

Διπλωματική Εργασία

Σχεδίαση και υλοποίηση βοηθού ρομποτικού βραχίονα
με φωνητική αλληλεπίδραση και χρήση μηχανικής
όρασης

Νάντζιος Γεώργιος

Επιβλέπων καθηγητής: Δρ.Μηνάς Δασυγένης

ΕΡΓΑΣΤΗΡΙΟ ΡΟΜΠΟΤΙΚΗΣ ΕΝΣΩΜΑΤΩΜΕΝΩΝ ΚΑΙ ΟΛΟΚΛΗΡΩΜΕΝΩΝ ΣΥΣΤΗΜΑΤΩΝ

<https://arch.ece.uowm.gr/>

Οκτώβριος 2021

Δομή Παρουσίασης

- ❖ Εισαγωγή
- ❖ Παρόμοιες Υλοποιήσεις
- ❖ Θεωρητικό Υπόβαθρο
- ❖ Υλικό Μέρος
- ❖ Λογισμικό Μέρος
- ❖ Σενάρια & Εκτέλεση Πειραμάτων
- ❖ Συμπεράσματα & Μελλοντικές Επεκτάσεις



ΕΙΣΑΓΩΓΗ

Παρουσίαση συστήματος

Ολοκληρωμένο σύστημα ρομποτικού βραχίονα μεταβίβασης αντικειμένων στο χρήστη με φωνητική αλληλεπίδραση

- Αναγνώριση βάσει :
 - Σχήματος & χρώματος
 - Αλγορίθμου αναγνώρισης του είδους αντικειμένων YOLO
 - QR code
- Φωνητική αλληλεπίδραση
 - Ελληνικά
 - Αγγλικά
- Βασικές λειτουργίες
 - Εύρεσης και ανάγνωσης των αντικειμένων που το περιβάλλουν
 - Παράδοση των αντικειμένων στο χέρι του χρήστη
 - Λειτουργεί σε συνθήκες πραγματικού χρόνου

Εφαρμογές και κίνητρα

- Συνδυαστική εφαρμογή διαφορετικών τεχνολογιών
- Βοήθεια στην καθημερινότητα των ανθρώπων
- Πολλαπλές χρήσεις:
 - Υγείας (ιατρικών επεμβάσεων & περίθαλψη ασθενών)
 - Βιομηχανίας (γραμμή παραγωγής, ταξινόμηση και οργάνωση αντικειμένων)
 - Οικιακό περιβάλλον (εργασίες του νοικοκυριού)
 - Παροχή βοήθειας σε ΑμΕΑ και ηλικιωμένους

Παρόμοιες Υλοποιήσεις

- Περιορισμένο εύρος αναγνώρισης αντικειμένων
- Δυσχρηστίας της εφαρμογής και απαίτηση για φωνητική καθοδήγηση
- Μικρή επιφάνεια τοποθέτησης αντικειμένων
- Υψηλό ποσοστό αστοχίας, προσέγγισης αντικειμένου, έλλειψη ανάδρασης επενεργητών
- Λειτουργία υπό συγκεκριμένες συνθήκες





ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

Θεωρητικό υπόβαθρο

- Εργαλεία ανάπτυξης λογισμικού

- PyCharm (IDE)
- Notepad++
- Niryo-One Studio



Γλώσσα και Βιβλιοθήκες

- Python
- Μηχανική όραση

- OpenCV
- Zbar
- Yolo Algorithm

- Φωνητική αναγνώριση

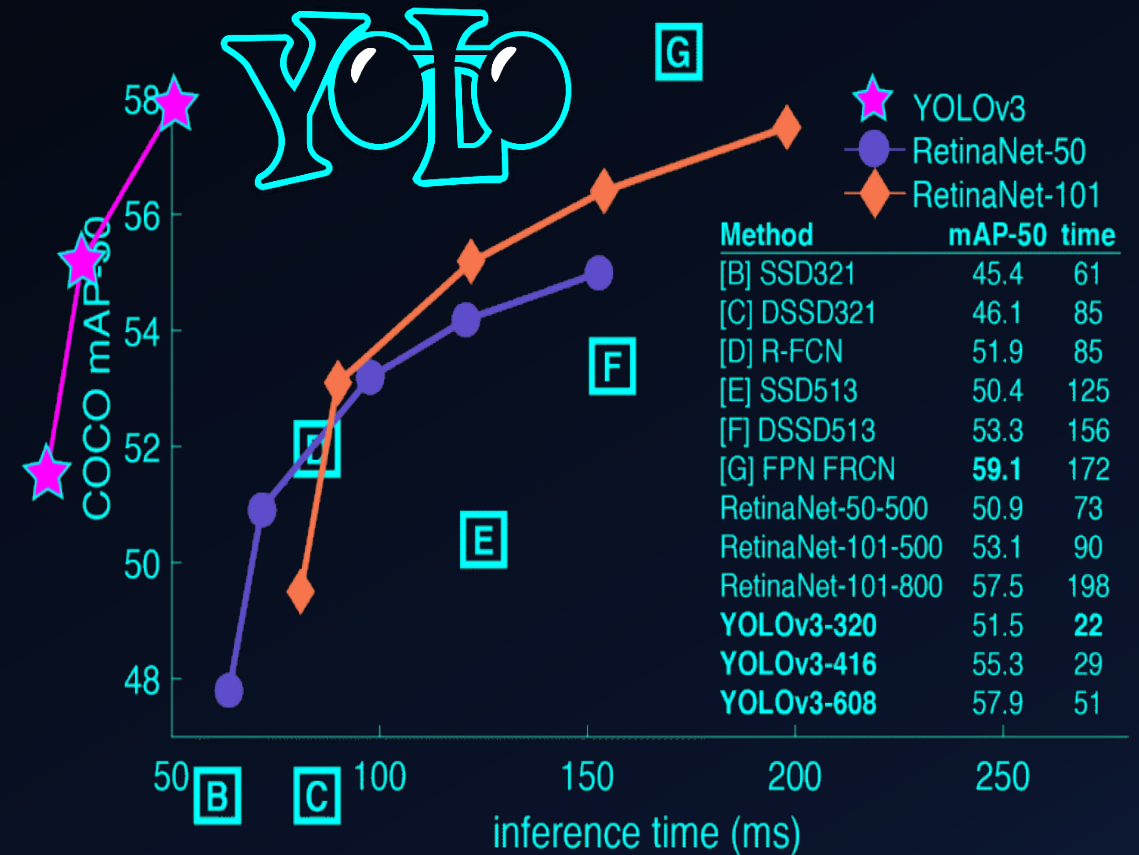
- Google voice Recognition

- Φωνητική απάντηση

- pyttsx3

Γιατί Yolo;

- You Only Look Once
- Πλήρες συνδεδεμένο νευρωνικό δίκτυο
- Καλή ισορροπία ακρίβειας και ταχύτητας
- Ταυτόχρονος εντοπισμός διαφορετικών αντικειμένων σε μία εικόνα
- Εντοπισμός θέσης αντικειμένων στο χώρο

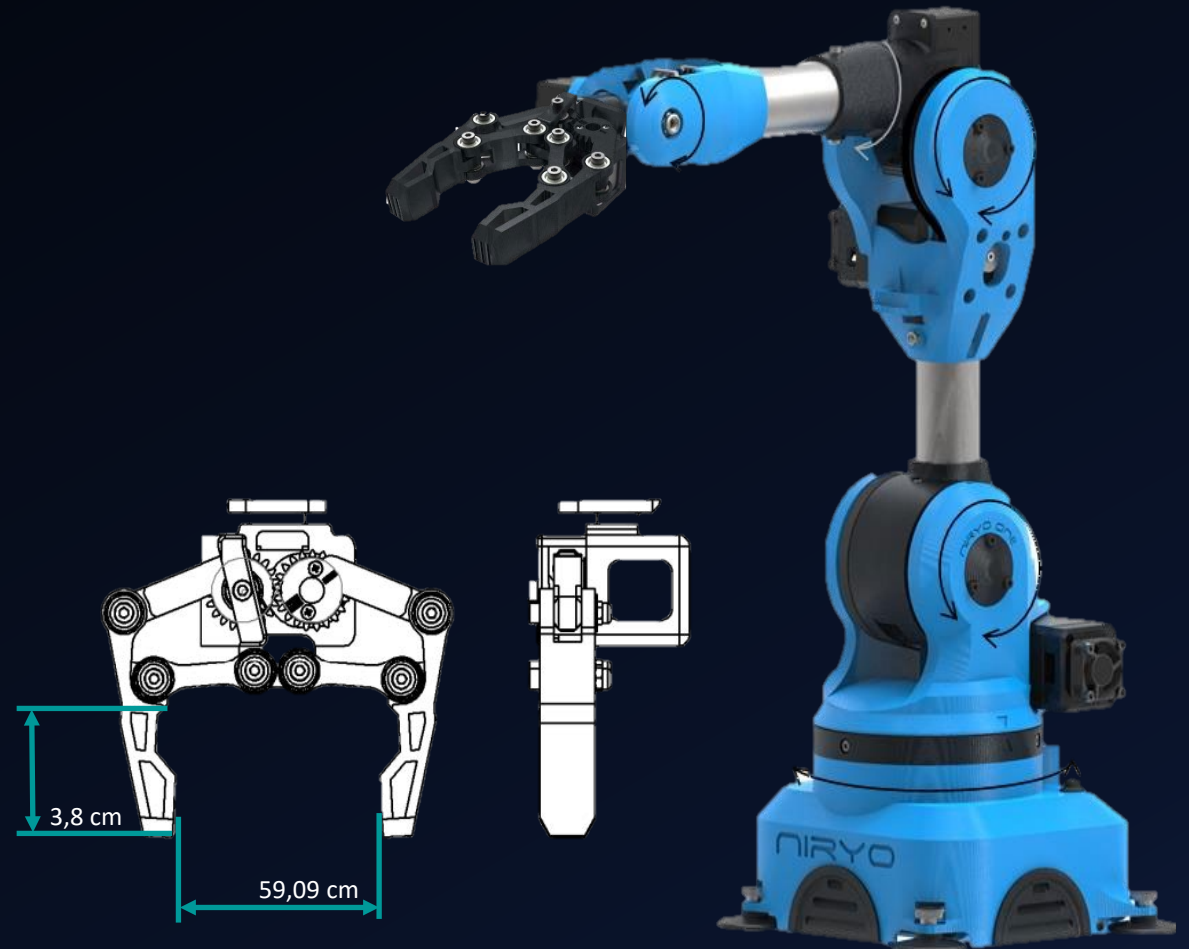




ΥΛΙΚΟ ΜΕΡΟΣ

Niryo-One

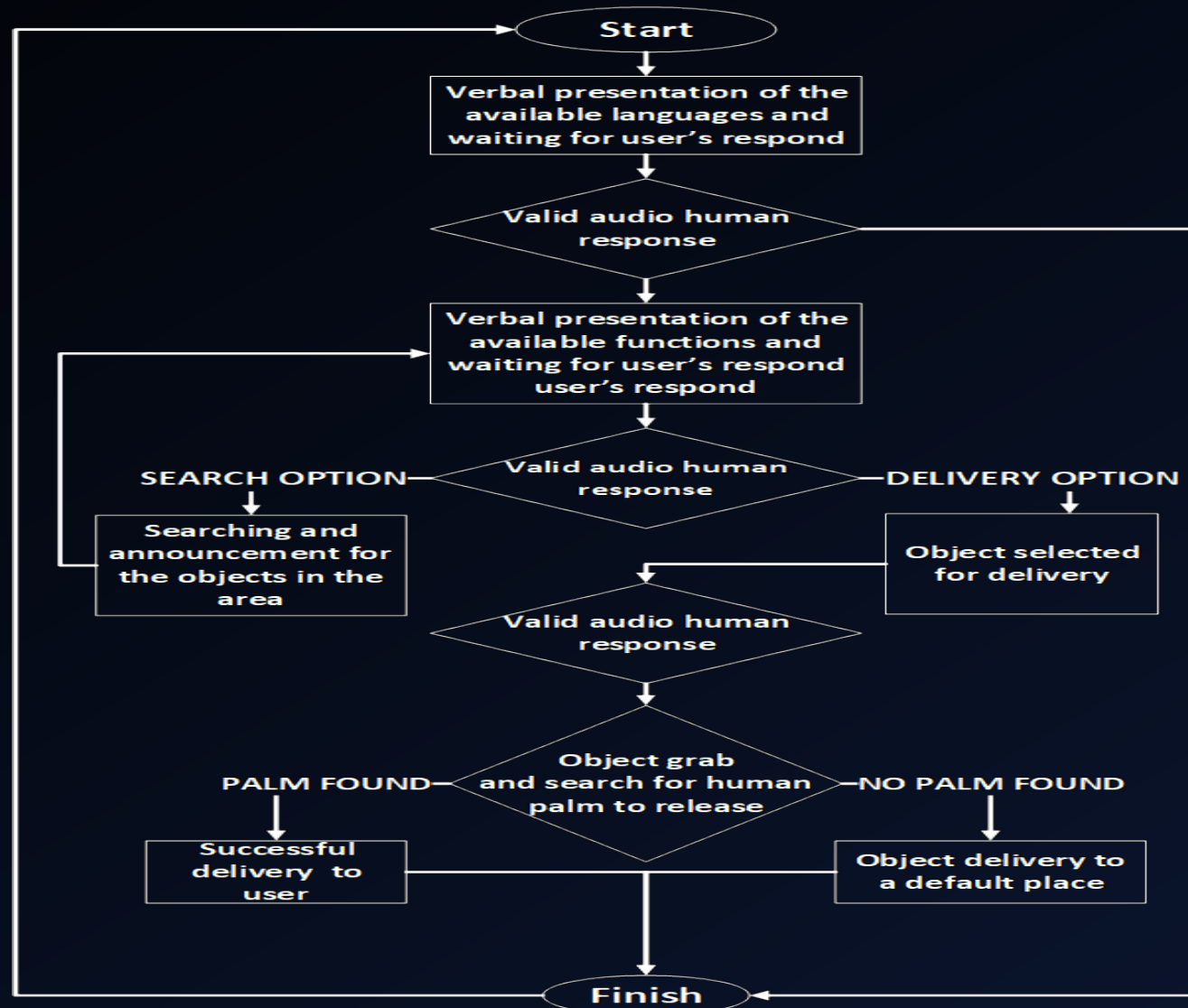
- Βραχίονας 6 βαθμών ελευθερίας
- Υποστήριξη Python API
- Μέγιστη απόσταση αρπαγής **440mm**
- Βάρος **3.3kg**
- Ανοιχτού κώδικα και διαθέσιμα STL αρχεία



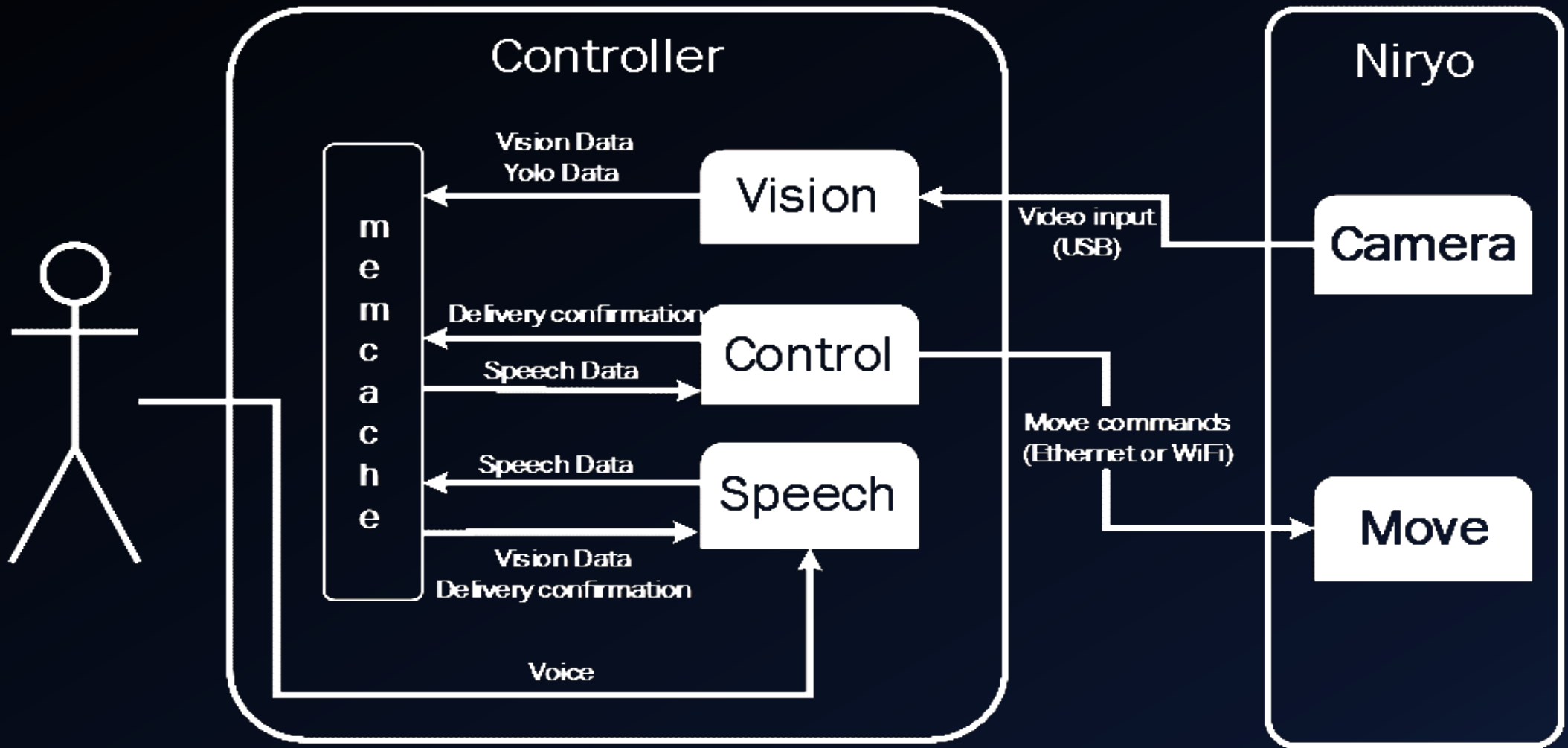


ΛΟΓΙΣΜΙΚΟ ΜΕΡΟΣ

Διάγραμμα λειτουργίας

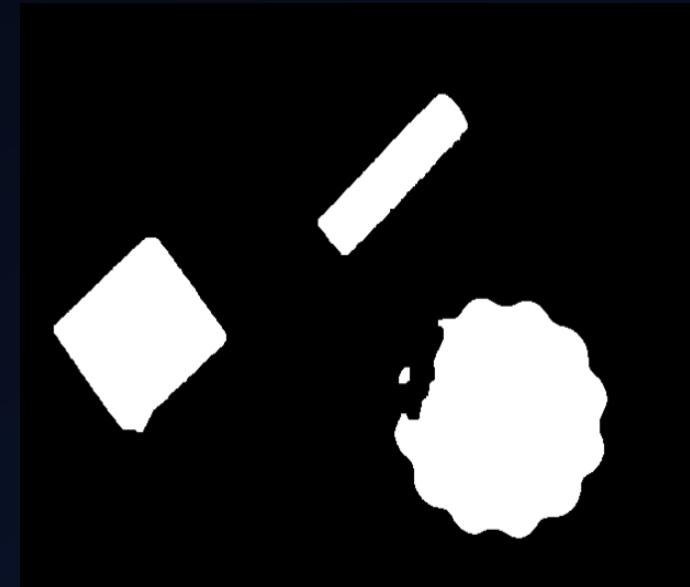
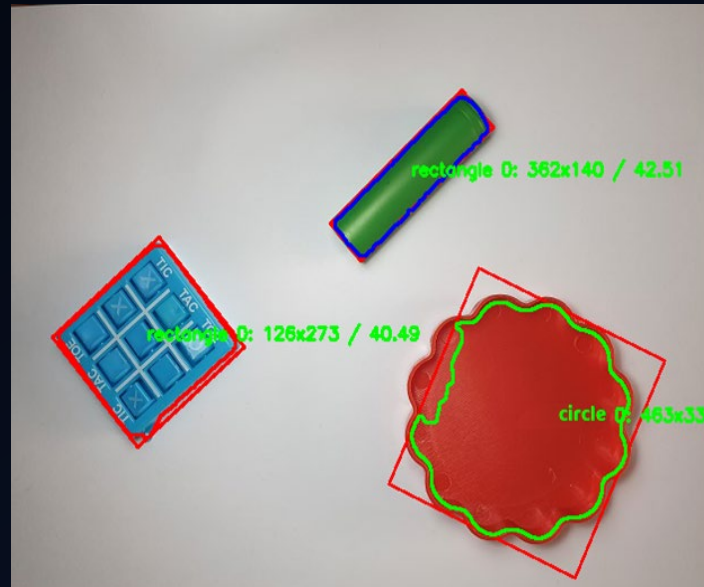


Αρχιτεκτονική συστήματος



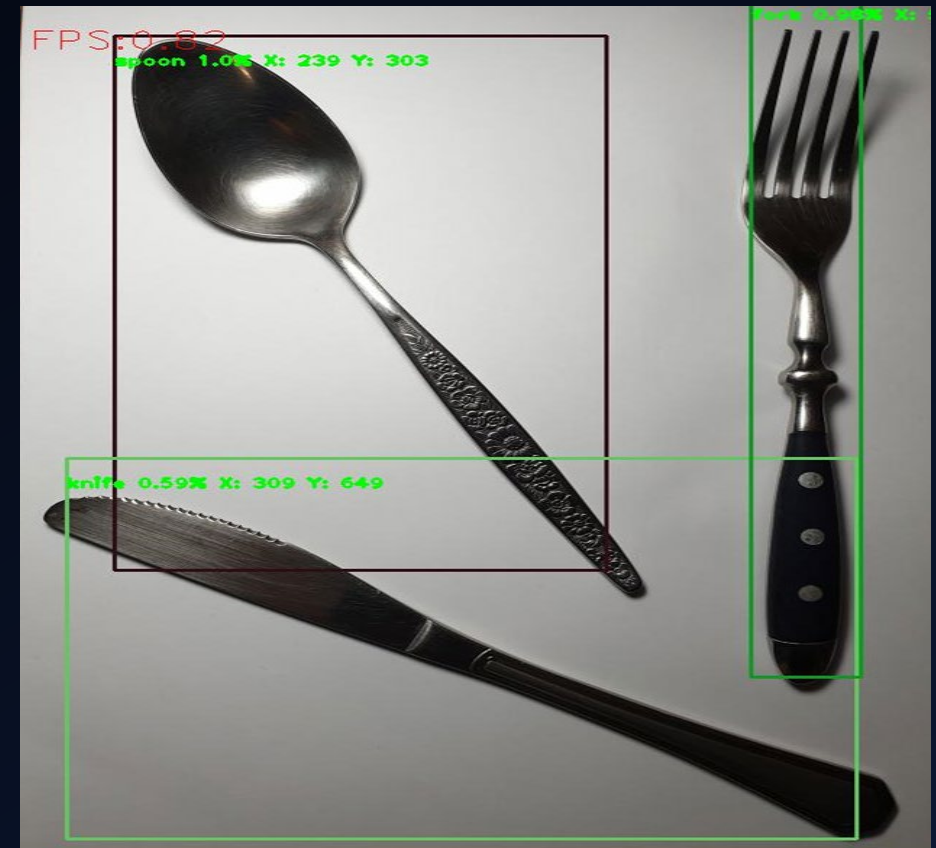
Όραση: Εντοπισμός σχήματος και χρώματος

- Δυνατότητα εντοπισμού οποιουδήποτε χρώματος
- Εντοπισμός τετράγωνο, ορθογώνιο, τρίγωνο, πεντάγωνο και κύκλος
- Εύρεση των περιγραμμάτων
- Υπολογισμός κέντρου μάζας, γωνίας κλίσης συντεταγμένων χώρου και σχήματος



Όραση: Αναγνώριση με YOLO

- Αναγνώριση πλήθος καθημερινών αντικειμένων χωρίς την εκπαίδευση από το χρήστη
- Χρήση OpenCV ως framework
- Χρήση COCO dataset ως pretrained model με 80 διαφορετικά ονόματα αντικειμένων
- Εύρεση ετικέτας κατηγορίας, βαθμολογίας εμπιστοσύνης, και συντεταγμένες κέντρου



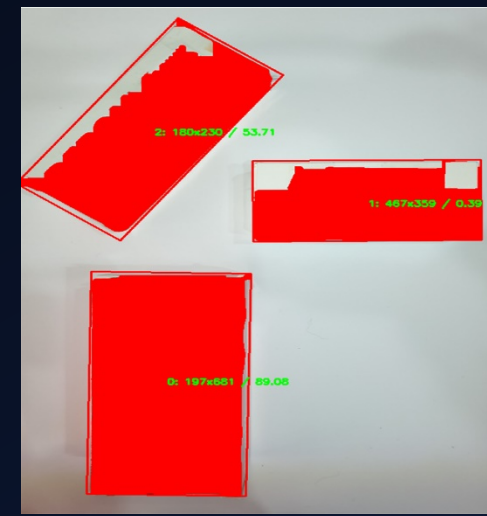
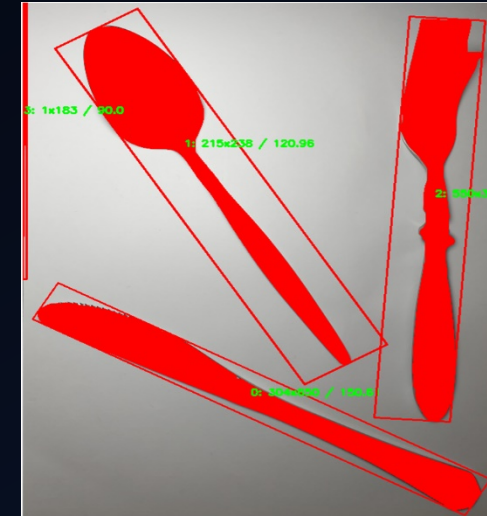
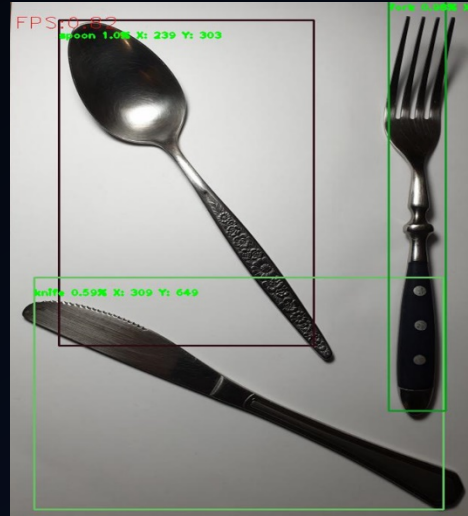
Όραση: Εντοπισμός και Αποκωδικοποίηση QR code

- Απλή χρήση και δημιουργία
- Εξατομικευμένη ονοματοδοσία
- Εντοπισμός οποιουδήποτε αντικειμένου
- Μικρή επιφάνεια κάλυψης
- Χαμηλό ποσοστό εσφαλμένης ανάγνωσης
- Δυνατότητα ανάγνωσης από οποιαδήποτε γωνία
- Εξαγωγή ετικέτας ονόματος και συντεταγμένες κέντρου QR code



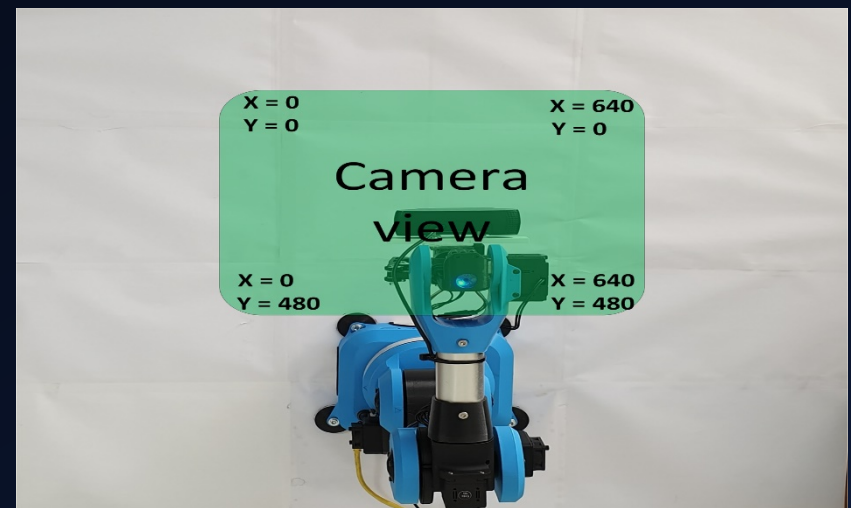
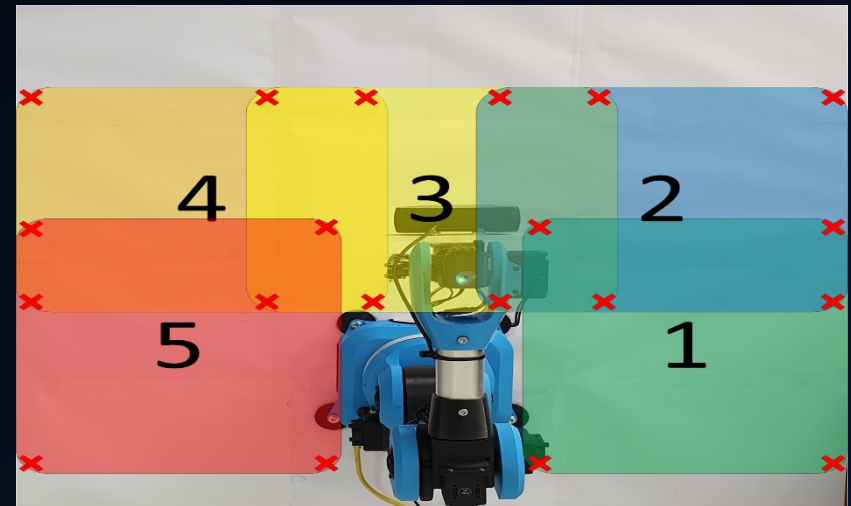
Όραση: Εντοπισμός περιγραμμάτων αντικειμένων

- Αναγκαία η ανάπτυξη του για τον εντοπισμό κλίσης γωνίας σε YOLO και QR code
- Εντοπισμός οποιοδήποτε αντικειμένου στο χώρο
- Εντοπισμός κέντρου και γωνίας



Μετατροπή συντεταγμένων

- Δημιουργία 5 θέσεων περιμετρικά για αναζήτηση
- Χειρισμός μέσω των τιμών x , y , z , roll, pitch, yaw



Μετρικά

Λειτουργία	Χρόνος σε sec	Σενάρια κώδικα	Γραμμές κώδικα	Μέγεθος (bytes)	Συναρτήσεις	Βιβλιοθήκες
Χρόνος διεκπεραίωσης λειτουργίας αναζήτησης αντικειμένων	15	Speech	538	22.777	13	11
Χρόνος διεκπεραίωσης λειτουργίας αναζήτησης χρήστη	3-15	Vision	553	24.576	12	11
Χρόνος προσέγγισης και αρπαγής	7	Control	332	21.675	9	6
Χρόνος παράδοσης	5	Σύνολο	1423	469.028	34	19

A dark blue background with a faint image of a hand holding a pen. The hand is positioned in the upper right quadrant, with the pen pointing downwards. The background has a subtle gradient and some abstract shapes. On the left side, there are several vertical lines of varying lengths, resembling a stylized 'E' or a decorative element.

Σενάρια & Εκτέλεση Πειραμάτων

Πειραματική Διαδικασία



- CPU: Intel Core i5-7200U dual core 2.5GHz (up to 3.1GHz)
- GPU: Nvidia GeForce MX150 2GB GDDR5
- Memory: 8GB DDR4
- Storage: 2x 256GB Pcle SSD
- Operating System: Windows 10 Pro 64-bit

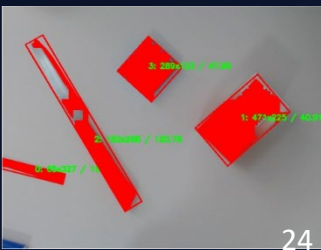
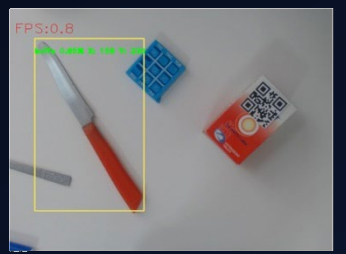
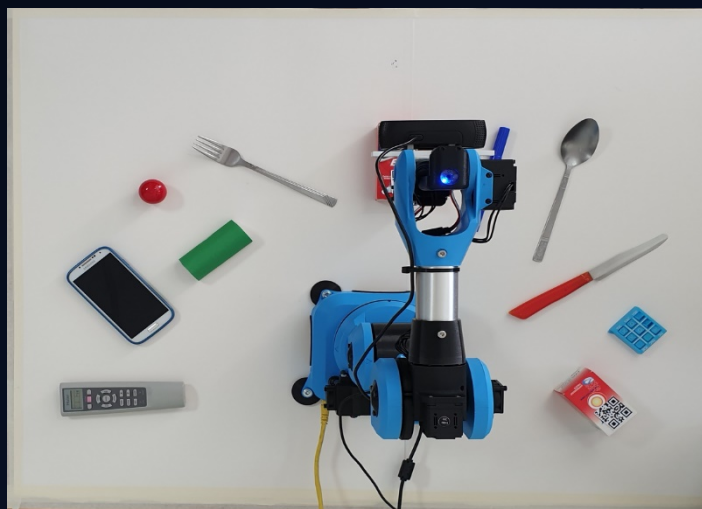
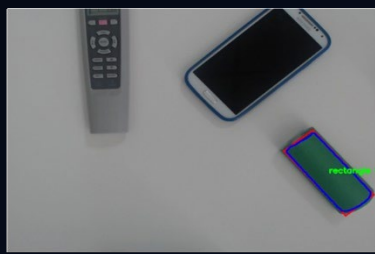
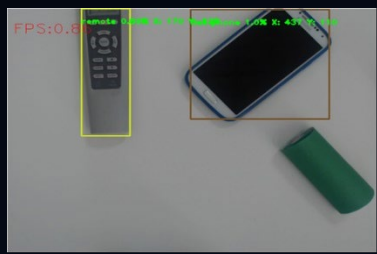
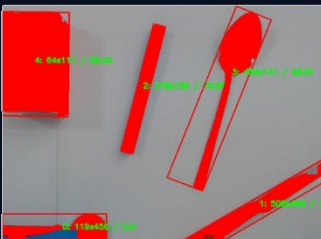
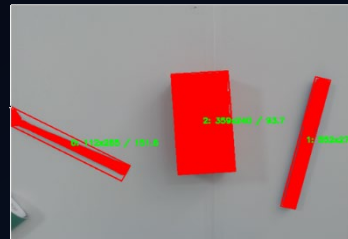
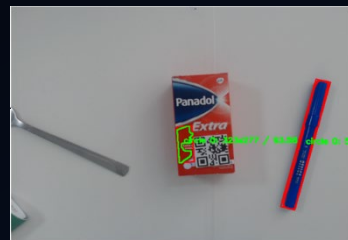
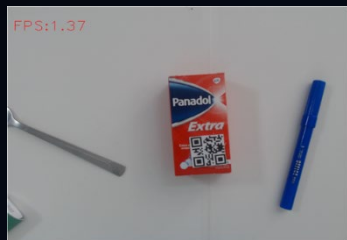
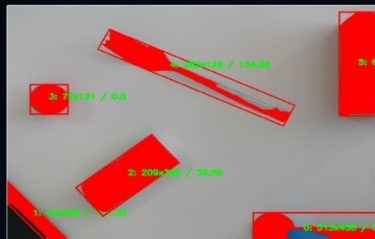
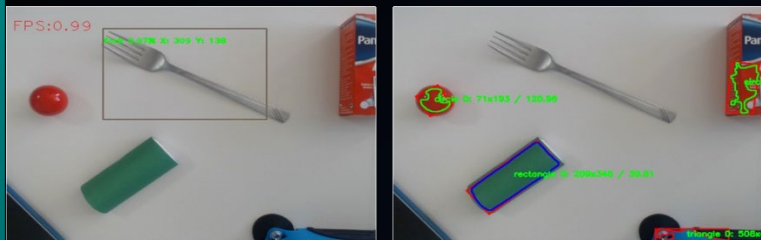


- SoC: Broadcom BCM2711B0 quadcore A72 (ARMv8-A) 64-bit @1.5GHz
- GPU: Broadcom VideoCore VI
- Networking: 2.4 GHz and 5 GHz 802.11b/g/n/ac wireless LAN
- RAM: 4GB LPDDR4 SDRAM
- Bluetooth: Bluetooth 5.0, Bluetooth Low Energy (BLE)
- Ports: 2 × micro-HDMI 2.0, 3.5 mm analogue audio-video jack, 2 × USB 2.0, 2 × USB 3.0, Gigabit Ethernet
- Operating System: Raspbian
- Dimensions: 88 mm × 58 mm × 19.5 mm, 46 g

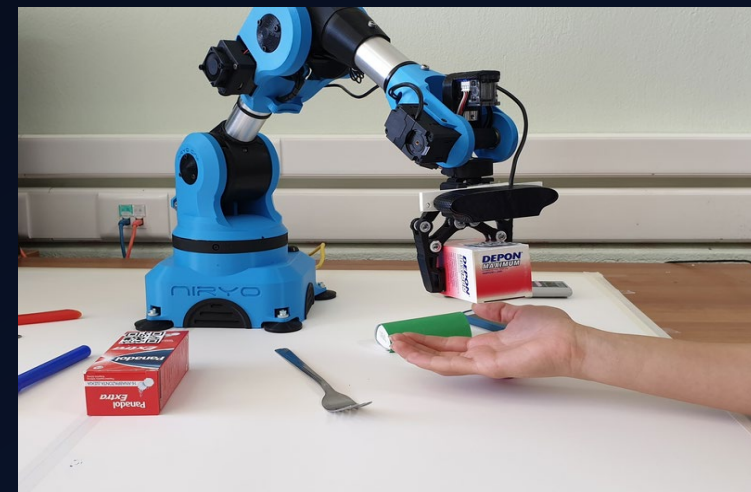
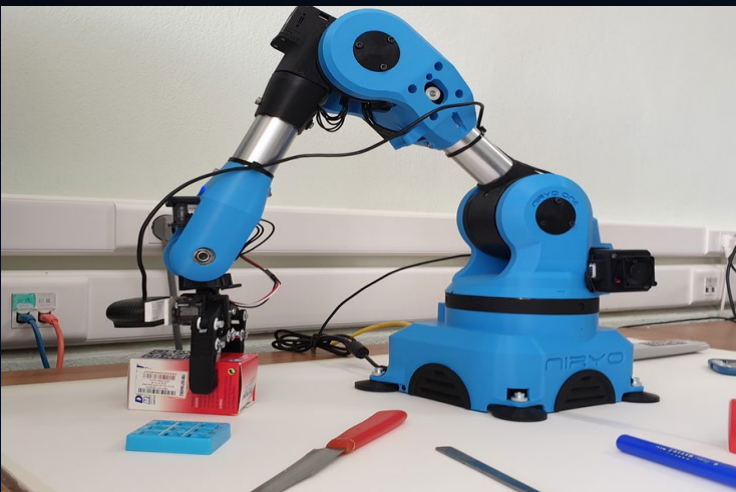
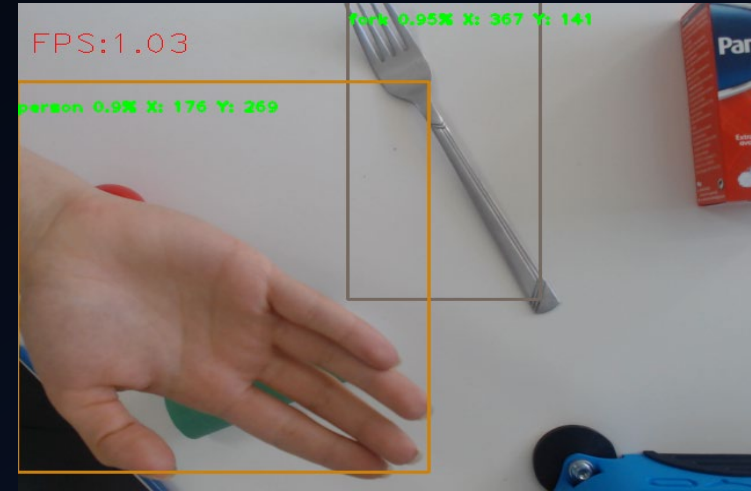
Σενάριο πειραμάτων

	Πείραμα 1	Πείραμα 2	Πείραμα 3	Πείραμα 4
Γλώσσα αλληλεπίδρασης	Ελληνικά	Αγγλικά	Ελληνικά	Αγγλικά
Φωνή χρήστη	Ανδρική	Ανδρική	Γυναικεία	Γυναικεία
Λειτουργία αναζήτησης	Ναι	Όχι	Όχι	Ναι
Τρόπος παράδοσης:	Χέρι χρήστη	Χέρι χρήστη	Προκαθορισμένο σημείο	Χέρι χρήστη
Λειτουργία αναγνώρισης αντικειμένου	Υolo	Υolo	Σχήμα & Χρώμα	QR code

Εκτέλεση πειράματος Μέρος Α




Εκτέλεση πειράματος Μέρος Β





Video επίδειξης λειτουργίας



Συμπεράσματα & Μελλοντικές Επεκτάσεις

Αποτελέσματα

	Φορητός Η/Υ				Raspberry Pi 4			
	Πείραμα 1	Πείραμα 2	Πείραμα 3	Πείραμα 4	Πείραμα 1	Πείραμα 2	Πείραμα 3	Πείραμα 4
Μέση τιμή fps κατά την εκτέλεση του Yolo	1,51	1,45	1,42	1,42	0,46	0,42	0,36	0,38
Αποτυχημένες προσπάθειες αναγνώριση φωνητικής εντολής	0	1	1	0	0	1	2	0
Αριθμός λανθασμένων ή μη αναγνωρισμένων αντικειμένων από σύνολο 11 αντικειμένων	1	0	1	0	2	2	3	2
Συνολικός χρόνος αλληλεπίδρασης	90	53	32	92	104	58	42	101

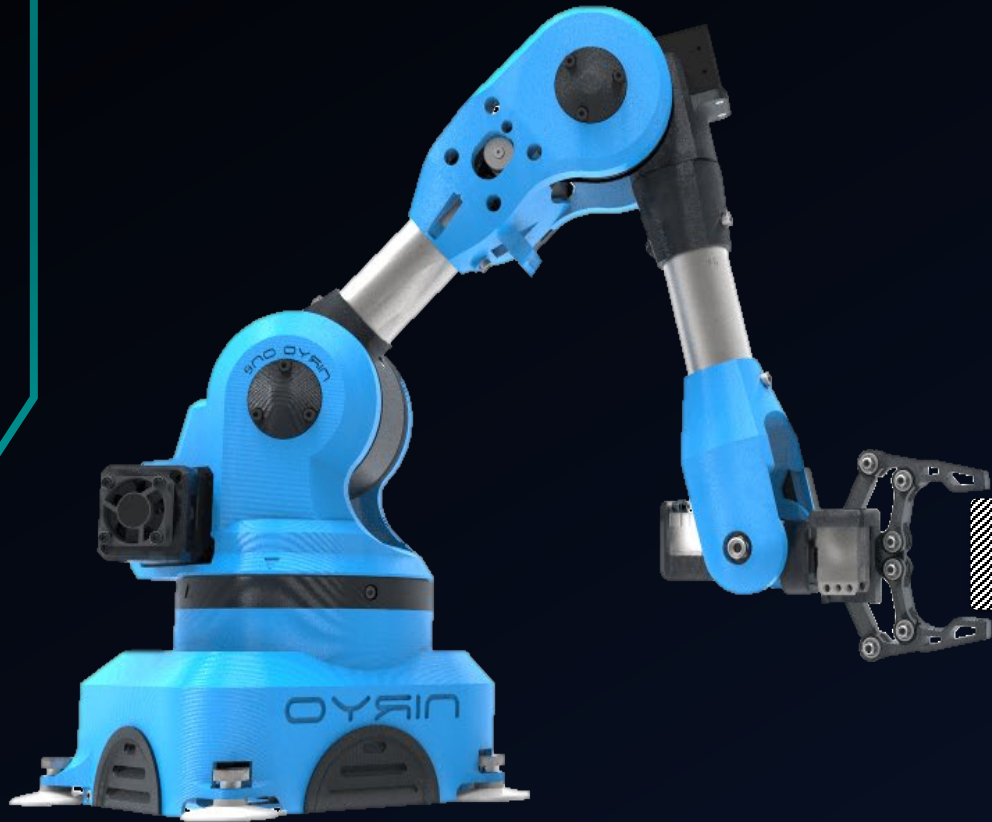
Συμπεράσματα

- Εύκολη και άμεση επικοινωνία με το χρήστη
- Ευρύ πλήθος αναγνώρισης αντικειμένων
- Δυνατότητα προσωπικής παραμετροποίησης
- Προσθήκη εύρεσης γωνίας σε αντικείμενα Yolo & Qr code
- Ομαλή λειτουργία στο Raspberry Pi 4
- Ποσοστό επιτυχημένης αναγνώρισης αντικειμένων
 - Φορητό υπολογιστή 95,45%
 - Raspberry Pi 4 79,54%.

Μελλοντικές Επεκτάσεις

- Προσθήκη πηγής φωτός LED στο τελικό εργαλείο δράσης
- Ενσωμάτωση μαγνητικού διακόπτη στο τελικό εργαλείο δράσης
- Αντικατάσταση κάμερας με στεροσκοπική ή προσθήκη ultrasonic αισθητήρα
- Εκτέλεση σε πλακέτα επεξεργασίας Nvidia Jetson Nano
- Ανάπτυξη εφαρμογής για smartphone

ΕΥΧΑΡΙΣΤΩ



ΕΡΩΤΗΣΕΙΣ ;

Αρχιτεκτονική συστήματος

