



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΙΩΑΝΝΙΝΩΝ

ΣΧΟΛΗ ΟΙΚΟΝΟΜΙΚΩΝ ΚΑΙ ΔΙΟΙΚΗΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

**ΥΛΟΠΟΙΗΣΗ ΑΝΙΧΝΕΥΣΗΣ ΚΑΙ ΑΝΑΓΝΩΡΙΣΗΣ ΠΡΟΣΩΠΟΥ ΜΕ
ΤΗΝ ΧΡΗΣΗ RASPBERRY PI**

Παπαγεωργίου Ελευθέριος Α.Μ. 1413

Επιβλέπων: Καρβέλης Πέτρος

Επίκουρος Καθηγητής

ΑΡΤΑ, 2024

IMPLEMENTATION OF FACE DETECTION AND RECOGNITION USING RASPBERRY PI

Εγκρίθηκε από τριμελή εξεταστική επιτροπή

Αρτα 2024

ΕΠΙΤΡΟΠΗ ΑΞΙΟΛΟΓΗΣΗΣ

1. Επιβλέπων καθηγητής

Πέτρος Καρβέλης

2. Μέλος επιτροπής

Γιαννακέας Νικόλαος

3. Μέλος επιτροπής

Τζάλλας Αλέξανδρος

© Παπαγεωργίου Ελευθέριος 2023.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Δήλωση μη λογοκλοπής

Δηλώνω υπεύθυνα και γνωρίζοντας τις κυρώσεις του Ν. 2121/1993 περί Πνευματικής Ιδιοκτησίας, ότι η παρούσα μεταπτυχιακή εργασία είναι εξ ολοκλήρου αποτέλεσμα δικής μου ερευνητικής εργασίας, δεν αποτελεί προϊόν αντιγραφής ούτε προέρχεται από ανάθεση σε τρίτους. Όλες οι πηγές που χρησιμοποιήθηκαν (κάθε είδους, μορφής και προέλευσης) για τη συγγραφή της περιλαμβάνονται στη βιβλιογραφία.

Παπαγεωργίου Ελευθέριος

Υπογραφή

ΠΕΡΙΛΗΨΗ

Στην παρούσα πτυχιακή εργασία θα αναφερθούμε στο Raspberry Pi, την μηχανική όραση, τα συστήματα μηχανικής μάθησης καθώς και θα δημιουργήσουμε μια εφαρμογή η οποία αξιοποιώντας την υπολογιστική όραση μέσω μιας κάμερας και την χρήση τεχνητή νοημοσύνη θα καταφέρει να εντοπίσει και να αναγνωρίσει ανθρώπινα πρόσωπα. Για την υλοποίηση του έργου θα χρησιμοποιήσουμε το Raspberry Pi 4 Model B. Η εφαρμογή με όνομα “Face Recognition” έχει ως σκοπό τον εντοπισμό και την αναγνώριση προσώπων σε πραγματικό χρόνο που έχουν καταχωρηθεί στο σύστημα. Με την χρήση της Υπολογιστικής Όρασης και τεχνολογίες αναγνώρισης προσώπων, θα καταφέρουμε σε πραγματικό χρόνο να εντοπίσουμε και αναγνωρίσουμε το πρόσωπο που βρίσκεται στο οπτικό πεδίο της κάμερας. Για την εκπαίδευση του μοντέλου θα χρειαστούμε ένα δείγμα φωτογραφιών με αρκετές γωνίες του προσώπου, οι οποίες θα δοθούν ως είσοδος για την εκπαίδευση του μοντέλου. Για την αναγνώριση των προσώπων θα χρησιμοποιήσουμε ένα προεκπαιδευμένο μοντέλο τύπου Haar Cascade και HOG (Histogram of Oriented Gradients)

Λέξεις-κλειδιά: Raspberry Pi, Face Recognition, Face Detection, Haar Cascade

ABSTRACT

In this thesis we will refer to the Raspberry Pi, machine vision, machine learning systems and we will create an application that utilizes computer vision through a camera and the use of artificial intelligence to identify and recognize human faces. For the implementation of the project we will use Raspberry Pi 3 Model B. The application named "Face Recognition" aims to detect and recognize faces in real time that have been registered in the system. By using Computer Vision and face recognition technologies, we will be able to detect and recognize in real time the face that is in the field of view of the camera. To train the model we will need a sample of photographs with several angles of the face, which will be provided as input for training the model. For face recognition we will use a pre-trained Haar Cascade and HOG (Histogram of Oriented Gradients)

Keywords: Raspberry Pi, Face Recognition, Face Detection, Haar Cascade

ΠΕΡΙΕΧΟΜΕΝΑ

1.	Εισαγωγή στο Raspberry Pi.....	8
1.1	Περιγραφή Raspberry Pi.....	8
1.2	Ιστορική αναδρομή.....	9
1.3	Εκδόσεις.....	10
1.4	Λειτουργικά συστήματα.....	13
2.	Μηχανική όραση.....	15
2.1	Περιγραφή μηχανικής όρασης.....	15
2.2	Ιστορική αναδρομή.....	16
2.3	Μηχανική όραση και τεχνητή νοημοσύνη.....	17
3.	Συστήματα μηχανικής μάθησης.....	19
3.1	Λειτουργίες.....	19
3.2	Τεχνικές εντοπισμού αντικειμένων-προσώπου.....	20
3.3	Τεχνικές ταυτοποίησης προσώπων.....	21
4.	Εργαλεία ανάπτυξης.....	22
4.1	Περιγραφή OpenCV.....	23
4.2	Περιγραφή Python.....	25
4.3	Περιγραφή και εφαρμογές Haar Cascade.....	27
5.	Περιγραφή αλγορίθμου εκπαίδευσης.....	29
5.1	Εφαρμογή του αλγορίθμου HOG.....	30
5.2	Υλοποίηση συστήματος.....	32

ΕΙΣΑΓΩΓΗ

Ο εντοπισμός και η αναγνώριση προσώπου αποτελούν σημαντικές εφαρμογές της τεχνητής νοημοσύνης καθώς έχουν ευρύ φάσμα χρήσεων, όπως η ασφάλεια, αναγνώριση προσώπων για πρόσβαση σε συστήματα και περιβάλλοντα, αυτόματη ταξινόμηση φωτογραφιών και πολλά άλλα. Το Raspberry Pi ως ένας μικροϋπολογιστής παρέχει ένα ισχυρό και προσιτό εργαλείο για την ανάπτυξη συστημάτων εντοπισμού και αναγνώρισης προσώπου. Με την σωστή χρήση των αισθητήρων του και την εφαρμογή αλγορίθμων επεξεργασίας εικόνας, το Raspberry Pi μπορεί να ανιχνεύσει πρόσωπα σε πραγματικό χρόνο και να τα αναγνωρίσει με ακρίβεια. Στην παρούσα εργασία, εξετάζουμε τη χρήση του Raspberry Pi για την υλοποίηση ενός συστήματος εντοπισμού και αναγνώρισης προσώπου, παρέχοντας μια επισκόπηση των απαιτούμενων εργαλείων, τεχνικών και προκλήσεων που συναντάμε σε αυτό το πεδίο.

Κεφάλαιο 1: Εισαγωγή στο Raspberry Pi

1.1 Περιγραφή Raspberry Pi

Το Raspberry Pi είναι μια σειρά υπολογιστών μονής πλακέτας που ξεκίνησε από το Raspberry Pi Foundation το 2012. Ο σκοπός του ήταν να παρέχει μια προσιτή, εκπαιδευτική και πειραματική πλατφόρμα που να ενθαρρύνει τον προγραμματισμό και την εξοικείωση με την τεχνολογία. Από την αρχική του έκδοση έχει εξελιχθεί σε διάφορα μοντέλα με διαφορετικές δυνατότητες και προδιαγραφές.



Εικόνα 1.1 Λογότυπο Raspberry Pi

Κάθε μοντέλο του Raspberry Pi περιλαμβάνει μια μικρή πλακέτα που περιέχει έναν επεξεργαστή, μνήμη RAM, θύρες επικοινωνίας, όπως USB και Ethernet, θύρα για σύνδεση με οθόνη, θύρες GPIO για σύνδεση με αισθητήρες και περιφερειακά, και μια θύρα κάρτας SD για την αποθήκευση του λειτουργικού συστήματος και των αρχείων. Οι περιφερειακές θύρες GPIO είναι ιδιαίτερα σημαντικές, καθώς επιτρέπουν τη σύνδεση με εξωτερικές συσκευές και την επέκταση των δυνατοτήτων του Raspberry Pi. Ένα από τα μεγάλα πλεονεκτήματα του Raspberry Pi είναι η ευελιξία του λειτουργικού συστήματος που μπορεί να υποστηρίξει. Οι περισσότεροι χρήστες επιλέγουν την χρήση του Raspbian, μια έκδοση του Linux που έχει σχεδιαστεί ειδικά για το Raspberry Pi. Ωστόσο, υπάρχουν επίσης άλλα λειτουργικά συστήματα που μπορούν να χρησιμοποιηθούν, όπως το Ubuntu, το Windows 10 IoT Core και πολλά άλλα.

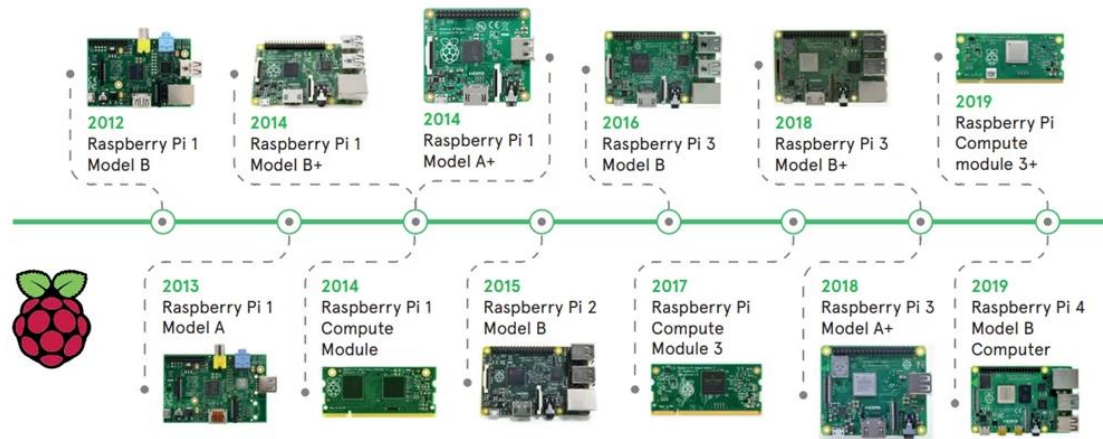
Με την χρήση του Raspberry Pi, μπορούν να αναπτυχθούν διάφορες εφαρμογές και έργα. Από την ανάπτυξη προγραμμάτων και τη δημιουργία δικτυακών εφαρμογών μέχρι την κατασκευή ρομπότ, το Raspberry Pi προσφέρει ένα ευέλικτο περιβάλλον για την εξερεύνηση και την καινοτομία. Δίνει την δυνατότητα να αναπτυχθεί συστήματα αυτοματισμού για το σπίτι, επιτρέπει την διασύνδεση με αισθητήρες που μετρούν την θερμοκρασία, την υγρασία και την κίνηση ώστε να δημιουργηθούν συστήματα ελέγχου φωτισμού και πολλά άλλα. Το Raspberry Pi διαθέτει επίσης μια ζωντανή και ενεργή κοινότητα χρηστών και προγραμματιστών όπου υπάρχουν πληθώρα πόρων, παραδειγμάτων και προγραμμάτων που μπορούν να αξιοποιηθούν, καθώς και φόρουμ υποστήριξης και να ανταλλαγής ιδεών με άλλους χρήστες. Έχει αναδειχθεί ως ένα ισχυρό εργαλείο για την εκμάθηση και την εξερεύνηση της τεχνολογίας. Είναι κατάλληλο για εκπαιδευτικούς, φοιτητές, ερευνητές και ερασιτέχνες που επιθυμούν να εξασκηθούν στον προγραμματισμό, την ηλεκτρονική και την κατασκευή πρωτοτύπων.

Συνοψίζοντας, το Raspberry Pi είναι μια εξαιρετικά ευέλικτη πλατφόρμα υπολογιστή που συνδυάζει την προσιτή τιμή με τις εξελισσόμενες δυνατότητες. Είναι ιδανικό για προγραμματισμό, πειραματισμό και δημιουργική εξέλιξη σε ποικίλους τομείς της τεχνολογίας. [1]

1.2 Ιστορική αναδρομή

Η ιστορία του Raspberry Pi ξεκίνησε το 2006, όταν η εταιρεία Cambridge Computer Laboratory του Πανεπιστημίου του Cambridge συνειδητοποίησε ότι οι νέοι φοιτητές που εισήχθησαν στο πανεπιστήμιο δεν είχαν την απαραίτητη γνώση και εμπειρία στον τομέα της πληροφορικής και του προγραμματισμού. Επίσης, παρατηρήθηκε ότι οι υπάρχουσες πλατφόρμες εκπαίδευσης ήταν ακριβές και περιορισμένες σε λειτουργίες. Αυτή η ανάγκη για μια προσιτή και ευέλικτη πλατφόρμα οδήγησε στη δημιουργία του Raspberry Pi.

Η πρώτη πραγματική αναφορά έγινε το 2008, όταν ο Eben Upton, ένας από τους ιδρυτές της ομάδας, εκδήλωσε την ανησυχία του για την μείωση της απήχησης των φοιτητών στην πληροφορική. Το 2009, η ομάδα του Raspberry Pi ξεκίνησε την ανάπτυξη της πλατφόρμας με στόχο να δημιουργήσει ένα μικρό, οικονομικό και προγραμματιζόμενο υπολογιστή που θα ήταν προσιτός για εκπαιδευτικούς σκοπούς. Στις 29 Φεβρουαρίου 2012, το πρώτο μοντέλο, το Raspberry Pi Model B, κυκλοφόρησε στην αγορά. Η ανταπόκριση ήταν εντυπωσιακή και η πλατφόρμα απέκτησε γρήγορα μεγάλη δημοτικότητα. Το Raspberry Pi Model B ήταν ένας μικρός υπολογιστής με μια επεξεργαστική μονάδα ARM, μνήμη RAM, θύρες USB, έξοδο HDMI και άλλες συνδέσεις, παρέχοντας τη δυνατότητα για εγκατάσταση λειτουργικού συστήματος Linux για προγραμματισμό. Η τιμή του ήταν πολύ προσιτή, περίπου 35 δολάρια ΗΠΑ, κάτι που το έκανε προσιτό για εκπαιδευτικούς οργανισμούς και ατομικούς χρήστες. Η επιτυχία του πρώτου μοντέλου οδήγησε στην κυκλοφορία πολλών επόμενων εκδόσεων του Raspberry Pi, κάθε μια με βελτιωμένες προδιαγραφές και νέες λειτουργίες. Οι εκδόσεις που ακολούθησαν περιλάμβαναν το Raspberry Pi Model A, το Raspberry Pi Model B+, το Raspberry Pi 2 Model B, το Raspberry Pi 3 Model B, το Raspberry Pi Zero και το Raspberry Pi 4 Model B.



Εικόνα 1.2 Ιστορική Αναδρομή Raspberry Pi

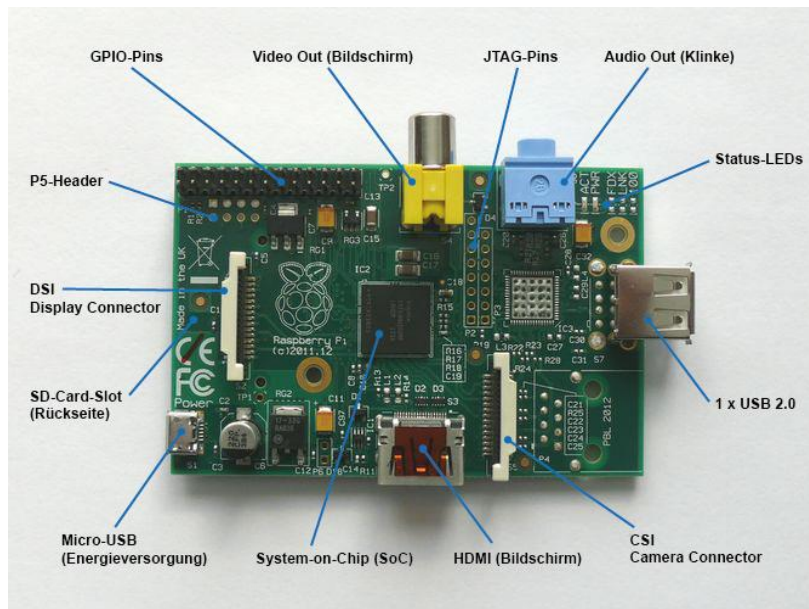
Καθώς περνούσαν τα χρόνια, η πλατφόρμα εξελίσσονταν και αποκτούσε όλο και περισσότερες δυνατότητες. Οι βελτιώσεις σε επεξεργαστές, μνήμη, γραφικά και συνδεσιμότητα δίνουν στους χρήστες τη δυνατότητα να χρησιμοποιήσουν το Raspberry Pi για ευρύ φάσμα εφαρμογών, από εκπαιδευτικούς σκοπούς έως προγραμματισμό, ρομποτική, δημιουργία εφαρμογών IoT και πολλά άλλα.

[2]

1.3 Οι εκδόσεις του Raspberry Pi

Το Raspberry Pi έχει γνωρίσει αρκετές εκδόσεις κατά τη διάρκεια της εξέλιξής του. Ας εξετάσουμε λεπτομερέστερα μερικές από τις πιο σημαντικές εκδόσεις του Raspberry Pi και τις αναβαθμίσεις που προσέφεραν:

Raspberry Pi 1 Model A/B: Η πρώτη έκδοση του Raspberry Pi που κυκλοφόρησε το 2012. Βασίζεται σε έναν επεξεργαστή ARM 11 στα 700 MHz, διαθέτει μόλις 256MB μνήμη RAM, θύρες USB 2.0 και υποδοχή κάρτας SD για αποθήκευση του λειτουργικού συστήματος. . Αυτή η έκδοση ήταν ιδανική για αρχάριους χρήστες και για εκπαιδευτικούς σκοπούς.



Εικόνα 1.3 Raspberry Pi 1 Model A/B

Raspberry Pi 1 Model A+/B+: Αυτές οι εκδόσεις κυκλοφόρησαν το 2014 και προσέφεραν μια σειρά από βελτιώσεις σε σχέση με τις αρχικές εκδόσεις. Η μνήμη RAM αυξήθηκε σε 512MB, διέθεταν περισσότερες θύρες GPIO, παρέχοντας περισσότερες επιλογές συνδεσιμότητας όπως 4 USB ports, Ethernet port, 3.5mm jack for audio & composite video out καθώς και HDMI.



Εικόνα 1.4 Raspberry Pi 1 Model A+/B+

Raspberry Pi 2 Model B: Η συγκεκριμένη έκδοση κυκλοφόρησε το 2015 και αντιπροσώπευε μια σημαντική αναβάθμιση. Διέθετε τετραπύρηνο επεξεργαστή ARM Cortex-A7 στα 900 MHz και 1GB μνήμη RAM. Η απόδοση αυξήθηκε σημαντικά σε σχέση με τις προηγούμενες εκδόσεις, επιτρέποντας την χρήση για πιο απαιτητικές εφαρμογές και εργαλεία ανάπτυξης.



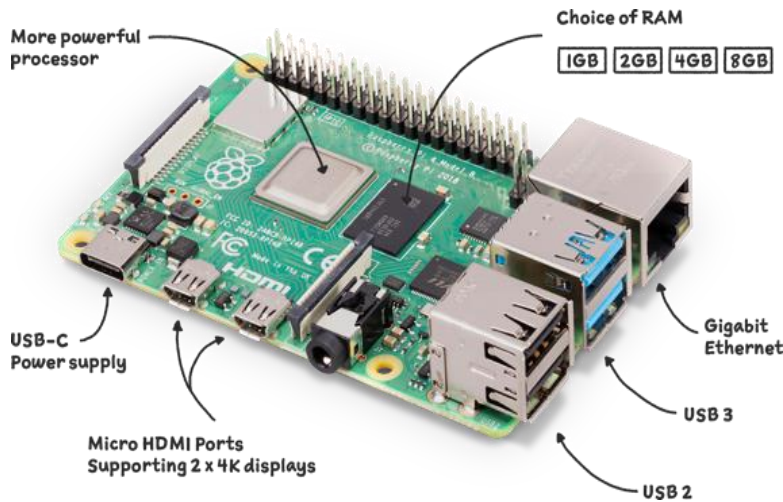
Εικόνα 1.5 Raspberry Pi 2 Model B

Raspberry Pi 3 Model B/B+: Οι εκδόσεις αυτές παρουσιάστηκαν το 2016 και το 2018 αντίστοιχα. Ενσωμάτωσαν τετραπύρηνο επεξεργαστή ARM Cortex-A53 στα 1.2 GHz και διέθεταν 1GB ή 2GB μνήμη RAM. Η σημαντική βελτίωση ήταν η ενσωματωμένη υποστήριξη Wi-Fi και Bluetooth, η οποία επέτρεπε ευκολότερη σύνδεση και επικοινωνία με άλλες συσκευές.



Εικόνα 1.6 Raspberry Pi 3 Model B/B+

Raspberry Pi 4 Model B: Η πιο πρόσφατη έκδοση που κυκλοφόρησε το 2019. Διαθέτει τετραπύρηνο επεξεργαστή ARM Cortex-A72 στα 1.5 GHz επιλογή για 2,4 ή 8 GB μνήμης RAM, προσφέροντας ακόμα μεγαλύτερη απόδοση. Επιπλέον, προστέθηκε υποστήριξη για Gigabit Ethernet, USB 3.0 θύρες και υποστήριξη διπλής οθόνης 4K.



Εικόνα 1.7 Raspberry Pi 4 Model B

Οι εκδόσεις του Raspberry Pi αντικατοπτρίζουν τη συνεχή ανάπτυξη της πλατφόρμας και την επιθυμία για νέες και καλύτερες λειτουργίες. Αυτές οι εκδόσεις παρέχουν μεγαλύτερη απόδοση, αυξημένη συνδεσιμότητα και περισσότερες δυνατότητες για εφαρμογές και έργα. Επιπλέον, η ανοιχτή αρχιτεκτονική του Raspberry Pi επιτρέπει στους χρήστες να προσαρμόσουν και να επεκτείνουν το σύστημα με περιφερειακά και αισθητήρες, επιτρέποντας την πραγματοποίηση διαφόρων εφαρμογών σε πολλούς τομείς της τεχνολογίας.

[3]

1.4 Λειτουργικά συστήματα

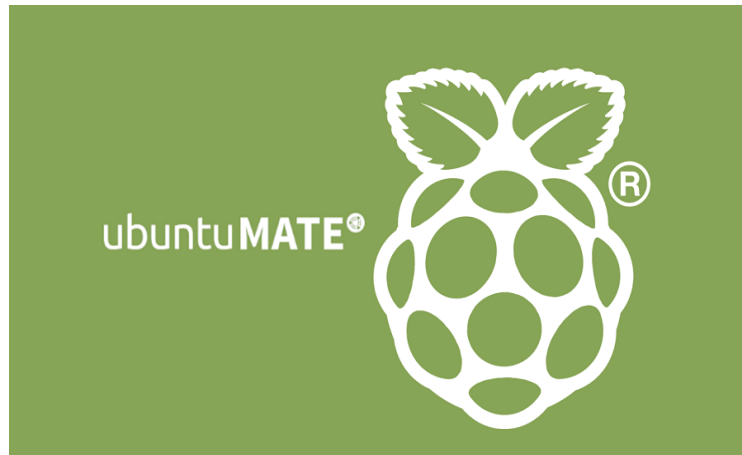
Τα λειτουργικά συστήματα που εκτελούνται στο Raspberry Pi παίζουν σημαντικό ρόλο στην επίτευξη της πλήρους δυναμικότητας αυτού του μικροϋπολογιστή. Λόγω της ευελιξίας του οι χρήστες έχουν τη δυνατότητα να εγκαταστήσουν διάφορα λειτουργικά συστήματα ανάλογα με τις ανάγκες και τις προτιμήσεις τους. Ανεξάρτητα από το αν ο χρήστης θέλει να δημιουργήσει ένα περιβάλλον εργασίας, ένα media center ή μια εφαρμογή IoT, υπάρχει ένα λειτουργικό σύστημα που καλύπτει τις ανάγκες του καθενός. Σε αυτήν την ενότητα θα εξετάσουμε μερικά από τα πιο δημοφιλή λειτουργικά συστήματα που χρησιμοποιούνται στο Raspberry Pi, καθώς και τις δυνατότητες τους.

Raspbian:

Το Raspbian είναι ένα από τα πιο δημοφιλή λειτουργικά συστήματα που χρησιμοποιούνται στο Raspberry Pi. Βασίζεται στο Debian Linux και έχει σχεδιαστεί ειδικά για να εκτελείται στο Raspberry Pi. Είναι πλήρως βελτιστοποιημένο για τους πόρους του Raspberry Pi, καθιστώντας το γρήγορο και αποδοτικό. Έχει προεγκατεστημένο περιβάλλον επιφάνειας εργασίας PIXEL (Pi Improved X-Window Environment, Lightweight), το οποίο παρέχει μια ευχρηστία περιβάλλοντος γραφικής διεπαφής για τους χρήστες. Το Raspbian προσφέρει επίσης πρόσβαση σε ένα ευρύ φάσμα πακέτων και εφαρμογών που υποστηρίζονται από το Debian, επιτρέποντας στους χρήστες να προσαρμόσουν το Raspberry Pi στις ανάγκες τους.

Ubuntu MATE:

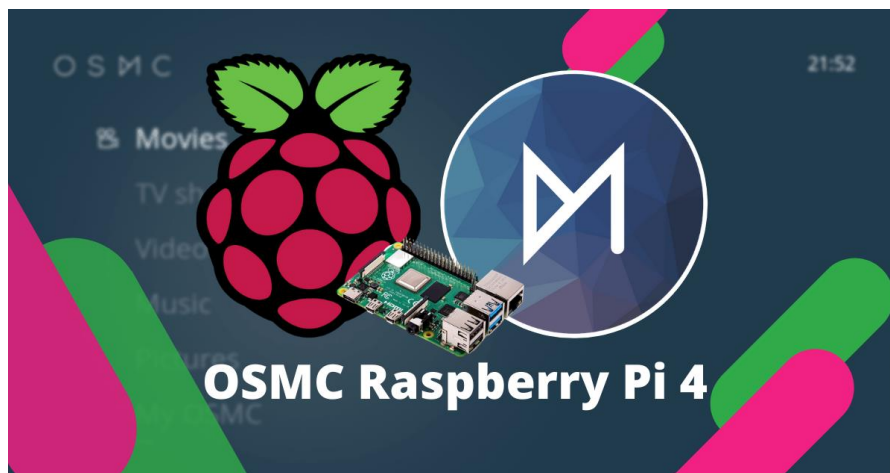
Το Ubuntu MATE είναι μια επίσημη έκδοση του Ubuntu Linux που έχει προσαρμοστεί για να λειτουργεί στο Raspberry Pi. Το Ubuntu MATE προσφέρει μια πλήρη και φιλική προς τον χρήστη επιφάνεια εργασίας, που είναι παρόμοια με το γνωστό περιβάλλον εργασίας Gnome 2. Παρέχει ένα πλήθος εφαρμογών και εργαλείων για τους χρήστες, επιτρέποντας τους να επεκτείνουν τις δυνατότητες του Raspberry Pi. Οι χρήστες μπορούν να χρησιμοποιήσουν το Ubuntu MATE ώστε να αναπτύξουν εφαρμογές, να δημιουργήσουν media centers ή ακόμη και να εκτελέσουν περιβάλλοντα εργασίας γραφείου.



Εικόνα 1.8 Λογότυπο λειτουργικού Ubuntu Mate

OSMC:

Το OSMC είναι μια έκδοση που βασίζεται στο Kodi (πρώην XBMC) και αποτελεί ένα πλήρες λειτουργικό σύστημα για πολυμέσα, σχεδιασμένο ειδικά για το Raspberry Pi. Με το OSMC, μπορείτε να μετατρέψετε το Raspberry Pi σας σε ένα πλήρες κέντρο πολυμέσων που μπορεί να αναπαράγει ποικίλα αρχεία πολυμέσων, όπως βίντεο, μουσική και φωτογραφίες. Είναι εύκολο στη χρήση και προσφέρει πολλές επιλογές προσαρμογής για τους χρήστες που θέλουν να δημιουργήσουν το δικό τους καταναμημένο σύστημα πολυμέσων.

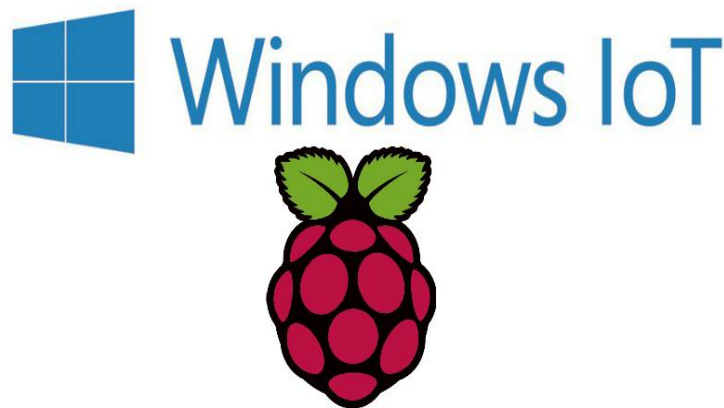


Εικόνα 1.9 Λειτουργικό OSMC

Windows 10 IoT Core:

Το Windows 10 IoT Core είναι μια έκδοση του λειτουργικού συστήματος Windows 10 που είναι σχεδιασμένη για το Internet of Things (IoT). Υποστηρίζει το Raspberry Pi και παρέχει ένα ευέλικτο περιβάλλον για ανάπτυξη εφαρμογών IoT. Οι χρήστες μπορούν να χρησιμοποιήσουν το Visual Studio και τη γλώσσα προγραμματισμού C# για να δημιουργήσουν εφαρμογές που αλληλεπιδρούν με αισθητήρες, μονάδες ελέγχου και άλλες συσκευές, αξιοποιώντας τις δυνατότητες του Raspberry Pi στο πλαίσιο του IoT.

[4]



Εικόνα 1.10 Λειτουργικό Windows IoT

Κεφάλαιο 2 Μηχανική Όραση

2.1 Η διαδικασία της μηχανικής όρασης περιλαμβάνει την ανάλυση, την επεξεργασία και την ερμηνεία των εικόνων ή των βίντεο. Αυτό περιλαμβάνει την αναγνώριση αντικειμένων, τον εντοπισμό προσώπων, την αναγνώριση χαρακτηριστικών, την κατηγοριοποίηση εικόνων και πολλές άλλες εργασίες. Για να επιτευχθεί αυτό, η μηχανική όραση χρησιμοποιεί διάφορες τεχνικές και αλγορίθμους, όπως:

Εξαγωγή χαρακτηριστικών: Πρόκειται για τη μετατροπή μιας εικόνας σε μια αναπαράσταση χαρακτηριστικών, που μπορεί να περιλαμβάνει πληροφορίες για τις γωνίες, τις ακμές, τα χρώματα και άλλα στοιχεία.

Ανίχνευση αντικειμένων: Περιλαμβάνει την ανίχνευση και τον εντοπισμό των αντικειμένων σε μια εικόνα ή ένα βίντεο. Αυτό μπορεί να γίνει μέσω της χρήσης αλγορίθμων όπως το Haar cascades ή το YOLO (You Only Look Once).

Αναγνώριση προσώπων: Περιλαμβάνει την αναγνώριση και τον εντοπισμό προσώπων σε εικόνες ή βίντεο. Συχνά χρησιμοποιούνται μέθοδοι όπως οι επεκτάσεις του Eigenfaces, οι Local Binary Patterns (LBP) ή οι Convolutional Neural Networks (CNNs).

Σημασιολογική συμπερασματολογία: Αφορά την κατανόηση του περιεχομένου των εικόνων και την ερμηνεία τους σε έννοιες και σημασίες. Αυτό μπορεί να γίνει μέσω της εφαρμογής τεχνικών όπως το image captioning ή τα αναπαραστατικά μοντέλα γνώσης.

Η μηχανική όραση έχει εφαρμογές σε πολλούς τομείς, όπως η αυτόνομη οδήγηση, η ασφάλεια, η ιατρική διάγνωση, η ρομποτική, η αναγνώριση χειρονομιών, η επεξεργασία εικόνας και βίντεο, και πολλές άλλες.



Εικόνα 2.1

2.2. Ιστορική αναδρομή της Μηχανικής Όρασης

Η μηχανική όραση αποτελεί έναν σημαντικό κλάδο της τεχνητής νοημοσύνης που ασχολείται με την ανάπτυξη συστημάτων που μπορούν να επεξεργάζονται και να κατανοούν εικόνες και βίντεο όπως το ανθρώπινο μάτι. Η ιστορία της μηχανικής όρασης ξεκινά από την αρχαιότητα, όπου οι άνθρωποι αναγνώριζαν και ερμήνευαν τις εικόνες για να κατανοήσουν τον κόσμο γύρω τους. Οι πρώτες προσπάθειες στη μηχανική όραση χρονολογούνται στα μέσα του 20ο αιώνα, όταν οι επιστήμονες άρχισαν να αναπτύσσουν αλγορίθμους και τεχνικές για την αυτόματη αναγνώριση και επεξεργασία εικόνων. Από τη δεκαετία του 1960 και μετά, η μηχανική όραση απέκτησε ολοένα και μεγαλύτερη σημασία και προσέελκυσε το ενδιαφέρον της επιστημονικής κοινότητας. Μία από τις πρώτες επιτυχίες στον τομέα της μηχανικής όρασης ήταν η ανάπτυξη του αλγορίθμου της ανίχνευσης ακμών από τον John Canny το 1986. Ο αλγόριθμος του Canny αποτελεί ακόμα έναν από τους πιο γνωστούς αλγορίθμους ανίχνευσης ακμών που χρησιμοποιούνται στην επεξεργασία εικόνας. Την τελευταία δεκαετία, με την αύξηση της υπολογιστικής ισχύος και την ανάπτυξη τεχνολογιών όπως η ευρεία χρήση των νευρωνικών δικτύων, η μηχανική όραση έχει σημειώσει σημαντική πρόοδο. Παραδείγματα εφαρμογών της μηχανικής όρασης περιλαμβάνουν την αναγνώριση προσώπων, την αυτόματη οδήγηση οχημάτων και την ανίχνευση αντικειμένων σε εικόνες και βίντεο.

[5]

2.3 Η σχέση μεταξύ Μηχανικής Όρασης και Τεχνητής Νοημοσύνης

Η Μηχανική Όραση και η Τεχνητή Νοημοσύνη αποτελούν δύο στενά συνδεδεμένους τομείς της επιστήμης των υπολογιστών που επιδιώκουν να αποδώσουν στις μηχανές την ικανότητα να αντιλαμβάνονται, να κατανοούν και να αλληλεπιδρούν με τον κόσμο όπως οι άνθρωποι. Η Μηχανική Όραση ασχολείται με την ανάπτυξη αλγορίθμων και τεχνικών για την αναγνώριση και επεξεργασία εικόνων και βίντεο, ενώ η Τεχνητή Νοημοσύνη ασχολείται με την ανάπτυξη ευφυών συστημάτων που μπορούν να λαμβάνουν αποφάσεις, να μάθουν και να εκτελούν εργασίες που απαιτούν ανθρώπινη νοημοσύνη. Η σχέση μεταξύ της Μηχανικής Όρασης και της Τεχνητής Νοημοσύνης είναι στενή και αλληλεπιδραστική. Η Μηχανική Όραση παρέχει τη βάση για την αντίληψη των εικόνων και των βίντεο από μηχανές, ενώ η Τεχνητή Νοημοσύνη εκμεταλλεύεται την ανάλυση των εικόνων για να πραγματοποιήσει πιο πολύπλοκες λειτουργίες. Η Μηχανική Όραση παρέχει τα εργαλεία και τις μεθόδους για την αναγνώριση αντικειμένων, την ανίχνευση χαρακτηριστικών και την ανάλυση της δομής των εικόνων, ενώ η Τεχνητή Νοημοσύνη αξιοποιεί αυτές τις πληροφορίες για τη λήψη αποφάσεων, την αυτόματη εκμάθηση και την αλληλεπίδραση με το περιβάλλον.



Εικόνα 2.2 Διασύνδεση μηχανικής όρασης και τεχνητής νοημοσύνης

Ένας σημαντικός τομέας που συνδέει τη Μηχανική Όραση με την Τεχνητή Νοημοσύνη είναι η αναγνώριση αντικειμένων. Οι αλγόριθμοι της Μηχανικής Όρασης μπορούν να αναγνωρίσουν αντικείμενα σε εικόνες και βίντεο με βάση τα χαρακτηριστικά τους, όπως το σχήμα, το χρώμα και η δομή. Η Τεχνητή Νοημοσύνη μπορεί να χρησιμοποιήσει αυτή την πληροφορία για να αναλύσει το περιεχόμενο των εικόνων και να κατανοήσει το περιβάλλον τους. Παραδείγματα εφαρμογών που συνδυάζουν τη Μηχανική Όραση και τη Τεχνητή Νοημοσύνη περιλαμβάνουν την αυτόνομη οδήγηση αυτοκινήτων, την αναγνώριση προσώπων και την επεξεργασία της φυσικής γλώσσας. Η ιστορία της Μηχανικής Όρασης και της Τεχνητής Νοημοσύνης χρονολογείται από τη δεκαετία του 1950, όταν ξεκίνησαν οι πρώτες προσπάθειες για την ανάπτυξη συστημάτων που να μπορούν να "βλέπουν" και να "κατανοούν"

εικόνες. Οι πρώτοι αλγόριθμοι που αναπτύχθηκαν εστίαζαν στην αναγνώριση απλών σχημάτων και την εξαγωγή βασικών χαρακτηριστικών από τις εικόνες. Καθώς η τεχνολογία εξελισσόταν, αναπτύχθηκαν πιο πολύπλοκοι αλγόριθμοι για την αναγνώριση προηγμένων αντικειμένων και την επεξεργασία εικόνων υψηλής ανάλυσης. Συνοψίζοντας, η Μηχανική Όραση και η Τεχνητή Νοημοσύνη αποτελούν στενά συνδεδεμένους τομείς της επιστήμης των υπολογιστών που επιδιώκουν να αναπτύξουν την ικανότητα των μηχανών να αντιλαμβάνονται και να αλληλεπιδρούν με τον κόσμο μέσω των εικόνων. Η ανάπτυξη στον τομέα αυτό έχει επιτρέψει τη δημιουργία προηγμένων συστημάτων αυτόνομης οδήγησης, αναγνώρισης προσώπων και πολλών άλλων εφαρμογών. Η συνεχής εξέλιξη και βελτίωση αυτών των τεχνολογιών ανοίγει νέους ορίζοντες για εφαρμογές όπως η ρομποτική, η αυτόνομη ναυσιπλοΐα, η ασφάλεια και η υγεία.

[6]

Κεφάλαιο 3: Συστήματα μηχανικής μάθησης

Τα συστήματα μηχανικής μάθησης έχουν επαναστατήσει στον τρόπο που αντιμετωπίζουμε τα προβλήματα και τις προκλήσεις στον κλάδο της τεχνητής νοημοσύνης. Βασίζόμενα σε αλγορίθμους και μοντέλα που επιτρέπουν την αυτόματη εκμάθηση από δεδομένα, τα συστήματα μηχανικής μάθησης επιτυγχάνουν αυξημένη ακρίβεια και απόδοση σε προβλήματα πρόβλεψης, αναγνώρισης προτύπων, ταξινόμησης και άλλων πολύπλοκων εργασιών. Σε αυτήν την ενότητα, θα εξετάσουμε τα θεμέλια των συστημάτων μηχανικής μάθησης, τις διάφορες τεχνικές που χρησιμοποιούνται και τις εφαρμογές τους σε διάφορους τομείς.

3.1 Λειτουργίες μηχανικής μάθησης

Οι βασικές έννοιες που διέπουν τα συστήματα μηχανικής μάθησης είναι οι εξής: οι αλγόριθμοι μάθησης, τα δεδομένα εκπαίδευσης και οι συναρτήσεις κόστους. Οι αλγόριθμοι μάθησης περιλαμβάνουν τον αλγόριθμο εκπαίδευσης και τον αλγόριθμο πρόβλεψης. Ο αλγόριθμος εκπαίδευσης εκπαιδεύει το μοντέλο μηχανικής μάθησης με βάση τα δεδομένα εκπαίδευσης, ενώ ο αλγόριθμος πρόβλεψης χρησιμοποιεί το εκπαιδευμένο μοντέλο για να κάνει προβλέψεις ή να παράγει αποτελέσματα. Τα δεδομένα εκπαίδευσης αποτελούνται από είσοδο και τις αντίστοιχες ετικέτες ή το επιθυμητό αποτέλεσμα. Οι συναρτήσεις κόστους μετρούν την απόδοση του μοντέλου με βάση την απόκλιση μεταξύ της πραγματικής και της προβλεπόμενης εξόδου.

Υπάρχουν διάφορες τεχνικές μηχανικής μάθησης που εφαρμόζονται, όπως οι αλγόριθμοι ταξινόμησης, οι αλγόριθμοι συστάσεων, οι αλγόριθμοι συσταδοποίησης και οι αλγόριθμοι ενίσχυσης. Οι αλγόριθμοι ταξινόμησης χρησιμοποιούνται για να ταξινομήσουν δεδομένα σε διάφορες κατηγορίες ή κλάσεις. Οι αλγόριθμοι συστάσεων προτείνουν αντικείμενα ή πληροφορίες που ενδιαφέρουν έναν χρήστη βάσει των προτιμήσεών του. Οι αλγόριθμοι συσταδοποίησης ομαδοποιούν δεδομένα με βάση τις ομοιότητες τους, ενώ οι αλγόριθμοι ενίσχυσης μαθαίνουν να προβλέπουν την καλύτερη δράση σε ένα περιβάλλον με ανταμοιβή.

Τα συστήματα μηχανικής μάθησης έχουν εφαρμογές σε πολλούς τομείς, όπως η ιατρική, η ρομποτική, η αυτόνομη οδήγηση, η αναγνώριση φωνής και εικόνας, ο χρηματοοικονομικός τομέας και πολλοί άλλοι. Στην ιατρική, οι αλγόριθμοι μηχανικής μάθησης χρησιμοποιούνται για τη διάγνωση ασθενειών, την πρόβλεψη πιθανών παθήσεων και τη βελτίωση της θεραπευτικής απόδοσης. Στη ρομποτική, οι αλγόριθμοι μηχανικής μάθησης βοηθούν τα ρομπότ να αντιλαμβάνονται και να αλληλεπιδρούν με το περιβάλλον τους. Στην αυτόνομη οδήγηση, οι αλγόριθμοι μηχανικής μάθησης επιτρέπουν στα οχήματα να αναγνωρίζουν σήματα, εμπόδια και να προβλέπουν συμπεριφορές άλλων οδηγών.

Συνοψίζοντας, τα συστήματα μηχανικής μάθησης αποτελούν ένα σημαντικό πεδίο της τεχνητής νοημοσύνης που επιτρέπει την αυτόματη εκμάθηση από δεδομένα. Με τη χρήση αλγορίθμων μάθησης, δεδομένων εκπαίδευσης και συναρτήσεων κόστους, αυτά τα συστήματα μπορούν να επιτύχουν ακρίβεια και απόδοση σε πολύπλοκες εργασίες. Οι εφαρμογές τους είναι πολλές και εκτείνονται από την ιατρική και τη ρομποτική μέχρι την αυτόνομη οδήγηση και την αναγνώριση φωνής και εικόνας.

[7]

3.2 Τεχνικές Εντοπισμού Αντικειμένων-Προσώπου

Ο εντοπισμός αντικειμένων-προσώπου είναι ένα σημαντικό πρόβλημα στον τομέα της υπολογιστικής όρασης. Ο στόχος είναι να αναγνωριστούν και να τοποθετηθούν αυτόματα πρόσωπα και αντικείμενα σε εικόνες και βίντεο. Ο εντοπισμός αντικειμένων-προσώπου έχει εφαρμογές σε πολλούς τομείς, όπως η ασφάλεια, η παρακολούθηση, η ρομποτική και οι εικονικές πραγματικότητες. Σε αυτήν την ενότητα, θα εξετάσουμε διάφορες τεχνικές εντοπισμού αντικειμένων-προσώπου και τις εφαρμογές τους.

Υπολογιστικής Όρασης: Ο εντοπισμός αντικειμένων-προσώπου βασίζεται σε αλγόριθμους υπολογιστικής όρασης. Μια από τις κύριες προσεγγίσεις είναι η χρήση συνδυασμού χαρακτηριστικών και αλγορίθμων μηχανικής μάθησης, όπως τα Νευρωνικά Δίκτυα Συνέλιξης (Convolutional Neural Networks - CNNs). Τα CNNs αναλύουν τις εικόνες σε διάφορα επίπεδα χαρακτηριστικών και εξάγουν πληροφορίες για την τοποθεσία των αντικειμένων-προσώπων.

Μέθοδοι Προσδιορισμού Περιοχών-Ενδιαφέροντος (Region of Interest - ROI): Οι μέθοδοι ROI εστιάζουν σε συγκεκριμένες περιοχές της εικόνας που πιθανόν να περιέχουν αντικείμενα-πρόσωπα. Αυτές οι περιοχές εξετάζονται από εξειδικευμένους αλγορίθμους για την αναγνώριση και τον προσδιορισμό των αντικειμένων-προσώπων.

Μέθοδοι Βαθιάς Μάθησης: Οι τεχνικές βαθιάς μάθησης επιτρέπουν την εκπαίδευση μοντέλων που μπορούν να εντοπίζουν αντικείμενα-πρόσωπα με μεγάλη ακρίβεια. Η χρήση νευρωνικών δικτύων όπως τα Υπεράκτια Νευρωνικά Δίκτυα (Faster R-CNNs) και οι Μονάδες Εντοπισμού Αντικειμένων που έχουν επιτύχει εξαιρετικά αποτελέσματα στον εντοπισμό αντικειμένων-προσώπων.

Μέθοδοι Χρήσης Αναλυτικών Μοντέλων: Ορισμένες τεχνικές επικεντρώνονται στη χρήση αναλυτικών μοντέλων για τον εντοπισμό αντικειμένων-προσώπων. Αυτές οι μέθοδοι περιλαμβάνουν τη χρήση μοντέλων όπως τα Χαρτογραφικά Μοντέλα Βάθους (Depth Map Models) και τα Μοντέλα Σημείων Περιγραφής (Feature Point Models) για τον εντοπισμό αντικειμένων-προσώπων με βάση την αναγνώριση χαρακτηριστικών.

Εφαρμογές του Εντοπισμού Αντικειμένων-Προσώπου μπορούμε να συναντήσουμε :

Ασφάλεια: Ο εντοπισμός αντικειμένων-προσώπου χρησιμοποιείται σε συστήματα ασφαλείας για την αναγνώριση και την καταγραφή προσώπων, είτε για λόγους ασφαλείας εσωτερικών χώρων είτε για τον εντοπισμό ανεπιθύμητων ή επικίνδυνων αντικειμένων.

Παρακολούθηση κίνησης: Ο εντοπισμός αντικειμένων-προσώπου είναι σημαντικός στην παρακολούθηση κίνησης σε βίντεο. Χρησιμοποιείται για τον εντοπισμό και την ανίχνευση αντικειμένων-προσώπων, που μπορεί να χρησιμοποιηθούν για την αναγνώριση προσώπων, την ανάλυση κίνησης ή την ανίχνευση ασυνήθιστης συμπεριφοράς.

Ρομποτική: Ο εντοπισμός αντικειμένων-προσώπου είναι σημαντικός στη ρομποτική για τον αυτόνομο προσδιορισμό των αντικειμένων γύρω από το ρομπότ και την αλληλεπίδραση με αυτά.

Εικονική Πραγματικότητα (VR): Ο εντοπισμός αντικειμένων-προσώπου χρησιμοποιείται σε εικονικές πραγματικότητες για την ανίχνευση και την τοποθέτηση εικονικών αντικειμένων-προσώπων σε πραγματικά περιβάλλοντα.

[8]

3.3 Τεχνικές Ταυτοποίησης Προσώπων

Η ταυτοποίηση προσώπων αποτελεί μία από τις σημαντικότερες εφαρμογές της τεχνητής νοημοσύνης και της επεξεργασίας εικόνας. Με τη χρήση συγκεκριμένων τεχνικών, είναι δυνατή η αναγνώριση και ταυτοποίηση ατόμων βάσει των χαρακτηριστικών του προσώπου τους. Στην παρούσα ενότητα, θα εξετάσουμε μερικές από τις κύριες τεχνικές ταυτοποίησης προσώπων και τις εφαρμογές τους.

Ανίχνευση Προσώπων (Face Detection): Η ανίχνευση προσώπων είναι το πρώτο στάδιο της ταυτοποίησης προσώπων και αφορά τον εντοπισμό των περιοχών των προσώπων σε μια εικόνα ή ένα βίντεο. Υπάρχουν πολλές μέθοδοι ανίχνευσης προσώπων, μεταξύ των οποίων ξεχωρίζουν οι μέθοδοι βασισμένες σε χαρακτηριστικά (feature-based methods), οι μέθοδοι βασισμένες σε μοντέλα (model-based methods) και οι μέθοδοι βασισμένες σε νευρωνικά δίκτυα (deep learning-based methods).

Εξαγωγή Χαρακτηριστικών (Feature Extraction): Αφού εντοπιστούν τα πρόσωπα, ακολουθεί η εξαγωγή χαρακτηριστικών από τις αντίστοιχες περιοχές των προσώπων. Αυτά τα χαρακτηριστικά μπορούν να περιλαμβάνουν τη μορφολογία του προσώπου (όπως η τοπολογία των ματιών, της μύτης και του στόματος) και τα χρωματικά χαρακτηριστικά (όπως η αναλογία χρωμάτων και οι αποχρώσεις). Οι μέθοδοι εξαγωγής χαρακτηριστικών περιλαμβάνουν τη χρήση τεχνικών όπως οι Γεωμετρικοί Περιγραφείς (Geometric Descriptors), οι Περιγραφείς Εικονοστοιχείων (Pixel Descriptors) και οι Μετασχηματισμοί Χαρακτηριστικών (Feature Transformations).

Αντιστοίχιση (Matching): Στο στάδιο της αντιστοίχισης, τα χαρακτηριστικά που εξήχθησαν από τα πρόσωπα σε μια εικόνα συγκρίνονται με μια βάση δεδομένων προσώπων για να εντοπιστούν οι αντιστοιχίες. Αυτή η διαδικασία μπορεί να γίνει με τη χρήση μεθόδων όπως οι Περιγραφείς Τενσοριακού Πεδίου (Tensor Field Descriptors), οι Περιγραφείς Χαρακτηριστικών Κλειδιών (Keypoint Descriptors) και οι Μετρικές Απόστασης (Distance Metrics).

Αναγνώριση (Recognition): Το τελικό στάδιο της ταυτοποίησης προσώπων είναι η αναγνώριση των προσώπων βάσει των αντιστοιχιών που βρέθηκαν στο προηγούμενο στάδιο. Αυτή η διαδικασία μπορεί να γίνει με τη χρήση μεθόδων όπως οι Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines), τα Νευρωνικά Δίκτυα (Neural Networks) και οι Μέθοδοι Συγκριτικής Ανάλυσης (Comparative Analysis Methods).

Οι τεχνικές ταυτοποίησης προσώπων έχουν ευρεία εφαρμογή σε πολλούς τομείς, όπως:

Ασφάλεια: Η ταυτοποίηση προσώπων χρησιμοποιείται για τον έλεγχο πρόσβασης σε κτίρια, ηλεκτρονικές συσκευές και πλατφόρμες, εξασφαλίζοντας υψηλό επίπεδο ασφάλειας.

Επιβολή Νόμου: Οι αρχές επιβολής του νόμου χρησιμοποιούν την ταυτοποίηση προσώπων για την αναγνώριση εγκληματοφόρων ή για την επίλυση εγκληματικών πράξεων μέσω της αντιστοίχισης μεταξύ προσώπων και εικόνων από κάμερες ασφαλείας.

Ηλεκτρονική Πληρωμή: Οι τεχνικές ταυτοποίησης προσώπων χρησιμοποιούνται για ασφαλείς και γρήγορες συναλλαγές μέσω της αναγνώρισης προσώπων για την εξουσιοδότηση πληρωμών.

Επισκόπηση Πολυμέσων: Οι υπηρεσίες επισκόπησης πολυμέσων χρησιμοποιούν τα τεχνικά συστήματα ταυτοποίησης προσώπων για την αυτόματη αναγνώριση και ομαδοποίηση προσώπων σε φωτογραφίες και βίντεο.

Ψυχαγωγία και Υπηρεσίες Προσωπικού Συνόδου: Οι τεχνικές ταυτοποίησης προσώπων είναι ευρέως χρησιμοποιούμενες σε εφαρμογές ψυχαγωγίας, όπως οι ψηφιακές μάσκες και οι φίλτρα προσώπου σε κοινωνικές πλατφόρμες.

[9]

Η ανίχνευση αντικειμένων είναι μια σημαντική εργασία στον τομέα της υπολογιστικής όρασης, η οποία αποσκοπεί στον εντοπισμό και την αναγνώριση αντικειμένων σε εικόνες ή βίντεο. Στην παρούσα ενότητα, θα εξετάσουμε και θα αναλύσουμε τις πιο γνωστές τεχνικές ανίχνευσης αντικειμένων και τις εφαρμογές τους σε διάφορους τομείς.

Μετασχηματισμός Haar-like Features:

Ο μετασχηματισμός Haar-like features αποτελεί μια από τις πρώτες τεχνικές ανίχνευσης αντικειμένων που εφαρμόστηκαν με επιτυχία. Αυτή η τεχνική βασίζεται σε ορισμένα χαρακτηριστικά Haar, όπως οριζόντιες και κάθετες γραμμές, που εξετάζονται σε διάφορες περιοχές της εικόνας. Τα χαρακτηριστικά αυτά χρησιμοποιούνται σε έναν ταξινομητή, όπως ένα AdaBoost classifier, για την ανίχνευση των αντικειμένων.

Τεχνικές βασισμένες στο Histogram of Oriented Gradients (HOG):

Η τεχνική HOG είναι μια δημοφιλής μέθοδος ανίχνευσης αντικειμένων, ιδίως για την ανίχνευση προσώπων. Η τεχνική HOG θα αναλυθεί περαιτέρω στην συνέχεια.

Νευρωνικά Δίκτυα:

Τα νευρωνικά δίκτυα έχουν αποδειχθεί πολύ αποτελεσματικά στην ανίχνευση αντικειμένων. Ένα δημοφιλές παράδειγμα είναι το συνελκτικό νευρωνικό δίκτυο (CNN). Τα CNN είναι σε θέση να μάθουν από μεγάλα σύνολα δεδομένων και να αναγνωρίζουν αντικείμενα με εξαιρετική ακρίβεια. Ένα διάσημο παράδειγμα αποτελεί η αρχιτεκτονική του YOLO (You Only Look Once), η οποία παρέχει γρήγορη και ακριβή ανίχνευση αντικειμένων.

[10]

ΚΕΦΑΛΑΙΟ 4 ΕΡΓΑΛΕΙΑ ΑΝΑΠΤΥΞΗΣ

Η ανάπτυξη λογισμικού και εφαρμογών απαιτεί τη χρήση εξειδικευμένων εργαλείων που να παρέχουν τις απαραίτητες δυνατότητες και λειτουργίες για την επίτευξη των στόχων του έργου. Στον ευρύτερο τομέα της όρασης υπολογιστών και της ανίχνευσης αντικειμένων, τρία από τα πιο σημαντικά εργαλεία ανάπτυξης που έχουν επιτύχει μεγάλη αναγνώριση είναι το OpenCV, η Python και το Haar Cascade. Το OpenCV αποτελεί μία ισχυρή βιβλιοθήκη ανοιχτού κώδικα για την επεξεργασία εικόνας και τον υπολογισμό της όρασης των υπολογιστών. Είναι γνωστό για την ευελιξία και τις πολλές λειτουργίες που προσφέρει, επιτρέποντας την ανάπτυξη προηγμένων αλγορίθμων ανίχνευσης και επεξεργασίας εικόνας. Η Python, από την άλλη πλευρά, είναι μία γλώσσα προγραμματισμού υψηλού επιπέδου που προσφέρει ευκολία χρήσης και γρήγορη ανάπτυξη εφαρμογών. Τέλος, το Haar Cascade είναι μία τεχνική ανίχνευσης αντικειμένων που βασίζεται στη χρήση χαρακτηριστικών Haar και έχει εφαρμοστεί με επιτυχία σε πολλά προβλήματα αναγνώρισης και ανίχνευσης προσώπων.

4.1 ΠΕΡΙΓΡΑΦΗ OPENCV

Το OpenCV (Open Source Computer Vision Library) είναι μια ανοικτού κώδικα βιβλιοθήκη υπολογιστικής όρασης (computer vision) που παρέχει ένα ευέλικτο και ισχυρό πλαίσιο για την ανάπτυξη εφαρμογών όρασης μηχανής (machine vision) και ρομποτικής. Το OpenCV έχει γραφεί σε C++ και μπορεί να χρησιμοποιηθεί σε πολλές γλώσσες προγραμματισμού, όπως C++, Python, Java και MATLAB, καθιστώντας το προσβάσιμο σε μια ευρεία κοινότητα προγραμματιστών.

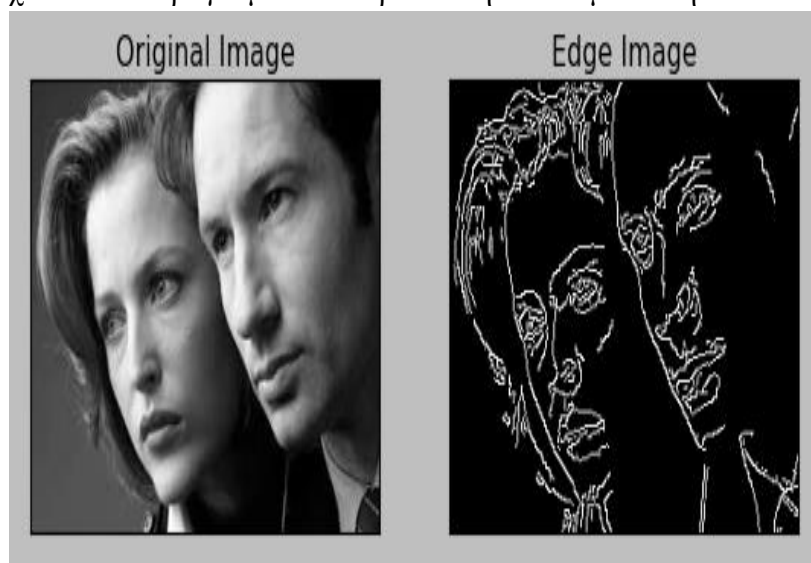
Το OpenCV προσφέρει μια πληθώρα αλγορίθμων και λειτουργιών για την επεξεργασία, την ανάλυση και την αναγνώριση εικόνων και βίντεο. Αυτή η βιβλιοθήκη έχει ευρεία εφαρμογή σε πολλούς τομείς, όπως η ρομποτική, η αυτόνομη οδήγηση, η ασφάλεια, η αναγνώριση προτύπων, η επεξεργασία εικόνας και η ανάλυση βίντεο.

Η βιβλιοθήκη παρέχει μια πληθώρα λειτουργιών, συμπεριλαμβανομένων των εξής:

Ανάγνωση και εγγραφή εικόνων και βίντεο: Το OpenCV μπορεί να διαχειριστεί διάφορα είδη αρχείων εικόνας και βίντεο, όπως JPEG, PNG, BMP, AVI, κ.λπ. Παρέχει επίσης δυνατότητες για την ανάγνωση και εγγραφή από συσκευές εισόδου, όπως κάμερες.

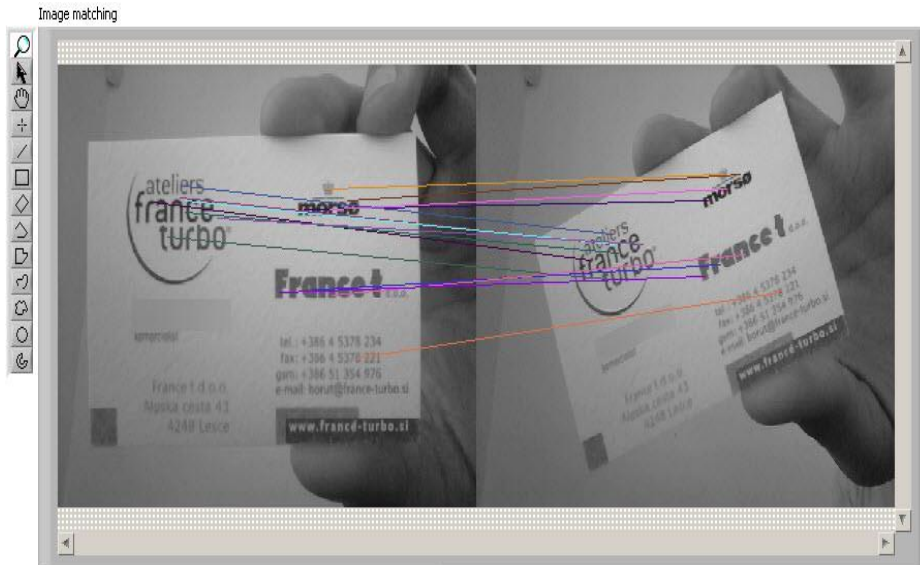
Βασικές λειτουργίες επεξεργασίας εικόνας: Το OpenCV παρέχει απλές λειτουργίες επεξεργασίας εικόνας, όπως αλλαγή μεγέθους, αντιθεση, φωτεινότητα, αποκοπή, περιστροφή, αναστροφή και αλλαγή χρωματικού χώρου.

Ανίχνευση ακμών: Ο αλγόριθμος Canny που υποστηρίζεται από το OpenCV είναι μία από τις πιο γνωστές τεχνικές ανίχνευσης ακμών. Αυτή η τεχνική εντοπίζει τις αλλαγές στην ένταση των εικονοστοιχείων και παράγει μια αναπαράσταση των ακμών στην εικόνα.



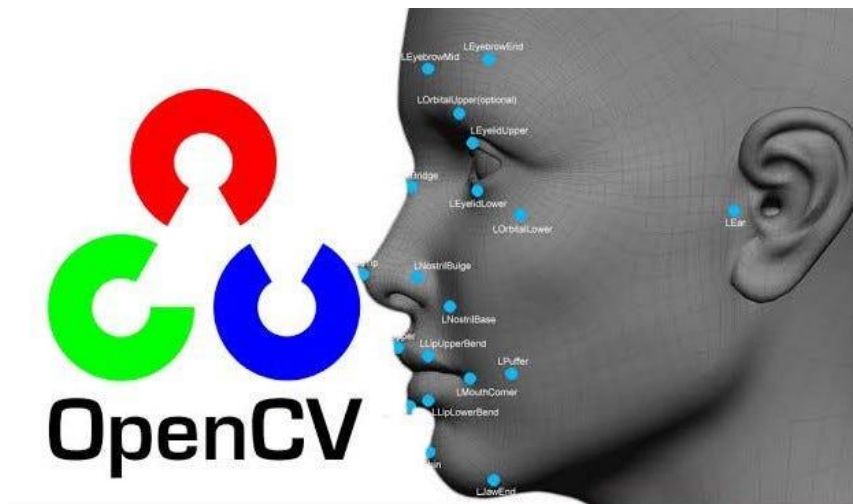
Εικόνα 4.1 Ανίχνευση Ακμών

Ανίχνευση χαρακτηριστικών: Το OpenCV παρέχει αλγορίθμους όπως το SIFT (Scale-Invariant Feature Transform) και το SURF (Speeded Up Robust Features), οι οποίοι είναι κορυφαίες τεχνικές για την ανίχνευση και την περιγραφή χαρακτηριστικών σε εικόνες. Αυτές οι τεχνικές επιτρέπουν τον εντοπισμό και τον προσδιορισμό μοναδικών σημείων ενδιαφέροντος, τα οποία μπορούν να χρησιμοποιηθούν για αναγνώριση, ευθυγράμμιση και αναπαράσταση εικόνων.



Εικόνα 4.2 Αλγόριθμος SIFT

Ανίχνευση προσώπων: Το OpenCV παρέχει αλγορίθμους για την ανίχνευση προσώπων σε εικόνες και βίντεο. Οι αλγόριθμοι Haar Cascade και HOG (Histogram of Oriented Gradients) είναι δύο από τις πιο γνωστές τεχνικές που χρησιμοποιούνται για την ανίχνευση προσώπων.



Εικόνα 4.3 Open CV

Εφαρμογές επαυξημένης πραγματικότητας (AR): Το OpenCV παρέχει εργαλεία και αλγορίθμους για την ανίχνευση και την αντίστοιχη τοποθέτηση 3D αντικειμένων σε εικόνες και βίντεο, δημιουργώντας έτσι εφαρμογές επαυξημένης πραγματικότητας.

[11]

4.2 ΠΕΡΙΓΡΑΦΗ ΡΥΤΗΘΝ

Η Python είναι μια υψηλού επιπέδου γλώσσα προγραμματισμού που δημιουργήθηκε από τον Guido van Rossum και παρουσιάστηκε για πρώτη φορά το 1991. Είναι μια γλώσσα γενικού σκοπού, διερμηνευόμενη, δυναμική και με ευανάγνωστη σύνταξη. Έχει γίνει ιδιαίτερα δημοφιλής λόγω της απλότητάς της, της ευελιξίας και της ικανότητάς της να επεκτείνεται με πρόσθετες βιβλιοθήκες και πακέτα.

Τα βασικά χαρακτηριστικά της Python είναι τα εξής:

Ευανάγνωστη σύνταξη: Η σύνταξη της Python είναι σχεδιασμένη να είναι απλή και ευανάγνωστη, με χρήση κεντρικής σημασίας (indentation) για την οργάνωση των κομματιών κώδικα. Αυτό καθιστά τον κώδικα πιο ευανάγνωστο και εύκολο στη συντήρηση.

Δυναμική τύπων: Στην Python, δεν χρειάζεται να δηλώνετε τον τύπο μιας μεταβλητής κατά τη δήλωσή της. Η Python αναγνωρίζει αυτόματα τον τύπο της μεταβλητής με βάση την τιμή που της ανατίθεται. Αυτό καθιστά την ανάπτυξη γρηγορότερη και ευέλικτη.

Υψηλού επιπέδου δομές δεδομένων: Η Python παρέχει ενσωματωμένες δομές δεδομένων, όπως λίστες, πλειάδες, λεξικά και σύνολα, που επιτρέπουν την αποθήκευση και την οργάνωση δεδομένων με έναν ευέλικτο τρόπο. Αυτές οι δομές δεδομένων διευκολύνουν την επεξεργασία και την ανάλυση των δεδομένων.



Εικόνα 4.4 Χαρακτηριστικά Python

Μεγάλη κοινότητα και οικοσύστημα: Η Python έχει μια μεγάλη και ενεργή κοινότητα προγραμματιστών που υποστηρίζει και συνεισφέρει σε πολλά έργα ανοιχτού κώδικα. Υπάρχουν χιλιάδες πακέτα και βιβλιοθήκες που είναι διαθέσιμες για χρήση, όπως το NumPy για αριθμητικούς υπολογισμούς, το Pandas για ανάλυση δεδομένων, το Django για ανάπτυξη ιστοσελίδων, και πολλά άλλα.

Πλατφόρμες και πολυπλοκότητα: Η Python είναι διαθέσιμη σε πολλές πλατφόρμες, συμπεριλαμβανομένων των Windows, mac OS και Linux. Μπορεί να χρησιμοποιηθεί για ανάπτυξη εφαρμογών επιφάνειας εργασίας, διαδικτυακών εφαρμογών, επιστημονικών υπολογισμών, μηχανικής μάθησης και πολλών άλλων εφαρμογών.

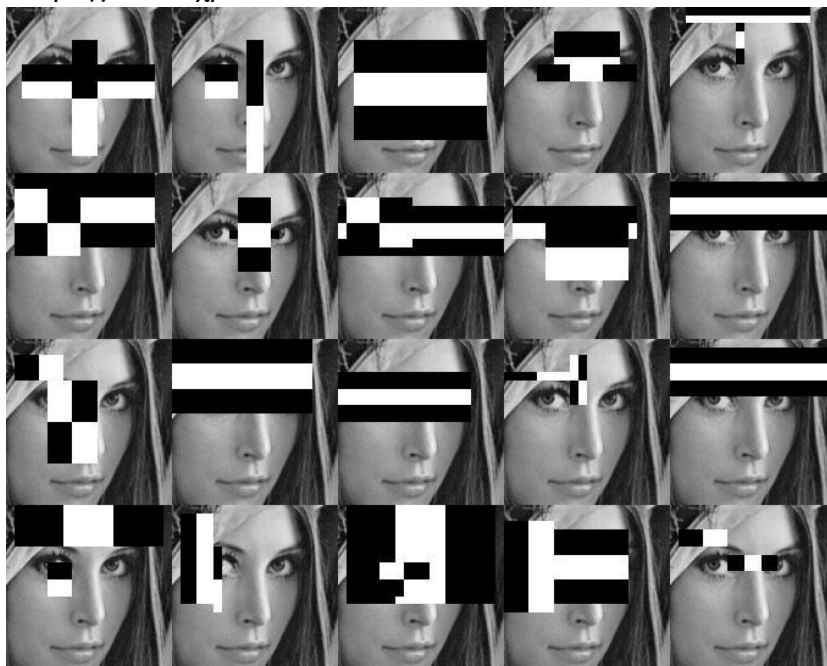
Η Python έχει πολλές άλλες λειτουργίες και χαρακτηριστικά που την καθιστούν μια ισχυρή γλώσσα προγραμματισμού για αρχάριους και έμπειρους προγραμματιστές. Η ευελιξία, η απλότητα και η δυνατότητα επέκτασης της Python την καθιστούν ιδανική για πολλούς τομείς προγραμματισμού.

[12]

4.3 ΠΕΡΙΓΡΑΦΗ HAAR CASCADE

Η ανίχνευση αντικειμένων είναι ένα σημαντικό πρόβλημα στον τομέα της υπολογιστικής όρασης και της επεξεργασίας εικόνας. Ένας από τους πιο γνωστούς αλγόριθμους για την ανίχνευση αντικειμένων είναι το Haar Cascade. Σε αυτήν την ενότητα θα εξετάσουμε το Haar Cascade, περιγράφοντας το τι είναι, πώς λειτουργεί και ποιες είναι οι εφαρμογές του.

Το Haar είναι ένας τύπος χαρακτηριστικών που χρησιμοποιούνται στην ανίχνευση αντικειμένων σε εικόνες και βίντεο. Τα χαρακτηριστικά αυτά αναλύουν τις διαφορές στη φωτεινότητα μεταξύ συνεχών περιοχών της εικόνας. Κάθε χαρακτηριστικό αποτελείται από θετικές και αρνητικές περιοχές, και η αξία του υπολογίζεται από τη διαφορά των γκρι επιπέδων μεταξύ αυτών των περιοχών. Τα χαρακτηριστικά Haar είναι γνωστά για την αποτελεσματικότητά τους και χρησιμοποιούνται ευρέως στην ανίχνευση προσώπων και άλλων αντικειμένων. Η χρήση των χαρακτηριστικών Haar επιτρέπει την ταχεία και ακριβή ανίχνευση αντικειμένων σε πραγματικό χρόνο.



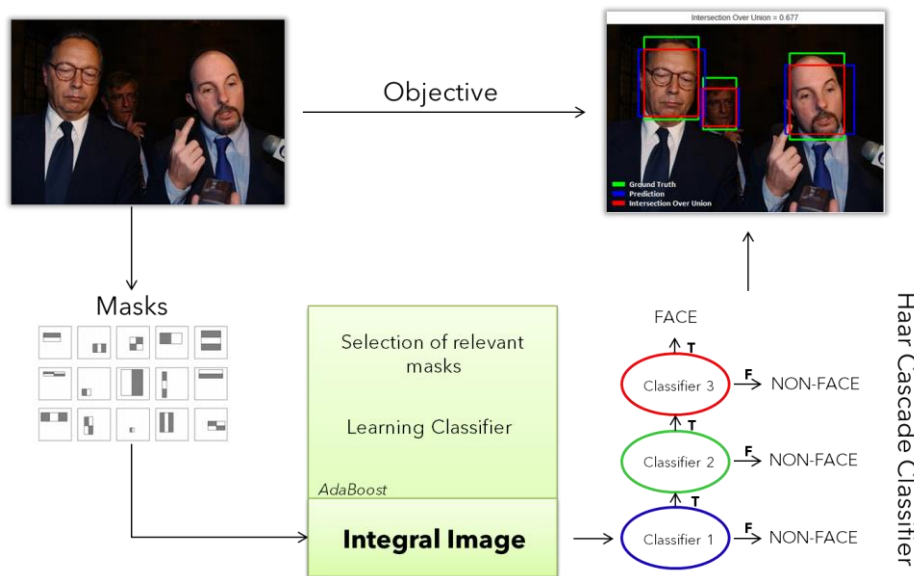
Εικόνα 4.5 Ανίχνευση Προσώπου με Haar Cascade

Το Haar Cascade είναι μια τεχνική ανίχνευσης αντικειμένων που αναπτύχθηκε από τους Paul Viola και Michael Jones το 2001. Βασίζεται στη χρήση ειδικά σχεδιασμένων χαρακτηριστικών Haar για την ανίχνευση αντικειμένων σε εικόνες. Τα χαρακτηριστικά Haar είναι παραλληλόγραμμες περιοχές που υπολογίζουν τη διαφορά των γκρι επιπέδων εντός αυτών των περιοχών. Οι τιμές αυτών των χαρακτηριστικών χρησιμοποιούνται στη συνέχεια για την κατηγοριοποίηση των περιοχών της εικόνας.

Η διαδικασία λειτουργίας του Haar Cascade περιλαμβάνει τα ακόλουθα βασικά βήματα:

Εκπαίδευση: Κατά τη φάση της εκπαίδευσης, ο αλγόριθμος εκπαιδεύεται με ένα μεγάλο σύνολο εικόνων που περιέχουν το επιθυμητό αντικείμενο και εικόνες που δεν περιέχουν το αντικείμενο. Κατά τη διάρκεια αυτής της διαδικασίας, το Haar Cascade εκτελεί αρκετές προσαρμογές και επιλογές χαρακτηριστικών για να επιλέξει τα βέλτιστα χαρακτηριστικά που μπορούν να διαχωρίσουν το αντικείμενο από το φόντο.

Εκπαίδευση του Κατηγορητή: Αφού ολοκληρωθεί η εκπαίδευση, δημιουργείται ένα σύνολο κατηγορητών (classifiers). Ο καθένας από αυτούς τους κατηγορητές είναι ένας απλός αλγόριθμος μηχανικής μάθησης, όπως ο AdaBoost, που βασίζεται στα χαρακτηριστικά Haar που επιλέχθηκαν κατά την εκπαίδευση.



Εικόνα 4.6 Βήματα Haar Cascade

Ταξινόμηση: Κατά τη διάρκεια της διαδικασίας ταξινόμησης, η εικόνα διαβιβάζεται σε ένα παράθυρο που εκτελείται πάνω από την εικόνα και συγκρίνεται με τους κατηγορητές που έχουν εκπαιδευτεί. Αν οι τιμές των χαρακτηριστικών Haar που εξάγονται από το παράθυρο ταιριάζουν με τα κατηγορητές, το παράθυρο ταξινομείται ως πιθανή περιοχή ανίχνευσης.

Επικύρωση: Για την εξακρίβωση της ακρίβειας των αποτελεσμάτων, το Haar Cascade χρησιμοποιεί μια σειρά σταδίων επικύρωσης. Τα παράθυρα που θεωρούνται πιθανές περιοχές ανίχνευσης περνούν από αυτά τα στάδια για να ελεγχθούν περαιτέρω. Σε κάθε στάδιο, μπορεί να εφαρμοστεί ένα κατηγορητής, ο οποίος απορρίπτει τα παράθυρα που δεν αντιστοιχούν στο

επιθυμητό αντικείμενο. Τέλος εφόσον ολοκληρωθούν τα παραπάνω βήματα και αντιστοιχηθεί το επιθυμητό αντικείμενο εμφανίζεται ένα πλαίσιο γύρω από το αντικείμενο.

Το Haar Cascade έχει ευρεία εφαρμογή στην υπολογιστική όραση και την ανίχνευση αντικειμένων. Οι κυριότερες εφαρμογές περιλαμβάνουν:

Ανίχνευση προσώπων: Το Haar Cascade χρησιμοποιείται για την ανίχνευση προσώπων σε εικόνες και βίντεο. Αποτελεί τη βάση για πολλές εφαρμογές αναγνώρισης προσώπου.

Ανίχνευση αντικειμένων: Εκτός από την ανίχνευση προσώπων, το Haar Cascade μπορεί να χρησιμοποιηθεί για την ανίχνευση άλλων αντικειμένων, όπως αυτοκίνητα, ποδήλατα, ζώα κ.λπ.

Εφέ σε πραγματικό χρόνο: Το Haar Cascade είναι αποτελεσματικός για εφαρμογές σε πραγματικό χρόνο, καθώς μπορεί να παρέχει γρήγορη ανίχνευση αντικειμένων σε πραγματικό χρόνο.

Συμπεράσματα:

Το Haar Cascade είναι μια ισχυρή τεχνική ανίχνευσης αντικειμένων που βασίζεται στη χρήση χαρακτηριστικών Haar. Η μεγάλη του δυνατότητα είναι ότι μπορεί να ανιχνεύσει αντικείμενα με αξιόλογη ακρίβεια και σε πραγματικό χρόνο. Με εφαρμογές στην ανίχνευση προσώπων και σε άλλα αντικείμενα, το Haar Cascade έχει αναδειχθεί ως μια αξιόπιστη και αποτελεσματική μέθοδος ανίχνευσης αντικειμένων σε εικόνες και βίντεο.

[13]

ΚΕΦΑΛΑΙΟ 5

ΠΕΡΙΓΡΑΦΗ ΑΛΓΟΡΙΘΜΟΥ ΕΚΠΑΙΔΕΥΣΗΣ

Για την αναγνώριση προσώπων θα χρησιμοποιήσουμε την μέθοδο HOG (Histogram of Oriented Gradients) η οποία αποτελεί μια προηγμένη τεχνική στον χώρο της αναγνώρισης προσώπου, εξειδικευμένη στην εξαγωγή και ανάλυση των χαρακτηριστικών ενός αντικειμένου. Αποτελεί ένα ευέλικτο εργαλείο που έχει εφαρμοστεί επιτυχώς σε ποικίλες εφαρμογές όρασης υπολογιστών. Εστιάζει στην ανάλυση των κατευθύνσεων των οριογραμμών σε μια εικόνα, εκτιμώντας την κατεύθυνση και την ένταση των pixel, με στόχο την ανίχνευση συγκεκριμένων δομικών στοιχείων.

Η διαδικασία ξεκινά με τη διαίρεση της εικόνας σε μικρά κελιά. Για κάθε κελί, υπολογίζονται η ένταση και η κατεύθυνση των pixel που περιλαμβάνει. Αυτά τα προσανατολισμένα οριογράμματα ομαδοποιούνται σε κύτταρα, με κάθε κύτταρο να περιέχει ένα ιστόγραμμα των κατευθύνσεών τους. Η κανονικοποίηση των ιστογραμμάτων διασφαλίζει την ανεξαρτησία από τις διακυμάνσεις στη φωτεινότητα.

Το HOG αποκτά την ιδιαιτερότητα του όταν τα κύτταρα οργανώνονται σε μπλοκ. Κάθε μπλοκ περιέχει πολλά κύτταρα, και η συνολική πληροφορία από τα κύτταρα εντός του μπλοκ συνδυάζεται. Αυτό το διαφορετικό επίπεδο οργάνωσης παρέχει περισσότερη πληροφορία σχετικά με τη δομή των αντικειμένων.

Η σημαντικότητα του HOG στην αναγνώριση προσώπου οφείλεται στο ότι τα πρόσωπα έχουν εκφραστική δομή, η οποία είναι καλά περιγράψιμη μέσω των προσανατολισμένων οριογραμμμάτων. Οι αλγόριθμοι μηχανικής μάθησης, όπως οι μηχανές διανυσμάτων υποστήριξης (SVM), εκπαιδεύονται με βάση το τελικό διάνυσμα χαρακτηριστικών για την ταξινόμηση προσώπων.

Το HOG έχει εφαρμοστεί σε ερευνητικές εργασίες και σε εφαρμογές πρακτικής, όπως συστήματα παρακολούθησης και ασφαλείας, διότι προσφέρει αξιόπιστη αναγνώριση προσώπων σε ποικίλες συνθήκες φωτισμού και γωνίας λήψης.

Στο πεδίο της αναγνώρισης προσώπου, η μέθοδος HOG συνδυάζεται άριστα με την μέθοδο Haar Cascade η οποία αναλύθηκε παραπάνω επίσης η χρήση νευρωνικών δικτύων προτείνεται για ακόμη καλύτερα αποτελέσματα. [14]

5.1 Εφαρμογή του αλγορίθμου HOG

Υπολογισμός κλίσης

Το πρώτο βήμα υπολογισμού σε πολλούς ανιχνευτές χαρακτηριστικών στην προεπεξεργασία εικόνας είναι η εξασφάλιση κανονικοποιημένων τιμών χρώματος και γάμμα. Ωστόσο, όπως επισημαίνουν οι Dalal και Triggs, αυτό το βήμα μπορεί να παραλειφθεί στον υπολογισμό του περιγραφέα HOG, καθώς η επακόλουθη κανονικοποίηση του περιγραφέα επιτυγχάνει ουσιαστικά το ίδιο αποτέλεσμα. Η προεπεξεργασία της εικόνας παρέχει επομένως μικρή επίδραση στην απόδοση. Αντ' αυτού, το πρώτο βήμα υπολογισμού είναι ο υπολογισμός των τιμών κλίσης. Η πιο συνηθισμένη μέθοδος είναι η εφαρμογή της 1-D κεντραρισμένης, σημειακής μάσκας διακριτής παραγώγου σε μία ή και στις δύο οριζόντιες και κάθετες κατευθύνσεις. Συγκεκριμένα, η μέθοδος αυτή απαιτεί φιλτράρισμα των δεδομένων χρώματος ή έντασης της εικόνας με τους ακόλουθους πυρήνες φίλτρων:

$[-1,0,1]$ και $[-1,0,1]^T$.

Οι Dalal και Triggs δοκίμασαν άλλες, πιο σύνθετες μάσκες, όπως η μάσκα Sobel 3x3 ή οι διαγώνιες μάσκες, αλλά αυτές οι μάσκες είχαν γενικά χειρότερες επιδόσεις στην ανίχνευση ανθρώπων σε εικόνες. Πειραματίστηκαν επίσης με την εξομάλυνση Gauss πριν από την εφαρμογή της μάσκας παραγώγων, αλλά ομοίως διαπίστωσαν ότι η παράλειψη οποιασδήποτε εξομάλυνσης είχε καλύτερη απόδοση στην πράξη.

Διαχωρισμός προσανατολισμού

Το δεύτερο βήμα του υπολογισμού είναι η δημιουργία των ιστογραμμάτων κυττάρων. Κάθε εικονοστοιχείο εντός του κελιού δίνει μια σταθμισμένη ψήφο για ένα δυαδικό ιστόγραμμα προσανατολισμού με βάση τις τιμές που βρέθηκαν στον υπολογισμό της κλίσης. Τα ίδια τα κελιά μπορούν να έχουν είτε ορθογώνιο είτε ακτινωτό σχήμα και τα κανάλια ιστογράμματος κατανέμονται ομοιόμορφα στις 0 έως 180 μοίρες ή στις 0 έως 360 μοίρες, ανάλογα με το αν η κλίση είναι "χωρίς πρόσημο" ή "με πρόσημο". Οι Dalal και Triggs διαπίστωσαν ότι οι μη προσημασμένες κλίσεις που χρησιμοποιήθηκαν σε συνδυασμό με 9 κανάλια ιστογράμματος είχαν τις καλύτερες επιδόσεις στα πειράματά τους για την ανίχνευση ανθρώπων, ενώ σημείωσαν ότι οι προσημασμένες κλίσεις οδηγούν σε σημαντικές βελτιώσεις στην αναγνώριση ορισμένων άλλων κατηγοριών αντικειμένων, όπως τα αυτοκίνητα ή οι μοτοσικλέτες. Όσον αφορά το βάρος της ψήφου, η συνεισφορά των εικονοστοιχείων μπορεί να είναι είτε το ίδιο το μέγεθος της κλίσης είτε κάποια συνάρτηση του μεγέθους. Στις δοκιμές, το ίδιο το μέγεθος της κλίσης παράγει γενικά τα καλύτερα αποτελέσματα. Άλλες επιλογές για το βάρος της ψήφου θα μπορούσαν να περιλαμβάνουν την τετραγωνική ρίζα ή το τετράγωνο του μεγέθους της κλίσης, ή κάποια αποκομμένη έκδοση του μεγέθους

Μπλοκ περιγραφής

Για να ληφθούν υπόψη οι αλλαγές στο φωτισμό και την αντίθεση, οι δυνάμεις της κλίσης πρέπει να κανονικοποιηθούν τοπικά, πράγμα που απαιτεί την ομαδοποίηση των κυττάρων σε μεγαλύτερα, χωρικά συνδεδεμένα μπλοκ. Ο περιγραφέας HOG είναι στη συνέχεια το συνυφασμένο διάλυμα των συνιστωσών των κανονικοποιημένων ιστογραμμάτων κυττάρων από όλες τις περιοχές μπλοκ. Αυτά τα μπλοκ συνήθως επικαλύπτονται, πράγμα που σημαίνει ότι κάθε κύτταρο συνεισφέρει περισσότερες από μία φορές στον τελικό περιγραφέα. Υπάρχουν δύο κύριες γεωμετρικές μπλοκ: ορθογώνια μπλοκ R-HOG και κυκλικά μπλοκ C-HOG. Τα μπλοκ R-HOG είναι γενικά τετραγωνικά πλέγματα, τα οποία αναπαρίστανται με τρεις παραμέτρους: τον αριθμό των κελιών ανά μπλοκ, τον αριθμό των εικονοστοιχείων ανά κελί και τον αριθμό των καναλιών ανά ιστόγραμμα κελιού. Στο πείραμα ανίχνευσης ανθρώπων των Dalal και Triggs, οι βέλτιστες παράμετροι βρέθηκαν να είναι τέσσερα κελιά 8x8 pixels ανά μπλοκ (16x16 pixels ανά μπλοκ) με 9 κανάλια ιστογράμματος. Επιπλέον, διαπίστωσαν ότι θα μπορούσε να επιτευχθεί κάποια μικρή βελτίωση της απόδοσης με την εφαρμογή ενός χωρικού παραθύρου Gauss σε κάθε μπλοκ πριν από την καταγραφή των ψήφων ιστογράμματος, ώστε να σταθμιστούν λιγότερο τα εικονοστοιχεία γύρω από την άκρη των μπλοκ. Τα μπλοκ R-HOG φαίνονται αρκετά παρόμοια με τους περιγραφείς μετασχηματισμού χαρακτηριστικών κλίμακας (SIFT), ωστόσο, παρά τον παρόμοιο σχηματισμό τους, τα μπλοκ R-HOG υπολογίζονται σε πυκνά πλέγματα σε κάποια ενιαία κλίμακα χωρίς ευθυγράμμιση προσανατολισμού, ενώ οι περιγραφείς SIFT υπολογίζονται συνήθως σε αραιά, αναλλοίωτα σε κλίμακα βασικά σημεία της εικόνας και περιστρέφονται για να ευθυγραμμιστεί ο προσανατολισμός. Επιπλέον, τα μπλοκ R-HOG χρησιμοποιούνται σε συνδυασμό για την κωδικοποίηση πληροφοριών χωρικής μορφής, ενώ οι περιγραφείς SIFT χρησιμοποιούνται μεμονωμένα.

Τα κυκλικά μπλοκ HOG (C-HOG) μπορούν να βρεθούν σε δύο παραλλαγές: εκείνα με ένα μόνο, κεντρικό κελί και εκείνα με ένα γωνιακά διαιρεμένο κεντρικό κελί. Επιπλέον, αυτά τα

μπλοκ C-HOG μπορούν να περιγραφούν με τέσσερις παραμέτρους: τον αριθμό των γωνιακών και ακτινικών κυψελών, την ακτίνα της κεντρικής κυψέλης και τον παράγοντα επέκτασης για την ακτίνα των πρόσθετων ακτινικών κυψελών. Οι Dalal και Triggs διαπίστωσαν ότι οι δύο κύριες παραλλαγές παρείχαν ίση απόδοση και ότι δύο ακτινικά bins με τέσσερα γωνιακά bins, ακτίνα κέντρου 4 pixels και συντελεστή επέκτασης 2 παρείχαν την καλύτερη απόδοση στον πειραματισμό τους. Επίσης, η γκαουσιανή στάθμιση δεν παρείχε κανένα όφελος όταν χρησιμοποιήθηκε σε συνδυασμό με τα μπλοκ C-HOG. Τα μπλοκ C-HOG φαίνονται παρόμοια με τους περιγραφείς πλαισίου σχήματος, αλλά διαφέρουν έντονα στο ότι τα μπλοκ C-HOG περιέχουν κελιά με πολλά κανάλια προσανατολισμού, ενώ τα πλαίσια σχήματος χρησιμοποιούν μόνο έναν αριθμό παρουσίας ακμών στη διατύπωσή τους

Αναγνώριση αντικειμένων

Οι περιγραφείς HOG μπορούν να χρησιμοποιηθούν για την αναγνώριση αντικειμένων παρέχοντάς τους ως χαρακτηριστικά σε έναν αλγόριθμο μηχανικής μάθησης. Οι Dalal και Triggs χρησιμοποίησαν περιγραφείς HOG ως χαρακτηριστικά σε μια μηχανή διανυσμάτων υποστήριξης (SVM) ωστόσο, οι περιγραφείς HOG δεν συνδέονται με έναν συγκεκριμένο αλγόριθμο μηχανικής μάθησης.[15]

Κωδικοποίηση

Στο κομμάτι της κωδικοποίησης και του Dataset το αρχείο `train_model.py` θα αναλύσει τις φωτογραφίες που βρίσκονται στον φάκελο `dataset` και θα δημιουργήσει ένα αρχείο με όνομα `encodings.pickle` το οποίο θα περιέχει τα κριτήρια για την αναγνώριση των προσώπων. Προτείνεται οι φωτογραφίες να οργανώνονται σε φακέλους με βάση το όνομα του ατόμου. Για παράδειγμα, δημιουργήστε ένα νέο φάκελο με το όνομα `Paul` και τοποθετήστε όλες τις φωτογραφίες του προσώπου του `Paul` στο φάκελο `Paul` μέσα στο φάκελο `dataset`.

Εκπαίδευση μοντέλου

Το αρχείο `train_model.py` χρησιμοποιεί τις βιβλιοθήκες (OpenCV, `face_recognition` και `Imutils`). Σκοπός του είναι να φορτώσει το περιεχόμενο του φακέλου “`dataset`”, να το αναλύσει και να δημιουργήσει τις κωδικοποιήσεις για κάθε πρόσωπο ώστε να μπορεί να το αναγνωρίσει. Χρησιμοποιώντας το `Imutils` μπορεί να εντοπίσει τη διαδρομή που έχουμε ορίσει. Επιπλέον, χρησιμοποιώντας την βιβλιοθήκη `pickle` πραγματοποιεί σειριοποίηση και από-σειροποίηση μιας δομής ενός αντικειμένου στην Python. Συγκεκριμένα, τη χρησιμοποιούμε για να δημιουργήσουμε και να γράψουμε στο δίσκο το αρχείο “`encodings.pickle`”. Το OpenCV χρησιμοποιείται για την μετατροπή της εικόνας από RGB του OpenCV σε RGB του `dlib`. Το `dlib` είναι μια εργαλειοθήκη για τη δημιουργία μηχανικής μάθησης σε πραγματικό κόσμο καθώς και ανάλυση δεδομένων. Έπειτα, με την χρήση της βιβλιοθήκης `face_recognition` εντοπίζονται οι συντεταγμένες του κάθε προσώπου της εικόνας εισόδου και δημιουργούνται τα κριτήρια για τον εντοπισμό του μέσω του Haar Cascade. Τα κριτήρια αυτά, με την ολοκλήρωση της διαδικασίας, αποθηκεύονται στο αρχείο “`encodings.pickle`”.

Αναγνώριση

Προσώπου

Το αρχείο `face_req.py` είναι υπεύθυνο για την αναγνώριση των προσώπων σε πραγματικό

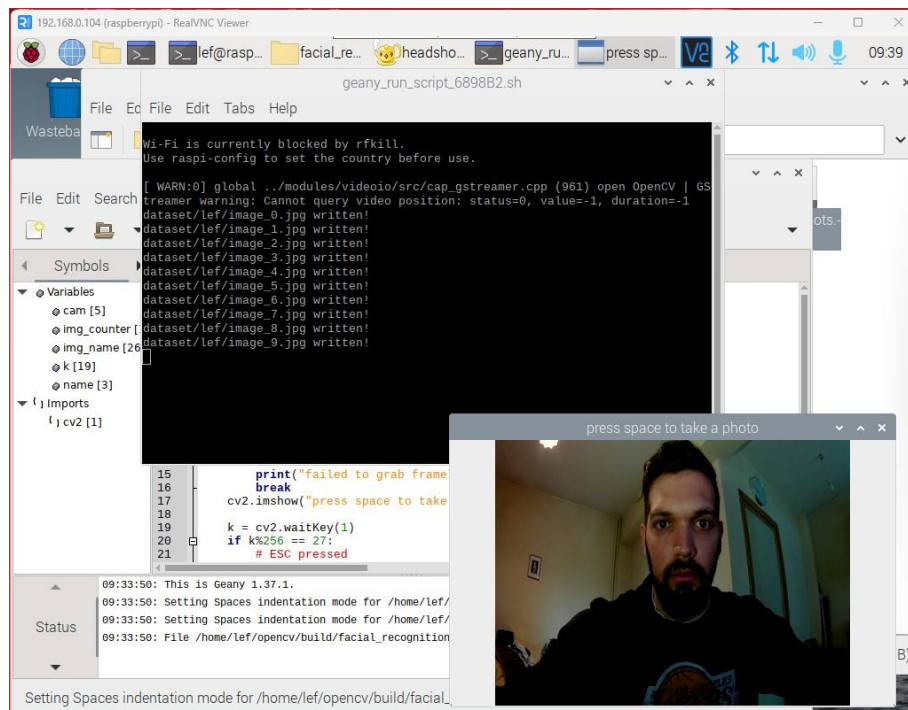
χρόνο μέσω κάμερας. Με την εκτέλεση του script θα φορτωθούν τα αρχεία “encodings.pickle” και “haarcascade_frontalface_default.xml”, όπου το πρώτο περιλαμβάνει τις κωδικοποιήσεις των προσώπων για την αναγνώρισή τους και το δεύτερο είναι υπεύθυνο για τον εντοπισμό τους. Το πρόγραμμα θα εκτελείται συνεχώς έως ότου πατήσουμε το κουμπί Esc για να διακόψουμε την λειτουργία του. Ολοκληρώνοντας τις απαραίτητες μετατροπές στο κάθε καρέ μόλις εντοπιστεί ένα πρόσωπο, το περιβάλλει με ένα κίτρινο τετράγωνο όπου στο πάνω μέρος αναγράφεται το όνομα του προσώπου στην περίπτωση αναγνώρισης διαφορετικά εμφανίζεται η λέξη Unknown.

ΥΛΟΠΟΙΗΣΗ ΣΥΣΤΗΜΑΤΟΣ

Για την υλοποίηση του συστήματος εντοπισμού και ανίχνευσης προσώπων με την χρήση Raspberry Pi θα χρειαστούμε ένα Raspberry Pi 4, μια κάμερα και μια micro SD των 16Gb

Αρχικά από την ιστοσελίδα του Raspberry Pi (<https://www.raspberrypi.com/>) θα κατεβάσουμε και θα εγκαθιστούμε το λειτουργικό σύστημα raspberry Pi Os στην κάρτα SD

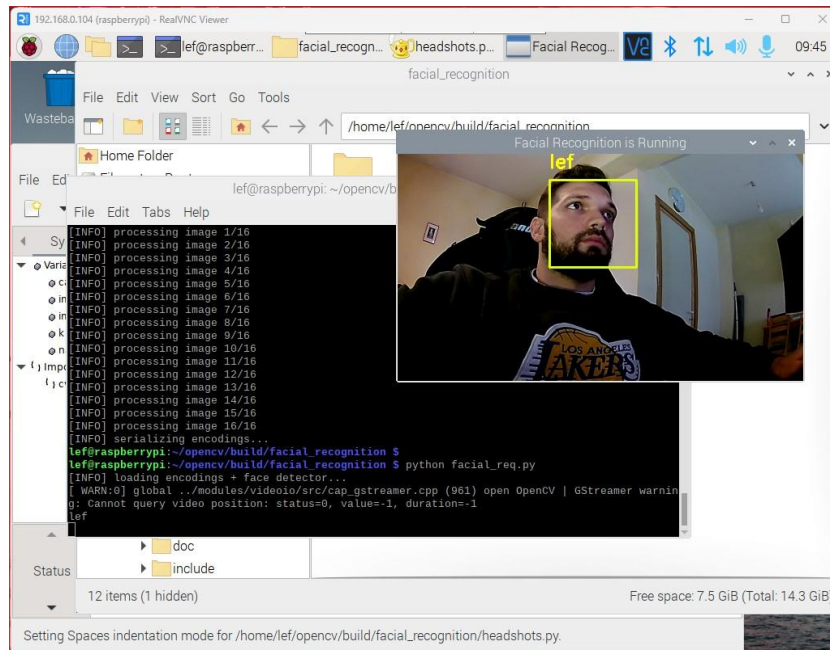
στην συνέχεια με την χρήση του Raspberry θα εγκαταστήσουμε τα προγράμματα : OpenCV, face_recognition και Imutils.



Εικόνα 5.21 Λήψη φωτογραφιών για την εκπαίδευση του μοντέλου

Εγκαθιστώντας τον βασικό κώδικα από την ιστοσελίδα https://github.com/carolinedunn/face_recognition και εκπαιδευοντας το μοντέλο με

φωτογραφίες ατόμων οι οποίες λήφθηκαν με την χρήση της κάμερας το πρόγραμμα αναγνωρίζει πρόσωπα σε πραγματικό χρόνο.



Εικόνα 5.22 Εκτέλεση αρχείου facial_req.py και αναγνώριση προσώπου

[16]

ΠΗΓΕΣ :

[1]

Raspberry Pi Foundation: <https://www.raspberrypi.org/>
"Raspberry Pi For Dummies" από Sean McManus και Mike Cook, Wiley, 2019.
"Exploring Raspberry Pi: Interfacing to the Real World with Embedded Linux" από Derek Molloy, Wiley, 2016.
Επίσημη σελίδα του Raspberry Pi στο GitHub: <https://github.com/raspberrypi>

[2]

Upton, Eben, et al. "The Raspberry Pi Foundation: Pi in the Sky." Computer, vol. 48, no. 8, 2015,
"Raspberry Pi User Guide" από Eben Upton και Gareth Halfacree, Wiley, 2016.
"The Raspberry Pi story: From hobbyists to industry" από Sam Gibbs, Ars Technica, 2017.

[3]

"Raspberry Pi For Dummies" , Sean McManus , Mike Cook, Wiley, 2019.
"Exploring Raspberry Pi: Interfacing to the Real World with Embedded Linux" , Derek Molloy, Wiley, 2016.

[4]

Επίσημη ιστοσελίδα του Raspberry Pi Foundation (<https://www.raspberrypi.org/>)
Επίσημη ιστοσελίδα του Ubuntu MATE (<https://ubuntu-mate.org/>)
Επίσημη ιστοσελίδα του OSMC (<https://osmc.tv/>)
Επίσημη ιστοσελίδα της Microsoft για το Windows IoT Core (<https://docs.microsoft.com/en-us/windows/iot-core/>)
"Raspberry Pi Cookbook" από Simon Monk

[5]

Szeliski, R. (2010). Computer Vision: Algorithms and Applications. Springer.
Forsyth, D. A., & Ponce, J. (2012). Computer Vision: A Modern Approach. Prentice Hall.
Ballard, D. H., & Brown, C. M. (1982). Computer Vision. Prentice Hall.
Image and Vision Computing Journal - Special Issue on Computer Vision History. Volume 27, Issue 11, October 2009.
IEEE Computer Society Technical Committee on Pattern Analysis and Machine Intelligence (PAMI) - History of Computer Vision.
<https://www.computer.org/technical-committees/pami/history-of-computer-vision/>

[6]

Forsyth, D., & Ponce, J. (2012). Computer Vision: A Modern Approach. Pearson Education.

- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
- Russell, S. J., & Norvig, P. (2016). *Artificial Intelligence: A Modern Approach*. Pearson.
- Fei-Fei, L., & Perona, P. (2005). A Bayesian Hierarchical Model for Learning Natural
- [7]
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.
- Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. MIT Press.
- [8]
- Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems*
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*.
- Zhang, Z., Zhou, P., Lin, Z., & Zhang, C. (2018). Face detection: A survey. *IEEE transactions on pattern analysis and machine intelligence*.
- [9]
- Zhao, W., Chellappa, R., Phillips, P.J., & Rosenfeld, A. (2003). Face recognition: A literature survey. *ACM Computing Surveys*, 35(4), 399-458.
- Li, S.Z., & Jain, A.K. (2011). *Handbook of Face Recognition*. Springer.
- Jain, A.K., Ross, A., & Prabhakar, S. (2004). An Introduction to Biometric Recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1), 4-20.
- Viola, P., & Jones, M. (2004). Robust real-time face detection. *International Journal of Computer Vision*, 57(2), 137-154.
- Parkhi, O.M., Vedaldi, A., Zisserman, A., & Jawahar, C.V. (2015). Deep Face Recognition. *British Machine Vision Conference (BMVC)*, 6(1), 41.
- [10]
- Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2001*, 1, I-I.
- Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005*, 1, 886-893.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, 779-788.

[11]

Bradski, G., & Kaehler, A. (2008). Learning OpenCV: Computer vision with the OpenCV library. O'Reilly Media.

OpenCV documentation: <https://docs.opencv.org/>

Szeliski, R. (2010). Computer vision: algorithms and applications. Springer Science & Business Media.

Forsyth, D. A., & Ponce, J. (2012). Computer vision: a modern approach. Pearson Education.

[12]

Python.org: <https://www.python.org/doc/>

Python 3 Documentation: <https://docs.python.org/3/>

Python Wikipedia: [https://en.wikipedia.org/wiki/Python_\(programming_language\)](https://en.wikipedia.org/wiki/Python_(programming_language))

[13]

Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1, I-511-I-518.

OpenCV Documentation: Haar Cascade Classifiers.

Bradski, G., & Kaehler, A. (2008). Learning OpenCV: Computer Vision with the OpenCV Library. O'Reilly Media.

[14]

Dalal, N., & Triggs, B. (2005). Histograms of Oriented Gradients for Human Detection. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). doi: 10.1109/CVPR.2005.177

Dalal, N., Triggs, B., Schmid, C. (2006). Human Detection Using Oriented Histograms of Flow and Appearance. European Conference on Computer Vision (ECCV'06). doi: 10.1007/11744023_35

[15]

Histograms of Oriented Gradients for Human Detection Navneet Dalal and Bill Triggs

[16] Raspberry Pi 4 Facial Recognition https://github.com/carolinedunn/facial_recognition