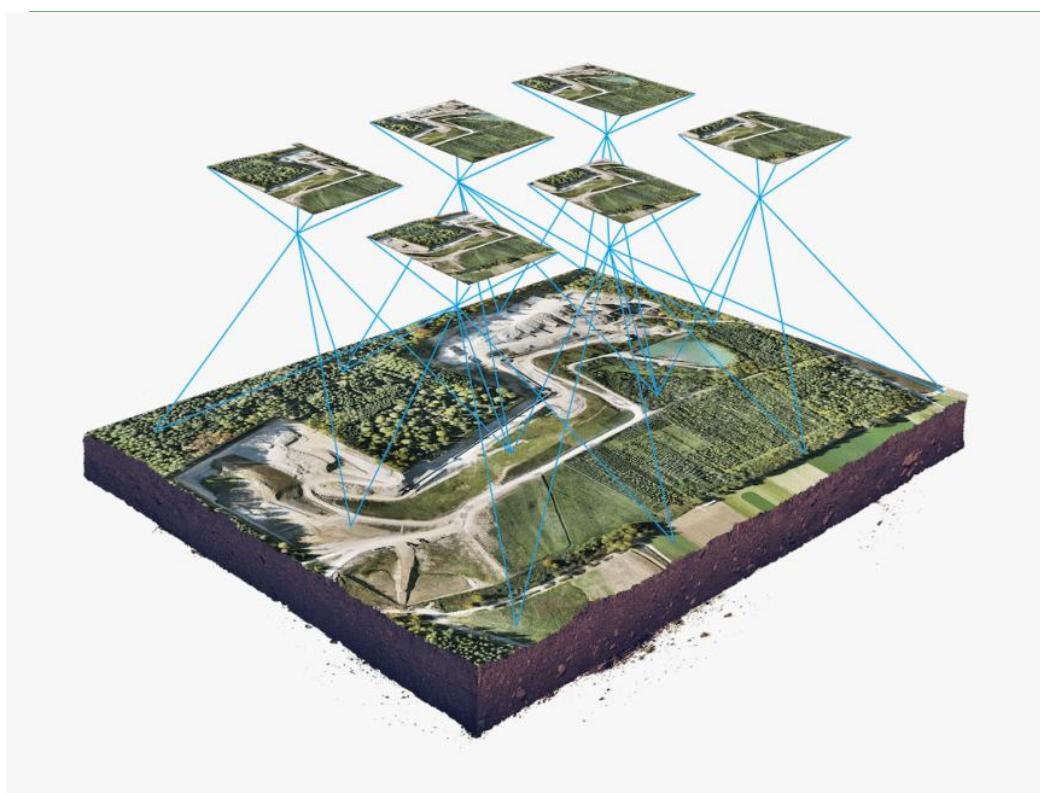




ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΑΡΥΜΑ ΑΘΗΝΑΣ  
ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΦΑΡΜΟΓΩΝ  
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΤΟΠΟΓΡΑΦΙΑΣ  
ΚΑΙ ΓΕΩΠΛΗΡΟΦΟΡΙΚΗΣ

**ΗΜΙΑΥΤΟΜΑΤΗ ΑΝΑΓΝΩΡΙΣΗ  
ΚΩΔΙΚΟΠΟΙΗΜΕΝΩΝ ΣΤΟΧΩΝ  
ΣΤΙΣ ΕΦΑΡΜΟΓΕΣ ΤΗΣ ΦΩΤΟΓΡΑΜΜΕΤΡΙΑΣ**



Πτυχιακή Εργασία  
Σμαραγδή Ελένη  
Α.Μ.: 10070

Επιβλέπων Καθηγητής:  
Λάζαρος Γραμματικόπουλος

Ηράκλειο  
Μάρτιος 2024

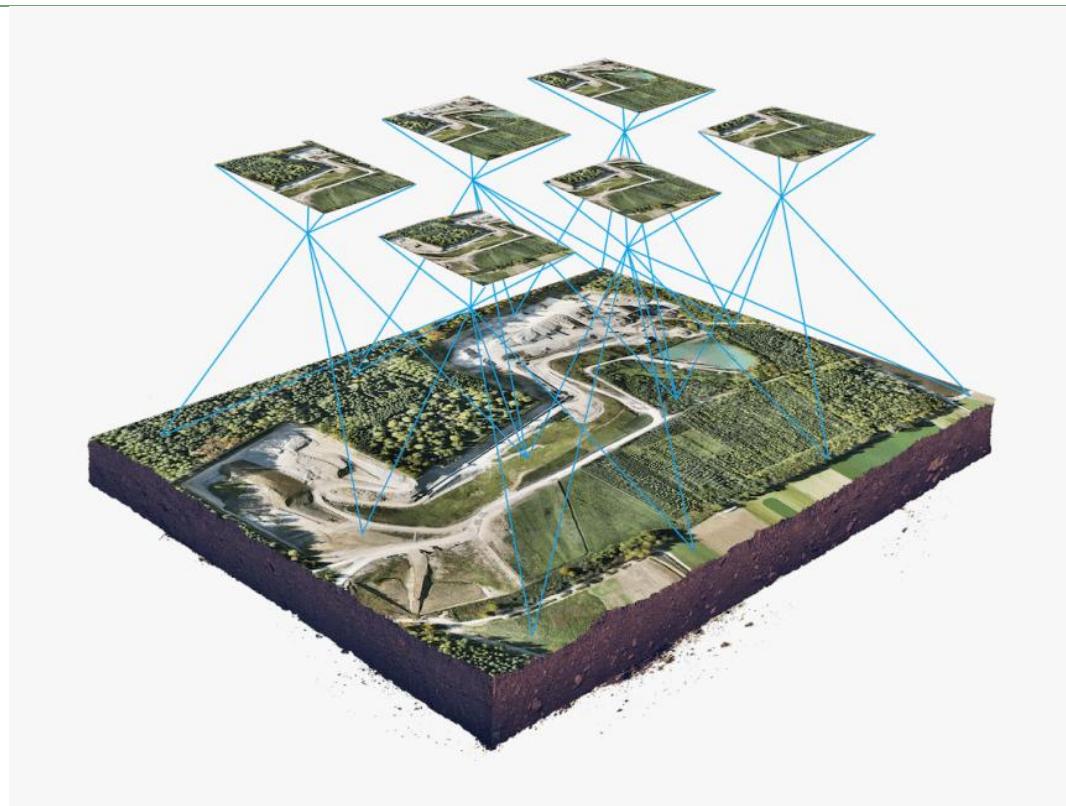


ATHENS UNIVERSITY OF APPLIED SCIENCES  
FACULTY OF TECHNOLOGICAL APPLICATIONS  
DEPARTMENT OF SURVEYING AND GEOINFOMATICS  
ENGINEERING

---

***SEMI-AUTOMATIC RECOGNITION OF  
CODEDTARGETS IN PHOTOGRAVIMETRY  
APPLICATIONS***

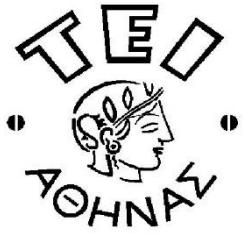
---



**Diploma Thesis  
Smaragdi Eleni  
R.N.: 10070**

**Supervisor:  
Lazaros Grammatikopoulos**

**Heraklion  
March 2024**



ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΑΡΥΜΑ ΑΘΗΝΑΣ  
ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΦΑΡΜΟΓΩΝ  
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΤΟΠΟΓΡΑΦΙΑΣ  
ΚΑΙ ΓΕΩΠΛΗΡΟΦΟΡΙΚΗΣ

ΗΜΙΑΥΤΟΜΑΤΗ ΑΝΑΓΝΩΡΙΣΗ ΚΩΔΙΚΟΠΟΙΗΜΕΝΩΝ ΣΤΟΧΩΝ  
ΣΤΙΣ ΕΦΑΡΜΟΓΕΣ ΤΗΣ ΦΩΤΟΓΡΑΜΜΕΤΡΙΑΣ

Η πτυχιακή/διπλωματική εργασία εξετάστηκε επιτυχώς από την κάτωθι Εξεταστική Επιτροπή:

A/α	ΟΝΟΜΑΕΠΩΝΥΜΟ	ΒΑΘΜΙΔΑ/ΙΔΙΟΤΗΤΑ	ΨΗΦΙΑΚΗΠΟΓΡΑΦΗ
	Λάζαρος Γραμματικόπουλος	Αναπληρωτής Καθηγητής	
	Ελένη Πέτσα	Καθηγήτρια	
	Γεώριος Σφήκας	Επίκουρος Καθηγητής	

## **ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ**

Η κάτωθι υπογεγραμμένη Σμαραγδή Ελένη του Νικολάου, με αριθμό μητρώου 10070, φοιτήτρια του Τεχνολογικού Εκπαιδευτικού Ιδρύματος Αθήνας, της Σχολής Τεχνολογικών Εφαρμογών, του Τμήματος Μηχανικών Τοπογραφίας και Γεωπληροφορικής, δηλώνω υπεύθυνα ότι:

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

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

Η Δηλούσα

*ΥΠΟΓΡΑΦΗ*

*Στις ψυχές που γρήγορα έφυγαν..*

*Μα και σ' αυτήν που θα ρθει..*

## **ΕΥΧΑΡΙΣΤΙΕΣ**

Ολοκληρώνοντας την πτυχιακή μου εργασία αισθάνομαι βαθιά ευγνωμοσύνη, καθώς είχα την τύχη και την ευτυχία να συνεργαστώ με τόσο αξιόλογους ανθρώπους, όπως τους καθηγητές μου, Κα Έλλη Πέτσα και Κο Λάζαρο Γραμματικόπουλο, Επίκουρο καθηγητή του Πανεπιστημίου Δυτικής Αττικής, οι οποίοι φρόντισαν να εξοπλιστώ με γνώση και αγάπη για την Επιστήμη της Φωτογραφικής και κατά συνέπεια της Επιστήμης των Υπολογιστών. Σπουδαία θέση κατείχαν στην προσπάθειά μου και οι συνάδελφοι που προσέθεσαν το πολύτιμο λιθαράκι τους, χαρίζοντάς μου νέες πληροφορίες και δίνοντάς μου δύναμη να φτάσω έως το τέλος. Ένας από εκείνους είναι και ο Βασίλειος Τσιρώνης, Διπλωματούχος Αγρονόμος Τοπογράφος Μηχανικός Ε.Μ.Π..

Κυρίως όμως νιώθω τυχερή, διότι δίπλα σε αυτούς τους ανθρώπους αποκόμισα νέες εμπειρίες που με «μπόλιασαν» με προσόντα για την επαγγελματική μου σταδιοδρομία, αλλά και για την ίδια μου τη ζωή. Καθώς με δίδαξαν ότι, παρ' όλες τις δυσκολίες και τις διακυμάνσεις, δεν πρέπει ποτέ να σταματάς να προσπαθείς, να ονειρεύεσαι, να μαθαίνεις.

Χωρίς την πίστη και τη βοήθεια της οικογένειάς μου και του συζύγου μου, δε θα διαβάζατε αυτό το κείμενο, ούτε θα ήμουν ο άνθρωπος που είμαι σήμερα.

Τη μέχρι σήμερα επαγγελματική μου σταδιοδρομία στον κλάδο της Τοπογραφίας την οφείλω κατ' αποκλειστικότητα στον Κύριο Γραμματικόπουλο. Χωρίς εκείνον επίσης, δε θα μπορούσα να «αγγίξω» το άγνωστο, που μεγαλύτερο δώρο από αυτό, θαρρώ πως δεν υπάρχει..

Για όλους αυτούς τους λόγους- και άλλους πολλούς,

σας ευχαριστώ όλους ξεχωριστά.

Έλενα Ν. Μ. Σμαραγδή

## **ΠΕΡΙΛΗΨΗ**

Στην προκείμενη εργασία γίνεται προσπάθεια να προσεγγιστούν οι σύγχρονες Εφαρμογές της Φωτογραμμετρίας και συγκεκριμένα, σε συνδυασμό με την Επιστήμη της Τεχνητής Νοημοσύνης. Αυτό διότι, οι αυτοματοποιημένες μέθοδοι που προσφέρονται πλέον μέσω του τομέα της Όρασης Υπολογιστών και της Μηχανικής Μάθησης, αλλάζει άρδην τα μέχρι πρότινος δεδομένα για τις εργασίες πάνω στις Φωτογραμμετρικές Εφαρμογές.

Στόχος λοιπόν, ήταν η γνωριμία με τις προαναφερθείσες Επιστήμες και τις εφαρμογές τους, καθώς και η δημιουργία ενός κώδικα, ικανού να ανιχνεύει κωδικοποιημένους στόχους σε εικόνες, με σκοπό τη χρήση τους σε Φωτογραμμετρικές Εφαρμογές. Με την ανίχνευση αυτών των στόχων, μπορούν να απλουστευθούν εργασίες, που υπό κανονικές συνθήκες αποτελούν πολύ χρονοβόρες και δαπανηρές διαδικασίες στις εφαρμογές της Φωτογραμμετρίας, ενώ επιπλέον μπορεί να απαιτούσαν την επίβλεψη κάποιου χρήστη και ίσως αντίστοιχη χειρωνακτική εργασία.

Το αποτέλεσμα της εργασίας είναι η επιτυχής λειτουργία του κώδικα ημιαυτόματης αναγνώρισης στόχων σε εικόνες, γεγονός που γεννά την προσδοκία για την ενσωμάτωση του κώδικα σε πολύπλοκες εφαρμογές, αλλά και για την περαιτέρω εξέλιξή του.

Ελπίζω λοιπόν, μέσω της εργασίας αυτής, να μεταφερθεί στους αναγνώστες η γνώση που προσωπικά είχα την τύχη να αποκομίσω και ενδεχομένως να συμβάλει, έστω και στο ελάχιστο δυνατόν, στην εξέλιξη της τεχνολογίας που προσφέρεται στο πεδίο της Φωτογραμμετρίας και κατά συνέπεια, της Επιστήμης της Τοπογραφίας.

## **ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ**

Φωτογραμμετρία, Τεχνητή Νοημοσύνη, Όραση Υπολογιστών, Μηχανική Μάθηση, Κωδικοποιημένοι Στόχοι, Ημιαυτόματη Αναγνώρισης Στόχων.

**SEMI-AUTOMATIC RECOGNITION OF  
CODED TARGETS IN PHOTOGRAMMETRY  
APPLICATIONS**

**Diploma Thesis**

Smaragdi Eleni

March 2024

**ABSTRACT**

In this paper, an attempt is made to approach the modern Applications of Photogrammetry, in conjunction with the Science of Artificial Intelligence. The automated methods now offered by the Computer Vision and Machine Learning fields, are changing dramatically the way things worked by now on Photogrammetry Applications.

Therefore, the aim was to become acquainted with the aforementioned Sciences and their applications and also to create a code capable of detecting coded targets on images, with the purpose to use them in Photogrammetric Applications. The detection of those targets can simplify some tasks that, under normal circumstances, are time-consuming and expensive procedures in Photogrammetry applications. Also, they may have required a user to supervise the procedures and perhaps do some manual works too.

The result of this paper is a successful operation of the semi-automatic target recognition code on images, a fact that gives rise to the expectation of the integration of the code in more complex applications and also, of its further development.

I hope that, through this experiment, the knowledge that I personally had the chance to obtain, might be transferred to the readers and that it may contribute, even to the minimum, to the evolution of the technology offered in the field of Photogrammetry and therefore, of the Scientific field of Topography.

**KEY WORDS**

Photogrammetry, Artificial Intelligence, Computer Vision, Machine Learning, Coded Targets, Semi- Automatic Recognition of Targets.

## **ΠΕΡΙΕΧΟΜΕΝΑ**

ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ .....	1
ΕΥΧΑΡΙΣΤΙΕΣ .....	3
ΠΕΡΙΛΗΨΗ .....	3
ΕΙΣΑΓΩΓΗ .....	1
1 ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ, ΟΡΑΣΗ ΥΠΟΛΟΓΙΣΤΩΝ ΚΑΙ ΦΩΤΟΓΡΑΜΜΕΤΡΙΑ .....	2
1.1 ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ (ARTIFICIAL INTELLIGENCE).....	3
1.1.1 ΟΡΙΣΜΟΙ .....	4
1.1.2 ΙΣΤΟΡΙΚΗ ΑΝΑΔΡΟΜΗ .....	5
ΤΟ «ΠΑΙΧΝΙΔΙ ΜΙΜΗΣΗΣ»ΤΟΥ ALAN TURING (TURING TEST):.....	5
1.1.3 ΜΕΛΛΟΝΤΙΚΕΣ ΠΡΟΒΛΕΨΕΙΣ ΚΑΙ ΚΙΝΔΥΝΟΙ.....	8
1.2 ΟΡΑΣΗ ΥΠΟΛΟΓΙΣΤΩΝ (COMPUTER VISION).....	9
1.2.1 ΟΡΙΣΜΟΙ .....	10
1.2.2 ΕΦΑΡΜΟΓΕΣ ΤΗΣ ΟΡΑΣΗΣ ΥΠΟΛΟΓΙΣΤΩΝ .....	10
1.3 ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ (MACHINE LEARNING).....	12
1.3.1 ΟΡΙΣΜΟΙ .....	12
2 ΑΝΙΧΝΕΥΣΗ ΑΝΤΙΚΕΙΜΕΝΩΝ ΣΕ ΕΙΚΟΝΕΣ.....	14
2.1 ΑΛΓΟΡΙΘΜΟΣ VIOLA- M. JONES .....	14
2.1.1 ΦΙΛΤΡΑ ΤΥΠΟΥ HAAR .....	14
2.1.2 ΠΙΝΑΚΑΣ ΠΡΟΣΤΙΘΕΜΕΝΟΥ ΕΜΒΑΔΟΥ (INTEGRAL IMAGE) .....	15
2.1.3 ΑΛΓΟΡΙΘΜΟΣ ADABoost .....	19
2.1.4 ΑΛΛΗΛΟΥΧΙΑ ΤΑΞΙΝΟΜΗΤΩΝ .....	21
3 ΚΩΔΙΚΟΠΟΙΗΜΕΝΟΙ ΣΤΟΧΟΙ .....	24
Για τη χρήση των κωδικοποιημένων στόχων, υπάρχουν λογισμικά όπως το PhotoModeler, AutodeskReCap, 3DF Zephyr, καθώς και το AgisoftPhotoScan, με το οποίο και θα ασχοληθούμε παρακάτω.....	26
3.1 ΤΡΙΣΔΙΑΣΤΑΤΗ ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΑΝΤΙΚΕΙΜΕΝΟΥ ΜΕ ΤΗ ΧΡΗΣΗ ΚΩΔΙΚΟΠΟΙΗΜΕΝΩΝ ΣΤΟΧΩΝ .....	26
3.1.1 ΔΗΜΙΟΥΡΓΙΑ ΚΑΙ ΤΟΠΟΘΕΤΗΣΗ ΚΩΔΙΚΟΠΟΙΗΜΕΝΩΝ ΣΤΟΧΩΝ .....	27
3.1.2 ΛΗΨΕΙΣ ΕΙΚΟΝΩΝ .....	28
3.1.3 ΕΠΕΞΕΡΓΑΣΙΑ ΕΙΚΟΝΩΝ .....	29
3.1.4 ΑΠΟΤΕΛΕΣΜΑΤΑ.....	31
4 ΠΕΙΡΑΜΑΤΙΚΗ ΕΦΑΡΜΟΓΗ: ΗΜΙΑΥΤΟΜΑΤΗ ΑΝΑΓΝΩΡΙΣΗ ΚΩΔΙΚΟΠΟΙΗΜΕΝΩΝ ΣΤΟΧΩΝ .....	33
4.1 ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΣ ΕΡΓΑΣΙΩΝ .....	34

4.1.1	ΕΠΙΛΟΓΗ ΛΟΓΙΣΜΙΚΩΝ .....	34
4.1.2	ΛΗΨΕΙΣ ΦΩΤΟΓΡΑΦΙΩΝ .....	35
4.1.3	ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ ΕΙΚΟΝΩΝ.....	36
4.2	ΕΠΙΣΗΜΑΝΣΗ ΣΤΟΧΩΝ .....	38
4.3	ΔΗΜΙΟΥΡΓΙΑ ΚΩΔΙΚΑ .....	39
4.3.1	ΠΕΡΙΓΡΑΦΗ ΚΩΔΙΚΑ.....	40
4.3.2	ΑΠΟΤΕΛΕΣΜΑΤΑ ΚΩΔΙΚΑ .....	44
5	ΣΥΜΠΕΡΑΣΜΑΤΑ .....	46
	ΒΙΒΛΙΟΓΡΑΦΙΑ .....	49
	ΠΑΡΑΡΤΗΜΑ Α: ΚΩΔΙΚΑΣ .....	51
	ΠΑΡΑΡΤΗΜΑΒ: ΕΙΚΟΝΕΣ ΕΚΠΑΙΔΕΥΣΗΣ ΑΛΓΟΡΙΘΜΟΥ.....	55
	ΠΑΡΑΡΤΗΜΑ Γ: ΠΙΝΑΚΑΣ ΑΚΡΩΝΥΜΩΝ .....	70

## ΕΙΣΑΓΩΓΗ

Η ανάγκη για εμβάθυνση στη γνώση της Επιστήμης των Υπολογιστών και της Φωτογραμμετρίας αποτέλεσε βασικό κίνητρο για τη συγγραφή της προκείμενης εργασίας. Στόχος της είναι η ανάπτυξη κατάλληλου λογισμικού, ικανού να εντοπίζει με ημιαυτόματο τρόπο σε μία τυχαία εικόνα, κωδικοποιημένους φωτογραμμετρικούς στόχους και να αναγνωρίζει την ταυτότητά τους, με σκοπό τη μελλοντική τους χρήση σε Εφαρμογές της Φωτογραμμετρίας. Η μεθοδολογία που ακολουθείται ξεκινά με τη προσέγγιση των όρων Τεχνητή Νοημοσύνη, Όραση Υπολογιστών, Μηχανική Μάθηση και κωδικοποιημένοι στόχοι, υπό την ευρεία τους έννοια. Γίνεται νύξη δε, στον αλγόριθμο των Viola και Jones, περί ανίχνευσης αντικειμένων σε εικόνες και εν κατακλείδι παρουσιάζεται μία εφαρμογή ημιαυτόματης εύρεσης κωδικοποιημένων στόχων.

Συγκεκριμένα:

Στο **Κεφάλαιο 1**, γίνεται αναφορά στους όρους της Τεχνητής Νοημοσύνης, της Όρασης Υπολογιστών και της Μηχανικής Μάθησης και στη σύνδεσή τους με την Επιστήμη της Φωτογραμμετρίας και της Τοπογραφίας.

Στο **Κεφάλαιο 2**, που αφορά την ανίχνευση αντικειμένων, παρουσιάζεται ο αλγόριθμος Viola – Jones και τα βήματα λειτουργίας του.

Στο **Κεφάλαιο 3**, γίνεται λόγος για τους κωδικοποιημένους στόχους. Αναφέρεται συνοπτικά μία πειραματική εφαρμογή κωδικοποιημένων στόχων, που αφορούν στην κατασκευή τρισδιάστατων μοντέλων αντικειμένων.

Στο **Κεφάλαιο 4** παρατίθεται μία πειραματική εφαρμογή ημιαυτόματης ανίχνευσης κωδικοποιημένων στόχων στο προγραμματιστικό περιβάλλον Matlab και συγκεκριμένα: ο τρόπος συλλογής δεδομένων και η επεξεργασία τους, η ανάπτυξη του κώδικα (και επεξήγησή του) με σκοπό τη μηχανική μάθηση του αλγόριθμου, η επεξήγηση της σχετικής θεωρίας και τα προκύπτοντα από την εφαρμογή δεδομένα.

Σημειώνεται, ότι τα αποτελέσματα του πειράματος ήταν θετικά, με επιτυχία κατά 90%, όσον αφορά την αναγνώριση συγκεκριμένου κωδικοποιημένου στόχου σε μία τυχαία εικόνα.

# **1 ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ, ΟΡΑΣΗ ΥΠΟΛΟΓΙΣΤΩΝ ΚΑΙ ΦΩΤΟΓΡΑΜΜΕΤΡΙΑ**

Ταξιδεύοντας πίσω στο χρόνο, αντιλαμβάνεται κανείς ότι το ανθρώπινο είδος, από τα πρώτα του μόλις «βήματα» και καθ' όλη την πορεία του, εκτός από την προσπάθειά του να καλύψει τις βασικές ανάγκες του, είχε στρέψει το ενδιαφέρον του και σε άλλα ενδιαφέροντα, το οποίο εξελίχθηκε σε νέα ανάγκη. Την ανάγκη για μάθηση και εξερεύνηση του περιβάλλοντός του, που μεγεθυνόταν ανά τους αιώνες. Μέχρι και στις μέρες μας, που η τεχνολογική ακμή ξεπερνάει τις παλαιότερες προβλέψεις και προσδοκίες και απορρέουν ολοένα και περισσότερες απαντήσεις από τις Επιστήμες, ο άνθρωπος φαίνεται να θρέφει όλο και περισσότερο την περιέργειά του, εξακολουθώντας να προσπαθεί να πλησιάσει, το «άγνωστο». Η ανθρώπινη φύση λοιπόν, είναι συνυφασμένη, θα έλεγε κανείς, με την ανάγκη για απαντήσεις σε φιλοσοφικά ερωτήματα, αλλά και τη βαθιά επιθυμία για εξέλιξη και ευελιξία.

Στην προσπάθειά του λοιπόν να ερμηνεύσει τον κόσμο, ο άνθρωπος ανέπτυξε τις Επιστήμες. Οι Επιστήμες για τις οποίες θα γίνει λόγος στην παρούσα εργασία, έχουν μεγάλο εύρος εφαρμογών και συνδέονται ουσιαστικά με άλλους κλάδους και Επιστήμες, όπως αυτή της Τοπογραφίας.

Πρώτη και κύρια, η Επιστήμη της Φωτογραμμετρίας, είναι αυτή που «τροφοδοτεί» με πληροφορίες και λύσεις -μεταξύ άλλων- το πεδίο της Τοπογραφίας. Τα τελευταία χρόνια, η Φωτογραμμετρία σχετίζεται άμεσα με τους κλάδους της Τεχνητής Νοημοσύνης, της Όρασης των Υπολογιστών και της Μηχανικής Μάθησης, οι οποίες αποτελούν παρακλάδια της Επιστήμης των Υπολογιστών. Οι τομείς αυτοί αναπτύσσονται με τεράστια άλματα κατά τις τελευταίες δεκαετίες και μπορούν να παράξουν πλήθος πληροφοριών, που είναι χρήσιμες για την επίλυση Φωτογραμμετρικών προβλημάτων.

Ως εκ τούτου, όπως μπορεί να αντιληφθεί κανείς, η κλασσική Τοπογραφία που ασχολείται με τις μεθόδους και τις τεχνικές μετρήσεων, την επεξεργασία τους, την απεικόνιση τόπων, τις χαράξεις κ.α., αναγάγεται πλέον σε μία Επιστήμη, η οποία έχει τη δυνατότητα εξαγωγής δεδομένων και περαιώσης εργασιών, με αυτοματοποιημένο τρόπο (π.χ. αυτόματη αποτύπωση και τρισδιάστατη μοντελοποίηση πεδίων). Προκειμένου όμως, να αντιληφθούμε πλήρως το πεδίο εφαρμογών των ανωτέρω Επιστημών, θα πρέπει απαραιτήτως να γίνει μία νύξη, για καθεμιά από αυτές χωριστά.

## 1.1 ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ (ARTIFICIAL INTELLIGENCE)



Πηγή:[1]

*“By far, the greatest danger of artificial intelligence is that people conclude too early that they understand it.”*

Eliezer Yudkowsky

Ως Τεχνητή Νοημοσύνη (Artificial Intelligence ή AI) νοείται ο κλάδος της Επιστήμης των Υπολογιστών, ο οποίος έχει ως αντικείμενο, εκτός από την κάλυψη ανθρώπινων φιλοσοφικών ερωτημάτων περί της δυνατότητας του ανθρώπου για δημιουργία ευφυών μηχανών, την αυτοματοποιημένη εκτέλεση εργασιών από τις μηχανές αυτές, που υπό κανονικές συνθήκες θα επωμιζόταν ο άνθρωπος.

Ουσιαστικά, πρόκειται για τον τομέα που ως στόχο έχει τη σχεδίαση υπολογιστικών συστημάτων, ικανά να μιμούνται τον άνθρωπο, ως προς τις γνωστικές του ικανότητες, με σκοπό την επίλυση προβλημάτων. Υπάρχει δε, η προσδοκία του να καταφέρει μία τέτοια μηχανή να φτάσει στο επίπεδο σκέψης του ανθρώπινου νου, αλλά ακόμη και να υπερτερήσει αυτού. Καθ' αυτόν τον τρόπο ο άνθρωπος δύναται να αποδεσμευτεί από χρονοβόρες και δαπανηρές εργασίες και ενδεχομένως να παρέχεται αυτόματα η επίλυση προβλημάτων, στα οποία ο ίδιος αδύνατεί να αντεπεξέλθει.

Η TN αποτελεί κομβικό σημείο διάφορων Επιστημών, όπως η Ψυχολογία, τα Μαθηματικά, η Φιλοσοφία, η Νευρολογία, η Γλωσσολογία, η Οικονομία και η Ηλεκτρική Εφαρμοσμένη Μηχανική, οι οποίες και «γέννησαν» τα ερεθίσματα για τη δημιουργία της, γι' αυτό και τις τροφοδοτεί με χρήσιμα εργαλεία.

[1]:[https://www.engineering.columbia.edu/files/seas/styles/816x460/public/content/cs\\_image/2022/28/robot-hero.jpg?itok=TrT\\_MOb-](https://www.engineering.columbia.edu/files/seas/styles/816x460/public/content/cs_image/2022/28/robot-hero.jpg?itok=TrT_MOb-)

### 1.1.1 ΟΡΙΣΜΟΙ

Η Τεχνητή Νοημοσύνη ως όρος εμφανίστηκε για πρώτη φορά στο συνέδριο του Dartmouth College το 1956, μέλη του οποίου ήταν επιφανείς ερευνητές όπως ο John McCarthy -ο οποίος και οργάνωσε το συνέδριο- οι Marvin Minsky, Julian Bigelow, Allen Newell, Herbert Simon κ.ά.. (Wikipedia)

Μετέπειτα, πολλοί ερευνητές έκαναν προσπάθεια να ορίσουν την έννοια της TN, μερικοί από τους οποίους είναι:

- Barr & Feigenbaum: «Τεχνητή Νοημοσύνη είναι ο τομέας της Επιστήμης των Υπολογιστών, που ασχολείται με τη σχεδίαση ευφυών υπολογιστικών συστημάτων, δηλαδή συστημάτων που επιδεικνύουν χαρακτηριστικά που σχετίζονται με τη νοημοσύνη στην ανθρώπινη συμπεριφορά- την κατανόηση της γλώσσας, τη μάθηση, τη συλλογιστική ικανότητα, τη λύση προβλημάτων κ.ο.κ..»(Avron Barr, 1981)
- Patrick Henry Winston: «TN είναι η μελέτη των υπολογιστικών μεθόδων που καθιστούν δυνατά: την αντίληψη, το συλλογισμό και την δράση.»(Winston, 1976)
- Marvin Minsky: «TN είναι η Επιστήμη που κάνει τις μηχανές να κάνουν πράγματα που θα απαιτούσαν ευφυΐα αν γινόταν από έναν άνθρωπο.» (AI)
- Jubraj R., Graham T., Ryan E.:«Η TN ορίζεται ως ένα υπολογιστικό σύστημα που μπορεί να αισθανθεί, να κατανοήσει, να δράσει και να μάθει. Ένα σύστημα που μπορεί να αντιληφθεί τον κόσμο γύρω του, να αναλύσει και να κατανοήσει τις πληροφορίες που λαμβάνει, να αναλάβει δράσεις με βάση αυτή την κατανόηση και να βελτιώσει τις επιδόσεις του, μαθαίνοντας από αυτό που συνέβη.» (Jubraj R., 2018)
- Rich & Knight: «TN είναι μελέτη του πώς να κάνουμε τους Ηλεκτρονικούς Υπολογιστές να κάνουν πράγματα, για τα οποία προς το παρόν, οι άνθρωποι είναι καλύτεροι.» (Elaine Rich, 1991)

Παρόλο που έχουν διθεί πολλοί ορισμοί όπως τους ανωτέρω, θα έλεγε κανείς ότι ακόμη δεν υπάρχει ακριβής ορισμός για την TN, λόγω της πολύπλοκης φύσης της και της δυσκολίας προς την κατανόηση της ανθρώπινης νοημοσύνης, γενικότερα.

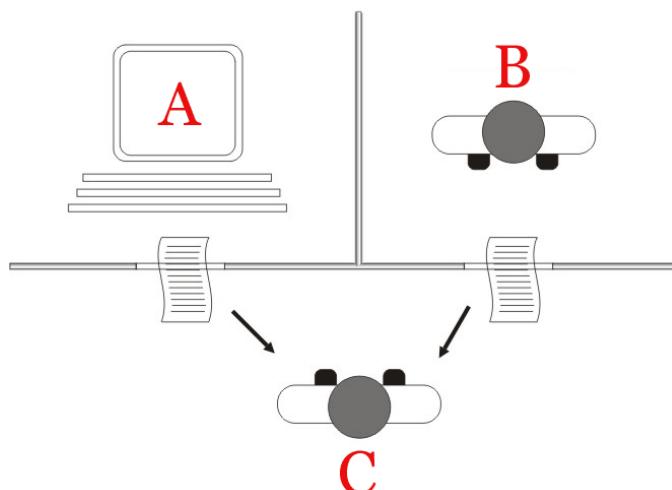
### 1.1.2 ΙΣΤΟΡΙΚΗ ΑΝΑΔΡΟΜΗ

#### ΤΟ «ΠΑΙΧΝΙΔΙ ΜΙΜΗΣΗΣ»ΤΟΥ ALAN TURING (TURING TEST):

Ο Alan Turing, φημισμένος μαθηματικός, θεωρητικός της λογικής, κρυπτογράφος και θεωρητικός βιολόγος από πολλούς θεωρείται ο «πατέρας» της Τεχνητής Νοημοσύνης. Είναι ευρέως γνωστός για τη συνεισφορά του στην Επιστήμη των Υπολογιστών και την Τεχνητή Νοημοσύνη, μέσω των έργων του «Μηχανή Turing», την επέκταση αυτής, λεγόμενη ως «Καθολική Μηχανή Turing» (στον Β' Παγκόσμιο Πόλεμο), αλλά και το test Turing- ή αλλιώς «Παιχνίδι Μίμησης» ή «Δοκιμή Turing».

Το τελευταίο του έργο αναπτύχθηκε κατά το 1950, στην προσπάθειά του να απαντήσει στο ερώτημα αν οι μηχανές μπορούν να σκεφτούν. Καθώς δεν έχει διευκρινιστεί πλήρως ο τρόπος λειτουργίας του ανθρώπινου νου, ήταν ακατόρθωτο να απαντηθεί θεωρητικά αν μία υπολογιστική μηχανή μπορεί να έχει νοημοσύνη, παρόμοια με την ανθρώπινη. Σκέφτηκε όμως, ότι θα μπορούσε να παρατηρηθεί η συμπεριφορά μίας τέτοιας μηχανής και να συγκριθεί με εκείνη ενός ανθρώπου.

Στην παρατήρηση αυτή αναπτύχθηκε και το Παιχνίδι Μίμησης, στο οποίο παίρνουν μέρος ένας κριτής, ένα φυσικό πρόσωπο και μία υπολογιστική μηχανή. Τα τρία αυτά μέλη βρίσκονται σε διαφορετικούς χώρους, ενώ ο κριτής απευθύνει μια σειρά ερωτήσεων και δέχεται τις απαντήσεις με τέτοιο τρόπο, ώστε να είναι αδύνατο να αντιληφθεί ποιός του απαντά κάθε φορά. Από τον τρόπο με τον οποίο δίδεται η απάντηση, ο κριτής οφείλει να συμπεράνει, αν αυτός που απάντησε ήταν το πρόσωπο ή η μηχανή (Εικόνα 1). (ΓΕΩΡΓΟΥΛΗ)



Εικόνα 1: Αναπαράσταση του Τεστ Turing

Πηγή:[https://upload.wikimedia.org/wikipedia/commons/5/55/Turing\\_test\\_diagram.png](https://upload.wikimedia.org/wikipedia/commons/5/55/Turing_test_diagram.png)

Έπειτα από τη δημιουργία του τεστ Turing, έχουν εξελιχθεί πιο σύγχρονες προσεγγίσεις επί του θέματος. Μία από αυτές είναι το «Total Turing Test», όπου η μηχανή καλείται επιπλέον να γνωρίσει αντικείμενα και εικόνες, όπως θα έκανε ο άνθρωπος, καθώς και να χειριστεί τα αντικείμενα στο χώρο. Για το Total Turing Test λοιπόν, απαραίτητη είναι η χρήση της Όρασης Υπολογιστών (Μηχανική Όραση) και της Ρομποτικής, αντίστοιχα.

Για πρώτη φορά το 2014, ένα πρόγραμμα υπολογιστή, το Eugene Goostman, πέρασε το πλήρες τεστ Turing του 2014 που πραγματοποιήθηκε στη Royal Society του Λονδίνου, αφού κατάφερε να «ξεγελάσει» το 33% των κριτών. (ΓΕΩΡΓΟΥΛΗ)

Άλλες δοκιμασίες, βασισμένες στο τεστ Turing (TuringTest) είναι: α) το Αντίστροφο τεστ Turing, β) το Marcus Test, γ) το Lovelace Test 2.0 και δ) το Τεστ Ελάχιστου Ευφυούς Σήματος.



"We'll know whether to treat it with any special moral consideration when we see if it passes the Turing test."

Πηγή: <https://i0.wp.com/www.chrismadden.co.uk/images/cartoons/turing-test-robot-baby-cartoon-chris-madden.gif?resize=500%2C676>

Έπειτα από το έργο του Alan Turing, παρατηρήθηκε η ενασχόληση όλο και περισσότερων Επιστημόνων με την TN και τις εφαρμογές της. Με την πάροδο του χρόνου ακολούθησε μεγάλο ενδιαφέρον, το οποίο ακόμη και στις μέρες μας δεν έχει εκλείψει, καθώς οι δυνατότητες της TN αναπτύσσονται εκθετικά, σε αναλογία με την πρόοδο της Επιστήμης των Υπολογιστών και τις δυνατότητες των υπολογιστικών μέσων και προγραμμάτων. Ακολουθεί σχετικός Πίνακας με τις φάσεις- ορόσημα της TN (Πίνακας 1).

1943	Οι McCulloch και Pitts προτείνουν ένα μοντέλο τεχνητών νευρώνων που έχει τη δυνατότητα να μαθαίνει και να υπολογίζει κάθε υπολογίσιμη συνάρτηση.
1950	Ο Alan Turing, που θεωρείται ο πατέρας της TN, εμπνέεται το τεστ της μίμησης (τεστ Τούρινγκ) για την αναγνώριση ευφυών μηχανών.
1951	Οι Minsky και Edmonds υλοποιούν το πρώτο νευρωνικό δίκτυο, το SNARC (Stochastic Neural Analog Reinforcement Calculator), το οποίο έχει 40 νευρώνες και χρησιμοποιεί 3000 λυχνίες.
1956	Συνάντηση στο Dartmouth College ερευνητών από το χώρο των Μαθηματικών, της Ηλεκτρονικής και Ψυχολογίας (McCarthy, AllenNewell, Herbert Simon, Marvin Minsky) με κοινό στόχο τη μελέτη δυνατοτήτων χρήσης των υπολογιστών για την προσομοίωση της ανθρώπινης νοημοσύνης.
1958	Δημιουργία της γλώσσας Lisp από τον McCarthy.
1966	Μετά από έρευνα γύρω από την κατανόηση γλώσσας και την αντίληψη μηχανής, ο Weizenbaum δημιουργεί το ELIZA.
1977	Δημιουργία των πρώτων εμπείρων συστημάτων: DENDRAL (1971), MYCIN (1975), Prospector (1977).
1972	<b>α.</b> Οι Colmerauer και Roussel από το Πανεπιστήμιο της Μασσαλίας σε συνεργασία με τον R. Kowalski από το Πανεπιστήμιο του Εδιμβούργου καταλήγουν στη δημιουργία της γλώσσας λογικού προγραμματισμού PROLOG. <b>β.</b> Ο Winograd εμβαθύνει στην κατανόηση φυσικής γλώσσας.
1975 & 1977	Ο M. Minsky δημιουργεί κεφάλαια περί αναπαράστασης της γνώσης σε βιβλία.
1976	Οι Newell & Simon υποστηρίζουν την υπόθεση ότι ένα φυσικό συμβολικό σύστημα διαθέτει τα απαραίτητα χαρακτηριστικά για νοήμονες ενέργειες.
1970-	Ανάπτυξη εξελικτικών αλγόριθμων. Εκδίδονται βιβλία με μελέτες:
1973	του Rechenberg για τη βελτιστοποίηση των τεχνικών συστημάτων και τις αρχές της βιολογικής εξέλιξης.
1975	του Holland για την προσαρμοστικότητα στα φυσικά και τεχνητά συστήματα.
1992	του Koza, για το Γενετικό Προγραμματισμό (Genetic Programming).
1995	του Fogel για τον Εξελικτικό Υπολογισμό (Evolutionary Computation).
1980-90	Αναγέννηση των Τεχνητών Νευρωνικών Δικτύων.
1986	Οι Rumelhart and McClelland περιγράφουν τη δημιουργία προσομοιώσεων της αντίληψης στον υπολογιστή.
1987	1 <sup>ο</sup> Διεθνές Συνέδριο για τα Νευρωνικά Δίκτυα του IEEE (Institute of Electrical and Electronics Engineers).
1965 & 1968	Πρώτος ο Zadeh εισάγει τους όρους «Ασαφή Σύνολα» (Fuzzy Sets, 1965) και «Ασαφείς Αλγόριθμοι» (Fuzzy Algorithms, 1968).
1983	Ο Sugeno διατυπώνει την «Ασαφή Θεωρία».
1992	1ο Συνέδριο του IEEE για τα Ασαφή Σύνολα.
1990 -	Δημιουργία <b>α.</b> υπολογιστικών συστημάτων και μηχανών που βασίζονται σε αρχές της TN και τα οποία παρουσιάζουν τάσεις προσαρμογής στο περιβάλλον τους (π.χ. ρομπότ) και <b>β.</b> εφαρμογών που τείνουν να «μαθαίνουν» από την εμπειρία τους: Νοήμονες πράκτορες, Μηχανές Αναζήτησης στο διαδίκτυο, Περιρρέουσα Νοημοσύνη.

Πίνακας 1: Οι Ημερομηνίες - Ορόσημα της TN.

Πηγή: [http://repfiles.kallipos.gr/html\\_books/93/00e-introduction.html](http://repfiles.kallipos.gr/html_books/93/00e-introduction.html)

### **1.1.3 ΜΕΛΛΟΝΤΙΚΕΣ ΠΡΟΒΛΕΨΕΙΣ ΚΑΙ ΚΙΝΔΥΝΟΙ**

Ήδη στις μέρες μας η Τεχνητή Νοημοσύνη εφαρμόζεται κατά κόρων σε πολλούς τομείς. Μερικοί από αυτούς είναι η Οικονομία, τα Μαθηματικά, η Αστρονομία, η Στατιστική, η Ιατρική και τα Συστήματα Υγείας, ενώ επίσης μπορεί να γίνει χρήση της από οποιαδήποτε επιχείρηση ή οργανισμό (από αυτοκινητικές εταιρείες, μέχρι και σε γεωπονικές μελέτες). Η φύσης της την καθιστά εφαρμόσιμη σχεδόν σε κάθε πεδίο που μπορούμε να φανταστούμε, όπως λ.χ. το χώρο των Τεχνών. Προδιαγράφεται λοιπόν ένα νέο μέλλον, διαφορετικό από αυτό των περασμένων αιώνων, γεγονός που εγείρει τόσο ερωτήματα, όσο και ανησυχίες.

Το θέμα που τίθεται συνηθέστερα γύρω από την TN, είναι το ηθικό πλαίσιο της προσαρμογής της, τα «όρια» δηλαδή που τίθενται κατά την εναρμόνισή της με την ανθρώπινη φύση και τους κινδύνους γύρω από την ανεξαρτησία που επιδεικνύουν οι υπολογιστικές μηχανές που βασίζονται σε αυτήν.

Πολλοί κάνουν λόγο για τον κώδικα ηθικής που θα πρέπει να εφαρμόζεται στην εκάστοτε εφαρμογή, όπως π.χ. κατά τη δημιουργία ενός ρομπότ με λογική κρίση. Άλλοι είναι υποστηρικτές της ιδέας περί του κινδύνου «καταπάτησης» του κώδικα αυτού, βασισμένοι στο ότι ένα ρομπότ που σκέφτεται, σχεδόν όπως ο άνθρωπος, μπορεί να αναπτύξει τη δική του ηθική, είτε αυτή είναι σωστή, είτε λανθασμένη.

Τέλος, ανησυχία προκαλεί το ενδεχόμενο της επί σκοπού λανθασμένης χρήσης λογισμικών TN, αποσκοπώντας σε προσωπικά οφέλη, όπως για παράδειγμα η παραπληροφόρηση και κοινωνική χειραγώγηση, η καταπάτηση των προσωπικών δεδομένων, η εξάπλωση του hacking, κ.ά..

Κρίσιμο είναι λοιπόν, να δημιουργείται ένα σύστημα αυστηρού ελέγχου και να εφαρμόζεται απαραιτήτως ένα προσαρμοσμένο νομικό πλαίσιο, στην εκάστοτε εφαρμογή, βασισμένη στην TN.

Να προστεθεί όμως, ότι μελλοντικές προβλέψεις θέλουν την TN να ακμάζει κατά τις επόμενες δεκαετίες, κυρίως σε συνδυασμό με τους κβαντικούς υπολογιστές, οι οποίοι ανέτρεψαν άρδην τα μέχρι τώρα δεδομένα γύρω από τις δυνατότητες των υπολογιστικών μηχανών. Ο συνδυασμός αυτών των πολλά υποσχόμενων «εργαλείων», πιστεύεται ότι θα αλλάξει τον κόσμο που μέχρι τώρα γνωρίζουμε. Και φυσικά, όπως σε κάθε τέτοια ανατροπή, έτσι και σε αυτήν την περίπτωση, οι απόψεις τείνουν να διίστανται.

## 1.2 ΟΡΑΣΗ ΥΠΟΛΟΓΙΣΤΩΝ (COMPUTER VISION)



Πηγή:[2]

*“Computers are able to see, hear and learn.  
Welcome to the future.”*

Dave Waters

Η Όραση των Υπολογιστών, ή Μηχανική Όραση έγκειται στο πεδίο της Τεχνητής Νοημοσύνης, καθώς πρόκειται για την εκμάθηση των υπολογιστικών μηχανών, να αναγνωρίζει αντικείμενα, ανθρώπους, μοτίβα, κ.ο.κ.. Για να καταφέρουμε να εκπαιδεύσουμε μία υπολογιστική μηχανή να «βλέπει» και να αντιλαμβάνεται τις πληροφορίες που προκύπτουν, χρειάζεται πολύ μεγάλος όγκος δεδομένων κατά την εκπαίδευσή του. Στις μέρες μας υπάρχει η δυνατότητα εκμετάλλευσης πληθώρας πληροφοριών, που αντλείται από απλά προσωπικά μέσα (π.χ. κινητά τηλέφωνα και προσωπικοί υπολογιστές), μέχρι και από υπερμεγέθεις συλλογές δεδομένων («Big Data»), γεγονός που αποτέλεσε βασικό κίνητρο, ώστε να εξελιχθεί και ο συγκεκριμένος κλάδος.

Η Όραση Υπολογιστών βασίζεται σε διάφορες μεθόδους, όπως:

[2]:[https://assets-global.website-files.com/63f902d79a33f7ff016cde0b/63f902d89a33f74fa16ce756\\_6390a10ecc27e94dcd043dbc\\_computer-vision-street.webp](https://assets-global.website-files.com/63f902d79a33f7ff016cde0b/63f902d89a33f74fa16ce756_6390a10ecc27e94dcd043dbc_computer-vision-street.webp)

- Η επεξεργασία της εικόνας: αναζήτηση αιχμών και σχημάτων βάσει των τιμών των pixels (π.χ. διαφοροποιήσεων χρωματισμού), εφαρμογή φίλτρων κ.ά.
- Η Μηχανική και η Βαθιά Μάθηση (Machine Learning και Deep Learning)
- Τα Νευρωνικά Δίκτυα, καθώς και συνδυασμό των παραπάνω.

### **1.2.1 ΟΡΙΣΜΟΙ**

Σύμφωνα με τους Dana H. Ballard και Christopher M. Brown, «Η Όραση Υπολογιστών είναι η κατασκευή σαφών, ουσιαστικών περιγραφών φυσικών αντικειμένων από εικόνες.» (Dana H. Ballard, 1982)

«Όραση Υπολογιστών είναι η Επιστήμη που παρέχει στους υπολογιστές ή σε άλλες μηχανές την όραση ή την ικανότητα να βλέπουν.» (Learned-Miller, 2012)

Αναφέρει χαρακτηριστικά ο Richard Szeliski : «Στην Όραση Υπολογιστών προσπαθούμε να περιγράψουμε τον κόσμο που εμείς βλέπουμε, σε μία ή περισσότερες εικόνες και να ανακατασκευάσουμε τις ιδιότητές του, όπως το σχήμα, το φωτισμό, και τις χρωματικές διανομές.» (Szeliski, 2022)

Ουσιαστικά, με την Όραση Υπολογιστών παρέχεται η δυνατότητα στις υπολογιστικές μηχανές να αναγνωρίζουν αντικείμενα στο χώρο και να αντιλαμβάνονται περαιτέρω τις ιδιότητες του, να παράγεται δηλαδή από αυτή την αναγνώριση, κάποιο λογικό συμπέρασμα. Γίνεται προσπάθεια δηλαδή, για την προσέγγιση του ανθρώπινου τρόπου αντίληψης του περιβάλλοντος, μέσω μία εικόνας ή συλλογής εικόνων, ή ακόμη και βίντεο (σε πραγματικό χρόνο).

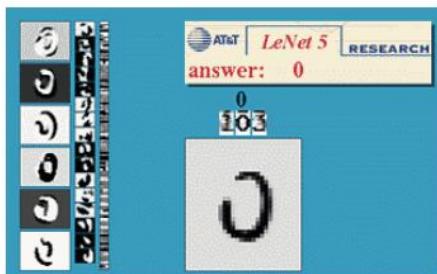
### **1.2.2 ΕΦΑΡΜΟΓΕΣ ΤΗΣ ΟΡΑΣΗΣ ΥΠΟΛΟΓΙΣΤΩΝ**

Ένα μεγάλο εύρος εφαρμογών καλύπτεται από το εν λόγω επιστημονικό πεδίο, το οποίο καθιστά, θα έλεγε κανείς, ευκολότερη την εργασία του ανθρώπου σε διάφορες περιπτώσεις. Μερικές από αυτές είναι οι ακόλουθες:

- Οπτική αναγνώριση χαρακτήρων (OCR): ανάγνωση χειρόγραφων ταχυδρομικών κωδίκων σε γράμματα και αυτόματη αναγνώριση πινακίδων κυκλοφορίας (ANPR).(Εικόνα 2.a)
- Επιθεώρηση μηχανημάτων: ταχεία επιθεώρηση εξαρτημάτων, για τη διασφάλιση ποιότητας, με χρήση στερεοσκοπικής όρασης με εξειδικευμένο φωτισμό, σε αεροσκάφη ή αυτοκίνητα.(Εικόνα 2.b)

- Εργασίες αποθήκης: αυτόνομες εργασίες και εργασίες από ρομποτικούς χειριστές. (Εικόνα 2.c)
- Ιατρική απεικόνιση: καταγραφή προ-εγχειρητικών εικόνων ή διεξαγωγή μακροχρόνιων μελετών της μορφολογίας του εγκεφάλου των ανθρώπων καθώς γερνούν.(Εικόνα 2.d)
- Αυτοδηγούμενα οχήματα: ικανά να οδηγούν μεταξύ πόλεων, καθώς και αυτόνομη πτήση.(Εικόνα 2.e)
- Κατασκευή τρισδιάστατων μοντέλων (φωτογραμμετρία): πλήρως αυτοματοποιημένη κατασκευή τρισδιάστατων μοντέλων από αεροφωτογραφίες και φωτογραφίες από drones.(Εικόνα 2.f)
- Αναγνώριση δακτυλικών αποτυπωμάτων και βιομετρικά στοιχεία: για τον αυτόματο έλεγχο ταυτότητας πρόσβασης και για εγκληματολογικές εφαρμογές.

(Szeliski, 2022)



(a)



(b)



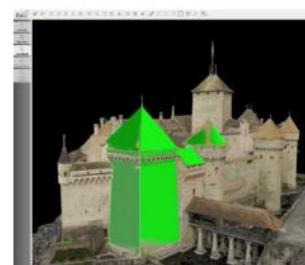
(c)



(d)



(e)

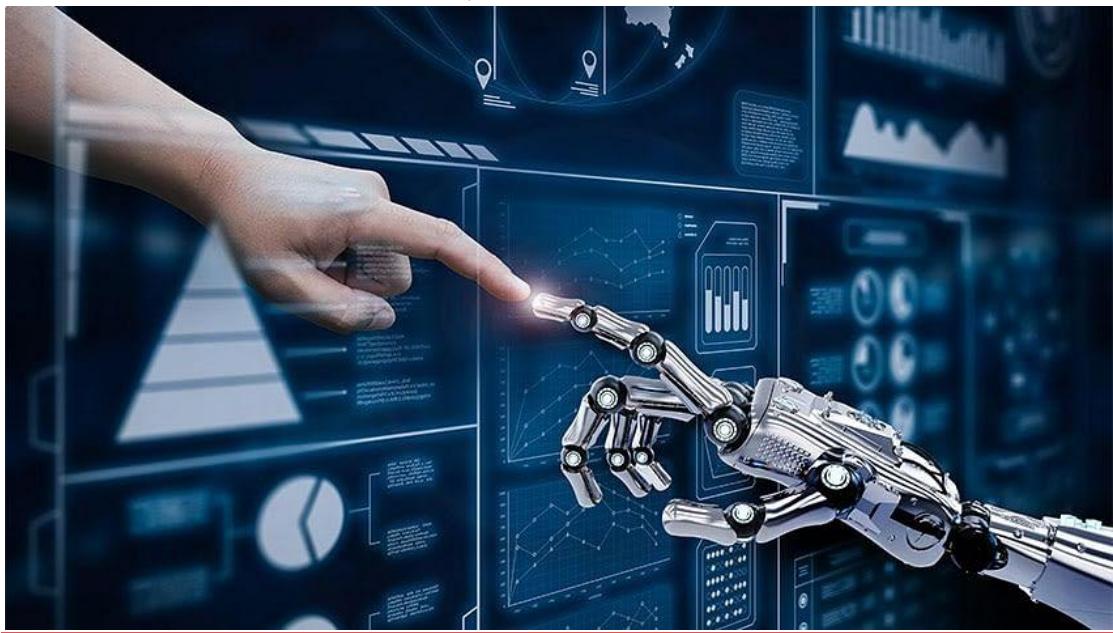


(f)

Εικόνα 2: Εφαρμογές Της Όρασης Υπολογιστών.

Πηγή: (Szeliski, 2022)

### 1.3 ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ (MACHINE LEARNING)



Πηγή:[3]

*“A Breakthrough In Machine Learning  
Would Be Worth Ten Microsofts.”*

*Bill Gates*

Χρήσιμο είναι να γίνει νύξη και στο πεδίο που αναφέρθηκε σε προηγούμενη ενότητα, αυτό της Μηχανικής Μάθησης, η οποία σχετίζεται άμεσα με την Όραση Υπολογιστών και χρησιμοποιείται στα επόμενα κεφάλαια της εργασίας, για την εφαρμογή της στο πεδίο της Φωτογραμμετρίας.

#### 1.3.1 ΟΡΙΣΜΟΙ

«Ένα πρόγραμμα υπολογιστή λέγεται ότι μαθαίνει από εμπειρία E ως προς μια κλάση εργασιών T και ένα μέτρο επίδοσης P, αν η επίδοσή του σε εργασίες της κλάσης T, όπως αποτιμάται από το μέτρο P, βελτιώνεται με την εμπειρία E» (Tom M. Mitchell)

«Μηχανική Μάθηση ονομάζεται η ικανότητα ενός υπολογιστικού συστήματος να δημιουργεί μοντέλα ή πρότυπα από ένα σύνολο δεδομένων.» (ΓΕΩΡΓΟΥΛΗ)

«Η μηχανική μάθηση μπορεί να οριστεί ως ένα σύστημα το οποίο βελτιώνει την απόδοση του κατά την εκτέλεση εργασιών, χωρίς την ανάγκη επαναπρογραμματισμού.» (Βίκτωρ, 2021)

[3]:<https://www.simplilearn.com/machine-learning-vs-deep-learning-major-differences-you-need-to-know-article>

"Το πεδίο μελέτης που δίνει στους υπολογιστές την ικανότητα να μαθαίνουν, χωρίς να έχουν ρητά προγραμματιστεί." (Arthur Samuel )

Η Μηχανική Μάθηση (ή Machine Learning) είναι ένας τομέας της Τεχνητής Νοημοσύνης, ο οποίος εξελίσσεται ολοένα και περισσότερο τις τελευταίες δεκαετίες και εφαρμόζεται σε πολλά και διαφορετικά πεδία.

Ουσιαστικά, η Μηχανική Μάθηση διερευνά τη μελέτη και την κατασκευή αλγόριθμων που μπορούν να μαθαίνουν από τα δεδομένα και να κάνουν προβλέψεις σχετικά με αυτά.

Παραδείγματα εφαρμογών αποτελούν τα φίλτρα spam (spam filtering), η οπτική αναγνώριση χαρακτήρων (OCR), οι μηχανές αναζήτησης και η Υπολογιστική Όραση (wikipedia), η οποία προαναφέρθηκε και θα μας απασχολήσει σε επόμενα κεφάλαια.

## **2 ΑΝΙΧΝΕΥΣΗ ΑΝΤΙΚΕΙΜΕΝΩΝ ΣΕ ΕΙΚΟΝΕΣ**

### **2.1 ΑΛΓΟΡΙΘΜΟΣ VIOLA- M. JONES**

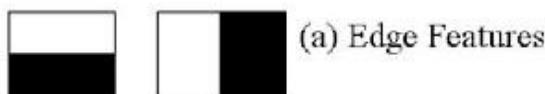
Τα τελευταία χρόνια, η ραγδαία εξέλιξη της τεχνολογίας επηρέασε και τόσο το χώρο της Φωτογραμμετρίας, όσο και της Όρασης των Υπολογιστών. Πολλοί ερευνητές ασχολήθηκαν με την ανάπτυξη λογισμικών που αφορούν στον αυτόματο εντοπισμό προσώπων ή αντικειμένων σε εικόνες. Στη συνέχεια, οι αλγόριθμοι αυτοί χρησιμοποιήθηκαν και στο πεδίο της Φωτογραμμετρίας με σκοπό την αυτοματοποίηση φωτογραμμετρικών διαδικασιών, έτσι ώστε να διασφαλιστεί η καλή ποιότητα των αποτελεσμάτων τους και ο σύντομος χρόνος διεκπεραίωσής τους, άρα κατά συνέπεια και το μικρό κόστος κατά τις φωτογραμμετρικές μεθόδους.

Ένας από τους αλγόριθμους που συνέβαλαν στον εντοπισμό αντικειμένων και άλλαξε άρδην τα μέχρι τότε δεδομένα ήταν αυτός των Viola και Jones (Paul Viola, 2001). Ο αλγόριθμος αυτός έχει ως αντικείμενο την ανίχνευση προσώπων και ήταν ο πρώτος που, εκτός από το να επιτυγχάνει αξιόπιστα αποτελέσματα, κατάφερε και να λειτουργεί σε πραγματικό χρόνο.

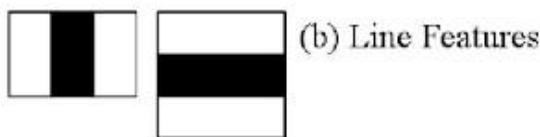
Η διαδικασία που ακολουθείται για τον εντοπισμό των προσώπων, αποτελείται από τα τέσσερα παρακάτω στάδια.

#### **2.1.1 ΦΙΛΤΡΑ ΤΥΠΟΥ HAAR**

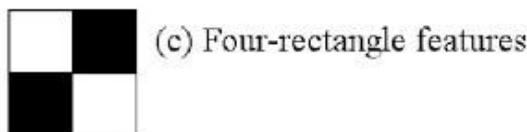
Για τον εντοπισμό των ζητούμενων αντικειμένων στις εικόνες (σε αυτήν την περίπτωση ανθρώπινα χαρακτηριστικά) γίνεται χρήση των φίλτρων τύπου Haar. Τα φίλτρα αυτά αποτελούνται από 2, 3 ή 4 ορθογώνια, όπως φαίνεται στην Εικόνα 3. Για κάθε περιοχή, υπολογίζεται το άθροισμα των τιμών των εικονοστοιχείων κάθε ορθογωνίου, ενώ η τιμή του εκάστοτε χαρακτηριστικού προκύπτει από την αφαίρεση του συνόλου του άσπρου ορθογωνίου από αυτή του μαύρου (ισοδύναμο με συνελικτικό φίλτρο όπου το λευκό είναι +1, το μαύρο -1).



(a) Edge Features



(b) Line Features



(c) Four-rectangle features

Εικόνα 3: Φίλτρα Τύπου Haar.  
Πηγή: <https://www.youtube.com/watch?v=Wwn81tVIR10>

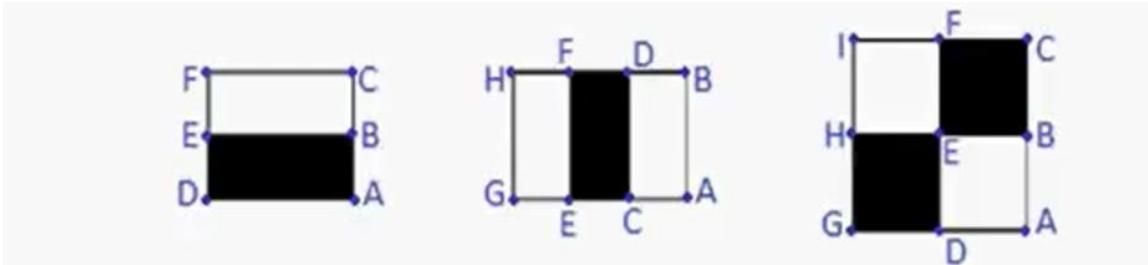
Εκτός από τα κλασσικά φίλτρα Haar, για τις οριζόντιες και κατακόρυφες ακμές που βρίσκονται στην πρώτη κατηγορία (a) της Εικόνας 3, οι Viola και Jones προσέθεσαν επίσης τα φίλτρα εύρεσης ράβδων και γραμμών (b), καθώς και τα στραμμένα κατά  $45^\circ$  που αφορούν σε διαγώνια σχήματα (c). Ως εκ τούτου, το σύνολο των χαρακτηριστικών που προκύπτει στην εικόνα είναι πολύ μεγαλύτερο από τον αριθμό των εικονοστοιχείων της. Ο ανιχνευτής που χρησιμοποιείται είναι διαστάσεων 24px x 24px, δηλαδή για μία περιοχή με 576 εικονοστοιχεία, τα χαρακτηριστικά φτάνουν πάνω από 180.000. (Paul Viola, 2001) Το μειονέκτημα της παραπάνω διαδικασίας είναι ότι απαιτείται αρκετός χρόνος για την εύρεση των ζητούμενων χαρακτηριστικών, το οποίο όμως αντιμετωπίζεται με τη χρήση του Πίνακα Προστιθέμενου Εμβαδού (γνωστή στη βιβλιογραφία και ως Εικόνα Ολοκλήρωμα).

(Paul Viola, 2001) (Βαρυτιμίδης, 2008) (ΓΕΩΡΓΙΟΥ, 2014)

## 2.1.2 ΠΙΝΑΚΑΣ ΠΡΟΣΤΙΘΕΜΕΝΟΥ ΕΜΒΑΔΟΥ (INTEGRAL IMAGE)

Όπως αναφέρθηκε και στην προηγούμενη υποενότητα, ο υπολογισμός των τιμών των χαρακτηριστικών Haar γίνεται μέσω της αφαίρεσης του συνολικού αριθμού, μεταξύ των ορθογωνίων. Καθώς όμως τα χαρακτηριστικά που προκύπτουν είναι πάνω από 180.000 μόνο για μια εικόνα (με μέγεθος φίλτρου ανιχνευτή 24x24), είναι αυτονόητο ότι η διαδικασία αυτή απαιτεί πάρα πολύ χρόνο, άρα δε θα έχει πρακτική εφαρμογή σε κανονικού μεγέθους εικόνες. Για να αποφευχθεί αυτό το πρόβλημα, χρησιμοποιείται ο Πίνακας Προστιθέμενου Εμβαδού (Integral Image) και ο

αντίστοιχος Στραμμένος Πίνακας Προστιθέμενου Εμβαδού. Με τους πίνακες αυτούς αποφεύγεται η χρονοβόρα διαδικασία της άθροισης των τιμών σε κάθε ορθογώνιο του φίλτρου, αντιθέτως, χρησιμοποιούνται μόνο τέσσερις τιμές του πίνακα για τον υπολογισμό του αθροίσματος. Συγκεκριμένα, στα φίλτρα που περιέχουν δύο τετράγωνα (ακμών), κανονικά θα χρειαζόταν οκτώ τιμές. Επειδή όμως τα ορθογωνία γειτνιάζουν χρειάζονται μόνο έξι τιμές, οκτώ στην περίπτωση των τριών ορθογωνίων και εννιά στην περίπτωση των τεσσάρων (Εικόνα 4).



Εικόνα 4: Υπολογισμός τιμών ορθογωνίων.

Πηγή: <https://www.youtube.com/watch?v=Wwn81tVIR10>

Αναλυτικότερα, ο Πίνακας Προστιθέμενου Εμβαδού (II) έχει τις ίδιες διαστάσεις με την αρχική εικόνα και σε κάθε θέση  $(x,y)$  περιλαμβάνεται το άθροισμα όλων των τιμών επάνω και αριστερά αυτής, καθώς και της ίδιας της  $(x,y)$ . Γενικά ισχύει:

$$II(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y')$$

όπου  $i(x', y')$  η τιμή του εικονοστοιχείου στη θέση  $(x', y')$  της αρχικής εικόνας.

Για να κατασκευαστεί λοιπόν ολόκληρος ο Π.Π.Ε. αρκεί μόνο ένα πέρασμα στην αρχική εικόνα, από αριστερά προς τα δεξιά και από πάνω προς τα κάτω, χρησιμοποιώντας την ακόλουθη σχέση:

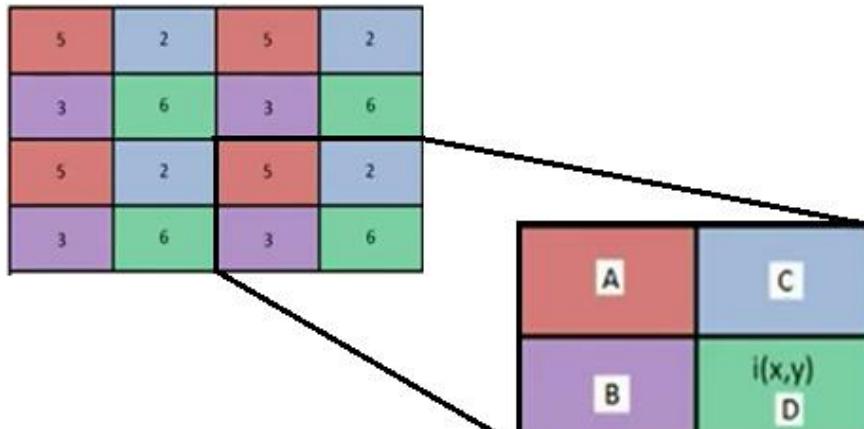
$$II(x, y) = II(x, y - 1) + II(x - 1, y) - II(x - 1, y - 1) + i(x, y),$$

Έχοντας τις αρχικές τιμές  $II(-1, y) = 0$  και  $II(x, -1) = 0$ .

Για παράδειγμα, έστω ότι θέλουμε να μετατρέψουμε και το τελευταίο εικονοστοιχείο (D) της Εικόνας 5, στον Π.Π.Ε..

Τότε ισχύει:

$$II(x, y) = D + sum(C) + sum(B) - sum(A) = 6 + 46 + 48 - 36 = 64.$$



(α) Αρχική εικόνα

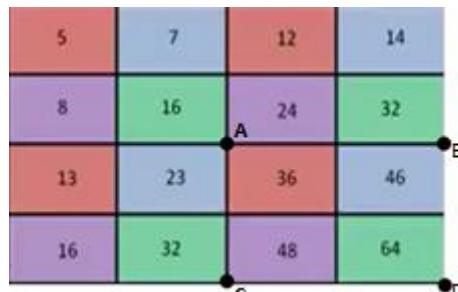
5	7	12	14
8	16	24	32
13	23	36	46
16	32	48	64

β) Integral Image

Εικόνα 5: Μετατροπή εικόνας σε Π.Π.Ε  
Πηγή: <https://www.youtube.com/watch?v=Wwn81tVIR1O>

Αφού δημιουργηθεί ο Πίνακας Προστιθέμενου Εμβαδού, χρησιμοποιείται για την εύρεση του συνόλου των τιμών των εικονοστοιχείων για τα ζητούμενα ορθογώνια των φίλτρων. Στην προαναφερθείσα περίπτωση ο υπολογισμός του συνόλου ( $S$ ) των τιμών του ορθογωνίου με κορυφές τα σημεία  $A(x_a, y_a)$ ,  $B(x_b, y_b)$ ,  $C(x_c, y_c)$ ,  $D(x_d, y_d)$  γίνεται βάσει της σχέσης:

$$S = II(xd, yd) - II(xb, yb) - II(xc, yc) + II(xa, ya) = 64 - 32 - 32 + 16 = 16$$



Εικόνα 6: Υπολογισμός Συνόλου  $S$  του Ορθογωνίου.

Πηγή: <https://www.youtube.com/watch?v=Wwn81tVIR1O>

Για τα στραμμένα κατά  $45^\circ$  χαρακτηριστικά δημιουργείται ο Στραμμένος Πίνακας Προστιθέμενου Εμβαδού ( $\Pi_R$ ), για τον οποίο ισχύουν οι ίδιες συνθήκες με τον Π.Π.Ε: έχει ίδιες διαστάσεις με την αρχική εικόνα και σε κάθε θέση του  $(x,y)$  περιλαμβάνεται το άθροισμα όλων των τιμών επάνω και αριστερά αυτής, καθώς και της ίδιας της  $(x,y)$ . Γενικά ισχύει:

$$IIR(x,y) = \sum_{y' \leq y, y' \leq y-x-x'} i(x',y')$$

Για να κατασκευαστεί λοιπόν ολόκληρος ο Στραμμένος Πίνακας Προστιθέμενου Εμβαδού ( $\Pi_R$ ) αρκεί μόνο ένα πέρασμα στην αρχική εικόνα, από αριστερά προς τα δεξιά και από πάνω προς τα κάτω, χρησιμοποιώντας την ακόλουθη σχέση:

$$IIR(x,y) = IIR(x-1,y-1) + IIR(x+1,y-1) - IIR(x,y-2) + i(x,y) + i(x,y-1),$$

Έχοντας τις αρχικές τιμές

$$IIR(-1,y) = IIR(x,-1) = IIR(x,-2) = 0 \text{ και } IIR(x,-1) = 0.$$

Αφού δημιουργηθεί ο Στραμμένος Πίνακας Προστιθέμενου Εμβαδού, χρησιμοποιείται για την εύρεση του συνόλου των τιμών των εικονοστοιχείων για τα ζητούμενα στραμμένα ορθογώνια των φίλτρων. Για την αντίστοιχη περίπτωση της Εικόνας 5 (αλλά αυτή τη φορά με στραμμένο ορθογώνιο), ο υπολογισμός των συνόλου (S) των τιμών του ορθογωνίου με κορυφές τα σημεία A( $x_a, y_a$ ), B( $x_b, y_b$ ), C( $x_c, y_c$ ), D( $x_d, y_d$ ) γίνεται όπως και στον Π.Π.Ε. βάσει της σχέσης:

$$S = IIR(xd,yd) - IIR(xb,yb) - IIR(xc,yc) + IIR(xa,ya),$$

Όπως έχει ήδη αναφερθεί, από την εφαρμογή των φίλτρων τύπου Haar προκύπτουν πολυάριθμα αποτελέσματα, αφού η εφαρμογή αυτή γίνεται σε όλη την εικόνα και σε κάθε δυνατή κλίμακα. Αυτός ήταν και ο λόγος που οι Viola και Jones προσέθεσαν έναν ταξινομητή χρησιμοποιώντας τον αλγόριθμο AdaBoost, έτσι ώστε να απορρίπτεται η πλειοψηφία των εξαχθέντων χαρακτηριστικών και να χρησιμοποιείται μόνο μία ομάδα με τα πιο κρίσιμα. Έτσι, η τελική εκπαίδευση του αλγόριθμου γίνεται με τη χρήση λιγότερων χαρακτηριστικών, άρα σε σαφώς πιο μειωμένο χρόνο, όπως θα δούμε στις επόμενες ενότητες.

(Paul Viola, 2001) (Βαρυτιμίδης, 2008) (ΓΕΩΡΓΙΟΥ, 2014)

### 2.1.3 ΑΛΓΟΡΙΘΜΟΣ ADABoost

Με τη χρήση του Πίνακα Προστιθέμενου Εμβαδού και του αντίστοιχου Στραμμένου Πίνακα, τα χαρακτηριστικά τύπου Haar εντοπίζονται με μεγάλη ταχύτητα, καθώς όμως το σύνολο των χαρακτηριστικών στην εικόνα είναι πολύ μεγάλο (υπερπολλαπλάσιο από αυτό των εικονοστοιχείων της), η μετέπειτα χρήση τους θα καθιστούσε τη συνολική διαδικασία πολύ χρονοβόρα. Γι' αυτό το λόγο επιλέγεται ένας μικρός αριθμός χαρακτηριστικών, βάσει των οποίων κατασκευάζεται ένας ταξινομητής δεδομένων.

Ο ταξινομητής αυτός λαμβάνει ως δεδομένα ένα μεγάλο αριθμό με «θετικές» και «αρνητικές» εικόνες, δηλαδή τα δείγματα στα οποία περιέχεται ή όχι ένα (ή και περισσότερα) αντικείμενο. Με βάση αυτές τις εικόνες λοιπόν, εκπαιδεύεται στον εντοπισμό του συγκεκριμένου αντικειμένου σε νέες εικόνες, όπως θα δούμε παρακάτω.

Ενώ ο αλγόριθμος AdaBoost στην αρχική του μορφή χρησιμοποιείται για την ενίσχυση της λειτουργίας ενός απλού ταξινομητή (ο οποίος αναφέρεται και ως αδύναμος), στη περίπτωσή μας χρησιμοποιείται και για τη διαλογή των σημαντικότερων χαρακτηριστικών της εικόνας. Ο απλός ταξινομητής χαρακτηρίζεται ως αδύναμος καθώς οι πιθανότητες της σωστής ταξινόμησης των δεδομένων μπορεί να είναι ελαφρά μεγαλύτερη του 50%, δηλαδή της τυχαίας ταξινόμησης. Για να προκύψει ένας ισχυρότερος ταξινομητής λοιπόν, γίνεται συνδυασμός πολλών ασθενών. Αναλυτικότερα η διαδικασία έχει ως ακολούθως:

Καταρχάς, δίνονται δείγματα εικόνων, θετικά και αρνητικά, για την εκπαίδευση του αλγόριθμου, δηλαδή εικόνες  $(x_1, y_1), \dots, (x_n, y_n)$  με  $y_i=0$  για αρνητικά παραδείγματα και  $y_i=1$  για θετικά. Στη συνέχεια, αρχικοποιείται μία τιμή βάρους ( $\omega$ ) σε όλα τα δείγματα (i) των εικόνων, είτε θετικά είτε αρνητικά, η οποία και είναι η ίδια για όλα:

$$\omega_{1,i} = \frac{1}{2m} \frac{1}{2l'}$$

όπου m το σύνολο των αρνητικών παραδειγμάτων και l το σύνολο των θετικών.

Έτσι, πραγματοποιείται η αρχική εκπαίδευση (αρχικοποίηση) του ταξινομητή μας, γίνεται δηλαδή η πρώτη ταξινόμηση όλων των παραδειγμάτων.

Έπειτα, για ένα διάστημα  $t=1$  έως  $T$  επαναλήψεων (όσοι θα είναι και οι ταξινομητές που θα προκύψουν), επαναλαμβάνονται τα παρακάτω βήματα:

- Πραγματοποιείται κανονικοποίηση των βαρών, ώστε για κάθε χαρακτηριστικό  $j$  με βάρος  $\omega$ , να ισχύει:

$$\omega_{t,i} \leftarrow \frac{\omega_{t,i}}{\sum_{j=1}^n \omega_{t,j}},$$

- Εκπαιδεύεται αντίστοιχα κάθε ταξινομητής  $h_j$  με σφάλμα:

$$e_j = \sum_i \omega_i |h_j(x_i) - y_i|,$$

- Επιλέγεται ο ταξινομητής με το χαμηλότερο σφάλμα  $e_t$ ,
- Ενημερώνονται τα βάρη, ώστε:

$$\omega_{t+1,i} = \omega_{t,i} \beta t^{1-e_i},$$

όπου  $e_i = 0$  σε περίπτωση που το παραδειγμα  $x_i$  έχει ταξινομηθεί σωστά και  $e_i = 1$  στην περίπτωση λανθάνουσας ταξινόμησης, ενώ:  $\beta t = \frac{et}{1-et}$ .

Με αυτόν τον τρόπο δίνεται μεγαλύτερο βάρος στα παραδείγματα που ταξινομήθηκαν λανθασμένα, έτσι ώστε με την επανάληψη των παραπάνω βημάτων, να μπορούμε να διακρίνουμε τους καλύτερους ασθενείς ταξινομητές. Οι ταξινομητές αυτοί χρησιμοποιούνται ώστε να προκύψει ο τελικός (ισχυρός ταξινομητής):

$$h(x) = 1, \text{ εάν } \sum_{t=1}^T \omega_t h_t(x) \geq 0.5, \sum_{t=1}^T \omega_t,$$

διαφορετικά  $h(x) = 0$ , όπου  $\omega_t = \log$ .

Αθροίζονται δηλαδή τα αποτελέσματα από τους καλύτερους ασθενείς ταξινομητές και στην περίπτωση που σύνολό τους περνά ένα προκαθορισμένο κατώφλι, συμπεραίνεται ότι το εκάστοτε χαρακτηριστικό έχει κριθεί σωστά, αλλιώς λανθασμένα.

Καθώς στην παραπάνω μέθοδο χρησιμοποιείται ένα μόνο χαρακτηριστικό για κάθε στάδιο του αλγόριθμου, ο αδύναμος ταξινομητής επιλέγει τα απλά χαρακτηριστικά, που βοηθούν στη διάκριση των θετικών από τα αρνητικά παραδείγματα. Έτσι, για κάθε χαρακτηριστικό  $f_j(x)$  ο ταξινομητής  $h_j(x)$  καθορίζει το βέλτιστο κατώφλι ( $\theta_j$ ), για τη διαδικασία της ταξινόμησης, έτσι ώστε να γίνει εσφαλμένος εντοπισμός σε όσο το δυνατόν λιγότερα παραδείγματα. Ισχύει:

$$h_j(x) = \begin{cases} 1, & p_{jfi}(x) < pj\theta_j \\ 0, & p_{jfi}(x) \geq pj\theta_j \end{cases}$$

όπου  $p_j$  μία ισοτιμία που δείχνει την κατεύθυνση της ανισότητας.

Ο αλγόριθμος AdaBoost είναι γνωστός τόσο για τη χρηστικότητα του, λόγω της υψηλούς ταχύτητας λειτουργίας του, όσο και για την ορθότητα των αποτελεσμάτων του, καθώς έχει αποδειχθεί ότι το σφάλμα ταξινόμησης του ισχυρού ταξινομητή που προκύπτει από την εφαρμογή του, τείνει προς το μηδέν εκθετικά, ως προς τον αριθμό των κύκλων εκπαίδευσης. (Paul Viola, 2001) Έτσι καθίσταται χρήσιμο εργαλείο για την ορθή ανίχνευση αντικειμένων σε εικόνες, με μεγάλες ταχύτητες.

(Paul Viola, 2001) (Βαρυτιμίδης, 2008) (ΓΕΩΡΓΙΟΥ, 2014)

#### 2.1.4 ΑΛΛΗΛΟΥΧΙΑ ΤΑΞΙΝΟΜΗΤΩΝ

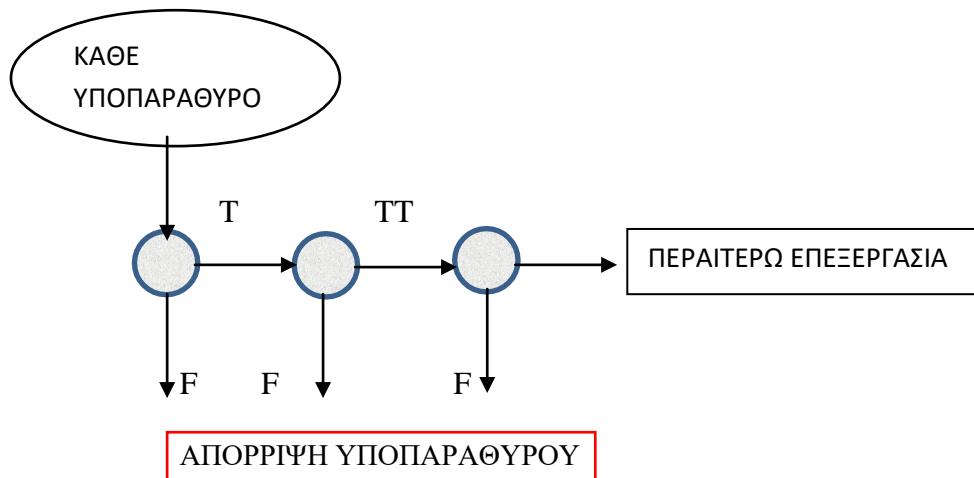
Στην προηγούμενη ενότητα παρουσιάστηκαν οι ταξινομητές των δειγμάτων για την ανίχνευση αντικειμένων και ως εκ τούτου καταλήξαμε στην αναγκαιότητά τους, καθώς καθιστούν τη διαδικασία ανίχνευσης αρκετά γρήγορη. Η ταχύτητα αυτή όμως δεν είναι αρκετά μεγάλη όταν στοχεύουμε στην ανίχνευση σε πραγματικό χρόνο. Γι' αυτό το λόγο κατασκευάστηκε ένα πιο σύνθετο σύστημα ταξινόμησης, που μέσω της αλληλουχίας μερικών ταξινομητών (cascade of classifiers) μπορεί να «φιλτράρει» ακόμη γρηγορότερα και με μεγαλύτερη ακρίβεια τα αντικείμενα στα οποία έχει εκπαιδευτεί ο αλγόριθμος. Σύμφωνα με τους Viola και Jones, το κλειδί για την επίτευξη αυτού του στόχου είναι η δημιουργία μικρών, άρα και αποδοτικών ταξινομητών, οι οποίοι μπορούν να απορρίπτουν πολλά από τα λανθασμένα παραδείγματα, ενώ ταυτόχρονα εντοπίζουν σχεδόν όλα τα θετικά.

Γενικά, η αρχή που ακολουθείται είναι οι απλοί αυτοί ταξινομητές να εντοπίζουν και να απορρίπτουν γρήγορα την πλειοψηφία των υποπαραθύρων, που τελικά δεν αντιστοιχούν σε κάποιο ζητούμενο αντικείμενο, ενώ πιο περίπλοκοι και χρονοβόροι ταξινομητές να αξιοποιούνται έπειτα, ώστε να επιτυγχάνουν λιγότερους εσφαλμένους εντοπισμούς (ψευδών θετικών) δειγμάτων. Η δομή του συστήματος των διαδοχικά συνδεδεμένων ταξινομητών έχει ως εξής:

Για κάθε εξεταζόμενο υποπαράθυρο της εικόνας, ο πρώτος ταξινομητής κάνει τον πρώτο έλεγχο και το κατατάσσει ως θετικό ή αρνητικό. Στην περίπτωση που είναι

θετικό, περνά ως είσοδος στο δεύτερο ταξινομητή. Εάν κριθεί ξανά ως θετικό, περνάει στον τρίτο και η διαδικασία αυτή επαναλαμβάνεται για όλους τους ταξινομητές, ενώ σε οποιοδήποτε στάδιο κριθεί ως αρνητικό, σταματάει άμεσα η διαδικασία και το δείγμα απορρίπτεται (βλ. Σχήμα 1).

Όπως είναι προφανές, η προαναφερθείσα διαδικασία είναι παρόμοια με αυτή που ακολουθεί κάθε απλός ταξινομητής, με τη βασική διαφορά όμως, ότι ο τελευταίος δίνει τα συνολικά αποτελέσματα από όλη τη διαδικασία, αφού ολοκληρώσει την εργασία του. Αντίθετα, κατά τη διάρκεια της λειτουργίας των διαδοχικά συνδεδεμένων ταξινομητών, γίνεται συχνός έλεγχος του συνόλου των τιμών των χαρακτηριστικών ώστε να καταλήξουμε εάν το εκάστοτε υποπαράθυρο θα πρέπει, ή όχι, να απορριφθεί (βλ. Σχήμα 1).



Σχήμα 1: Στάδια επιλογής υποπαραθύρων.

Κατά τη λειτουργία του παραπάνω συστήματος, κάθε ταξινομητής λαμβάνει ως αρνητικά δείγματα, εκείνα που οι προηγούμενοι εσφαλμένα έκριναν ως θετικά. Έτσι, σε αντίθεση με τα θετικά παραδείγματα, το σύνολο των αρνητικών αυξάνεται κατακόρυφα. Το γεγονός αυτό δυσχεραίνει την εκπαίδευση των πιο περίπλοκων ταξινομητών, καθώς όσο πιο δύσκολη γίνεται η ταξινόμηση, τόσο αυξάνονται τα λανθασμένα αποτελέσματα. Ταυτόχρονα όμως, οδηγεί στην αύξηση της τελικής επίδοσης του συστήματος.

Οι αρχικοί ταξινομητές πρέπει να έχουν χαμηλό κατώφλι, ώστε να μη χάνεται η πληροφορία των πραγματικών αντικειμένων. Για τη σωστή λειτουργία του

συστήματος όμως, πρέπει επίσης να αυξηθούν οι θετικές ταξινομήσεις. Για να το επιτύχουμε αυτό, οι ταξινομητές εκπαιδεύονται μέσω του αλγόριθμου Adaboost και με τον τρόπο αυτό ρυθμίζεται η τιμή του κατωφλίου τους. Να σημειωθεί όμως, ότι ένα χαμηλό κατώφλι, ενώ οδηγεί σε υψηλότερο αριθμό εντοπισμένων στοιχείων, προκαλεί επίσης περισσότερα ψευδώς θετικά δείγματα.

(Paul Viola, 2001) (Βαρυτιμίδης, 2008) (ΓΕΩΡΓΙΟΥ, 2014)

Ο αλγόριθμος των Viola και Jones αποτέλεσε επανάσταση στο χώρο της όρασης των Υπολογιστών και της Μηχανικής Μάθησης. Πάνω σε αυτόν στηρίζονται μέχρι και σήμερα πολλές εφαρμογές ανίχνευσης αντικειμένων, είτε αμιγώς με αυτά τα βήματα, είτε σε συνδυασμό με άλλες μεθόδους (π.χ. Φίλτρα H.O.G.)

### 3 ΚΩΔΙΚΟΠΟΙΗΜΕΝΟΙ ΣΤΟΧΟΙ

Στο κεφάλαιο αυτό, παρουσιάζεται η θεωρία γύρω από τους κωδικοποιημένους στόχους (coded targets), οι οποίοι χρησιμοποιούνται κατά κόρων στις μέρες μας και γίνεται λόγος για τους τρόπους δημιουργίας τους, τις μεθόδους αποκωδικοποίησής τους και τις εφαρμογές στις οποίες γίνεται χρήση τους.

Ο κώδικας QR (Quick Response ή Γρήγορης Απόκρισης) είναι ένας γραμμωτός κώδικας (barcode) δύο διαστάσεων (wikipedia), που δημιουργήθηκε με σκοπό την ταχεία αποκωδικοποίηση των δεδομένων, τα οποία εμπεριέχει. Η χρήση του είναι ευρεία, καθώς το συναντάμε καθημερινά σε διαφόρων ειδών έντυπα- από αεροπορικά εισιτήρια, μέχρι διαφημιστικά φυλλάδια και επαγγελματικές κάρτες- αλλά και σε ψηφιακή μορφή (π.χ. σε αρχείο εικόνας). Με τη σάρωσή του, αντλούμε πληροφορίες σχετικά με το εκάστοτε αντικείμενο. Η σάρωση αυτή γίνεται, είτε μέσω της κάμερας ψηφιακού κινητού τηλεφώνου, είτε μέσω άλλων ψηφιακών συσκευών με τη χρήση του κατάλληλου λογισμικού (Εικόνα 7). Κλασσικό παράδειγμα αποτελεί πλέον η ευρεία χρήση που είχαν οι στόχοι αυτοί, επί της Πανδημίας COVID-19, όπου ήταν αναγκαία η επίδειξή τους, είτε σε έντυπη μορφή, είτε σε ψηφιακή, ώστε να επιτραπεί η είσοδος σε κάποιο χώρο (π.χ. εργασιακό), με σκοπό δηλαδή την ενημέρωση του αν κάποιος ήταν εμβολιασμένος ή όχι κατά του ιού.

Ο κώδικας QR αποτελείται από έναν τετραγωνικό πίνακα μαύρων και λευκών τετραγώνων που αναπαριστούν τα κωδικοποιημένα δεδομένα σε δυαδική μορφή. Μια ειδική σήμανση σε τρεις από τις τέσσερις γωνίες του τετραγώνου παρέχει προσανατολισμό. Τα δεδομένα στον κώδικα QR επεκτείνονται με έναν κώδικα διόρθωσης σφαλμάτων. Αυτό σημαίνει ότι η απώλεια έως και 30% του κώδικα είναι ανεκτή, δηλαδή μπορεί ακόμα να αποκωδικοποιηθεί. (wikipedia)



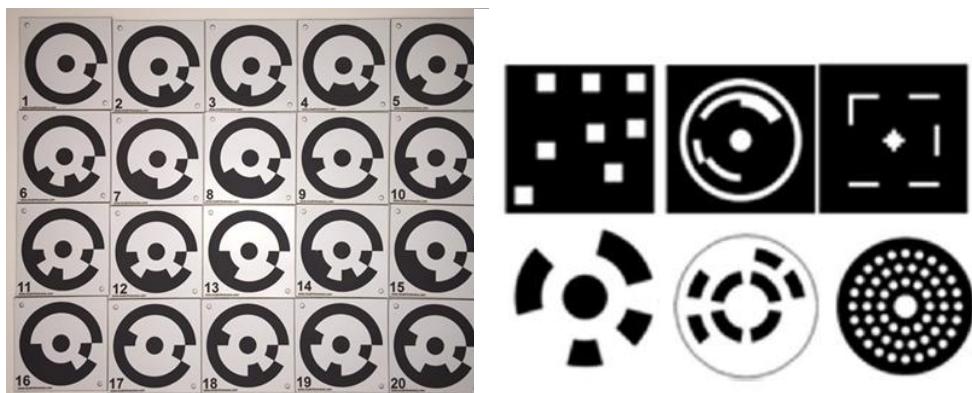
Εικόνα 7: Κώδικας QR και σάρωσή του.

Πηγές: <https://www.erasmus.gr/microsites/1012/ti-einai-qr-code>

<https://www.barcodesinc.com/news/new-barcode-workflows-for-nurses/>

Τα τελευταία χρόνια όμως, η χρήση του κώδικα QR έφτασε και στο χώρο της Φωτογραμμετρίας, καθώς χρησιμοποιείται πλέον για την αυτόματη αναζήτηση και ταχεία αποκωδικοποίηση των προσημασμένων σημείων στο χώρο, μέσω της λήψης φωτογραφιών. Συγκεκριμένα, γίνεται χρήση ειδικών κωδικοποιημένων στόχων, συνήθως στρόγγυλης ή τετράγωνης μορφής, οι οποίοι τοποθετούνται συνήθως γύρω από το αντικείμενο που ζητείται να αποτυπωθεί.

Ο εντοπισμός των στόχων αυτών προκύπτει από αυτοματοποιημένες διαδικασίες, αποφεύγοντας έτσι τον κίνδυνο χονδροειδών σφαλμάτων και τις χρονοβόρες εργασίες. Αυτός είναι και ο λόγος που καθίσταται χρήσιμη η συγκεκριμένη μέθοδος στο πεδίο της Φωτογραμμετρίας, καθώς μπορεί να παραχθεί εύκολα και γρήγορα η ζητούμενη πληροφορία ανά εφαρμογή (π.χ. τρισδιάστατη απεικόνιση κτισμάτων). Δίνεται λοιπόν, η δυνατότητα του αυτόματου εντοπισμού, της αναγνώρισης και της μέτρησης των στόχων, που σηματοδοτούν τα σημεία ενδιαφέροντος.



Εικόνα 8: Παραδείγματα Κωδικοποιημένων Στόχων

Πηγή: (Ran Chen K. Z.)

Ειδικότερα, στις εφαρμογές της φωτογραμμετρίας που σχετίζονται με τα μετρητικά συστήματα τριών διαστάσεων, ένας από τους κυρίαρχους, αλλά και δυσκολότερους στόχους, είναι η εύρεση των ομόλογων σημείων σε πολλαπλές εικόνες- το λεγόμενο «πρόβλημα της αντιστοίχησης των σημείων». Επιλύοντας αυτό το πρόβλημα, επιτυγχάνεται η ακριβής μέτρηση των σημείων στις εικόνες και η αντιστοίχησή τους, απαλλαγμένη από σφάλματα, γεγονός που επηρεάζει άμεσα την ποιότητα των αποτελεσμάτων των τρισδιάστατων μετρήσεων. Στην επίλυση του προβλήματος αυτού συντελεί η παρουσία των κωδικοποιημένων στόχων.

Κατά τους Chen, Zhong, Li, Liu και Zhan (Ran Chen K. Z., 2016), η διανομή κωδικοποιημένων στόχων στο μετρούμενο αντικείμενο είναι μια συνηθισμένη και αξιόπιστη μέθοδος για το βέλτιστο εντοπισμό της θέσης των στόχων και την αξιόπιστη αντιστοίχηση των στόχων σε πολλαπλές εικόνες. Με τη χρήση των σύγχρονων ψηφιακών φωτογραφικών μηχανών, ο αριθμός αυτών των εικόνων μπορεί να ανέρχεται σε μερικές εκατοντάδες, όταν αφορά σε μεγάλα και πολύπλοκα αντικείμενα. Σε αυτές τις περιπτώσεις δεν είναι πρακτικό να χρησιμοποιούνται μόνο χειροκίνητα μέσα για τον εντοπισμό των ομόλογων σημείων στις εικόνες.

Γίνεται εύκολα αντιληπτό λοιπόν, ότι με τη χρήση των κωδικοποιημένων στόχων στις φωτογραμμετρικές εφαρμογές επιτυγχάνεται τόσο η ποιότητα των μετρήσεων, όσο και η εξοικονόμηση χρόνου και κόστους.

(Sung JoonAhn, 2016) (SungJoonAhn, 1999) (Ran Chen K. Z., 2016)

Οι κωδικοποιημένοι στόχοι θα πρέπει να ικανοποιούν συγκεκριμένες προδιαγραφές, μερικές από τις οποίες είναι οι ακόλουθες:

- Να είναι αμετάβλητοι στις περιστροφές και στις αλλαγές της κλίμακας.
- Δυνατότητα ανίχνευσης στόχου στην εικόνα, όταν υπάρχει ανομοιόμορφη αντίθεση.
- Ελάχιστος χρόνος επεξεργασίας.
- «Συμπυκνωμένο» σχήμα.

(Vladimir A. Knyaz, 1998)

Για τη χρήση των κωδικοποιημένων στόχων, υπάρχουν λογισμικά όπως το PhotoModeler, AutodeskReCap, 3DF Zephyr, καθώς και το AgisoftPhotoScan, με το οποίο και θα ασχοληθούμε παρακάτω.

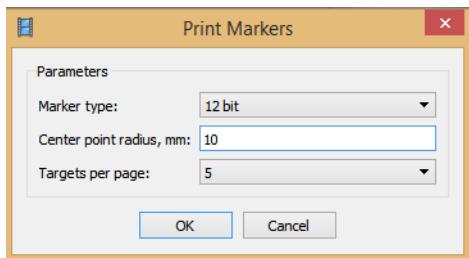
### **3.1 ΤΡΙΣΔΙΑΣΤΑΤΗ ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΑΝΤΙΚΕΙΜΕΝΟΥ ΜΕ ΤΗ ΧΡΗΣΗ ΚΩΔΙΚΟΠΟΙΗΜΕΝΩΝ ΣΤΟΧΩΝ**

Μία από τις εφαρμογές των κωδικοποιημένων στόχων στο πεδίο της Φωτογραμμετρίας είναι και αυτή της τρισδιάστατης μοντελοποίησης αντικειμένων, με τη μέθοδο Structure From Motion (SFM). Παρακάτω γίνεται μία σύντομη περιγραφή της διαδικασίας αυτής, στο λογισμικό AgisoftPhotoScan.

Στις εφαρμογές όπου χρησιμοποιούνται κωδικοποιημένοι στόχοι, σκοπός είναι ο αυτόματος εντοπισμός τους, αλλά και η επιτυχής αποκωδικοποίησή τους. Η ακρίβεια αυτών όμως, εξαρτάται καταρχήν από τη σωστή κωδικοποίησή τους, καθώς θα πρέπει κάθε στόχος να είναι μοναδικός και «ανθεκτικός» στις εκάστοτε συνθήκες λήψης των φωτογραφιών.

### 3.1.1 ΔΗΜΙΟΥΡΓΙΑ ΚΑΙ ΤΟΠΟΘΕΤΗΣΗ ΚΩΔΙΚΟΠΟΙΗΜΕΝΩΝ ΣΤΟΧΩΝ

Για τη δημιουργία και την εκτύπωση των στόχων, μπορεί εύκολα κάποιος να χρησιμοποιήσει τα εργαλεία από το περιβάλλον του Agisoft Photo Scan, όπου μπορεί να επιλεγεί ο τύπος, το μέγεθος του κέντρου και ο αριθμός των στόχων (Εικόνα 9).



Εικόνα 9: Επιλογές Παραμέτρων Στόχων

Πηγή: [https://www.agisoft.com/pdf/PS\\_1.1\\_Tutorial%20\(IL\)%20-%20Coded%20Targets%20and%20Scale%20Bars.pdf](https://www.agisoft.com/pdf/PS_1.1_Tutorial%20(IL)%20-%20Coded%20Targets%20and%20Scale%20Bars.pdf)

Κατά το αντίστοιχο εγχειρίδιο της Agisoft, ο σωστός τρόπος τοποθέτησης των στόχων είναι ο εξής, όπως φαίνεται στις Εικόνα 10.



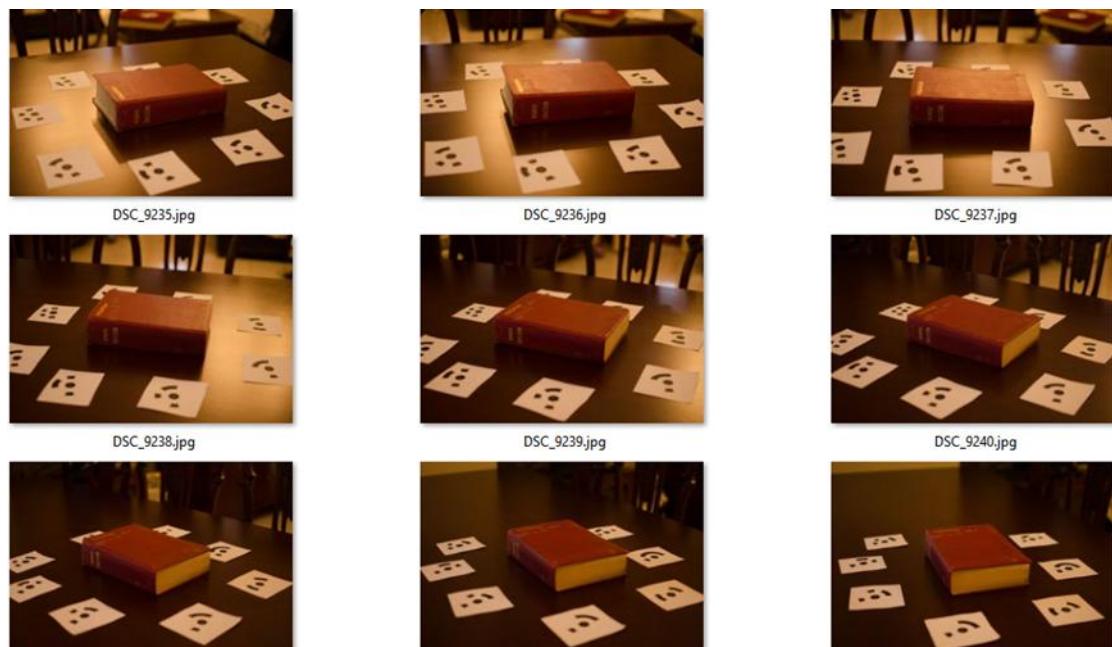
Εικόνα 10: Μεθοδολογία τοποθέτησης των στόχων

Πηγή: [https://www.agisoft.com/pdf/PS\\_1.1\\_Tutorial%20\(IL\)%20-%20Coded%20Targets%20and%20Scale%20Bars.pdf](https://www.agisoft.com/pdf/PS_1.1_Tutorial%20(IL)%20-%20Coded%20Targets%20and%20Scale%20Bars.pdf)

Αφού λοιπόν εκτυπωθούν όλοι οι στόχοι, τοποθετούνται διαδοχικά και ομοιόμορφα γύρω από το αντικείμενο ενδιαφέροντος, έτσι ώστε να είναι όλοι ευδιάκριτοι, σε τουλάχιστον δύο φωτογραφίες.

### 3.1.2 ΛΗΨΕΙΣ ΕΙΚΟΝΩΝ

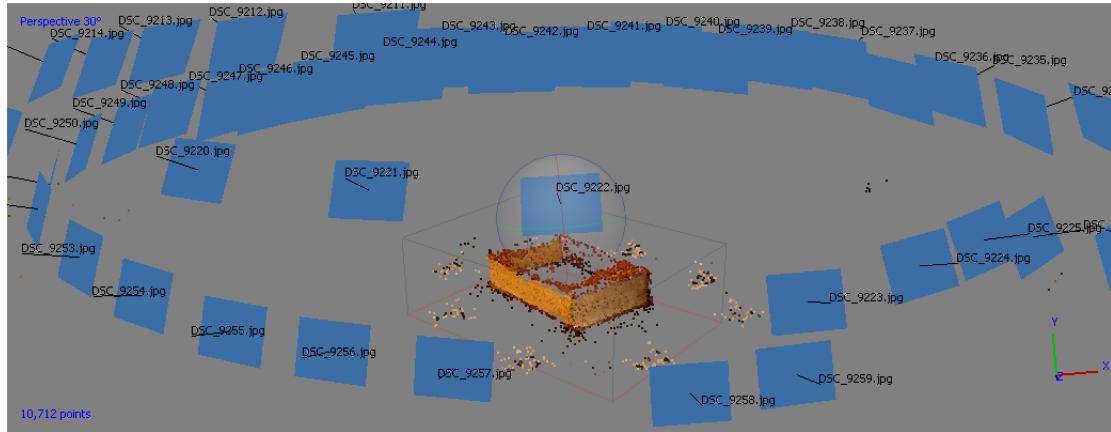
Ξεκινώντας τη διαδικασία για την αποτύπωση αντικειμένου τοποθετούνται οι προσημασμένοι στόχοι γύρω του, όπως περιγράφηκε παραπάνω και γίνεται λήψη των φωτογραφιών. Για την εξασφάλιση σωστών αποτελεσμάτων, ορθό είναι να υπάρχει εξαρχής ένα πλάνο, όσον αφορά τις θέσεις της φωτογραφικής μηχανής, ανάλογα με τον εκάστοτε χώρο και φυσικά το ζητούμενο αντικείμενο. Για παράδειγμα, σύμφωνα με το εγχειρίδιο της Agisoft, για την αποτύπωση ενός τμήματος αρχαιολογικού χώρου μετά την ανασκαφή του (συνήθως 4mx 4m περίπου), θα πρέπει να γίνει λήψη 10 περίπου φωτογραφιών κατά μήκος κάθε πλευράς, μίας τουλάχιστον φωτογραφίας πάνω από κάθε πλευρά, δηλαδή σε μεγαλύτερο ύψος από τις προηγούμενες και μίας σε καθεμιά από τις τέσσερεις γωνίες του πλαισίου. Επίσης, πολύ σημαντικός παράγοντας είναι η μηχανή να βρίσκεται σε όσο το δυνατόν πιο κατακόρυφη θέση, καθώς και οι εικόνες να έχουν επικάλυψη περίπου 60-80% μεταξύ τους.



Εικόνα 11: Δείγμα λήψεων φωτογραφιών

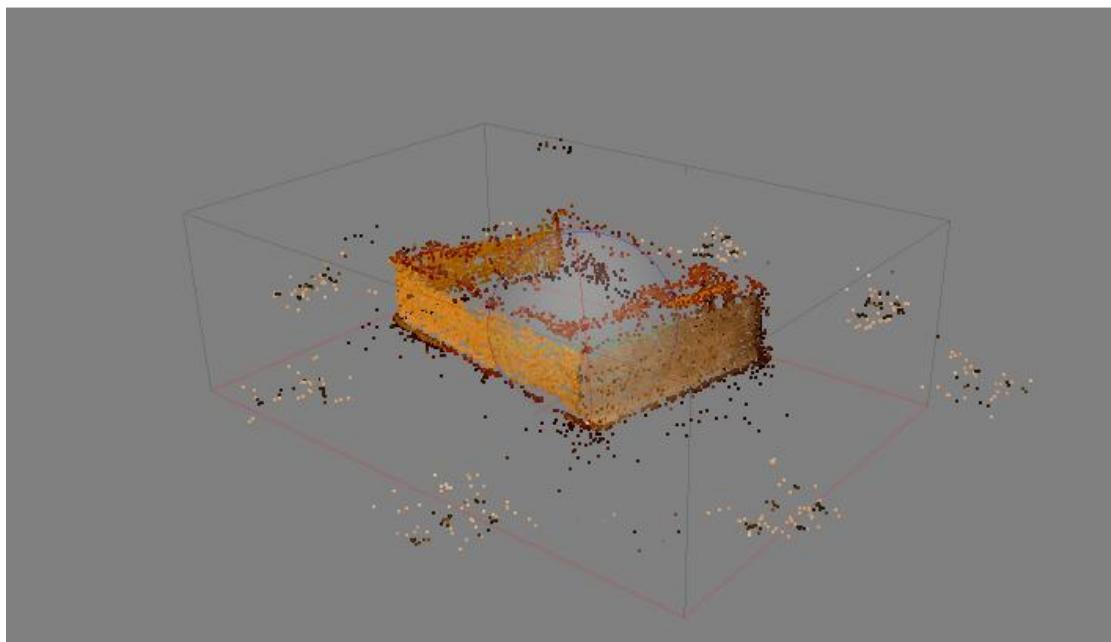
### 3.1.3 ΕΠΕΞΕΡΓΑΣΙΑ ΕΙΚΟΝΩΝ

Στη συνέχεια, αφού εισαγθούν οι φωτογραφίες στο περιβάλλον του Agisoft Photo Scan και επιλεγούν αυτές που θεωρούνται κατάλληλες βάσει των παραπάνω κανόνων, πραγματοποιείται η ευθυγράμμισή τους (Align Photos), όπως στην Εικόνα 12.



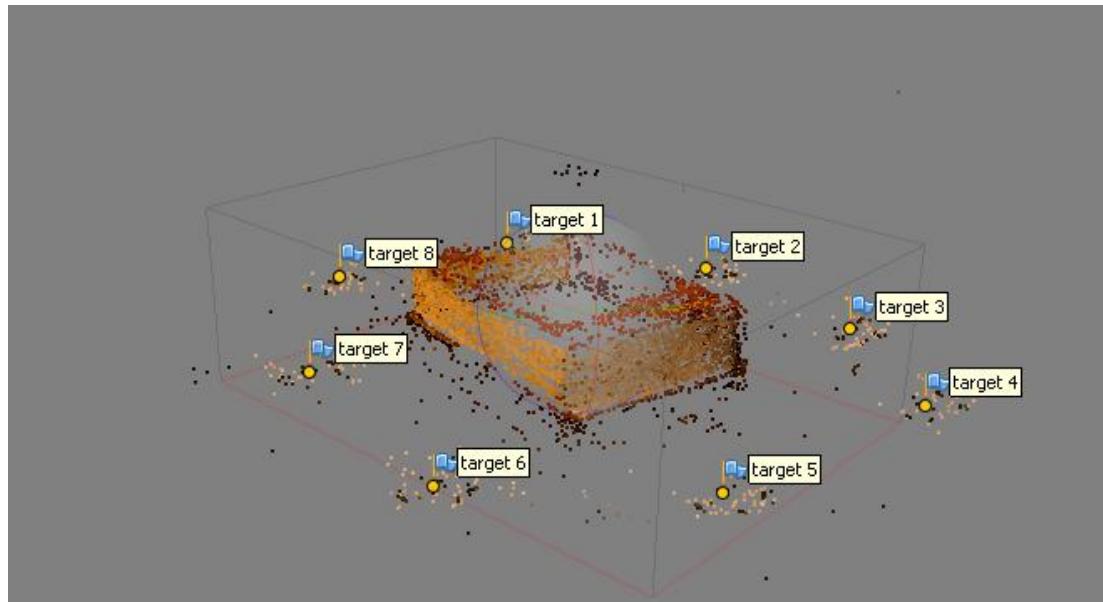
Εικόνα 12: Ευθυγράμμιση εικόνων και προσανατολισμός τους

Έτσι, υπολογίζονται οι θέσεις λήψης των εικόνων και γίνεται η πρώτη αντιστοίχηση των κοινών σημείων τους. Σε περίπτωση που κρίνεται απαραίτητο, μπορούν να αφαιρεθούν εικόνες που δε θεωρούνται κατάλληλες για την εργασία, ενώ επίσης και να επαναληφθεί η διαδικασία του προσανατολισμού τους. Έπειτα, όπως φαίνεται στην Εικόνα 13, δημιουργείται αυτόματα ένα αραιό «σύννεφο» σημείων, ώστε να υπάρξει μία πρώτη φωτορεαλιστική άποψη του αντικειμένου.



Εικόνα 13: Πρώτη φωτορεαλιστική άποψη

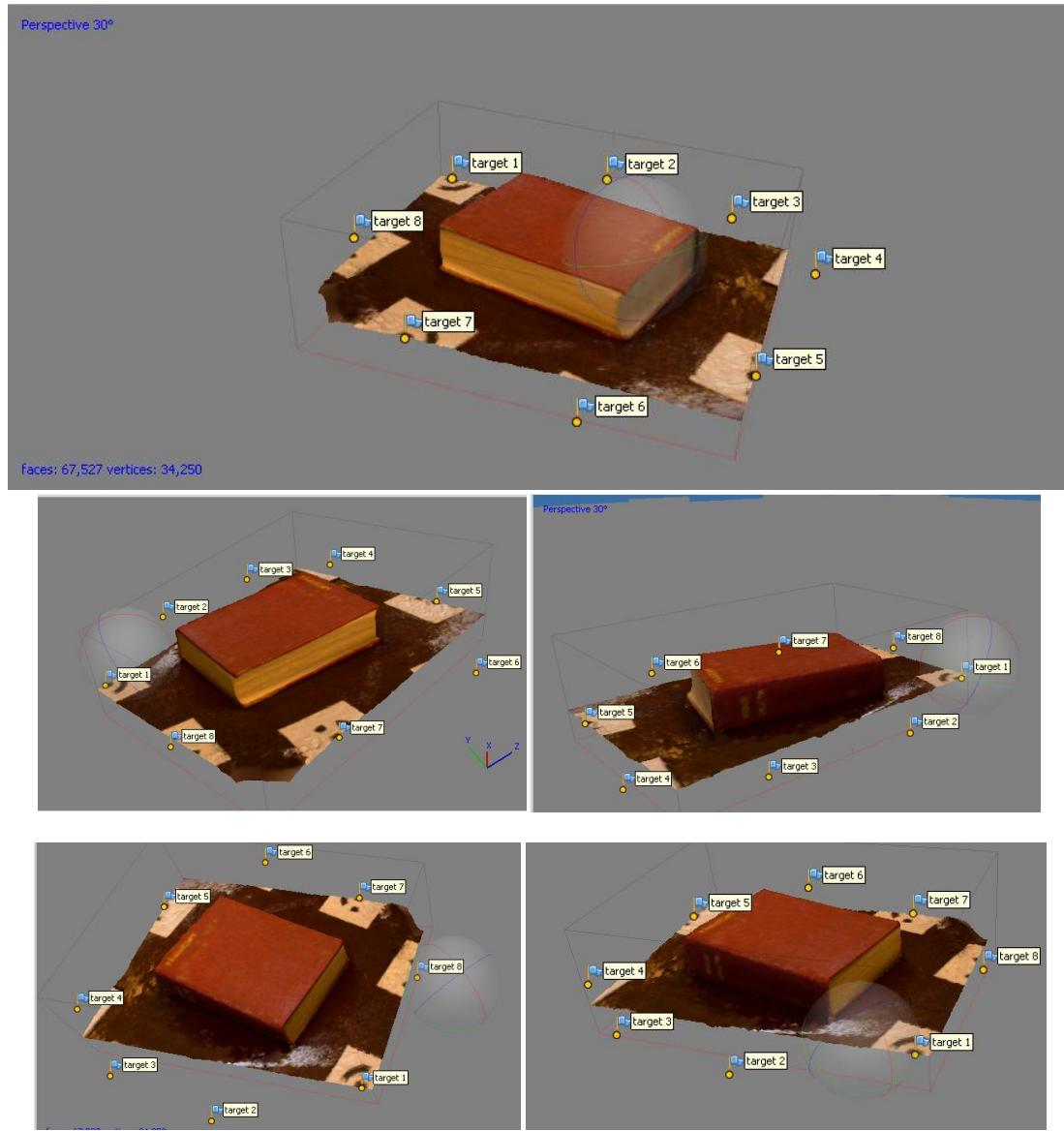
Αφού ολοκληρωθεί και ο προσανατολισμός των εικόνων, ακολουθεί ο αυτόματος «χαρακτηρισμός» των προσημασμένων στόχων (Detect Markers), του οποίου η ανεκτικότητα σε σφάλματα, ορίζεται εξ αρχής (Εικόνα 14).



Εικόνα 14: Απεικόνιση προσημασμένων σημείων

### 3.1.4 ΑΠΟΤΕΛΕΣΜΑΤΑ

Τέλος, στο περιβάλλον του Agisoft Photo μπορεί να παραχθεί ένα πυκνότερο δίκτυο σημείων, μέσω της επιλογής Build Dense Cloud, καθώς και να δοθεί «υφή» στο μοντέλο του αντικείμενο (Build Mesh & Build Texture),(Εικόνα 15).



Εικόνα 15: Τελικό μοντέλο με φωτογραφική υφή

Για να επιτευχθεί η εργασία αυτή με μεγαλύτερη ακρίβεια, καθώς και για να παραχθεί ένα πλήρες (φωτορεαλιστικά) μοντέλο του αντικειμένου, απαραίτητη είναι η μέτρηση φωτοσταθερών σημείων στο χώρο, με τη χρήση GPS ή Total Station. Να σημειωθεί όμως, ότι συγκριτικά με άλλες (παραδοσιακές) μεθόδους, η χρήση της μεθόδου SfM (Structure From Motion) που ακολουθείται στο Photo Scan, προσφέρει αρκετά ακριβή αποτελέσματα, με τυπικό σφάλμα 6-10mm. Με τέτοιες προϋποθέσεις είναι δυνατόν να υπάρξουν παρόμοιες εργασίες και σε αντικείμενα μεγαλύτερου

μεγέθους ή σε ευρύτερους χώρους, χρησιμοποιώντας για παράδειγμα ορθοφωτογραφίες μεγάλης κλίμακας.

## **4 ΠΕΙΡΑΜΑΤΙΚΗ ΕΦΑΡΜΟΓΗ: ΗΜΙΑΥΤΟΜΑΤΗ ΑΝΑΓΝΩΡΙΣΗ ΚΩΔΙΚΟΠΟΙΗΜΕΝΩΝ ΣΤΟΧΩΝ**

Στο κεφάλαιο αυτό, στόχος είναι η παρουσίαση ενός κώδικα που δημιουργήθηκε με σκοπό την αυτοματοποίηση της εύρεσης και του χαρακτηρισμού κωδικοποιημένων στόχων σε εικόνες, οι οποίοι δύναται να χρησιμοποιηθούν είτε ως σημεία αναφοράς, είτε ως φωτοσταθερά σημεία στο χώρο, σε περαιτέρω φωτογραμμετρικές εφαρμογές. Η εφαρμογή στηρίχθηκε στη θεωρία των Viola και Jones (βλ. Κεφάλαιο 2).

Η δυνατότητα της αυτόματης αναγνώρισης των στόχων στις εικόνες, θα μπορούσε να οδηγήσει στην απλοποίηση των εργασιών της φωτογραμμετρίας, όπως αυτή της τρισδιάστατης μοντελοποίησης αντικειμένων (π.χ. κτηρίων) και της παραγωγής ορθοφωτογραφιών και ψηφιακών μοντέλων εδάφους (DTM), εξασφαλίζοντας ταυτόχρονα, πολύ μεγάλη ακρίβεια, σε περίπτωση που πρόκειται για φωτογραμμετρικούς στόχους με γνωστές συντεταγμένες (GCPs).

Στην περίπτωση όμως, που δε διατίθεται η πληροφορία των συντεταγμένων στο χώρο, όπως σε δύσβατα σημεία, αρχαιολογικά μνημεία, πολύ μεγάλες εκτάσεις κλπ., οι στόχοι μπορούν να χρησιμοποιηθούν για την αυτόματη εύρεση και σηματοδότηση κοινών σημείων σε αλληλεπικαλυπτόμενες περιοχές διαδοχικών εικόνων, για την παραγωγή στερεοσκοπικής εικόνας (π.χ. αεροφωτογραφίες μεγάλων εκτάσεων). Επίσης, μπορεί να γίνει χρήση τους βοηθητικά, για την παραγωγή του πυκνού νέφους σημείων- σε περίπτωση τρισδιάστατου μοντέλου-, τον προσδιορισμό της θέσης λήψης των φωτογραφιών, χωρίς γνωστό εξωτερικό προσανατολισμό της φωτογραφικής μηχανής και τον αρχικό προσανατολισμό των εικόνων (image alignment). Ως εκ τούτου, προκύπτουν και τα σημεία σύνδεσης (tie points) των εικόνων.

Σε γενικές γραμμές, οι λειτουργίες αυτές θα μπορούσαν να πραγματοποιηθούν και με απλούς φωτογραμμετρικούς στόχους. Ιδιαίτερο ενδιαφέρον όμως, φαίνεται να έχουν στην περίπτωσή μας οι κωδικοποιημένοι στόχοι, μιας και φέρουν μία επιπλέον πληροφορία, σε σχέση με τους κοινούς στόχους, αυτή της «ταυτότητάς» τους. Με αυτό τον τρόπο μπορούν να αποφευχθούν χονδροειδή σφάλματα, ενώ ταυτόχρονα καταργείται η χειροκίνητη εύρεση στόχων στις εικόνες, ως ομόλογα σημεία. Γι' αυτό και η παρούσα εφαρμογή στηρίχθηκε στη χρήση κωδικοποιημένων στόχων, όπως αναλύεται στις ακόλουθες παραγράφους.

## **4.1 ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΣ ΕΡΓΑΣΙΩΝ**

Όταν αναφερόμαστε σε εργασίες MachineLearning και DeepLearning, αυτομάτως πρέπει να γίνεται αντιληπτή η αναγκαιότητα για συλλογή πληθώρας δεδομένων, τα οποία μετά την επεξεργασία τους μπορούν να «εκπαιδεύσουν» έναν πρόγραμμα. Στην προκειμένη περίπτωση, σκοπός ήταν η επιτυχής αναγνώριση κωδικοποιημένων στόχων, συνεπώς απαραίτητη ήταν η συλλογή μεγάλου πλήθους φωτογραφιών, τόσο «θετικών» όσο και «αρνητικών» δειγμάτων τους, ώστε να μπορεί να ολοκληρωθεί σωστά η εκμάθηση της υπολογιστικής μηχανής.

Οι εφαρμογές που αφορούν σε μεγάλο πλήθος δεδομένων είναι και αυτές που χρήζουν μεγαλύτερης ανάγκης από σωστή προετοιμασία και προγραμματισμό εργασιών. Με αυτόν τον τρόπο, αποφεύγονται χονδροειδή σφάλματα κατά τη συλλογή των δεδομένων και ως εκ τούτου, εξασφαλίζεται η εξοικονόμηση χρόνου καθώς και το επιθυμητό, όσον αφορά την ποιότητά του, αποτέλεσμα.

Στην παρούσα εργασία, σκοπός ήταν η δημιουργία ενός αλγόριθμου, ικανού να αναγνωρίζει 20 φωτοσταθερά σημεία συγκεκριμένων προτύπων μέσα σε εικόνες, στις οποίες συμπεριλαμβάνονται κι άλλες, τυχαίες, λεπτομέρειες οποιασδήποτε μορφής. Δεδομένου λοιπόν, ότι χρειαζόταν ένας μεγάλος αριθμός φωτογραφιών, καθώς και η απαραίτητη επεξεργασία τους, ώστε να προκύψει το επιθυμητό αποτέλεσμα, ακολούθησε ο κάτωθι προγραμματισμός.

### **4.1.1 ΕΠΙΛΟΓΗ ΛΟΓΙΣΜΙΚΩΝ**

Αρχικά έγινε χρήση του λογισμικού της Agisoft , «AgisoftPhotoscanPro», ώστε να παραχθούν οι 20 κωδικοποιημένοι στόχοι, οι οποίοι δημιουργήθηκαν με ακτίνα 3.5mm και με πλαίσιο χαρτιού μεγέθους 10cm x 10cm περίπου.

Σημειώνεται εδώ, ότι θα μπορούσαν να χρησιμοποιηθούν οποιοιδήποτε άλλοι κωδικοποιημένοι στόχοι, καθώς ο αλγόριθμος πρόκειται να εκπαιδευτεί πάνω στα δεδομένα που του δίνονται και είναι ανεξάρτητος οποιουδήποτε προγράμματος.

Έπειτα, επιλέχθηκε το προγραμματιστικό περιβάλλον «Matlab» της MathWorks, μέσα στο οποίο δύναται να γίνει ο χαρακτηρισμός των στόχων και η δημιουργία του κώδικα, που επρόκειτο να εκπαιδευτεί για την αναγνώρισή τους στις εκάστοτε εικόνες.

#### 4.1.2 ΛΗΨΕΙΣ ΦΩΤΟΓΡΑΦΙΩΝ

Στην πρώτη φάση της διαδικασίας, χρησιμοποιήθηκαν οι στόχοι του λογισμικού AgisoftPhotoscanPro, το μέγεθος των οποίων επιλέχθηκε με βάση το πλήθος των στόχων, την έκταση της επιφάνειας εργασίας και την απόσταση λήψης των φωτογραφιών. Σκοπός της προεργασίας αυτής ήταν οι στόχοι να είναι ευδιάκριτοι σε όλες τις εικόνες, ακόμη κι όταν επρόκειτο να φωτογραφηθούν όλοι ταυτόχρονα.

Έτσι, τοποθετήθηκαν οι 20 στόχοι σε έξι διαφορετικές, σχεδόν επίπεδες επιφάνειες, κατανεμημένοι με διαφορετική διάταξη σε κάθε περίπτωση και πάρθηκαν 298 φωτογραφίες, εκ των οποίων και μερικές με «καθαρές» επιφάνειες, δηλαδή χωρίς στόχους, καθώς και άλλες με «σβησμένους» στόχους, ώστε να μη φαίνεται το μοτίβο τους.



Εικόνα 16: Λήψη Φωτογραφιών με Στόχους.

Να σημειωθεί, ότι λήψη των περισσότερων φωτογραφιών έγινε σε θέση σχεδόν παράλληλη της καταγραφικής πλακέτας της φωτογραφικής μηχανής με αυτήν της εκάστοτε επιφάνειας εργασίας (π.χ. Εικόνα 16). Πάρθηκαν όμως και φωτογραφίες με διαφορετικές γωνίες λήψης, με έντονη προοπτική και πολλά σημεία φυγής (Εικόνα 17).

Καθ' αυτόν τον τρόπο, γίνεται προσπάθεια το πρόγραμμα που επρόκειτο να δημιουργηθεί στην πορεία, να καθίσταται ανθεκτικό σε μελλοντικές περιπτώσεις με έντονες γωνίες λήψης των εικόνων, με αποτέλεσμα να καλύπτει δύσκολες εφαρμογές,

όσον αφορά την πρόσβαση στις ζητούμενες περιοχές και τυχόν έντονο ανάγλυφο με μεγάλη ακτινική εκτροπή.



Εικόνα 17: Φωτογραφίες με Έντονη Προοπτική

#### 4.1.3 ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ ΕΙΚΟΝΩΝ

Εν συνεχείᾳ, οι εικόνες καταχωρίστηκαν σε «θετικές» και «αρνητικές».

Ως θετικές εικόνες νοούνται εκείνες, που χρησιμοποιούνται στον αλγόριθμο για να υποδείξουν θετικά δείγματα, οι οποίες λειτουργούν δηλαδή ως βάση για το λογισμικό που εκπαιδεύει το πρόγραμμα, ώστε να αναγνωρίζει τους στόχους σε μία εικόνα. Οι εικόνες που επιλέχθηκαν ως θετικές, ήταν εκείνες που λήφθηκαν σε διάφορα περιβάλλοντα και σε διαφορετικές ημέρες και ώρες, ενώ συμπεριλήφθηκαν κάποιες με έντονη προοπτική, αλλά και κάποιες με επιφάνεια σχεδόν παράλληλη της καταγραφικής πλακέτας της φωτογραφικής μηχανής. Διότι, όπως προλέχθηκε, κύριος στόχος ήταν το πρόγραμμα να μπορεί να εκπαιδευτεί σε όλες τις πιθανές συνθήκες που μπορεί να ληφθεί μία τέτοια φωτογραφία με στόχους. Πλήθος από αυτές τις εικόνες παρουσιάζεται στο Παράρτημα B.



Εικόνα 18: Δείγμα θετικών εικόνων

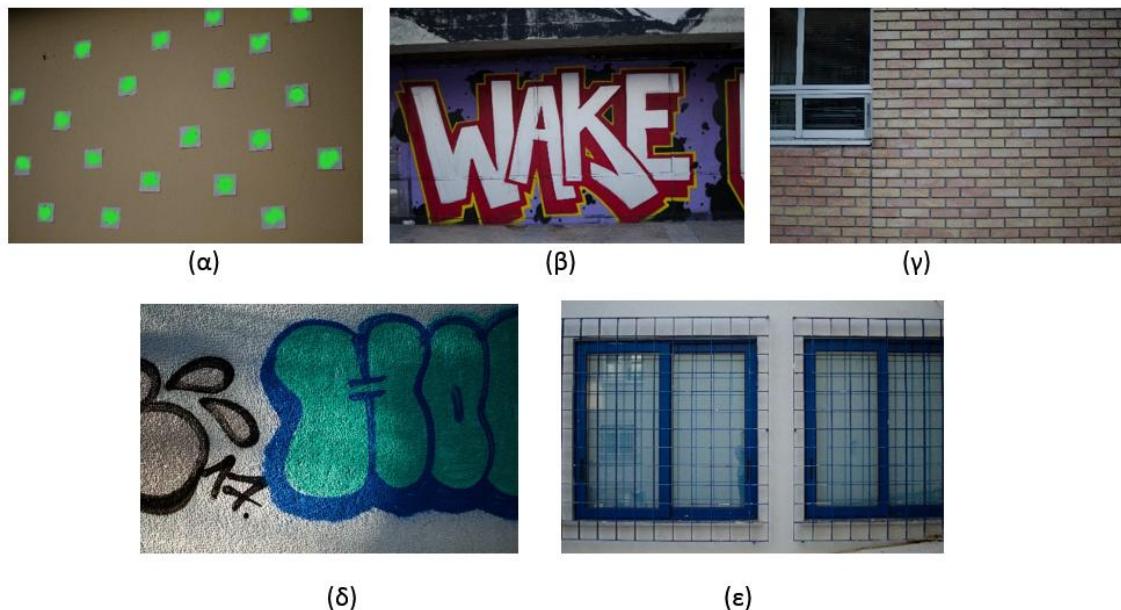


Εικόνα 19: Δείγμα θετικών εικόνων

Αντιθέτως, ως αρνητικές ορίζονται οι εικόνες που δεν περιέχουν την κατάλληλη πληροφορία, στην προκειμένη περίπτωση σβησμένοι στόχοι και επιφάνειες χωρίς στόχους, ώστε να εκπαιδευτεί ο αλγόριθμος να απορρίπτει τέτοιους είδους δεδομένα. Σε αυτές λοιπόν, απεικονίζονται οι ίδιες ή και διαφορετικές επιφάνειες με αυτές των θετικών εικόνων (Εικόνα 20 β,γ), εκ των οποίων και:

- επιφάνειες με στόχους, οι οποίοι έχουν υποστεί επεξεργασία επί της εικόνας μετά τη λήψη, ώστε να μη φαίνεται το σχήμα τους (Εικόνα 20 α),
- επιφάνειες χωρίς στόχους (Εικόνα 20 β,γ,δ,ε)

(Για περισσότερες εικόνες βλ. Παράρτημα B.)



Εικόνα 20: Δείγμα αρνητικών εικόνων

## 4.2 ΕΠΙΣΗΜΑΝΣΗ ΣΤΟΧΩΝ

Κατόπιν ολοκλήρωσης της προετοιμασίας των δεδομένων, αφετηρία της εργασίας ήταν η χρήση του προγράμματος Matlab και ειδικότερα της εφαρμογής Image labeler, η οποία μας επιτρέπει να επισημάνουμε θετικά δείγματα αντικειμένων (στην προκειμένη περίπτωση στόχων), σε μια ομάδα εικόνων. Ουσιαστικά «μαρκάρουμε» σε κάθε εικόνα με ένα πλαίσιο τετράγωνου ή ορθογώνιου σχήματος, κάθε αντικείμενο που επιθυμούμε να ονοματοδοτηθεί. Έτσι, ανά εικόνα προκύπτει μία λίστα με «ταμπέλες», οι οποίες περιγράφουν την ονομασία του εκάστοτε αντικειμένου (βλ. Παράρτημα Β, «Εικόνες Επισημασμένων Στόχων»). Ταυτόχρονα, το πρόγραμμα αυτό μας παρέχει την πληροφορία με τις εικονοσυντεταγμένες και το ακριβές μέγεθος του εκάστοτε στόχου, σε κάθε εικόνα.

Στη συγκεκριμένη περίπτωση, εισήχθησαν 208 θετικές εικόνες, οι οποίες κρίθηκαν κατάλληλες για να υποδειχθούν οι στόχοι. Οι εικόνες αυτές προέκυψαν από διάφορες αποστάσεις από την επιφάνεια των στόχων, με διαφορετικό φωτισμό, σε διαφορετικά περιβάλλοντα, ενώ αρκετές από εκείνες είναι έντονης προοπτικής. Απαραιτήτως όμως όλες οι εικόνες συμπεριελάμβαναν έστω έναν στόχο.

Έτσι, κατά το τέλος της επεξεργασίας των εικόνων εξήχθηκε ένας πίνακας διαστάσεων 208x20 (δηλαδή Εικόνες x Στόχοι), ενώ η πληροφορία κάθε κελιού έχει ως εξής:

$$k = (j, i, w, h), \text{όπου:}$$

- $k$  - τυχαία μεταβλητή του εκάστοτε κελιού,
- $j, i$  - οι συντεταγμένες του εικονοστοιχείου (στήλη, γραμμή)
- $w, h$  - το πλάτος και ύψος του στόχου.

Προφανώς, ο πίνακας συμπεριλαμβάνει και κελιά με μηδενική τιμή, καθώς σε αρκετές εικόνες εμφανίζονται μερικοί μόνο στόχοι και όχι όλοι. Ο πίνακας αυτός φέρει πληροφορία, απαραίτητη για τα επόμενα βήματα, όσον αφορά την εκπαίδευση του προγράμματος στην ημιαυτόματη αναγνώριση αυτών των στόχων σε άλλες, τυχαίες εικόνες.

### 4.3 ΔΗΜΙΟΥΡΓΙΑ ΚΩΔΙΚΑ

Αφότου ολοκληρώθηκε η συλλογή και κατηγοριοποίηση των εικόνων και η επισήμανση των στόχων, δημιουργήθηκε κώδικας στο προγραμματιστικό περιβάλλον του MATLAB, στην έκδοση R2013b. Ο κώδικας αυτός παρατίθεται αναλυτικά στο Παράρτημα Α.

Κύριος σκοπός ήταν η εκπαίδευση ενός ταξινομητή εικόνων, ειδικά στους ζητούμενους στόχους, ώστε εν συνεχεία να προκύπτει, η ημιαυτόματη εύρεση και αναγνώριση των στόχων σε τυχαίες εικόνες. Έγινε λοιπόν, χρήση του αλγόριθμου Viola-Jones, μέσω των δύο βασικών εντολών: `trainCascadeObjectDetector` και `vision.CascadeObjectDetector`.

Η εντολή «`trainCascadeObjectDetector`» παρέχει έναν ταξινομητή, ο οποίος εκπαιδεύεται σε συγκεκριμένες θετικές και αρνητικές εικόνες, ώστε να αναγνωρίζει μία κατηγορία αντικειμένων. Ο ταξινομητής αυτός αποτελείται από στάδια (ο αριθμός των οποίων ορίζεται από τον προγραμματιστή), όπου κάθε στάδιο είναι ένα σύνολο άλλων «αδύναμων» ταξινομητών (στη βιβλιογραφία συναντάται ο όρος «Ensemble»). Σε κάθε στάδιο λοιπόν, ο ταξινομητής ενισχύεται με τη μέθοδο «Boosting», η οποία του παρέχει υψηλή ακρίβεια αποτελεσμάτων, λαμβάνοντας ένα σταθμισμένο μέσο όρο των αποφάσεων από τους αδύναμους ταξινομητές.

Αναλυτικότερα, πάνω στην εκάστοτε εικόνα, γίνεται χρήση ενός «συρόμενου» παραθύρου, που σαρώνει ολόκληρη την εικόνα ώστε να εντοπίσει το στόχο. Κάθε στάδιο του ταξινομητή χαρακτηρίζει την τρέχουσα περιοχή ως θετική ή αρνητική, υποδεικνύοντας, εάν το παράθυρο περιέχει το αντικείμενο ενδιαφέροντος ή όχι. Εάν το αποτέλεσμα είναι αρνητικό, η ταξινόμηση της περιοχής έχει ολοκληρωθεί και το παράθυρο μετακινείται στην επόμενη θέση. Αντιθέτως, σε περίπτωση που το αποτέλεσμα είναι θετικό, ο ταξινομητής περνάει την περιοχή στο επόμενο στάδιο του. Ακλουθώντας αυτή τη διαδικασία, ο «ανιχνευτής» που έχει δημιουργηθεί, ώστε να εντοπίσει το αντικείμενο, αναφέρει ότι αυτό έχει βρεθεί, όταν στο τελικό στάδιο του ταξινομητή κριθεί ως θετική η συγκεκριμένη τοποθεσία της εικόνας.

Να σημειωθεί ότι κατά τη σάρωση αυτή, το μέγεθος του παραθύρου ποικίλλει για την ανίχνευση αντικειμένων σε διαφορετικές κλίμακες, αλλά ο λόγος των διαστάσεων του πλαισίου παραμένει σταθερός.

Η βασική ιδέα είναι να απορρίπτονται όσο το δυνατόν γρηγορότερα τα αρνητικά δείγματα, καθώς η υπόθεση είναι ότι στην πλειοψηφία των παραθύρων δεν περιέχεται το αντικείμενο ενδιαφέροντος. Με αυτόν το γνώμονα, τα θετικά δείγματα είναι σπάνια και αξίζει να αφιερώνεται χρόνος για την επαλήθευσή τους.

Γενικά, τα αποτελέσματα των δειγμάτων κατατάσσονται σε τρείς κατηγορίες:

- a) Τα πραγματικά θετικά δείγματα: όταν ένα θετικό δείγμα ταξινομείται ορθά.
- b) Τα ψευδώς θετικά δείγματα: όταν ένα αρνητικό δείγμα ταξινομείται λανθασμένα ως θετικό.
- c) Τα ψευδώς αρνητικά δείγματα: όταν ένα θετικό δείγμα ταξινομείται λανθασμένα ως αρνητικό.

Γίνεται εύκολα αντιληπτό, ότι στην τελευταία περίπτωση, η πληροφορία του θετικού δείγματος χάνεται, χωρίς να μπορεί αυτό να διορθωθεί σε επόμενο στάδιο, καθώς σταματά η ταξινόμηση. Για να λειτουργήσει λοιπόν ορθά ο ταξινομητής κρίνεται αναγκαίο σε κάθε στάδιο του να ορίζεται ένα χαμηλό ποσοστό ψευδώς αρνητικών δειγμάτων. Ωστόσο, το ποσοστό των ψευδώς θετικών δειγμάτων μπορεί να είναι υψηλό, καθώς μπορεί να διορθωθεί το σφάλμα κατά τα επόμενα στάδια του ταξινομητή.

Ο κώδικας αυτός λοιπόν, έχοντας ως δεδομένα τους στόχους που ανιχνεύονται, καθίσταται δυνατός να συγκρίνει έναν οποιοδήποτε (από τους είκοσι) στόχο, με μία δοθείσα εικόνα που περιέχει «νέους» στόχους και εμφανίζει το ποσοστό συσχέτισής τους. Έτσι, σε περίπτωση που βρεθεί ο συγκεκριμένος στόχος και στη νέα αυτή εικόνα, το ποσοστό συσχέτισης είναι μεγαλύτερο απ' ό, τι αυτό των υπολοίπων, πράγμα που υποδεικνύει ότι ο ζητούμενος στόχος έχει ανιχνευτεί επιτυχώς. Δηλαδή, γνωρίζουμε πλέον την «ταυτότητά» του.

#### 4.3.1 ΠΕΡΙΓΡΑΦΗ ΚΩΔΙΚΑ

Αρχικά, στο πρόγραμμα εισάγεται ο πίνακας που προέκυψε στο προηγούμενο στάδιο, από την εφαρμογή Image labeler και δημιουργείται βάσει αυτού, νέος πίνακας- στήλη που περιλαμβάνει τις πληροφορίες των στόχων, δηλαδή την προέλευση της εικόνας, τις συντεταγμένες και τις διαστάσεις κάθε μαρκαρισμένου στόχου.

Από αυτόν τον πίνακα που προκύπτει λοιπόν (βλ. Παράρτημα A, «labels\_oneTarget9.mat»), αντλούνται οι εικόνες ως θετικά δείγματα, ενώ για τις

αρνητικές εικόνες χρησιμοποιούνται απλά οι φωτογραφίες που δεν περιέχουν στόχους, ή έχουν υποστεί επεξεργασία ώστε να μη φαίνονται (όπως ορίζονται οι αρνητικές εικόνες στην υποενότητα 4.1.3).

Ως εκ τούτου, τα δεδομένα στα οποία ο αλγόριθμος εκπαιδεύτηκε στην εφαρμογή αυτή, ήταν 208 θετικές εικόνες, με 2.892 μαρκαρισμένους στόχους, καθώς και 90 αρνητικές εικόνες.

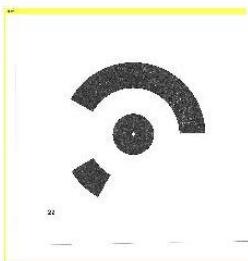
Με την εντολή `trainCascadeObjectDetector`, δημιουργήθηκε ο ταξινομητής με ονομασία «`TargetDetector.xml`», εισάγοντας ως δεδομένα τις θετικές και αρνητικές εικόνες και ορίζοντας δύο παραμέτρους:

- Τον αποδεκτό αριθμό ψευδών συναγερμών κάθε σταδίου της εκπαίδευσης, δηλαδή το κλάσμα αρνητικών δειγμάτων που έχουν ταξινομηθεί λανθασμένα ως θετικά δείγματα. Το εύρος τιμών που μπορεί να δοθεί είναι 0 έως 1 και στην προκειμένη επιλέχθηκε η τιμή 0.1.
- Τον αριθμό διαδοχικών σταδίων εκπαίδευσης. Με την αύξηση του αριθμού των σταδίων μπορεί να δημιουργηθεί ένας ανιχνευτής μεγαλύτερης ακρίβειας, αλλά αυξάνεται επίσης ο χρόνος της εκπαίδευσης του. Ενδεχομένως να χρειάζεται και μεγαλύτερος αριθμός εικόνων, καθώς σε κάθε στάδιο εξαλείφεται ένας αριθμός θετικών και αρνητικών δειγμάτων, όπως ορίζεται από τον αποδεκτό αριθμό ψευδών συναγερμών. (MathWorks) Ο αριθμός διαδοχικών σταδίων εκπαίδευσης που ορίστηκε στο πρόγραμμα ήταν τα πέντε (5) στάδια.

Δεδομένου λοιπόν του ταξινομητή αυτού, γίνεται χρήση του ανιχνευτή, μέσω της εντολής `vision.CascadeObjectDetector`, ώστε να εντοπιστεί ο στόχος σε μία εικόνα.

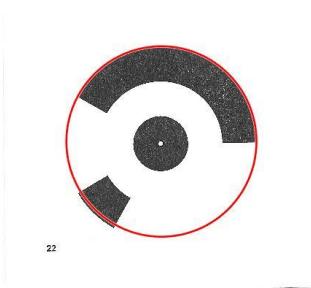
Στην προκειμένη περίπτωση δόθηκε εικόνα με το στόχο N.22. Αφού μετατράπηκε σε ασπρόμαυρη, ανιχνεύτηκε ο στόχος πάνω στην εικόνα.

Ως αποτέλεσμα, εμφανίζεται η δοθείσα εικόνα και επ' αυτής πλαισιώνεται ο ανιχνευμένος στόχος, περιβαλλόμενος από ένα προκαθορισμένου σχήματος πλαίσιο (εδώ παραλληλόγραμμο), συνοδευόμενο από αντίστοιχη λεζάντα («target»). (Εικόνα 21)



Εικόνα 21: Αρχική Εύρεση Στόχου 22 .

Στο επόμενο βήμα, δίδεται η εντολή για την εύρεση κύκλου στην εικόνα και τη σχεδίασή του, απ' όπου προκύπτει και η ακτίνα του, οι συντεταγμένες του κέντρου του στην εικόνα και οι διαστάσεις του. (Εικόνα 22) Βάσει αυτών των δεδομένων, γίνεται περικοπή του στόχου από την υπόλοιπη –αρχική- εικόνα. Με τον τρόπο αυτό εμφανίζεται μόνο το μοτίβο του στόχου, χωρίς το πλαίσιο που το περιβάλλει. (Εικόνα 23).



Εικόνα 22: Εύρεση κύκλου και ακτίνας στόχου

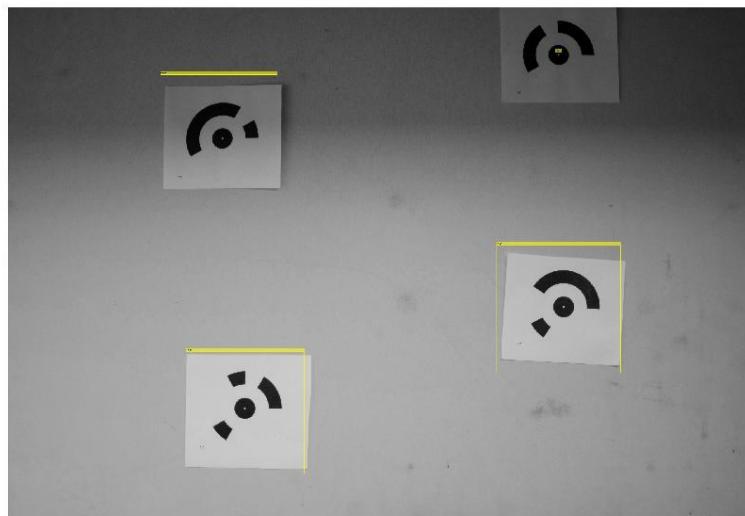


Εικόνα 23: Τελική μορφή στόχου

Στην συνέχεια, δίδεται μία «δοκιμαστική» εικόνα (Εικόνα 24), όπου ο αλγόριθμος καλείται να ανιχνεύσει τυχόν υπάρχοντες στόχους (Εικόνα 25), να τους σχεδιάσει και να τους αποκόψει, όπως ακριβώς λειτούργησε με το στόχο N22.



Εικόνα 24: Αρχική Εικόνα Τεσσάρων Στόχων



Εικόνα 25: Εντοπισμός στόχων.

Τέλος, συγκρίνει κάθε εικόνα από αυτές, με την εικόνα του δεδομένου στόχου (N.22), περιστρέφοντας την εικόνα του στόχου (N.22), και αλλάζοντας το μέγεθος της εκάστοτε εικόνας, ώστε να ταυτίζεται με αυτό του στόχου. Κατόπιν αυτού, κρατάει την καλύτερη συσχέτιση (από τις διαφορετικές γωνίες περιστροφής του στόχου) και ως αποτέλεσμα προβάλλει την εικόνα κάθε στόχου, συνοδευόμενη από μία λεζάντα που αναγράφει το ποσοστό της συσχέτισής τους (Εικόνα 26).



Εικόνα 26: Αποτελέσματα συσχέτισης στόχων (%).

Συνοψίζοντας λοιπόν, για να λειτουργήσει ο κώδικας αυτός, απαιτείται η εισαγωγή:

1. θετικών δειγμάτων εικόνων,
2. αρνητικών δειγμάτων εικόνων,
3. μίας εικόνας που περιέχει μόνο το ζητούμενο στόχο,
4. μίας εικόνας, στην οποία πρόκειται να εκτελεστεί η αναζήτηση του ζητούμενου στόχου.

#### 4.3.2 ΑΠΟΤΕΛΕΣΜΑΤΑ ΚΩΔΙΚΑ

Όπως αναφέρθηκε σε προηγούμενη υποενότητα, στην εφαρμογή που εκτελέστηκε ο ανωτέρω αλγόριθμος, έγινε αναζήτηση του στόχου N.22 σε εικόνα που εμπεριείχε, εκτός αυτού, τρείς ακόμη στόχους. Ως αποτέλεσμα της συσχέτισής τους, ο αντίστοιχος στόχος N.22 της δοθείσας εικόνας είχε συντελεστή συσχέτισης **0,90%**, ενώ οι άλλοι δύο έφτασαν στο **0,58%** και **0,71%** (Εικόνα 26).

Το ποσοστό που αντιστοιχεί στο ζητούμενο στόχο, δεν απέχει δραματικά από αυτό των άλλων στόχων, πράγμα που προκαλεί προβληματισμό, όσον αφορά την ορθότητα ανίχνευσης της ταυτότητας του στόχου.

Το αποτέλεσμα αυτό όμως, είναι αρκετά ικανοποιητικό, εάν σκεφτεί κανείς ότι η λήψη των αρχικών φωτογραφιών και η μετέπειτα επεξεργασία τους θα μπορούσε να αλλάξει άρδην το αποτέλεσμα του αλγόριθμου.

Να σημειωθεί ότι:

- Σε ορισμένες περιπτώσεις, η φωτογραφίες των στόχων είχαν ληφθεί επί σκοπού σε ανεπαρκώς φωτισμένο χώρο και σε φόντο παρόμοιο με το πλαίσιο των στόχων.

- Το δείγμα των εικόνων για την εκπαίδευση του αλγόριθμου ανερχόταν μόνο στις 298, ενώ για τη διαδικασία αυτή χρειάζονται- ιδανικά- χιλιάδες φωτογραφίες και ιδιαίτερα για τις αρνητικές εικόνες, όπως προτείνεται από τη Mathworks(MathWorks, 2013)
- Για την εκπαίδευση θα μπορούσαν να χρησιμοποιηθούν αρνητικές εικόνες μόνο με σβησμένους στόχους (παρόμοιας μορφολογίας με των κανονικών) και μόνο με φόντο παρόμοιο των θετικών, όπως προτείνεται από τη Mathworks (MathWorks)
- Θα μπορούσε να οριστεί ως αποδεκτός αριθμός ψευδών συναγερμών, κάποια τιμή μικρότερη από 0.1 και να αυξηθούν τα στάδια εκπαίδευσης.
- Θα μπορούσαν να χρησιμοποιηθούν τα φίλτρα Haar, LBP και HOG, σε συνδυασμό με την εκπαίδευση του ταξινομητή.

Στη συγκεκριμένη εφαρμογή ήταν σκόπιμη η αυστηρή επιλογή εικόνων και τα στάδια εκπαίδευσης σε συνάρτηση με τον αποδεκτό αριθμό ψευδών συναγερμών, ώστε να ελεγχθεί η αντοχή του αλγόριθμου. Στόχος ήταν επίσης, η επίτευξη σύντομων χρόνων λειτουργίας του προγράμματος, με τη σκέψη μελλοντικών εφαρμογών ανίχνευσης στόχων σε πραγματικό χρόνο και εφαρμογών με μεγάλο όγκο δεδομένων. Ως εκ τούτου δεν έγινε χρήση των φίλτρων Haar, LBP και HOG, καθώς η διαδικασία αυτή θα αύξανε εκθετικά το χρόνο περαιώσης της ανίχνευσης.

Υπάρχει η παραδοχή λοιπόν, για τη δυνατότητα μελλοντικής βελτίωσης του αλγόριθμου, τροποποιώντας τις ανωτέρω παραμέτρους. Κρίνεται αναγκαίο δε, να λαμβάνεται υπόψιν η εκάστοτε εφαρμογή στην οποία πρόκειται να γίνει χρήση του, καθώς σε κάθε περίπτωση αλλάζουν οι ανάγκες, με βάση τα αποτελέσματα που αποζητούνται. Για να αποδειχθούν όμως τα προαναφερθέντα, σίγουρα χρειάζονται περαιτέρω εφαρμογές πάνω στον συγκεκριμένο κώδικα.

## 5 ΣΥΜΠΕΡΑΣΜΑΤΑ

Συνοψίζοντας, η Τεχνητή Νοημοσύνη έχει αρχίσει να εισέρχεται στην καθημερινότητά μας, από απλές εφαρμογές (όπως η σάρωση κωδικοποιημένων στόχων για την παραπομπή μας σε κάποια ιστοσελίδα), έως τη χρήση λογισμικών για ιατρικές επεμβάσεις μέσω της ρομποτικής και την αυτόνομη οδήγηση οχημάτων.

Δεν άργησε όμως να εισέλθει και στο πεδίο της Φωτογραμμετρίας, καθώς ήδη η Μηχανική Όραση και η Μηχανική Μάθηση έχουν προσφέρει σε εφαρμογές της την αυτοματοποίηση εργασιών, όπως η αυτόματη δημιουργία τρισδιάστατων μοντέλων αντικειμένων, η κατασκευή Ψηφιακών Μοντέλων Εδάφους από ορθοφωτογραφίες, ο αυτόματος σχετικός προσανατολισμός και πολυεικονικός προσανατολισμός, η μέθοδος SFM (Structure from Motion) με την ανίχνευση και συνταύτιση χαρακτηριστικών σημείων και πολλά άλλα.

Στην παρούσα εργασία, έγινε μία πειραματική προσέγγιση για την αυτοματοποίηση της διαδικασίας ανίχνευσης και ταυτοποίησης κωδικοποιημένων στόχων σε τυχαίες εικόνες. Οι στόχοι αυτοί δύναται να χρησιμοποιηθούν είτε ως σημεία αναφοράς, είτε ως φωτοσταθερά σημεία στο χώρο, σε περαιτέρω φωτογραμμετρικές εφαρμογές.

Συγκεκριμένα, η δυνατότητα της αυτόματης αναγνώρισης των στόχων στις εικόνες, θα μπορούσε να οδηγήσει και στην αυτόματη (τρισδιάστατη) μοντελοποίηση αντικειμένων (π.χ. κτηρίων), στην αυτόματη παραγωγή ορθοφωτογραφιών και ψηφιακών μοντέλων εδάφους (DTM), κ.ά. φωτογραμμετρικές μεθόδους που ήδη αναφέρθηκαν και που κανονικά θα απαιτούσαν χειρωνακτική εργασία ή επίβλεψη.

Οι κωδικοποιημένοι στόχοι, διαφέρουν από τους κανονικούς φωτογραμμετρικούς στόχους, από την άποψη ότι φέρουν μία επιπλέον πληροφορία, αυτή της «ταυτότητάς» τους. Με αυτό τον τρόπο μπορούν να αποφευχθούν χονδροειδή σφάλματα, ενώ ταυτόχρονα καταργείται η χειροκίνητη εύρεση στόχων στις εικόνες, ως οιμόλογα σημεία.

Να σημειωθεί επίσης, ότι ενδεχομένως να υπήρχε η δυνατότητα ενσωμάτωσης του κώδικα αυτού (κατόπιν εξέλιξής του φυσικά) σε προγράμματα φωτογραμμετρικών εφαρμογών, όπως το AgisoftMetashape και το Drone2Map για το ArcGIS της Esri, όπου χρησιμοποιούνται φωτοσταθερά σημεία (αρχικά ως tie points κι έπειτα ως

GCPs), αλλά σχεδόν απαραίτητη είναι η χειροκίνητη διόρθωση σφαλμάτων επί των εικόνων από τον χρήστη. Καθ' αυτόν τον τρόπο, θα μπορούσε να αποφευχθεί ένα παραπάνω στάδιο εργασίας που, αδιαμφισβήτητα, καθιστά τη διαδικασία αρκετά χρονοβόρα, ενώ ταυτόχρονα θα μπορούσε η πληροφορία της ταυτότητας των στόχων, να διασφαλίζει τη σωστή εύρεσή τους.

Δημιουργήθηκε λοιπόν ένας κώδικας αυτόματης ανίχνευσης και αναγνώρισης είκοσι προκαθορισμένων, κωδικοποιημένων στόχων στο περιβάλλον του Matlab της MathWorks. Για να λειτουργήσει ο κώδικας αυτός, απαιτείται η προετοιμασία των δεδομένων που πρόκειται να λάβει. Χρησιμοποιήθηκαν λοιπόν 298 φωτογραφίες, από τις οι οποίες οι 208 εμπεριείχαν κωδικοποιημένους στόχους, που υποδείχθηκαν (μαζί με την ονομασία τους) σε καθεμιά από αυτές. Αυτές λοιπόν, αποτελούν τα λεγόμενα θετικά δείγματα εικόνων. Οι υπόλοιπες 90 εικόνες, που είτε δεν εμπεριείχαν καθόλου στόχους, είτε έγινε επεξεργασία τους ώστε να μην είναι ευκρινείς οι στόχοι, αποτελούν τα αρνητικά δείγματα εικόνων.

Με δεδομένα τα θετικά και αρνητικά δείγματα εικόνων, ο κώδικας εκπαιδεύει έναν ταξινομητή, ο οποίος στηρίζεται στη θεωρία των Viola και Jones και καθίσταται δυνατός να αναγνωρίζει σε μία εικόνα εάν η εκάστοτε περιοχή μέσα σε αυτήν εμπεριέχει κάποιο στόχο ή όχι. Έτσι, παρέχοντας μία εικόνα με έναν συγκεκριμένο στόχο της επιλογής του χρήστη, προκύπτει η ανίχνευσή του, ενώ μετέπειτα ο κώδικας μπορεί να εντοπίσει το στόχο αυτόν σε μία δεύτερη εικόνα, με τυχαίους στόχους. Στην εφαρμογή χρησιμοποιήθηκε ο στόχος με τον κωδικό N.22 και εντοπίστηκε σε τυχαία εικόνα που εμπεριείχε τρεις ακόμη στόχους, με επιτυχία 90%.

Ο αλγόριθμος που προέκυψε είναι σε θέση να εντοπίσει στόχους διαφορετικών μεγεθών (από την άποψη της διαφορετικής απόστασης λήψης των φωτογραφιών), σε διαφορετικά περιβάλλοντα, και επίσης υπό μεγάλες κλίσεις. Έτσι, η διαδικασία καθίσταται ανθεκτική στις αλλαγές της απόστασης της φωτογραφικής μηχανής από τους στόχους, από τη γωνία λήψης των εικόνων, καθώς και από τις διαφοροποιήσεις των χρωμάτων και της φωτεινότητας του εκάστοτε περιβάλλοντος. Ένα ακόμη προτέρημα αυτού του αλγόριθμου είναι η χρήση μη βαθμονομημένης φωτογραφικής μηχανής και μη συγκεκριμένου φωτογραφικού εξοπλισμού γενικότερα, καθώς αυτό συνήθως αποτελεί μία χρονοβόρα, αλλά και δαπανηρή διαδικασία.

Τα μειονεκτήματα του αλγόριθμου από την άλλη, είναι ότι για την καλύτερη δυνατή λειτουργία του, απαιτείται πολύ μεγάλο πλήθος εικόνων, καθώς και η σήμανση κάποιων από αυτές, με το ζητούμενο αντικείμενο. Αυτή η διαδικασία βεβαία, ακολουθείται μία μόνο φορά, σε περίπτωση που το ίδιο αντικείμενο πρόκειται να χρησιμοποιηθεί σε πολλές εφαρμογές, όπως για παράδειγμα αυτοί οι είκοσι στόχοι, πάνω στους οποίους έγινε η εκπαίδευση του ταξινομητή. Επίσης, άλλη μία ατέλεια του κώδικα αυτού, είναι ότι κατά τη λειτουργία του πρέπει ο χρήστης να παρέχει μία εικόνα που περιέχει μόνο τον ζητούμενο στόχο και μία εικόνα στην οποία πρόκειται να εκτελεστεί η αναζήτηση του στόχου αυτού. Τέλος, θεωρώ ότι σημαντική δυσκολία αποτελεί ο ορισμός του πλήθους των διαδοχικών σταδίων εκπαίδευσης του ταξινομητή και του αποδεκτού αριθμού ψευδών συναγερμών κάθε σταδίου από αυτά, (αν και παρέχονται ήδη προκαθορισμένες τιμές), καθώς αυτομάτως απαιτείται η γνώση του χειριστή στο συγκεκριμένο θεωρητικό πεδίο, ίσως και η εμπειρία πάνω στα προγράμματα αυτά, καθώς η ισορροπία των δύο ανωτέρων τιμών είναι σχετική, ανάλογα με την εκάστοτε εφαρμογή.

Συνεπώς, το αποτέλεσμα τα δημιουργίας του συγκεκριμένου κώδικα αφ' ενός, φέρει αρκετές δυσκολίες όσον αφορά το πρακτικό κομμάτι του, αφ' εταίρου δε, δίνει τη βάση για απλές φωτογραμμετρικές εφαρμογές, ενώ ταυτόχρονα εγείρει το ενδιαφέρον ως προς την εξέλιξή του σε ένα δυνατό λογισμικό, ευκολότερο στη χρήση και εφαρμόσιμο σε όλο και περισσότερες και πιο απαιτητικές εργασίες.

Κλείνοντας λοιπόν, εύχομαι η προσπάθεια που έγινε σε αυτήν την εργασία για την εξοικείωση με τα επιστημονικά πεδία της Τεχνητής Νοημοσύνης και της Φωτογραμμετρίας και τη σύμπραξη των μεθόδων και λειτουργιών τους για την πειραματική προσέγγιση που διατελέστηκε, να έφερε τα επιθυμητά αποτελέσματα στους αναγνώστες και να έθεσε ένα μικρό λιθαράκι στην εξέλιξη των εφαρμογών της Φωτογραμμετρίας.

## ΒΙΒΛΙΟΓΡΑΦΙΑ

- (n.d.). Ανάκτηση από <https://www.jstor.org/stable/resrep30590.7?seq=1>
- (n.d.). Ανάκτηση από <https://www.investopedia.com/terms/t/turing-test.asp>
- Avron Barr, E. A. (1981). The Handbook of Artificial Intelligence: Volume 1. In E. A. Avron Barr, *The Handbook of Artificial Intelligence: Volume 1*. (p. 408). William Kaufmann Inc.
- Dana H. Ballard, C. M. (1982). *Computer Vision*. Νέα Υόρκη: Prentice- Hall Inc.
- Elaine Rich, K. K. (1991). *Artificial Intelligence*. McGraw-Hill.
- Jubraj R., G. T. (2018). Ανάκτηση από <https://pdfcoffee.com/accents-redefine-banking-pdf-pdf-free.html>
- Learned-Miller, E. G. (2012). *Introduction to Computer Vision*. Μασαχουσέτη.
- MathWorks. (n.d.). Ανάκτηση από <https://www.mathworks.com/help/vision/ref/vision.cascadeobjectdetector-system-object.html>
- MathWorks. (2013). *MathWorks*. Ανάκτηση από <https://www.mathworks.com/help/vision/ref/traincascadeobjectdetector.html>
- Minsky, M. (1986). *The Society of Mind*. Νέα Υόρκη: Published by Simon & Schuster, Inc.
- Paul Viola, M. J. (2001). Rapid Object Detection using a Boosted Cascade of Simple Features.
- Ran Chen, K. Z. *An accurate and reliable circular coded target detection algorithm*.
- Ran Chen, K. Z. (2016). *An accurate and reliable circular coded target detection algorithm*.
- Sung JoonAhn, M. S. (2016). *A NEW CIRCULAR CODED TARGET FOR THE AUTOMATION OF PHOTOGAMMETRIC 3D-SURFACE MEASUREMENTS*.
- SungJoonAhn, W. R. (1999). *Circular Coded Landmark for Optical 3D-Measurement and Robot Vision*.
- Szeliski, R. (2022). *Computer Vision:Algorithms and Applications*.
- Vladimir A. Knyaz, A. V. (1998). on Contact 3D Model Reconstruction Using Coded Targets. Moscow, Russia.
- wikipedia. (n.d.). Ανάκτηση από [https://el.wikipedia.org/wiki/%CE%9A%CF%8E%CE%B4%CE%B9%CE%BA%CE%B1%CF%82\\_QR](https://el.wikipedia.org/wiki/%CE%9A%CF%8E%CE%B4%CE%B9%CE%BA%CE%B1%CF%82_QR)
- wikipedia. (n.d.). Ανάκτηση από [https://el.wikipedia.org/wiki/%CE%9C%CE%B7%CF%87%CE%B1%CE%BD%CE%B9%CE%BA%CE%AE\\_%CE%BC%CE%AC%CE%B8%CE%B7%CF%83%CE%B7](https://el.wikipedia.org/wiki/%CE%9C%CE%B7%CF%87%CE%B1%CE%BD%CE%B9%CE%BA%CE%AE_%CE%BC%CE%AC%CE%B8%CE%B7%CF%83%CE%B7)

*Wikipedia*. (n.d.). Ανάκτηση από [https://en.wikipedia.org/wiki/Dartmouth\\_workshop](https://en.wikipedia.org/wiki/Dartmouth_workshop)

Winston, P. H. (1976). *Artificial Intelligence*.

Βαρυτιμίδης, Χ. Ι. (2008). Ανίχνευση αντικειμένων και ημιαυτόματος.

Βίκτωρ, Μ. (2021). Ανάκτηση από

<https://polynoe.lib.uniwa.gr/xmlui/bitstream/handle/11400/364/ntst44.pdf?sequence=1&isAllowed=y>

ΓΕΩΡΓΙΟΥ, Χ. Ι. (2014). Αναγνώριση ταυτότητας προσώπου από βιντεοσκοπήσεις.

ΓΕΩΡΓΟΥΛΗ, Κ. (n.d.). Ανάκτηση από [http://repfiles.kallipos.gr/html\\_books/93/00e-introduction.html](http://repfiles.kallipos.gr/html_books/93/00e-introduction.html)

Κατερίνα, Γ. (n.d.). Ανάκτηση από [http://repfiles.kallipos.gr/html\\_books/93/00e-introduction.html](http://repfiles.kallipos.gr/html_books/93/00e-introduction.html)

## ΠΑΡΑΡΤΗΜΑ Α: ΚΩΔΙΚΑΣ

%ΠΡΟΕΤΟΙΜΑΣΙΑ ΔΕΔΟΜΕΝΩΝ

%Εκκαθάριση Επιφάνειας εργασίας του Matlab (Καρτέλα Command Window)

```
clear;
```

```
clc;
```

%Φόρτωσε τον πίνακα που προέκυψε από την εφαρμογή ImageLabeler  
%«labels9».

%Δημιουργησε ένα νέο πίνακα-στήλη που περιέχει κάθε στόχο (χωρίς κελιά), με  
%την ονομασία «targetBBs».

```
load('labels9.mat');  
fileNames =labels9{:,1};  
for i=1:size(labels9,1)  
    row = cell2mat(labels9{i,2:end});  
    targetBBs{i,1} = row;  
end
```

%Αποθήκευσε εκ νέου τα δεδομένα στον πίνακα labels9, αυτή τη φορά με την  
%πληροφορία «όνομα φακέλου εικόνας» και «στόχου» δηλ. προέλευση και  
%ονομασία κάθε εικόνας και συντεταγμένες-διαστάσεις κάθε στόχου.

```
labels9 = table(fileNames,targetBBs,'VariableNames',{ 'imageFilenames','targets'});
```

%Αποθήκευσε τον τελικό πίνακα, με την ονομασία «labels\_oneTarget9.mat»  
save('labels\_oneTarget9.mat','labels9')

%Φόρτωσε τον πίνακα που προέκυψε.

```
load('labels_oneTarget9.mat');
```

%Δημιουργησε νέο πίνακα με όνομα «positiveInstances», με όλες τις

μαρκαρισμένες- θετικές- εικόνες

```
positiveInstances = labels9(:,1:2);
```

%Προσδιόρισε τη διαδρομή για το φάκελο των θετικών εικόνων

```
imDir = fullfile('C:\','Users\','Elena\','Desktop\','tcd\','+p');
```

```
addpath(imDir);
```

%Προσδιόρισε τη διαδρομή για το φάκελο των αρνητικών εικόνων

```
negativeFolder = fullfile('C:\','Users\','Elena\','Desktop\','tcd\','-n');
```

%Δημιουργησε ένα αρχείο imageDatastore με ονομασία «negativeImages», το

οποίο θα περιέχει τις αρνητικές εικόνες (για εξοικονόμηση μνήμης).

```
negativeImages = imageDatastore(negativeFolder);
```

```

%ΕΚΠΑΙΔΕΥΣΗ ΤΑΞΙΝΟΜΗΤΗ
%Εκπαίδευσε έναν ταξινομητή ανιχνευτών αντικειμένων με ονομασία
%'TargetDetector.xml', με 0.1 αποδεκτό αριθμό ψευδών συναγερμών και 5
%στάδια εκπαίδευσης.
trainCascadeObjectDetector('TargetDetector.xml',positiveInstances, ...
    negativeFolder,'FalseAlarmRate',0.1,'NumCascadeStages',5);

%Χρησιμοποίησε το νέο -εκπαίδευμένο- ταξινομητή, ώστε ο ανιχνευτής να
εντοπίσει το στόχο σε μία εικόνα
detector = vision.CascadeObjectDetector('TargetDetector.xml');

%ΑΝΙΧΝΕΥΣΗ ΣΤΟΧΟΥ «22».
%Διάβασε την εικόνα με το δεδομένο στόχο (22) και μετάτρεψε την σε
%ασπρόμαυρη.
T22=imread('target22.jpg');
T22=rgb2gray(T22);

%Ανίχνευσε ένα στόχο στην εικόνα και δείξε τον.
bb2 = step(detector,T22);
img_labeled2 = insertObjectAnnotation(T22,'rectangle',bb2,'target');
figure
imshow(img_labeled2)

%Κόψε την εικόνα, ώστε να φαίνεται μόνο το πλαίσιο του στόχου
x2=bb2(1);
y2=bb2(2);
w2=bb2(3);
h2=bb2(4);
Ic2=T22(y2:y2+h2,x2:x2+w2,:);
figure
imshow(Ic2)

%Βρες τους κύκλους στην εικόνα.
r2=round((min(w2,h2))/2);
[ce,ra,me]=
imfindcircles(Ic2,[floor(r2/2),r2],'ObjectPolarity','dark','Sensitivity',0.99);

```

```

%Διάλεξε τον καταλληλότερο κύκλο και σχεδίασέ τον
[~,max_index2] = max(me);
hold on
viscircles(ce(max_index2,:),ra(max_index2));
hold off

%Κόψε την εικόνα με βάση τη διάμετρο του κύκλου και εμφάνισε την τελική
Ic2=Ic2((ce(2)-ra):(ce(2)+ra),(ce(1)-ra):(ce(1)+ra),:);
figure
imshow(Ic2)

% ΑΝΙΧΝΕΥΣΗ ΣΤΟΧΟΥ ΣΤΗΝ ΕΙΚΟΝΑ-ΤΕΣΤ
%Διάβασε και εμφάνισε τη νέα εικόνα για το τεστ
img = imread('test.jpg');
img=img(:,:,3);
imshow(img)

%Ανίχνευσε τους στόχους στην εικόνα και εμφάνισέ τους
bbox = step(detector,img);
img_labeled = insertObjectAnnotation(img,'rectangle',bbox,'target');
imshow(img_labeled)

%Κόψε την εικόνα, στα πλαίσια κάθε στόχου
rows=size(bbox,1);
figure
for i=1:rows
    bb=bbox(i,:);
    x=bb(1);
    y=bb(2);
    w=bb(3);
    h=bb(4);
    Ic=img(y:y+h,x:x+w,:);

    %Βρες τους κύκλους στην παραχθείσα εικόνα
    r=round((min(w,h))/2);
    [centers,radii,metric]=
    imfindcircles(Ic,[floor(r/2),r], 'ObjectPolarity','dark','Sensitivity',0.99);

```

```

%Διάλεξε τον καταλληλότερο
[~,max_index] = max(metric);

%hold on
%viscircles(centers(max_index,:),radii(max_index));
%hold off

%Κόψε την εικόνα βάσει της διαμέτρου του κύκλου
cee=centers;
raa=radii;
mee=metric;
Ic=Ic((cee(2)-raa):(cee(2)+raa),(cee(1)-raa):(cee(1)+raa),:);

%figure
%imshow(Ic)

%ΣΥΣΧΕΤΙΣΗ ΕΙΚΟΝΩΝ
%Bρες τη συγχέτιση για κάθε στόχο της εικόνας και εμφάνισε το ποσοστό της
%πάνω από τους στόχους
max_corr(i)=0;
for theta=(15:-0.5:-15)
    %Περίστρεψε την εικόνα του στόχου 22 και άλλαξε το μέγεθος της νέας εικόνας,
    %ώστε να συμπίπτει με αυτό του στόχου 22.
    %disp(theta)
    %B(i) = imrotate(Ic2,theta,'loose',bbox);
    B=imrotate(Ic2,theta);
    s=size(B);
    N=imresize(Ic,s);
    %Για κάθε περιστροφή σύγκρινε τις εικόνες και κράτα την καλύτερη συγχέτιση
    corr=corr2(B,N);
    if corr>max_corr(i)
        max_theta(i)=theta;
    end
    max_corr(i)=max(max_corr(i),corr);
end
%Για την καλύτερη συγχέτιση δείξε την εικόνα και το αντίστοιχο ποσοστό της.
subplot(1,rows,i)
imshow(N)
title(num2str(max_corr(i)))
end

```

## ΠΑΡΑΡΤΗΜΑΤΑ: ΕΙΚΟΝΕΣ ΕΚΠΑΙΔΕΥΣΗΣ ΑΛΓΟΡΙΘΜΟΥ

Θετικές Εικόνες:





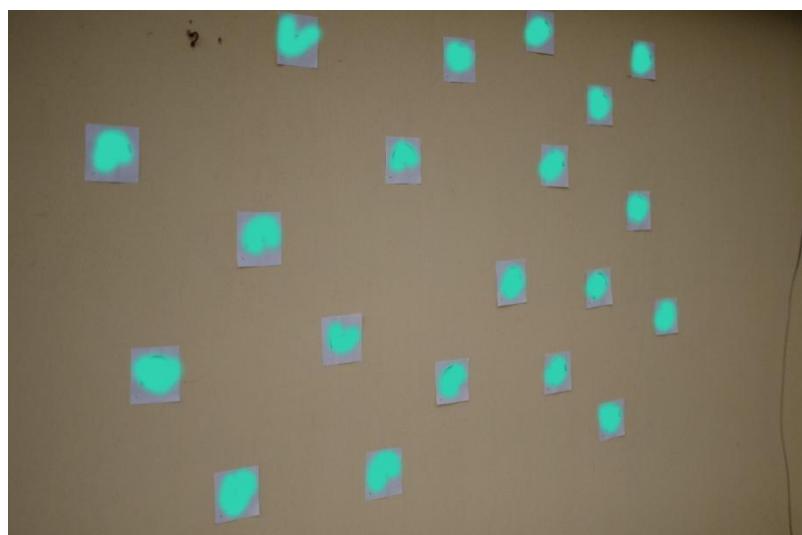


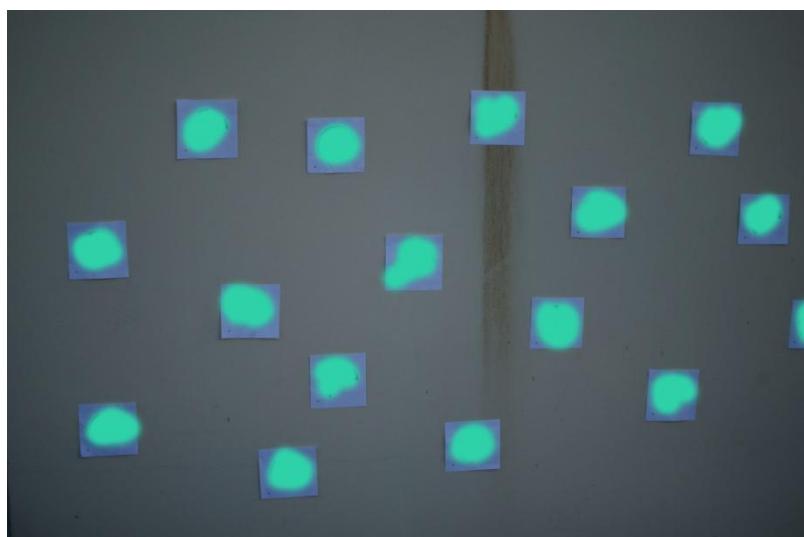
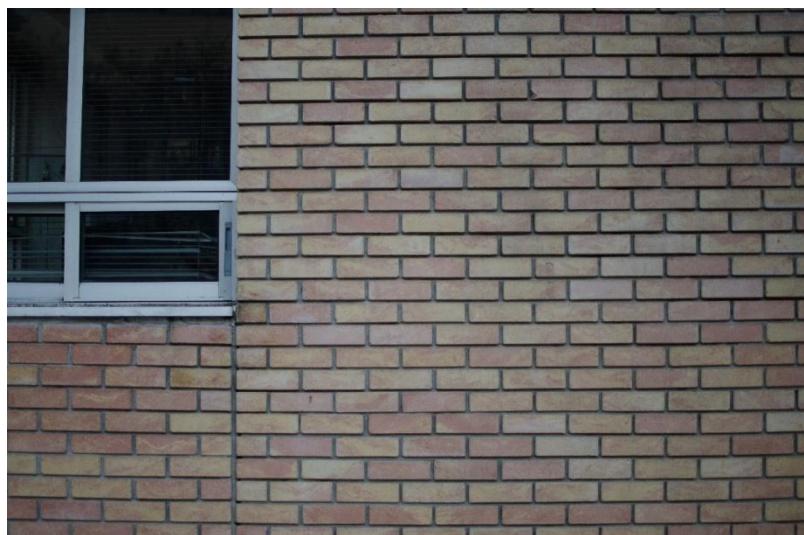


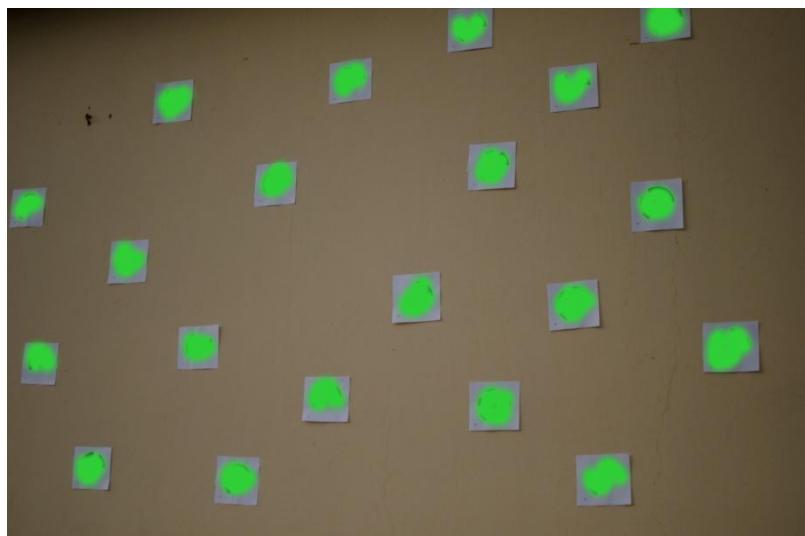




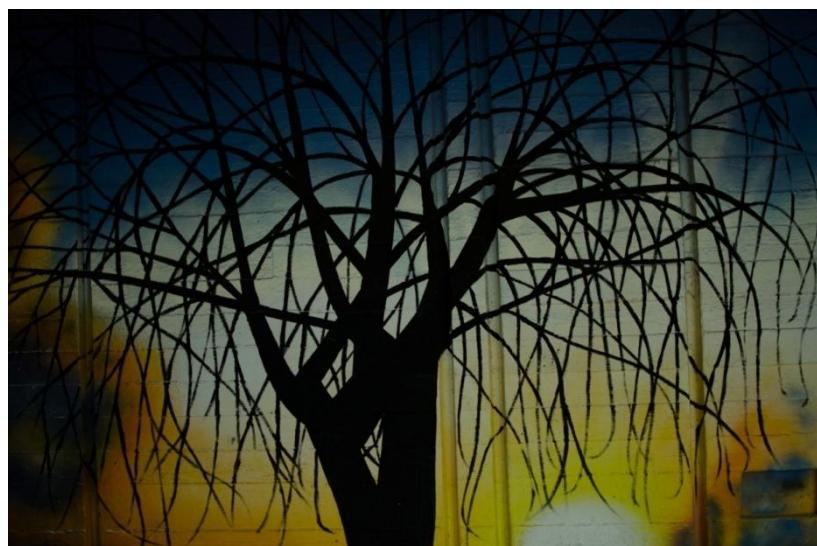
## Αρνητικές Εικόνες:





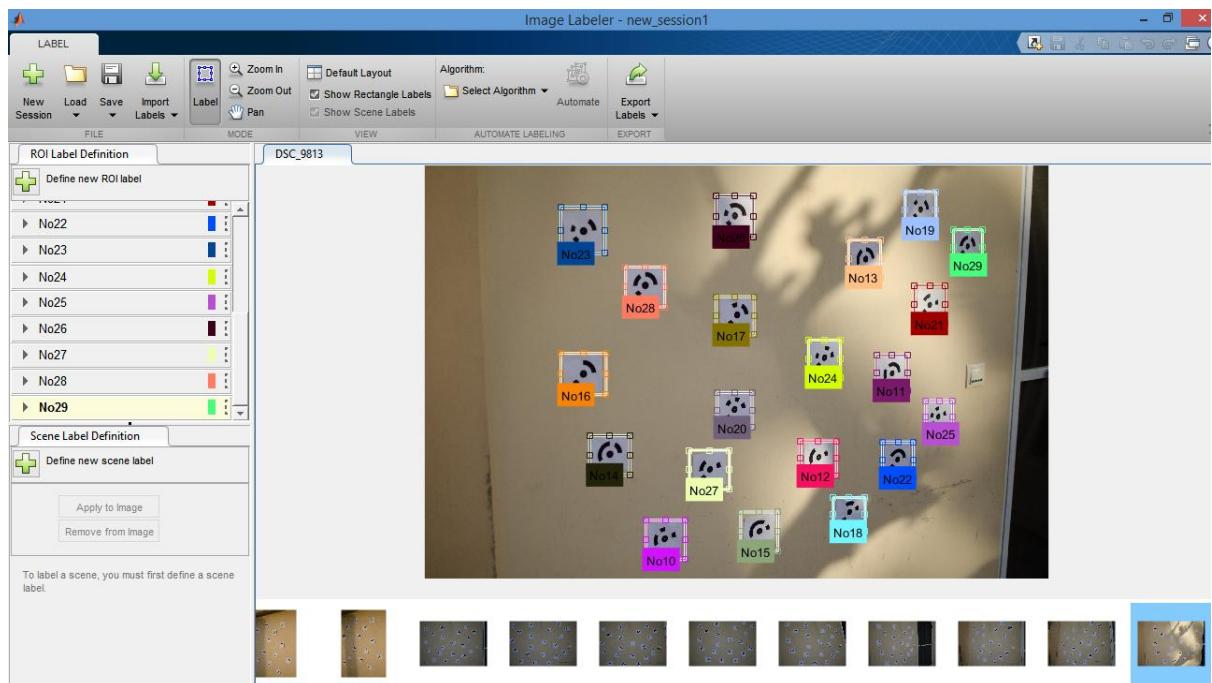
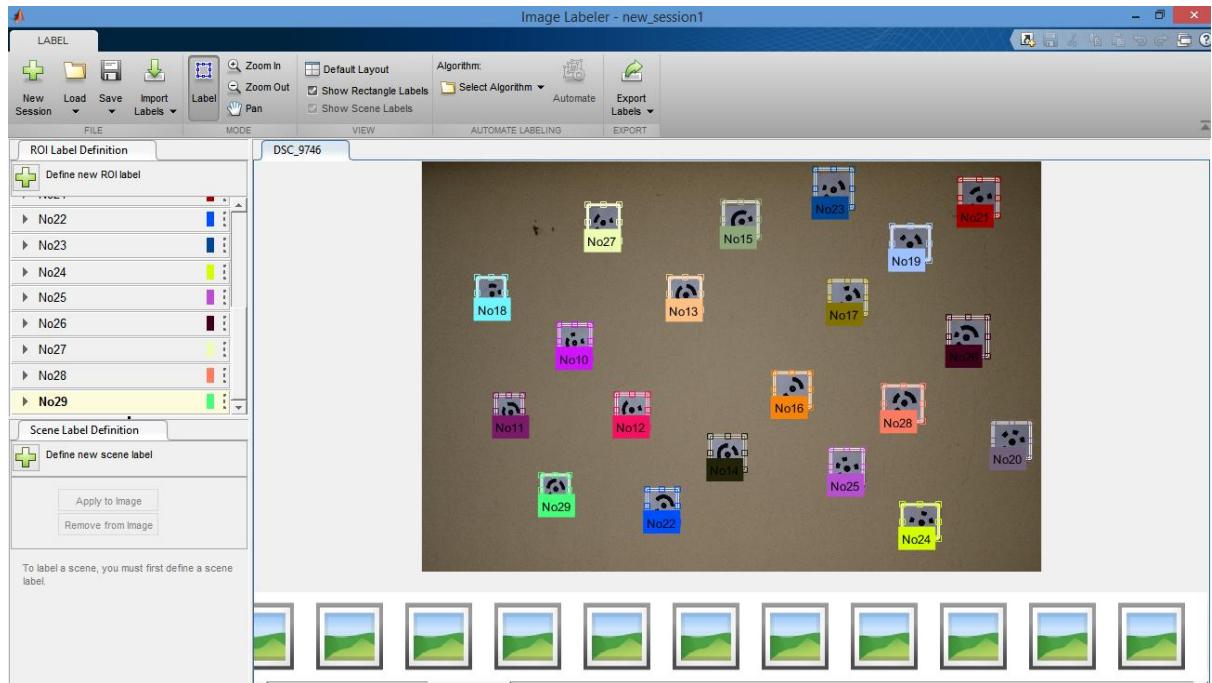


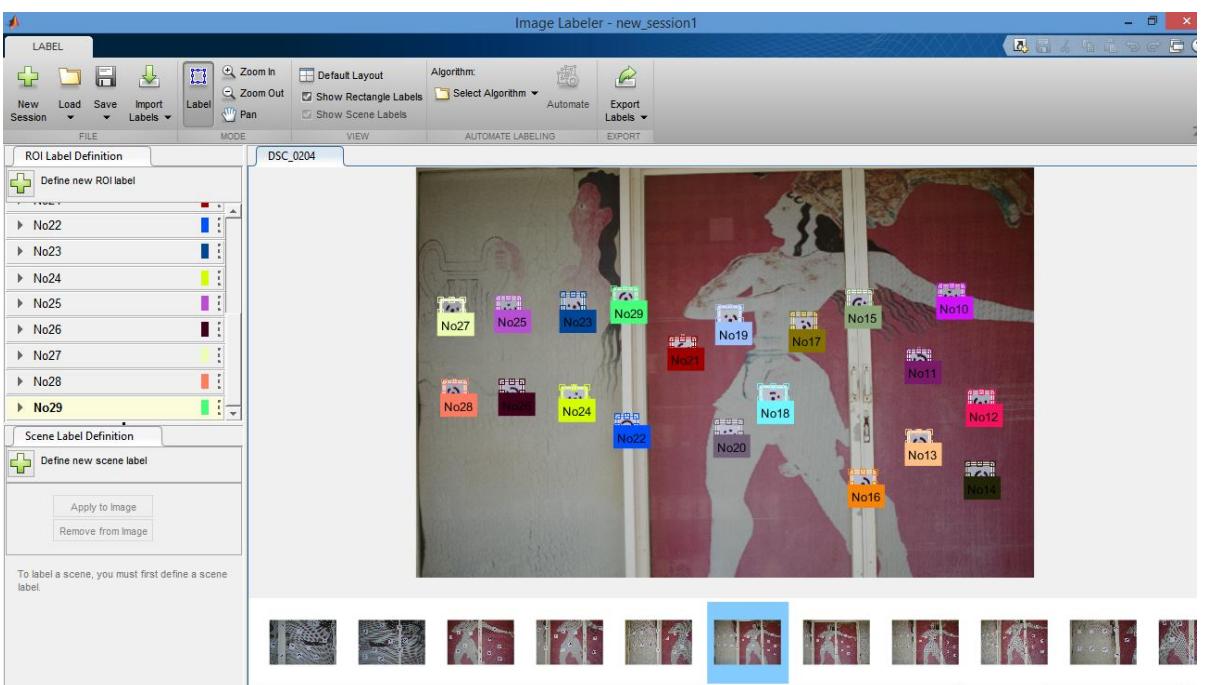
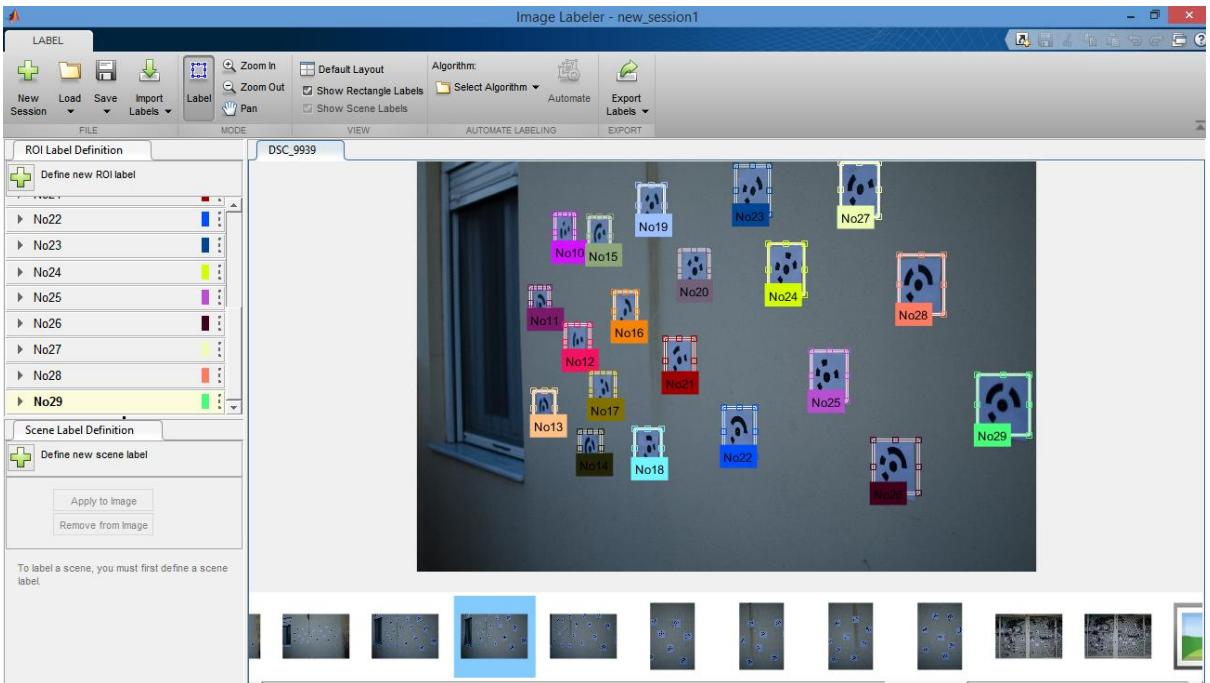


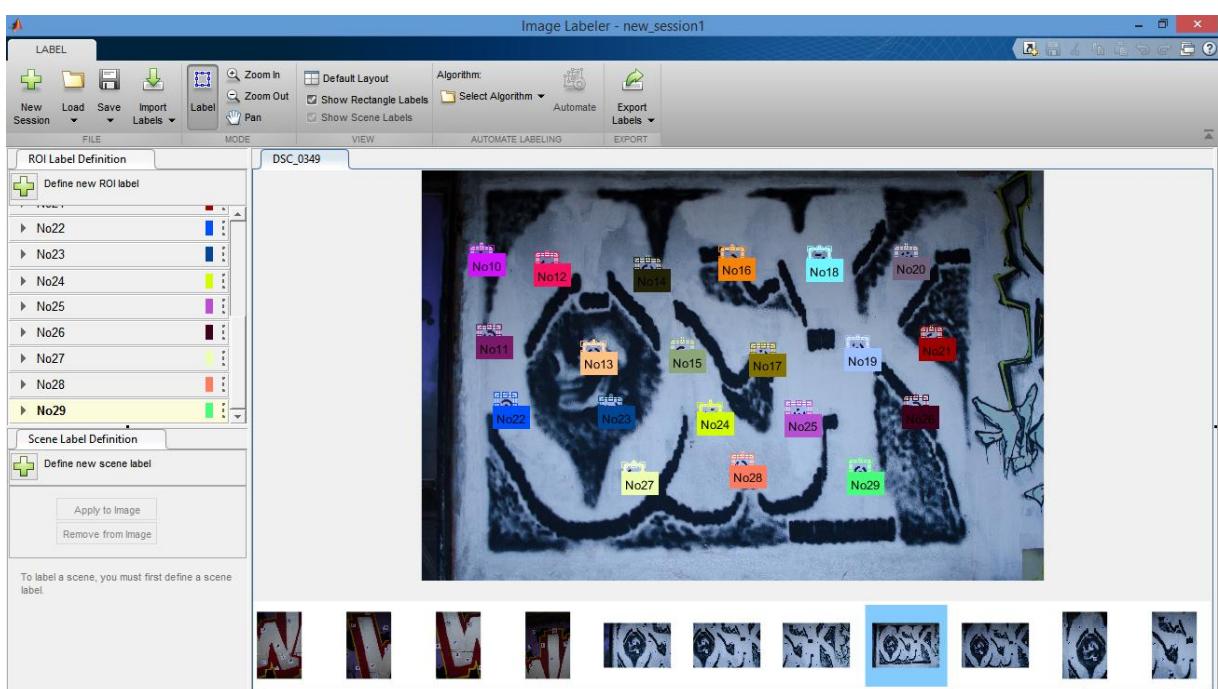
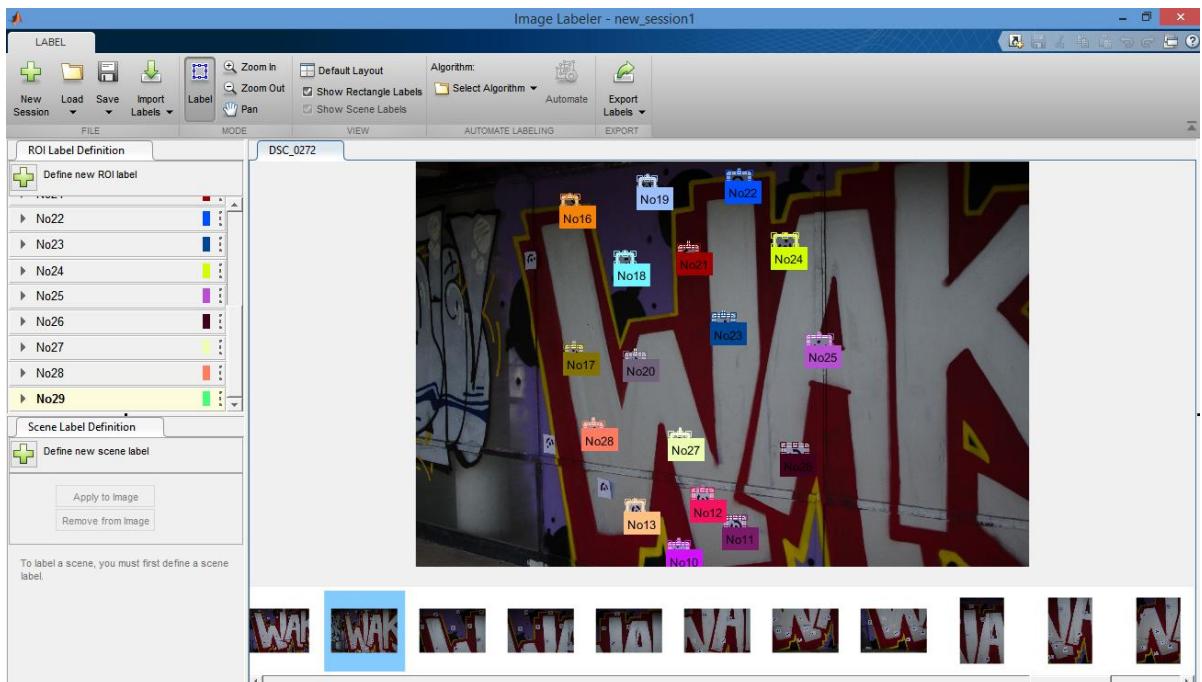




## Εικόνες Επισημασμένων Στόχων:







### **ΠΑΡΑΡΤΗΜΑ Γ: ΠΙΝΑΚΑΣ ΑΚΡΩΝΥΜΩΝ**

<b>AI</b>	<b>ARTIFICIAL INTELLIGENCE</b>
<b>TN</b>	<b>ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ</b>
<b>QR</b>	<b>QUICK RESPONSE</b>
<b>II</b>	<b>INTERNAL IMAGE</b>
<b>HOG</b>	<b>HISTOGRAM OF ORIENTED GRAPHS</b>
<b>GCP</b>	<b>GROUND CONTROL POINT</b>