Kyriazidis, G. Voutsakelis, S. Kontogiannis, and G. Kokkonis, “Interaction with Tactile Paving in a Virtual Reality Environment: Simulation of an Urban Environment for People with Visual Impairments,” *MTI*, vol. 9, no. 7, p. 71, Jul. 2025, doi: 10.3390/mti9070071.
- [33] G. Voutsakelis, I. Dimkaros, N. Tzimos, G. Kokkonis, and S. Kontogiannis, “Development and Evaluation of a Tool for Blind Users Utilizing AI Object Detection and Haptic Feedback,” *Machines*, vol. 13, no. 5, p. 398, May 2025, doi: 10.3390/machines13050398.
- [34] “Roboflow Documentation | Roboflow Docs.” Accessed: Aug. 13, 2026. [Online]. Available: <https://docs.roboflow.com/>
- [35] “ONNX | Home.” Accessed: Aug. 13, 2026. [Online]. Available: <https://onnx.ai/>
- [36] K. Hyodo, *Tools to convert ONNX files (NCHW) to TensorFlow format (NHWC)*. (Sep. 2022). Python. Accessed: Aug. 13, 2026. [Online]. Available: <https://github.com/PINTO0309/onnx2tf>
- [37] “CameraX overview | Android media | Android Developers.” Accessed: Aug. 13, 2026. [Online]. Available: <https://developer.android.com/media/camera/camerax?authuser=3>
- [38] “Pedestrian-Light-Crossing Model > Overview.” Accessed: Aug. 19, 2026. [Online]. Available: <https://universe.roboflow.com/pedestrian-light-crossing/pedestrian-light-crossing-uoz6o>
- [39] “TextToSpeech | API reference | Android Developers.” Accessed: Aug. 13, 2026. [Online]. Available: <https://developer.android.com/reference/android/speech/tts/TextToSpeech>
- [40] “Vibrator | API reference | Android Developers.” Accessed: Aug. 13, 2026. [Online]. Available: <https://developer.android.com/reference/android/os/Vibrator>
- [41] “VibrationEffect | API reference | Android Developers.” Accessed: Aug. 13, 2026. [Online]. Available: <https://developer.android.com/reference/android/os/VibrationEffect>
- [42] “About the Unified Modeling Language Specification Version 2.5.1.” Accessed: Aug. 13, 2026. [Online]. Available: <https://www.omg.org/spec/UML/2.5.1/About-UML>
- [43] G. Melissourgou and C. Katsanos, “UEQ-GR and UEQ-S-GR: Towards a Greek Adaptation of the User Experience Questionnaire and its Short Version,” in *Proceedings of the 3rd International Conference of the ACM Greek SIGCHI Chapter*, Syros Greece: ACM, Sep. 2025, pp. 203–211. doi: 10.1145/3749012.3749059.

- [44] “ISO 9241-11:2018 - Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts.” Accessed: Aug. 13, 2026. [Online]. Available: <https://www.iso.org/standard/63500.html>

ΠΑΡΑΡΤΗΜΑ Α : Πηγαίος Κώδικας , Καταγραφή Μετρήσεων και Παράμετροι Εφαρμογής

Στο παράρτημα αυτό παρατίθεται ο πηγαίος κώδικας δύο συστατικών της εφαρμογής, τα οποία τεκμηριώνουν άμεσα ευρήματα και ισχυρισμούς που αναπτύχθηκαν στα Κεφάλαια 3 και 7.

A.1 Καταγραφή μετρήσεων (MetricsLogger.kt).

Η κλάση αυτή αποτελεί το εργαλείο καταγραφής με το οποίο συλλέχθηκαν όλα τα πειραματικά δεδομένα πεδίου που παρουσιάζονται στις Ενότητες 5.2 και 5.3: ο χρόνος εξαγωγής προβλέψεων ανά καρέ, ο συνολικός χρόνος επεξεργασίας καρέ, το πλήθος και το είδος των ανιχνεύσεων, η σταθεροποιημένη χωρική ζώνη, και το τυχόν εκφωνηθέν μήνυμα (μέθοδος logInference), καθώς και οι σημάνσεις εδάφους αναφοράς μέσω των φυσικών πλήκτρων της συσκευής, στις οποίες βασίστηκε η αξιολόγηση ορθότητας του Πίνακα 5.7 (μέθοδος logEvent). Η εγγραφή σε ουρά μνήμης με περιοδικό άδειασμα (FLUSH_THRESHOLD) διασφαλίζει ότι το I/O δίσκου δεν επηρεάζει τις ίδιες τις μετρήσεις καθυστέρησης που καταγράφονται.

```
package com.example.blindnav

import android.content.Context
import android.os.Build
import java.io.File
import java.text.SimpleDateFormat
import java.util.Date
import java.util.Locale
import java.util.concurrent.ConcurrentLinkedQueue
import java.util.concurrent.atomic.AtomicBoolean

/** * Καταγραφή μετρήσεων για την πειραματική αξιολόγηση (Κεφ. 4-5).
* * Γράφει δύο αρχεία CSV στον φάκελο εξωτερικών αρχείων της
εφαρμογής: * - inference_<timestamp>.csv : μία γραμμή ανά εξαγωγή
προβλέψεων * - events_<timestamp>.csv : σημάνσεις εδάφους αναφοράς
(ground truth) * * Πρόσβαση στα αρχεία χωρίς root: * adb pull
/sdcard/Android/data/com.example.blindnav/files/metrics/ * ή μέσω
Android Studio: View > Tool Windows > Device Explorer * *
ΣΗΜΕΙΩΣΗ: η εγγραφή γίνεται σε ουρά στη μνήμη και αδειάζει σε
παρτίδες, * ώστε το I/O να μην επηρεάζει τη μέτρηση της καθυστέρησης.
*/

class MetricsLogger(context: Context) {

    private val dir: File = File(context.getExternalFilesDir(null),
"metrics").apply { mkdirs() }

    private val stamp =
SimpleDateFormat("yyyyMMdd_HHmms", Locale.US).format(Date())
    private val inferenceFile = File(dir, "inference_$stamp.csv")

    private val eventsFile = File(dir, "events_$stamp.csv")

    private val queue = ConcurrentLinkedQueue<String>()

    private val isFlushing = AtomicBoolean(false)

    /** Χρόνος έναρξης της συνεδρίας, για σχετικές
χρονοσημάνσεις σε ms. */

    private val sessionStart = System.currentTimeMillis()

    init {

        inferenceFile.writeText("t_ms, mode, inference_ms, total_ms,
detections, top_class, confidence, zone, announced, message, all_classes\n"
)

        eventsFile.writeText("t_ms, event, note\n")

        // Καταγραφή συσκευής, για την περιγραφή του πειραματικού
περιβάλλοντος

        File(dir, "device_$stamp.txt").writeText(
"model=${Build.MANUFACTURER} ${Build.MODEL}\n" +
"android=${Build.VERSION.RELEASE} (API
${Build.VERSION.SDK_INT})\n" +
"session_start=${Date(sessionStart)}\n" ) }
}
```

```
private fun now() = System.currentTimeMillis() - sessionStart

/** * Καλείται μετά από κάθε εξαγωγή προβλέψεων.
* * @param mode "tactile" ή "traffic" *

@param inferenceMs καθαρός χρόνος του
interpreter.run() *

@param totalMs συνολικός χρόνος καρέ
(προεπεξεργασία + inference + pipeline) *

@param detections πλήθος ανιχνεύσεων μετά το
φίλτρο εμπιστοσύνης *

@param topClass επικρατέστερη κλάση (ή null) *

@param confidence εμπιστοσύνη της επικρατέστερης
ανίχνευσης *

@param zone σταθεροποιημένη χωρική ζώνη από τον
TemporalSmoother *

@param announced αν ο StateTracker ενέκρινε
ανακοίνωση *

@param message το κείμενο που εκφωνήθηκε (κενό
αν δεν εκφωνήθηκε) *

@param allClasses όλες οι ανιχνεύσεις σε μορφή
κειμένου */

fun logInference(

    mode: String,

    inferenceMs: Long,

    totalMs: Long,

    detections: Int,

    topClass: String?,

    confidence: Float,

    zone: String?,

    announced: Boolean,

    message:

    String = "", allClasses: String = "" )

{ queue.add(

    listOf(

        now(),

        mode,

        inferenceMs,

        totalMs,

        detections,

        topClass ?: "none",
String.format(Locale.US, "%.3f", confidence), zone
?: "none",

        if (announced) 1 else 0,
message.replace(" ", ";").replace("\n", " "),
allClasses.replace(" ", ";").joinToString(", ") ) )

    if (queue.size >= FLUSH_THRESHOLD) flush() }

/** * Σήμανση εδάφους αναφοράς.
Καλείται από κουμπί στην οθόνη ή από φυσικό *
πλήκτρο έντασης, τη στιγμή που ο παρατηρητής
βλέπει το πραγματικό γεγονός. * Επιτρέπει τη
σύγκριση «τι ανακοίνωσε η εφαρμογή» με «τι υπήρχε
πραγματικά». * * Παραδείγματα: "warning_tile",
"directional_tile", "light_red", * "light_green",
"crossing_start", "false_alarm" */
```

Code 1 Koutline

```
fun logEvent(event: String, note: String = "") {
    eventsFile.appendText("${now()}, $event, ${note.replace(", ", ";")}\n")
}
/** Άδειασμα της ουράς στο δίσκο. Ασφαλές για κλήση από οποιοδήποτε νήμα. */
fun flush()
    { if (!isFlushing.compareAndSet(false, true)) return
    try {
        val batch = StringBuilder()
        while (true) {
            val line = queue.poll() ?: break
            batch.append(line).append('\n')
        }
        if (batch.isNotEmpty()) inferenceFile.appendText(batch.toString())
    } finally {
        isFlushing.set(false)
    }
}
/** Καλείται σε onPause()/onDestroy() ώστε να μη χαθούν εγγραφές. */
fun close() = flush()
companion object {
    private const val FLUSH_THRESHOLD = 50
}
}
```

A.2 Κεντρικές παράμετροι (AppConfig.kt)

Η κλάση αυτή συγκεντρώνει όλες τις ρυθμιζόμενες παραμέτρους της εφαρμογής σε ένα σημείο, ώστε η βαθμονόμηση του συστήματος να είναι κεντρική και τεκμηριώσιμη. Οι τιμές της, το ενιαίο κατώφλι εμπιστοσύνης 0, 45, το διάστημα εξαγωγής προβλέψεων 220 ms, το κατώφλι κλίσης -70° , το μέγεθος παραθύρου εξομάλυνσης 5 καρτέ, και το ελάχιστο διάστημα δόνησης 150 ms, παρουσιάζονται και αιτιολογούνται στην Ενότητα 3.8.

<pre> package com.example.blindnav.config /** * thresholds της εφαρμογής. * Όλα τα configurable values μαζεμένα εδώ για εύκολο calibration */ object AppConfig { // — Inference _____ /** Χρονικό interval μεταξύ consecutive inferences (ms) — 100ms για ~5 frames/500ms */ const val INFERENCE_INTERVAL_MS = 220L /** Ελάχιστο confidence score για non-tile detections (εμπόδια, διαβάσεις) */ const val MIN_CONFIDENCE_THRESHOLD = 0.45f /** Confidence threshold για πλακίδια — χαμηλότερο γιατί ο νέος YOLO αναγνωρίζει καλύτερα */ const val TILE_CONFIDENCE_THRESHOLD = 0.45f /** Confidence threshold για το μοντέλο φαναριών (υψηλότερο λόγω none class) */ const val TRAFFIC_LIGHT_CONFIDENCE_THRESHOLD = 0.45f // — Spatial Zones (Χωρικές Ζώνες) _____ /** Όριο αριστερής ζώνης (normalized x coordinate) */ const val ZONE_LEFT_BOUNDARY = 0.33f /** Όριο δεξιάς ζώνης (normalized x coordinate) */ const val ZONE_RIGHT_BOUNDARY = 0.66f // — Obstacle Proximity _____ /** Μέγιστη εκτιμώμενη απόσταση (μέτρα) για να θεωρηθεί obstacle CRITICAL */ const val OBSTACLE_DISTANCE_THRESHOLD_METERS = 1.5f /** Fallback: ελάχιστο bounding box width (normalized) αν depth δεν είναι διαθέσιμο */ const val OBSTACLE_PROXIMITY_THRESHOLD = 0.20f // — State Change / Cooldown _____ /** Ελάχιστος χρόνος (ms) πριν επαναληφθεί το ίδιο μήνυμα */ const val MIN_REPEAT_INTERVAL_MS = 5000L /** Τα CRITICAL μηνύματα μπορούν να επαναληφθούν πιο γρήγορα (ms) */ const val CRITICAL_REPEAT_INTERVAL_MS = 2000L // — Haptic / Vibration _____ /** Threshold γωνίας (μοίρες) για "ευθεία" στο heading lock */ const val HEADING_STRAIGHT_THRESHOLD_DEGREES = 10f /** Ελάχιστο interval μεταξύ δονήσεων (ms) για αποφυγή σύγχυσης */ const val HAPTIC_MIN_INTERVAL_MS = 150L // — Temporal Smoothing _____ /** Μέγεθος παραθύρου μνήμης frames */ const val SMOOTHING_WINDOW_SIZE = 5 </pre>	<pre> /** Ελάχιστα frames για σταθεροποίηση event (εμπόδιο κλπ) */ const val EVENT_MIN_FRAMES = 2 /** Ελάχιστα frames για σταθεροποίηση πλακιδίων — 2 γιατί ο νέος YOLO είναι πιο αξιόπιστος */ const val TILES_MIN_FRAMES = 2 /** Μέγεθος παραθύρου για φανάρια (μικρότερο = ταχύτερη απόκριση) */ const val TRAFFIC_SMOOTHING_WINDOW_SIZE = 3 /** Ελάχιστα frames για φανάρι */ const val TRAFFIC_MIN_FRAMES = 2 // — Device Orientation _____ /** Pitch angle threshold (μοίρες) για εναλλαγή walking/traffic light mode */ const val PITCH_THRESHOLD_DEGREES = -70f // — Model Files _____ /** Filename του TFLite μοντέλου πλακιδίων στο assets folder */ const val TACTILE_MODEL_FILENAME = "tactile_model.tflite" /** Filename του TFLite μοντέλου φαναριών στο assets folder */ const val TRAFFIC_MODEL_FILENAME = "traffic_model.tflite" /** Input size που περιμένει το YOLO11n μοντέλο (pixels) */ const val MODEL_INPUT_SIZE = 640 } </pre>
--	--