



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ**

**ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ**

**Τμ. ΜΗΧΑΝΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΗΣ ΣΧΕΔΙΑΣΗΣ ΚΑΙ ΠΑΡΑΓΩΓΗΣ**

***Π.Μ.Σ. «ΜΗ ΕΠΑΝΑΡΩΜΕΝΑ ΑΥΤΟΝΟΜΑ ΚΑΙ  
ΤΗΛΕΚΑΤΕΥΘΥΝΟΜΕΝΑ ΣΥΣΤΗΜΑΤΑ»***

---



**«ΣΧΕΔΙΑΣΗ ΣΥΣΤΗΜΑΤΟΣ ΣΤΑΘΜΟΥ ΕΓΚΑΙΡΗΣ ΠΡΟΕΙΔΟΠΟΙΗΣΗΣ ΜΕ  
ΧΡΗΣΗ ΜΗΧΑΝΙΚΗΣ ΟΡΑΣΗΣ ΚΑΙ ΑΥΤΟΝΟΜΗΣ ΠΛΟΗΓΗΣΗΣ UV»**

**“DESIGN AND IMPLEMENTATION OF A UNIFIED SURVEILLANCE SYSTEM  
WITH COMPUTER VISION, INCLUDING A SURVEILLANCE STATION AND AN  
AUTONOMOUS UV”**

**Σπουδαστής: Καρανάτσιος Δημήτριος(ΑΜ 8096606)**

**Υπεύθυνος Καθηγητής: Παπακίτσος Ευάγγελος**

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

**ΑΘΗΝΑ ΣΕΠΤΕΜΒΡΙΟΣ 2023**

**BLANK PAGE / ΣΕΛΙΔΑ ΚΕΝΗ**

## Μέλη Εξεταστικής Επιτροπής

1. Παπακίτσος Ευάγγελος
2. Παπουτσιδάκης Μιχαήλ
3. Χατζόπουλος Αβραάμ

1.

## ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

Ο κάτωθι υπογεγραμμένος Καρανάτσιος Δημήτριος του Γεωργίου, με αριθμό μητρώου 8096606 φοιτητής του Προγράμματος Μεταπτυχιακών Σπουδών «Μη Επανδρωμένα Αυτόνομα και Τηλεκατευθυνόμενα Συστήματα» του Τμήματος Μηχανικών Βιομηχανικής Σχεδίασης και Παραγωγής της Σχολής Μηχανικών Πανεπιστημίου Δυτικής Αττικής, δηλώνω υπεύθυνα ότι:

«Είμαι συγγραφέας αυτής της μεταπτυχιακής εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, οι όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε ακριβώς είτε παραφρασμένες, αναφέρονται στο σύνολό τους, με πλήρη αναφορά στους συγγραφείς, τον εκδοτικό οίκο ή το περιοδικό, συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο. Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από μένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου, όσο και του Ιδρύματος.

Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του διπλώματός μου».

Ο Δηλών

Καρανάτσιος Δημήτριος

Ημερομηνία

Σεπτέμβριος 2023



**«ΣΧΕΔΙΑΣΗ ΣΥΣΤΗΜΑΤΟΣ ΣΤΑΘΜΟΥ ΕΓΚΑΙΡΗΣ ΠΡΟΕΙΔΟΠΟΙΗΣΗΣ  
ΜΕ ΧΡΗΣΗ ΜΗΧΑΝΙΚΗΣ ΟΡΑΣΗΣ ΚΑΙ ΑΥΤΟΝΟΜΗΣ ΠΛΟΗΓΗΣΗΣ UV»**

**ΚΑΡΑΝΑΤΣΙΟΣ ΔΗΜΗΤΡΙΟΣ**

Μεταπτυχιακή διπλωματική εργασία που υποβάλλεται στο καθηγητικό σώμα για την μερική εκπλήρωση των υποχρεώσεων απόκτησης του μεταπτυχιακού τίτλου του Προγράμματος Μεταπτυχιακών Σπουδών «Μη Επανδρωμένα Αυτόνομα και Τηλεκατευθυνόμενα Συστήματα» του Τμήματος Μηχανικών Βιομηχανικής Σχεδίασης και Παραγωγής του Πανεπιστημίου Δυτικής Αττικής.

## Περίληψη

Στην συγκεκριμένη διπλωματική εργασία, παρουσιάζεται ένα μοντέλο ενοποιημένου συστήματος σταθμού έγκαιρης προειδοποίησης – μη επανδρωμένου οχήματος. Βασικό χαρακτηριστικό του συστήματος αυτού, αποτελεί η χρήση μηχανικής όρασης για τον εντοπισμό των αντικειμένων – στόχων. Για την πληρέστερη παρουσίαση των δυνατοτήτων του συστήματος, η ανάπτυξή του έχει γίνει σε μορφή σταθερού και κινητού σταθμού.

Αρχικά, γίνεται μια θεωρητική παρουσίαση και ανάλυση του τομέα της μηχανικής μάθησης καθώς και των βασικών μοντέλων και εργαλείων που χρησιμοποιήθηκαν. Ακολουθεί η ανάλυση λειτουργίας του συστήματος και του υλικού που χρησιμοποιήθηκε για τον σταθμό αναγνώρισης αλλά και την εξομοίωση των βασικών λειτουργιών του μη επανδρωμένου συστήματος. Τέλος, ακολουθεί η αναλυτική παρουσίαση των μορφών μηχανικής όρασης που αναπτύχθηκαν, οι οποίες καταδεικνύουν τις δυνατότητες του συστήματος, αλλά και τα πεδία πολιτικών και στρατιωτικών εφαρμογών τους.

**Λέξεις κλειδιά:** σταθμός παρατήρησης, λογισμικό, προγραμματισμός, Python, Android, Tensorflow, Tkinter, OpenCV 2, Raspberry Pi, γραφικό περιβάλλον

## **Abstract**

This thesis presents a unified surveillance system which consists of a surveillance station and an unmanned vehicle (UV). Main feature of this system, is the use of computer vision for the detection of target objects. For a better presentation of its capabilities, two different modes of surveillance stations have been developed, one of fixed position and one mobile.

Initially, there is a theoretical presentation of computer vision and especially of its main models and tools which have been used in this application. What follows is the analysis of main system functionality and the hardware which has been used for the surveillance station and the simulation of the unmanned vehicle. Finally, there is the presentation of all computer vision modes which were developed and are indicative for the capabilities of the system, as well as its possible civilian and military implementation.

**Key words:** surveillance station, software, programming, Python, Android, TensorFlow, Tkinter, OpenCV 2, Raspberry Pi, user interface

## ΠΕΡΙΕΧΟΜΕΝΑ

Περίληψη.....	6
Abstract.....	7
Περιεχόμενα.....	8-9
Κατάλογος πινάκων.....	9
Κατάλογος εικόνων.....	9
Κατάλογος διαγραμμάτων.....	10
1. Εισαγωγή.....	11
2. Μηχανική μάθηση.....	12-26
2.1. Νευρωνικά δίκτυα (Neural networks).....	14-18
2.2. Βαθιά μάθηση (Deep learning).....	19
2.3. Συνελκτικά νευρωνικά δίκτυα (CNN).....	19-20
2.4. TensorFlow.....	21-23
2.5. Keras.....	23-25
2.6. OpenCV 2.....	25-26
3. Ανάλυση συστήματος.....	27-31
4. Υλικό.....	32-38
4.1. Raspberry Pi 3 Model B.....	32-34
4.2. Κινητήρας servo (FS90MG).....	34-36
4.3. Samsung Galaxy A135G-SMA136B/DSN.....	37
4.4. Microsoft LifeCam Cinema USB Camera.....	37-38
4.5. Μηχανισμός στήριξης κινητήρων servo.....	38
5. Γραφικό περιβάλλον εφαρμογής.....	39-40
6. Περιγραφή αποστολής μη επανδρωμένου οχήματος (UV).....	41
7. Υποστηριζόμενες μορφές μηχανικής όρασης.....	41-50
7.1 Μη επανδρωμένο όχημα.....	42-46
7.1.1. Εντοπισμός - αναγνώριση αντικειμένων.....	42-43
7.1.2. Εντοπισμός - αναγνώριση προσώπου.....	44
7.1.3. Υπολογισμός απόστασης προσώπου.....	45
7.1.4. Εντοπισμός κίνησης.....	46
7.1.5. Ζωντανή ροή εικόνας από σταθερό σταθμό επιτήρησης.....	46
7.2. Κινητός σταθμός επιτήρησης.....	47
7.3. Σταθερός σταθμός επιτήρησης.....	47-50
7.3.1 Χειρισμός κίνησης κάμερας μέσω ζωντανής εικόνας.....	47-48
7.3.2. Εντοπισμός αντικειμένου και εστίαση σε αυτό.....	48-49



8. Συμπερασματικός επίλογος – Μελλοντική έρευνα.....	49-50
Βιβλιογραφία.....	51

### **Κατάλογος πινάκων**

<b>Πίνακας 1.</b> Διαφορές νευρωνικών δικτύων – υπολογιστών von Neumann.....	16
<b>Πίνακας 2.</b> Τεχνικά χαρακτηριστικά Raspberry Pi 3 Model B.....	32
<b>Πίνακας 3.</b> Τεχνικά χαρακτηριστικά κινητήρα servo (FS90MG).....	34-35
<b>Πίνακας 4.</b> Τεχνικά χαρακτηριστικά Samsung Galaxy A135G-SMA136B/DSN.....	37
<b>Πίνακας 5.</b> Χαρακτηριστικά λειτουργίας κάμερας Microsoft LifeCam.....	37

### **Κατάλογος εικόνων**

<b>Εικόνα 1.</b> Αποτέλεσμα εντοπισμού αντικειμένων.....	13
<b>Εικόνα 2.</b> Αρχιτεκτονική Tensorflow.....	23
<b>Εικόνα 3.</b> Δομή Keras.....	24
<b>Εικόνα 4.</b> Άνω όψη του Raspberry Pi 3 Model B.....	32
<b>Εικόνα 5.</b> Ακροδέκτες Raspberry Pi 3 Model B.....	33
<b>Εικόνα 6.</b> Κινητήρας servo FS90MG.....	35
<b>Εικόνα 7.</b> Παλμοί PWM και αντιστοιχία κύκλων εργασίας με τάσεις εξόδου.....	35
<b>Εικόνα 8.</b> Σχέση διάρκειας παλμών PWM με την γωνία άξονα κινητήρα servo.....	36
<b>Εικόνα 9.</b> Microsoft LifeCam Cinema USB Camera.....	38
<b>Εικόνα 10.</b> Μηχανισμός στήριξης - κίνησης κάμερας.....	38
<b>Εικόνα 11.</b> Γραφικό περιβάλλον εφαρμογής.....	39
<b>Εικόνα 12.</b> Εντοπισμός-αναγνώριση αντικειμένων “βάρκα” και “άνθρωπος”.....	42
<b>Εικόνα 13.</b> Εντοπισμός-αναγνώριση αντικειμένων “βιβλίο”.....	43
<b>Εικόνα 14.</b> Στιγμιότυπο αναγνώρισης προσώπου.....	44
<b>Εικόνα 15.</b> Στιγμιότυπο υπολογισμού απόστασης προσώπου.....	45
<b>Εικόνα 16.</b> Στιγμιότυπο εντοπισμού κίνησης.....	46
<b>Εικόνα 17.</b> Στιγμιότυπο απομακρυσμένου χειρισμού κίνησης κάμερας.....	48
<b>Εικόνα 18.</b> Στιγμιότυπο face tracking.....	49

## **Κατάλογος Διαγραμμάτων**

<b>Σχήμα 1.</b> Φάσεις ανάλυσης και επεξεργασίας πληροφοριών μέσω τεχνητής όρασης.....	14
<b>Σχήμα 2.</b> Παράδειγμα τεχνητού νευρωνικού δικτύου.....	15
<b>Σχήμα 3.</b> Φάσεις μηχανικής μάθησης.....	17
<b>Σχήμα 4.</b> Διάγραμμα αυτοκωδικοποιητή.....	18
<b>Σχήμα 5.</b> Λειτουργία συνελκτικών νευρωνικών δικτύων.....	20
<b>Σχήμα 6.</b> Ροή γενικής λειτουργίας συστήματος.....	28
<b>Σχήμα 7.</b> Ροή λειτουργίας σταθμού επιτήρησης.....	29
<b>Σχήμα 8.</b> Ροή λειτουργίας κατάστασης αναμονής μη επανδρωμένου οχήματος.....	30
<b>Σχήμα 9.</b> Ροή καταγραφής μεταβλητών λειτουργίας μη επανδρωμένου οχήματος.....	31
<b>Σχήμα 10.</b> Ροή ανανέωσης συντεταγμένων θέσης μη επανδρωμένου οχήματος.....	31

## **1. Εισαγωγή**

Δύο τομείς της τεχνολογίας που έχουν γνωρίσει αλματώδη ανάπτυξη τα τελευταία χρόνια, είναι αυτοί των μη επανδρωμένων οχημάτων και της μηχανικής όρασης. Η συνδυαστική τους χρήση μπορεί να βρει σημαντικές εφαρμογές στην υλοποίηση ολοκληρωμένων μηχανισμών έγκαιρης αναγνώρισης και αντιμετώπισης κινδύνων, τόσο στον πολιτικό, αλλά και στον στρατιωτικό τομέα.

Στόχος της συγκεκριμένης διπλωματικής εργασίας, είναι η ανάπτυξη και παρουσίαση ενός ενοποιημένου συστήματος παρατήρησης και εντοπισμού, με χρήση τεχνικών μηχανικής όρασης. Οι βασικές λειτουργίες των συστημάτων αυτού του είδους, μπορούν να συνοψιστούν στις εξής:

1. Εντοπισμός του στόχου ή απειλής.
2. Η ενέργεια αντίδρασης μετά τον εντοπισμό.

Η εμπειρία από τα σύγχρονα πεδία πολεμικών συγκρούσεων, καταδεικνύει πως είναι απολύτως εφικτή η κατασκευή τέτοιων αξιόπιστων συστημάτων, με χρήση συσκευών χαμηλού κόστους και εύκολα διαθέσιμων στο ευρύ κοινό, όπως είναι για παράδειγμα οι συσκευές κινητών τηλεφώνων. Ειδικά για χώρες με μικρούς οικονομικούς πόρους και έλλειψη στρατηγικού βάθους, όπως είναι η Ελλάδα, η χρήση αξιόπιστων μηχανισμών έγκαιρης προειδοποίησης χαμηλού κόστους αποτελεί μονόδρομο. Βάσει των παραπάνω, η παρούσα διπλωματική εργασία υλοποιήθηκε έχοντας τους εξής βασικούς στόχους:

1. Ανάπτυξη συστήματος εντοπισμού – έγκαιρης προειδοποίησης και ανάλυση των βασικών τεχνολογιών από τις οποίες διέπεται.
2. Ανάλυση των μέσων που μπορούν να χρησιμοποιηθούν για την κατασκευή ενός απλού και αξιόπιστου συστήματος.
3. Ανάδειξη των τομέων που μπορούν να ερευνηθούν περαιτέρω, προκειμένου να ενισχυθεί η εξειδίκευση και η ακρίβεια ενός συστήματος έγκαιρης προειδοποίησης, ειδικά για στρατιωτικούς σκοπούς.

## **2. Μηχανική μάθηση**

Η μηχανική όραση, αποτελεί ένα τμήμα του πολυδιάστατου τομέα της μηχανικής μάθησης. Ως μηχανική μάθηση, μπορεί να οριστεί το σύνολο των τεχνικών με τις οποίες μία μηχανή, όπως για παράδειγμα ένας υπολογιστής, μαθαίνει να αναλύει ορισμένα δεδομένα και να λαμβάνει τεκμηριωμένες αποφάσεις βάσει αυτών, χωρίς όμως να έχει προγραμματιστεί άμεσα και αποκλειστικά για αυτό.

Η μηχανική όραση αξιοποιείται σε μεγάλο βαθμό στην ιατρική, στη βιομηχανία και στα στρατιωτικά συστήματα. Στον στρατιωτικό τομέα, πεδίο εφαρμογής της αποτελούν ο εντοπισμός εχθρικών μονάδων και η καθοδήγηση πυραύλων σε επιλεγμένο στόχο με βάση δεδομένα εικόνας που λαμβάνονται πριν ή μετά την εκτόξευσή τους. Παράλληλα, εφαρμόζεται σε μη επανδρωμένα συστήματα, τα οποία λειτουργούν είτε αυτόνομα είτε βοηθητικά προς τον άνθρωπο χειριστή του οχήματος. Προκειμένου να καταστεί εφικτή η εκτέλεση των εργασιών αυτών, πρέπει πρώτα να ληφθούν οι πληροφορίες που πρόκειται να υποβληθούν σε επεξεργασία. Για να υλοποιηθεί αυτό ακολουθείται μια σειρά φάσεων, προκειμένου να ληφθούν οι πληροφορίες και να ακολουθήσει η ανάλυση τους. Οι φάσεις αυτές είναι οι εξής:

**1.Λήψη εικόνας:** Λήψη της εικόνας ή σειράς των εικόνων που θέλουμε να αναλύσουμε.

**2.Προεπεξεργασία:** Πριν εφαρμοστεί κάποια μέθοδος στην εικόνα που δημιουργήθηκε, η εικόνα πρέπει να ελεγχθεί εάν τηρεί κάποιες βασικές προδιαγραφές που θέτει η προς εφαρμογή μέθοδος. Ο έλεγχος περιλαμβάνει λειτουργίες, όπως επαναληπτική δειγματοληψία για να διασφαλιστεί ότι το σύστημα συντεταγμένων της εικόνας είναι το κατάλληλο και μείωση θορύβου για να διασφαλιστεί η απουσία λανθασμένης και άχρηστης πληροφορίας. Τέλος, ενισχύονται οι αντιθέσεις της εικόνας ώστε η ενδιαφέρουσα πληροφορία να εντοπίζεται ευκολότερα.

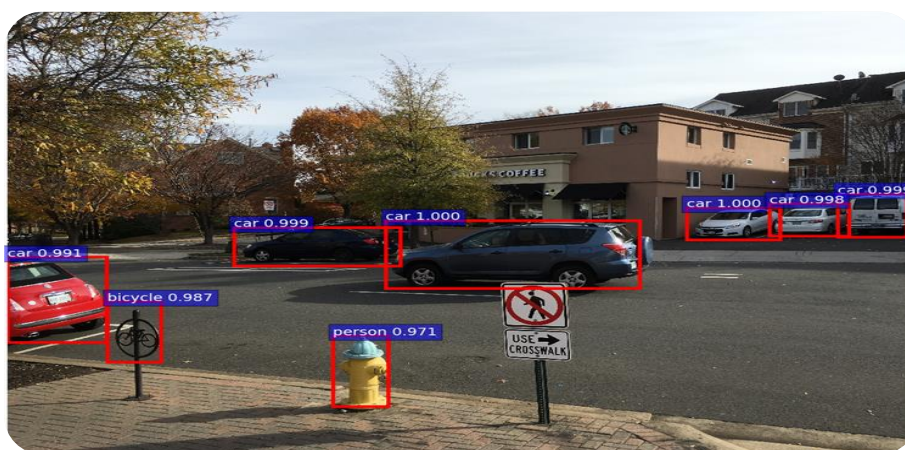
**3. Εξαγωγή χαρακτηριστικών:** Χαρακτηριστικά που υπάρχουν μέσα στην εικόνα σε διάφορους βαθμούς λεπτομέρειας, εξάγονται από την εικόνα. Τέτοια γνωρίσματα μπορεί να είναι σχήματα όπως γραμμές, ακμές, σημεία, γωνίες και πολλά άλλα.

**4. Εντοπισμός και τεμαχισμός:** Η εφαρμογή εφόσον εντοπίσει τα ενδιαφέροντα χαρακτηριστικά της εικόνας, τα εξάγει από την υπόλοιπη εικόνα και τα προωθεί σε άλλα υποσυστήματα της εφαρμογής.

5. Επεξεργασία υψηλού επιπέδου: Σε αυτό το σημείο, η είσοδος είναι ένα μικρό σύνολο δεδομένων, για παράδειγμα ένα σύνολο σημείων ή μια περιοχή της εικόνας, ή ένα συγκεκριμένο αντικείμενο. Ακολουθεί η επεξεργασία του συνόλου αυτού, όπως για παράδειγμα η εκτίμηση του μεγέθους του.

6. Αναγνώριση εικόνας: Ακολουθεί η κατηγοριοποίηση των αντικειμένων της εικόνας σε κάποιες προκαθορισμένες κατηγορίες.

7. Καταχώρηση εικόνας: Πρόκειται για τη σύγκριση και το συνδυασμό διαφορετικών όψεων των ίδιων αντικειμένων. Διαφορετικά ορίζεται και ως ταξινόμηση εικόνας.



**Εικόνα 1. Αποτέλεσμα εντοπισμού αντικειμένων**

(Moses Olafenwa, (16 June 2018), Object Detection with 10 lines of code, [medium.com, https://towardsdatascience.com/object-detection-with-10-lines-of-code-d6cb4d86f606](https://towardsdatascience.com/object-detection-with-10-lines-of-code-d6cb4d86f606))

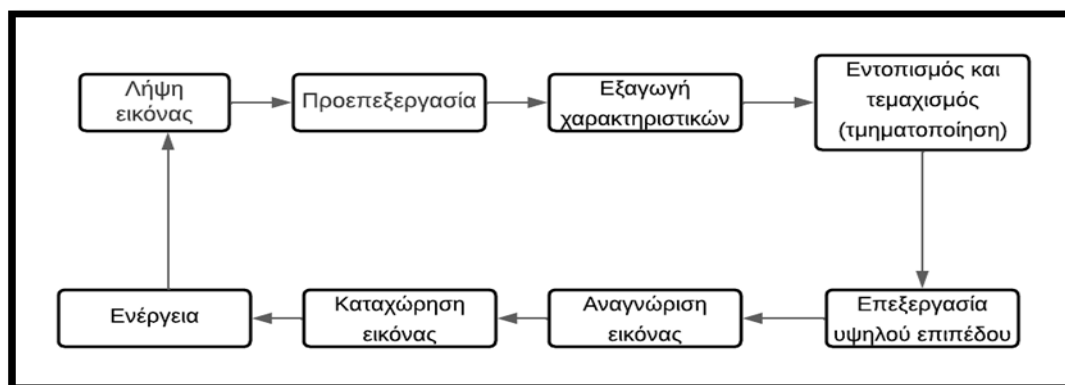
8. Ενέργεια: Με το αποτέλεσμα της ανίχνευσης μπορούμε να εφαρμόσουμε την ενέργεια για την οποία έχουμε αναπτύξει το σύστημά μας. Ορισμένες ενέργειες είναι :

8.1. Εντοπισμός αντικειμένου (object detection): Καθώς ο αριθμός των αντικειμένων που θα περιέχει μια εικόνα δεν είναι εκ των προτέρων γνωστός, με το object detection επιτυγχάνεται η εύρεση και η ταξινόμηση ενός μεταβλητού αριθμού αντικειμένων.

8.2. Ταυτοποίηση αντικειμένου (object identification): Σε αυτήν την περίπτωση, στόχος είναι η εύρεση συγκεκριμένων αντικειμένων σε εικόνες. Συνεπώς, προσδιορίζεται το αν εμφανίζεται το αντικείμενο σε μια εικόνα (frame). Αν αυτό συμβαίνει, εντοπίζεται και το σημείο της εικόνας στο οποίο εμφανίζεται το αντικείμενο.

8.3 Τμηματοποίηση αντικειμένου (object segmentation): Πρόκειται για τη δημιουργία μίας μάσκας για κάθε ανιχνευόμενο αντικείμενο, με τη μεγαλύτερη δυνατή ακρίβεια.

8.4 Ιχνηλάτηση αντικειμένου (object tracking): Ο σκοπός της ενέργειας αυτής είναι η παρακολούθηση ενός αντικειμένου κινούμενου με τη πάροδο του χρόνου, χρησιμοποιώντας διαδοχικά καρέ από βίντεο ως είσοδο. Είναι μια διαδικασία απαραίτητη για τον σχεδιασμό διαδρομών των αυτόνομων οχημάτων.



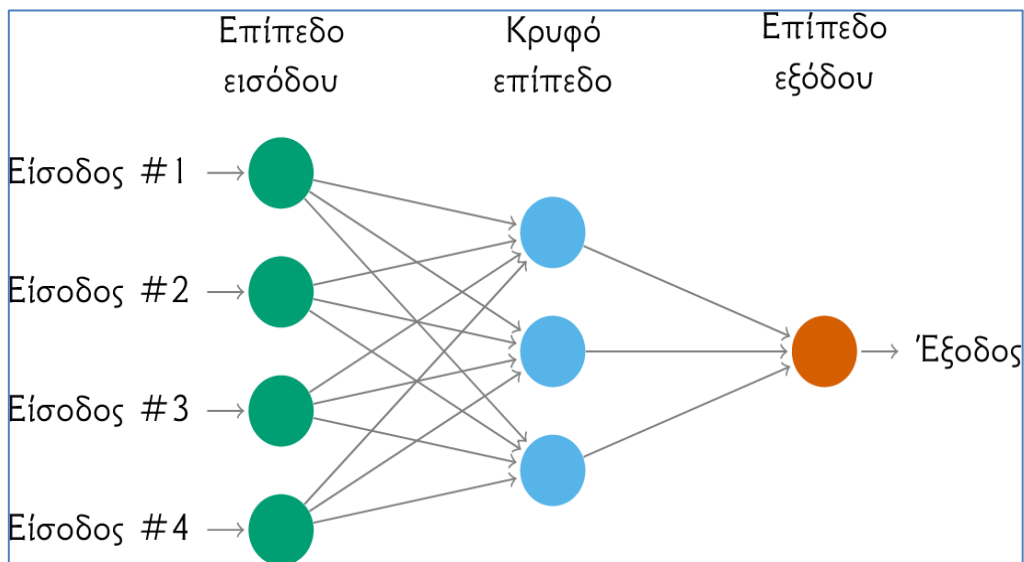
**Σχήμα 1.**  
**Φάσεις ανάλυσης και επεξεργασίας πληροφοριών μέσω τεχνητής όρασης**

## **2.1 Νευρωνικά δίκτυα – Neural networks**

Σημαντικότερη εξέλιξη στον τομέα της μηχανικής μάθησης αποτελεί η χρήση των νευρωνικών δικτύων, τα οποία προσπαθούν να προσομοιάσουν τον τρόπο λειτουργίας τους ανθρώπινου εγκεφάλου, κάτι που κρίθηκε απαραίτητο προκειμένου να αναπαραχθεί νοημοσύνη με τεχνητό τρόπο. Ο ανθρώπινος εγκέφαλος αποτελείται από δισεκατομμύρια νευρικά κύτταρα, που ονομάζονται νευρώνες. Κάθε νευρώνας, έχει την δυνατότητα μετάδοσης ενός ηλεκτροχημικού σήματος και είναι ενωμένος με χιλιάδες άλλους νευρώνες, μέσω ειδικών συνάψεων. Οι συνάψεις μεταξύ δύο νευρώνων σχηματίζονται μεταξύ της εξόδου του ενός (άξονας) και της εισόδου του άλλου (δενδρίτης). Ένας νευρώνας ενεργοποιείται μόνο όταν ο δενδρίτης του δεχθεί κάποιο σήμα το οποίο υπερβαίνει μία συγκεκριμένη τιμή.

Κατά όμοιο τρόπο, τα τεχνητά νευρωνικά δίκτυα αποτελούνται από υπολογιστικούς κόμβους που είναι διασυνδεδεμένοι μεταξύ τους και ονομάζονται νευρώνες. Αυτοί, με βάση τη λειτουργία τους, μπορούν να ταξινομηθούν σε τρεις βασικές κατηγορίες:

- 1) Νευρώνες εισόδου: Λαμβάνουν τις τιμές εισόδου από το περιβάλλον και τις διοχετεύουν στους υπολογιστικούς νευρώνες. Δεν εκτελούν κανέναν υπολογισμό.
- 2) Υπολογιστικοί ή κρυμμένοι νευρώνες: Πολλαπλασιάζουν κάθε είσοδο με το αντίστοιχο συναπτικό βάρος, υπολογίζοντας έτσι το ολικό άθροισμα των γινομένων. Κάθε υπολογιστικός νευρώνας υλοποιεί μία εσωτερική συνάρτηση ενεργοποίησης, η οποία τροφοδοτείται με το προηγούμενο άθροισμα. Το αποτέλεσμα αυτής της συνάρτησης αποτελεί την έξοδο του νευρώνα για την συγκεκριμένη τιμή εισόδου και βάρος.
- 3) Νευρώνες εξόδου: Αποδίδουν στο εξωτερικό περιβάλλον τα αποτελέσματα των υπολογιστικών νευρώνων.



**Σχήμα 2. Παράδειγμα τεχνητού νευρωνικού δικτύου**

(Big Blue Data Academy, (03 March 2023), Νευρωνικά Δίκτυα (Neural Networks): Ορισμός & Εφαρμογές, <https://bigblue.academy/gr/neuronika-diktua>)

Αν η  $i$ -οστή είσοδος του νευρώνα  $k$  είναι η  $x_{ki}$ , το  $i$ -οστό συναπτικό βάρος του νευρώνα  $k$  είναι το  $w_{ki}$ , και  $\varphi(\cdot)$  η συνάρτηση ενεργοποίησης του νευρωνικού δικτύου, τότε η έξοδος  $y_k$  είναι η:

$$y_k = \varphi \left( \sum_{i=0}^N x_{ki} w_{ki} \right)$$

Στον  $k$ -οστό νευρώνα υπάρχει το βάρος  $w_k = 1$ , το οποίο ονομάζεται πόλωση ή κατώφλι. Σε περίπτωση που το συνολικό άθροισμα των εισόδων ενός νευρώνα ξεπερνά την τιμή αυτή, τότε αυτός ενεργοποιείται, ενώ σε οποιαδήποτε άλλη περίπτωση παραμένει ανενεργός.

Η σωστή απεικόνιση των διανυσμάτων εισόδου σε διανύσματα εξόδου, συνεπάγεται και την ορθή λειτουργία του νευρωνικού δικτύου, όπως για παράδειγμα η σωστή αναγνώριση εικόνων. Ένα από τα βασικότερα χαρακτηριστικά όμως των νευρωνικών δικτύων είναι η ικανότητα εκπαίδευσης και μάθησης, δηλαδή η σταδιακή βελτίωση της ικανότητας επίλυσης συγκεκριμένων προβλημάτων. Η μάθηση προϋποθέτει την επαναληπτική διαδικασία προσαρμογής των παραμέτρων, όπως τα βάρη και η πόλωση, ώστε να επιλύεται με μεγαλύτερη επιτυχία κάποιο συγκεκριμένο πρόβλημα. Τέλος, βασικό χαρακτηριστικό ενός επιτυχούς νευρωνικού δικτύου αποτελεί η ικανότητα γενίκευσης, δηλαδή η επιτυχής επίλυση του προβλήματος, ακόμη και με δεδομένα εισόδου τα οποία δεν έχουν επεξεργαστεί από το δίκτυο σε προηγούμενες εκτελέσεις του. Η ικανότητα γενίκευσης μέσω εκμάθησης, αποτελεί μία από τις δομικές διαφορές των νευρωνικών δικτύων, σε σχέση με τον τρόπο λειτουργίας και επίλυσης προβλημάτων από τους υπολογιστές, η φιλοσοφία των οποίων έχει οριστεί από τον von Neumann. Οι διαφορές αυτές συνοψίζονται στον παρακάτω πίνακα:

	<b>Νευρωνικά δίκτυα</b>	<b>Υπολογιστής von Neumann</b>
1.	Παράλληλη επεξεργασία	Σειριακή επεξεργασία
2.	Εκπαίδευση με παραδείγματα, αλλάζοντας το βάρος των συνδέσεων τους.	Προγραμματισμός με εντολές λογικού χαρακτήρα (if – then).
3.	Ανοχή στα σφάλματα.	Καμία ανοχή στα σφάλματα.
4.	Η μνήμη, τα δίκτυα και οι μονάδες λειτουργίας συνυπάρχουν	Η μνήμη και η επεξεργασία πληροφορίας χωρίζονται.
5.	Αυτοοργάνωση κατά την διαδικασία της εκπαίδευσης.	Απόλυτη εξάρτηση από το λογισμικό.
6.	Η πληροφορία αποθηκεύεται στα βάρη των συνδέσεων.	Η πληροφορία αποθηκεύεται σε μονάδες μνήμης.
7.	Ο χρόνος ενός κύκλου είναι της τάξης των msec.	Ο χρόνος ενός κύκλου είναι της τάξης των nsec.

### **Πίνακας 1.**

#### **Διαφορές νευρωνικών δικτύων – υπολογιστών von Neumann**



Η εκπαίδευση των νευρωνικών δικτύων επιτυγχάνεται με την ανταλλαγή βαρών και τιμών εισόδου. Με τον τρόπο αυτό καθίσταται εφικτή η σταδιακή λήψη της απαραίτητης πληροφορίας που αργότερα θα πρέπει να ανακτηθεί κατά την λήψη μιας απόφασης. Οι μέθοδοι μάθησης που χρησιμοποιούνται, μπορούν να διαχωριστούν σε δύο μεγάλες κατηγορίες:

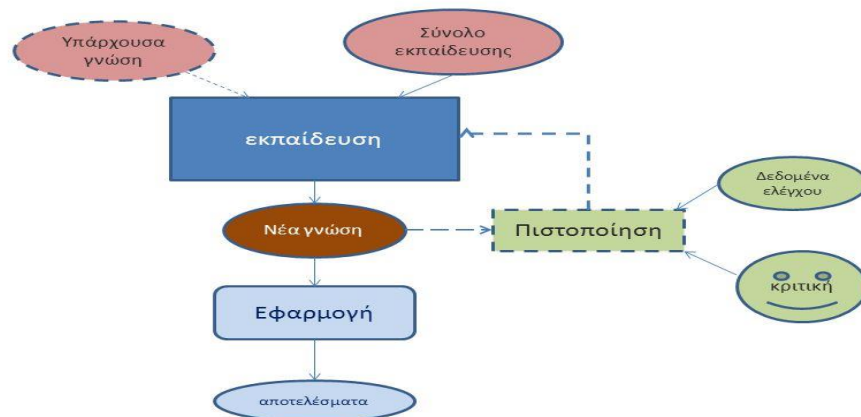
1. Μάθηση με επίβλεψη (supervised learning): Ο αλγόριθμος κατασκευάζει μία συνάρτηση που απεικονίζει δεδομένες εισόδους σε γνωστές επιθυμητές εξόδους, με στόχο να γενικευθεί αυτή η συνάρτηση και για εισόδους με άγνωστη έξοδο. Μπορεί να χρησιμοποιηθεί σε προβλήματα ταξινόμησης (classification), πρόγνωσης (prediction) και διερμηνείας (interpretation).

2. Μάθηση χωρίς επίβλεψη (unsupervised learning): Ο αλγόριθμος κατασκευάζει ένα μοντέλο για κάποιο σύνολο εισόδων σε μορφή παρατηρήσεων, χωρίς να γνωρίζει τις επιθυμητές εξόδους. Μπορεί να χρησιμοποιηθεί σε προβλήματα ανάλυσης συσχετισμών (association analysis) και ομαδοποίησης (clustering).

3. Μάθηση με ενίσχυση (graded learning): Ο αλγόριθμος μαθαίνει μια στρατηγική ενεργειών μέσα από άμεση αλληλεπίδραση με το περιβάλλον. Χρησιμοποιείται κυρίως σε προβλήματα σχεδιασμού (planning), όπως για παράδειγμα ο έλεγχος κίνησης αυτόνομων οχημάτων.

Στο σχήμα 3 που ακολουθεί παρουσιάζεται μία τυπική διαδικασία μηχανικής μάθησης:

### Φάσεις Μηχανικής Μάθησης



**Σχήμα 3. Φάσεις μηχανικής μάθησης**

(Γεωργούλη Κατερίνα, Τεχνητή Νοημοσύνη, Ανοικτά Ακαδημαϊκά Μαθήματα ΤΕΙ Αθήνας,

<https://slideplayer.gr/slide/11418311/>)

Σημαντικός τομέας των νευρωνικών δικτύων είναι αυτός των αυτοκωδικοποιητών (autoencoders). Στόχος αυτών, είναι η αναπαράσταση των τιμών εισόδου κατά την έξοδο με τον κατάλληλο τρόπο, έτσι ώστε να κωδικοποιούνται οι κατά περίπτωση χρήσιμες ιδιότητες των δεδομένων. Κάθε αυτοκωδικοποιητής αποτελείται από ένα ζεύγος νευρωνικών δικτύων, τα οποία είναι τα εξής:

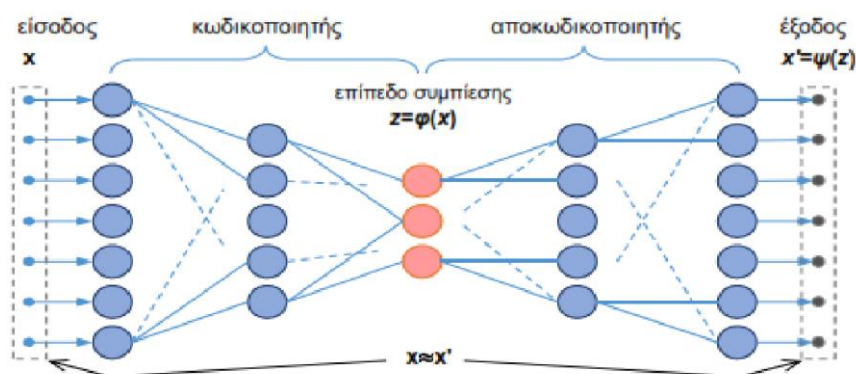
- 1) Κωδικοποιητής: Λαμβάνει τα δεδομένα εισόδου και προχωρά σε εμφάθυνση της κωδικοποίησης (code embedding), κατά έναν παράγοντα κωδικοποίησης:

$$z = \varphi(x).$$

- 2) Αποκωδικοποιητής: Λαμβάνει το κωδικοποιημένο διάνυσμα  $z$  και προχωρά σε μετασχηματισμό του, προβάλλοντας το στον χώρο διαστάσεων της εισόδου, με τελικό στόχο οι τιμές  $x$  και  $x'$ . Η διαδικασία αυτή αποτελεί την αντιστροφή της διαδικασίας κωδικοποίησης και μπορεί να αποδοθεί με την εξίσωση:

$$x' = \psi(z).$$

Βασικό χαρακτηριστικό των αυτοκωδικοποιητών αποτελεί ο παράγοντας συμπίεσης, μεταξύ κωδικοποιητή και αποκωδικοποιητή. Οι διαστάσεις του θα πρέπει να είναι μικρότερες από αυτές της εισόδου, ενώ η έξοδος θα πρέπει να έχει τις ίδιες διαστάσεις με την είσοδο. Η διαδικασία της εκπαίδευσης αποτελεί τον καθοριστικό παράγοντα προκειμένου η τιμές εξόδου να είναι όσο το δυνατόν πλησιέστερες με τις τιμές εισόδου. Η λειτουργία ενός αυτοκωδικοποιητή παρουσιάζεται στο Σχήμα 4:



**Σχήμα 4. Διάγραμμα αυτοκωδικοποιητή**

Steven Flores, (05 April 2019) , Variational Autoencoders are Beautiful,

<https://www.compthree.com/blog/autoencoder/>

## **2.2 Βαθιά μάθηση (Deep learning)**

Η βαθιά μάθηση είναι ένα υποσύνολο της μηχανικής μάθησης και αποτελεί ένα νευρωνικό δίκτυο με τρία ή περισσότερα επίπεδα. Σκοπός αυτών είναι να προσομοιώσουν την συμπεριφορά του ανθρώπινου εγκεφάλου, επιτρέποντας του να «μαθαίνει» μέσα από μεγάλες ποσότητες δεδομένων. Ένα νευρωνικό δίκτυο με ένα στρώμα μπορεί να εκτελέσει κατά προσέγγιση προβλέψεις, ενώ με επιπλέον κρυμμένα επίπεδα επιτυγχάνεται βελτιστοποίηση και ακρίβεια.

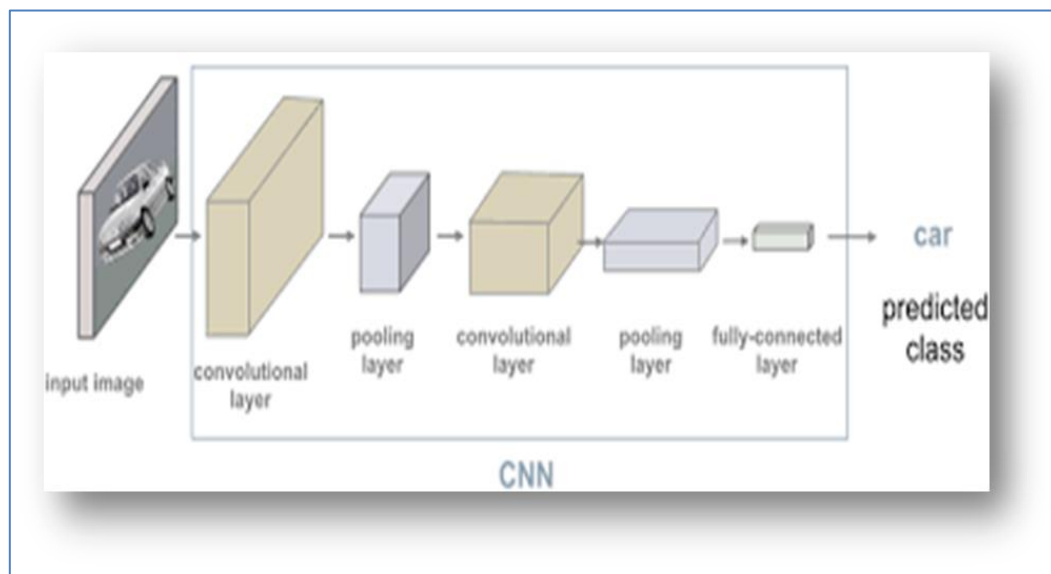
Η βαθιά μάθηση αξιοποιείται σε εφαρμογές και υπηρεσίες τεχνητής νοημοσύνης (AI) που βελτιώνουν την αυτοματοποίηση, εκτελώντας αναλυτικές και φυσικές εργασίες χωρίς την ανθρώπινη παρέμβαση. Η τεχνολογία βαθιάς εκμάθησης είναι πίσω από τα καθημερινά προϊόντα και υπηρεσίες (όπως ψηφιακοί βοηθοί, τηλεχειριστήρια με φωνητική λειτουργία), καθώς και από τις αναδυόμενες τεχνολογίες (όπως αυτοκινούμενα οχήματα).

## **2.3 Συνελικτικά νευρωνικά δίκτυα (Convolutional neural networks - CNN)**

Στη βαθιά μάθηση, ένα συνελικτικό νευρωνικό δίκτυο (CNN) είναι μια κλάση βαθιά νευρωνικών δικτύων το οποίο εφαρμόζεται κυρίως στην ανάλυση οπτικών εικόνων. Πρόκειται για πλήρως συνδεδεμένα δίκτυα ή αλλιώς κανονικοποιημένες εκδόσεις πολλαπλών στρώσεων, δηλαδή κάθε νευρώνας σε ένα στρώμα συνδέεται με όλους τους νευρώνες στο επόμενο στρώμα. Η συνέλιξη αυτή είναι ένα είδος γραμμικής λειτουργίας. Είναι δηλαδή απλά νευρωνικά δίκτυα, τα οποία αξιοποιούν την συνέλιξη, προκειμένου να επιτύχουν τον πολλαπλασιασμό τους σε ένα επίπεδο. Αποτελούν κανονικοποιημένες εκδόσεις πολλαπλών στρώσεων που είναι πλήρως συνδεδεμένα μεταξύ τους. Αυτή η πλήρης σύνδεση, όμως, τα καθιστά επιρρεπή στην υπερφόρτωση δεδομένων, γεγονός που αντιμετωπίζεται με τη μέτρηση βαρών προς τη συνάρτηση απώλειας. Ωστόσο, το CNN υιοθετεί μια διαφορετική προσέγγιση όσον αφορά την κανονικοποίηση, μιας και εκμεταλλεύεται το ιεραρχικό πρότυπο των δεδομένων και συγκεντρώνει τα πιο περίπλοκα μοτίβα χρησιμοποιώντας μικρότερα και απλούστερα πρότυπα.

Τα δίκτυα αυτά είναι γνωστά και ως αμετάβλητα ως προς την μετατόπιση ή τον χώρο τεχνητά νευρωνικά δίκτυα (shift invariant or space invariant artificial neural

networks - SIANN) και βασισμένα στην αρχιτεκτονική των κοινών βαρών τους και στα χαρακτηριστικά αμετάβλητης μετάφρασης. Τα CNN σε σύγκριση με άλλους αλγόριθμους ταξινόμησης εικόνων χρησιμοποιούν σχετικά μικρή προεπεξεργασία. Αυτό σημαίνει ότι το δίκτυο μαθαίνει τα φίλτρα. Η ανεξαρτησία από προηγούμενες γνώσεις και ανθρώπινη προσπάθεια στον σχεδιασμό χαρακτηριστικών αποτελεί σημαντικό πλεονέκτημα. Ένα CNN λειτουργεί σε τρία στάδια. Το πρώτο στάδιο είναι μια συνέλιξη (convolution), όπου μερικά pixels της εικόνας σαρώνονται κάθε φορά και δημιουργείται ένας χάρτης χαρακτηριστικών με τις πιθανότητες του κάθε χαρακτηριστικού να ανήκει στην αντίστοιχη κατηγορία. Το δεύτερο στάδιο είναι η ομαδοποίηση (pooling) η οποία συγκεντρώνει τα πιο σημαντικά χαρακτηριστικά της εικόνας. Τέλος, όταν τα χαρακτηριστικά βρίσκονται στον σωστό βαθμό ανάλυσης, το νευρωνικό δίκτυο εισέρχεται στο τρίτο στάδιο, και καθίσταται πλήρως νευρωνικό δίκτυο, το οποίο αναλύει τις πιθανότητες και αποφασίζει σε ποια κατηγορία ανήκει η εικόνα. Το βήμα αυτό μπορεί επίσης να χρησιμοποιηθεί και για άλλες εργασίες, όπως η δημιουργία κειμένου ή η αυτόματη δημιουργία λεζάντας για εικόνες.



**Σχήμα 5. Λειτουργία συνελκτικών νευρωνικών δικτύων**

(Ράπτης Σάββας, (10 Οκτωβρίου 2021) , Επιβλεπόμενη μηχανική μάθηση και τεχνητά νευρωνικά δίκτυα – Μέρος Β: “Neural Networks”, <https://2science.gr/machine-learning-2/>)

## 2.4 Tensorflow

Η μηχανική μάθηση περιλαμβάνει αρκετά περίπλοκες και χρονοβόρες διαδικασίες όπως η απόκτηση των δεδομένων εισόδου, η εκπαίδευση του εξειδικευμένου μοντέλου αναγνώρισης, η δημιουργία μηχανισμού σωστών προβλέψεων και η σωστή επεξεργασία των αποτελεσμάτων. Για αυτόν τον σκοπό, τα τελευταία χρόνια έχουν δημιουργηθεί εξειδικευμένα λογισμικά, με στόχο την απλοποίηση και συντόμευση της απαιτητικής διαδικασίας της μηχανικής μάθησης.

Ένα από αυτά είναι και το Tensorflow, που είναι λογισμικό ανοικτού κώδικα και δημιουργήθηκε από την Google το 2015. Βασική δομή του Tensorflow, αποτελεί το διάγραμμα ροής (dataflow graph). Το διάγραμμα ροής καθορίζει τον τρόπο με τον οποίο τα δεδομένα διαρρέουν ένα γράφημα, το οποίο σχηματίζεται από μία σειρά κόμβων. Η χρήση γραφήματος παρέχει ένα πλήθος πλεονεκτημάτων, τα οποία είναι τα εξής (Μουμουλίδης και Τζούμα, 2018):

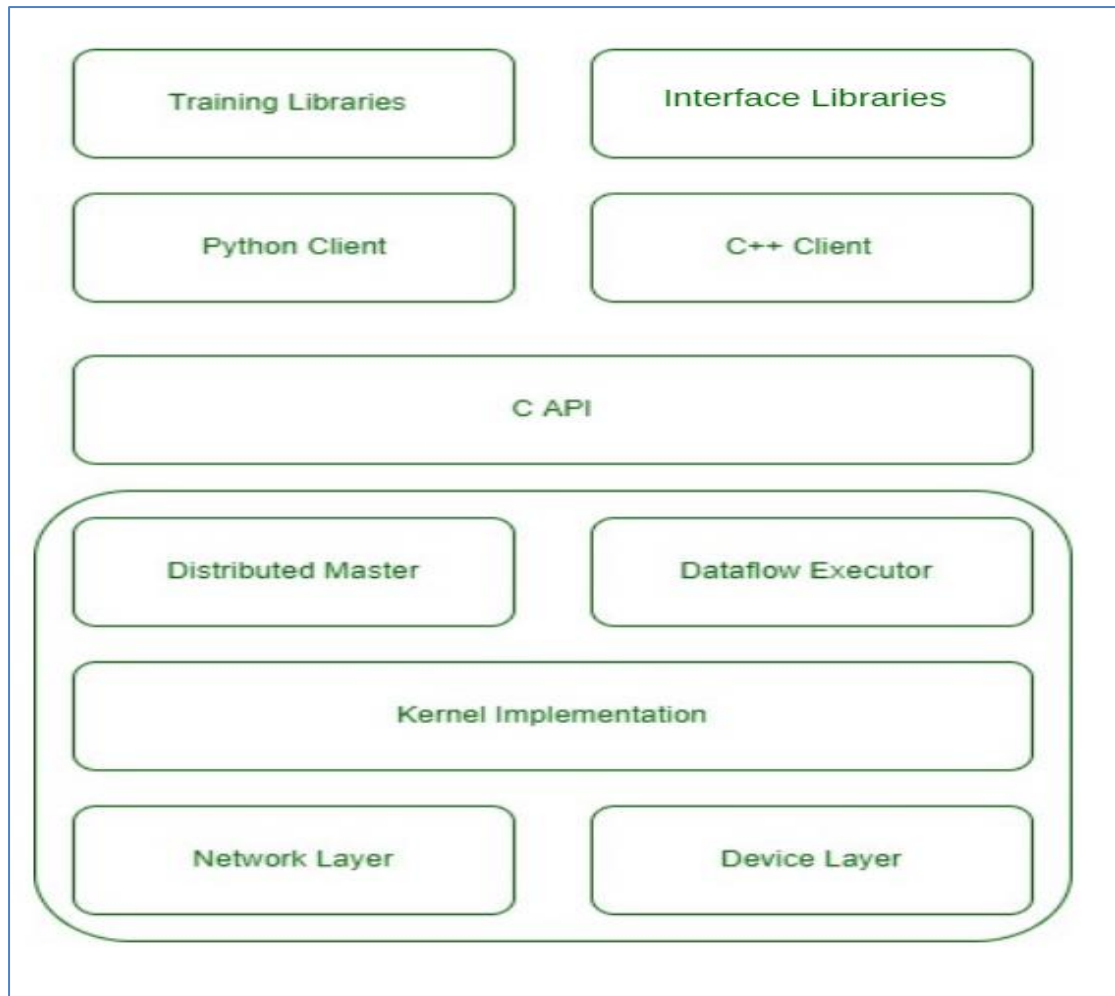
- 1) Καθιστά εφικτή την εκτέλεση σε κατανομημένα συστήματα. Σε συνδυασμό με το API υψηλού επιπέδου, το περιβάλλον ανάπτυξης καθίσταται ευέλικτο και εύκολα προσαρμόσιμο, με ισχυρές δυνατότητες.
- 2) Ο γράφος μπορεί εύκολα να παραμετροποιηθεί και να μεταβληθεί, με αποτέλεσμα να μπορεί να χρησιμοποιηθεί σε διαφορετικές πλατφόρμες και συσκευές. Επιπλέον μπορεί να αποθηκευτεί και να χρησιμοποιηθεί αργότερα είτε στην ίδια, είτε σε διαφορετική συσκευή. Σε αυτό συμβάλλει και η εύκολη παραμετροποίηση της απαιτούμενης μνήμης και υπολογιστικής ισχύος που απαιτείται για την εκτέλεση του Tensorflow.

Κάθε κόμβος του γραφήματος αποτελεί έναν μαθηματικό μετασχηματισμό, ενώ κάθε ακμή μεταξύ κόμβων του γραφήματος αποτελεί έναν πολυδιάστατο πίνακα δεδομένων, που ονομάζεται τανυστής (tensor). Τα βασικά χαρακτηριστικά που προσδιορίζουν έναν τανυστή είναι τα εξής:

- 1) Τύπος (type): Εκφράζει τον τύπο δεδομένων του τανυστή (int, float, κλπ.).
- 2) Σχήμα (shape): Εκφράζει το μέγεθος κάθε διάστασης του τανυστή (σε έναν πίνακα τριών διαστάσεων περιγράφει τον αριθμό των γραμμών, των στηλών και της τρίτης διάστασης).
- 3) Βαθμός (rank): Εκφράζει την διάσταση του τανυστή. Για παράδειγμα ένας τρισδιάστατος πίνακας είναι ένας τανυστής με βαθμό 3.

Η αρχιτεκτονική λειτουργίας του TensorFlow μπορεί να αναλυθεί σε έξι βασικά επίπεδα:

- 1) Επίπεδο συσκευής (device layer) και επίπεδο δικτύου (network layer): Το επίπεδο συσκευής είναι υπεύθυνο για την σωστή επικοινωνία του Tensorflow με το λειτουργικό σύστημα και το υλικό της συσκευής όπως CPU, GPU, TPU. Το επίπεδο δικτύου ελέγχει την δικτυακή επικοινωνία με άλλες συσκευές και πρωτόκολλα, σε περίπτωση που απαιτείται απομακρυσμένη εκτέλεση σε αυτές.
- 2) Υλοποίηση πυρήνα (kernel implementation): Ελέγχει την υλοποίηση και λειτουργία εφαρμογών μηχανικής μάθησης.
- 3) Ελεγκτής κατανομής (distributed master) και ελεγκτής ροής δεδομένων (dataflow executor): Ο ελεγκτής κατανομής κατανέμει την εργασία στην συσκευή εκτέλεσης, ενώ ο ελεγκτής ροής δεδομένων είναι υπεύθυνος για την ροή των δεδομένων στον γράφο με τον βέλτιστο τρόπο.
- 4) Διεπαφή προγραμματισμού εφαρμογών C (C API): Περιλαμβάνει όλο το απαραίτητο λογισμικό των λειτουργιών του TensorFlow. Έχει υλοποιηθεί σε γλώσσα C, λόγω της ταχύτητας και της αξιοπιστίας της.
- 5) Υλοποίηση εφαρμογής επιπέδου χρήστη (python client – C++ client): Ελέγχει την διασύνδεση του Tensorflow με την εφαρμογή σε επίπεδο χρήστη. Έχει υλοποιηθεί σε διάφορες γλώσσες προγραμματισμού, με κύριες τις python και C++.
- 6) Βιβλιοθήκες εκπαίδευσης - μάθησης (training – interface libraries): Περιλαμβάνει όλες τις βιβλιοθήκες που χρησιμοποιούνται για την εκπαίδευση και βελτιστοποίηση των απαραίτητων μοντέλων, κατά την διαδικασία της μηχανικής μάθησης.



**Εικόνα 2. Αρχιτεκτονική Tensorflow**

Τέλος η λειτουργία του Tensorflow μπορεί να χωριστεί σε τρία διακριτά μέρη, τα οποία είναι τα εξής:

- 1) Αρχική επεξεργασία δεδομένων.
- 2) Σχηματισμός του μοντέλου αναγνώρισης.
- 3) Εκπαίδευση, βελτίωση και αξιολόγηση χρησιμοποιούμενου μοντέλου.

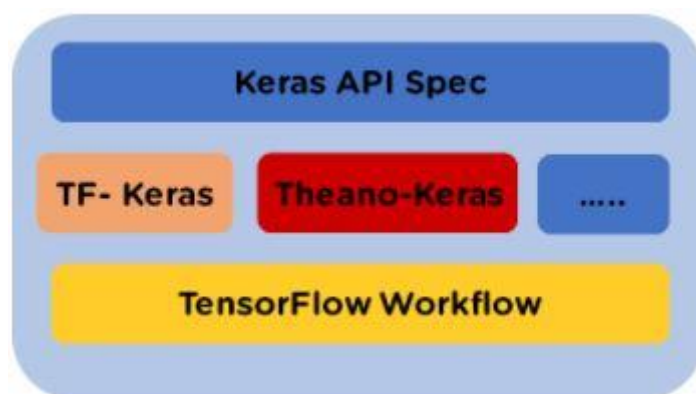
## **2.5 Keras**

Το Tensorflow, ως επίσημο API υψηλού επιπέδου υποστηρίζει το Keras, το οποίο είναι ένα API βαθιάς εκμάθησης (deep learning) υψηλού επιπέδου, υλοποιημένο σε γλώσσα python. Ως API υψηλού επιπέδου, το Keras μπορεί να χρησιμοποιηθεί σε συνδυασμό με τα παρακάτω API χαμηλού επιπέδου:

- 1) Tensorflow

- 2) Theano
- 3) PlaidML
- 4) MXNet
- 5) CNTK (Microsoft Cognitive Toolkit)

Όταν το Keras χρησιμοποιείται σε συνδυασμό με το Tensorflow, χρησιμοποιείται για την επιτάχυνση της διαδικασίας εκμάθησης, ενώ το τελευταίο για όλους τους υπολογισμούς που έχουν να κάνουν με τους τανυστές και τους γραφήματα ροής δεδομένων.



**Εικόνα 3. Δομή Keras**

((07 November 2023) , What Is Keras: The Best Introductory Guide To Keras,  
<https://www.simplilearn.com/tutorials/deep-learning-tutorial/what-is-keras>)

Η χρήση του Keras παρουσιάζει τα εξής πλεονεκτήματα:

- 1) Η εκμάθηση του είναι ιδιαίτερα εύκολη, ακόμη και από χρήστες με μηδενική εμπειρία στην επεξεργασία εικόνας και στην μηχανική όραση. Το Keras API χαρακτηρίζεται από απλότητα, ευχρηστία, ελαχιστοποίηση των ενεργειών του χρήστη και ευκρινή καταγραφή των σφαλμάτων εκτέλεσης.
- 2) Η κοινότητα ανάπτυξης και υποστήριξης του Keras είναι πολύ μεγάλη, κάτι που καθιστά την απόκτηση πληροφοριών και τεχνικών εξαιρετικά εύκολη.
- 3) Προσφέρει μικρότερο χρόνο εκτέλεσης σε σχέση με τα υπόλοιπα API υψηλού επιπέδου, ιδιαίτερα όταν χρησιμοποιείται σε συνδυασμό με το Tensorflow.
- 4) Μπορεί να εκτελεστεί τόσο στην κεντρική μονάδα επεξεργασίας (CPU), όσο και στον επεξεργαστή των γραφικών (GPU).



- 5) Υποστηρίζει σχεδόν όλα τα μοντέλα νευρωνικών δικτύων (Tensorflow, Theano, PlaidML, MXNet, CNTK).

Όλα αυτά τα πλεονεκτήματα καθιστούν το Keras ιδανικό για τις εξής περιπτώσεις:

- 1) Δημιουργία πολύπλοκων νευρωνικών δικτύων.
- 2) Εφαρμογές τεχνητής νοημοσύνης που απαιτούν επεξεργασία μεγάλου όγκου δεδομένων.
- 3) Δημιουργία και επεξεργασία μοντέλων που απαιτούν μικρή ταχύτητα εκτέλεσης.

Παρόλα αυτά οι υποψήφιοι χρήστες του Keras, θα πρέπει να έχουν υπόψιν τα εξής:

- 1) Η ταχύτητα εκτέλεσης είναι αυξημένη, λόγω της βαρύτητας που έχει δοθεί στην φιλικότητα προς τον χρήστη.
- 2) Δεν υποστηρίζεται η δημιουργία δυναμικών γράφων.
- 3) Δεν είναι κατάλληλο για αλγορίθμους και μοντέλα μάθησης όπως το PCM (principal component analysis).

## **2.6 OpenCV 2**

Το OpenCV 2 (open source computer vision) αποτελεί μία βιβλιοθήκη μηχανικής όρασης αληθινού χρόνου (real time), ανοικτού κώδικα. Βασικό χαρακτηριστικό του είναι η διαλειτουργικότητά του, διότι μπορεί να χρησιμοποιηθεί από διαφορετικά λειτουργικά συστήματα, κάτι που το καθιστά μία από τις δημοφιλέστερες βιβλιοθήκες στον τομέα της μηχανικής όρασης. Τα λειτουργικά συστήματα με τα οποία είναι συμβατό είναι τα ακόλουθα: Windows, Linux, macOS, FreeBSD, Android, iOS, Maemo - BlackBerry 10 και QNX. Η ανάπτυξη του OpenCV 2 έχει γίνει σε C++, αλλά έχουν δημιουργηθεί API υποστήριξης και άλλων γλωσσών προγραμματισμού, όπως python, java και javascript.

Οι βασικότεροι στόχοι που έχουν επιτευχθεί από την χρήση του OpenCV 2 είναι οι ακόλουθοι:

- 1) Ανάπτυξη του τομέα της μηχανικής όρασης, χάρη στον ανοικτό αλλά και πλήρως βελτιστοποιημένο κώδικα που παρέχει το OpenCV 2.
- 2) Ραγδαία ανάπτυξη των εμπορικών εφαρμογών μηχανικής όρασης, λόγω του ότι ο κώδικας είναι αποδοτικός και σε κινητές συσκευές.

- 3) Διάδοση του τομέα της μηχανικής όρασης, παρέχοντας ένα ομογενοποιημένο και ενιαίο περιβάλλον ανάπτυξης στους προγραμματιστές.

Ακολουθούν τα στοιχεία που αποτελούν την αρχιτεκτονική δομή του OpenCV 2:

- 1) Λειτουργικότητα πυρήνα (core functionality).
- 2) Επεξεργασία εικόνας (image processing).
- 3) Ανάλυση βίντεο (video analysis).
- 4) Στάθμιση κάμερας και τρισδιάστατη απεικόνιση (camera calibration and 3D reconstruction).
- 5) Δισδιάστατη απεικόνιση (2D features framework).
- 6) Εντοπισμός αντικειμένου (object detection).
- 7) Γραφικό περιβάλλον (high-level GUI)
- 8) Είσοδος – έξοδος βίντεο (video I/O).

### **3. Ανάλυση συστήματος**

Το σύστημα που αναπτύχθηκε στα πλαίσια της συγκεκριμένης διπλωματικής εργασίας αποτελείται από δύο ανεξάρτητα μέρη τα οποία είναι ο σταθμός εντοπισμού ή σταθμός επιτήρησης και το αυτόνομο όχημα (Σχήμα 6).

Βασικός σκοπός του σταθμού επιτήρησης είναι ο εντοπισμός ενός ή περισσότερων αντικειμένων με χρήση τεχνικών μηχανικής όρασης. Τα αντικείμενα στόχοι μπορούν να διαφέρουν, ανάλογα με το μοντέλο μηχανικής μάθησης και τον τύπο εκπαίδευσης που έχει προηγηθεί σε αυτό. Το μεγάλο πλήθος μοντέλων και μεθόδων εκπαίδευσης καθιστούν τον σταθμό επιτήρησης ικανό να χρησιμοποιηθεί σε ένα μεγάλο εύρος στρατιωτικών και πολιτικών εφαρμογών. Ο σταθμός επιτήρησης υλοποιήθηκε σε δύο διαφορετικές μορφές, με τις ίδιες όμως δυνατότητες αναγνώρισης:

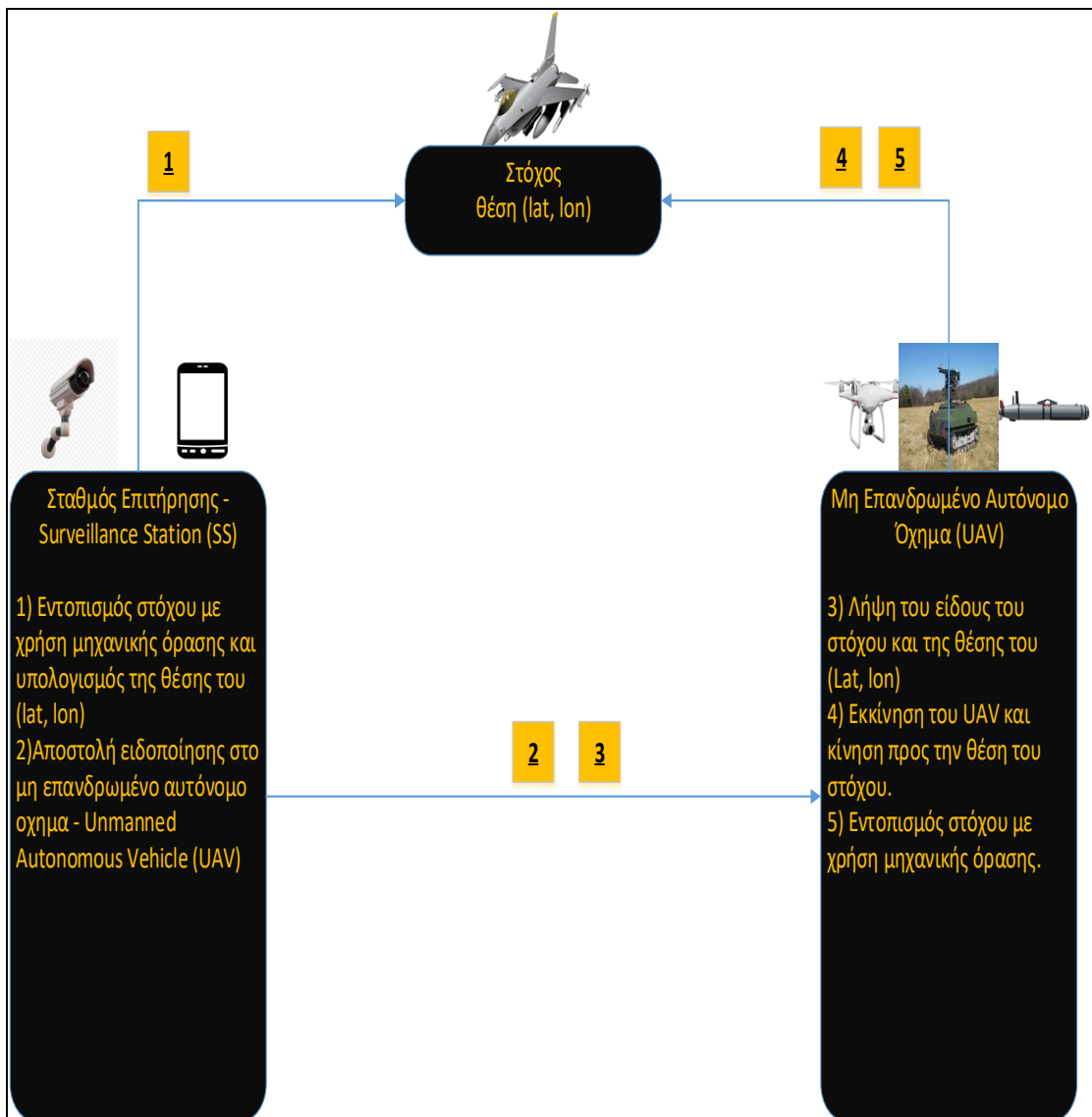
1. Σταθερός σταθμός επιτήρησης. Για την υλοποίησή του χρησιμοποιήθηκε ένας μικροϋπολογιστής Raspberry Pi 3 και μία κάμερα USB. Η κίνηση της κάμερας, χάρη σε ειδική βάση στήριξης, είναι δυνατή τόσο στο οριζόντιο όσο και στο κατακόρυφο επίπεδο.
2. Κινητός σταθμός επιτήρησης. Υλοποιήθηκε σε μία συσκευή κινητού τηλεφώνου και για την μηχανική αναγνώριση χρησιμοποιεί την κάμερα της συσκευής.

Βασικός στόχος της ανάπτυξης των δύο διαφορετικών μορφών σταθμών επιτήρησης, ήταν η κατάδειξη της ευκολίας και του χαμηλού κόστους για την δημιουργία συστημάτων τέτοιου είδους με επαρκείς δυνατότητες.

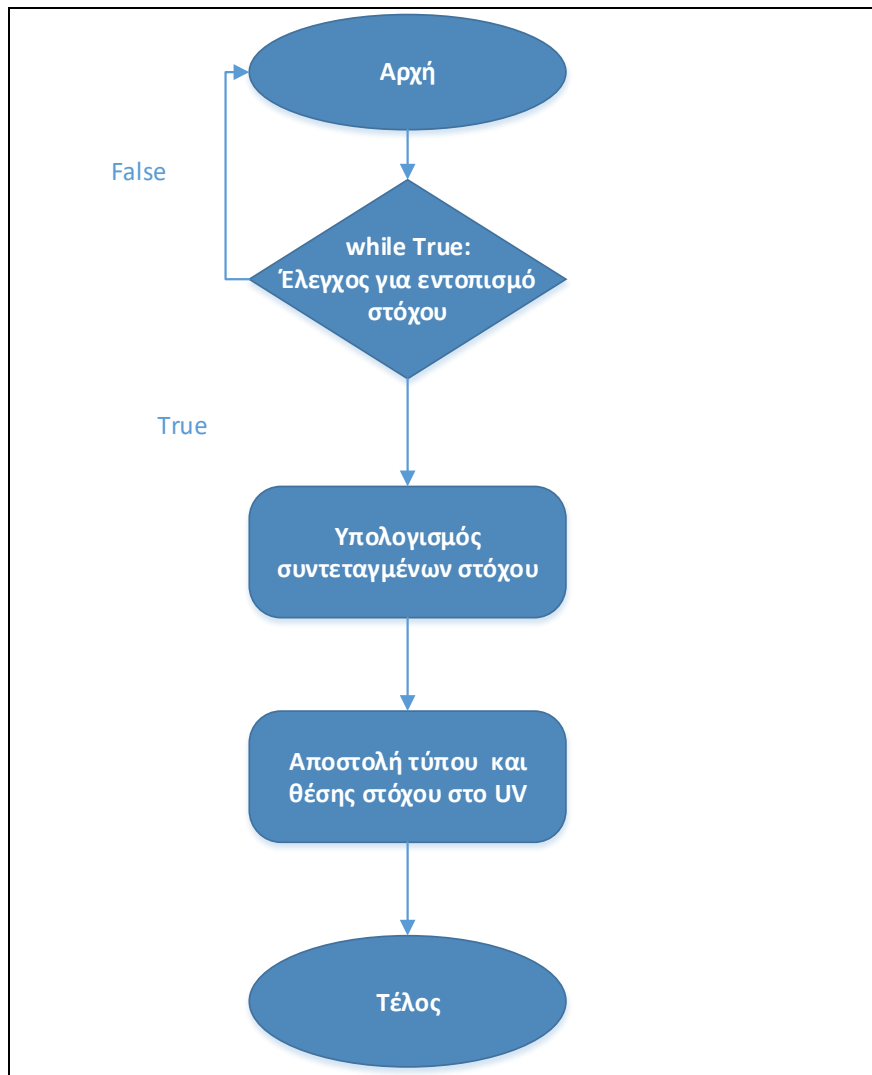
Όταν ο σταθμός εντοπίσει το αντικείμενο στόχο αποστέλλει ένα μήνυμα ειδοποίησης προς το μη επανδρωμένο όχημα (Σχήμα 7), το οποίο μέχρι στιγμής βρίσκεται σε κατάσταση αναμονής (Σχήμα 8). Με τον τρόπο αυτό ενημερώνει για τον τύπο του στόχου αλλά και για τις συντεταγμένες θέσης αυτού.

Σε περίπτωση θετικής απόκρισης του χειριστή του οχήματος, αρχίζει η εκτέλεση της αποστολής (Σχήματα 9 & 10). Βάσει των συντεταγμένων θέσης του μη επανδρωμένου οχήματος και του στόχου, υπολογίζεται η πορεία που θα ακολουθηθεί από το πρώτο προς εντοπισμό του δεύτερου. Θα πρέπει να σημειωθεί πως στα πλαίσια της συγκεκριμένης διπλωματικής δεν έγινε ανάπτυξη ενός πραγματικού μη επανδρωμένου οχήματος, αλλά πραγματοποιήθηκε η

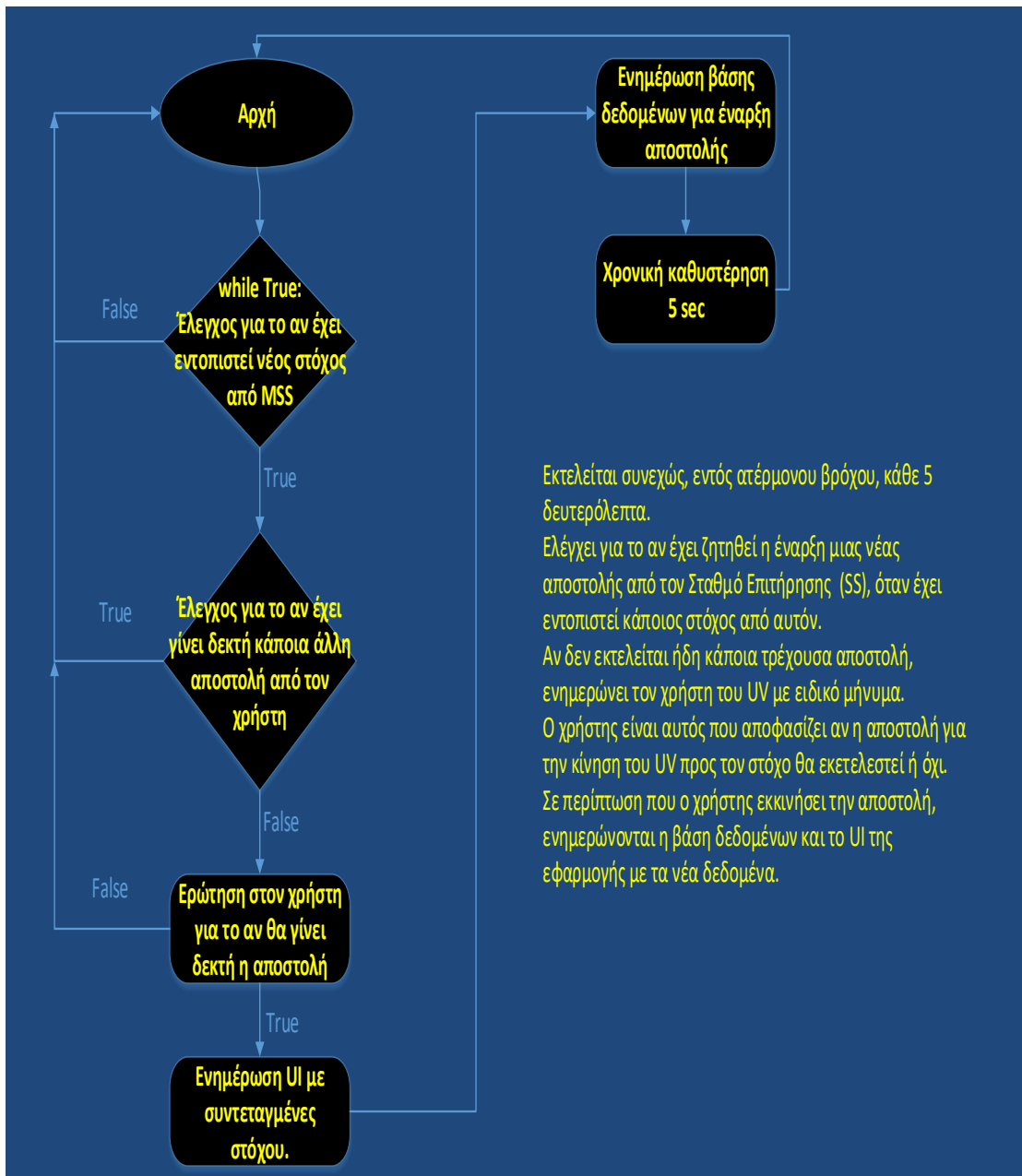
εξομοίωση του (αποστολή συντεταγμένων θέσης) από μία συσκευή κινητού τηλεφώνου με λειτουργικό android.



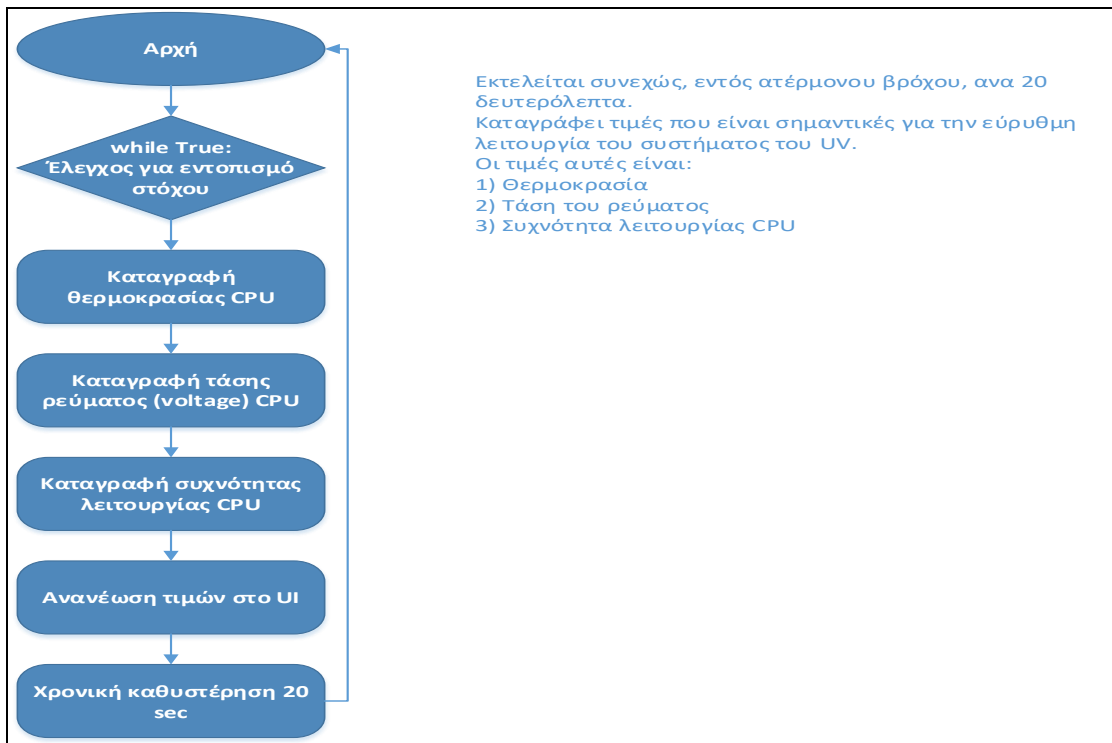
**Σχήμα 6. Ροή γενικής λειτουργίας συστήματος**



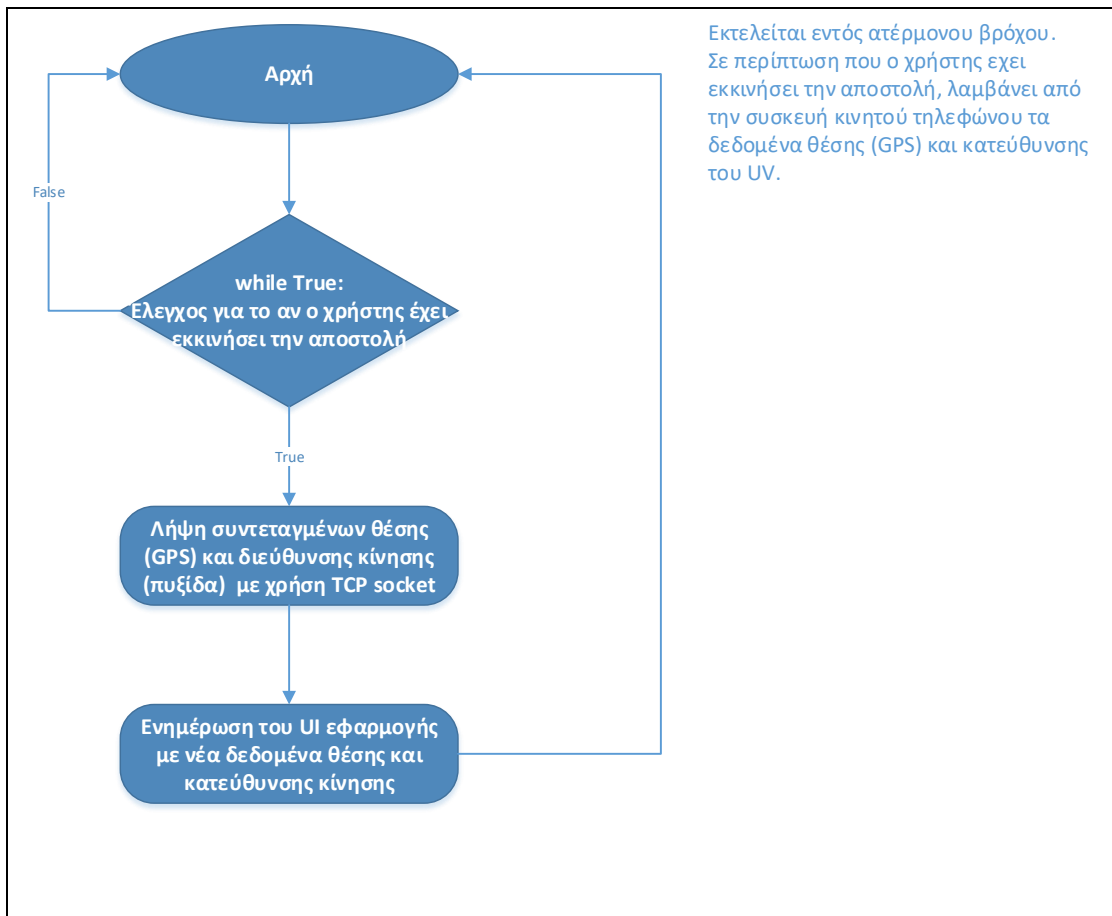
**Σχήμα 7. Ροή λειτουργίας σταθμού επιτήρησης**



**Σχήμα 8.**  
**Ροή λειτουργίας κατάστασης αναμονής μη επανδρωμένου οχήματος**



**Σχήμα 9.**  
**Ροή καταγραφής μεταβλητών λειτουργίας μη επανδρωμένου οχήματος**



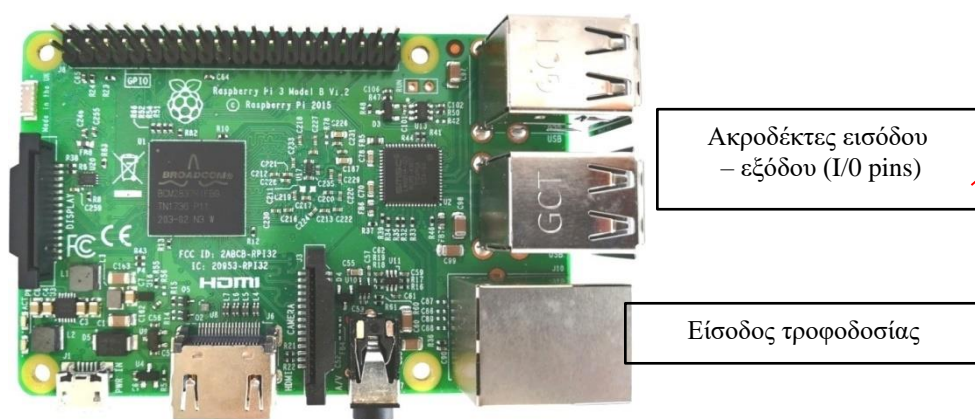
**Σχήμα 10.** Ροή ανανέωσης συντεταγμένων θέσης μη επανδρωμένου οχήματος

## 4.Υλικό

### 4.1 Raspberry Pi 3 Model B

System Chip (SoC)	Broadcom BCM2837
CPU	1.2GHz 64-bit Quad-Core ARM Cortex-A53 (ARMv8)
GPU	Broadcom VideoCore IV 3D
RAM	1GB LPDDR2
Power	Micro USB socket 5V DC 2.5A
Operating System	Raspbian GNU/Linux 10 (buster)
Wireless	2.4GHz 802.11b/g/n Wireless LAN
Ethernet	10/100 Ethernet
Video Output	1 × full size HDMI
Audio Output	3.5mm analogue audio-video jack, HDMI, 4 × USB 2.0
Camera Connector	15-pin MIPI Camera Serial Interface (CSI-2)
Display Connector	Display Serial Interface (DSI) 15 way Flexible Flat Cable (FFC) connector with two data lanes and a clock lane
GPIO Connector	40-pin GPIO header
Memory Card Slot	Micro SD

**Πίνακας 2. Τεχνικά χαρακτηριστικά Raspberry Pi 3 Model B**



**Εικόνα 4. Άνω όψη του Raspberry Pi 3 Model B**

Το Raspberry Pi 3 διαθέτει 40 ακροδέκτες εισόδου – εξόδου (GPIO pins), η ύπαρξη των οποίων καθιστά δυνατή την εύκολη διασύνδεσή του με



αισθητήρες, περιφερειακά και μικροελεγκτές. Πιο συγκεκριμένα, δύο από τους ακροδέκτες παρέχουν τάση 3.3V DC, άλλοι δύο τάση 5V DC και τέλος 8 ακροδέκτες παρέχουν γείωση. Από τους υπόλοιπους ακροδέκτες, κάποιοι μπορούν να χρησιμοποιηθούν για την μετάδοση σημάτων συγκεκριμένων προτύπων και πρωτοκόλλων όπως :

- Μετάδοση σημάτων διαμόρφωσης εύρους παλμού (pulse width modulation - PWM).
- Διασύνδεση συσκευών με χρήση SPI (serial peripheral interface).
- Διασύνδεση συσκευών με χρήση πρωτοκόλλου I<sup>2</sup>C (inter integrated circuit).
- Διασύνδεση συσκευών με χρήση πρωτοκόλλου UART (universal asynchronous receiver-transmitter).

Raspberry Pi 3 GPIO Header				
Pin#	NAME		NAME	Pin#
01	3.3v DC Power	⬇️⬆️	DC Power 5v	02
03	GPIO02 (SDA1 , I <sup>2</sup> C)	⬇️⬆️	DC Power 5v	04
05	GPIO03 (SCL1 , I <sup>2</sup> C)	⬇️⬆️	Ground	06
07	GPIO04 (GPIO_GCLK)	⬇️⬆️	(TXD0) GPIO14	08
09	Ground	⬇️⬆️	(RXD0) GPIO15	10
11	GPIO17 (GPIO_GEN0)	⬇️⬆️	(GPIO_GEN1) GPIO18	12
13	GPIO27 (GPIO_GEN2)	⬇️⬆️	Ground	14
15	GPIO22 (GPIO_GEN3)	⬇️⬆️	(GPIO_GEN4) GPIO23	16
17	3.3v DC Power	⬇️⬆️	(GPIO_GEN5) GPIO24	18
19	GPIO10 (SPI_MOSI)	⬇️⬆️	Ground	20
21	GPIO09 (SPI_MISO)	⬇️⬆️	(GPIO_GEN6) GPIO25	22
23	GPIO11 (SPI_CLK)	⬇️⬆️	(SPI_CE0_N) GPIO08	24
25	Ground	⬇️⬆️	(SPI_CE1_N) GPIO07	26
27	ID_SD (I <sup>2</sup> C ID EEPROM)	⬇️⬆️	(I <sup>2</sup> C ID EEPROM) ID_SC	28
29	GPIO05	⬇️⬆️	Ground	30
31	GPIO06	⬇️⬆️	GPIO12	32
33	GPIO13	⬇️⬆️	Ground	34
35	GPIO19	⬇️⬆️	GPIO16	36
37	GPIO26	⬇️⬆️	GPIO20	38
39	Ground	⬇️⬆️	GPIO21	40

**Εικόνα 5. Ακροδέκτες Raspberry Pi 3 Model B**

Στην συγκεκριμένη διπλωματική εργασία το Raspberry Pi χρησιμοποιήθηκε στις παρακάτω περιπτώσεις:

- 1) Σταθερός σταθμός επιτήρησης: Έλεγχος της κίνησης των κινητήρων servo, για την ανύψωση και την περιστροφή της κάμερας αναγνώρισης.

Επιπλέον σε αυτό εκτελείται το λογισμικό και ο κώδικας (Tensorflow, OpenCV 2) της μηχανικής όρασης.

- 2) Μη επανδρωμένο αυτόνομο όχημα (UV): Έλεγχος των λειτουργιών του οχήματος, όπως είναι η επικοινωνία με τον σταθμό επιτήρησης και η μηχανική όραση.

#### **4.2 Κινητήρας servo (FS90MG)**

Στην συγκεκριμένη διπλωματική εργασία χρησιμοποιήθηκαν 2 κινητήρες αυτού του τύπου, οι οποίοι χρησιμοποιούνται για την περιστροφή και την ανύψωση της κάμερας του σταθερού σταθμού επιτήρησης.

Ο περιστροφικός αυτός κινητήρας, χάρη στον ελεγκτή στροφών που διαθέτει, μπορεί να ελέγχει με μεγάλη ακρίβεια μεγέθη όπως η γωνιακή ταχύτητα, η γωνιακή επιτάχυνση και η γωνία περιστροφής. Η γωνία περιστροφής κυμαίνεται από 0° έως 180°, ενώ η τάση λειτουργίας από 4.8V έως 6V. Τα βασικά χαρακτηριστικά λειτουργίας του κινητήρα φαίνονται στον πίνακα 3 που ακολουθεί:

Operating Voltage	4.8V – 6V
Idle Current (at stopped)	5mA – 6mA
No Load Speed	0.12sec/60degree
Running Current (at no load)	120mA
Maximum Stall Torque	1.5kg.cm
Control System	Analog
Command Signal	PWM
Operating Angle	120 degree
Required Pulse	900 to 2100 us
Neutral Position	1500 us
Rotating Direction	CW (1500 to 900 us) CCW (1500 to 2100 us)
Motor Type	Metal brush motor
Gear Type	Metal Gear
Bearing Type	None
Horn Gear Spline	20T (4.8 mm)
Horn Type	Plastic, POM
Case	ABS
Connector Wire Length	200 +5 mm
Operating Temperature	-15 to 70°C

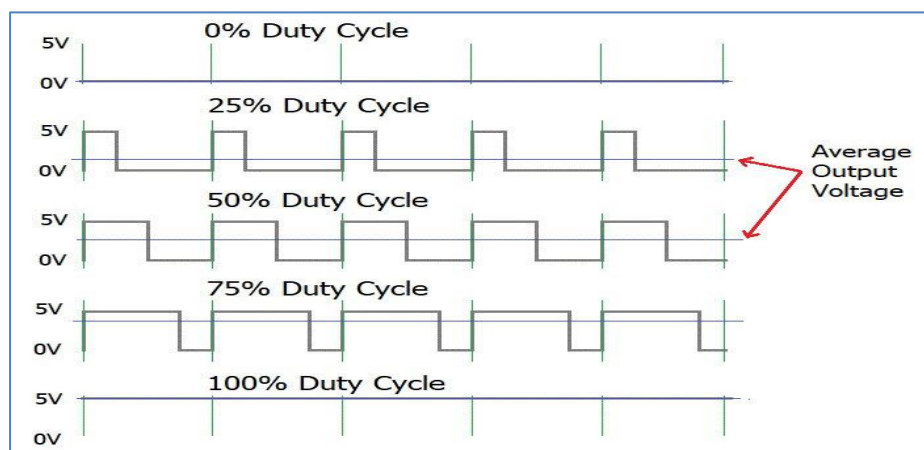
Size	23.2 × 12.5 × 22 mm
Weight	11gr

**Πίνακας 3.**  
**Τεχνικά χαρακτηριστικά κινητήρα servo (FS90MG)**



**Εικόνα 6. Κινητήρας servo FS90MG**

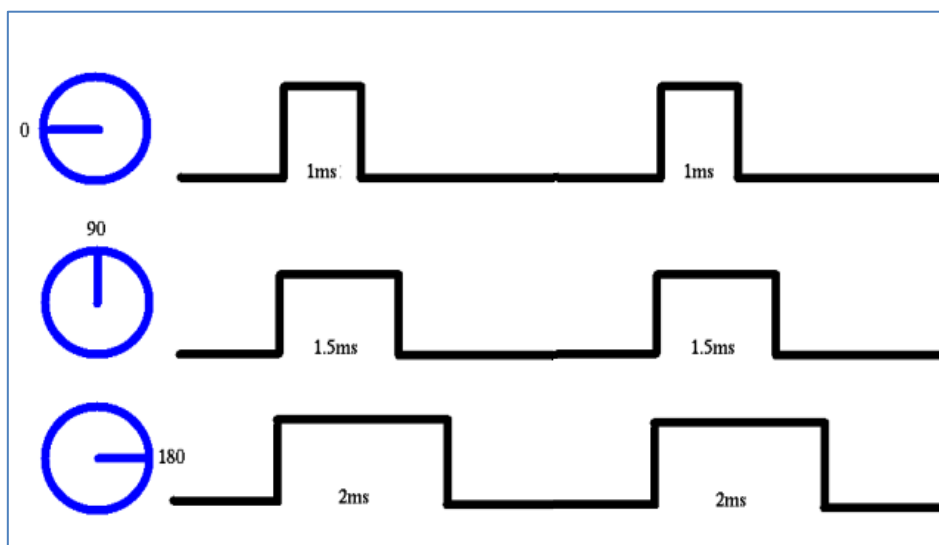
Πριν προχωρήσουμε στην λειτουργία του κινητήρα, θα πρέπει να αναλύσουμε την κυματομορφή PWM. Η κυματομορφή PWM, είναι περιοδική και μπορεί να λάβει δύο χαρακτηριστικές τιμές, τις ON και OFF. Η ON αποτελεί την μέγιστη δυνατή τιμή που μπορεί να λάβει, ενώ η OFF αντιστοιχεί στην τιμή 0. Ως κύκλος εργασίας (duty cycle) ορίζεται το ποσοστό της περιόδου κατά το οποίο η κυματομορφή βρίσκεται στο θετικό μέτωπο (ON). Η εφαρμογή μιας κυματομορφής PWM κατά την τροφοδοσία ενός φορτίου, επιτρέπει τον καλύτερο έλεγχο της ισχύος που ασκείται. Στην περίπτωση ενός κινητήρα, αυτό σημαίνει ακριβέστερο έλεγχο των στροφών του.



**Εικόνα 7.**  
**Παλμοί PWM και αντιστοιγία κύκλων εργασίας με τάσεις εξόδου**

Ένας κινητήρας servo αποτελείται από τρία βασικά τμήματα τα οποία είναι ένα κύκλωμα ελέγχου, ένα σύστημα γραναζιών και έναν κινητήρα συνεχούς τάσης. Αυτά καθιστούν τον κινητήρα αυτού του είδους ικανό να στρέφει τον άξονά του μεταξύ ενός συγκεκριμένου εύρους γωνιών, το οποίο καθορίζεται από τη διάρκεια των παλμών PWM που του ασκείται. Επιπλέον, χάρη στην ύπαρξη ενός ποτενσιόμετρου, το κύκλωμα ελέγχου του κινητήρα παρακολουθεί συνεχώς την τρέχουσα γωνία του άξονα. Όσο αυτή διαφέρει από την επιθυμητή τελική γωνία, ο κινητήρας συνεχούς ρεύματος περιστρέφει τον άξονα προς την απαιτούμενη κατεύθυνση, έως ότου η γωνία περιστροφής του ταυτιστεί με την επιθυμητή γωνία. Σύμφωνα με τα προηγούμενα η συνολική γωνία κατά την οποία περιστρέφεται ο άξονας, εξαρτάται από την θέση της αρχικής του γωνίας σε σχέση με την επιθυμητή. Κατά συνέπεια, η τάση που εφαρμόζεται στον κινητήρα είναι ανάλογη της απόστασης αυτών των γωνιών.

Η περίοδος  $T$  των παλμών PWM που δέχεται ο κινητήρας είναι 20 msec, ενώ η ελάχιστη και μέγιστη διάρκειά τους είναι 1 msec και 2 msec αντίστοιχα. Η διάρκεια του παλμού είναι αυτή που καθορίζει την γωνιακή θέση του άξονα του κινητήρα. Τέλος, όπως προαναφέραμε η γωνία περιστροφής μπορεί να κυμανθεί από  $0^\circ$  έως  $180^\circ$ . Συνεπώς, όπως φαίνεται και στην εικόνα X, παλμός διάρκειας 1 msec οδηγεί τον άξονα του κινητήρα στην μικρότερη δυνατή γωνία  $0^\circ$ , ενώ παλμός διάρκειας 2 msec στην μεγαλύτερη δυνατή γωνία  $180^\circ$ . Τέλος, παλμός διάρκειας 1.5 msec, συνεπάγεται τελική γωνία  $90^\circ$ .



**Εικόνα 8.**  
**Σχέση διάρκειας παλμών PWM με την γωνία άξονα κινητήρα servo.**

### **4.3 Samsung Galaxy A135G-SMA136B/DSN**

Αποτελεί την συσκευή κινητού τηλεφώνου στην οποία εκτελούνται όλες οι λειτουργίες του κινητού σταθμού επιτήρησης. Τα χαρακτηριστικά λειτουργίας του ακολουθούν στον πίνακα 4:

Memory	microSDXC (dedicated slot) Internal 128GB 4GB RAM
OS	Android 11, upgradable to Android 13, One UI 5.1
Chipset	Mediatek MT6833 Dimensity 700 (7 nm)
CPU	Octa-core (2x2.2 GHz Cortex-A76 & 6x2.0 GHz Cortex-A55)
GPU	Mali-G57 MC2
Camera	50 MP, f/1.8, 26mm (wide), PDAF 2 MP, f/2.4, (macro) 2 MP, f/2.4, (depth) LED flash, panorama, HDR Video 1080p@30fps
Sensors	Fingerprint (side-mounted), accelerometer, gyro, compass, barometer Virtual proximity sensing

**Πίνακας 4.**  
**Τεχνικά χαρακτηριστικά Samsung Galaxy A135G-SMA136B/DSN**

### **4.4 Microsoft LifeCam Cinema USB Camera**

Η κάμερα χρησιμοποιείται για την αναγνώριση εικόνας τόσο στο σταθερό σταθμό επιτήρησης, όσο και στο μη επανδρωμένο όχημα (UV). Τα βασικά χαρακτηριστικά λειτουργίας του φαίνονται στον πίνακα 5 που ακολουθεί:

<i>Image Sensor</i>	<i>CMOS 2 megapixel</i>
<i>Diagonal Field</i>	<i>73°</i>
<i>Resolution</i>	<i>1280x780 video</i>
<i>Autofocus</i>	<i>4'' - ∞</i>

**Πίνακας 5.**  
**Χαρακτηριστικά λειτουργίας κάμερας Microsoft LifeCam**



**Εικόνα 9. Microsoft LifeCam Cinema USB Camera**

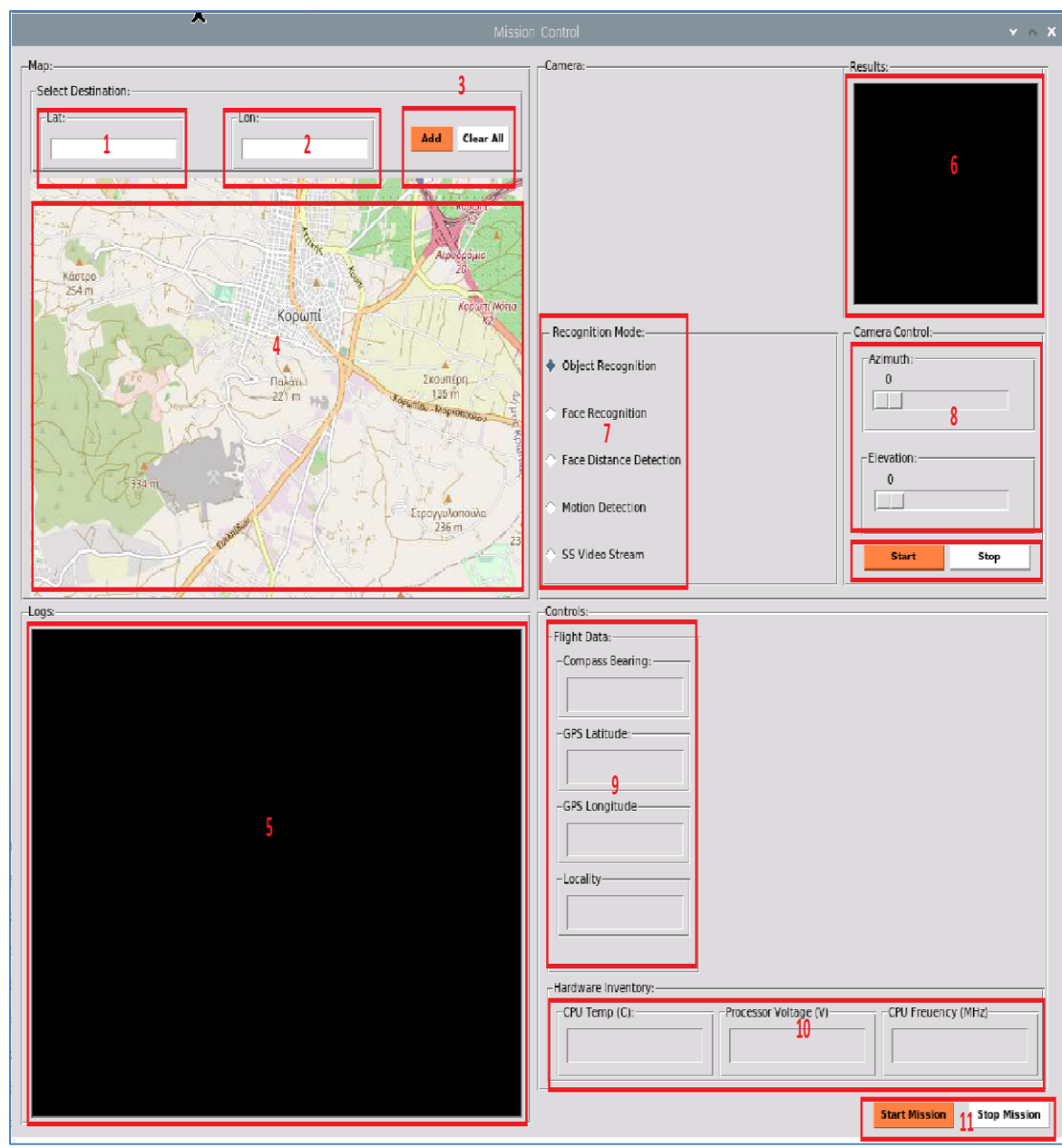
#### **4.5 Μηχανισμός στήριξης κινητήρων για κίνηση 3 διαστάσεων (PT Pan/Tilt camera platform / servo mount for servo SG90)**

Για την κίνηση των κινητήρων servo στο οριζόντιο και κατακόρυφο επίπεδο χρησιμοποιήθηκε η διάταξη που φαίνεται στην εικόνα 10 που ακολουθεί:



**Εικόνα 10. Μηχανισμός στήριξης - κίνησης κάμερας**

## 5.Γραφικό περιβάλλον εφαρμογής (GUI)



**Εικόνα 11. Γραφικό περιβάλλον εφαρμογής**

Ακολουθεί η ανάλυση των επιμέρους τμημάτων (components) του γραφικού περιβάλλοντος της εφαρμογής, καθώς και η σύντομη επεξήγηση της λειτουργίας τους:

1. Πεδίο στο οποίο εισάγεται από τον χρήστη το γεωγραφικό πλάτος του σημείου στόχου και χρησιμοποιείται κυρίως για έλεγχο εκσφαλμάτωσης (debug).
2. Πεδίο στο οποίο εισάγεται από τον χρήστη το γεωγραφικό μήκος του σημείου στόχου και χρησιμοποιείται κυρίως για έλεγχο εκσφαλμάτωσης (debug).
3. Κουμπιά που χρησιμοποιούνται για την εισαγωγή νέων σημείων στον χάρτη, αλλά και για την διαγραφή όλων των σημείων που ήδη έχουν εισαχθεί σε αυτόν.
4. Χάρτης: Παρουσιάζει την τοποθεσία του στόχου, βάσει των συντεταγμένων που έχει αποστείλει ο SS, αλλά και τα σημεία από τα οποία έχει διέλθει το μη επανδρωμένο όχημα. Η ανανέωση του χάρτη με νέα στοιχεία είναι αυτόματη. Διαθέτει δυνατότητα zoom in – zoom out με χρήση του ποντικιού.
5. Πεδίο στο οποίο καταγράφονται όλα τα δεδομένα του μη επανδρωμένου οχήματος καθώς και όλα τα σφάλματα κατά την εκτέλεση του κώδικα αυτού (debug exceptions).
6. Πεδίο στο οποίο καταγράφονται τα αποτελέσματα της αναγνώρισης εικόνας μέσω της κάμερας του μη επανδρωμένου οχήματος.
7. Επιλογή του τύπου αναγνώρισης εικόνας (recognition mode) από τον χρήστη (Αναγνώριση Αντικειμένου – Object Recognition, Υπολογισμός Απόστασης Αντικειμένου – Face Distance Detection, Εντοπισμός Κίνησης – Motion Detection, Ζωντανή Ροή Εικόνας από SS – SS Video Stream).
8. Έλεγχος από τον χρήστη της γωνίας ανύψωσης και περιστροφής της κάμερας του μη επανδρωμένου οχήματος.
9. Πεδίο στο οποίο εμφανίζονται τα δεδομένα θέσης του μη επανδρωμένου οχήματος (Ενδειξη Πυξίδας – Compass Bearing, Γεωγραφικό Πλάτος – GPS Latitude, Γεωγραφικό Μήκος – GPS Longitude, Χώρα - Locality).
10. Πεδίο στο οποίο εμφανίζονται ενδείξεις λειτουργίας του μη επανδρωμένου οχήματος (Θερμοκρασία Επεξεργαστή – CPU Temp, Τάση Λειτουργίας Επεξεργαστή – Processor Voltage, Συχνότητα Λειτουργίας Επεξεργαστή - CPU Frequency).
11. Κουμπιά που χρησιμοποιούνται για την έναρξη και παύση της εφαρμογής.



## **6. Περιγραφή αποστολής μη επανδρωμένου οχήματος (UV)**

Ακολουθεί μια όσο το δυνατόν πιο σύντομη και απλή περιγραφή του τρόπου λειτουργίας της αποστολής του μη επανδρωμένου οχήματος:

1. Ο χρήστης πατώντας το κουμπί ‘Start Mission’ (Γραφικό Περιβάλλον 11), ενεργοποιεί το μη επανδρωμένο όχημα, το οποίο μπαίνει σε κατάσταση αναμονής για πιθανή λήψη σήματος εντοπισμού στόχου από τον σταθμό επιτήρησης. Ταυτόχρονα ξεκινά και η δειγματοληψία των βασικών ενδείξεων λειτουργίας του Raspberry Pi. Στο γραφικό περιβάλλον της εφαρμογής γίνεται η εμφάνιση των τιμών αυτών των ενδείξεων (Γραφικό Περιβάλλον 10).
2. Όταν ο σταθμός επιτήρησης εντοπίσει κάποιον στόχο, τότε αποστέλλει το μήνυμα εντοπισμού στο μη επανδρωμένο όχημα. Ο χειριστής ενημερώνεται με ένα popup μήνυμα πως έχει εντοπιστεί κάποιος στόχος και αποφασίζει αν θα ξεκινήσει η αποστολή του μη επανδρωμένου οχήματος.
3. Σε περίπτωση θετικής απόκρισης του χειριστή, ο χάρτης του γραφικού περιβάλλοντος (Γραφικό περιβάλλον 4) ενημερώνεται με τις συντεταγμένες του και η αποστολή ξεκινά.
4. Το μη επανδρωμένο όχημα, κατά την μετακίνησή του, ενημερώνει ανά περιοδικά διαστήματα τον χάρτη του γραφικού περιβάλλοντος και τα κατάλληλα πεδία (Γραφικό Περιβάλλον 9) με τις τρέχουσες συντεταγμένες θέσεις και την κατεύθυνσή του αντίστοιχα. Ως αισθητήρας GPS, χρησιμοποιείται μία συσκευή κινητού τηλεφώνου.
5. Σε όλη την διάρκεια της αποστολής ο χειριστής μπορεί να ενεργοποιήσει όλες τις υποστηριζόμενες μορφές αναγνώρισης εικόνας, μέσω της κάμερας USB του μη επανδρωμένου οχήματος (Γραφικό Περιβάλλον 7).
6. Η αποστολή μπορεί να τερματιστεί από τον χειριστή χρησιμοποιώντας το κουμπί ‘Stop’ (Γραφικό Περιβάλλον 11).

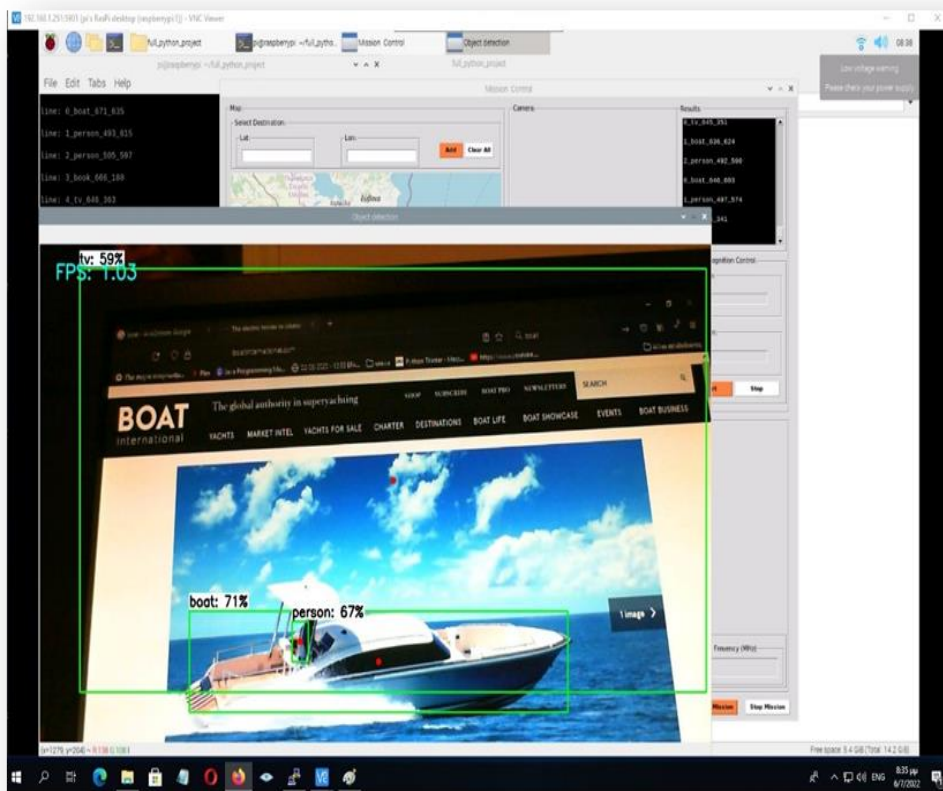
## **7. Υποστηριζόμενες μορφές μηχανικής όρασης**

Όπως έχει αναφερθεί, στο μη επανδρωμένο όχημα όσο και στον σταθμό επιτήρησης έχει γίνει η ανάπτυξη συγκεκριμένων μοντέλων και μηχανισμών αναγνώρισης. Θα πρέπει να σημειωθεί πως τα μοντέλα που αναπτύχθηκαν στο μη

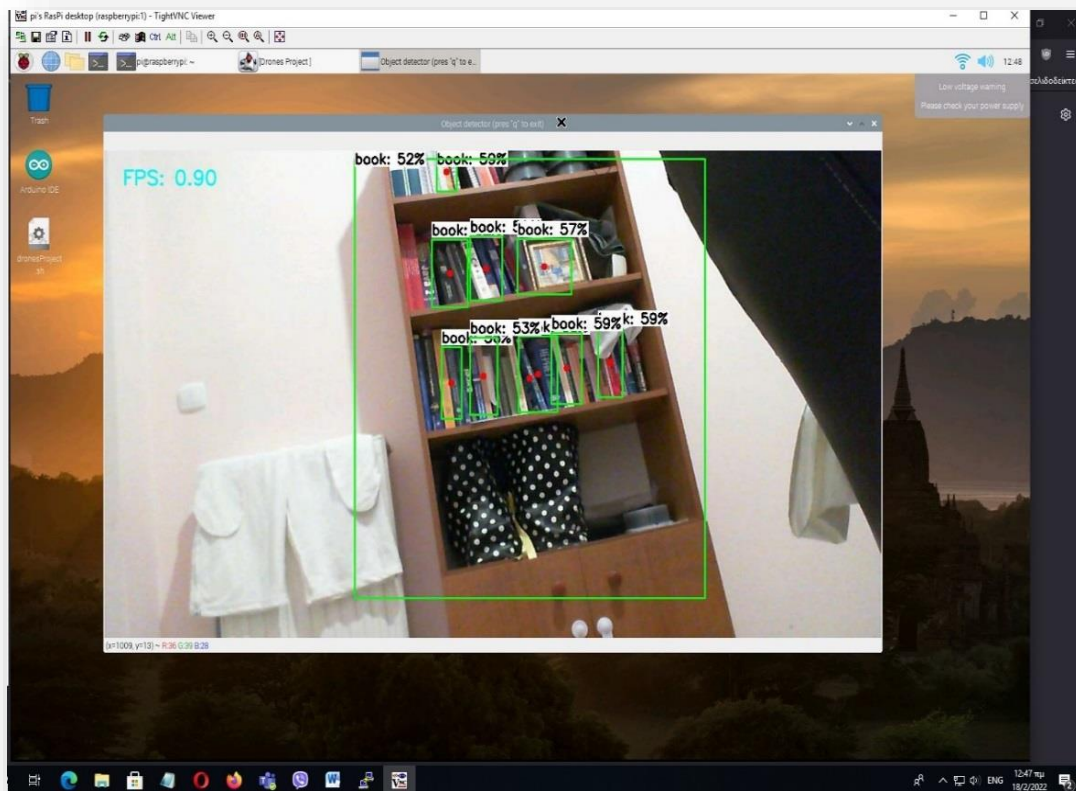
επανδρωμένο όχημα μπορούν να αναπτυχθούν με μεγάλη ευκολία και στον σταθμό επιτήρησης, τόσο στην σταθερή όσο και στην κινητή του μορφή και αντίστροφα. Ακολουθεί η ανάλυση της λειτουργίας των μηχανισμών μηχανικής όρασης που αναπτύχθηκαν:

## 7.1 Μη επανδρωμένο όχημα

### 7.1.1. Εντοπισμός - αναγνώριση αντικειμένων (object detection - recognition)



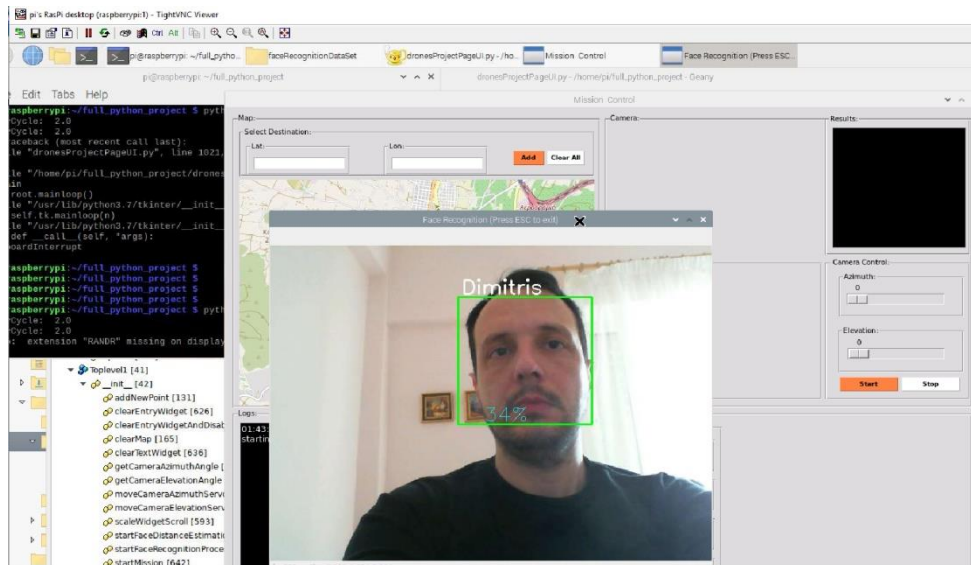
Εικόνα 12. Εντοπισμός-αναγνώριση αντικειμένων “βάρκα” και “άνθρωπος”



**Εικόνα 13. Εντοπισμός-αναγνώριση αντικειμένων “βιβλίο”**

Ο μηχανισμός αυτός επιτρέπει τον εντοπισμό και την αναγνώριση συγκεκριμένων αντικειμένων. Στην περίπτωση του Σταθερού Σταθμού Επιτήρησης, η αναγνώριση γίνεται με χρήση της βιβλιοθήκης OpenCV 2 και της πλατφόρμας TensorFlow, το οποίο χρησιμοποιεί το μοντέλο εκπαίδευσης coco\_ssd\_mobilenet\_v1. Το συγκεκριμένο μοντέλο έχει βελτιστοποιηθεί προκειμένου να αναγνωρίζει σε πολύ μικρό χρόνο έως και 100 διαφορετικά είδη αντικειμένων.

### 7.1.2. Εντοπισμός - αναγνώριση προσώπου (face detection - recognition)

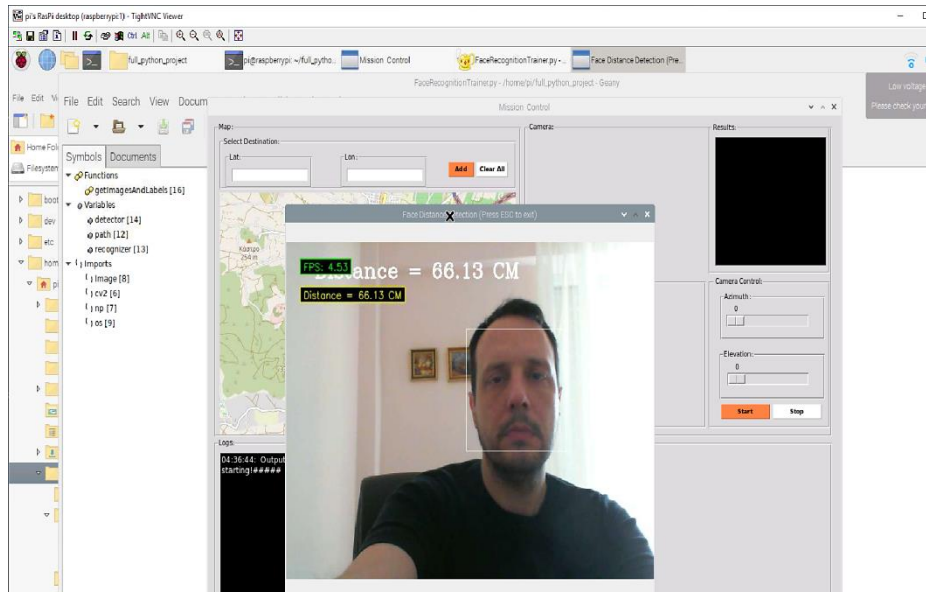


**Εικόνα 14. Στιγμιότυπο αναγνώρισης προσώπου**

Ο συγκεκριμένος μηχανισμός που αναπτύχθηκε επιτρέπει όχι μόνο την αναγνώριση προσώπων σε πραγματικό χρόνο (live), μέσω της κάμερας αναγνώρισης, αλλά και την ταύτιση των προσώπων με συγκεκριμένα άτομα. Και σε αυτήν την περίπτωση η αναγνώριση γίνεται με χρήση της βιβλιοθήκης OpenCV 2. Πιο συγκεκριμένα χρησιμοποιούνται οι αλγόριθμοι LBPH (Local Binary Patterns Histograms) και Haar-Cascade.

Σημαντικό και απαραίτητο βήμα για την επιτυχή αναγνώριση, αποτελεί η διαδικασία της εκπαίδευσης, η οποία γίνεται με την εισαγωγή φωτογραφιών των προς αναγνώριση ατόμων. Ακολουθεί η επεξεργασία των εισαγόμενων δεδομένων και η παραμετροποίηση του τελικού αρχείου που περιέχει τις απαραίτητες τιμές για την κατηγοριοποίηση και διαχωρισμό των προς αναγνώριση προσώπων.

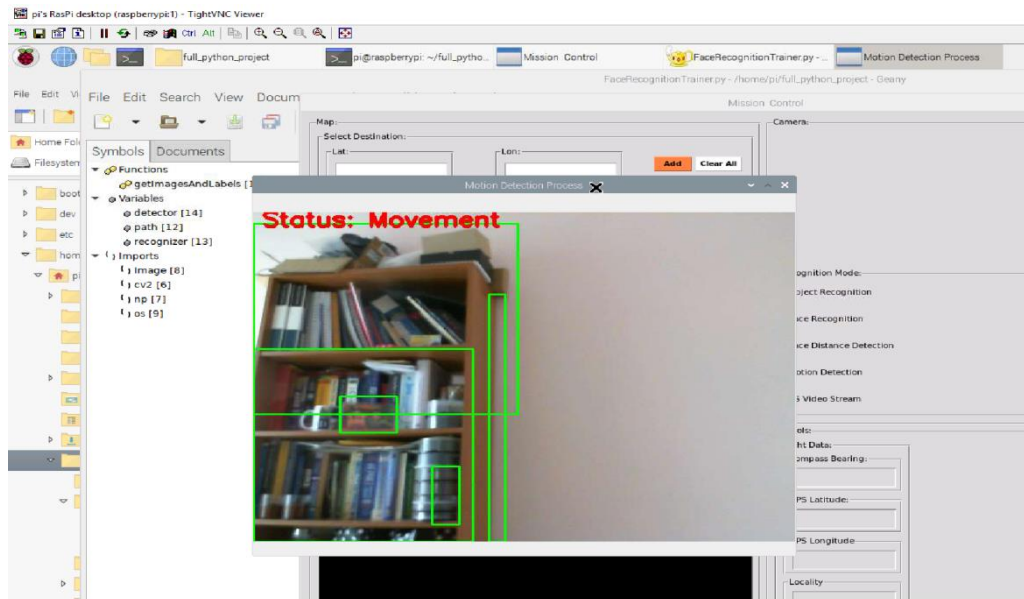
### 7.1.3. Υπολογισμός απόστασης προσώπου (face distance detection)



**Εικόνα 15. Στιγμιότυπο υπολογισμού απόστασης προσώπου**

Ο συγκεκριμένος μηχανισμός που αναπτύχθηκε επιτρέπει τον εντοπισμό προσώπου και τον υπολογισμό της απόστασής του από την κάμερα, σε πραγματικό χρόνο (live) και με μεγάλο βαθμό ακρίβειας. Όπως είναι εύκολα κατανοητό, τη θέση του προς αναγνώριση αντικείμενου μπορεί να λάβει οποιοδήποτε άλλο αντικείμενο αν χρησιμοποιηθεί το κατάλληλο μοντέλο εκπαίδευσης. Η αναγνώριση προσώπου γίνεται και εδώ με χρήση της βιβλιοθήκης OpenCV 2 και του αλγορίθμου Haar-Cascade, αφού δοθούν οι αναμενόμενες μέσες διαστάσεις πλάτους και ύψους ενός προσώπου. Θα πρέπει να σημειωθεί πως ο υπολογισμός της απόστασης γίνεται με πολύ μεγάλη ακρίβεια σε περίπτωση που το προς εντοπισμό αντικείμενο, όπως είναι το πρόσωπο, βρίσκεται στο ίδιο ύψος με την κάμερα εντοπισμού. Σε αντίθετη περίπτωση, για την εξαγωγή σωστών αποτελεσμάτων θα πρέπει να γίνει προσθήκη τεχνικών φωτογραμμετρίας.

#### 7.1.4. Εντοπισμός κίνησης (motion detection)



**Εικόνα 16. Στιγμιότυπο εντοπισμού κίνησης**

Με τον μηχανισμό αυτό καθίσταται εφικτός ο εντοπισμός οποιασδήποτε κίνησης σε πραγματικό χρόνο (live), μέσω της κάμερας αναγνώρισης. Ο εντοπισμός γίνεται με χρήση της βιβλιοθήκης OpenCV 2. Σε περίπτωση εντοπισμού κίνησης από οποιοδήποτε αντικείμενο εντός του οπτικού πεδίου της κάμερας αναγνώρισης, γίνεται η επισήμανση των αντίστοιχων frames της εικόνας και εμφανίζεται και το κατάλληλο μήνυμα, όπως φαίνεται στην εικόνα 16.

#### 7.1.5. Ζωντανή ροή εικόνας από σταθερό σταθμό επιτήρησης

Στην περίπτωση αυτή, πραγματοποιείται ζωντανή μετάδοση της εικόνας της κάμερας του σταθερού σταθμού επιτήρησης στο μη επανδρωμένο όχημα, χωρίς την χρήση τεχνικών μηχανικής όρασης. Με αυτόν το τρόπο οι χρήστες που είναι συνδεδεμένοι στο μη επανδρωμένο όχημα, μπορούν να παρακολουθούν την εικόνα που τους μεταδίδεται από τον σταθμό.

## **7.2. Κινητός σταθμός επιτήρησης**

Ως κινητός σταθμός επιτήρησης χρησιμοποιήθηκε μία συσκευή κινητού τηλεφώνου. Η επιλογή του κινητού τηλεφώνου δεν είναι τυχαία για τους εξής λόγους:

1. Το κόστος του είναι χαμηλό.
2. Είναι εύκολη η χρήση του.
3. Η χρήση του είναι άκρως διαδεδομένη, αφού σήμερα σχεδόν όλοι έχουν τουλάχιστον μία.
4. Είναι εύκολος ο προγραμματισμός του, ειδικά στους απαραίτητους τομείς της μηχανικής όρασης και του γεωεντοπισμού.

Όλοι οι προαναφερθέντες λόγοι καθιστούν εύκολη την ανάπτυξη ενός συστήματος επιτήρησης και εντοπισμού ειδικού σκοπού, (φύλαξη συνόρων, παρατήρηση εχθρικού εδάφους κλπ.), το οποίο θα μπορούσε κατά κύριο λόγο να βασιστεί στις ατομικές συσκευές κινητών τηλεφώνων μη στρατιωτικού προσωπικού, όπως είναι για παράδειγμα οι κάτοικοι των παραμεθόριων περιοχών. Η πρόκληση σε αυτήν την περίπτωση είναι η ανάπτυξη των κατάλληλων και εξειδικευμένων μοντέλων αναγνώρισης τα οποία απαιτούνται για την ορθή λειτουργία τέτοιου είδους εφαρμογών.

Στα πλαίσια της συγκεκριμένης διπλωματικής εργασίας, η αναγνώριση αντικειμένων υλοποιήθηκε χρησιμοποιώντας την πλατφόρμα TensorFlow Lite, ενώ ως αντικείμενο στόχος ορίστηκε η συσκευή ποντικιού (mouse) του ηλεκτρονικού υπολογιστή.

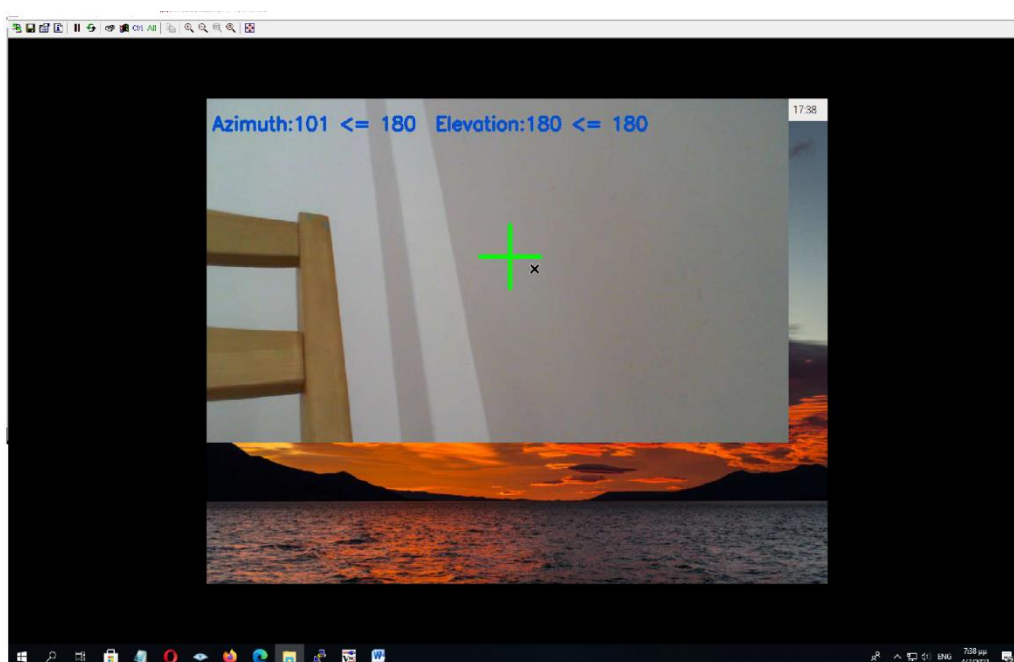
## **7.3. Σταθερός σταθμός επιτήρησης**

### **7.3.1. Χειρισμός κίνησης κάμερας μέσω ζωντανής εικόνας (manual handling)**

Μία από τις βασικές απαιτήσεις των συστημάτων παρατήρησης, είναι η δυνατότητα απομακρυσμένου χειρισμού της κίνησης της κάμερας, σε ζωντανό χρόνο, έτσι ώστε να γίνεται ευκολότερος ο εντοπισμός και η ανάλυση ενδεχόμενων στόχων. Στην περίπτωση του σταθερού σταθμού επιτήρησης, υλοποιήθηκε η ανάπτυξη αυτού του χειρισμού. Πάνω στην προβολή της ζωντανής εικόνας από την κάμερα, ο χρήστης, χρησιμοποιώντας το ποντίκι του υπολογιστή, μπορεί να ελέγχει τον μηχανισμό κίνησης της κάμερας (εικόνα 17) στο τρισδιάστατο επίπεδο.



Το κέντρο της εικόνας αποτελεί την αρχή των αξόνων κίνησης  $O(0,0)$  και εμφανίζεται με πράσινο σταυρό. Όπως έχουμε προαναφέρει, το εύρος της κίνησης των κινητήρων servo που χρησιμοποιήθηκαν είναι από  $0^\circ$  έως  $180^\circ$ . Η απόσταση της θέσης του κέρσορα του ποντικιού σε pixel από το σημείο  $O$ , μετατρέπεται σε γωνία κίνησης, η οποία και εκτελείται από τον μηχανισμό κίνησης. Φυσικά, αν η γωνία της κίνησης υπερβαίνει τη μέγιστη δυνατή γωνία των κινητήρων servo, τότε αυτή δεν πραγματοποιείται. Μετά από κάθε κίνηση, η τρέχουσα γωνία τόσο στο οριζόντιο όσο και στο κατακόρυφο επίπεδο εμφανίζονται στην οθόνη (ενδείξεις Azimuth και Elevation). Ένα στιγμιότυπο αυτής της διαδικασίας εμφανίζεται στην εικόνα 17 που ακολουθεί:



**Εικόνα 17. Στιγμιότυπο απομακρυσμένου χειρισμού κίνησης κάμερας**

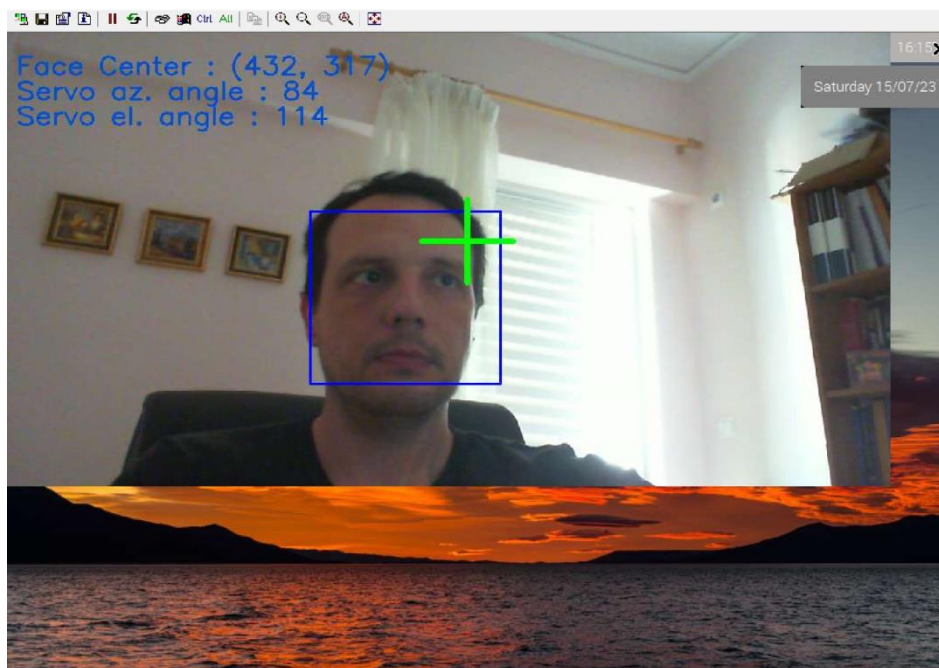
### **7.3.2. Εντοπισμός αντικειμένου και εστίαση σε αυτό (face detection and tracking)**

Όπως έχουμε δει, μία από τις ικανότητες που έχουν αναπτυχθεί στα πλαίσια αυτής της διπλωματικής είναι ο εντοπισμός συγκεκριμένων αντικειμένων ή προσώπων. Μία επιπλέον ικανότητα που μπορεί να πολλαπλασιάσει τις ικανότητες ενός συστήματος εντοπισμού, είναι και η αυτόματη εστίαση σε αυτό, ακόμη και σε περίπτωση κίνησής του.



Σύμφωνα με τον αλγόριθμο λειτουργίας, όταν εντοπιστεί ο επιθυμητός τύπος αντικειμένου, λαμβάνονται οι συντεταγμένες των pixel του περιγράμματος του (παράθυρο). Για κάθε frame που λαμβάνεται από την ζωντανή ροή εικόνας της κάμερας, γίνεται ο υπολογισμός της κίνησης, ως προς το οριζόντιο και κατακόρυφο επίπεδο, προκειμένου το κέντρο της εικόνας της κάμερας (σταυρός), να κινηθεί προς το παράθυρο του εντοπισμένου αντικειμένου. Όταν το κέντρο κάμερας ταυτιστεί με οποιοδήποτε σημείο από το εσωτερικό του περιγράμματος του αντικειμένου, τότε η κίνηση του μηχανισμού στήριξης της κάμερας σταματά. Στην συγκεκριμένη περίπτωση, ως επιθυμητό αντικείμενο προς εντοπισμό έχει οριστεί το ανθρώπινο πρόσωπο.

Θα πρέπει να σημειωθεί πως η αποτελεσματικότητα και η ταχύτητα του αλγορίθμου μπορεί να βελτιωθεί κατά πολύ, αν χρησιμοποιηθούν τεχνικές εκτίμησης των μεταβλητών κατάστασης του συστήματος, όπως για παράδειγμα το φίλτρο Kalman. Ένα στιγμιότυπο της διαδικασίας ακολουθεί στην εικόνα 18.



**Εικόνα 18. Στιγμιότυπο face tracking**

## **8. Συμπερασματικός επίλογος – Μελλοντική Έρευνα**

Τα συστήματα εντοπισμού και έγκαιρης προειδοποίησης αποτελούν ένα πεδίο με ευρύτατο φάσμα πολιτικών και στρατιωτικών εφαρμογών. Θεωρώντας την

συγκεκριμένη διπλωματική ως βάση, ανοίγονται πολλοί τομείς έρευνας για το μέλλον, οι οποίοι θα διευρύνουν το πεδίο εφαρμογής συστημάτων αυτού του είδους.

Κάποιες προτάσεις είναι οι ακόλουθες:

1. Δημιουργία και ανάπτυξη νέων μοντέλων αναγνώρισης, προκειμένου να αναπτυχθούν συστήματα με πολύ εξειδικευμένες ικανότητες, όπως για παράδειγμα αναγνώριση του τύπου στρατιωτικών οχημάτων.
2. Ανάπτυξη συστημάτων αναγνώρισης με συνδυαστική χρήση πραγματικών αυτόνομων οχημάτων, στα πλαίσια ενός συστήματος οδηγού – ακόλουθου (master - slave). Κατά την άποψη του συγγραφέα, η επίτευξη του συγκεκριμένου στόχου είναι εύκολα υλοποιήσιμη με πολύ λίγες αλλαγές και προσθήκες στον κώδικα της συγκεκριμένης διπλωματικής.
3. Ανάπτυξη συστημάτων άλλου είδους για στρατιωτικούς σκοπούς, όπως για παράδειγμα η κατασκευή αυτόνομου οπλικού σταθμού με χρήση μηχανικής όρασης, κάτι το οποίο κατά την άποψη του συγγραφέα αποτελεί ένα πρώτο βήμα για την ανάπτυξη πιο πολύπλοκων κατασκευών, όπως για παράδειγμα ένας αυτόνομος πύργος άρματος.

## **Βιβλιογραφία**

- [1] Villan Alberto Fernandez, (2019), *Mastering OpenCV 4 with Python: A practical guide covering topics from image processing, augmented reality to deep learning with OpenCV 4 and Python 3.7*, Birmingham, United Kingdom, Packt
- [2] Minichino Joe, Howse Joseph, (2019), *Learning OpenCV 4 Computer Vision with Python 3, Get to Grips with Tools, Techniques, and Algorithms for Computer Vision and Machine Learning*, Birmingham, United Kingdom, Packt
- [3] Rajalingappaa Shanmugamani, (2018), *Deep Learning for Computer Vision, Expert Techniques to Train Advanced Neural Networks Using TensorFlow and Keras*, Birmingham, United Kingdom, Packt
- [4] Blokdyk Geraldus, (2018), *TensorFlow: A Complete Guide*, New York, 5STARCOOKS
- [5] Moroney Lawrence, (2021), *AI and Machine Learning for On-Device Development*, Sebastopol, California, O'Reilly
- [6] Sufyan bin Uzayr, (2023), *Mastering Android Studio: A Beginner's Guide (Mastering Computer Science)*, Boca Raton, Florida, CRC Press
- [7] Daniel Wilczak, (25 April 2021), *Control a Stepper Motor using Python and a Raspberry Pi*, medium.com, <https://danielwilczak101.medium.com/control-a-stepper-motor-using-python-and-a-raspberry-pi-11f67d5a8d6d>
- [8] Liz Miller, (25 April 2021), *How to Control a Servo with Raspberry Pi*, <https://www.learnrobotics.org/blog/raspberry-pi-servo-motor/>
- [9] Adrian Rosebrock, (30 March 2015), *Accessing the Raspberry Pi Camera with OpenCV and Python*, pyimagesearch.com, <https://pyimagesearch.com/2015/03/30/accessing-the-raspberry-pi-camera-with-opencv-and-python/>
- [10] Monk Simon, (2022), *Raspberry Pi Cookbook*, Sebastopol, California, O'Reilly
- [11] Burger Wilhelm, Burge Mark, (2016), *Digital Image Processing*, London, United Kingdom, Springer
- [12] Dey Sandipan, Burge Mark, (2016), *Hands-On Image Processing with Python: Expert techniques for advanced image analysis and effective interpretation of image data*, Birmingham, United Kingdom, Packt
- [13] Ravishankar Chityala, Sridevi Pudipeddi, (2021), *Image Processing and Acquisition using Python*, Boca Raton, Florida, CRC Press

