



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΟΛΟΓΩΝ ΚΑΙ ΑΕΡΟΝΑΥΠΗΓΩΝ ΜΗΧΑΝΙΚΩΝ
[ΚΑΤΑΣΚΕΥΑΣΤΙΚΟΣ ΤΟΜΕΑΣ]
[ΕΡΓΑΣΤΗΡΙΟ ΜΗΧΑΝΟΥΡΓΙΚΗΣ ΤΕΧΝΟΛΟΓΙΑΣ ΚΑΙ
ΑΥΤΟΜΑΤΙΣΜΟΥ]

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ – ΕΛΛΗΝΙΚΗ ΠΕΡΙΛΙΨΗ

Ανάπτυξη μεθόδου δημιουργίας και χρήσης Συνθετικών Δεδομένων για εφαρμογές Τεχνητής Νοημοσύνης σε Συστήματα Παραγωγής – εφαρμογή στη οπτική αναγνώριση

Μανέττας Χρήστος

1054510

Παναγιώτης Σταυρόπουλος – Επίκουρος Καθηγητής

Διπλωματική εργασία υποβληθείσα στο Τμήμα Μηχανολόγων και Αεροναυπηγών του

Πανεπιστημίου Πατρών

Πανεπιστήμιο Πατρών, [Τμήμα]

[Χρήστος Μανέττας]

© [2022] – Με την επιφύλαξη παντός δικαιώματος



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ

ΤΜΗΜΑ ΜΗΧΑΝΟΛΟΓΩΝ ΚΑΙ ΑΕΡΟΝΑΥΠΗΓΩΝ ΜΗΧΑΝΙΚΩΝ

[ΚΑΤΑΣΚΕΥΑΣΤΙΚΟΣ ΤΟΜΕΑΣ]

**[ΕΡΓΑΣΤΗΡΙΟ ΜΗΧΑΝΟΥΡΓΙΚΗΣ ΤΕΧΝΟΛΟΓΙΑΣ ΚΑΙ
ΑΥΤΟΜΑΤΙΣΜΟΥ]**

Η παρούσα διπλωματική εργασία παρουσιάστηκε

από τον

Μανέττα Χρήστο

1054510

την [18/07/2022]

Η έγκριση της διπλωματικής εργασίας δεν υποδηλοί την αποδοχή των γνώμων του συγγραφέα.
Κατά τη συγγραφή τηρήθηκαν οι αρχές της ακαδημαϊκής δεοντολογίας.

ΠΙΝΑΚΕΣ

Πίνακας 1: Πλήρης οδηγός για τα συνθετικά σύνολα δεδομένων με σκοπό την εκπαίδευση	5
Πίνακας 2: Πλήρης οδηγός για τα πραγματικά σύνολα δεδομένων με σκοπό την επαλήθευση	5
Πίνακας 3: Ακρίβεια ταξινόμησης ανά μοντέλο για δεδομένα με άνω προσανατολισμό	5
Πίνακας 4: Ακρίβεια ταξινόμησης ανά μοντέλο για δεδομένα με κάτω προσανατολισμό.....	5
Πίνακας 5: Ακρίβεια ταξινόμησης ανά μοντέλο για δεδομένα με άνω και κάτω προσανατολισμό	5
Πίνακας 6: Παράμετροι εκπαίδευσης κατά την ανίχνευση χρώματος.....	7
Πίνακας 7: Μεταβλητές εκπαίδευσης.....	10
Πίνακας 8: Εκτιμώμενες και πραγματικές συντεταγμένες των μπουκαλιών.....	11
Πίνακας 9: Μεταβλητές εκπαίδευσης ελέγχου ποιότητας	13

EIKONES

Εικόνα 1: Βήματα συνθετικής παραγωγής εικόνων	2
Εικόνα 2: Τεχνική καλιμπραρίσματος κάμερας μέσω της βιβλιοθήκης OpenCV	2
Εικόνα 3: Μοντέλο pinhole	3
Εικόνα 4: Τοπικό και γενικό σύστημα συντεταγμένων (2)	3
Εικόνα 5: Δείγματα από διαφορετικές κατηγορίες και διαφορετικούς προσανατολισμούς (συνθετικές εικόνες με άνω και κάτω προσανατολισμό του εξαρτήματος)	4
Εικόνα 6: Δείγματα από διαφορετικές κατηγορίες και διαφορετικούς προσανατολισμούς (πραγματικές εικόνες με άνω και κάτω προσανατολισμό του εξαρτήματος)	4
Εικόνα 7: Πειραματική διάταξη ανίχνευσης χρώματος.....	6
Εικόνα 8: Δομή αρχείων κειμένου YOLO.....	7
Εικόνα 9: Δείγματα του συνθετικού σετ δεδομένων ανίχνευσης χρώματος.....	7
Εικόνα 10: Αποτελέσματα ανίχνευσης χρώματος	8
Εικόνα 11: Αναβαθμισμένη πειραματική διάταξη - αρχικά αποτελέσματα	8
Εικόνα 12: Αναβαθμισμένη πειραματική διάταξη - ενισχυμένη εικόνα.....	8
Εικόνα 13: Δείγμα του συνθετικού σετ δεδομένων	9
Εικόνα 14: Πειραματική διάταξη	10
Εικόνα 15: Παραδείγματα κακής ανίχνευσης.....	10
Εικόνα 16: Τελική πειραματική διάταξη	11
Εικόνα 17: Ιδανική θέση βραχίονα (επάνω) και η αντίστοιχη εκτιμώμενη θέση (κάτω)	12
Εικόνα 18: Δείγματα συνθετικών δεδομένων για την αναγνώριση χρώματος	12
Εικόνα 19: Κάμερα ανίχνευσης τύπου μπουκαλιών.....	13
Εικόνα 20: Παραδείγματα ανίχνευσης από τις δύο κάμερες του ελέγχου ποιότητας	14
Εικόνα 21: Φωτογραφία πριν και μετά τις βελτιώσεις	14

ΠΙΝΑΚΑΣ ΟΡΟΛΟΓΙΩΝ

MM – Μηχανική Μάθηση

BMM – Βαθιά Μηχανική Μάθηση

ConvNets – Νευρωνικά Δίκτυα Συνέλιξης

TNΔ – Τεχνητά Νευρωνικά Δίκτυα

CNN – Νευρωνικά Δίκτυα Συνέλιξης

RGB – Red Green Blue (είδος φίλτρων χρώματος μιας κάμερας)

ReLU – Rectified Linear Unit (είδος συνάρτησης μεταφοράς)

RoI – Region of Interest (περιοχή ενδιαφέροντος)

CAD – Computer-aided Design (Υπολογιστικά Αρχεία Σχεδιασμού)

et.al - Μεταξύ άλλων (λατινικά)

Tensorboard – Εφαρμογή της Keras

Xception - Μοντέλο Νευρωνικού Δικτύου Συνέλιξης

InceptionV3 - Μοντέλο Νευρωνικού Δικτύου Συνέλιξης

VGG16 – Μοντέλο Νευρωνικού Δικτύου Συνέλιξης

Keras – Βιβλιοθήκη μοντέλων Τεχνητής Νοημοσύνης

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΠΙΝΑΚΕΣ.....	III
ΕΙΚΟΝΕΣ.....	IV
ΠΙΝΑΚΑΣ ΟΡΟΛΟΓΙΩΝ.....	V
ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ.....	VI
1 ΕΙΣΑΓΩΓΗ ΚΑΙ ΑΝΑΣΚΟΠΗΣΗ ΒΙΒΛΙΟΓΡΑΦΙΑΣ.....	1
2 ΠΕΡΙΓΡΑΦΗ ΜΕΘΟΔΟΥ.....	1
3 ΕΡΕΥΝΗΤΙΚΑ ΠΕΙΡΑΜΑΤΑ.....	2
3.1 ΤΑΞΙΝΟΜΗΣΗ ΠΕΡΙΣΤΡΟΦΗΣ ΑΝΤΙΚΕΙΜΕΝΟΥ.....	3
3.2 ΑΝΑΓΝΩΡΙΣΗ ΧΡΩΜΑΤΟΣ.....	5
3.3 ΑΝΙΧΝΕΥΣΗ ΑΝΤΙΚΕΙΜΕΝΩΝ ΚΑΙ ΥΠΟΛΟΓΙΣΜΟΣ ΠΡΑΓΜΑΤΙΚΩΝ ΣΥΝΤΕΤΑΓΜΕΝΩΝ.....	9
3.4 ΕΛΕΓΧΟΣ ΠΟΙΟΤΗΤΑΣ	12
4 ΣΥΖΗΤΗΣΗ ΚΑΙ ΣΥΜΠΕΡΑΣΜΑΤΑ.....	14
5 ΑΝΑΦΟΡΕΣ.....	15

1. ΕΙΣΑΓΩΓΗ ΚΑΙ ΑΝΑΣΚΟΠΗΣΗ ΒΙΒΛΙΟΓΡΑΦΙΑΣ

Ο στόχος αυτής της μελέτης δεν είναι μόνο να δημιουργήσει ένα συνθετικό σύνολο δεδομένων ικανό να αντικαταστήσει πραγματικές εικόνες για εκπαίδευση, αλλά και να διερευνήσει τις βασικές πτυχές που επηρεάζουν τη δημιουργία εικόνων. Για ερευνητικές μελέτες όπως αυτή, υπάρχει αρκετός χρόνος για να δημιουργηθεί το σωστό σύνολο δεδομένων που απαιτείται. Ωστόσο, για να εφαρμοστούν καινοτομίες όπως αυτή σε βιομηχανικές εφαρμογές, η δημιουργία εικόνων πρέπει να είναι εύκολη, γρήγορη και ακριβής χωρίς να χρειάζεται μεγάλη τεχνογνωσία από τον χειριστή. Επομένως, η επισήμανση των παραγόντων που την επηρεάζουν περισσότερο είναι χρήσιμη.

Οι βιομηχανικές εφαρμογές, λαμβάνοντας υπόψη το πεδίο της αναγνώρισης εικόνας, απαιτούν τις περισσότερες φορές τεχνικές ταξινόμησης εικόνας και αλγόριθμους ανίχνευσης αντικειμένων. Η ταξινόμηση εικόνας, η οποία μπορεί να οριστεί ως το έργο της κατηγοριοποίησης των εικόνων σε μία από τις πολλές προκαθορισμένες κατηγορίες, είναι ένα θεμελιώδες πρόβλημα στην Μηχανική Όραση. Αποτελεί τη βάση για άλλες εφαρμογές οπτικής αναγνώρισης με υπολογιστή, όπως εντοπισμός, ανίχνευση και τμηματοποίηση (Rawat et al., 2017). Ωστόσο, για να αποκτήσουμε μια πλήρη κατανόηση της εικόνας, δεν πρέπει να επικεντρωθούμε μόνο στην ταξινόμηση διαφορετικών εικόνων, αλλά και να προσπαθήσουμε να εκτιμήσουμε με ακρίβεια την κλάση και την θέση των αντικειμένων που περιέχονται σε κάθε εικόνα. Αυτή η εργασία αναφέρεται ως ανίχνευση αντικειμένων (Zhao et al., 2019).

Η σύγχρονη λύση για την αντιμετώπιση αυτών των αναγκών είναι η Βαθεία Μηχανική Μάθηση (BMM). Η BMM είναι ένας νέος τομέας μηχανικής μάθησης που επιτρέπει την επεξεργασία δεδομένων σε πολλαπλά επίπεδα επεξεργασίας προς εξαιρετικά μη γραμμικές και πολύπλοκες αναπαραστάσεις χαρακτηριστικών. Κύρια συνεισφορά στον τομέα προσφέρουν οι εφαρμογές οπτικής αναγνώρισης και εφαρμογές επεξεργασίας γλώσσας (LeCun, Bengio, & Hinton, 2015).

Η Συνθετική Μεγέθυνση Συνόλου Δεδομένων είναι ένας αναβαθμισμένος τύπος αύξησης δεδομένων, έχοντας ένα σημαντικό πλεονέκτημα, οι πρόσθετες εικόνες είναι συνθετικά παραγμένες, μοιάζουν πολύ με τις πραγματικές εικόνες εισόδου και δεν αποτελούν μεταμορφωμένες εικόνες εισόδου (Mikolajczyk et al., 2018). Κατέχοντας αυτήν την τεχνογνωσία, μπορεί να δημιουργηθεί απεριόριστος αριθμός εικόνων.

Η χρήση μόνο συνθετικών δεδομένων για εκπαιδευτικούς σκοπούς στα CNN είναι μια σχετικά νέα και πολλά υποσχόμενη ιδέα. Ο K. Bousmalis (et al., 2018) και ο K. Fang (et al., 2018) έχουν ασχοληθεί με την οπτική αναγνώριση για την αρπαγή αντικειμένων από ρομποτικούς βραχίονες, με βάση τη συνθετική αύξηση δεδομένων. Η απόδοση του μοντέλου βελτιώθηκε και η ανάγκη για πραγματικά δεδομένα μειώθηκε δραματικά. Ο M. Andulkar (et al., 2018) εξέτασε την ανίχνευση αντικειμένων σε βιομηχανικά περιβάλλοντα χρησιμοποιώντας ConvNets αποκλειστικά εκπαιδευμένα σε συνθετικά δεδομένα. Οι δοκιμές σε πραγματικές, βιομηχανικές συνθήκες ήταν επιτυχημένες, αναφέροντας ότι η εφαρμογή των CNN σε βιομηχανικά περιβάλλοντα μπορεί να είναι ευπροσάρμοστη, εύκολη και αξιόπιστη.

2. ΠΕΡΙΓΡΑΦΗ ΜΕΘΟΔΟΥ

Σε αυτό το κεφάλαιο, δύο προτεινόμενες μέθοδοι παρουσιάζονται, αναφορικά με την δημιουργία συνθετικών δεδομένων και τον υπολογισμό πραγματικών συντεταγμένων. Στη συνέχεια, αυτές οι μέθοδοι θα εφαρμοστούν σε ερευνητικά πειράματα με σκοπό την διερεύνηση της αποτελεσματικότητας τους και των δυνατοτήτων τους σε εφαρμογές μηχανικής όρασης.

Η παραγωγή συνθετικών δεδομένων, όπως αναφέρεται επίσης στην ενότητα 1, αναφέρεται στη δημιουργία δεδομένων ή εικόνων από το μηδέν ή με βάση ορισμένα δείγματα δεδομένων. Η μέθοδος που προτείνεται σε αυτό το κεφάλαιο εστιάζει στη δημιουργία συνθετικών εικόνων από την αρχή, χωρίς να περιλαμβάνει πραγματικά δεδομένα για να βασιστεί. Η διαδικασία δημιουργίας εικόνας μπορεί να συνοψιστεί ως εξής: Επιλέγεται μια μηχανή γραφικών για να αναπαρασταθεί μια πραγματική ή μια φανταστική, επιθυμητή σκηνή. Μετά την προσαρμογή της σκηνής στις ανάγκες του εκάστοτε προβλήματος, η απόδοση και η δημιουργία αρχείων πραγματοποιείται σε συνδυασμό με αυτόματη επισήμανση και άλλους, επιπλέον βαθμούς ελευθερίας (αυτοματοποίηση μεγέθυνσης του συνόλου δεδομένων σύμφωνα με παραμετροποίηση κάποιων τιμών). Έτσι, δημιουργείται ένα ολοκαίνουργιο σύνολο δεδομένων. Τα βήματα αυτής της διαδικασίας μπροούν να αναπαρισταθούν σχηματικά στην εικόνα 1.



Εικόνα 1: Βήματα συνθετικής παραγωγής εικόνων

ΥΠΟΛΟΓΙΣΜΟΣ ΠΡΑΓΜΑΤΙΚΩΝ ΣΥΝΤΕΤΑΓΜΕΝΩΝ

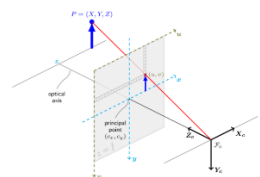
Ρύθμιση και βαθμονόμηση: Επιλέγεται μια σταθερή θέση για την κάμερα, καθώς η κάμερα πρέπει να είναι ακίνητη καθ' όλη τη διάρκεια της διαδικασίας. Αυτό μπορεί να ελεγχθεί εύκολα, αφού η κάμερα είναι τοποθετημένη στον ρομποτικό βραχίονα, επομένως αρκεί μόνο η αποθήκευση των επιθυμητών συντεταγμένων της θέσης βαθμονόμησης. Η παραμόρφωση του φακού μπορεί να αφαιρεθεί με την εκτίμηση των παραμέτρων του φακού και του αισθητήρα της κάμερας που είναι γνωστή ως βαθμονόμηση/καλιμπράρισμα κάμερας. Η βαθμονόμηση θα πραγματοποιηθεί σύμφωνα με τις συστάσεις της βιβλιοθήκης OpenCV, χρησιμοποιώντας την ανίχνευση γωνιών μιας σκακιέρας. Λαμβάνονται πενήντα εικόνες της σκακιέρας μπροστά από την κάμερα, με μεταβλητή θέση (Εικ.2).



Εικόνα 2: Τεχνική καλιμπραρίσματος κάμερας μέσω της βιβλιοθήκης OpenCV

Βαθμονόμηση προοπτικής: Επιλέγεται ένα σταθερό σημείο αναφοράς. Σε αυτήν την περίπτωση, αυτό το σημείο θα είναι η βάση του ρομποτικού βραχίονα. Στο ίδιο επίπεδο όπου στέκονται τα μπουκάλια, σημειώνονται 10 σημεία. Για κάθε σημείο, οι συντεταγμένες x και y μετρώνται σε χιλιοστά σε σχέση με το σταθερό σημείο αναφοράς. Στη συνέχεια μετράται η απόσταση από την κάμερα – το λεγόμενο βάθος. Επομένως, ένα διάνυσμα 1×3 προκύπτει για κάθε σημείο (x , y , βάθος). Αντίστοιχα, καταγράφονται οι τιμές των pixel για κάθε σημείο, όπως φαίνονται στην

κάμερα. Η σχέση, μεταξύ των τιμών των pixel και των συντεταγμένων του πραγματικού κόσμου, περιγράφεται από το μοντέλο κάμερας pinhole και την αντίστοιχη εξίσωσή του στην εικόνα 3.



$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Εικόνα 3: Μοντέλο pinhole

Όπου: x, y, z είναι οι πραγματικές συντεταγμένες ενός σημείου στο χώρο, u, v οι αντίστοιχες τιμές pixel του σημείου όπως φαίνονται στην κάμερα, $[r, t]$ το μητρώο 3×4 της κάμερας, c_x, c_y το κέντρο της εικόνας σε τιμές pixel, f_x, f_y τα μήκη εστίασης της κάμερας. Σε αυτό το σημείο αξίζει να σημειωθεί ότι η συντεταγμένη z των μπουκαλιών θα θεωρείται γνωστή και σταθερή, καθώς όλα τα μπουκάλια βρίσκονται στο ίδιο ύψος.

3. ΕΡΕΥΝΗΤΙΚΑ ΠΕΙΡΑΜΑΤΑ

Σε αυτό το κεφάλαιο οι μέθοδοι που περιγράφονται στο κεφάλαιο 3 εφαρμόζονται σε τέσσερις (4) διαφορετικές περιπτώσεις ερευνητικών πειραμάτων και συγκεκριμένα: α) ταξινόμηση περιστροφής αντικειμένου, β) αναγνώριση χρώματος γ) ανίχνευση αντικειμένου και υπολογισμός πραγματικών συντεταγμένων και δ) έλεγχος ποιότητας.

3.1. ΤΑΞΙΝΟΜΗΣΗ ΠΕΡΙΣΤΡΟΦΗΣ ΑΝΤΙΚΕΙΜΕΝΟΥ

Περιγραφή προβλήματος και δημιουργία συνθετικών δεδομένων

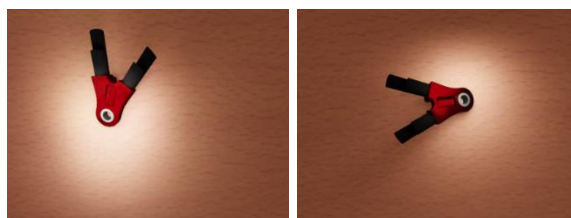
Η παρούσα μελέτη επιχειρεί να εξετάσει την εφαρμογή συνθετικών δεδομένων για εφαρμογές οπτικής αναγνώρισης. Θα εξεταστεί η περιστροφή γύρω από τον άξονα z (Εικ. 4) ενός κατασκευασμένου εξαρτήματος, με την ταξινόμηση να γίνεται σε 18 κατηγορίες (βήμα 20° για κάθε κατηγορία). Το κομμάτι θα εξεταστεί τόσο προς τον άνω όσο και προς τον κάτω προσανατολισμό. Το έργο Μηχανικής Μάθησης που μελετήθηκε θεωρείται εργασίας ταξινόμησης.



Εικόνα 4: Τοπικό και γενικό σύστημα συντεταγμένων (2)

Η διαδικασία δημιουργίας συνθετικών εικόνων είναι αυτή που περιγράφηκε στο κεφάλαιο 2.1. Για κάθε κατηγορία δημιουργούνται συνολικά 300, έγχρωμες εικόνες RGB. Η ίδια διαδικασία θα πραγματοποιηθεί και για τους δύο προσανατολισμούς του εξαρτήματος με τον ίδιο αριθμό εικόνων ανα κατηγορία. Η ανάλυση που επιλέχθηκε ήταν 900×675 . Παραδείγματα από διάφορες κατηγορίες γωνιών και προσανατολισμών φαίνονται στην εικόνα 5.





Εικόνα 5: Δείγματα από διαφορετικές κατηγορίες και διαφορετικούς προσανατολισμούς (συνθετικές εικόνες με άνω και κάτω προσανατολισμό του εξαρτήματος)

Πραγματικά δεδομένα για επαλήθευση

Έχοντας δημιουργήσει τα δεδομένα εκπαίδευσης, εικόνες από το πραγματικό εξάρτημα για κάθε κατηγορία ελήφθησαν χρησιμοποιώντας μια επαγγελματική κάμερα Nikon d5000. Η ανάλυση των εικόνων είναι 4000x3000 pixel. Ομοίως, η περιστροφή θα μετρηθεί γύρω από τον άξονα z με βήμα 20° ($0^\circ - 340^\circ$, συνολικά 18 κατηγορίες). Η περιστροφή γωνίας γίνεται και μετρείται από ένα κλασικό μοιρογνώνιο. Οι εικόνες λαμβάνονται παράλληλα πάνω από το τμήμα. Και οι δύο προσανατολισμοί θα εξεταστούν. Τα παραδείγματα φαίνονται στην εικόνα 6.



Εικόνα 6: Δείγματα από διαφορετικές κατηγορίες και διαφορετικούς προσανατολισμούς (πραγματικές εικόνες με άνω και κάτω προσανατολισμό του εξαρτήματος)

Ταξινόμηση μέσω εκμάθησης με μεταφορά και παράμετροι εκπαίδευσης

Πολλά μοντέλα Νευρωνικών Δικτύων Συνέλιξης προσφέρονται για ανίχνευση και ταξινόμηση αντικειμένων όπως τα ResNet, InceptionV3, VGG16, Xception, MobileNet και DenseNet121, τα οποία έχουν λάβει μέρος στην πρόκληση οπτικής αναγνώρισης μεγάλης κλίμακας ImageNet (ILSVRC). Επομένως, όλα τους είναι ήδη προ-εκπαιδευμένα στο σύνολο δεδομένων ImageNet που περιέχει περισσότερες από 1,2 εκατομμύρια φυσικές εικόνες και περισσότερες από 1.000 διαφορετικές κατηγορίες (O. Russakovsky et al., 2015). Σε αυτή τη μελέτη θα χρησιμοποιηθούν τα VGG16, MobileNet και DenseNet121, τα οποία είναι όλα διαθέσιμα μέσω της βιβλιοθήκης Keras Python API (Keras, 2021). Η επιλογή έγινε μετά τη δοκιμή πολλών άλλων μοντέλων, με βάση την υψηλότερη ακρίβεια που παρείχαν. Όλα τα μοντέλα θα εκπαιδευτούν με βάση το ίδιο σύνολο δεδομένων. Τα μεγέθη παρτίδων, ο αριθμός επαναλήψεων, οι βελτιστοποιητικοί παράγοντες και οι τιμές απώλειας ανα επανάληψη θα είναι πανομοιότυπες. Συγκεκριμένα, επιλέχθηκε ένα μέγεθος παρτίδας 32, κατηγορική συνάρτηση απώλειας, Adam ως βελτιστοποιητής και 6 επαναλήψεις για εκπαίδευση. Τα βάρη θα αρχικοποιηθούν από τη βιβλιοθήκη βαρών "ImageNet". Το μέγεθος της εικόνας θα είναι η προεπιλεγμένη επιλογή για κάθε μοντέλο. Σε αυτήν την περίπτωση, το προεπιλεγμένο μέγεθος και για τα τρία δίκτυα είναι (224,224,3). Έτσι, RGB, έγχρωμες εικόνες 3 καναλιών με μέγεθος (224.224) θα τροφοδοτηθούν στο δίκτυο.

Αποτελέσματα

Στο κύριο μέρος της παρούσας έρευνας, η εκπαίδευση και η αξιολόγηση πραγματοποιήθηκαν σε τρία σενάρια:

- Πρώτον, εκπαίδευση με συνθετικές εικόνες με προσανατολισμό προς τα πάνω και αξιολόγηση της ακρίβειας περιστροφής με πραγματικές εικόνες με προσανατολισμό προς τα πάνω.
- Δεύτερον, εκπαίδευση με συνθετικές εικόνες με κατεύθυνση προς τα κάτω και αξιολόγηση της ακρίβειας περιστροφής με πραγματικές εικόνες με κατεύθυνση προς τα κάτω .

Τέλος, εκπαίδευση τόσο με συνθετικές εικόνες με προσανατολισμό προς τα πάνω όσο και προς τα κάτω και αξιολόγηση με πραγματικές εικόνες μεικτού προσανατολισμού (πάνω και κάτω).

Πίνακας 1: Πλήρης οδηγός για τα συνθετικά σύνολα δεδομένων με σκοπό την εκπαίδευση

Σύνολο δεδομένων άνω προσανατολισμού	5400 εικόνες (300 ανά κατηγορία)
Σύνολο δεδομένων κάτω προσανατολισμού	5400 εικόνες (300 ανά κατηγορία)
Σύνολο δεδομένων μεικτού προσανατολισμού	10800 εικόνες (600 ανά κατηγορία), είναι το άθροισμα των δύο προηγούμενων συνόλων

Πίνακας 2: Πλήρης οδηγός για τα πραγματικά σύνολα δεδομένων με σκοπό την επαλήθευση

Σύνολο δεδομένων άνω προσανατολισμού	180 εικόνες (10 ανά κατηγορία)
Σύνολο δεδομένων κάτω προσανατολισμού	180 εικόνες (100 ανά κατηγορία)
Σύνολο δεδομένων μεικτού προσανατολισμού	360 εικόνες (20 ανά κατηγορία), είναι ουσιαστικά το άθροισμα των δύο προηγούμενων συνόλων

Ακολουθούν τα αποτελέσματα ακρίβειας κάθε μοντέλου βάσει των τριών σεναρίων που περιεγράφηκαν προηγουμένως:

Πίνακας 3: Ακρίβεια ταξινόμησης ανά μοντέλο για δεδομένα με άνω προσανατολισμό

Μοντέλο:	MobileNet	DenseNet121	VGG16
Ακρίβεια (%)	85.5	73.8	96.5

Πίνακας 4: Ακρίβεια ταξινόμησης ανά μοντέλο για δεδομένα με κάτω προσανατολισμό

Μοντέλο:	MobileNet	DenseNet121	VGG16
Ακρίβεια (%)	71.7	65.9	92.0

Πίνακας 5: Ακρίβεια ταξινόμησης ανά μοντέλο για δεδομένα με άνω και κάτω προσανατολισμό

Μοντέλο:	MobileNet	DenseNet121	VGG16
Ακρίβεια (%)	89.8	74.5	95.6

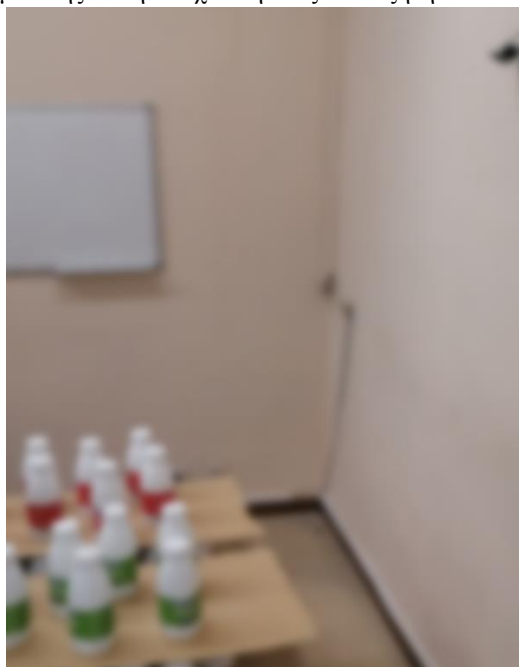
3.2. ΑΝΑΓΝΩΡΙΣΗ ΧΡΩΜΑΤΟΣ

Περιγραφή προβλήματος και πειραματική διάταξη

Στόχος του ερευνητικού πειράματος είναι η διερεύνηση της εφαρμογής συνθετικών συνόλων δεδομένων σε εργασίες αναγνώρισης αντικειμένων. Στις εργασίες Αναγνώρισης Αντικειμένων, εντοπίζονται μεμονωμένα Τμήμα Μηχανολόγων και Αεροναυπηγών Μηχανικών – [Κατασκευαστικός Τομέας]

αντικείμενα σε κάθε εικόνα, επομένως δεν χρησιμοποιείται η ταξινόμηση ολόκληρης της εικόνας μεταξύ κατηγοριών. Δεδομένου ότι το πρώτο μέρος της μελέτης κάλυψε τον τομέα της ταξινόμησης, έχει γίνει προσπάθεια σε αυτήν τη δεύτερη μελέτη να εξεταστεί η αποτελεσματικότητα των συνθετικών δεδομένων σε εργασίες αναγνώρισης αντικειμένων.

Το πρόβλημα που αντιμετωπίζεται σε αυτήν την μελέτη θεωρείται αρκετά απλό: Δύο σετ μπουκαλιών με διαφορετικές ετικέτες (το ένα με πράσινες και το άλλο με κόκκινες) τοποθετούνται δίπλα-δίπλα. Μία κάμερα θα είναι υπεύθυνη για να αναγνωρίσει ποια από τις δύο είναι η παλέτα με τα μπουκάλια με την πράσινη ετικέτα και ποια η κόκκινη. Μια αντίστοιχη εικόνα της διάταξης παρουσιάζεται στην Εικ. 7. Η κάμερα που χρησιμοποιείται θα είναι μια εμπορική κάμερα RGB 720p, καθώς δεν είναι απαραίτητη η υψηλή ανάλυση. Αυτή η εργασία θα περιλαμβάνει ένα μέρος της διαδικασίας της τρίτης μελέτης και η συσχέτισή τους θα συζητηθεί σε επόμενη ενότητα.



Εικόνα 7: Πειραματική διάταξη ανίχνευσης χρώματος

Δημιουργία συνθετικών εικόνων και παράμετροι εκπαίδευσης

Για να επιτευχθεί αυτό, δημιουργήθηκε ένα μικρό, συνθετικό σύνολο δεδομένων. Δημιουργήθηκαν 100 εικόνες και τα αντίστοιχα αρχεία κειμένου τους, ακολουθώντας την ίδια διαδικασία που περιγράφεται παρακάτω. Η επισημάνση στην ανίχνευση αντικειμένων χαρακτηρίζεται από δύο τιμές: κατηγορία και πλαίσιο οριοθέτησης. Αυτές οι τιμές συνήθως περιέχονται σε αρχείο κειμένου με το ίδιο όνομα με την εικόνα. Δεδομένου ότι το YOLO είναι το CNN που θα χρησιμοποιηθεί για εκπαίδευση και ανίχνευση, η δημιουργία αρχείων κειμένου θα ακολουθήσει τη μορφοποίηση YOLO. Το YOLO υποστηρίζει αρχεία .txt με την ακόλουθη μορφή (Εικ. 8)

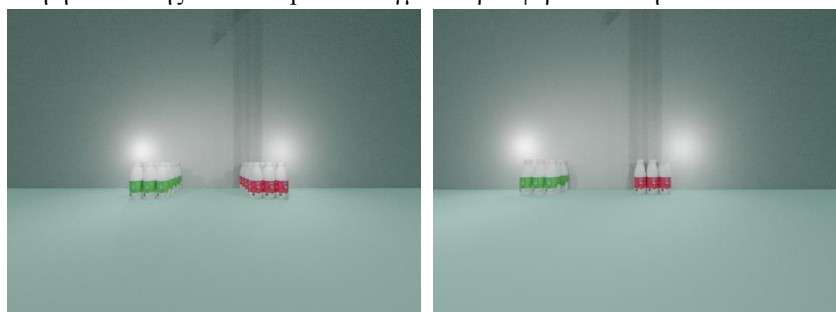
```

File Edit Format View Help
0 0.389524 0.510000 0.029524 0.080000
0 0.406667 0.548333 0.033333 0.095000
0 0.384762 0.548333 0.037143 0.095000
0 0.426667 0.548333 0.028571 0.095000
1 0.488571 0.548333 0.020000 0.095000
1 0.509524 0.548333 0.020000 0.095000
1 0.530476 0.548333 0.021905 0.095000
1 0.468571 0.548333 0.021905 0.095000

```

Εικόνα 8: Δομή αρχείων κειμένου YOLO

Η επιλεγμένη ανάλυση ήταν επίσης 900x675 pixel. Δείγματα προσφέρονται στην εικόνα 9.

**Εικόνα 9: Δείγματα του συνθετικού σετ δεδομένων ανίχνευσης χρώματος**

Οι παράμετροι εκπαίδευσης που χρησιμοποιήθηκαν προσφέρονται στον πίνακα 6.

Επιλεγμένα μοντέλα	Αριθμός συνόλων δεδομένων	Μέγεθος συνόλων δεδομένων και περιεχόμενο	Μέγεθος παρτίδας	Υπομέγεθος παρτίδας	Ανάλυση εικόνας	Διαχωρισμός δεδομένων εκπαίδευσης και επαλήθευσης (%)
YOLOv4	1	110 εικόνες με ανάλυση 900x675 χρησιμοποιήθηκαν για την ανίχνευση χρώματος.	64	16	480x480	90/10

Πίνακας 6: Παράμετροι εκπαίδευσης κατά την ανίχνευση χρώματος

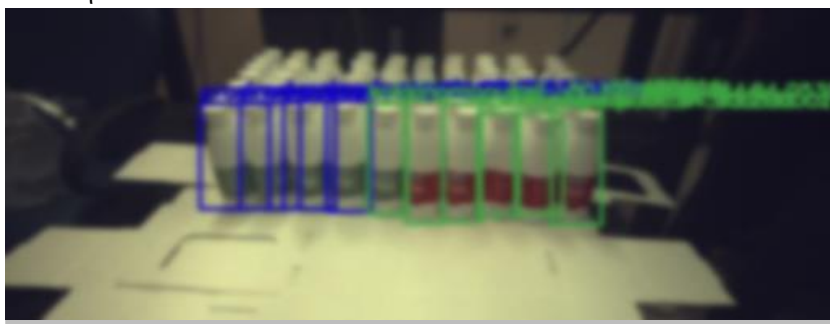
Αποτελέσματα

Η ανίχνευση χρώματος δεν διερευνήθηκε διεξοδικά καθώς η εργασία του αλγορίθμου δεν ήταν τόσο απαιτητική. Το επιθυμητό αποτέλεσμα του αλγόριθμου ήταν να δείξει ποια παλέτα περιέχει τα μπουκάλια με πράσινη ετικέτα και ποια τα κόκκινα. Για να επιτευχθεί αυτό, είναι απαραίτητος ο εντοπισμός τουλάχιστον μιας φιάλης από κάθε παλέτα (μία με κόκκινη ετικέτα και μία πράσινη). Αναγνωρίζοντας ότι η κάμερα είναι τοποθετημένη σχεδόν στο κέντρο της παλέτας, μια απλή σύγκριση μεταξύ των τιμών του κεντρικού pixel των δύο πλαισίων οριοθέτησης είναι υπεραρκετή για να προσδιοριστεί εάν η κόκκινη παλέτα είναι η δεξιά ή η αριστερή. Ένα δείγμα των αποτελεσμάτων ανίχνευσης προσφέρονται στην Εικ. 10.



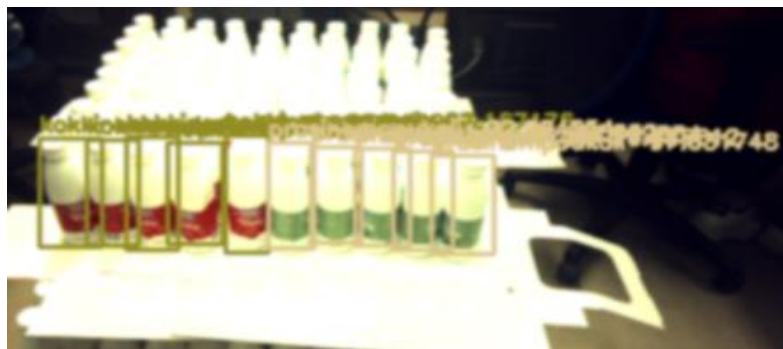
Εικόνα 10: Αποτελέσματα ανίχνευσης χρώματος

Σε μια προσπάθεια να διερευνηθεί η απόδοση του αλγορίθμου σε μια πιο απαιτητική εργασία, δημιουργήθηκε το ακόλουθο πείραμα. 10 μπουκάλια τοποθετήθηκαν το ένα δίπλα στο άλλο, με πολύ σφιχτή τοποθέτηση μεταξύ τους. Η κάμερα ρυθμίστηκε επίσης σε μεγαλύτερη απόσταση και σε συνδυασμό με διαφορετική γωνία, ενώ η κατεύθυνση της πηγής φωτισμού άλλαξε, με αποτέλεσμα ένα πιο σκοτεινό περιβάλλον. Μετά από συνδυασμό όλων των παραπάνω αλλαγών, δημιουργήθηκε ένα πολύ πιο απαιτητικό περιβάλλον. Τα αποτελέσματα ανίχνευσης παρουσιάζονται στην Εικ. 11.



Εικόνα 11: Αναβαθμισμένη πειραματική διάταξη - αρχικά αποτελέσματα

Αν και ο αλγόριθμος φαινόταν να αντιμετωπίζει επιτυχώς ακόμη και αυτό το πρόβλημα, όπως φαίνεται στην Εικ. 11, παρουσιάστηκε σφάλμα σε ένα στιγμιότυπο που λήφθηκε. Είναι προφανής μια εσφαλμένη ανίχνευση φιάλης με πράσινη ετικέτα.



Εικόνα 12: Αναβαθμισμένη πειραματική διάταξη - ενισχυμένη εικόνα

Ο φωτισμός του δωματίου με μια πρώτη ματιά φαίνεται να είναι ανεπαρκής και χαμηλός, ειδικά για έναν αλγόριθμο που αντιμετωπίζει ένα τόσο δύσκολο έργο με τόσα πολλά αντικείμενα σε μία εικόνα. Ως εκ τούτου, η βελτίωση του φωτισμού ήταν η πρώτη ενέργεια που έγινε. Τα αποτελέσματα βελτιώθηκαν αμέσως (Εικ. 12).

3.3. ΑΝΙΧΝΕΥΣΗ ΑΝΤΙΚΕΙΜΕΝΩΝ ΚΑΙ ΥΠΟΛΟΓΙΣΜΟΣ ΠΡΑΓΜΑΤΙΚΩΝ ΣΥΝΤΕΤΑΓΜΕΝΩΝ

Περιγραφή προβλήματος και δημιουργία συνθετικών εικόνων

Στόχος της τρίτης μελέτης είναι η διερεύνηση της εφαρμογής συνθετικών συνόλων δεδομένων σε εργασίες αναγνώρισης/ανίχνευσης αντικειμένων. Η αναγνώριση αντικειμένου θα προσφέρεται στη συνέχεια για μια εργασία "Pick and Place" σε συνεργασία με έναν αλγόριθμο υπολογισμού συντεταγμένων στον πραγματικό κόσμο. Τα αντικείμενα ενδιαφέροντος θα είναι δύο τύποι μπουκαλιών γάλακτος (ένα κόκκινο και ένα πράσινο). Όσον αφορά τη δημιουργία εικόνας, τα βήματα θα είναι ελαφρώς διαφορετικά από αυτά που περιγράφονται στην Εικ. 1 (ενότητα 2.1). Ειδικότερα:

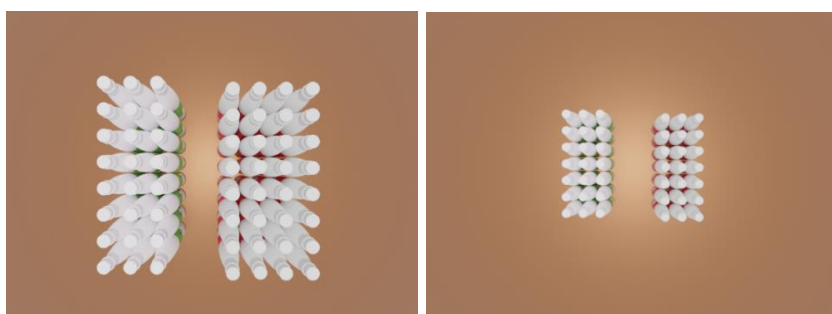
- Τα σχέδια CAD των φιαλών, δεν ήταν διαθέσιμα και, ως εκ τούτου, δημιουργήθηκαν χρησιμοποιώντας διαστάσεις που προσφέρει η εταιρεία. Το λογισμικό CATIA προσφέρθηκε για αυτήν την εργασία.
- Το CAD στη συνέχεια εισήχθη σε μια σκηνή του λογισμικού Blender. Οι ετικέτες πράσινων και κόκκινων φιαλών δόθηκαν ως εικόνες JPEG. Στη συνέχεια αναπτύχθηκε η χαρτογράφηση UV προκειμένου να εφαρμοστούν οι ετικέτες στα μπουκάλια.
- Επιλογή σωστού φωτισμός και θέσης κάμερας.
- Επιλογή περιβάλλοντα χώρου, τοποθέτηση και χρωματισμός. Το σκηνικό είναι πλέον έτοιμο.

Το σκηνικό θα αποτελείται από ποικίλο αριθμό μπουκαλιών, από 40 έως 60 στο σύνολο, χωρισμένα σε 2 παλέτες με βάση το χρώμα του μπουκαλιού. Η εικόνα 18 καθιστά το πρόβλημα πιο κατανοητό.

Η δημιουργία και η αύξηση των δεδομένων θα πραγματοποιηθεί με βάση τις ακόλουθες μεταβλητές: θέση των δύο σετ μπουκαλιών στον χώρο, γωνία κάμερας, ύψος κάμερας, φωτισμός. Το τελευταίο βήμα πριν την απόδοση της σκηνής και τη δημιουργία του συνόλου δεδομένων είναι η επισήμανση των μπουκαλιών. Η διαδικασία επισήμανσης ήταν πανομοιότυπη με αυτήν που παρουσιάστηκε στην ενότητα 3.2.1.

Δεδομένου ότι στόχος αυτής της μελέτης δεν είναι μόνο η ανίχνευση αλλά και ο πραγματικός υπολογισμός συντεταγμένων για τη συλλογή αντικειμένων, αξίζει να εξεταστεί ποια είναι τα σημαντικά σημεία ενός κουτιού οριοθέτησης για την επιλογή ενός μπουκαλιού. Η αρπαγή της φιάλης υποτίθεται ότι θα γίνει από το καπάκι της. Με άλλα λόγια, η ρομποτική λαβή θεωρείται πως κινείται ακριβώς πάνω από το μπουκάλι και εκτελεί μια γραμμική, προς τα κάτω κίνηση μέχρι να μπορέσει να πιάσει με ασφάλεια το μπουκάλι. **Για να επιτευχθεί αυτό, το κέντρο του πώματος της φιάλης είναι το κύριο σημείο ενδιαφέροντος για κάθε φιάλη.**

Η αλληλουχία δημιουργίας συνόλου δεδομένων είναι πλέον έτοιμη. 900x675 pixels είναι η ανάλυση που επιλέχθηκε. Δημιουργήθηκαν 200 εικόνες με μεταβλητή θέση, ύψος κάμερας, φωτισμό, αριθμό μπουκαλιών κ.λπ. Αντίστοιχα δημιουργήθηκαν 200 αρχεία κειμένου με το ίδιο όνομα. Δείγματα δεδομένων διαφόρων θέσεων, φωτισμού και αριθμού φιαλών παρέχονται στο σχήμα 13.



Εικόνα 13: Δείγμα του συνθετικού σετ δεδομένων

Πειραματική διάταξη και παράμετροι εκπαίδευσης

Η διάταξη αποτελείται από δύο κάμερες RGB και τα μπουκάλια που προσφέρονται από την εταιρεία. Το σκηνικό περιλαμβάνει επίσης έναν ρομποτικό βραχίονα της Universal Robots, ο οποίος θα είναι το σημείο αναφοράς μας στην εκτίμηση της θέσης και ένας πάγκος εργασίας όπου βρίσκεται το ρομπότ. Τα μπουκάλια θα στοιβάζονται

σε πολλαπλά στρώματα, πάνω σε μια εμπορική παλέτα. Η εικόνα 14 είναι αντιπροσωπευτική της διάταξης. Τα μπουκάλια θα χωριστούν επίσης σε δύο ξεχωριστές παλέτες με βάση το χρώμα τους. Η κύρια κάμερα που χρησιμοποιείται θα τοποθετηθεί στο τελικό σημείο του ρομποτικού βραχίονα και θα είναι υπεύθυνη για την ανίχνευση φιάλης και τον υπολογισμό των συντεταγμένων (Εικ.14). Μια συμπληρωματική κάμερα RGB έχει τοποθετηθεί στον τοίχο, ανιχνεύοντας έτσι το χρώμα κάθε παλέτας που δεν είναι αναγνωρίσιμο από την κάμερα που είναι τοποθετημένη πάνω στο ρομπότ. Η κάμερα αυτή είναι η ίδια με την προηγούμενη πειραματική έρευνα (3.2).



Εικόνα 14: Πειραματική διάταξη

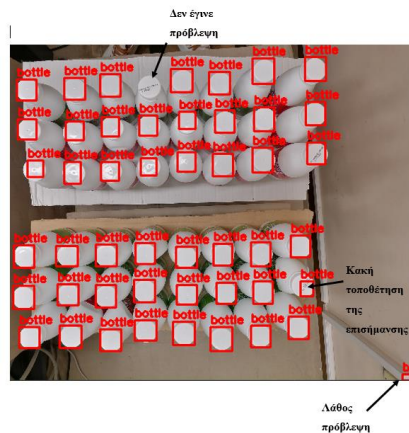
Το μοντέλο που επιλέχθηκε για την εκπαίδευση και ανίχνευση των αντικειμένων, ήταν το Yolov4. Στον παρακάτω πίνακα παρουσιάζονται οι κυριότερες μεταβλητές που αφορούν την εκπαίδευση του μοντέλου.

Πίνακας 7: Μεταβλητές εκπαίδευσης

Επιλεγμένα μοντέλα	Αριθμός συνόλων δεδομένων	Μέγεθος συνόλων δεδομένων και περιεχόμενο	Μέγεθος παρτίδας	Ανάλυση εικόνας	Διαχωρισμός δεδομένων εκπαίδευσης και επαλήθευσης (%)
YOLOv4	1	220 εικόνες με ανάλυση 900x675 χρησιμοποιήθηκαν για την ανίχνευση αντικειμένων.	64	480x480	90/10

Αποτελέσματα

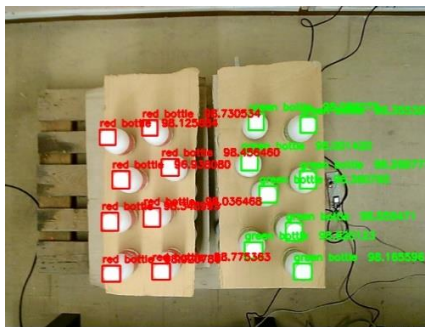
Εμφανίστηκαν διάφορα είδη αστοχιών ανά διάφορες λήψεις που έγιναν και με διάφορες κάμερες.



Εικόνα 15: Παραδείγματα κακής ανίχνευσης

Το σενάριο που απεικονίστηκε προηγουμένως παρουσιάζει πιθανές αστοχίες τις οποίες μπορεί να εμφανίσει ο αλγόριθμος. Το μέγεθος εικόνας, η τοποθέτηση της κάμερας και ο φωτισμός άλλαξαν προκειμένου να δημιουργήσουμε αυτό το πιο απαιτητικό περιβάλλον. Συγκεκριμένα η παραπάνω φωτογραφία αιχμαλωτίστηκε με ανάλυση 4000x3000 pixel. Για να εισαχθεί στον αλγόριθμο απαιτήθηκε σμίκρυνση περίπου 10 φορές από το αρχικό της μέγεθος, με αποτέλεσμα να καταστραφεί χρήσιμη πληροφορία για τον αλγόριθμο και επομένως η απόδοση του

να είναι πολλή χειρότερη. Με χρήση μικρότερης ανάλυσης για την αιχμαλώτιση της φωτογραφίας (640x480 pixel), τα αποτελέσματα ήταν αψεγάδιαστα.



Εικόνα 16: Τελική πειραματική διάταξη

Για να εξεταστεί η ακρίβεια του αλγόριθμου υπολογισμού συντεταγμένων, επιλέχθηκε μικρότερος αριθμός μπουκαλιών. Συγκεκριμένα, 16 μπουκάλια (8 κόκκινα και 8 πράσινα) τοποθετήθηκαν σε τυχαίες θέσεις πάνω από την παλέτα όπως φαίνεται στην εικόνα 16. Για κάθε μπουκάλι, υπολογίζεται το κέντρο του πλαισίου επισήμανσης και οδηγείται στον αλγόριθμο υπολογισμού. Οι συντεταγμένες που προκύπτουν από τον αλγόριθμο αποτυπώνονται στον Πίνακα 8.

Πίνακας 8: Εκτιμώμενες και πραγματικές συντεταγμένες των μπουκαλιών

Αριθμός μπουκαλιού και τύπος	Εκτιμώμενη θέση (mm)	Πραγματική θέση (mm)	Απόκλιση συντεταγμένων (mm)	Συνολική απόκλιση (mm)	Αρπαγή αντικειμένου
1 - Κόκκινο	-835, -369	-840, -363	5, 6	7.8	Επιτυχημένη
2 - Κόκκινο	-813, -465	-819, -466	6, 1	6.1	Επιτυχημένη
3 - Κόκκινο	-716, -433	-722, -435	6, 2	6.3	Επιτυχημένη
4 - Κόκκινο	-745, -320	-748, -316	3, 4	5	Επιτυχημένη
5 - Κόκκινο	-630, -360	-634, -360	4, 0	4	Επιτυχημένη
6 - Κόκκινο	-620, -457	-625, -462	5, 5	7.1	Επιτυχημένη
7 - Κόκκινο	-501, -451	-497, -457	4, 6	7.2	Επιτυχημένη
8 - Κόκκινο	-506, -337	-508, -333	2, 4	4.5	Επιτυχημένη
9 - Πράσινο	-558, -125	-560, -120	2, 5	5.4	Επιτυχημένη
10 - Πράσινο	-509, -10	-509, -3	0, 7	7	Επιτυχημένη
11 - Πράσινο	-600, -32	-600, -23	0, 9	9	Επιτυχημένη
12 - Πράσινο	-720, -2	-720, 7	0, 9	9	Επιτυχημένη
13 - Πράσινο	-686, -93	-691, -87	5, 6	7.8	Επιτυχημένη
14 - Πράσινο	-761, -138	-758, -129	3, 9	9.4	Επιτυχημένη
15 - Πράσινο	-851, -124	-856, -112	5, 12	13	Επιτυχημένη
16 - Πράσινο	-845, -4	-850, 9	5, 13	13.9	Επιτυχημένη

Αντίστοιχα, προσφέρονται μερικές εικόνες (Εικ. 17) τόσο της εκτιμώμενης όσο και της ιδανικής τοποθέτησης της αρπάγης του ρομποτικού βραχίονα.



Πρόβλεψη θέσης
vs
Πραγματική θέση

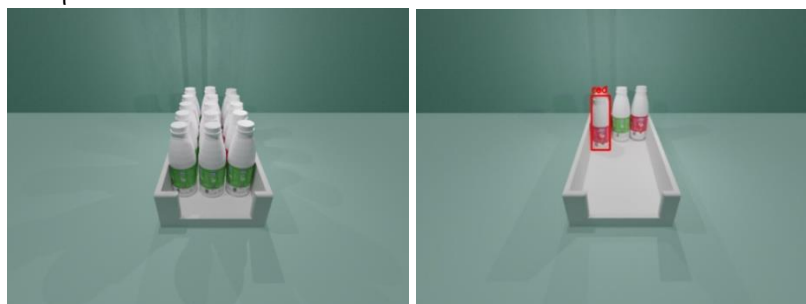
Εικόνα 17: Ιδανική θέση βραχίονα (επάνω) και η αντίστοιχη εκτιμώμενη θέση (κάτω)

3.4. ΕΛΕΓΧΟΣ ΠΟΙΟΤΗΤΑΣ

Περιγραφή προβλήματος και δημιουργία συνθετικών δεδομένων

Ο απώτερος στόχος της εφαρμογής Ανίχνευσης Αντικειμένων (παράγραφος 3.3.1) είναι ο υπολογισμός των συντεταγμένων των φιαλών στον πραγματικό κόσμο. Αυτές οι συντεταγμένες χρησιμοποιήθηκαν αργότερα σε μια διαδικασία αρπαγής και τοποθέτησης αντικειμένων, με στόχο τη δημιουργία μικτών συσκευασιών από κόκκινα και πράσινα μπουκάλια με βάση μια παραγγελία που δόθηκε από έναν χρήστη. Επομένως, πρέπει να εφαρμοστεί ένα σύστημα ποιοτικού ελέγχου προκειμένου να παρατηρηθεί και να διασφαλιστεί η σωστή προετοιμασία της μεικτής συσκευασίας. Το προτεινόμενο λογισμικό θα αποτελείται από δύο αλγόριθμους ανίχνευσης: α) ένας που ανιχνεύει συνεχώς τον συνολικό αριθμό φιαλών μέσα στο κουτί και β) ένας που ανιχνεύει το χρώμα κάθε φιάλης αφού μπει στο κουτί. Για να επιτευχθεί αυτό, πρέπει να δημιουργηθούν νέα συγκεκριμένα σύνολα δεδομένων, που να πληρούν κάθε πιθανό σενάριο δημιουργίας μικτών πακέτων. Οι διαθέσιμες συσκευασίες είναι διαστάσεων 6x3, με μέγιστο αριθμό 18 φιαλών. Δείγμα ενός κουτιού μικτής παραγγελίας δίδεται στο σχήμα 18

Η διαδικασία δημιουργίας συνθετικών δεδομένων θα είναι ακριβώς η ίδια όπως και για τον αλγόριθμο Ανίχνευσης Αντικειμένων (ενότητα 3.2.1). Θα χρησιμοποιηθούν τα ίδια CAD φιαλών, με τις ίδιες ετικέτες να εφαρμόζονται μέσω χαρτογράφησης UV. Η επισήμανση των μπουκαλιών θα ακολουθήσει επίσης τη μορφή μοντέλου YOLO. **Η κύρια διαφορά στη δημιουργία δεδομένων σε αντίθεση με τα σύνολα δεδομένων που περιγράφηκαν προηγουμένως, είναι ότι θα ανιχνευθούν μόνο τα μπουκάλια της πρώτης σειράς.** Ένα παράδειγμα θα καταστήσει το πρόβλημα πιο κατανοητό.



Εικόνα 18: Δείγματα συνθετικών δεδομένων για την αναγνώριση χρώματος

Αλγόριθμος ανίχνευσης αριθμού μπουκαλιών

Ο δεύτερος αλγόριθμος, που είναι υπεύθυνος για την καταμέτρηση του συνολικού αριθμού φιαλών μέσα στο κουτί, θα είναι ο ίδιος με αυτόν που χρησιμοποιήθηκε για την ανίχνευση των καπακιών όλων των φιαλών στην παλέτα (παράγραφος 3.3). Επομένως, δεν απαιτείται δημιουργία δεδομένων για αυτόν τον αλγόριθμο.

Πειραματική διάταξη και παράμετροι εκπαίδευσης

Η πειραματική διάταξη αποτελείται από δύο κάμερες RGB, το κουτί συσκευασίας και τα μπουκάλια που προσφέρει η εταιρεία. Στόχος της διαδικασίας είναι η επιθεώρηση του τύπου και του αριθμού των φιαλών που

τοποθετούνται στο κουτί, σε σχέση με μια δεδομένη παραγγελία. Η πρώτη κάμερα είναι υπεύθυνη για την αναγνώριση του τύπου των φιαλών που εισάγονται στο κουτί, ενώ η δεύτερη είναι υπεύθυνη για την ανίχνευση ολόκληρου του αριθμού των φιαλών μέσα στο κουτί (Εικ. 19).



Εικόνα 19: Κάμερα ανίχνευσης τύπου μπουκαλιών

Για την εκπαίδευση του νέου μοντέλου βάσει των νέων συνθετικών δεδομένων, χρησιμοποιήθηκαν οι παρακάτω μεταβλητές.

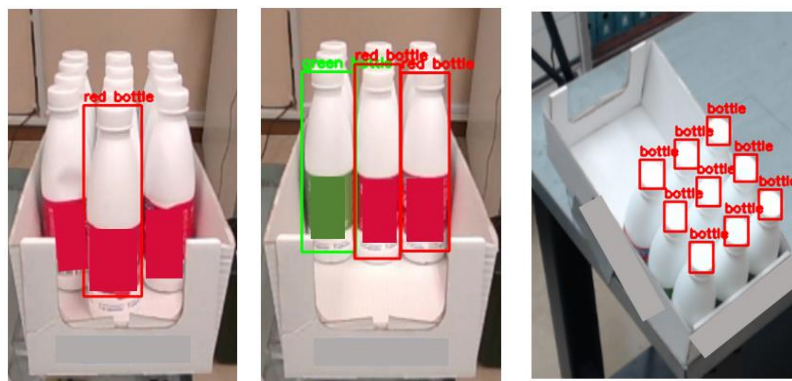
Πίνακας 9: Μεταβλητές εκπαίδευσης ελέγχου ποιότητας

Επιλεγμένα μοντέλα	Αριθμός συνόλων δεδομένων	Μέγεθος συνόλων δεδομένων και περιεχόμενο	Μέγεθος παρτίδας	Ανάλυση εικόνας	Διαχωρισμός δεδομένων εκπαίδευσης και επαλήθευσης (%)
YOLOv4	2	*	64	480x480	90/10

*200 εικόνες με ανάλυση 900x675 χρησιμοποιήθηκαν για την ανίχνευση χρώματος και μπουκαλιών. Η ανίχνευση καπακιών πραγματοποιήθηκε με το σύνολο δεδομένων του προηγούμενου πειράματος (ενότητα 3.3.3)

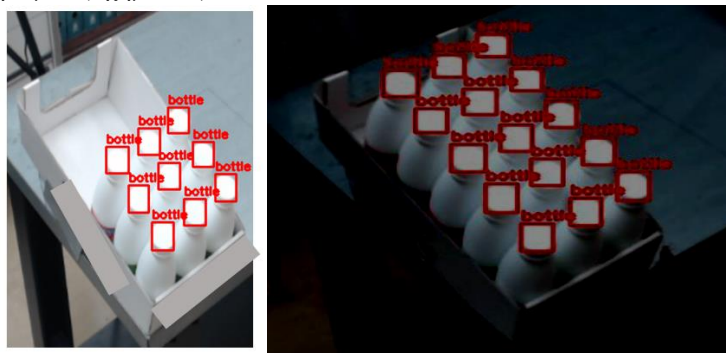
Αποτελέσματα

Το εξάρτημα ποιοτικού ελέγχου ακολουθεί την ακόλουθη διαδικασία. Η κάμερα μέτρησης αριθμού μπουκαλιών καταγράφει εικόνες κάθε 20 δευτερόλεπτα. Ένα μπουκάλι τοποθετείται στο κουτί. Το μπουκάλι ανιχνεύεται από την κάμερα μέτρησης αριθμού μπουκαλιών. Η έγχρωμη κάμερα ενεργοποιείται και ανιχνεύει το μπουκάλι του χρώματος. Η ίδια διαδικασία επαναλαμβάνεται κάθε φορά που τοποθετείται ένα νέο μπουκάλι. Προκειμένου να αποφευχθεί ο εντοπισμός του χρώματος του ίδιου μπουκαλιού δύο φορές (στην ίδια σειρά), η έγχρωμη κάμερα έχει ρυθμιστεί ώστε να αποθηκεύει δεδομένα μία φορά κάθε τρία μπουκάλια που ανιχνεύονται. Μόλις η παραγγελία πλησιάσει την τελευταία σειρά των αναμενόμενων φιαλών (σύμφωνα με την παραγγελία που δίνεται), τότε επιτρέπεται στον αλγόριθμο ανίχνευσης χρώματος να ανιχνεύει και να αποθηκεύει δεδομένα για λιγότερα από 3 μπουκάλια.



Εικόνα 20: Παραδείγματα ανίχνευσης από τις δύο κάμερες του ελέγχου ποιότητας

Η παραπάνω εικόνα παρέχει μια προφανή απόδειξη της δυσκολίας του εντοπισμού των καπακιών των φιαλών. Ο φωτισμός και η αντίθεση καθιστούν την αναγνώριση των φιαλών δύσκολη ακόμα και για το ανθρώπινο μάτι, καθώς οι λευκές επιφάνειες είναι ομοιόμορφες. Οι αλλαγές εικόνες ήταν απαραίτητες για την παροχή σταθερής και αξιόπιστης απόδοσης αλγορίθμου (Σχήμα. 21).



Εικόνα 21: Φωτογραφία πριν και μετά τις βελτιώσεις

4. ΣΥΖΗΤΗΣΗ ΚΑΙ ΣΥΜΠΕΡΑΣΜΑΤΑ

Οι δοκιμές σε πραγματικά δεδομένα εμφάνισαν πολύ ενδιαφέροντα αποτελέσματα. Ο Πίνακας 3 αντιπροσωπεύει τα αποτελέσματα των δοκιμών που βασίζονται μόνο σε εικόνες ανω- προσανατολισμένες. Τα VGG16 και MobileNet απέδωσαν ικανοποιητικά, ενώ το DenseNet121 είχε μέτρια απόδοση. Στη συνέχεια, ο Πίνακας 4 απεικονίζει τα αποτελέσματα των δοκιμών που βασίζονται μόνο σε εικόνες με κατεύθυνση προς τα κάτω. Παρόλο που ο φωτισμός κατευθύνθηκε προς το κέντρο του τμήματος, παρέχοντας καλύτερη προβολή, αυτό το σύνολο δεδομένων φαίνεται να είναι πιο απαιτητικό για όλα τα μοντέλα. Η απόδοση στο VGG16 είναι σταθερή, ενώ το MobileNet παρουσιάζει σημαντική μείωση της ακρίβειας. Το DenseNet121 είχε ήδη την πιο αδύναμη απόδοση μεταξύ των τριών μοντέλων και επιδεινώθηκε ακόμη περισσότερο όταν αντιμετώπισε το δεύτερο σύνολο δεδομένων. Ο κάτω προσανατολισμός αναμενόταν να είναι πιο δύσκολο να ταξινομηθεί λόγω της γωνίας περιστροφής του κομματιού κατά τον άξονα x, όταν βρίσκεται στον πάγκο εργασίας. Δοκιμάστηκαν πολλές γωνίες, παρέχοντας ελαφρώς καλύτερα αποτελέσματα. Τέλος, οι επιδόσεις για κάθε CNN στα συνδυασμένα σύνολα δεδομένων φαίνονται στον Πίνακα 5. Το VGG16 φάνηκε να είναι αξιόπιστο και σε αυτό το σενάριο, το MobileNet απέδωσε ακόμη καλύτερα και από το πρώτο σενάριο πειραμάτων, ενώ το DenseNet121 διατήρησε την ακρίβειά του. Είναι μια διαφορετική εργασία και η αρχιτεκτονική των μοντέλων διαδραματίζει σημαντικό ρόλο στην απόδοση του κάθε μοντέλου. Όπως αναφέρεται από την βιβλιογραφία της Keras, το DenseNet121 είχε την καλύτερη απόδοση, μεταξύ των τριών μοντέλων, στο σύνολο δεδομένων ImageNet. Παρ όλα ταύτα, σε αυτό το σενάριο παρουσίασε την χαμηλότερη απόδοση. Η απλή και βαθιά αρχιτεκτονική του VGG16 (βασίζεται κυρίως σε στρώματα Συνέλιξης και Maxpooling) φαίνεται ότι είναι κατάλληλη για μοναδικά, προσαρμοσμένα σύνολα δεδομένων όπως αυτό που χρησιμοποιήθηκε σε αυτή τη μελέτη, καθώς

Τμήμα Μηχανολόγων και Αεροναυπηγών Μηχανικών – [Κατασκευαστικός Τομέας]

προέβλεψε σχεδόν κάθε πιθανή εικόνα με την οποία δοκιμάστηκε. Τα σφάλματα ήταν μόνο $\pm 20^\circ$ μακριά από τη σωστή κατηγορία, επισημαίνοντας ότι με βάση το απαιτούμενο βήμα γωνίας, η ακρίβεια μπορεί να είναι ακόμη μεγαλύτερη.

Στη δεύτερη και τρίτη μελέτη, η εφαρμογή των συνθετικών δεδομένων δοκιμάστηκε σε εργασίες ανίχνευσης αντικειμένων. Και οι δύο εργασίες ανίχνευσης φάνηκαν να είναι αποτελεσματικές και επιτυχημένες, εκπληρώνοντας έτσι τον ρόλο τους ως μέρος μιας μεγαλύτερης, πιο σύνθετης ιδέας.

Αν και αυτή η εργασία ήταν απλώς ένα προκαταρκτικό βήμα για την τρίτη μελέτη και, επομένως, ήταν ένα σχετικά απλό πρόβλημα, παρείχε ορισμένα σημαντικά συμπεράσματα και παρατηρήσεις σχετικά με τη βελτιστοποίηση ανίχνευσης αντικειμένων. Συγκεκριμένα: Η γωνία της κάμερας είναι σημαντική ειδικά εάν το σύνολο δεδομένων που δημιουργείται δεν είναι πλούσιο σε διαφοροποιήσεις γωνίας κάμερας. Ο φωτισμός είναι κρίσιμος και μπορεί να είναι περιοριστικός παράγοντας σε πολλές περιπτώσεις για την ενσωμάτωση διεργασιών σε βιομηχανικά περιβάλλοντα. Ρύθμιση πειράματος: Πολλά αντικείμενα και ειδικά με μικρές διαφορές, όπως οι δύο τύποι φιαλών στην περίπτωση μας, μπορεί να περιπλέξουν τον αλγόριθμο και να οδηγήσουν σε λάθος εντοπισμό. Το όριο ακρίβειας σχετίζεται στενά με την πολυπλοκότητα του πειράματος.

Ο αλγόριθμος αναγνώρισης αντικειμένων, σε συνδυασμό με τον αλγόριθμο υπολογισμού συντεταγμένων πραγματικού κόσμου, ανταποκρίθηκε στις προσδοκίες της εργασίας και πρόσφερε τις επιθυμητές, επιτυχημένες αρπαγές φιαλών που ήταν απαραίτητες για την εργασία πακεταρίσματος που εξετάστηκε. Συγκεκριμένα, η ανίχνευση ήταν άψογη, αφού προσδιορίστηκαν οι απαραίτητες παράμετροι, ενώ τα αποτελέσματα του υπολογισμού συντεταγμένων ήταν αρκετά ικανοποιητικά και επαρκή για ένα επιτυχημένο πιάσιμο από την αρπάγη του ρομπότ. Βασικές πτυχές που εξασφάλισαν υψηλή απόδοση είναι οι εξής μεταξύ άλλων: α) Η χαρτογράφηση ετικετών η οποία δημιούργησε ρεαλιστικά αντικείμενα σχεδόν πανομοιότυπα με τα πραγματικά, β) η απόφαση ανίχνευσης του καπακιού της φιάλης αντί για την ανίχνευση ολόκληρου του μπουκαλιού ήταν κρίσιμη για την υψηλή ακρίβεια του αλγόριθμου υπολογισμού συντεταγμένων, γ) σωστή βαθμονόμηση και τοποθέτηση της κάμερας σε μια ορισμένη απόσταση από τα αντικείμενα, δ) η ανάλυση εικόνας των πραγματικών εικόνων πρέπει να είναι παρόμοια με τις συνθετικά δημιουργούμενες, ε) η χαρτογράφηση εικονοστοιχείων πρέπει να είναι προσεκτική και ακριβής τόσο κατά τη μέτρηση των συντεταγμένων του πραγματικού κόσμου όσο και των τιμών των pixel.

Η ακρίβεια ανίχνευσης φαίνεται να είναι επαρκής για τους σκοπούς της μελέτης μας. Η ακρίβεια του κουτιού οριοθέτησης δεν είναι απαραίτητη εφόσον ανιχνεύονται οι σωστές φιάλες. Για παράδειγμα, κατά την ανάπτυξη της τελικής έκδοσης του αλγορίθμου, υπήρξαν περιστατικά εσφαλμένης ανίχνευσης. Η αύξηση του συνόλου δεδομένων ήταν η λύση σε αυτήν την πρόκληση. Βασικός παράγοντας για την επιτυχία του αλγορίθμου ήταν επίσης η περικοπή εικόνας. Η μεγέθυνση (zoom) και η περικοπή της εικόνας, προκειμένου να είναι όσο το δυνατόν πιο παρόμοια με εκείνα των συνθετικών δεδομένων, φάνηκε να μεγιστοποιεί την ακρίβεια και να παρέχει τα επιθυμητά αποτελέσματα.

Η μείωση της φωτεινότητας ήταν καθοριστική για την απόδοση του αλγορίθμου. Όπως φαίνεται στην Εικ. 28, η υψηλή φωτεινότητα σε συνδυασμό με το λευκό χρώμα του φωτισμού στο δωμάτιο, περιπλέκει την ανίχνευση του λευκού καπακιού της φιάλης, ακόμη και για το ανθρώπινο μάτι. Η μείωση της φωτεινότητας μαζί με τις αλλαγές αντίθεσης προσέφεραν το επιθυμητό αποτέλεσμα. Όπως απεικονίζεται στην Εικ. 21, μια εικόνα με χαμηλότερη φωτεινότητα έχει ως αποτέλεσμα υψηλότερη αντίθεση μεταξύ του πώματος της φιάλης και του περιβάλλοντος χώρου του, παρέχοντας έτσι μεγαλύτερη και καλύτερη ακρίβεια ανίχνευσης.

5. ΑΝΑΦΟΡΕΣ

Blender Online Community.2019.Blender - a 3D Modelling andRendering Package. Amsterdam: Stichting BlenderFoundation.<http://www.blender.org>

Efros, A. A., & Freeman, W. T. (2001, August). Image quilting for texture synthesis and transfer. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques* (pp. 341-346).

Mikołajczyk, A., & Grochowski, M. (2018, May). Data augmentation for improving deep learning in image classification problem. In *2018 international interdisciplinary PhD workshop (IIPhDW)* (pp. 117-122). IEEE.