

ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

«Κάμερα ασφαλείας με αναγνώριση προσώπου σε Raspberry Pi και χειρισμός με εφαρμογή Android»

Της φοιτήτριας
Καρυπίδου Ελλάδας
Αρ. Μητρώου: 123861

Επιβλέπων
Γιακουμής Άγγελος
Επίκουρος Καθηγητής

ΘΕΣΣΑΛΟΝΙΚΗ
ΦΕΒΡΟΥΑΡΙΟΣ 2023

Κάμερα ασφαλείας με αναγνώριση προσώπου
σε Raspberry Pi και χειρισμός με εφαρμογή Android

Κωδικός Δ.Ε. 19132

Φοιτήτρια: Καρυπίδου Ελλάδα
Εισηγητής: Άγγελος Γιακουμής
ΦΕΒΡΟΥΑΡΙΟΣ 2023

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

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

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

« Security camera with facial recognition on Rasberry Pi and control with
Android app »

Ellada Karypidou

Περιεχόμενα

Πρόλογος
Περίληψη
Abstract
Ευχαριστήρια
Κεφάλαιο 1ο: Raspberry Pi 4.....	11
1.1 Ιστορία.....	12
1.2 Λειτουργικό Σύστημα -Raspberry Pi OS.....	12
1.3 Debian Λειτουργικό Σύστημα.....	13
1.4 Συναρμολόγηση και εγκατάσταση λειτουργικού συστήματος.....	14
Κεφάλαιο 2ο: Βιβλιοθήκη OpenCV -(Open Source Computer Vision Library).....	15
Κεφάλαιο 3ο: Λειτουργικό Σύστημα Android.....	16
3.1 Εκδόσεις του Android [22].....	16
Κεφάλαιο 4ο: Γλώσσα προγραμματισμού Python.....	17
4.1 Χρήσεις της γλώσσας προγραμματισμού python [20].....	17
Κεφάλαιο 5ο: Τεχνητή Νοημοσύνη.....	18
5.1 Ορισμός TN.....	20
5.2 Σύντομη ιστορική αναδρομή.....	21
5.2.1 Σύντομο χρονοδιάγραμμα.....	23
5.3 Είδη Μάθησης.....	24
Κεφάλαιο 6ο: Computer Vision.....	24
6.1 Landmark Detection.....	25
Κεφάλαιο 7ο: Histograms of Oriented Gradients-HOG.....	25
7.1 Τι είναι ο «Feature Descriptor» περιγραφέας χαρακτηριστικών.....	26
7.2 Αλγόριθμος μάθησης για ταξινόμηση.....	27
Κεφάλαιο 8ο: Ο νευρώνας.....	29
Κεφάλαιο 9ο: Ανθρώπινο πρόσωπο.....	30
9.1 Ενδοατομικές και διατομικές παραλλαγές.....	31
9.2 Ανίχνευση προσώπου με βάση τη γνώση.....	32
9.3 Ανίχνευση προσώπου με βάση τα χαρακτηριστικά.....	33
9.4 Ανίχνευση προσώπου με βάση τα περιγράμματα (templates).....	33
9.5 Ανίχνευση προσώπου με χρήση καθολικών αλγορίθμων.....	34
9.6 Αναγνώριση προσώπου.....	34

9.7	Αναγνώριση φωνής.....	35
9.8	Δακτυλικό αποτύπωμα	35
9.9	Αναγνώριση βάση συμπεριφοράς	35
9.10	Δακτυλογράφηση.....	35
9.11	Αναγνώριση ίριδας.....	35
Κεφάλαιο 10ο:	Αναγνώριση προσώπου η υλοποίηση με raspberry pi 4	37
10.1	Εγκατάσταση βιβλιοθηκών	37
10.2	RaspController-Απομακρυσμένη διαχείριση του raspberry pi	39
10.3	Δημιουργία καταλόγων.....	43
10.4	Βήμα πρώτο -Αποθήκευση στιγμιότυπων προσώπου.....	44
10.5	Βήμα δεύτερο- Εκπαίδευση αλγορίθμου	46
10.6	Βήμα τρίτο- Αναγνώριση προσώπου	46
Κεφάλαιο 11ο:	Κώδικας υλοποίησης.....	48
11.1	facial_req.py.....	48
11.2	facial_req_email.py	51
11.3	headshots.py	55
11.4	headshots_picam.py	57
11.5	train_model.py.....	58
Κεφάλαιο 12ο:	Συμπεράσματα	60

Πρόλογος

Η εργασία πραγματοποιείται στο πλαίσιο της διπλωματικής εργασίας για την απόκτηση του πτυχίου μηχανικού πληροφορικής του Τμήματος Μηχανικών Πληροφορικής και Ηλεκτρονικών Συστημάτων του Διεθνούς Πανεπιστημίου Ελλάδος. Υλοποιήθηκε ο κώδικας για την αναγνώριση προσώπων με τη χρήση του raspberry pi 4 και τη χρήση της βιβλιοθήκης OpenCV, . Η υλοποίηση της πτυχιακής εργασίας απαιτούσε την προμήθεια των υλικών- hardware του raspberry pi 4, τη συναρμολόγηση και την εγκατάσταση του λειτουργικού συστήματος και των απαραίτητων βιβλιοθηκών, λεπτομέρειες θα δοθούν στις παρακάτω ενότητες και θα αναλυθεί όλη η τεχνολογία που χρησιμοποιήθηκε τόσο σε υλικό- hardware όσο και σε λογισμικό- software. Ο αλγόριθμος στη πρώτη φάση εκπαιδεύεται και σε δεύτερη φάση γίνεται η αναγνώριση προσώπου. Γίνεται η χρήση της κάμερας raspberry pi camera board v2 από την οποία αποθηκεύονται οι φωτογραφίες του προσώπου προς αναγνώριση και τέλος η εκτέλεση του αλγορίθμου με την οποία γίνεται η οροθέτηση του προσώπου και η αναγνώριση ή όχι. Στην περίπτωση που γίνει η αναγνωριστεί ενός προσώπου στέλνεται ηλεκτρονικό μήνυμα ειδοποίησης. Ο χειρισμός της εφαρμογής μπορεί να γίνει και απομακρυσμένα με κινητή συσκευή android.

Περίληψη

Τα τελευταία χρόνια ολοένα και περισσότερο με τη ραγδαία ανάπτυξη της τεχνολογίας επηρεάζεται η καθημερινότητα της ανθρώπινης ζωής. Εξαιρετικά σημαντικό και κρίσιμο ρόλο έπαιξε η τεχνολογία στην αντιμετώπιση της πανδημίας του COVID 19. Είδαμε να πραγματοποιούνται μεγάλες αλλαγές ώστε να μειωθεί η διασπορά του ιού και να σωθούν εκατομμύρια ανθρώπινες ζωές. Οι αλλαγές ήταν εξαιρετικά αισθητές στη ζωή των ανθρώπων. Χάρης την τεχνολογία είδαμε τα σχολεία και πανεπιστήμια να κάνουν αποκλειστικά μαθήματα εξ' αποστάσεως και να μην χάνεται η σχολική - ακαδημαϊκή χρονιά, είδαμε τις δημόσιες υπηρεσίες να προωθούν ακόμα περισσότερο την ηλεκτρονική συναλλαγή με τους πολίτες, την ιατρική να κάνει άλματα προόδου και πολλά ακόμα που μας οδήγησε στο σήμερα.

Από την άλλη συναντάμε την τεχνητή νοημοσύνη σε καθημερινές δραστηριότητες μας, χαρακτηριστικό παράδειγμα είναι η αναγνώριση ομιλίας που τη βλέπουμε στα κινητά τηλέφωνα και τις συσκευές, όπως οι είναι οι ψηφιακοί προσωπικοί βοηθοί. Έτσι δίνοντας φωνητικές εντολές, με μεγάλη αξιοπιστία μπορεί να αναπαραχθεί το αγαπημένο σου τραγούδι, να υποδειχθεί η διαδρομή για το σπίτι, να ενημερωθούμε έγκαιρα για τα καιρικά φαινόμενα, να πραγματοποιείται η μηχανική μετάφραση κειμένων με πολύ καλή ακρίβεια, να αναγνωρίζονται οι λωρίδες του αυτοκινητόδρομου βοηθώντας τους οδηγούς να αντιληφθούν τους πεζούς, τα άλλα αυτοκίνητα και τα διάφορα αντικείμενα που θα υπάρχουν ώστε να αποτρέψουν τα τροχαία ατυχήματα. Σε έρευνες που έχουν γίνει, έχει διαπιστωθεί ότι με την τεχνολογία της μηχανικής μάθησης στην αναγνώριση αντικειμένων που χρησιμοποιείται στα αυτοκίνητα έχουν μειωθεί τα τροχαία ατυχήματα κατά 40%.

Η εργασία πραγματεύεται τη δημιουργία ενός έργου αναγνώρισης προσώπων με τη χρήση του raspberry pi 4 και την βιβλιοθήκη OpenCV. Η εφαρμογή εκτελείται σε δύο φάσεις, η πρώτη είναι η εκπαίδευση του αλγορίθμου, στην οποία μέσω της κάμερας αποθηκεύονται στιγμιότυπα από τα πρόσωπα προς αναγνώριση, να αναφέρω ότι είναι πολύ σημαντικό να έχουμε πολλά δείγματα φωτογραφιών-datasets ώστε να μπορέσει να γίνει καλύτερη "εκπαίδευση", στην δεύτερη φάση εκτελείται η εφαρμογή στην οποία μέσω της κάμερας εμφανίζεται το πρόσωπο και εκτελείται ο αλγόριθμος για την αναγνώριση του προσώπου. "Όλη αυτή η τεχνολογία κάνει χρήση των αλγορίθμων μηχανικής μάθησης και θα γίνει η ανάλυση παρακάτω.

Λέξεις κλειδιά: αναγνώριση προσώπου, μηχανική μάθηση, raspberry pi

Abstract

The paper deals with creating a face recognition project using raspberry pi 4 and OpenCV library. The application is executed in two phases, the first is the training of the system, in which snapshots of the faces to be recognized are stored through the camera, the second phase - after there are enough recordings from the first step to have a large sample of photos - the application is executed in which the face and its recognition appear through the camera. The technology use the machine learning algorithms.

Ευχαριστίες

Καθώς ήμουν εργαζόμενη όλο το διάστημα της φοίτησής μου στη σχολή και κατά τη διάρκεια της ολοκλήρωσης της πτυχιακής εργασίας η ηθική στήριξη που πήρα ήταν καθοριστική για την ολοκλήρωση όλης αυτής της πορείας.

Το να είσαι φοιτητής σε μία σχολή τριτοβάθμιας εκπαίδευσης και παράλληλα να εργάζεσαι είναι ένας μαραθώνιος, όπως είχε πει ένα καθηγητής της σχολής μου “η ιδιότητα ενός φοιτητή είναι full time job”.

Το τελευταίο φιλικό «χτύπημα» στην πλάτη μου το πήρα από την φίλη μου Αλέκα Καρακίτσου, που μου είπε «'Αντε Ελλάδα, μπουρξείς», και την ευχαριστώ πολύ γι' αυτό.

Έτσι, ευχαριστώ την οικογένειά μου, τον Θεοφύλακτο που έκανε υπομονή τόσο καιρό, φυσικά τη φίλη μου Αλέκα Καρακίτσου και την Κατερίνα Πάτκου.

Τέλος, ευχαριστώ τον επιβλέποντα καθηγητή μου, για όλη τη βοήθειά του.

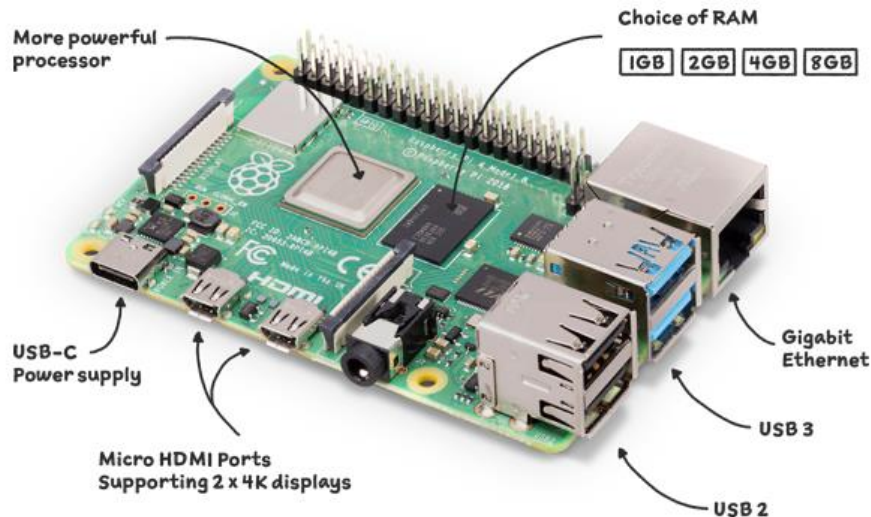
Ελλάδα Καρυπίδου

Θεσσαλονίκη 2023

Συντομογραφίες

Παρακάτω παρατίθενται τα κυριότερα ακρωνύμια της παρούσας διπλωματικής εργασίας:

Δ.Ε. Διπλωματική Εργασία
ΔΙΠΑΕ Διεθνές Πανεπιστήμιο Ελλάδος
Π.Ε. Πτυχιακή Εργασία
ΤΝ Τεχνητή Νοημοσύνη
ΑΙ Artificial Intelligence
ΝΝ Neural Network
ΜΛ Machine Learning
DL Deep Learning
RAM Random Access Memory
API Application Programming Interface
CPU Central Processing Unit
OS Operational System
Λ.Σ. Λειτουργικό Σύστημα
SSH Secure Shell
IP Internet Protocol
cmd Command Prompt
SVM Support Vector Machine



1. Εικόνα raspberry pi 4 [11]

Κεφάλαιο 1ο: Raspberry Pi 4

Raspberry Pi [11] είναι μία συστοιχία από small single-board computers (SBCs) που αναπτύχθηκαν στο Ηνωμένο Βασίλειο από τον Οργανισμό Raspberry Pi σε συνεργασία με την Broadcom [small single-board computers-SBCs - ένας υπολογιστής μονής πλακέτας είναι ένας πλήρης υπολογιστής κατασκευασμένος σε μία πλακέτα κυκλώματος, με μικροεπεξεργαστή, μνήμη, είσοδο/έξοδο και άλλα χαρακτηριστικά που απαιτούνται από έναν λειτουργικό υπολογιστή]

Το έργο Raspberry Pi αρχικά ήταν προς την προώθηση της διδασκαλίας της βασικής επιστήμης των υπολογιστών στα σχολεία και στις αναπτυσσόμενες χώρες. Έγινε αμέσως πολύ δημοφιλές το αρχικό μοντέλο

Το αρχικό μοντέλο έγινε πιο δημοφιλές από ό,τι αναμενόταν, πουλώντας εκτός της αγοράς-στόχου του για χρήσεις όπως η ρομποτική. Χρησιμοποιείται ευρέως σε πολλούς τομείς, όπως για την παρακολούθηση του καιρού. Συνήθως χρησιμοποιείται από χομπίστες υπολογιστών και ηλεκτρονικών, λόγω της υιοθέτησης των προτύπων HDMI και USB.

1.1 Ιστορία

Η ιδέα για την δημιουργία μικρού και οικονομικού υπολογιστή βρέθηκε από τους Eben Upton, Rob Mullins, Jack Lang και Alan Mycroft στο Εργαστήριο Υπολογιστών του Πανεπιστημίου του Cambridge το 2006, όταν διαπίστωσαν ότι υπήρχε αισθητή μείωση του αριθμούν και των δεξιοτήτων των φοιτητών που έκαναν αίτηση για ώστε να φοιτήσουν στο Τμήμα Επιστήμης Υπολογιστών. Αρκετές εκδόσεις των πρώιμων πρωτοτύπων Raspberry Pi σχεδιάστηκαν αλλά ήταν πολύ περιορισμένες λόγω του υψηλού κόστους και της χαμηλής ισχύος επεξεργαστών για κινητές συσκευές εκείνη την εποχή.

Το 2008, η ομάδα ξεκίνησε μια συνεργασία με τον Pete Lomas, MD της Norcott Technologies και τον David Braben, και δημιούργησε το Raspberry Pi Foundation. Τρία χρόνια αργότερα, γεννήθηκε το Raspberry Pi Model B και είχε πουλήσει πάνω από δύο εκατομμύρια μονάδες μέσα σε δύο χρόνια μαζικής παραγωγής.

Στις αρχές του 2013 ο οργανισμός χωρίστηκε σε δύο μέρη: το Raspberry Pi Foundation, το οποίο είναι υπεύθυνο για τις φιλανθρωπικές και εκπαιδευτικές δραστηριότητες, και Raspberry Pi Trading Ltd, υπεύθυνη για τις μηχανολογικές και εμπορικές δραστηριότητες. Η Raspberry Pi (Trading) Ltd είναι μια εξ ολοκλήρου θυγατρική του Raspberry Pi Foundation, με τα χρήματα που κερδίζονται από τις πωλήσεις των προϊόντων Raspberry Pi να χρησιμοποιούνται για τη χρηματοδότηση του φιλανθρωπικού έργου του Ιδρύματος. Το 2015 το Raspberry Pi Foundation απέκτησε το Code Club. Το 2017 το Raspberry Pi Foundation εξαγόρασε την CoderDojo.

1.2 Λειτουργικό Σύστημα -Raspberry Pi OS

Το Raspberry Pi OS [11] (πρώην Raspbian) είναι ένα λειτουργικό σύστημα που βασίζεται στο Debian. Από το 2013, παρέχεται επίσημα από το Ίδρυμα Raspberry Pi ως το κύριο λειτουργικό σύστημα για την οικογένεια συμπαγών υπολογιστών μιας πλακέτας Raspberry Pi.

Το Raspberry Pi OS αναπτύχθηκε για πρώτη φορά από τους Mike Thompson και Peter Green ως Raspbian. Η πρώτη έκδοση κυκλοφόρησε στις 15 Ιουλίου 2012. Καθώς το Raspberry Pi δεν είχε επίσημα παρεχόμενο λειτουργικό σύστημα εκείνη την εποχή, το Ίδρυμα Raspberry Pi αποφάσισε να βασιστεί στη δουλειά που έκανε το έργο Raspbian και άρχισε να παράγει και να κυκλοφορεί τη δική του έκδοση του λογισμικού. Η πρώτη κυκλοφορία του Raspbian από το Ίδρυμα, η οποία τώρα αναφερόταν τόσο στο κοινοτικό έργο όσο και στο επίσημο λειτουργικό σύστημα, ανακοινώθηκε στις 10 Σεπτεμβρίου 2013.

Το Raspberry Pi OS είναι εξαιρετικά βελτιστοποιημένο για τη σειρά συμπαγών υπολογιστών μονής πλακέτας Raspberry Pi με επεξεργαστές ARM. Λειτουργεί σε κάθε Raspberry Pi εκτός από τον μικροελεγκτή Pico. Το Raspberry Pi OS χρησιμοποιεί ένα τροποποιημένο LXDE ως περιβάλλον επιφάνειας εργασίας με τη διαχείριση παραθύρων στοίβαξης Openbox, μαζί με ένα μοναδικό θέμα. Η προεπιλεγμένη διανομή αποστέλλεται με ένα αντίγραφο του συστήματος άλγεβρας υπολογιστή Wolfram Mathematica, VLC και μια ελαφριά έκδοση του προγράμματος περιήγησης ιστού Chromium.

Το Raspberry Pi OS έχει ένα περιβάλλον επιφάνειας εργασίας, το PIXEL, που βασίζεται στο LXDE, το οποίο μοιάζει με πολλούς κοινούς επιτραπέζιους υπολογιστές, όπως το macOS και τα Microsoft Windows. Η επιφάνεια εργασίας έχει μια εικόνα φόντου. Μια γραμμή μενού βρίσκεται στην κορυφή και περιέχει ένα μενού εφαρμογής και συντομεύσεις για ένα πρόγραμμα περιήγησης ιστού (Chromium), τη διαχείριση αρχείων και το τερματικό. Το άλλο άκρο της γραμμής μενού εμφανίζει ένα μενού Bluetooth, μενού Wi-Fi, έλεγχο έντασης και ρολόι. Η επιφάνεια εργασίας μπορεί επίσης να αλλάξει από την προεπιλεγμένη εμφάνισή της, όπως επανατοποθέτηση της γραμμής μενού.

Το λειτουργικό σύστημα Raspberry Pi έχει τρεις εκδόσεις εγκατάστασης:

- Raspberry Pi OS Lite (32-bit & 64-bit)
- Λειτουργικό σύστημα Raspberry Pi με επιτραπέζιο υπολογιστή (32 bit & 64 bit)
- Λειτουργικό σύστημα Raspberry Pi με επιτραπέζιο υπολογιστή και προτεινόμενο λογισμικό (32 bit)

1.3 Debian Λειτουργικό Σύστημα



2. Λογότυπο του Debian λειτουργικού συστήματος [12]

Το Debian [12] επίσης γνωστό ως Debian GNU/Linux, είναι μια διανομή Linux που αποτελείται από ελεύθερο λογισμικό ανοιχτού κώδικα.

Η πρώτη έκδοση του Debian (0.01) κυκλοφόρησε στις 15 Σεπτεμβρίου 1993 και η πρώτη σταθερή έκδοση (1.1) κυκλοφόρησε στις 17 Ιουνίου 1996. Ο κλάδος του Debian Stable είναι η πιο δημοφιλής έκδοση για προσωπικούς υπολογιστές και διακομιστές. Το Debian είναι επίσης η βάση για πολλές άλλες διανομές, κυρίως το Ubuntu. Το Debian είναι ένα από τα παλαιότερα λειτουργικά συστήματα που βασίζονται στον πυρήνα του Linux. Το έργο συντονίζεται μέσω του Διαδικτύου από μια ομάδα εθελοντών. Από την ίδρυσή του, το Debian έχει αναπτυχθεί ανοιχτά και διανέμεται ελεύθερα.

1.4 Συναρμολόγηση και εγκατάσταση λειτουργικού συστήματος.

Για την υλοποίηση και ολοκλήρωση της εργασίας ήταν απαραίτητη η αγορά του raspberry pi και των περιφερειακών του όπως είναι η κάμερα. Η αγορά έγινε από εμένα, μέσω του επίσημη αντιπροσώπου του raspberry pi στην Ελλάδα NETTOP [25]

Θα κάνω μία μικρή παρένθεση και θα αναφέρω ότι όλη αυτή η διαδικασία της συναρμολόγησης του raspberry pi, εγκατάστασης λειτουργικού συστήματος κλπ ώστε να είναι πλήρως λειτουργικό ήταν μία εργασία που απαιτούσε αρκετό χρόνο (το raspberry pi συναρμολογήθηκε από την αρχή, ακόμα και ο ανεμιστήρας έπρεπε να βιδωθεί).

Από την επίσημη ιστοσελίδα του raspberry pi έγινε τη λήψη του εκτελέσιμου για την εγκατάσταση του Λ.Σ (<https://www.raspberrypi.com/software/>).

Τα υλικά που αγοράστηκαν ήταν:

- [Raspberry Pi 4 Model B/4GB](#)
- [Raspberry Pi 4 Official μετασχηματιστής στα 3A \(15,3W\) USB-C \(μαύρος\)](#)
- [Raspberry Pi 8MP Camera board v2](#)
- [Μαύρο - 1m Official καλώδιο για το Raspberry Pi 4 από Micro-HDMI to Standard HDMI \(type A\)](#)
- [Raspberry Pi 4 Heatsink/Ψυκτρα 40x30x5mm](#)
- [Integral microSDHC memory card 16 GB MSDHC UHS-1 U1 CL10 V10 A1 with adapter \(Raspberry Pi recommended\)](#)

Κεφάλαιο 2ο: Βιβλιοθήκη OpenCV -(Open Source Computer Vision Library)



3. Εικόνα με το λογότυπο του OpenCV [13]

Το OpenCV- (Βιβλιοθήκη Υπολογιστικής Όρασης Ανοιχτού Κώδικα) [13] είναι μια βιβλιοθήκη λογισμικού υπολογιστικής όρασης ανοιχτού κώδικα και μηχανικής μάθησης. Το OpenCV δημιουργήθηκε για να παρέχει μια κοινή υποδομή για εφαρμογές υπολογιστικής όρασης και να επιταχύνει τη χρήση της μηχανικής αντίληψης στα εμπορικά προϊόντα. Όντας ένα προϊόν με άδεια χρήσης Apache 2, το OpenCV διευκολύνει τις επιχειρήσεις να χρησιμοποιούν και να τροποποιούν τον κώδικα.

Η βιβλιοθήκη διαθέτει περισσότερους από 2500 βελτιστοποιημένους αλγόριθμους, οι οποίοι περιλαμβάνουν ένα ολοκληρωμένο σύνολο κλασικών και υπερσύγχρονων αλγορίθμων υπολογιστικής όρασης και μηχανικής μάθησης. Αυτοί οι αλγόριθμοι μπορούν να χρησιμοποιηθούν για την ανίχνευση και αναγνώριση προσώπων, την αναγνώριση αντικειμένων, την ταξινόμηση των ανθρώπινων ενεργειών σε βίντεο, την παρακολούθηση κινήσεων της κάμερας, την παρακολούθηση κινούμενων αντικειμένων, την εξαγωγή τρισδιάστατων μοντέλων αντικειμένων, την παραγωγή τρισδιάστατων σημείων σύννεφων από στερεοφωνικές κάμερες, τη συρραφή εικόνων για παραγωγή υψηλής ανάλυσης εικόνα μιας ολόκληρης σκηνής, εύρεση παρόμοιων εικόνων από βάση δεδομένων εικόνων, αφαίρεση κόκκινων ματιών από εικόνες που τραβήχτηκαν με φλας, ακολουθήστε τις κινήσεις των ματιών, αναγνωρίστε το τοπίο και δημιουργήστε δείκτες για να το επικαλύψετε με επαυξημένη πραγματικότητα κ.λπ. Το OpenCV έχει περισσότερους από 47 χιλιάδες χρήστες κοινότητα και ο εκτιμώμενος αριθμός λήψεων υπερβαίνει τα 18 εκατομμύρια. Η βιβλιοθήκη χρησιμοποιείται εκτενώς σε εταιρείες, ερευνητικές ομάδες και από κυβερνητικούς φορείς.

Μαζί με καθιερωμένες εταιρείες όπως η Google, η Yahoo, η Microsoft, η Intel, η IBM, η Sony, η Honda, η Toyota που χρησιμοποιούν τη βιβλιοθήκη, υπάρχουν πολλές νεοφυείς επιχειρήσεις όπως οι Applied Minds, VideoSurf και Zeitera, που κάνουν εκτενή χρήση του OpenCV. Οι εφαρμογές του OpenCV εκτείνονται από τη συρραφή εικόνων streetview, την ανίχνευση εισβολών σε βίντεο παρακολούθησης στο Ισραήλ, την παρακολούθηση εξοπλισμού ορυχείων στην Κίνα, τη βοήθεια ρομπότ στην πλοήγηση και την παραλαβή αντικειμένων στο Willow Garage, τον εντοπισμό ατυχημάτων πνιγμού σε πισίνες στην Ευρώπη, τη λειτουργία διαδραστικής τέχνης σε Ισπανία και Νέα Υόρκη, έλεγχος διαδρόμων για συντρίμια στην Τουρκία, επιθεώρηση ετικετών σε προϊόντα σε εργοστάσια σε όλο τον κόσμο για την ταχεία ανίχνευση προσώπου στην Ιαπωνία. Διαθέτει διεπαφές C++, Python, Java και MATLAB και υποστηρίζει Windows, Linux, Android και Mac OS. Το OpenCV κλίνει κυρίως σε εφαρμογές όρασης σε πραγματικό χρόνο και εκμεταλλεύεται τις οδηγίες MMX και SSE όταν είναι διαθέσιμες. Μια πλήρως

εξοπλισμένη διεπαφή CUDA και OpenCL αναπτύσσεται ενεργά αυτήν τη στιγμή. Υπάρχουν πάνω από 500 αλγόριθμοι και περίπου 10 φορές περισσότερες συναρτήσεις που συνθέτουν ή υποστηρίζουν αυτούς τους αλγόριθμους. Το OpenCV είναι γραμμένο εγγενώς σε C++ και έχει ένα πρότυπο διεπαφής που λειτουργεί άψογα με κοντύτερα STL.

Κεφάλαιο 3ο: Λειτουργικό Σύστημα Android



4. Λογότυπο του λειτουργικού συστήματος Android [21]

Το Android [21] είναι λειτουργικό σύστημα για συσκευές κινητής τηλεφωνίας το οποίο τρέχει τον πυρήνα του λειτουργικού Linux. Αρχικά αναπτύχθηκε από την Google και αργότερα από την Open Handset Alliance. Επιτρέπει στους κατασκευαστές λογισμικού να συνθέτουν κώδικα με την χρήση της γλώσσας προγραμματισμού Java, ελέγχοντας την συσκευή μέσω βιβλιοθηκών λογισμικού ανεπτυγμένων από την Google. Το Android είναι κατά κύριο λόγο σχεδιασμένο για συσκευές με οθόνη αφής, όπως τα έξυπνα τηλέφωνα και τα τάμπλετ, με διαφορετικό περιβάλλον χρήσης για τηλεοράσεις, αυτοκίνητα και ρολόγια χειρός. Παρόλο που έχει αναπτυχθεί για συσκευές με οθόνη αφής, έχει χρησιμοποιηθεί σε κονσόλες παιχνιδιών, ψηφιακές φωτογραφικές μηχανές, συνηθισμένους Η/Υ και σε άλλες ηλεκτρονικές συσκευές. Το Android είναι το πιο ευρέως διαδεδομένο λογισμικό στον κόσμο.

Η πρώτη παρουσίαση της πλατφόρμας Android έγινε στις 5 Νοεμβρίου 2007. Το λογότυπο για το λειτουργικό σύστημα Android είναι ένα ρομπότ σε χρώμα πράσινου μήλου και σχεδιάστηκε από τη γραφίστρια Ιρίνα Μπλοκ. Η τρέχουσα έκδοση είναι η Android 12L, η οποία κυκλοφόρησε στις 7 Μαρτίου του 2022. Όπως φαίνεται στον παρακάτω πίνακα, ορισμένες από τις εκδόσεις Android έχουν ονόματα γλυκού.

3.1 Εκδόσεις του Android [22]

- 11
- 10
- Pie
- Oreo

- Nougat
- Marshmallow
- Lollipop
- KitKat
- Jelly Bean

Κεφάλαιο 4ο: Γλώσσα προγραμματισμού Python

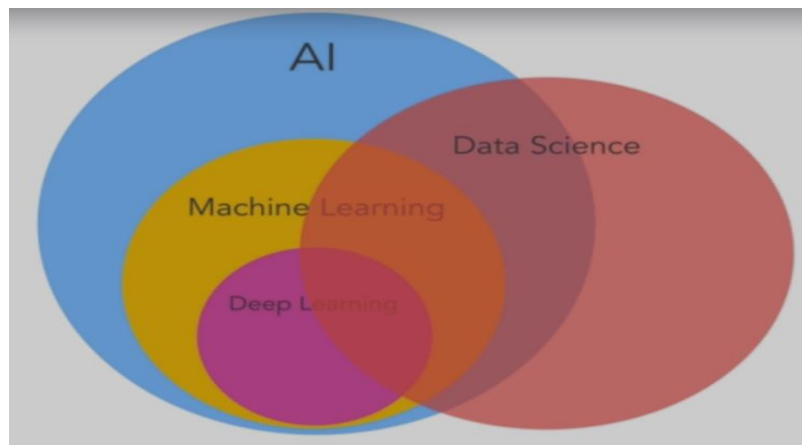
Η Python [20] είναι μια γλώσσα προγραμματισμού υπολογιστών που χρησιμοποιείται συχνά για την κατασκευή ιστοσελίδων και λογισμικού, την αυτοματοποίηση εργασιών και τη διεξαγωγή ανάλυσης δεδομένων. Η Python είναι μια γλώσσα γενικής χρήσης, που σημαίνει ότι μπορεί να χρησιμοποιηθεί για τη δημιουργία μιας ποικιλίας διαφορετικών προγραμμάτων και δεν είναι εξειδικευμένη για συγκεκριμένα προβλήματα. Αυτή η ευελιξία, μαζί με τη φιλικότητα προς τους αρχάριους, την έχει καταστήσει μία από τις πιο χρησιμοποιούμενες γλώσσες προγραμματισμού σήμερα. Είναι μία από τις πιο δημοφιλής γλώσσες προγραμματισμού.

Η Python είναι δημοφιλής για διάφορους λόγους. Έχει μια απλή σύνταξη που μιμείται τη φυσική γλώσσα, επομένως είναι πιο εύκολο να διαβαστεί και να κατανοηθεί. Αυτό καθιστά πιο γρήγορη την κατασκευή έργων και ταχύτερη τη βελτίωσή τους. Είναι ευέλικτο. Η Python μπορεί να χρησιμοποιηθεί για πολλές διαφορετικές εργασίες, από την ανάπτυξη ιστού έως τη μηχανική εκμάθηση. Είναι φιλικό προς τους αρχάριους, καθιστώντας το δημοφιλές για κωδικοποιητές αρχικού επιπέδου. Είναι ανοιχτού κώδικα, που σημαίνει ότι είναι δωρεάν για χρήση και διανομή, ακόμη και για εμπορικούς σκοπούς. Το αρχείο λειτουργιών και βιβλιοθηκών της Python —πακέτα κώδικα που έχουν δημιουργήσει τρίτοι χρήστες για να επεκτείνουν τις δυνατότητες της Python—είναι τεράστιο και αυξάνεται. Η Python έχει μια μεγάλη και ενεργή κοινότητα που συνεισφέρει στη δεξαμενή λειτουργιών και βιβλιοθηκών της Python και λειτουργεί ως χρήσιμος πόρος για άλλους προγραμματιστές. Η τεράστια κοινότητα υποστήριξης σημαίνει ότι εάν οι κωδικοποιητές βρεθούν σε εμπόδιο, η εύρεση λύσης είναι σχετικά εύκολη. κάποιος είναι βέβαιο ότι έχει αντιμετωπίσει το ίδιο πρόβλημα πριν.

4.1 Χρήσεις της γλώσσας προγραμματισμού python [20]

- Ανάλυση Δεδομένων και Μηχανική Μάθηση
- Ανάπτυξη web εφαρμογών, ιστοσελίδων (Web development)
- Αυτοματοποίηση (Automation)
- Δοκιμή λογισμικού (software testing)

Κεφάλαιο 5ο: Τεχνητή Νοημοσύνη



5. Τεχνητής Νοημοσύνης, της Μηχανικής Μάθησης και του Deep Learning [23]

Στόχος αυτού του κεφαλαίου είναι η εξερεύνηση της έννοιας του τεχνητού νευρωνικού δικτύου (Artificial Neural Network)-AI ή απλούστερα Neural Network, θα κάνω μια εισαγωγή στις έννοιες της τεχνητής νοημοσύνης-AI, της μηχανικής μάθησης-ML, της βαθιάς μηχανικής μάθησης-DL καθώς και μια σύντομη ιστορική αναδρομή.

ο Δαρβίνος μας έδειξε και μας έμαθε πως φτάσαμε ως εδώ στον δρόμο της εξέλιξης. Μία μεγάλη διαφορά του ανθρώπου με τα υπόλοιπα ζώα είναι η νοημοσύνη και η συνείδηση του. Τίθεται το ερώτημα εάν τελικά έχει μπει και ο τελευταίος κρίνος στην εξέλιξη. Σε αντίθεση με τα προηγούμενα βήματα, ίσως το επόμενο βήμα της εξέλιξης να γίνει από τον άνθρωπο χρησιμοποιώντας την τεχνολογία. Η σύγχρονη τεχνολογία έχει τη δυνατότητα να μας δώσει προσθετικά μέλη, συνθετικό δέρμα ή συνθετικά όργανα και να γίνουμε βιονική. Παρόλα αυτά δεν είναι εύκολο να φανταστούμε ότι οι υπολογιστές θα μπορούσαν να φτάσουν σε σημείο που θα αντικαταστήσουν το μυαλό του ανθρώπου. Αξιοσημείωτο είναι το 1997 που για πρώτη φορά οι υπολογιστές νίκησαν τον άνθρωπο πρωταθλητή του σκάκι Gary Kasparov (vs deep blue), Αντίστοιχα έχουν υπάρξει και άλλες φορές που οι υπολογιστές έχουν κερδίσει σε παιχνίδια που υπήρχε ακόμα και ο παράγοντας της τύχης. Όλα αυτά είναι παιχνίδια, όμως τι συμβαίνει με άλλα πιο σημαντικά και πιο σοβαρά θέματα;

Η τεχνητή νοημοσύνη έχει ενσωματωθεί στην καθημερινότητα μας, χαρακτηριστικό παράδειγμα είναι η αναγνώριση ομιλίας που τη συναντούμε στα κινητά τηλέφωνα, όπως οι είναι οι ψηφιακοί προσωπικοί βοηθοί - εφαρμογές αναγνώρισης φωνής και εντολών-. Ένας ψηφιακός βοηθός δίνοντας φωνητικές εντολές με μεγάλη αξιοπιστία μπορεί να παίζει το αγαπημένο μας τραγούδι, να μας δείξει τη διαδρομή για το σπίτι, ή και να μας προτείνει εναλλακτικό δρόμο στην περίπτωση κυκλοφοριακής συμφόρησης, να μας ενημερώσει για τον καιρό, να κάνει τη μηχανική μετάφραση κειμένων, να αναγνωρίσει τις λωρίδες του αυτοκινητόδρομου, να αναγνωρίσει τους πεζούς, τα άλλα αυτοκίνητα και τα διάφορα αντικείμενα που θα υπάρχουν ώστε να αποτρέψουν τα τροχαία ατυχήματα. (Σε έρευνες που έχουν γίνει, έχει διαπιστωθεί ότι με την τεχνολογία της τεχνητής νοημοσύνης χρησιμοποιείται στα αυτοκίνητα έχουν μειωθεί τα τροχαία ατυχήματα κατά 40%). Η λίστα δεν σταματά αλλά ολοένα και αυξάνεται. Ορισμένοι πολύ γνωστοί ψηφιακοί βοηθοί είναι η CORTANA της Microsoft, το SIRI της Apple και το Alexa της Amazon.

Η Τεχνητή Νοημοσύνη μπορεί να μοιάζει με έναν σύγχρονο όρο όμως τα θεμέλια για τα σημαντικά κατορθώματα της, ξεκίνησαν αρκετά χρόνια πριν. Ήδη από την αρχαιότητα στα θεμέλια της πρακτικής προσέγγισης της Τεχνητής Νοημοσύνης έκαναν συνεισφορά οι Συλλογισμοί του Αριστοτέλη (322-384 π.Χ.).

Ο τομέας της τεχνητής νοημοσύνης είναι ένας πολύ ευρύς τομέας όπως είναι η ανθρώπινη νοημοσύνη. Αναφέρεται στην ανάπτυξη υπολογιστικών συστημάτων τα οποία είναι ικανά να εκτελούν εργασίες που απαιτούν ανθρώπινη νοημοσύνη, όπως είναι η οπτική αντίληψη, η λήψη αποφάσεων και η αναγνώριση φωνής. Η μηχανική μάθηση αναφέρεται στον υποτομέα της τεχνητής νοημοσύνης όπου χρησιμοποιούνται αλγόριθμοι οι οποίοι καθιστούν δυνατή την "εκμάθηση" του υπολογιστή απευθείας από δεδομένα. Στη μηχανική μάθηση οι αλγόριθμοι που χρησιμοποιούνται είναι ικανοί να κάνουν τον υπολογιστή να μπορεί να γενικεύει από ένα σύνολο δεδομένων/παραδειγμάτων. Αυτό σημαίνει ότι δίνοντας αρκετά παραδείγματα στον αλγόριθμο, θα μπορεί να μάθει ποιες είναι οι σωστές ενέργειες από μόνο του χωρίς να χρειάζεται να μπούμε στη διαδικασία προγραμματισμού σαφών εντολών ανάλογα με τις συνθήκες.

Πολύ χαρακτηριστικές εφαρμογές μηχανικής μάθησης είναι σε ιατρικές διαγνώσεις, όπου από ένα σύνολο μετρήσεων ενός ασθενή μπορούμε να αποφανθούμε αν πάσχει από κάποια συγκεκριμένη ασθένεια σύμφωνα με δεδομένα που έχουμε από πραγματικούς ασθενείς. Το Deep Learning πρόκειται για το αποκορύφωμα των δυνατοτήτων των υπολογιστών στον τομέα της τεχνητής νοημοσύνης μέχρι στιγμής. Οι αλγόριθμοι που χρησιμοποιούνται στο Deep Learning έχουν ξεπεράσει κατά πολύ σε απόδοση παλαιότερες τεχνικές πάνω σε εφαρμογές αναγνώρισης εικόνας, κειμένου, φωνής και μας έχουν δώσει τη δυνατότητα να δημιουργήσουμε νέες, πολύ ενδιαφέρουσες, εφαρμογές οι οποίες πριν 15 χρόνια θεωρούνταν αδύνατες όπως αυτόνομα οχήματα που οδηγούνται μόνο τους, μετάφραση γλώσσας από ομιλία σε πραγματικό χρόνο, συστήματα τα οποία μπορούν να νικήσουν εύκολα παγκόσμιους πρωταθλητές σε παιχνίδια όπως σκάκι.

Αυτό που ξεχωρίζει το Deep Learning από το Machine Learning είναι η χρήση της έννοιας του τεχνητού νευρωνικού δικτύου. Η ιδέα των τεχνητών νευρωνικών δικτύων έχει εμπνευστεί από τους πραγματικούς νευρώνες του ανθρώπινου εγκεφάλου και τη δομή του εγκεφαλικού φλοιού. Τα τεχνητά νευρωνικά δίκτυα είναι μια προσέγγιση/μοντελοποίηση των πραγματικών νευρωνικών δικτύων που υπάρχουν στον ανθρώπινο εγκέφαλο. Στη μοντελοποίηση υπάρχουν στρώματα (layers) νευρώνων και η είσοδος του μοντέλου είναι το πρώτο στρώμα (input layer) ενώ η έξοδος είναι το τελευταίο στρώμα (output layer). Τα δεδομένα εισέρχονται στο input layer και η λειτουργία των μεσαίων στρωμάτων (hidden layers) είναι να εξακριβώσουν τη συνάρτηση μεταξύ της εισόδου και της εξόδου. Η κύρια διαφορά μεταξύ του μοντέλου αυτού και των πραγματικών νευρωνικών δικτύων είναι ότι στον ανθρώπινο εγκέφαλο όλο το σύστημα είναι ασύγχρονο, υπό την έννοια του ότι οι νευρώνες "ενεργοποιούνται" αυτόνομα και δεν λειτουργούν βάση των νευρωνικών στρωμάτων στα οποία ανήκουν.

Η πρόσφατες εξελίξεις στον τομέα της μηχανικής όρασης είναι μια από τις πιο σημαντικές και πιθανότατα επαναστατικές τεχνολογίες της επιστήμης της πληροφορικής. Ήδη είναι γεγονός τα αυτοκινούμενα οχήματα, δηλαδή τα αυτοκίνητα που κινούνται αυτόνομα, αλλά χωρίς την "έξυπνη" όραση δεν μπορούσε να είναι εφικτή η αντίληψη της διαφοράς μεταξύ μίας τσαλακωμένης πλαστικής σακούλας στον δρόμο, που ένα αυτοκίνητο μπορεί να πατήσει και να συνεχίσει το δρόμο του χωρίς να χρειαστεί να σταματήσει ή μιας πέτρας με το ίδιο μέγεθος που πρέπει να αποφύγει το αμάξι καθώς θα του προκαλέσει ζημιές και ατυχήματα.

Οι φωτογραφίες και τα βίντεο ολόένα και περισσότερο γίνονται πιο σημαντικά κομμάτια της παγκόσμιας ζωής. Παρατηρούμε λοιπόν, ότι ενώ υπάρχει η τεχνολογία με την οποία μπορούμε να παρακολουθούμε εξαιρετικής ποιότητας φωτογραφίες και βίντεο ακόμα και τα πιο προηγμένα

λογισμικά δυσκολεύονται ακόμα να “καταλάβουν” και να διαχειριστούν αυτές τις τεράστιες ποσότητες περιεχομένου. Οι πιο έξυπνες μηχανές μας είναι “τυφλές” βλέπουν αλλά δεν αντιλαμβάνονται κάτι που ολοένα βελτιώνεται χρησιμοποιώντας την τεχνητή νοημοσύνη. Είναι γεγονός ότι τα μάτια από μόνα τους δεν βλέπουν, αλλά μεταφέρουν την εικόνα στον εγκέφαλο του ανθρώπου και εκεί γίνεται η επεξεργασία και η αντίληψη της εικόνας. Καταλήγοντας, η μηχανική μάθηση μπορεί να παίξει εξαιρετικά σημαντικό ρόλο στην όραση και την αντίληψη μέσα από φωτογραφίες και βίντεο, σύμφωνα με τους επιστήμονες, το κλειδί δεν είναι μόνο η βελτίωση των αλγορίθμων αλλά και η ύπαρξη πολύ μεγάλων ποσοτήτων δεδομένων, έτσι θα είναι εφικτή σωστή διάγνωση και αντίληψη συναισθημάτων, αντικειμένων και προθέσεων, όπως ακριβώς ένα μικρό παιδί μαθαίνει μέσα από εμπειρίες του τα παραδείγματα του πραγματικού κόσμου.

Το επιστημονικό πεδίο της τεχνητής νοημοσύνης είναι αρκετά μεγάλο και πέρα από το αντικείμενο της παρούσας εργασίας, εδώ θα μείνουμε μόνο στο κομμάτι της υλοποίησης της ταυτοποίησης και αναγνώρισης προσώπου.

5.1 Ορισμός TN

Στο ερώτημα «Τι είναι Τεχνητή Νοημοσύνη;» οι ερευνητές του χώρου δίνουν πολλές διαφορετικές απαντήσεις, φαινόμενο που δεν απαντά σε άλλους επιστημονικούς χώρους, όπως η Φυσική, η Χημεία, η Ιατρική κ.ά. Ωστόσο, όλοι φαίνεται να συμφωνούν πως η TN είναι επιστήμη και όχι απλώς ένας κλάδος της τεχνολογίας λογισμικού.

Κατά τον Patrick Winston (1992), διευθυντής του εργαστηρίου TN του Πανεπιστημίου MIT, πρωταρχικός σκοπός της TN είναι «να κάνει τις μηχανές πιο έξυπνες» σε αυτό συμφωνούν οι περισσότεροι από τους ερευνητές που αντιμετωπίζουν την TN ως αναζήτηση μεθόδων οι οποίες θα κάνουν τους ηλεκτρονικούς υπολογιστές πιο έξυπνους και, συνεπώς, πιο χρήσιμους από όσο είναι σήμερα .

Όμως η TN έχει και πρακτικούς και επιστημονικούς στόχους.

Οι πρακτικοί στόχοι αποβλέπουν στο να επιλυθούν πραγματικά προβλήματα της ανθρώπινης επιβίωσης, με τη χρήση ιδεών της TN σχετικών με την αναπαράσταση και τη χρήση της απαραίτητης γνώσης. Ένας καλά διατυπωμένος ορισμός, που καλύπτει τους πρακτικούς στόχους της TN και, πιο συγκεκριμένα, του κλάδου της TN που καλείται Υπολογιστική Τεχνητή Νοημοσύνη, είναι ο ακόλουθος:

«Τεχνητή Νοημοσύνη είναι εκείνος ο κλάδος της επιστήμης των υπολογιστών που ασχολείται με το σχεδιασμό ευφών υπολογιστικών συστημάτων, δηλαδή συστημάτων με χαρακτηριστικά τα οποία σχετίζονται με την ευφυΐα στην ανθρώπινη συμπεριφορά (μάθηση, αιτίαση, επίλυση προβλημάτων, κατανόηση φυσικής γλώσσας, αναγνώριση αντικειμένων κτλ.)»

Παρόμοιος είναι και ο ακόλουθος ορισμός: «Τεχνητή Νοημοσύνη είναι η μελέτη τού πώς να κάνουμε τους υπολογιστές ικανούς να κάνουν πράγματα στα οποία προς το παρόν οι άνθρωποι τα καταφέρνουν καλύτερα» (Rich & Knight, 1990).

Από την πλευρά της, η TN μπορεί να προσφέρει στο χώρο της Γνωστικής Επιστήμης μεθοδολογίες ανάπτυξης υπολογιστικών συστημάτων που υλοποιούν ή ακόμα δίνουν και λύσεις σε θεωρητικά μοντέλα απαραίτητα για την προώθηση της έρευνας και τη διαμόρφωση νέων θεωριών σχετικά με τη λειτουργία της ανθρώπινης νόησης. Δηλαδή, η TN ρίχνει νέο φως σε παραδοσιακά ερωτήματα των ψυχολόγων, γλωσσολόγων και φιλοσόφων.

Με βάση τα παραπάνω, ένας πληρέστερος ορισμός της TN είναι ο ακόλουθος:

«Τεχνητή Νοημοσύνη είναι η επιστήμη που επιδιώκει να κατασκευάσει μηχανές οι οποίες όχι μόνο θα επιδεικνύουν ανθρώπινη συμπεριφορά, αλλά θα μπορούν, επίσης, να προσαρμόζονται στο περιβάλλον τους με τρόπο παρόμοιο με αυτόν των ανθρώπων».

Οι επιστημονικοί στόχοι αποβλέπουν στο να καθοριστούν ποιες από τις υπάρχουσες ιδέες περί αναπαράστασης και χρήσης της γνώσης είναι σε θέση να δώσουν απαντήσεις στο αιώνιο φιλοσοφικό ερώτημα:

«τι είναι ευφυΐα και πώς αυτή εκφράζεται;»

Τα μοντέλα που χρησιμοποιεί η ΤΝ, με βάση οποιονδήποτε από τους δυο παραπάνω ορισμούς, στηρίζονται στη χρήση πολύπλοκων ηλεκτρονικών συστημάτων ως μέσου υλοποίησης. Το φαινόμενο αυτό εξηγείται βάσει της θεμελιώδους παράδοσης της δυτικής φιλοσοφίας ότι η νοητική ικανότητα (σκέψη) του ανθρώπου είναι στην ουσία ένας λογικός χειρισμός νοητικών συμβόλων, δηλαδή ιδεών. Ο ηλεκτρονικός υπολογιστής, σε αντίθεση με άλλα μηχανικά κατασκευάσματα (π.χ. ρολόι), μπορεί να χειριστεί σύμβολα σε μορφή «χαρακτήρων», αφού πρώτα προγραμματιστεί κατάλληλα. Σύμφωνα με τη θεωρία χειρισμού συμβόλων, η νοημοσύνη εξαρτάται μόνο από την οργάνωση ενός συστήματος και τη λειτουργία του ως χειριστή συμβόλων και όχι από το υλικό κατασκευής των συμβόλων ή την ακριβή μορφή τους. Συνάγεται, λοιπόν, το συμπέρασμα πως η σύγχρονη τεχνολογία των υπολογιστών είναι προς το παρόν η αρμόζουσα, ώστε αυτοί να αποτελέσουν μοντέλα με δυνατότητες επίδειξης κάποιας μορφής τεχνητής νοημοσύνης. Ωστόσο, τίποτα δεν αποκλείει στο μέλλον κάποια άλλη τεχνολογία να αποδειχθεί καταλληλότερη για τον παραπάνω σκοπό.

Η παραπάνω θεώρηση οδηγεί στον ορισμό της λεγόμενης συμβολικής Τεχνητής Νοημοσύνης (symbolic Artificial Intelligence):

«Συμβολική Τεχνητή Νοημοσύνη είναι η επιστήμη που μελετά τη φύση της ανθρώπινης νοημοσύνης και στη συνέχεια τον τρόπο αναπαραγωγής της σε υπολογιστές με τη χρήση συμβόλων».

5.2 Σύντομη ιστορική αναδρομή

Οι απαρχές της Τεχνητής Νοημοσύνης ανάγονται στους «συλλογισμούς» του Αριστοτέλη (384-322 π.Χ.), οι οποίοι παρείχαν πρότυπα εκφράσεων που έδιναν πάντα σωστά συμπεράσματα από σωστές υποθέσεις (Αριστοτέλεια συλλογιστική). Στους νεότερους χρόνους, οι πρώτες σημαντικές στιγμές είναι το 1854, οπότε ο George Boole έθεσε τις βάσεις της προτασιακής λογικής, και το 1879, οπότε ο Gottlob Frege πρότεινε ένα σύστημα αυτοματοποιημένης συλλογιστικής και έθεσε τις βάσεις του κατηγορηματικού λογισμού (predicate calculus).

- Η πρώτη [5] μαθηματική μοντελοποίηση νευρωνικών δικτύων έγινε το 1943 από τους Walter Pitts και Warren McCulloch. Ο πρώτος ήταν λογιστής ενώ ο δεύτερος ήταν νευροεπιστήμονας και ο στόχος τους ήταν να δημιουργήσουν ένα συνδυασμό μαθηματικών μοντέλων και αλγορίθμων που να μιμούνται τον τρόπο με τον οποίο σκέφτονται οι άνθρωποι.
- Το 1950 υπήρξε η πρώτη πρόβλεψη για τη δημιουργία του τομέα του machine learning από τον διάσημο Alan Turing. Πρακτικά ανέφερε ότι είναι αναπόφευκτη η δημιουργία υπολογιστών που θα μαθαίνουν από δεδομένα.
- Το 1952 δημιουργήθηκε το πρώτο πρόγραμμα μηχανικής μάθησης από τον Arthur Samuel, που ήταν ένας μηχανικός στην αμερικανική εταιρία IBM. Το πρόγραμμα αυτό μάθαινε να παίζει ντάμα και σε κάθε παιχνίδι γινόταν όλο και καλύτερο. ΜΗ

- Το 1957, ο ψυχολόγος Frank Rosenblatt έστησε τα θεμέλια των deep neural networks (DNN), δηλαδή νευρωνικά δίκτυα με πολλά στρώματα, μέσω του δημοφιλούς paper “The Perceptron: A Perceiving and Recognizing Automaton”. Δημιούργησε τον perceptron, τον πρώτο αλγόριθμο Machine Learning όπου μπορούσε να μάθει να αναγνωρίζει μοτίβα σε δεδομένα ήχου, ηλεκτρικά και οπτικά.
- Το 1959 έγινε η ανακάλυψη των απλών και σύνθετων εγκεφαλικών κυττάρων από τους νευροφυσιολόγους David H. Hubel και Torsten Wiesel. Παρόλο που αυτή η ανακάλυψη δεν ανήκει στον τομέα του Deep Learning, η ιδέα των τεχνητών νευρωνικών δικτύων επηρεάστηκε πολύ από αυτές τις ανακαλύψεις στον τομέα της βιολογίας.
- Το 1960 δημοσιεύτηκε το paper “Gradient Theory of Optimal Flight Paths” από τον Henry J. Kelley όπου πρότεινε μια αρχική μορφή του Back Propagation αλγορίθμου που χρησιμοποιείται και σήμερα με τεχνικές Gradient Descent.
- Το 1965 δημιουργήθηκαν τα πρώτα DNN από τους Alexey Ivakhnenko και V.G. Lapa. Οι αλγόριθμοι εκμάθησης που αναπτύχθηκαν χρησιμοποιούσαν DNN με 8 layers και επιδείκνυαν τη λειτουργία ενός συστήματος αναγνώρισης όπου ονομάστηκε Alpha. Πολλοί τους θεωρούν ως τους ιδρυτές του τομέα του Deep Learning.
 - Το 1979 αναπτύσσεται από τον Kunihiko Fukushima το πρώτο τεχνητό νευρωνικό δίκτυο όπου αναγνωρίζει μοτίβα σε εικόνες ενώ επιπλέον χρησιμοποιήθηκε και σε εφαρμογές αναγνώρισης φυσικής γλώσσας. Η δουλειά του επηρεάστηκε πολύ από τη δουλειά των David H. Hubel και Torsten Wiesel που προαναφέρθηκαν και στο τέλος κατέληξε στη δημιουργία των πρώτων συνελκτικών νευρωνικών δικτύων (Convolutional Neural Networks).
 - Το 1982 δημιουργούνται τα πρώτα recurrent neural networks (RNN) από τον John Hopfield όπου πήραν και το όνομα του ως Hopfield Networks. Παρόλο που αυτός ο τύπος νευρωνικών δικτύων δεν αποτελεί αντικείμενο αυτής της διπλωματικής εργασίας, είναι σημαντικό να γίνει αναφορά για το που χρησιμοποιούνται. Τα RNN είναι πολύ χρήσιμα σε περιπτώσεις όπου έχουμε να κάνουμε με σειριακά δεδομένα όπως είναι τα κείμενα γιατί κάθε νευρώνας διαθέτει κάποιο τύπο μνήμης την οποία χρησιμοποιεί ώστε να παράξει την έξοδο του. Η πιο χαρακτηριστική σημερινή εφαρμογή που χρησιμοποιεί RNN είναι σε συνδυασμό με τον τομέα NLP (Natural Language Processing) όπου δοσμένου ενός κειμένου μπορεί το νευρωνικό δίκτυο να παράξει παρόμοια κείμενα.
 - Το 1985 αναπτύσσεται η πρώτη εφαρμογή που μετατρέπει κείμενο σε ομιλία από τον Terry Sejnowski.
 - Το 1986 γίνονται σημαντικές βελτιώσεις στον τομέα της πρόβλεψης λέξεων (word prediction) από τον David Rumelhart, Geoffrey Hinton, and Ronald J. Williams. Έγιναν βελτιώσεις στον αλγόριθμο Back Propagation των ήδη υπάρχοντων νευρωνικών δικτύων. Παραδείγματα εφαρμογών αυτού του τομέα σήμερα είναι το word prediction που έχει το Facebook messenger.
 - Το 1989 ο Yann LeCun συνδύασε CNN με τον Back Propagation αλγόριθμο ώστε να αναπτύξει ένα σύστημα όπου αναγνώριζε χειρόγραφα ψηφία από εικόνες.
 - Το 1989 ο Christopher Watkins εισήγαγε την ιδέα του Q-Learning με την PhD εργασία του “Learning from Delayed Rewards”. Αυτός ο αλγόριθμος πρότεινε ότι είναι δυνατόν να επιτευχθεί βέλτιστος έλεγχος απευθείας χωρίς την μοντελοποίηση πιθανοτήτων μετάβασης ή προβλεπόμενων ανταμοιβών όπως στα μοντέλα markov.
 - Το 1993 αναπτύσσεται το πρώτο λειτουργικό πολύ βαθύ RNN με 1000 στρώματα από τον Jürgen Schmidhuber. Ήταν ένα τεράστιο βήμα κυρίως για την αντιμετώπιση της υπολογιστικής πολυπλοκότητας και της ικανότητας νευρωνικών δικτύων με τόσα πολλά layers.
 - Το 1995 έγιναν τεράστιες βελτιώσεις στη μοντελοποίηση των Support Vector Machines (SVM) από τους Corinna Cortes και Vladimir Vapnik. Παρόλο που η ιδέα των SVM είχε ξεκινήσει τη δεκαετία του 1960, τα σημερινά μοντέλα βασίζονται στην δουλειά των παραπάνω.

- Το 1997 αναπτύσσεται ένα RNN framework που ονομάστηκε Long short-term memory (LSTM) από τους Jürgen Schmidhuber και Sepp Hochreiter όπου βελτιώνει την απόδοση των RNN αντιμετωπίζοντας το long-term dependency πρόβλημα, όπου οι σημαντικές πληροφορίες βρίσκονται πολύ "πίσω" στην μνήμη των νευρώνων, και κάνοντας αυτά τα νευρωνικά δίκτυα να μπορούν να "θυμούνται καλύτερα" παλαιότερες πληροφορίες. Με διάφορες βελτιώσεις από τότε, τα LSTM δίκτυα χρησιμοποιούνται σήμερα σε εφαρμογές αναγνώρισης ομιλίας της Google.
- Το 1998 αναπτύσσεται ο αλγόριθμος stochastic gradient descent (SGD) από τον Yann LeCun σε συνδυασμό με τον Back Propagation αλγόριθμο. Αυτοί οι 2 αλγόριθμοι είναι πολύ σημαντικοί και χρησιμοποιούνται ακόμα και σήμερα σε εφαρμογές Deep Learning.
- Το 2009 γίνεται η δημοσίευση του ImageNet από την Fei-Fei Li (<https://www.image-net.org/>) του πανεπιστημίου του Stanford στο ευρύ κοινό, μιας τεράστιας βάσης δεδομένων labelled εικόνων, δηλαδή εικόνων μαζί με τις αντίστοιχες ετικέτες/labels ώστε να μπορούν ερευνητές να κάνουν train μοντέλα Deep Learning για supervised learning σε πολύ μεγάλο όγκο δεδομένων. Το dataset αυτό περιείχε 14 εκατομμύρια εικόνες και σήμερα έχει πολύ περισσότερες και χρησιμοποιείται για να δοκιμαστούν αποδόσεις μοντέλων Deep Learning.
- Το 2011 δημιουργήθηκε η CNN αρχιτεκτονική AlexNet από τον Alex Krizhevsky. Η συγκεκριμένη αρχιτεκτονική νίκησε πολλά machine/deep learning βραβεία το 2011 και το 2012 και τα αποτελέσματα έφεραν μια αναγέννηση στην Deep Learning κοινότητα.
- Το 2014 δημιουργείται το πρώτο Deep Learning σύστημα με 97.35% ακρίβεια σε αναγνώριση προσώπου από τη facebook. Ήταν ένα τεράστιο βήμα στον τομέα γιατί η καλύτερη ακρίβεια που είχε επιτευχθεί μέχρι τότε ήταν 70% και γιατί ξεκινάει τα Deep Learning συστήματα να έχουν συγκρίσιμες επιδόσεις με τις ανθρώπινες δυνατότητες, όπου είναι περίπου 97.5%.
- Το 2014 γίνεται η εισαγωγή της έννοιας των Generative Adversarial Networks (GAN) από μία ομάδα ερευνητών με οδηγό τον Ian Goodfellow και τον Yann LeCun. Αυτός ο τύπος νευρωνικών δικτύων είναι πολύ χρήσιμος στην μάθηση χωρίς επιτήρηση (unsupervised learning) όπου είναι πρακτικά ο τελικός στόχος της τεχνητής νοημοσύνης. Στο unsupervised learning έχουμε συστήματα τα οποία μαθαίνουν από δεδομένα χωρίς να χρειάζεται να έχουμε ορίσει labels/ετικέτες. Τα GAN έχουν 2 ανταγωνιζόμενα δίκτυα, το πρώτο παίρνει τα δεδομένα και προσπαθεί να παράξει δυσδιάκριτα δεδομένα ενώ το δεύτερο παίρνει τα αρχικά δεδομένα και αυτά που παρήγαγε το πρώτο και προσπαθεί να μάθει αν κάθε δεδομένο είναι γνήσιο ή όχι. Μαθαίνοντας παράλληλα, αυτά τα 2 δίκτυα σπρώχνουν το ένα το άλλο ώστε να γίνουν "έξυπνα" γρήγορα.
- 2016 μέχρι σήμερα: Πλέον υπάρχουν πάρα πολλά προϊόντα μηχανικής μάθησης στην αγορά από επιχειρήσεις σε μία τεράστια γκάμα εφαρμογών. Οι βελτιώσεις στο hardware των υπολογιστών καθιστούν συστήματα που χρειάζονταν μήνες να γίνουν train δυνατά ώστε να τα καταφέρνουν μέσα σε ώρες.

5.2.1 Σύντομο χρονοδιάγραμμα

- 1960s: "Ρηγά" νευρωνικά δίκτυα (Shallow neural networks). [5]
- 1960-70s: Εισαγωγή της ιδέας του Back Propagation.
- 1974-80: Πρώτη περίοδος μείωσης του ερευνητικού ενδιαφέροντος στον τομέα AI.
- 1980s: Εισαγωγή της έννοιας του Convolutional Layer. • 1987-93: Δεύτερη περίοδος μείωσης του ερευνητικού ενδιαφέροντος στον τομέα AI.
- 1990s: Μηχανική μάθηση χωρίς επιτήρηση (Unsupervised Deep Learning).
- 1990s-2000s: Αναγέννηση του supervised Deep Learning.

- 2006s-σήμερα: Το μοντέρνο Deep Learning.

5.3 Είδη Μάθησης

Εν γένει, ο τομέας της Μηχανικής Μάθησης αναπτύσσει τρεις τρόπους μάθησης, ανάλογους με τους τρόπους με τους οποίους μαθαίνει ο άνθρωπος: επιβλεπόμενη μάθηση, μη επιβλεπόμενη μάθηση και ενισχυτική μάθηση. Πιο αναλυτικά:

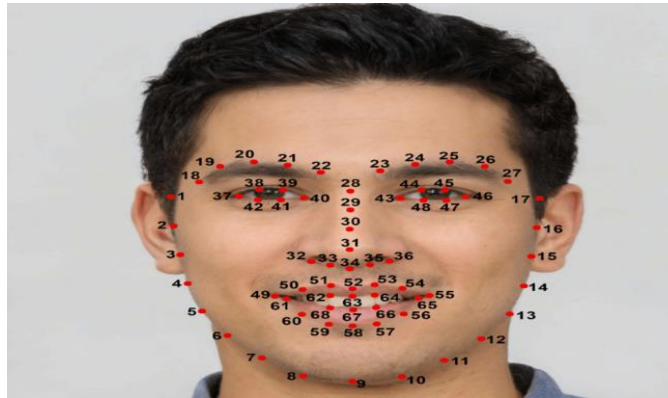
- **Επιβλεπόμενη Μάθηση (Supervised Learning)** είναι η διαδικασία όπου ο αλγόριθμος κατασκευάζει μια συνάρτηση που απεικονίζει δεδομένες εισόδους (σύνολο εκπαίδευσης) σε γνωστές επιθυμητές εξόδους, με απώτερο στόχο τη γενίκευση της συνάρτησης αυτής και για εισόδους με άγνωστη έξοδο. Χρησιμοποιείται σε προβλήματα:
 - Ταξινόμησης (Classification)
 - Πρόγνωσης (Prediction)
 - Διερμηνείας (Interpretation)
- **Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning)**, όπου ο αλγόριθμος κατασκευάζει ένα μοντέλο για κάποιο σύνολο εισόδων υπό μορφή παρατηρήσεων χωρίς να γνωρίζει τις επιθυμητές εξόδους. Χρησιμοποιείται σε προβλήματα:
 - Ανάλυσης Συσχετισμών (Association Analysis)
 - Ομαδοποίησης (Clustering)
- **Ενισχυτική Μάθηση (Reinforcement Learning)**, όπου 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 ρομπότ και η βελτιστοποίηση εργασιών σε εργοστασιακούς χώρους.

Για κάθε πρόβλημα προς επίλυση στο χώρο της Μηχανικής Μάθησης υπάρχει ένας κατάλληλος τρόπος μάθησης και για κάθε τρόπο μάθησης υπάρχει τουλάχιστον ένας κατάλληλος αλγόριθμος που μπορεί να χρησιμοποιηθεί.

Κεφάλαιο 6ο: Computer Vision

Ο τομέας του Computer Vision έχει δει ραγδαία πρόοδο τα τελευταία 10 χρόνια και συνεχίζει να παράγει αποτελέσματα όπου κανείς δεν είχε προβλέψει. Το γεγονός αυτό οφείλεται στην εισαγωγή του Deep Learning στο τομέα. Το Computer Vision έχει πλέον καταφέρει να υλοποιεί ικανοποιητικά εφαρμογές όπως να βοηθάει selfdriving αυτοκίνητα να εντοπίζουν τους πεζούς ή άλλα αυτοκίνητά στο δρόμο και να βελτιώσει σε τεράστιο βαθμό τις επιδόσεις face recognition εφαρμογών. Το Deep Learning έχει εισβάλλει σχεδόν σε όλους τους τομείς της καθημερινότητας μας, παρόλο που δεν είναι εμφανές στο ευρύ κοινό. Συνδυάζεται με υλοποιήσεις που έχουν να κάνουν με διαφημίσεις, με χρηματιστήριο, αναγνώριση φωνής, παραγωγή κειμένου, ακόμα και για μοντέλα που παράγουν καινούριες μορφές τέχνης όπως μουσική, εικόνες κτλ.

6.1 Landmark Detection



6. Εικόνα προσώπου ανθρώπου με landmarks [25]

Μία εντελώς παρόμοια εφαρμογή με object localization είναι το landmark detection. Μπορεί για κάποια εφαρμογή να μας ενδιαφέρει πέρα από το label να αναγνωρίζουμε συγκεκριμένα landmarks μέσα στις εικόνες. Σε μια face detection εφαρμογή μπορεί να θέλουμε επιπλέον να βρούμε landmarks ώστε να αναγνωρίζουμε το σχήμα των ματιών, του στόματος, της μύτης, των φρυδιών κ.ο.κ. ώστε για παράδειγμα να μπορούμε αργότερα να αναγνωρίσουμε αν το άτομο στην εικόνα έχει κλειστά τα μάτια, αν μιλάει, αν γελάει κ.ο.κ. Ο τρόπος με τον οποίο θα σχεδιάσουμε το μοντέλο είναι εντελώς αντίστοιχος με αυτό που είδαμε στο object localization. Η μόνη αλλαγή είναι ότι εδώ αντί να εισάγουμε τις 4 μεταβλητές που ορίζουν το bounding box θα εισάγουμε 2 μεταβλητές (l_{ix} , l_{iy}) για κάθε landmark που θέλουμε να εντοπίζουμε. Αν π.χ. θέλουμε ένα μοντέλο που να αναγνωρίζει 10 διαφορετικά άτομα (labels) και να εντοπίζει 64 face landmarks τότε πάλι θα προσθέσουμε ένα SoftMax(10) unit και 128 nodes για τα landmarks ενώ το target label θα είναι $y = (P_c \ l_{1x} \ l_{1y} \dots \ l_{64x} \ l_{64y} \ c_1 \dots \ c_{10})^T$ και θα έχει συνολικά 139 στοιχεία. Τα loss functions που θα χρησιμοποιήσουμε θα είναι πάλι log likelihood για τα c_i , MSE για τα (l_{ix} , l_{iy}) και logistic regression loss για το P_c . Θα πρέπει όμως να έχουμε ένα dataset στο οποίο κάποιος να έχει μπει στην πολύ επίπονη διαδικασία να κάνει το notation για τα landmarks σε κάθε εικόνα αλλά ευτυχώς σήμερα υπάρχουν πολλά τέτοια open source dataset στο internet.

Πολύ χαρακτηριστικό παράδειγμα εφαρμογής landmark detection που υπάρχει σήμερα είναι τα face filters που χρησιμοποιούν εφαρμογές όπως instagram, facebook messenger, snapchat κτλ. όπου εντοπίζουν τα landmarks μέσω CNN και βάση αυτών προσθέτουν επιπλέον γραφικά στις εικόνες προσώπων.

Κεφάλαιο 7ο: Histograms of Oriented Gradients-HOG

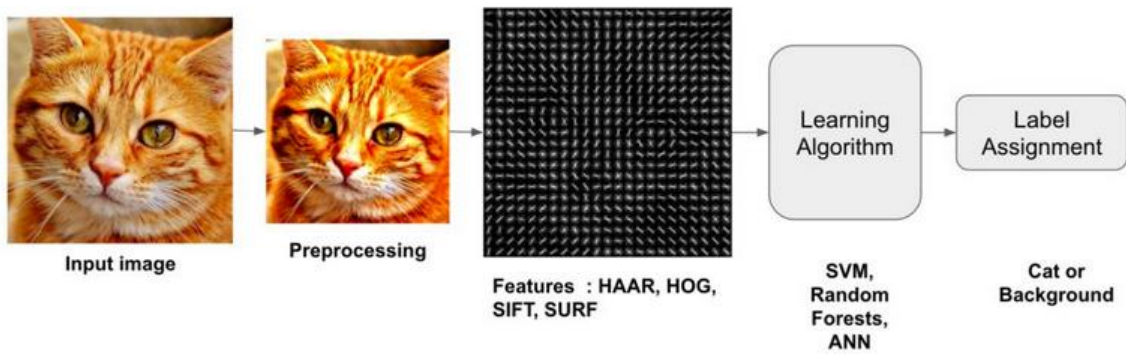
Επειδή ο κώδικας που υλοποιήθηκε για την αναγνώριση προσώπου βασίζεται στη λειτουργία της μεθόδου Histograms Oriented Gradients- HOG, παρακάτω θα αναλύσω τον τρόπο λειτουργίας της.

7.1 Τι είναι ο «Feature Descriptor» περιγραφέας χαρακτηριστικών.

Ένας[9] περιγραφέας χαρακτηριστικών είναι μια αναπαράσταση μιας εικόνας εξάγοντας χρήσιμες πληροφορίες και απορρίπτοντας ξένες πληροφορίες-«θορύβους» δηλαδή πληροφορίες που δεν χρειάζονται. Συνήθως, ένας περιγραφέας χαρακτηριστικών μετατρέπει μια εικόνα μεγέθους πλάτους x ύψους $x \cdot 3$ (κανάλια) σε διάνυσμα χαρακτηριστικών / πίνακα μήκους n . Ο περιγραφέας HOG μπορεί να υπολογιστεί για διάφορα μεγέθη. Τίθεται το ερώτημα τι είναι «χρήσιμο» και τι «ξένο»; Για να ορίσουμε το "χρήσιμο", πρέπει να γνωρίζουμε τι είναι "χρήσιμο"; Το διάνυσμα χαρακτηριστικών που παράγεται από αυτούς τους αλγόριθμους όταν τροφοδοτείται σε μια ταξινόμηση εικόνας, οι αλγόριθμοι όπως το Support Vector Machine (SVM) παράγουν καλά αποτελέσματα. Στον περιγραφέα χαρακτηριστικών HOG, η κατανομή (ιστογράμματα) των κατευθύνσεων των κλίσεων (προσανατολισμένες διαβαθμίσεις) χρησιμοποιούνται ως χαρακτηριστικά. Οι διαβαθμίσεις (παράγωγοι x και y) μιας εικόνας είναι χρήσιμες επειδή το μέγεθος των κλίσεων είναι μεγάλο γύρω από τις άκρες και τις γωνίες (περιοχές απότομων αλλαγών έντασης) και γνωρίζουμε ότι οι ακμές και οι γωνίες περιέχουν πολύ περισσότερες πληροφορίες σχετικά με το σχήμα του αντικειμένου από τις επίπεδες περιοχές. Ένας αλγόριθμος αναγνώρισης εικόνας (γνωστός και ως ταξινομητής εικόνας) λαμβάνει μια εικόνα (ή ένα patch μιας εικόνας) ως είσοδο και εξάγει αυτό που περιέχει η εικόνα. Με άλλα λόγια, η έξοδος είναι μια ετικέτα κλάσης (π.χ. "cat", "dog", "table" κ.λπ.). Πώς γνωρίζει ένας αλγόριθμος αναγνώρισης εικόνας τα περιεχόμενα μιας εικόνας; Εάν θέλουμε να βρούμε πχ τις γάτες σε εικόνες, πρέπει να εκπαιδεύσετε έναν αλγόριθμο αναγνώρισης εικόνων με χιλιάδες εικόνες γατών και χιλιάδες εικόνες φόντου που δεν περιέχουν γάτες.

Συχνά μια εικόνα εισόδου υποβάλλεται σε προεπεξεργασία για την ομαλοποίηση των εφέ αντίθεσης και φωτεινότητας. Ένα πολύ κοινό βήμα προεπεξεργασίας είναι η αφαίρεση του μέσου όρου των εντάσεων της εικόνας και η διαίρεση με την τυπική απόκλιση. Στις εικόνες που είναι με χρώματα ένας μετασχηματισμός χρωματικού χώρου (π.χ. χρωματικός χώρος RGB σε LAB) μπορεί να μας βοηθήσει να έχουμε καλύτερα αποτελέσματα.

Η εικόνα εισόδου έχει πάρα πολλές επιπλέον πληροφορίες που δεν είναι απαραίτητες για ταξινόμηση. Επομένως, το πρώτο βήμα στην ταξινόμηση της εικόνας είναι η απλοποίηση της εικόνας εξάγοντας τις σημαντικές πληροφορίες που περιέχονται στην εικόνα και αφήνοντας έξω τις υπόλοιπες. Για παράδειγμα, αν θέλουμε να βρούμε τα κουμπιά πουκάμισου και παλτό σε εικόνες, θα παρατηρήσουμε μια σημαντική διακύμανση στις τιμές των εικονοστοιχείων RGB. Ωστόσο, εκτελώντας έναν ανιχνευτή ακμών σε μια εικόνα μπορούμε να απλοποιήσουμε την εικόνα. Μπορούμε ακόμα να διακρίνουμε εύκολα το κυκλικό σχήμα των κουμπιών σε αυτές τις ακμές εικόνας και έτσι μπορούμε να συμπεράνουμε ότι η ανίχνευση άκρων διατηρεί τις βασικές πληροφορίες ενώ απορρίπτει μη βασικές πληροφορίες. Το βήμα ονομάζεται εξαγωγή χαρακτηριστικών. Στις παραδοσιακές προσεγγίσεις υπολογιστικής όρασης, ο σχεδιασμός αυτών των χαρακτηριστικών είναι ζωτικής σημασίας για την απόδοση του αλγορίθμου. Αποδεικνύεται ότι μπορούμε να κάνουμε πολύ καλύτερα από τον απλό εντοπισμό άκρων και να βρούμε λειτουργίες που είναι πολύ πιο αξιόπιστες. Στο παράδειγμά μας με τα κουμπιά πουκάμισου και παλτό, ένας καλός ανιχνευτής χαρακτηριστικών όχι μόνο θα καταγράψει το κυκλικό σχήμα των κουμπιών αλλά και πληροφορίες σχετικά με το πώς διαφέρουν τα κουμπιά από άλλα κυκλικά αντικείμενα, όπως τα ελαστικά αυτοκινήτου.



Ανατομία ταξινομητή εικόνας[9]

Πολλοί παραδοσιακοί αλγόριθμοι ταξινόμησης εικόνων ακολουθούν αυτήν τη γραμμή, ενώ οι αλγόριθμοι που βασίζονται στη βαθιά μάθηση παρακάμπτουν πλήρως το βήμα εξαγωγής χαρακτηριστικών. Ένας αλγόριθμος εξαγωγής χαρακτηριστικών μετατρέπει μια εικόνα σταθερού μεγέθους σε ένα διάνυσμα χαρακτηριστικών σταθερού μεγέθους. Το HOG βασίζεται στην ιδέα ότι η εμφάνιση τοπικού αντικειμένου μπορεί να περιγραφεί αποτελεσματικά από την κατανομή (ιστόγραμμα) των κατευθύνσεων των ακμών (προσανατολισμένες κλίσεις).



Υπολογισμός κλίσης : Υπολογισμός της κλίσης της εικόνας x και y , g_x και g_y , από την αρχική εικόνα. Χρησιμοποιώντας τις διαβαθμισμένες εικόνες g_x και g_y , μπορούμε να υπολογίσουμε το μέγεθος και τον προσανατολισμό της βαθμίδας χρησιμοποιώντας τις ακόλουθες εξισώσεις.

$$g = \sqrt{g_x^2 + g_y^2}$$

$$\theta = \arctan \frac{g_y}{g_x}$$

Οι υπολογιζόμενες κλίσεις είναι «μη υπογεγραμμένες» και επομένως το θ βρίσκεται στην περιοχή από 0 έως 180 μοίρες.

7.2 Αλγόριθμος μάθησης για ταξινόμηση

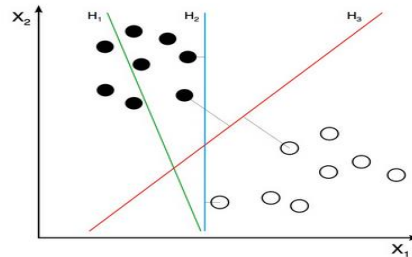
Για να μπορέσει ένας αλγόριθμος ταξινόμησης να κάνει τα μαγικά του, πρέπει να τον εκπαιδύσουμε δείχνοντας χιλιάδες παραδείγματα προσώπων και υποβάθρου. Διαφορετικοί αλγόριθμοι μάθησης μαθαίνουν διαφορετικά, αλλά η γενική αρχή είναι ότι οι αλγόριθμοι μάθησης αντιμετωπίζουν τα διανύσματα χαρακτηριστικών ως σημεία στον χώρο υψηλότερων διαστάσεων και προσπαθούν να βρουν επίπεδα/επιφάνειες που χωρίζουν τον χώρο υψηλότερων διαστάσεων με τέτοιο τρόπο ώστε όλα τα παραδείγματα που ανήκουν στην ίδια κλάση να είναι στη μία πλευρά του επιπέδου / επιφάνειας.

Για να απλοποιήσουμε τα πράγματα, θα αναφέρω παρακάτω τον αλγόριθμο εκμάθησης Support Vector Machines (SVM).

Πώς λειτουργεί το Support Vector Machine (SVM) για την ταξινόμηση εικόνων;

Το Support Vector Machine (SVM) είναι ένας από τους πιο δημοφιλείς εποπτευόμενους αλγόριθμους δυαδικής ταξινόμησης. Αν και οι ιδέες που χρησιμοποιούνται στο SVM υπάρχουν από το 1963, η τρέχουσα έκδοση προτάθηκε το 1995 από τους Cortes και Vapnik.

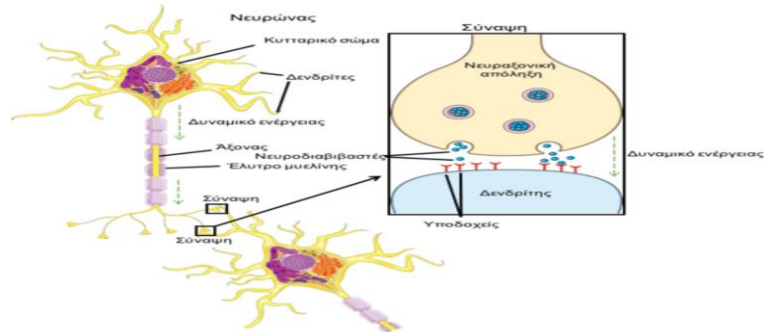
Παραπάνω ανέφερα ότι ο περιγραφέας HOG μιας εικόνας είναι ένα διάνυσμα χαρακτηριστικών. Μπορούμε να σκεφτούμε αυτό το διάνυσμα ως ένα σημείο σε ένα χώρο X διαστάσεων. Για ευκολία της οπτικοποίησης υποθέτουμε ότι το διάνυσμα χαρακτηριστικών είναι μόνο δισδιάστατο.



Στον απλοποιημένο κόσμο μας, έχουμε πλέον δισδιάστατα σημεία που αντιπροσωπεύουν τις δύο κατηγορίες (π.χ. γάτες και φόντο). Στην παραπάνω εικόνα, οι δύο κατηγορίες αντιπροσωπεύονται από δύο διαφορετικά είδη κουκκίδων. Όλες οι μαύρες κουκκίδες ανήκουν σε μια κατηγορία και οι λευκές κουκκίδες ανήκουν στην άλλη κατηγορία. Κατά τη διάρκεια της εκπαίδευσης, παρέχουμε στον αλγόριθμο πολλά παραδείγματα από τις δύο κατηγορίες. Με άλλα λόγια, λέμε στον αλγόριθμο τις συντεταγμένες των 2D κουκκίδων και επίσης αν η κουκκίδα είναι μαύρη ή άσπρη. Διαφορετικοί αλγόριθμοι μάθησης καταλαβαίνουν πώς να διαχωρίσουν αυτές τις δύο τάξεις με διαφορετικούς τρόπους. Το Linear SVM προσπαθεί να βρει την καλύτερη γραμμή που χωρίζει τις δύο κατηγορίες. Στο παραπάνω σχήμα, τα H_1 , H_2 και H_3 είναι τρεις γραμμές σε αυτόν τον 2D χώρο. Το H_1 δεν διαχωρίζει τις δύο κατηγορίες και επομένως δεν είναι καλός ταξινομητής. Το H_2 και το H_3 χωρίζουν και τα δύο τις δύο κατηγορίες, αλλά διαισθητικά φαίνεται ότι το H_3 είναι καλύτερος ταξινομητής από το H_2 επειδή το H_3 φαίνεται να διαχωρίζει τις δύο κατηγορίες πιο καθαρά. Γιατί? Επειδή το H_2 είναι πολύ κοντά σε μερικές από τις ασπρόμαυρες κουκκίδες. Από την άλλη πλευρά, το H_3 επιλέγεται έτσι ώστε να βρίσκεται σε μέγιστη απόσταση από τα μέλη των δύο τάξεων.

Τι γίνεται αν τα χαρακτηριστικά που ανήκουν στις δύο κατηγορίες δεν μπορούν να διαχωριστούν χρησιμοποιώντας ένα υπερεπίπεδο; Σε τέτοιες περιπτώσεις, το SVM εξακολουθεί να βρίσκει το καλύτερο υπερεπίπεδο λύνοντας ένα πρόβλημα βελτιστοποίησης που προσπαθεί να αυξήσει την απόσταση του υπερεπίπεδου από τις δύο κατηγορίες ενώ προσπαθεί να βεβαιωθεί ότι πολλά παραδείγματα εκπαίδευσης έχουν ταξινομηθεί σωστά. Αυτή η ανταλλαγή ελέγχεται από μια παράμετρο που ονομάζεται C . Όταν η τιμή του C είναι μικρή, επιλέγεται ένα υπερεπίπεδο μεγάλου περιθωρίου εις βάρος ενός μεγαλύτερου αριθμού εσφαλμένων ταξινομήσεων. Αντίθετα, όταν το C είναι μεγάλο, επιλέγεται ένα μικρότερο υπερεπίπεδο περιθωρίου που προσπαθεί να ταξινομήσει σωστά πολλά περισσότερα παραδείγματα.

Κεφάλαιο 8ο: Ο νευρώνας



7. Σχηματική απεικόνιση ενός νευρώνα (αριστερά) και της σύναψης (δεξιά)

Παρακάτω θα αναφέρω λίγες πληροφορίες για τον ανθρώπινο νευρώνα καθώς, η τεχνητή νοημοσύνη, όπως αναφέραμε και παραπάνω δανείζεται στον τρόπο λειτουργίας του σε αυτόν της λειτουργίας του ανθρώπινου νευρώνα.

Οι νευρώνες [24] είναι εξειδικευμένα κύτταρα που λαμβάνουν, επεξεργάζονται και διαδίδουν πληροφορίες μέσω του σχηματισμού ηλεκτρικού κυκλώματος όπως κάνει ένας ηλεκτρονικός υπολογιστής. Για να λάβουν μια πληροφορία έχουν μικρά άκρα που λέγονται δενδρίτες, οι όποιοι όπως και τα κλαδιά ενός δέντρου, απλώνονται προς άλλους νευρώνες και σχηματίζουν χιλιάδες συνδέσεις. Όταν ένας δενδρίτης λάβει ηλεκτρικά σήματα από έναν άλλο νευρώνα, αυτή η πληροφορία μετακινείται στο κέντρο του κυττάρου, όπου το σήμα επεξεργάζεται για να αποφασισθεί εάν και πως θα σταλεί στους άλλους νευρώνες.

Για να στείλουν τις πληροφορίες, οι νευρώνες αναπτύσσουν ένα πολύ εξειδικευμένο καλώδιο, τον άξονα. Οι άξονες από ένα μόνο κύτταρο μπορούν να έχουν μήκος από το μεγάλο δάχτυλο του ποδιού μέχρι το γοφό! Ο άξονας πρέπει να είναι προστατευμένος για να μεταβιβάσει το σήμα αποτελεσματικά. Γι' αυτό, είναι καλυμμένος από δομές που ονομάζονται στρώματα μυελίνης, τα οποία δρουν σαν ένα μονωτικό πλαστικό κάλυμμα σαν αυτό που καλύπτει τα κανονικά καλώδια. Επίσης μεταφέρουν τις ηλεκτρικές ωθήσεις πιο γρήγορα. Αυτό το στρώμα μυελίνης είναι πολύ σημαντικό για τη σωστή λειτουργία των νευρώνων, οι οποίοι είναι ευάλωτοι σε ασθένειες όπως την πολλαπλή σκλήρυνση, στην οποία τα στρώματα μυελίνης δέχονται επίθεση από το ίδιο μας το ανοσοποιητικό σύστημα.

Η επικοινωνία από έναν άξονα προς τους δενδρίτες άλλων νευρώνων συμβαίνει μέσα σε πολύ μικρούς χώρους στους οποίους το τελικό σημείο του άξονα ακουμπά τον γειτονικό δενδρίτη, όπως δυο δάχτυλα πολύ κοντά το ένα στο άλλο χωρίς όμως να ακουμπούν πραγματικά. Αυτές οι ενώσεις ονομάζονται συνάψεις και το σήμα μεταδίδεται από μόρια που ονομάζονται νευροδιαβιβαστές (όπως η ντοπαμίνη, η σεροτονίνη και το γλουταμινικό οξύ) οι οποίοι ταξιδεύουν μικρές αποστάσεις από τον ένα άξονα για να μεταβιβάσει την πληροφορία στον επόμενο νευρώνα.

Κεφάλαιο 9ο: Ανθρώπινο πρόσωπο

Η σημασία του ανθρώπινου προσώπου δεν είναι μόνο ανατομική και λειτουργική, αλλά έχει επιπλέον ένα γενικότερο ατομικό και διαπροσωπικό ρόλο. Αποτελεί για κάθε άνθρωπο ένα εντελώς ιδιαίτερο γνώρισμα, του οποίου η σπουδαιότητα έχει κοινωνικές, ψυχολογικές, καλλιτεχνικές και φιλοσοφικές προεκτάσεις. Η αναγνώριση προσώπου αποτελεί μια εναλλακτική πρόταση ταυτοποίησης ανθρώπων χωρίς τη χρήση εγγράφων, όπως είναι τα δακτυλικά αποτυπώματα, η ίριδα του ματιού κ.ά.

Οι δυνατότητες της γεωμετρίας του προσώπου για την αναγνώριση προσώπων και εκφράσεων ουσιαστικά άρχισαν να ερευνώνται τη τελευταία περίπου δεκαετία. Το υψηλό κόστος των αισθητήρων σε συνδυασμό με τη χαμηλή ποιότητα των δεδομένων που παρήγαγαν και το μεγάλο χρόνο απόκρισης οδήγησαν τη πλειοψηφία των ερευνητών στη χρήση έγχρωμων εικόνων ή εικόνων φωτεινότητας και την ανάπτυξη φωτομετρικών (photometry) μεθόδων αναγνώρισης.

Στον τομέα της αναγνώρισης προσώπων προτάθηκε ένας τεράστιος αριθμός τεχνικών, πολλές από τις οποίες έχουν ήδη χρησιμοποιηθεί για την ανάπτυξη εμπορικών συστημάτων.

Με την πρόοδο των αισθητήρων βάθους τα τελευταία χρόνια και τη διαθεσιμότητα δεδομένων στην επιστημονική κοινότητα για την διεξαγωγή πειραμάτων, η έρευνα γύρω από την 2D ή 3D αναγνώριση προσώπων και εκφράσεων σημείωσε σημαντική ώθηση. Όλα τα συστήματα αναγνώρισης προτύπων ακολουθούν μια τυπική αρχιτεκτονική που περιλαμβάνει τέσσερα βασικά στάδια επεξεργασίας:

1. την εκμάθηση του συστήματος μας από ένα dataset εικόνων,
2. την προ-επεξεργασία των εικόνων αυτών για τη μείωση του θορύβου ή την κανονικοποίηση των μετρήσεων (normalization, lighting correction, histogram equalization),
3. την εξαγωγή χαρακτηριστικών (features) με τα οποία γίνεται η διάκριση των προτύπων και
4. την ταξινόμηση, που περιλαμβάνει την κατηγοριοποίηση με βάση τα εξαγόμενα χαρακτηριστικά της εικόνας, του περιεχομένου της εικόνας σε γνωστές κατηγορίες (πρότυπα) που περιγράφουν το πρόβλημα.

Η πληθώρα των εφαρμογών της αναγνώρισης προσώπου, πολλές από τις οποίες έχουν σημαντικό κοινωνικό και ατομικό αντίκτυπο, καθώς και οι συνεχείς τεχνολογικές εξελίξεις ιδιαίτερα στην υπερσχή της περιβάλλουσας νοημοσύνης, καθιστούν την ανάπτυξη αξιόπιστων συστημάτων αναγνώρισης προσώπου που θα παρουσιάζουν σταθερά υψηλή επίδοση σε οποιοσδήποτε συνθήκες λειτουργίας, μία επιτακτική ανάγκη.

Παρατηρώντας την εκπληκτική ευκολία με την οποία οι άνθρωποι αναγνωρίζουν τα πρόσωπα γύρω τους, η υλοποίηση ενός συστήματος αυτόματης αναγνώρισης προσώπων που στηρίζεται στην καταγραφή και την ανάλυση εικόνων του προσώπου εμφανίζεται ως ένα σχετικά εύκολο εγχείρημα. Ωστόσο, όπως και στην περίπτωση των περισσότερων λειτουργιών της αντίληψης, η ευκολία αυτή είναι παραπλανητική.

Επιστημονικές έρευνες έχουν δείξει ότι πάνω από το 50% του φλοιού του εγκεφάλου παρουσιάζει αυξημένη δραστηριότητα κατά τη διάρκεια της επεξεργασίας οπτικών ερεθισμάτων, ενώ φαίνεται ότι ο ανθρώπινος εγκέφαλος αφιερώνει περισσότερους πόρους στην όραση παρά σε φαινομενικά πιο απαιτητικές νοητικές διεργασίες όπως η επίλυση μαθηματικών προβλημάτων. Το γεγονός αυτό είναι

ενδεικτικό του βαθμού δυσκολίας που εμπεριέχει η ανάπτυξη ενός αυτόματου συστήματος αναγνώρισης προσώπου.

Η βιομετρική τεχνολογία ασφάλειας δεν είναι πλέον ένα σενάριο επιστημονικής φαντασίας όπως συνηθίζαμε να βλέπουμε σε παλαιότερες ταινίες, αλλά ένα αρκετά ώριμο «προϊόν». Πρόκειται για μια βιώσιμη, οικονομικά αποδοτική λύση, ωστόσο η υιοθέτησή της, ιδιαίτερα στον επιχειρηματικό κόσμο, είναι χαμηλή. Παρόλο που η τεχνολογία εξακολουθεί να θεωρείται ως αντικατάσταση των κωδικών πρόσβασης, οι σημερινές εφαρμογές χρησιμοποιούν συχνά τα βιομετρικά στοιχεία ως δευτερεύουσα μορφή ταυτοποίησης που λειτουργεί σε συνδυασμό με την παραδοσιακή ασφάλεια.

Η βιομετρική τεχνολογία είναι σε θέση να αναγνωρίσει ένα άτομο με βάση τα μοναδικά χαρακτηριστικά του προσώπου, τα δακτυλικά αποτυπώματά του, την υπογραφή του, το DNA ή το σχήμα της ίριδας του ματιού του, και στη συνέχεια να μεταδώσει μια ασφαλή και εύχρηστη μέθοδο για σκοπούς ελέγχου ταυτότητας.

Με την βιομετρική τεχνολογία μπορούμε να μετράμε και να αναλύουμε στατιστικά τόσο τα φυσικά χαρακτηριστικά, όσο και τις συμπεριφορές κάθε ατόμου ξεχωριστά. Για παράδειγμα, τα συστήματα αναγνώρισης φωνής λειτουργούν μέσω μέτρησης των χαρακτηριστικών της ομιλίας ενός ατόμου, καθώς ο αέρας αποβάλλεται από τους πνεύμονες, διαμέσου του λάρυγγα και έξω από τη μύτη και το στόμα. Το λογισμικό επαλήθευσης της ομιλίας θα συγκρίνει αυτά τα χαρακτηριστικά με τα δεδομένα που είναι ήδη αποθηκευμένα και αν τα δύο αποτυπώματα της φωνής είναι αρκετά παρόμοια, τότε το βιομετρικό σύστημα ασφαλείας θα επιβεβαιώσει ότι πρόκειται για το ίδιο άτομο.

Υπάρχουν πολλές μορφές βιομετρικών, αλλά και άλλων τεχνολογιών που χρησιμοποιούνται σε όλο τον κόσμο.

9.1 Ενδοατομικές και διατομικές παραλλαγές

Η ύπαρξη ενδοατομικών παραλλαγών αποτελεί το μεγαλύτερο πρόβλημα της αναγνώρισης προσώπου, καθώς έχει αποδειχθεί ότι οι αλλαγές στην εμφάνιση ενός προσώπου που προκαλούνται από τις εκφράσεις, την πόζα ή το φωτισμό, είναι μεγαλύτερες από τις αντιληπτές διαφορές μεταξύ διαφορετικών προσώπων. Οι κύριοι παράγοντες από τους οποίους επηρεάζεται η εμφάνιση ενός προσώπου σε μία εικόνα θα μπορούσαν να συνοψιστούν στους παρακάτω

- **Οπτική γωνία.** Καθώς αλλάζει η θέση της κάμερας ή ο άνθρωπος στρέφει το κεφάλι του, η εμφάνιση του προσώπου στις εικόνες που καταγράφει η κάμερα μεταβάλλεται. Η μεταβολή αυτή οφείλεται αφενός μεν στην προβολική παραμόρφωση, η οποία οδηγεί σε επέκταση κάποιων τμημάτων του προσώπου και συμπύκνωση κάποιων άλλων και αφετέρου στην απόκρυψη περιοχών του προσώπου από το οπτικό πεδίο της κάμερας. Αν ο αλγόριθμος αναγνώρισης προσώπου έχει εκπαιδευτεί χρησιμοποιώντας εικόνες του προσώπου με συγκεκριμένη πόζα (pose), δηλαδή υπό συγκεκριμένη οπτική γωνία, τότε είναι γενικά δύσκολο να αναγνωρίσουμε το ίδιο πρόσωπο από μία διαφορετική οπτική γωνία.

- **Φωτισμός.** Όπως οι μεταβολές της πόζας, έτσι και οι μεταβολές του φωτισμού είναι αναπόφευκτες. Ο περιβαλλοντικός φωτισμός (ambient light) αλλάζει σημαντικά κατά τη διάρκεια της ημέρας, από ημέρα σε ημέρα, και μεταξύ διαφορετικών χώρων και τόπων (π.χ. σε ένα εσωτερικό περιβάλλον και έναν

εξωτερικό χώρο). Επιπλέον, λόγω της 3D δομής του προσώπου, πηγές φωτισμού κοντά στο πρόσωπο δημιουργούν έντονες σκιές και προκαλούν μεταβολές των φυσικών χρωμάτων της επιφάνειας του προσώπου, με αποτέλεσμα κάποια χαρακτηριστικά ή περιοχές του προσώπου να τονίζονται περισσότερο ή ακόμη και να αλλοιώνονται, ενώ οι περιοχές που βρίσκονται στη σκιά να μην είναι ορατές.

- Εκφράσεις. Το πρόσωπο είναι μία 3D επιφάνεια που περιέχει και παραμορφώσιμες περιοχές. Οι εκφράσεις του προσώπου που οφείλονται στα συναισθήματα και η ομιλία παραμορφώνουν ένα μεγάλο τμήμα του προσώπου και αλλοιώνουν δραστικά την εμφάνιση του.
- Απόκρυψη (occlusion). Πολλές φορές τμήματα του προσώπου αποκρύπτονται από άλλα αντικείμενα της σκηνής, από γυαλιά, από κασκόλ, σκούφους, καπέλα και άλλα παρόμοια αξεσουάρ.
- Χρόνος. Καθώς περνούν τα χρόνια, τα πρόσωπα αλλάζουν. Εκτός από τις αλλαγές που οφείλονται στη φυσιολογική γήρανση του ανθρώπου, η εμφάνιση του προσώπου μεταβάλλεται επίσης από παράγοντες όπως το χτένισμα, τα γένια, η χρήση προϊόντων καλλωπισμού και περιποίησης προσώπου που μεταβάλλουν π.χ. την υφή του δέρματος, το φυσικό μαύρισμα που προκαλείται από την έκθεση στον ήλιο, αλλά και οι αισθητικές επεμβάσεις.

Η μεγάλη πλειοψηφία των μεθόδων ανίχνευσης προσώπου αντιμετωπίζουν το πρόβλημα της αναγνώρισης προσώπων ως ένα πρόβλημα αναγνώρισης τρισδιάστατων αντικειμένων από δισδιάστατες (2D) εικόνες έντασης της φωτεινότητας (intensity images). Παρόλο που η αναγνώριση προσώπου από 2D εικόνες έχει φτάσει σε μία εποχή ωριμότητας παρουσιάζοντας υψηλές επιδόσεις σε ελεγχόμενες συνθήκες, οι 2D μέθοδοι αναγνώρισης προσώπου εξακολουθούν να επηρεάζονται σημαντικά από παράγοντες όπως ο φωτισμός, οι στροφές του κεφαλιού και οι εκφράσεις του προσώπου.

9.2 Ανίχνευση προσώπου με βάση τη γνώση

Οι μέθοδοι αυτής της κατηγορίας βασίζονται σε κανόνες που εξάγονται με βάση την ανθρώπινη γνώση σχετικά με την τυπική γεωμετρία του προσώπου και τις θέσεις των χαρακτηριστικών του. Δεδομένου ότι το ανθρώπινο πρόσωπο παρουσιάζει μία εγγενή συμμετρία, ενώ τα κύρια χαρακτηριστικά του εμφανίζονται σε μία φυσική αλληλουχία από πάνω προς τα κάτω και από αριστερά προς τα δεξιά, είναι σχετικά εύκολο να ορίσει κανείς κανόνες που περιγράφουν το σχήμα, το μέγεθος, την υφή και άλλα παρόμοια γνωρίσματα χαρακτηριστικών του προσώπου όπως τα μάτια, τα φρύδια, η μύτη, το πηγούνι, κτλ, αλλά και να περιγράψει τις μεταξύ τους χωρικές σχέσεις, δηλαδή τις σχετικές τους θέσεις και αποστάσεις. Συνήθως υιοθετείται μία ιεραρχική προσέγγιση που εξετάζει την εικόνα του προσώπου σε διαφορετικά επίπεδα ανάλυσης. Στα υψηλότερα επίπεδα οι υποψήφιες περιοχές ανιχνεύονται χρησιμοποιώντας μία χονδρική περιγραφή της γεωμετρίας του προσώπου. Σε χαμηλότερα επίπεδα, εξάγονται χαρακτηριστικά του προσώπου όπως αυτά που αναφέρθηκαν παραπάνω και μία περιοχή χαρακτηρίζεται ως "πρόσωπο" ή "μη-πρόσωπο" βάσει προκαθορισμένων κανόνων για τα χαρακτηριστικά του προσώπου και τις σχετικές τους θέσεις.

Το κύριο πρόβλημα των μεθόδων αυτής της κατηγορίας είναι η δυσκολία μετασχηματισμού της ανθρώπινης γνώσης σχετικά με τη γεωμετρία του προσώπου σε καλά ορισμένους κανόνες. Ένα άλλο πρόβλημα είναι ότι τέτοιες τεχνικές δεν αποδίδουν καλά όταν υπάρχουν περιστροφές του κεφαλιού.

9.3 Ανίχνευση προσώπου με βάση τα χαρακτηριστικά

Στόχος των μεθόδων αυτής της κατηγορίας είναι η ανάδειξη χαρακτηριστικών του προσώπου τα οποία παραμένουν σταθερά ακόμη και όταν η οπτική γωνία ή οι συνθήκες του φωτισμού μεταβάλλονται, και η χρησιμοποίηση των χαρακτηριστικών αυτών για τον εντοπισμό του προσώπου. Διάφορες μέθοδοι έχουν προταθεί, οι οποίες συνήθως χρησιμοποιούν χαρακτηριστικά όπως το χρώμα του δέρματος, η πληροφορία υφής ή τοπικά χαρακτηριστικά του προσώπου. Το χρώμα του δέρματος αποτελεί ένα πολύ ισχυρό εργαλείο για την αναζήτηση προσώπων, αφού η κατάτμηση έγχρωμων εικόνων είναι υπολογιστικά γρήγορη, ενώ ταυτόχρονα δεν επηρεάζεται ιδιαίτερα από την οπτική γωνία, την κλίμακα, τις σκιάσεις, τη μερική απόκρυψη του προσώπου και την ύπαρξη πολύπλοκου φόντου. Η προσέγγιση που βασίζεται στο χρώμα του δέρματος ταξινομεί κάθε εικονοστοιχείο ως "δέρμα" ή "μη-δέρμα" ανάλογα με την ομοιότητα του με το χρώμα του δέρματος, και στη συνέχεια χαρακτηρίζει μία υποπεριοχή της εικόνας ως "πρόσωπο", εάν περιλαμβάνει μία αρκετά μεγάλη συνεκτική ομάδα εικονοστοιχείων που έχουν χαρακτηριστεί ως "δέρμα". Το κύριο μειονέκτημα αυτής της προσέγγισης είναι ότι παρουσιάζει ευαισθησία στο φωτισμό, στην ύπαρξη άσχετων περιοχών που έχουν χρώμα παρόμοιο με αυτό του δέρματος και στις περιπτώσεις προσώπων που είναι το ένα δίπλα στο άλλο. Άλλες μέθοδοι αυτής της κατηγορίας χρησιμοποιούν τοπικά χαρακτηριστικά όπως τα μάτια, τα φρύδια, τη μύτη, το στόμα, κ.α. τα οποία εξάγονται χρησιμοποιώντας διαφορικά ή πολύ-διακριτά φίλτρα (derivative or multi-resolution filters), ανιχνευτές ακμών, μορφολογικές διεργασίες (morphological operations) ή τεχνικές threshold. Με βάση τα εξαγόμενα χαρακτηριστικά κατασκευάζονται στη συνέχεια στατιστικά μοντέλα, τα οποία περιγράφουν τις μεταξύ τους σχέσεις και χρησιμοποιούνται για να επαληθεύσουν την ύπαρξη ενός προσώπου. Για την επαλήθευση της ύπαρξης προσώπου έχουν επίσης προταθεί τεχνικές όπως τα νευρωνικά δίκτυα, η αντιστοίχιση γράφων και τα δέντρα αποφάσεων. Υπάρχουν επίσης τεχνικές που συνδυάζουν διαφορετικά χαρακτηριστικά για να βελτιώσουν την ακρίβεια της ανίχνευσης. Συνήθως αρχικά χρησιμοποιούνται στοιχεία όπως η υφή, το σχήμα και το χρώμα του προσώπου για να ανιχνευθούν υποψήφιες για "πρόσωπο" περιοχές της εικόνας, ενώ στη συνέχεια χρησιμοποιούνται τοπικά χαρακτηριστικά όπως τα μάτια, η μύτη και το στόμα για να επαληθευτεί η ύπαρξη του προσώπου. Οι τεχνικές που βασίζονται στην ανίχνευση σταθερών χαρακτηριστικών του προσώπου αποδεικνύονται προβληματικές στην περίπτωση που τα χαρακτηριστικά αυτά έχουν αλλοιωθεί λόγω φωτισμού, θορύβου ή απόκρυψης από κάποιο αντικείμενο.

9.4 Ανίχνευση προσώπου με βάση τα περιγράμματα (templates)

Η ανίχνευση του προσώπου με χρήση τεχνικών που βασίζονται στη σύγκριση περιγραμμάτων συνίσταται στα εξής τρία βήματα: αρχικά ανιχνεύεται το περίγραμμα του κεφαλιού χρησιμοποιώντας φίλτρα, ανιχνευτές ακμών ή σιλουέτες (silhouettes). Στη συνέχεια εξάγονται με τον ίδιο τρόπο τα

περιγράμματα τοπικών χαρακτηριστικών με βάση την εκ των προτέρων γνώση για τη γεωμετρία του προσώπου και των χαρακτηριστικών του. Τέλος, υπολογίζεται η συσχέτιση μεταξύ των χαρακτηριστικών που εξάγονται από την εικόνα εισόδου και αποθηκευμένων προκαθορισμένων περιγραμμάτων και αποφασίζεται εάν υπάρχει ή όχι πρόσωπο στην εικόνα. Οι μέθοδοι αυτές παρουσιάζουν ευαισθησία στην κλίμακα, το σχήμα και τις στροφές του κεφαλιού. Προκειμένου να αντιμετωπιστούν τέτοια προβλήματα προτάθηκαν τεχνικές που στηρίζονται σε παραμορφώσιμα περιγράμματα, οι οποίες μοντελοποιούν τη γεωμετρία του προσώπου χρησιμοποιώντας ελαστικά μοντέλα που μπορούν να μετατοπίζονται, να κλιμακώνονται ή να περιστρέφονται. Οι παράμετροι των μοντέλων αυτών είναι δυνατό να περιλαμβάνουν πληροφορία όχι μόνο για το σχήμα, αλλά και για την ένταση της φωτεινότητας των χαρακτηριστικών του προσώπου.

9.5 Ανίχνευση προσώπου με χρήση καθολικών αλγορίθμων

Ενώ οι τεχνικές της προηγούμενης κατηγορίας στηρίζονται σε ένα προκαθορισμένο περίγραμμα ή μοντέλο του προσώπου, οι καθολικές τεχνικές ανίχνευσης προσώπου χρησιμοποιούν μεγάλα σύνολα δειγμάτων (εικόνες προσώπων ή/και εικόνες χαρακτηριστικών του προσώπου) που περιέχουν διάφορες παραλλαγές στο σχήμα του προσώπου, το χρώμα του δέρματος, το χρώμα των ματιών, στις εκφράσεις (κλειστό/ανοιχτό στόμα) κτλ. Η ανίχνευση του προσώπου εξετάζεται σαν ένα πρόβλημα αναγνώρισης προτύπων με δύο κλάσεις: "πρόσωπο" και "μη πρόσωπο". Η κλάση "μη πρόσωπο" περιέχει εικόνες που απεικονίζουν οτιδήποτε εκτός από πρόσωπα, ενώ η κλάση "πρόσωπο" περιέχει μόνο εικόνες προσώπων. Για τον καθορισμό των κατανομών πυκνότητας πιθανότητας των δύο κλάσεων χρησιμοποιούνται κλασικές τεχνικές στατιστικής ανάλυσης και μηχανικής μάθησης. Προκειμένου να ανιχνευθεί η θέση του προσώπου στην εικόνα εισόδου, σαρώνεται ολόκληρη η εικόνα και οι περιοχές της ταξινομούνται ως "πρόσωπο" ή "μη πρόσωπο" με βάση τις συναρτήσεις πυκνότητας πιθανότητας των δύο κλάσεων. Γνωστές καθολικές τεχνικές που χρησιμοποιούνται για την ανίχνευση του προσώπου είναι τα ιδιοπρόσωπα (eigenfaces), ο αλγόριθμος LDA, τα νευρωνικά δίκτυα, οι μηχανές διανυσματικής υποστήριξης (Support Vector Machines) και τα κρυφά μοντέλα Markov.

9.6 Αναγνώριση προσώπου

Ένα Σύστημα Αναγνώρισης Προσώπου είναι μια εφαρμογή των συστημάτων αναγνώρισης προτύπων, που χρησιμοποιείται για αυτοματοποιημένη αναγνώριση ή επιβεβαίωση της ταυτότητας ενός ατόμου από μία ψηφιακή εικόνα ή ένα καρέ από βίντεο. Ένας τρόπος για να γίνει αυτό, είναι η σύγκριση χαρακτηριστικών προσώπου μεταξύ της εισόδου στο σύστημα (στατική εικόνα) και μιας βάσης δεδομένων χαρακτηριστικών. Η αναγνώριση προσώπου αποτελεί την πιο χρησιμοποιούμενη μέθοδο των βιομετρικών συστημάτων, γεγονός που εξηγεί γιατί εταιρείες όπως το Facebook και η Google έχουν επενδύσει σε αυτήν.

9.7 Αναγνώριση φωνής

Τα τελευταία χρόνια έχει επιτευχθεί σημαντική ανάπτυξη της βιομετρικής τεχνολογίας στους τομείς των τραπεζών, στο λιανικό εμπόριο και στην κινητή τηλεφωνία. Η τράπεζα HSBC ανακοίνωσε ότι θα λανσάρει υπηρεσίες ασφαλείας με αναγνώριση φωνής και αφής στη Βρετανία για 15 εκατομμύρια τραπεζικούς πελάτες της, ενώ και ο πάροχος ευρυζωνικών υπηρεσιών TalkTalk έχει επίσης ανακοινώσει τα σχέδια του για χρήση του ίδιου συστήματος.

9.8 Δακτυλικό αποτύπωμα

Το δάκτυλο του κάθε ανθρώπου έχει μοναδικό ανάγλυφο σχήμα στην επιφάνεια του και καθώς χρησιμοποιείται για να πιάνει πράγματα αφήνει μοναδικά σημάδια (αποτυπώματα) στα αντικείμενα τα οποία χειρίζεται. Τα δακτυλικά αποτυπώματα είναι τόσο μοναδικά που χρησιμοποιούνται ως αποδεικτικά στοιχεία από την εγκληματολογία. Η εν λόγω τεχνολογία έχει υιοθετηθεί σήμερα και χρησιμοποιείται ευρέως σε διάφορες συσκευές όπως smartphone, tablet, laptop κ.ά.

9.9 Αναγνώριση βάση συμπεριφοράς

Σε αυτό το πεδίο η ταυτοποίηση πραγματοποιείται με βάση τις κινήσεις και την ανθρώπινη δραστηριότητα των ανθρώπων. Ο όρος έρχεται σε αντίθεση με τη φυσική βιομετρία, η οποία περιλαμβάνει ανθρώπινα χαρακτηριστικά, όπως δακτυλικά αποτυπώματα ή μοτίβα ίριδας. Οι μέθοδοι βιομετρικής επαλήθευσης συμπεριφοράς περιλαμβάνουν τον τρόπο με τον οποίο κάποιος βαδίζει, τα αναγνωριστικά φωνής, την ανάλυση υπογραφών κ.ά. Τα βιομετρικά στοιχεία συμπεριφοράς χρησιμοποιούνται για ασφαλή έλεγχο ταυτότητας σε χρηματοπιστωτικά ιδρύματα, επιχειρήσεις, κυβερνητικές εγκαταστάσεις, σημεία πώλησης κ.ά.

9.10 Δακτυλογράφηση

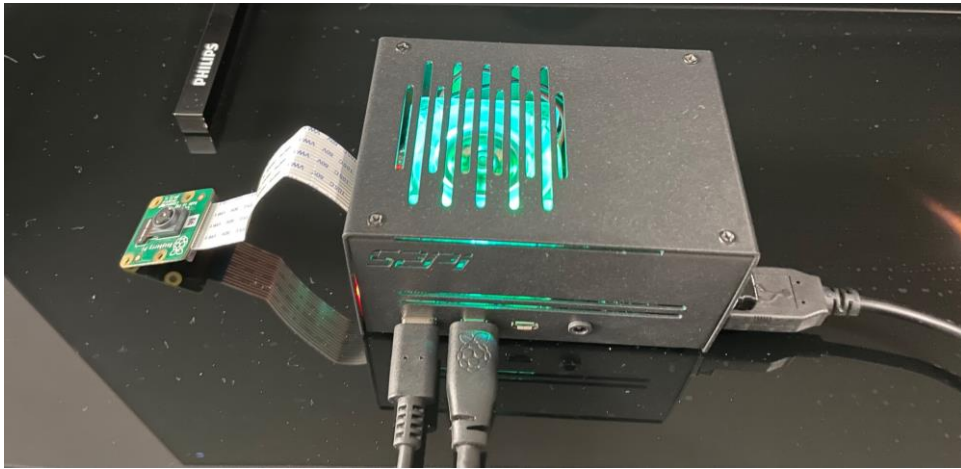
Ο ρυθμός δακτυλογράφησης και ο τρόπος με τον οποίο ο καθένας από εμάς χρησιμοποιεί το πληκτρολόγιο είναι μοναδικός και απίστευτα δύσκολος στην παραποίηση όταν χρησιμοποιείται σε συνδυασμό με άλλες μορφές ταυτοποίησης, όπως κωδικούς πρόσβασης. Οι ρυθμοί πληκτρολόγησης των χρηστών «μετρούνται» για να αναπτύξουν ένα μοναδικό βιομετρικό πρότυπο για τον χρήστη, ώστε να χρησιμοποιηθεί μελλοντικά για πιστοποίηση. Για παράδειγμα το σύστημα θα είναι σε θέση να διαχωρίσει το άτομο που πάτησε το πλήκτρο «Caps Lock», από το άτομο που είχε πατημένο το shift για να χρησιμοποιήσει κεφαλαία γράμματα.

9.11 Αναγνώριση ίριδας

Ανάμεσα σε όλες τις σχετικές διαθέσιμες τεχνολογίες βιομετρικών δεδομένων, η πιο διαδεδομένη και ταχέως εξελισσόμενη είναι η αναγνώριση της ίριδας του ματιού. Πρόκειται για μια τεχνολογία τόσο γρήγορη και απλή, όσο και το τράβηγμα μιας selfie, και δεν απαιτεί καμία φυσική επαφή με κάποιον αισθητήρα. Η ίριδα του κάθε ματιού είναι μοναδική, δεν αντιγράφεται, ούτε μπορεί να κλαπεί από κανέναν. Από την άλλη, τα μηχανήματα ανάγνωσης ίριδας δεν είναι δυνατό να «εξαπατηθούν» από χρήση μείκ-απ στο δέρμα, αλλαγή χτενίσματος ή κουρέματος μαλλιών ή μεταμφίεση σε ρούχα. Η

χρησιμοποίηση σκάνερ ίριδας στα έξυπνα κινητά βοήθησε επίσης στο να υπάρξει ευρεία αποδοχή της συγκεκριμένης τεχνολογίας. Επιπλέον, δε χρειάζεται καλωδιώσεις και αυτό είναι μια σημαντική παράμετρος που προσθέτει σημαντική ευκολία στη χρήση της. Τα υποκείμενα της αναγνώρισης μπορούν να προσεγγίζουν ένα συγκεκριμένο σημείο εισόδου, με τα χέρια τους φορτωμένα με αντικείμενα όπως βιβλία μέχρι αποσκευές, χωρίς να υποχρεούνται να επιδείξουν κάποια πιστοποιημένη κάρτα.

Κεφάλαιο 10ο: Αναγνώριση προσώπου η υλοποίηση με raspberry pi 4



8. Εικόνα του raspberry pi 4 πλήρως συναρμολογημένη και σε λειτουργία

Το παρακάτω κεφάλαιο θα αναλύσουμε όλη τη διαδικασία για την υλοποίηση της εφαρμογής αναγνώρισης προσώπου. Στα προηγούμενα κεφάλαια επικεντρωθήκαμε κυρίως στα θεωρητικά μέρη που κατά βάση είναι συγκεκριμένα και τυποποιημένα, είναι ορισμοί που κάθε επιστήμονας θα χρησιμοποιούσε αυτούσια. Θα ξεκινήσω από τα βασικά στοιχεία της εγκατάστασης των βιβλιοθηκών και θα συνεχίσω με την παρουσίαση του κώδικα και τέλος θα ολοκληρώσω με την εκτέλεση του κώδικα και την αντίστοιχη εμφάνιση των αποτελεσμάτων.

Για την αναγνώριση προσώπου με τη χρήση του Raspberry Pi 4, θα χρησιμοποιήσουμε τη βιβλιοθήκη OpenCV, face_recognition και imutils για να εκπαιδεύσουμε το Raspberry Pi με βάση ένα σύνολο εικόνων που θα συλλέγουμε και θα παρέχουμε ως σύνολο δεδομένων. Θα εκτελέσουμε train_model.py για να αναλύσουμε τις εικόνες στο σύνολο δεδομένων και να δημιουργήσουμε αντιστοίχιση μεταξύ ονομάτων και πρόσωπα με το αρχείο encodings.pickle.

Αφού ολοκληρώσουμε την εκπαίδευση του raspberry pi, θα εκτελέσουμε το αρχείο facial_req.py ώστε να εντοπίσουμε και να ταυτοποιήσουμε τα πρόσωπα. Τέλος θα αποστέλλεται ηλεκτρονικό μήνυμα στην περίπτωση αναγνώρισης προσώπου. Φυσικά θα μπορούσε να στέλνεται ηλεκτρονικό μήνυμα και στην περίπτωση εντοπισμού άγνωστου προσώπου.

10.1 Εγκατάσταση βιβλιοθηκών

Παρακάτω θα αναφέρω τα βήματα της εγκατάστασης των απαραίτητων βιβλιοθηκών και την προετοιμασία του raspberry pi για μηχανική εκμάθηση και αναγνώριση προσώπου ώστε τελικά να είναι εφικτή αναγνώριση προσώπων μέσω της κάμερας. Θα χρησιμοποιήσω το πακέτο Python face_recognition για τον υπολογισμό του πλαισίου οριοθέτησης γύρω από κάθε πρόσωπο. Το Imutils είναι μια σειρά λειτουργιών ευκολίας για την επιτάχυνση του υπολογισμού της βιβλιοθήκης OpenCV

στο Raspberry Pi, δεδομένου ότι έχουμε περιορισμένους πόρους, τέλος συνδέουμε στην θύρα USB την Raspberry Pi 8MP Camera board v2 και εκτελούμε τις παρακάτω εντολές στο τερματικό.

Για την εγκατάσταση της βιβλιοθήκης OpenCV θα πρέπει να εκτελεστούν στο τερματικό οι παρακάτω εντολές,

- sudo apt install cmake buildessential pkgconf git
- sudo apt install libjpeg-dev libtiff-dev libjasper-dev libpng-dev libwebp-dev libopenexr-dev
- sudo apt install libavcodec-dev libavformat-dev libswscale-dev libv4l-dev libxvidcore-dev libx264-dev libdc1394-22-dev libgstreamer-plugins-base1.0-dev libgstreamer1.0-dev
- sudo apt install libgtk-3-dev libqtgui4 libqtwebkit4 libqt4-test python3-pyqt5
- sudo apt install libatlas-base-dev liblapack-dev gfortran
- sudo apt install libhdf5-dev libhdf5-103
- sudo apt install python3-dev python3-pip python3-numpy
- sudo nano /etc/dphys-swapfile (αλλάζω τη γραμμή CONF_SWAPSIZE=100 σε CONF_SWAPSIZE=2048 ώστε να επεκτείνω το αρχείο)
- sudo systemctl restart dphys-swapfi (για να αποθηκευτούν οι αλλαγές μου, κάνω την επανέναρξη)
- git clone https://github.com/opencv/opencv.git
- git clone https://github.com/opencv/opencv_contrib.git
- mkdir ~/opencv/build
- cd ~/opencv/build
- cmake -D CMAKE_BUILD_TYPE=RELEASE \
- -D CMAKE_INSTALL_PREFIX=/usr/local \
- -D OPENCV_EXTRA_MODULES_PATH=~/opencv_contrib/modules \
- -D ENABLE_NEON=ON \
- -D ENABLE_VFPV3=ON \
- -D BUILD_TESTS=OFF \
- -D INSTALL_PYTHON_EXAMPLES=OFF \
- -D OPENCV_ENABLE_NONFREE=ON \
- -D CMAKE_SHARED_LINKER_FLAGS=-latomic \
- -D BUILD_EXAMPLES=OFF
- make -j\$(nproc)
- sudo make install

– sudo ldconfig

Αφού ολοκληρωθεί η εγκατάσταση του OpenCV, εκτελούμε τις παρακάτω εντολές στο τερματικό

```
sudo nano /etc/dphys-swapfile (αλλάζω ξανά τη γραμμή σε CONF_SWAPSIZE=100)
```

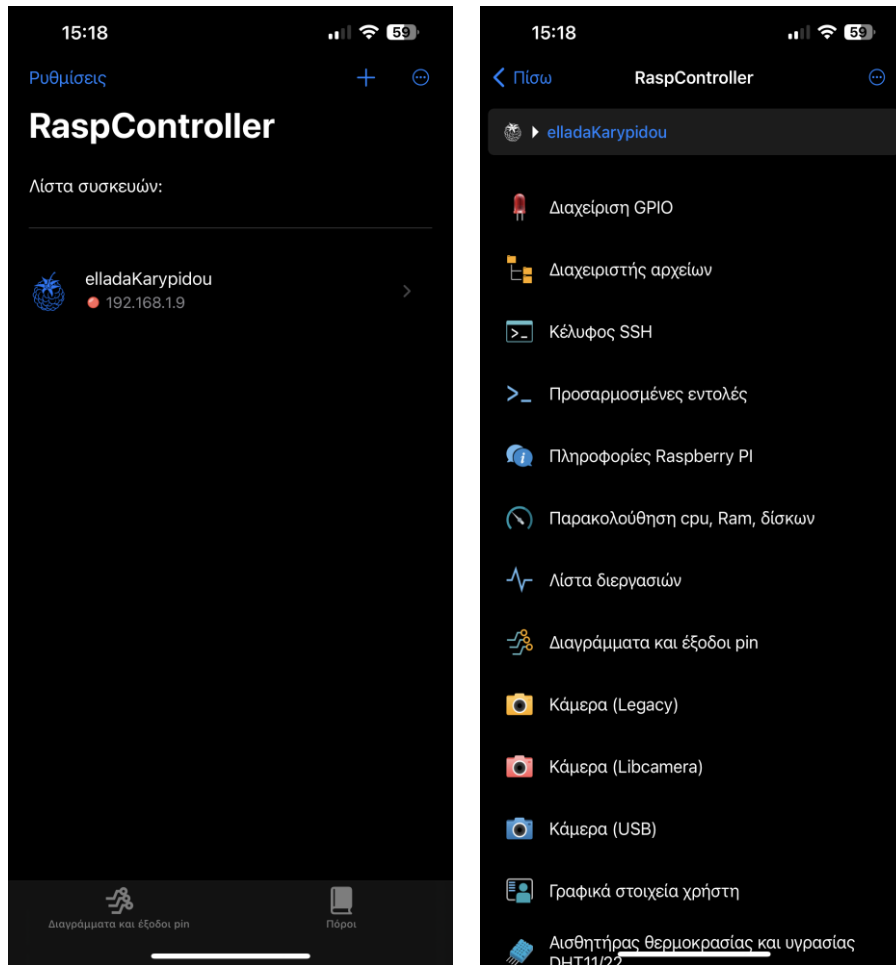
```
sudo systemctl restart dphys-swapfil (κάνω επανεκκίνηση ώστε να πάρει τις αλλαγές)
```

10.2 RaspController-Απομακρυσμένη διαχείριση του raspberry pi

Επειδή η εργασία έχει το κομμάτι της διαχείρισης του raspberry pi με συσκευές android, παρακάτω θα δούμε πως μπορούμε κάνοντας κάποιες τροποποιήσεις στο raspberry pi 4 να την συνδέσουμε με την mobile εφαρμογή RaspController. Με τη σύνδεση αυτή κάνουμε εφικτή την απομακρυσμένη διαχείριση του raspberry pi μέσα από android συσκευές . Η εφαρμογή προσφέρει εύκολη διαχείριση αρχείων, εκτέλεση του τερματικού κ.α., χρησιμοποιώντας το πρωτόκολλο SSH. Η δημιουργός της εφαρμογής είναι η ETTORE GALLINA. Υπάρχουν πολλές mobile εφαρμογές για την απομακρυσμένη διαχείριση του raspberry pi, εγώ θα μείνω εδώ καθώς μπορώ να κάνω τις λειτουργίες που χρειάζομαι. Παρακάτω θα δείξω φωτογραφίες από το interface της mobile εφαρμογής και ορισμένα παραδείγματα λειτουργίας, όπως της εκτέλεσης εντολών στο τερματικό.



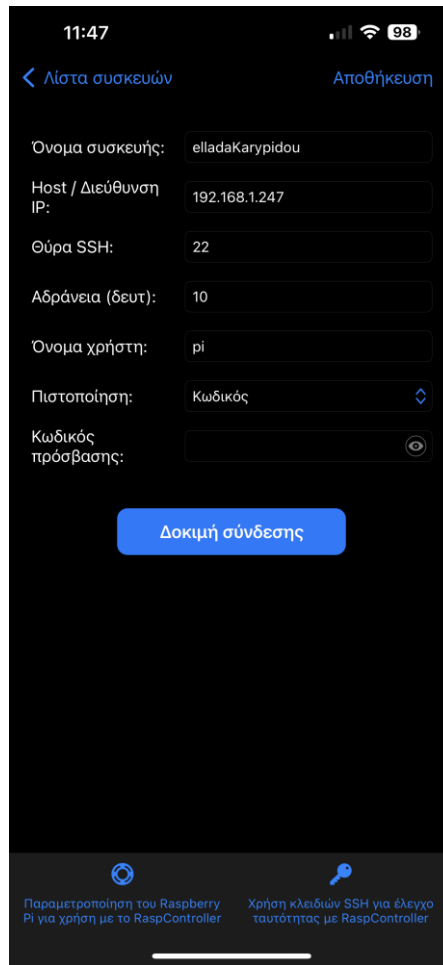
9. Εικόνα εφαρμογής RaspController



10. Εικόνα με στιγμιότυπα από την mobile εφαρμογή στο κινητό μου

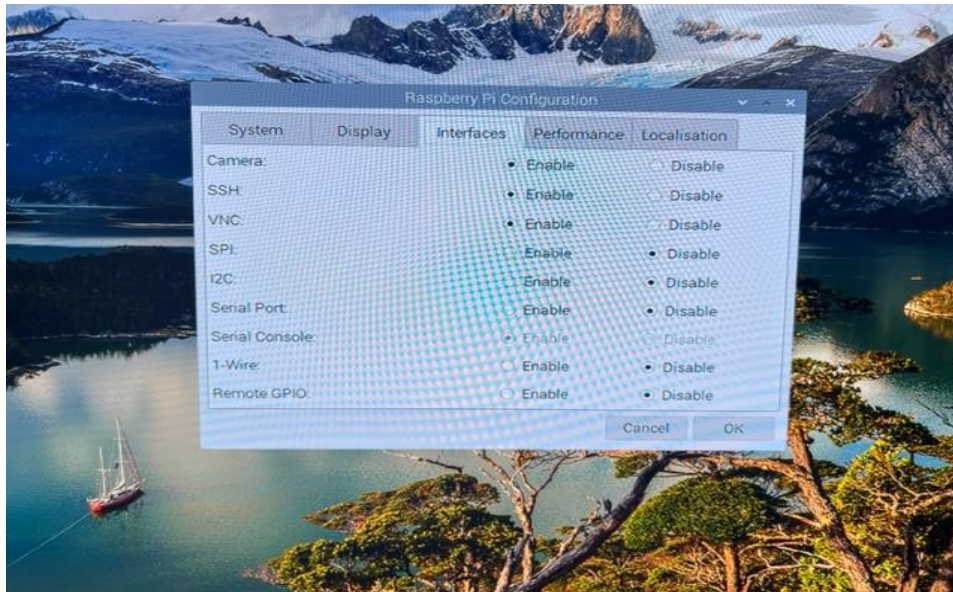
Παραπάνω φαίνονται δύο εικόνες για το πως είναι το interface της mobile εφαρμογής. Μετά από παραμετροποιήσεις του raspberry pi και της mobile εφαρμογής πλέον οι δύο συσκευές λειτουργούν και έχουν συνδεθεί. Να αναφέρω ότι για την ασφάλεια της σύνδεσης χρησιμοποίησα username και password, οπότε και στην περίπτωση που κάποιος επιτήδειος βρει την IP διεύθυνση της συσκευής δεν θα είναι εύκολο να βρεθούν τα υπόλοιπα στοιχεία.

Στην δεξιά εικόνα βλέπουμε ότι υπάρχει η δυνατότητα για απομακρυσμένη χρήση της κάμερας, του κελύφους SSH κ.α. Για να μπορέσει λοιπόν να γίνει η σύνδεση-επικοινωνία πρέπει γνωρίζουμε την IP διεύθυνση του raspberry pi, το username και το password. Υπάρχουν δύο τρόποι εύρεσης της IP, είτε εκτελώντας την εντολή ifconfig στο cmd, είτε δεξιά πάνω στην οθόνη του raspberry pi πατώντας στο σήμα του wifi. Σε λειτουργία που θα είναι μόνιμα σε συγκεκριμένο χώρο θα ήταν ιδανικό να ρυθμίσουμε τη συσκευή με στατική IP, εδώ για τις ανάγκες της εργασίας δεν θα την κάνω στατική διότι θα χρειαστεί να συνδεθώ σε άλλο δίκτυο, για την παρουσίαση.



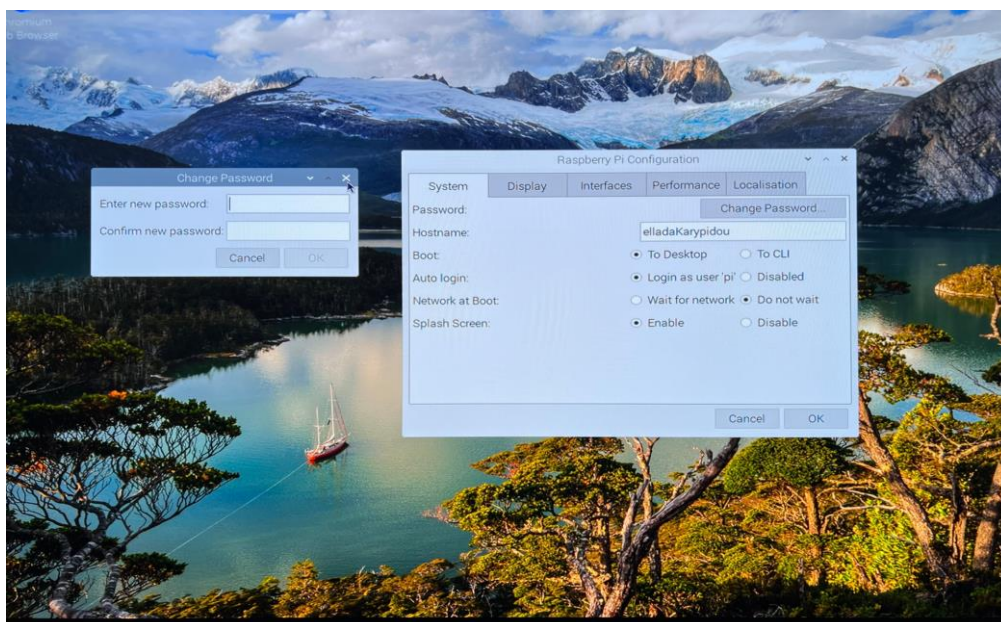
11. Εικόνα από το παράθυρο ρυθμίσεων του RaspController

Στην παραπάνω εικόνα βλέπουμε ότι εκτός από την IP διεύθυνση υπάρχει η πιστοποίηση με κωδικούς, πολύ σημαντικό διότι μας προσφέρει επιπλέον ασφάλεια. Για να δημιουργήσουμε την αντιστοιχία των κωδικών αυτών, στο παράθυρο του raspberry pi configuration- system γράφουμε ένα hostname και password. Αυτά θα χρησιμοποιήσουν στις ρυθμίσεις του RaspController. Οπότε ξέροντας την IP και τα στοιχεία hostname και password, είμαστε έτοιμη να συνδέσουμε τις δύο συσκευές. Η σύνδεση λειτουργεί και πλέον μπορούμε χειριστούμε απομακρυσμένα το raspberry pi.

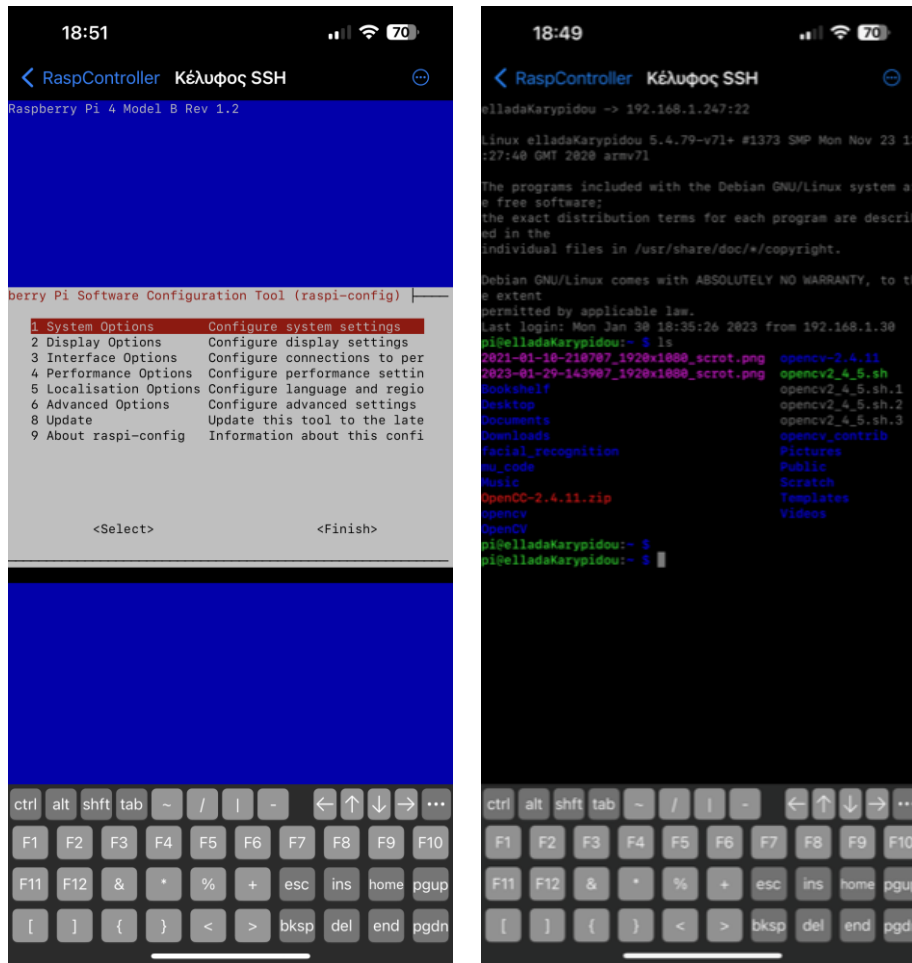


12. Εικόνα από το παράθυρο interfaces -configuration του Raspberry pi

Θα κάνουμε κάποιες τροποποιήσεις του raspberry pi configuration θα διαμορφώσουμε τα interfaces, θα πρέπει να τεθεί σε κατάσταση enable τα SSH, Camera VNC και serial console. Στην παρακάτω εικόνα δείχνω τις ρυθμίσεις των interfaces του raspberry pi.



13. Εικόνα με τα system- configuration του raspberry pi



14. Εικόνες από την εφαρμογή RaspController

Στις παραπάνω εικόνες βλέπουμε στιγμιότυπα από την εφαρμογή RaspController, έχω εκτελέσει στο κέλυφος SSH την εντολή `-ls` και βλέπουμε ότι εμφανίζεται η λίστα με τους καταλόγους που είχα δημιουργήσει. Υπάρχουν πολλές ακόμα δυνατότητες απομακρυσμένης διαχείρισης όπως να ανοίξω την κάμερα και να βλέπω ζωντανά την εικόνα του σπιτιού μου, να παρακολουθήσω τις διεργασίες που εκτελούνται, να παρακολουθήσω τη λειτουργία της CPU, RAM και των δίσκων κ.α.. Σε μελλοντικό χρόνο θα μπορούσα να συνδέσω αισθητήρες στο raspberry pi, όπως αισθητήρες θερμοκρασίας, υγρασίας, κίνησης κλπ ώστε να μπορώ να έχω επίσης τον έλεγχο αυτών.

10.3 Δημιουργία καταλόγων

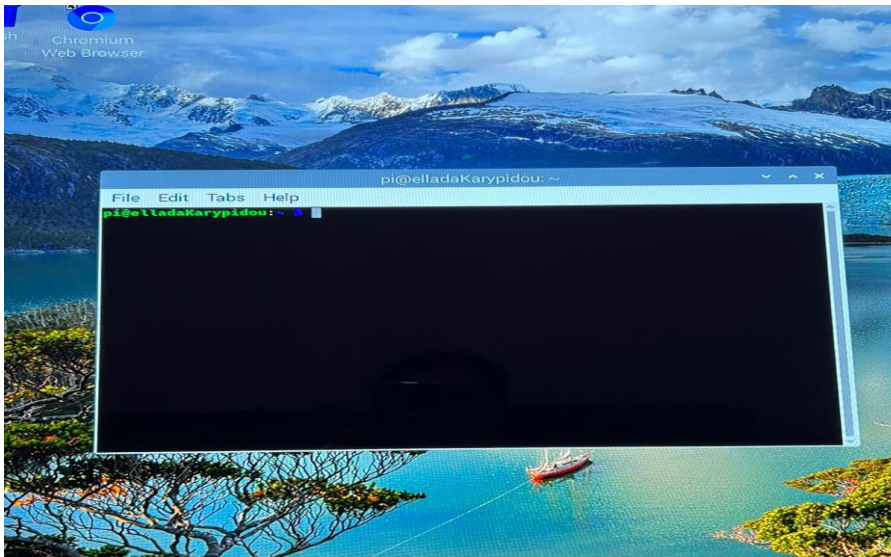
Αφού ολοκληρώθηκε επιτυχώς η εγκατάσταση βιβλιοθηκών, η προετοιμασία του raspberry pi, και η επιτυχής απομακρυσμένη διαχείριση, είμαστε έτοιμοι να προχωρήσουμε. Θα δημιουργήσουμε ένα φάκελο με το όνομα `facial_recognition` και μέσα υποφακέλους με τον κώδικα και τα dataset - στιγμιότυπα προσώπου. Θα δημιουργήσω ένα υποφάκελο με όνομα `dataset` και μέσα σε αυτό έναν

φάκελο με το όνομα μου, στον οποίο θα αποθηκεύσω στιγμιότυπα από το πρόσωπό μου. Πολύ σημαντικό είναι να έχω αποθηκεύσει αρκετές φωτογραφίες σε διαφορετικές οπτικές γωνίες ώστε να έχω καλό δείγμα φωτογραφιών για να έχουμε καλύτερη εκπαίδευση και αντίστοιχα καλύτερα αποτελέσματα στην αναγνώριση με λιγότερες αστοχίες. Αυτή η διαδικασία θα μπορούσε να γίνεται για περισσότερα άτομα αρκεί να εκπαιδεύσουμε τον αλγόριθμο με τα υπόλοιπα πρόσωπα. Θα συνδέσουμε στη USB θύρα την κάμερα.

Θα δημιουργήσω τους παρακάτω φακέλους και υποφακέλους:

facial_recognition

- dataset
 - Ellada_Karypidou (φάκελος στο οποίο αποθηκεύονται οι φωτογραφίες)
- facial_reg.py (αρχείο python που περιέχει τον κώδικα αναγνώρισης προσώπου)
- facial_reg_email.py (αρχείο python που περιέχει τον κώδικα αποστολής ειδοποίησης email)
- headshots.py (αρχείο python που περιέχει τον κώδικα αποθήκευσης φωτογραφιών)
- train_model.py (αρχείο python που εκπαιδεύει τον αλγόριθμο)



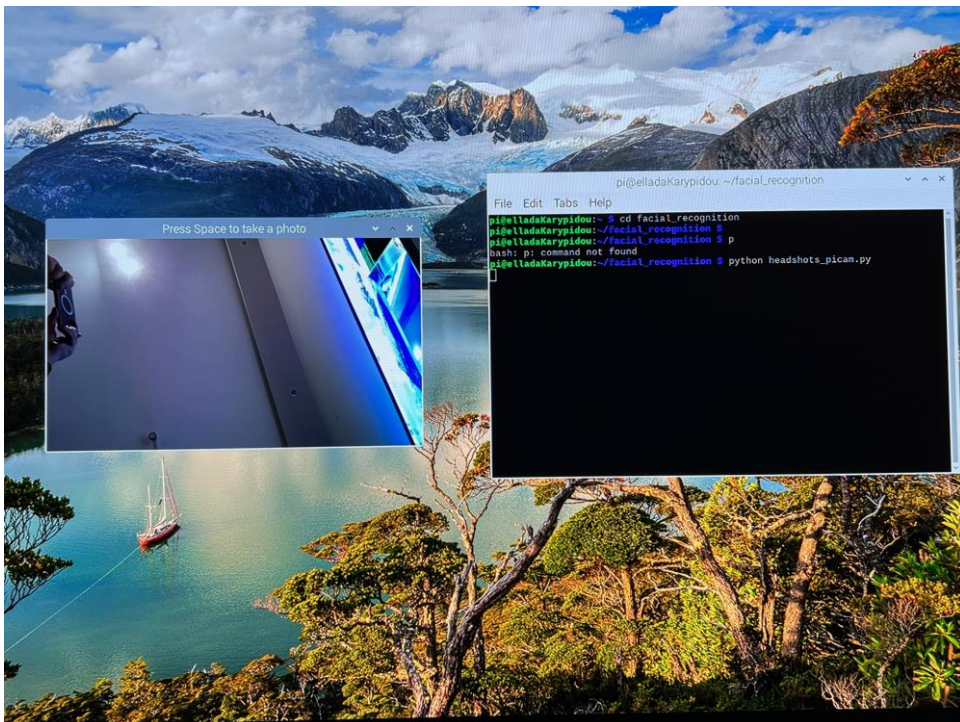
15. Εικόνα από το τερματικό του raspberry pi

10.4 Βήμα πρώτο -Αποθήκευση στιγμιότυπων προσώπου

Ξεκινάμε με την αποθήκευση φωτογραφιών και εκτελούμε με τη σειρά, η κάμερα ίσως κάνει λίγα δευτερόλεπτα ώστε να ανοίξει. Τοποθετώ την κάμερα σε σημείο που να φαίνεται το πρόσωπό μου και πατάω το πλήκτρο space ώστε να αποθηκεύονται οι φωτογραφίες-datasets. Είναι πολύ σημαντικό να

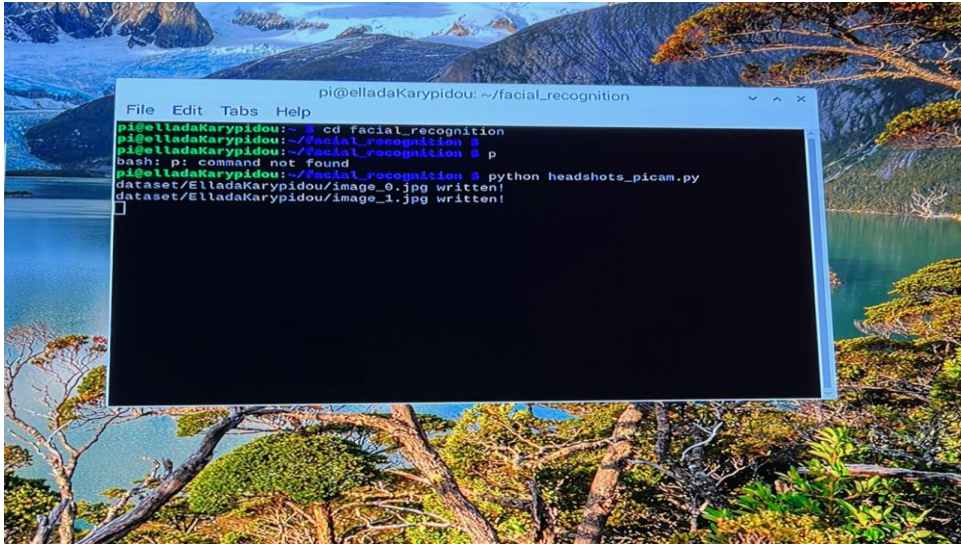
βγάλω αρκετές φωτογραφίες και σε διάφορες γωνίες του προσώπου, ώστε να υπάρχει μεγαλύτερο δείγμα άρα καλύτερη και ποσοστιαία πιθανότερη αναγνώριση του προσώπου. Όπως ανέφερα ήδη παραπάνω και σύμφωνα με την καθηγήτρια του Πανεπιστημίου Stanford, FEI FEILI (<https://engineering.stanford.edu/people/fei-fei-li>) δεν αρκεί να βελτιωθούν οι αλγόριθμοι αλλά πρέπει να υπάρχουν πολλά δεδομένα για να επιτύχουμε καλύτερα αποτελέσματα. Έτσι και εδώ, θα φροντίσω να έχω πολλές φωτογραφίες μου σε διαφορετικές γωνίες του πρόσωπό μου. Αν δεν είμαστε ήδη θα πρέπει να μεταφερθούμε στον κατάλογο γράφοντας στο τερματικό cd facial_recognition. Εκτελούμε τον κώδικα, γράφουμε στο τερματικό την παρακάτω εντολή:

- python headshots.py



16. Στην εικόνα βλέπουμε την εκτέλεση του κώδικα headshots.py

Βλέπουμε ότι με την εκτέλεση του κώδικα headshots.py ανοίγει στα αριστερά το παράθυρο της κάμερας. Εδώ λοιπόν, αν βάλουμε την κάμερα με κατεύθυνση στο πρόσωπό μας και πατήσουμε το πλήκτρο space θα αρχίσει να αποθηκεύει τις φωτογραφίες στο φάκελο.



17. Εικόνα του τερματικού που αποθηκεύει την φωτογραφία

Από τη στιγμή που πατήσουμε το space αποθηκεύεται το στιγμιότυπο του προσώπου στον φάκελο dataset, υποφάκελο Ellada Karypidou. Βλέπουμε στην εικόνα αυτήν ακριβώς τη διαδικασία. Στον φάκελο Ellada Karypidou, πλέον υπάρχουν όλες οι φωτογραφίες που έχουν αποθηκευτεί.

10.5 Βήμα δεύτερο- Εκπαίδευση αλγόριθμου

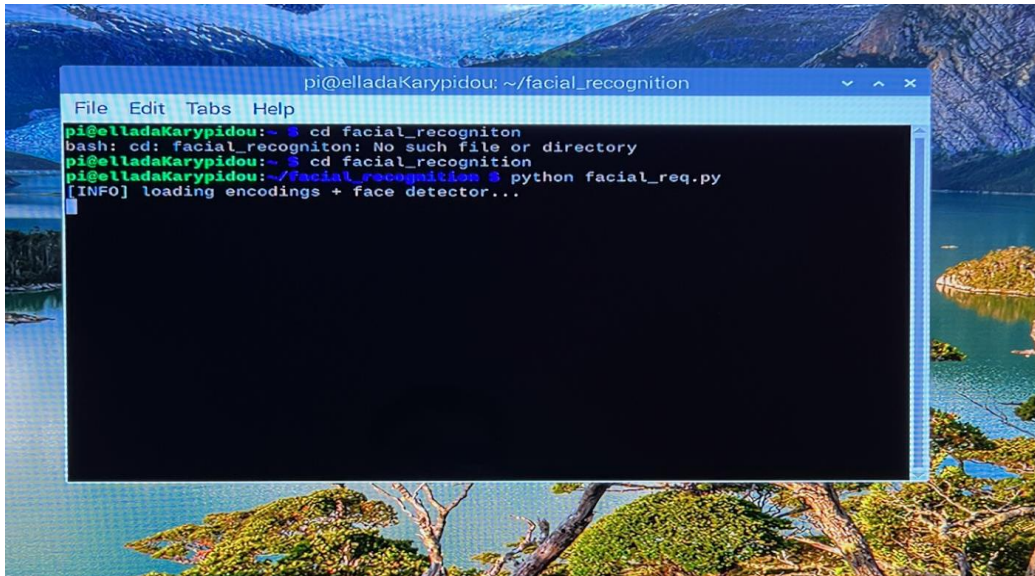
Τώρα μπορούμε να εκπαιδεύσουμε τον αλγόριθμό. Εκτελούμε τις παρακάτω εντολές στο τερματικό. Ίσως χρειαστούν λίγα λεπτά ώστε το raspberry pi να αναλύσει τις φωτογραφίες που αποθήκευσα και να δημιουργήσει το encodings.pickle, (το train_model.py θα δημιουργήσει ένα φάκελο encodings.pickle που θα περιέχει τα κριτήρια για την ταυτοποίηση προσώπων). Εκτελούμε τον κώδικα, γράφουμε στο τερματικό:

- `python train_model.py`

10.6 Βήμα τρίτο- Αναγνώριση προσώπου

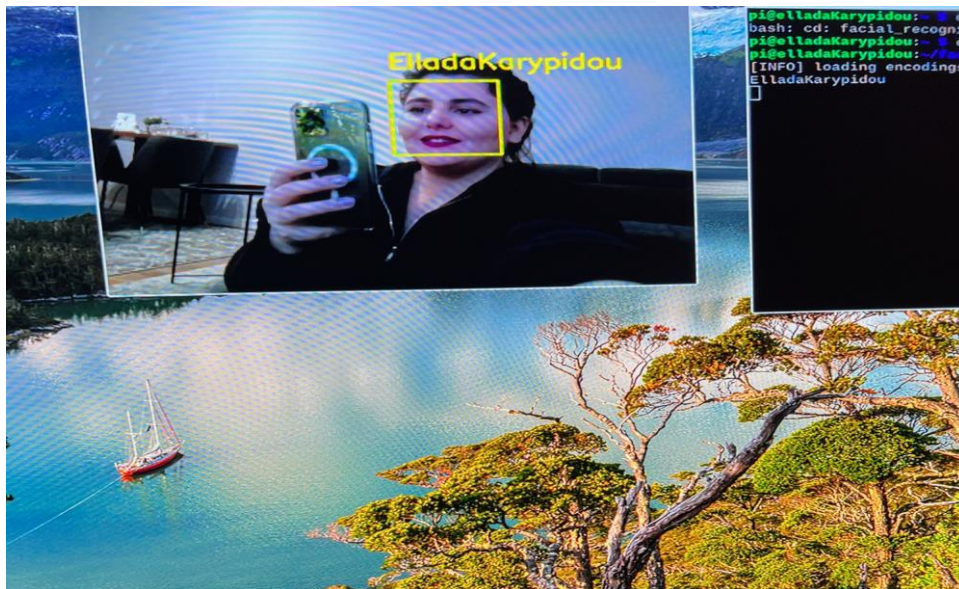
Τώρα είμαστε έτοιμοι να εκτελέσουμε το κύριο κομμάτι για την αναγνώριση προσώπου. Αν έχει εκπαιδευτεί σωστά το μοντέλο τότε θα πρέπει να υπάρχει η οριοθέτηση του προσώπου και αντίστοιχα να γράφει το όνομά μου, σε διαφορετική περίπτωση θα θεωρείται άγνωστο πρόσωπο. Γράφουμε στο τερματικό το παρακάτω

- `python facial_req.py`



18. Εικόνα του τερματικού εκτέλεσης της αναγνώρισης προσώπου

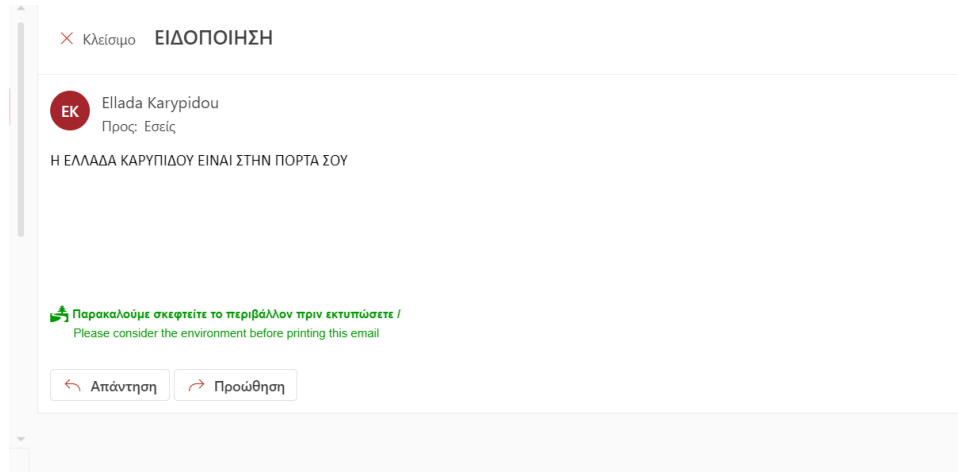
Στην εικόνα βλέπουμε το τελικό κομμάτι της εκτέλεσης του αλγορίθμου, της αναγνώρισης του προσώπου. Σε αυτήν την φάση ανοίγει πάλι η κάμερα ώστε να δεχτεί την εικόνα, οροθετείται το πρόσωπο και στην περίπτωση που αναγνωριστεί εμφανίζεται το όνομά μου, σε διαφορετική περίπτωση εμφανίζεται “UNKNOWN”. Επειδή έχω φροντίσει να έχω πολλές φωτογραφίες σε πολλές γωνίες το πιθανότερο είναι ο αλγόριθμος να με αναγνωρίσει.



19. Εικόνα εκτέλεσης του αλγορίθμου και αναγνώριση προσώπου

Στην παραπάνω φωτογραφία βλέπουμε ότι πράγματι ο αλγόριθμος με αναγνωρίζει, τοποθετεί την οριοθέτηση στο πρόσωπο και εμφανίζει το όνομά μου.

Στην περίπτωση που γίνει η αναγνώριση, στέλνεται email με την ειδοποίηση “Η ΕΛΛΑΔΑ ΚΑΡΥΠΙΔΟΥ ΕΙΝΑΙ ΣΤΗΝ ΠΟΡΤΑ ΣΟΥ”. Για τις ανάγκες της εργασίας στο email έχω ορίσει αποστολέα και παραλήπτη εμένα, φυσικά τα στοιχεία είναι παραμετροποιήσιμα και θα μπορούσε να σταλεί ειδοποίηση σε οποιοδήποτε email.



20. Εικόνα ειδοποίησης με ηλεκτρονικό μήνυμα για την αναγνώριση προσώπου

Κεφάλαιο 11ο: Κώδικας υλοποίησης

11.1 facial_req.py

```
#!/usr/bin/python
# ΕΙΣΑΓΩΓΗ ΑΠΑΡΑΙΤΗΤΩΝ ΒΙΒΛΙΟΘΗΚΩΝ

from imutils.video import VideoStream
from imutils.video import FPS
import face_recognition
import imutils
import pickle
import time
import cv2
```

```
#ΑΡΧΙΚΟΠΟΙΗΣΗ 'currentname' ΩΣΤΕ ΝΑ ΕΝΕΡΓΟΠΟΙΗΘΕΙ ΜΟΝΟ ΣΤΗΝ ΑΝΑΓΝΩΡΙΣΗ ΠΡΟΣΩΠΟΥ.
```

```
currentname = "unknown"
```

```
# ΚΑΘΟΡΙΣΜΟΣ ΠΡΟΣΩΠΩΝ ΑΠΟ ΤΟ ΑΡΧΕΙΟ encodings.pickle
```

```
encodingsP = "encodings.pickle"
```

```
#ΦΟΡΤΩΣΗ ΓΝΩΣΤΩΝ ΠΡΟΣΩΠΩΝ
```

```
print("[INFO] loading encodings + face detector...")
```

```
data = pickle.loads(open(encodingsP, "rb").read())
```

```
#ΑΡΧΙΚΟΠΟΙΗΣΗ ΤΟΥ VIDEO STREAMING ΚΑΙ ΠΡΟΕΤΟΙΜΑΣΙΑ ΤΗΣ ΚΑΜΕΡΑΣ
```

```
# src = 0 : ΠΗΓΗ ΚΑΜΕΡΑΣ ΤΟΥ LAPTOP
```

```
# src = 2 : ΠΗΓΗ ΚΑΜΕΡΑΣ USB WEBCAM
```

```
vs = VideoStream(src=2,framerate=10).start()
```

```
# ΕΝΑΡΞΗ FPS ΜΕΤΡΗΤΗ
```

```
fps = FPS().start()
```

```
while True:
```

```
# ΟΡΙΟΘΕΤΗΣΗ ΠΛΑΙΣΙΟΥ ΚΑΙ ΟΡΙΣΜΟΥ ΜΕΓΕΘΟΥΣ
```

```
frame = vs.read()
```

```
frame = imutils.resize(frame, width=500)
```

```
# ΕΝΤΟΠΙΣΜΟΣ ΤΟΥ ΠΛΑΙΣΙΟΥ ΠΡΟΣΩΠΟΥ
```

```
boxes = face_recognition.face_locations(frame)
```

```
encodings = face_recognition.face_encodings(frame, boxes)
```

```
names = []
```

```
for encoding in encodings:
```

```
# ΑΝΤΙΣΤΟΙΧΙΣΗ ΠΡΟΣΩΠΟΥ
```

```
matches = face_recognition.compare_faces(data["encodings"],
name = "Unknown" # ΜΗ ΑΝΑΓΝΩΡΙΣΗ ΠΡΟΣΩΠΟΥ

# ΕΛΕΓΧΟΣ ΕΥΡΕΣΗΣ ΑΝΤΙΣΤΟΙΧΙΑΣ ΠΡΟΣΩΠΟΥ
if True in matches:
#ΑΝ ΟΙ ΔΕΙΚΤΕΣ ΤΑΙΡΙΑΞΟΥΝ ->ΑΡΧΙΚΟΠΟΙΗΣΗ
matchedIdxs = [i for (i, b) in enumerate(matches) if b]
counts = {}

#ΑΡΙΘΜΗΣΗ ΣΤΙΣ ΦΩΤΟΓΡΑΦΙΕΣ ΠΟΥ ΑΝΑΓΝΩΡΙΣΤΗΚΑΝ
for i in matchedIdxs:
name = data["names"][i]
counts[name] = counts.get(name, 0) + 1

# ΚΑΘΟΡΙΣΜΟΣ ΑΝΑΓΝΩΡΙΣΜΕΝΟΥ ΠΡΟΣΩΠΟΥ ΜΕ ΤΗΝ ΜΕΓΑΛΥΤΕΡΗ
ΒΑΘΜΟΛΟΓΙΑ (ΣΕ ΠΕΡΙΠΤΩΣΗ ΙΣΟΒΑΘΜΙΑΣ Η ΠΡΩΤΗ ΕΓΓΡΑΦΗ )
name = max(counts, key=counts.get)

# ΣΤΗΝ ΠΕΡΙΠΤΩΣΗ ΑΝΑΓΝΩΡΙΣΗΣ ΕΜΦΑΝΙΣΗ ΟΝΟΜΑΤΟΣ ΣΤΗΝ ΟΘΟΝΗ
if currentname != name:
currentname = name
print(currentname)

# UPDATE ΤΗ ΛΙΣΤΑ ΟΝΟΜΑΤΩΝ
names.append(name)
for ((top, right, bottom, left), name) in zip(boxes, names):

#ΕΜΦΑΝΙΣΗ ΟΝΟΜΑΤΟΣ ΑΝΑΓΝΩΡΙΣΜΕΝΟΥ ΠΡΟΣΩΠΟΥ
cv2.rectangle(frame, (left, top), (right, bottom), (0, 255, 225), 2)
y = top - 15 if top - 15 > 15 else top + 15 cv2.putText(frame, name, (left, y),
cv2.FONT_HERSHEY_SIMPLEX, .8, (0, 255, 255), 2)
```

```
# ΕΜΦΑΝΙΣΗ ΕΙΚΟΝΑΣ ΣΤΗΝ ΟΘΟΝΗ
cv2.imshow("Facial Recognition is Running", frame)
key = cv2.waitKey(1) & 0xFF

# ΤΕΡΜΑΤΙΣΜΟΣ ΜΕ ΤΟ ΠΑΤΗΜΑ ΠΛΗΚΤΡΟΥ 'q'
if key == ord("q"):
    break

# UPDATE FPS ΜΕΤΡΗΤΗ
fps.update()

# ΤΕΡΜΑΤΙΣΜΟΣ ΤΟ ΧΡΟΝΟΜΕΤΡΟΥ FPS ΠΛΗΡΟΦΟΡΙΩΝ
fps.stop()
print("[INFO] elapsed time: {:.2f}".format(fps.elapsed()))
print("[INFO] approx. FPS: {:.2f}".format(fps.fps()))
cv2.destroyAllWindows()
vs.stop()
```

11.2 facial_req_email.py

```
#!/usr/bin/python
# ΕΙΣΑΓΩΓΗ ΑΠΑΡΑΙΤΗΤΩΝ ΠΑΚΕΤΩΝ
from imutils.video import VideoStream
from imutils.video import FPS
import face_recognition
import imutils
import pickle
import time
import cv2
import requests
```

```
#ΑΡΧΙΚΟΠΟΙΗΣΗ 'currentname'
currentname = "unknown"

# ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΠΡΟΣΩΠΟΥ ΑΠΟ ΤΟ ΑΡΧΕΙΟ encodings.pickle ΠΟΥ
ΔΗΜΙΟΥΡΓΗΘΗΚΕ ΑΠΟ ΤΟ train_model.py
encodingsP = "encodings.pickle"
cascade = "haarcascade_frontalface_default.xml"

#ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ EMAIL
def send_message(name):
return requests.post(

#ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ email ,
"https://api.google.com/v3/email@hostname.com/messages",
auth=("api", "my_API_KEY"),
files = [("attachment", ("image.jpg", open("image.jpg", "rb").read()))],
data={"from": 'hello@example.com',
"to": [MY_EMAIL_ADDRESS]},
"subject": "ΕΧΕΙΣ ΕΠΙΣΚΕΠΤΗ",
"html": "<html>" + name + " ΕΙΝΑΙ ΣΤΗΝ ΠΟΡΤΑ ΣΟΥ </html>"}

# ΦΟΡΤΩΣΗ ΚΑΙ ΕΝΣΩΜΑΤΩΣΗ ΠΡΟΣΩΠΩΝ
print("[INFO] loading encodings + face detector...")
data = pickle.loads(open(encodingsP, "rb").read())
detector = cv2.CascadeClassifier(cascade)

# ΑΡΧΙΚΟΠΟΙΗΣΗ ΤΟΥ ΒΙΝΤΕΟ, ΠΡΟΕΤΟΙΜΑΣΙΑ ΚΑΜΕΡΑΣ
print("[INFO] starting video stream...")
#vs = VideoStream(src=0).start()
vs = VideoStream(usePiCamera=True).start()
time.sleep(2.0)
```

```
# ΕΝΑΡΞΗ FPS ΜΕΤΡΗΤΗ
fps = FPS().start()

while True:
# ΑΠΟΘΗΚΕΥΣΗ ΣΤΙΓΜΙΟΤΥΠΟΥ ΠΡΟΣΩΠΟΥ ΑΠΟ ΤΟ ΒΙΝΤΕΟ
# WIDTH= 500px ΓΙΑ ΑΥΞΗΣΗ ΤΑΧΥΤΗΤΑΣ
frame = vs.read()
frame = imutils.resize(frame, width=500)

# ΑΛΛΑΓΗ ΧΡΩΜΑΤΩΝ ΣΕ ΚΛΙΜΑΚΑ ΤΟΥ ΓΚΡΙ ΚΑΙ ΤΟ ΑΝΤΙΣΤΡΟΦΟ
gray = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)
rgb = cv2.cvtColor(frame, cv2.COLOR_BGR2RGB)

# ΕΝΤΟΠΙΣΜΟΣ ΠΡΟΣΩΠΟΥ ΣΕ ΓΚΡΙ ΚΛΙΜΑΚΑ
rects = detector.detectMultiScale(gray, scaleFactor=1.1, minNeighbors=5, minSize=(30, 30),
flags=cv2.CASCADE_SCALE_IMAGE)

#ΤΟ OpenCV ΕΠΙΣΤΡΕΦΕΙ ΟΡΙΟΘΕΤΗΣΗ ΣΥΝΤΕΤΑΓΜΕΝΩΝ ΣΕ (x, y, w, h)
# ΟΜΩΣ ΕΜΕΙΣ ΧΡΕΙΑΖΟΜΑΣΤΕ ΣΕ top, right, bottom, left

boxes = [(y, x + w, y + h, x) for (x, y, w, h) in rects]

# ΕΝΣΩΜΑΤΩΣΗ ΠΛΑΙΣΙΟΥ ΟΡΙΟΘΕΤΗΣΗΣ ΠΡΟΣΩΠΟΥ
encodings = face_recognition.face_encodings(rgb, boxes)
names = []

for encoding in encodings:
matches = face_recognition.compare_faces(data["encodings"], encoding)
name = "Unknown"
```

```
# ΕΛΕΓΧΟΣ ΕΥΡΕΣΗΣ ΤΑΙΡΙΑΣΜΑΤΟΣ

if True in matches:
    matchedIdxs = [i for (i, b) in enumerate(matches) if b]
    counts = {}

    for i in matchedIdxs:
        name = data["names"][i]
        counts[name] = counts.get(name, 0) + 1

# ΚΑΘΟΡΙΣΜΟΣ ΑΝΑΓΝΩΡΙΣΜΕΝΟΥ ΠΡΟΣΩΠΟΥ ΜΕ ΤΗΝ ΜΕΓΛΥΤΕΡΗ
# ΒΑΘΜΟΛΟΓΙΑ (ΣΤΗΝ ΙΣΟΨΗΦΙΑ ΕΠΙΛΟΓΗ ΠΡΩΤΗΣ ΕΓΓΡΑΦΗΣ )
name = max(counts, key=counts.get)

# ΣΤΗΝ ΠΕΡΙΠΤΩΣΗ ΑΝΑΓΝΩΡΙΣΗΣ ΕΚΤΥΠΩΣΗ ΟΝΟΜΑΤΟΣ
if currentname != name:
    currentname = name
    print(currentname)

#ΦΩΤΟΓΡΑΦΙΣΗ ΚΑΙ ΑΠΟΣΤΟΛΗ ΗΛΕΚΤΡΟΝΙΚΟΥ ΜΗΝΥΜΑΤΟΣ

img_name = "image.jpg"
cv2.imwrite(img_name, frame)
print("Taking a picture.")

#ΑΠΟΣΤΟΛΗ ΗΛ. ΜΗΝΥΜΑΤΟΣ ΕΝΗΜΕΡΩΣΗΣ ΓΙΑ ΤΗΝ ΑΦΙΞΗ ΑΤΟΜΟΥ ΣΤΗΝ
#ΠΟΡΤΑ ΜΟΥ
request = send_message(name)
print ('Status Code: '+format(request.status_code))

#200 status code ΕΠΙΤΥΧΗΣ ΑΠΟΣΤΟΛΗ ΜΗΝΥΜΑΤΟΣ
# update ΤΗ ΛΙΣΤΑ ΜΕ ΤΑ ΟΝΟΜΑΤΑ
names.append(name)
```

```
# ΒΡΟΓΧΟΣ ΣΕ ΑΝΑΓΝΩΡΙΣΙΜΑ ΠΡΟΣΩΠΙΑ
for ((top, right, bottom, left), name) in zip(boxes, names):

# ΕΜΦΑΝΙΣΗ ΑΝΑΓΝΩΡΙΣΜΕΝΟΥ ΠΡΟΣΩΠΟΥ ΚΑΙ ΟΝΟΜΑΤΟΣ ΣΕ ΧΡΩΜΑΤΑ
BGR
cv2.rectangle(frame, (left, top), (right, bottom), (0, 255, 225), 2)
y = top - 15 if top - 15 > 15 else top + 15 cv2.putText(frame, name, (left, y),
cv2.FONT_HERSHEY_SIMPLEX,.8, (0, 255, 255), 2)

# ΕΜΦΑΝΙΣΗ ΠΡΟΣΩΠΟΥ ΣΤΗΝ ΟΘΟΝΗ
cv2.imshow("Facial Recognition is Running", frame)
key = cv2.waitKey(1) & 0xFF

# ΜΕ ΤΟ ΠΑΤΗΜΑ ΠΛΗΚΤΡΟΥ q ΣΤΑΜΑΤΑ Ο ΒΡΟΓΧΟΣ
if key == ord("q"):
break

# update FPS ΜΕΤΡΗΤΗ
fps.update()

# ΤΕΡΜΑΤΙΣΜΟΣ ΧΡΟΝΟΜΕΤΡΟΥ ΚΑΙ ΕΜΦΑΝΙΣΗ FPS ΠΛΗΡΟΦΟΡΙΩΝ
fps.stop()
print("[INFO] ΠΕΡΑΣΜΕΝΟΣ ΧΡΟΝΟΣ: {:.2f}".format(fps.elapsed()))
print("[INFO] ΠΕΡΙΠΟΙΟΥ FPS: {:.2f}".format(fps.fps()))
cv2.destroyAllWindows()
vs.stop()
```

11.3 headshots.py

```
import cv2

#ΒΑΖΟΥΜΕ ΤΟ ΟΝΟΜΑ ΠΡΟΣ ΑΝΑΓΝΩΡΙΣΗ
name = 'ElladaKarypidou'
```

```
cam = cv2.VideoCapture(0)

cv2.namedWindow("ΠΑΤΑ SPACE ΓΙΑ ΛΗΨΗ ΦΩΤΟΓΡΑΦΙΑΣ",
cv2.WINDOW_NORMAL)

cv2.resizeWindow("ΠΑΤΑ SPACE ΓΙΑ ΛΗΨΗ ΦΩΤΟΓΡΑΦΙΑΣ", 500, 300)

img_counter = 0

while True:

ret, frame = cam.read()

if not ret:

print("ΑΠΟΤΥΧΙΑ ΛΗΨΗΣ ΣΤΙΓΜΙΟΤΥΠΟΥ")

break

cv2.imshow(" ΠΑΤΑ SPACE ΓΙΑ ΛΗΨΗ ΦΩΤΟΓΡΑΦΙΑΣ ", frame)

k = cv2.waitKey(1)

if k%256 == 27:

# ESC ΠΑΤΗΜΑ

print("ΚΛΕΙΣΙΜΟ")

break

elif k%256 == 32:

# SPACE ΠΑΤΗΜΑ

img_name = "dataset/"+ name + "/image_{}.jpg".format(img_counter)

cv2.imwrite(img_name, frame)

print("{} written!".format(img_name))

img_counter += 1

cam.release()

cv2.destroyAllWindows()
```

11.4 headshots_picam.py

```
#ΕΙΣΑΓΩΓΗ ΑΠΑΡΑΙΤΗΤΩΝ ΒΙΒΛΙΟΘΗΚΩΝ
```

```
import cv2
```

```
from picamera import PiCamera
```

```
from picamera.array import PiRGBArray
```

```
#ΟΡΙΣΜΟΣ ΟΝΟΜΑΤΟΣ ΠΡΟΣ ΕΠΕΞΕΡΓΑΣΙΑ
```

```
name = 'ElladaKarypidou'
```

```
cam = PiCamera()
```

```
cam.resolution = (512, 304)
```

```
cam.framerate = 10
```

```
rawCapture = PiRGBArray(cam, size=(512, 304))
```

```
img_counter = 0
```

```
while True:
```

```
for frame in cam.capture_continuous(rawCapture, format="bgr", use_video_port=True):
```

```
image = frame.array
```

```
cv2.imshow("ΠΑΤΑ Space ΓΙΑ ΛΗΨΗ ΦΩΤΟΓΡΑΦΙΑΣ", image)
```

```
rawCapture.truncate(0)
```

```
k = cv2.waitKey(1)
```

```
rawCapture.truncate(0)
```

```
if k%256 == 27: # ESC pressed
```

```
break
```

```
elif k%256 == 32:
```

```
# ΠΑΤΗΜΑ SPACE
img_name = "dataset/"+ name +"/image_{}.jpg".format(img_counter)
cv2.imwrite(img_name, image)
print("{} written!".format(img_name))
img_counter += 1

if k%256 == 27:
print("ΚΛΕΙΣΙΜΟ")
break
cv2.destroyAllWindows()
```

11.5 train_model.py

```
#!/usr/bin/python
# import the necessary packages
from imutils import paths
import face_recognition
import pickle
import cv2
import os

print("ΕΝΑΡΞΗ ΕΠΕΞΕΡΓΑΣΙΑΣ ΠΡΟΣΩΠΩΝ.")
imagePaths = list(paths.list_images("dataset"))

# ΑΡΧΙΚΟΠΟΙΗΣΗ ΛΙΣΤΑΣ ΟΝΟΜΑΤΩΝ
knownEncodings = []
knownNames = []

# ΒΡΟΓΧΟΣ ΣΤΙΣ ΔΙΑΔΡΟΜΕΣ ΤΩΝ ΦΩΤΟΓΡΑΦΙΩΝ
for (i, imagePath) in enumerate(imagePaths):
```

```
# ΕΜΦΑΝΙΣΗΣ ΟΝΟΜΑΤΟΣ ΠΡΟΣΩΠΟΥ
print("[INFO] ΕΠΕΞΕΡΓΑΣΙΑ ΦΩΤΟΓΡΑΦΙΑΣ {}/{}".format(i + 1, len(imagePaths)))
name = imagePath.split(os.path.sep)[-2]

# ΦΟΡΤΩΣΗ ΦΩΤΟΓΡΑΦΙΑΣ ΚΑΙ ΜΕΤΑΤΡΟΠΗ ΣΕ RGB
image = cv2.imread(imagePath)
rgb = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)

# ΑΝΤΙΣΤΟΙΧΙΣΗ ΚΑΘΕ ΕΙΣΑΓΩΜΕΝΟΥ ΠΡΟΣΩΠΟΥ
boxes = face_recognition.face_locations(rgb, model="hog")

# ΥΠΟΛΟΓΙΣΜΟΣ ΠΡΟΣΘΗΚΗΣ ΠΡΟΣΩΠΟΥ
encodings = face_recognition.face_encodings(rgb, boxes)

for encoding in encodings:
    knownEncodings.append(encoding)
    knownNames.append(name)

print(" ΣΕΙΡΙΟΠΟΙΗΣΗ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ")
data = {"encodings": knownEncodings, "names": knownNames}

f = open("encodings.pickle", "wb")
f.write(pickle.dumps(data))
f.close()
```

Κεφάλαιο 12ο: Συμπεράσματα

Ο στόχος της διπλωματικής εργασίας ήταν η δημιουργία ενός έργου για την αναγνώριση προσώπου με κάμερα ασφαλείας και τη χρήση του raspberry pi. Το raspberry pi 4 διαμορφώθηκε και είναι πλήρως λειτουργικό. Έγιναν όλες οι παραμετροποιήσεις και εγκαταστάσεις βιβλιοθηκών. Ο αλγόριθμος εκτελείται και αναγνωρίζει το πρόσωπο στο οποίο έχει εκπαιδευτεί. Είναι δυνατή η απομακρυσμένη διαχείριση του raspberry pi από android συσκευή και τέλος, στην περίπτωση που αναγνωρίσει το πρόσωπο πραγματοποιείται η αποστολή ηλεκτρονικού μηνύματος ως ειδοποίηση.

Μελλοντικά, θα μπορούσε να γίνει επέκταση του έργου και να μπορέσει να γίνεται η αναγνώριση για περισσότερα πρόσωπα, βέβαια θα πρέπει αντίστοιχα να ολοκληρωθούν για κάθε νέο πρόσωπο όλα τα βήματα εκπαίδευσης του αλγορίθμου και αποθήκευσης δεδομένων.

Ακόμα, επειδή η χρήση της μάσκας είναι πλέον στην καθημερινότητα των ανθρώπων θα ήταν χρήσιμο να διαμορφωθεί ο αλγόριθμος και να χρησιμοποιηθούν τυχόν νέες βιβλιοθήκες ώστε η αναγνώριση προσώπου να γίνει και γι' αυτές τις περιπτώσεις. Μιας και αναφερόμαστε σε κάμερα ασφαλείας, η αναγνώριση αντικειμένων και κίνησης θα ήταν πολύ χρήσιμο. Αν λοιπόν είχαμε την κάμερα στην εξώπορτα του σπιτιού μας και εντόπιζε άγνωστα πρόσωπα, ή αντικείμενα όπως κουτιά θα ήταν χρήσιμο να έχουμε ειδοποίηση ώστε να λάβουμε τα μέτρα μας. Επίσης, εάν προσθέσουμε μικρόφωνο και ηχεία στο raspberry pi θα μπορούσαμε να τη χρησιμοποιήσουμε ως συσκευή παρακολούθησης και απομακρυσμένης επικοινωνίας, οπότε στην περίπτωση που δεν βρισκόμαστε στο σπίτι και είχαμε επισκέπτη στην πόρτα μας, θα είχαμε ειδοποιηθεί για την άφιξη προσώπου και ταυτόχρονα θα μπορούσαμε να είχαμε OnLine επικοινωνία μαζί του. Τέλος, στην περίπτωση που συνδέσουμε αισθητήρες στο raspberry pi θα ήταν εφικτό να ελέγχουμε τη θερμοκρασιών και υγρασία του σπιτιού. Όπως καταλαβαίνετε οι λειτουργίες που θα μπορούσαμε να έχουμε είναι πάρα πολλές.

Βιβλιογραφία - Πηγές ιστοσελίδων

1. Andrew Ng. www.deeplearning.ai deep learning open courses. URL <https://www.deeplearning.ai/>
2. MIT university machine/deep learning open courses. URL <https://ocw.mit.edu/index.htm>.
3. Stanford university machine/deep learning open courses. URL <https://online.stanford.edu/courses>.
4. Deeplizard Machine/Deep learning open courses. URL <https://deeplizard.com/>.
5. Ορέστη Νικόλα του Γεωργίου, Πανεπιστήμιο Πατρών, 2019
6. Histograms of Oriented Gradients for Human Detection Navneet Dalal and Bill Triggs
<http://lear.inrialpes.fr>
7. Γ. Σημαντήρης, Μαυρακάκης, Η Γκρίνιας, C Garcia, Γ. Τζιρίτας, ανίχνευση προσώπου και χαρακτηριστικών προσώπου, 2001 <https://www.csd.uoc.gr/~tziritas/papers/TR-01-02.pdf>
8. <https://www.deeplearning.ai/>
9. <https://learnopencv.com/histogram-of-oriented-gradients/>
10. <https://learnopencv.com/?s=raspberry+pi>
11. <https://www.raspberrypi.org/>
12. <https://www.debian.org/>
13. <https://opencv.org/>
14. https://brainbarriers4you.eu/the_neuron_el.html
15. <https://www.tomshardware.com>
16. <http://lear.inrialpes.fr/software/>
17. <http://cs231n.stanford.edu/>
18. <https://github.com/carol>
19. <https://www.hindawi.com/journals/cin/2018/7068349/>
20. <https://www.python.org/>
21. <https://www.android.com/>
22. https://en.wikipedia.org/wiki/Android_version_history
23. <https://lotuslabs.medium.com/clarifying-ai-machine-learning-deep-learning-data-science-with-venn-diagrams-c94198faa063>
24. https://brainbarriers4you.eu/the_neuron_el.html
25. https://www.researchgate.net/figure/sualizing-the-68-facial-landmark-coordinates-on-an-AI-generated-face-image_fig3_363553863
26. <https://www.nettop.gr/>
27. Histograms of Oriented Gradients for Human Detection Navneet Dalal and Bill Triggs
<http://lear.inrialpes.fr/people/triggs/pubs/Dalal-cvpr05.pdf>