



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Ανάπτυξη αλγορίθμου για το διαχωρισμό ξύλινων απορριμμάτων με τη χρήση τεχνικών τεχνητής νοημοσύνης

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ
Του
Βερέμη Αλέξανδρου

Επιβλέπων: Παναγιώτης Τσανάκας
Καθηγητής

Αθήνα, Ιούνιος 2023



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Ανάπτυξη αλγορίθμου για το διαχωρισμό ξύλινων απορριμμάτων με τη χρήση τεχνικών τεχνητής νοημοσύνης

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ
Του
Βερέμη Αλέξανδρου

Επιβλέπων: Παναγιώτης Τσανάκας
Καθηγητής

Εγκρίθηκε από την τριμελή επιτροπή την 18η Ιουλίου 2023.

.....
Παναγιώτης Τσανάκας
Καθηγητής

.....
Δημήτριος Σούντρης
Καθηγητής

.....
Άγγελος Αμδίτης
Ερευνητής ΕΠΙΣΕΥ

Αθήνα, Ιούνιος 2023

.....
Αλέξανδρος Βερέμης
Διπλωματούχος Ηλεκτρολόγος Μηχανικός
και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright© Αλέξανδρος Βερέμης, 2023
Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Οι πεπερασμένοι πόροι και τα κλιματικά ζητήματα έχουν καταστήσει αναγκαία τη μετάβαση στο μοντέλο της κυκλικής οικονομίας. Λόγω της ραγδαίας αύξησης της παραγωγής αστικών στερεών απορριμμάτων παγκοσμίως, η ΕΕ έχει θέσει νέους στόχους αποδοτικότερης ανακύκλωσης τους. Συνήθως τα αστικά απορρίμματα αποτελούν μια αρκετά πολύπλοκη ροή στερεών υλικών που περιλαμβάνει μέταλλα, ξύλο, πλαστικό, χαρτί, γυαλί, υλικά κατασκευών και κατεδαφίσεων κ.α. Η αποτελεσματική διαλογή των υλικών αποτελεί κρίσιμη προϋπόθεση για μία επιτυχημένη ανακύκλωση, συνεπώς η ερευνητική κοινότητα έχει στραφεί πλέον στις αυτοματοποιημένες τεχνικές διαλογής με στόχο τον εκσυγχρονισμό και την ψηφιοποίηση της ανακυκλωτικής διαδικασίας, όπου αισθητήρες χρησιμοποιούνται για την ανίχνευση των ανακυκλώσιμων υλικών στα απόβλητα με σκοπό την ανάπτυξη αυτόνομων συστημάτων για την ταξινόμησή τους.

Στο πλαίσιο αυτής της διπλωματικής εργασίας έγινε η ανάπτυξη ενός αλγορίθμου, βασισμένου σε τεχνικές βαθιάς μάθησης (deep learning), σε εργαστηριακό βιομηχανικό περιβάλλον, με σκοπό τη δημιουργία ενός αυτοματοποιημένου συστήματος διαχωρισμού ξύλινων απορριμμάτων. Η εκπαίδευση του μοντέλου πραγματοποιήθηκε επάνω σε δεδομένα που προέρχονταν από δύο πολυφασματικούς αισθητήρες (multispectral sensors) και μία βιομηχανική κάμερα. Με αυτόν τον τρόπο ήταν δυνατή η αναγνώριση, ο χαρακτηρισμός και η κατηγοριοποίηση των διαφόρων ειδών ξύλου, όπως επίσης και ο διαχωρισμός τους από άλλα απόβλητα, με βάση την σύνθεσή τους. Στη συνέχεια ακολούθησε η αντίστροφη διαδικασία της επαύξησης των εικόνων που είχαμε συλλέξει για τον έλεγχο ύπαρξης βελτίωσης στην ακρίβεια του μοντέλου μας, αξιοποιώντας τα ήδη υπάρχοντα δεδομένα. Ενώ, στο τέλος, υλοποιήθηκε μία έρευνα επάνω στις φασματικές ζώνες, στα πρότυπα της επιλογής χαρακτηριστικών (feature selection), αποσκοπώντας στη μελέτη του αντίκτυπου ενός εξοπλισμού κατώτερου επιπέδου και κατά συνέπεια στο ενδεχόμενο μιας πιο ευρείας αξιοποίησης του παραχθέντος μοντέλου.

Λέξεις Κλειδιά - Ξύλινα Απορρίμματα, Κυκλική Οικονομία, Τεχνητή Νοημοσύνη, Συνελικτικό Νευρωνικό Δίκτυο, Πολυφασματική Απεικόνιση.

Abstract

Finite resources and climate issues have necessitated the transition to the circular economy model. Due to the rapid increase in the production of municipal solid waste worldwide, the EU has set new targets for more efficient recycling. Typically, municipal waste is a fairly complex stream of solid materials comprising metal, wood, plastic, glass, construction and demolition materials, etc. The effective sorting of materials is a critical requirement for successful recycling and therefore the research community has now turned to automated sorting techniques to modernize and digitize the recycling process, where sensors are used to detect recyclable materials in the waste in order to develop autonomous systems for their sorting.

In the context of this thesis, an algorithm based on deep learning techniques was developed in a laboratory industrial environment, in order to create an automated system for the separation of wood waste. The training of the model was performed on data received from two multispectral sensors and an industrial camera. In this way it was possible to identify, characterize and categorize different types of wood, as well as to separate them from other waste based on their composition. Then followed, the reverse process of augmenting the images collected to check whether an improvement in the accuracy of our model was possible, using the existing data. In the end, a research on the spectral bands was implemented, in the standards of feature selection, aiming to study the impact of a lower-level equipment and consequently the possibility of a wider usage of the generated model.

Keywords - Wood Waste, Circular Economy, Artificial Intelligence, Convolutional Neural Network, Multispectral Imaging.

στη γιαγιά μου, Ασπασία

Ευχαριστίες

Καταρχάς, θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή μου κ. Παναγιώτη Τσανάκα, ο οποίος με εμπιστεύθηκε για την εκπόνηση αυτής της διπλωματικής εργασίας. Ακόμη, οφείλω ένα τεράστιο ευχαριστώ στους πάντα πρόθυμους κ. Σάββα Σιφναίο και κ. Γεώργιο Αρβανιτάκη που ήταν κάθε εβδομάδα παρόντες και στάθηκαν αρωγοί στην προσπάθειά μου. Χωρίς εκείνους η περάτωση της διπλωματικής μου θα καθίσταντο αρκετά δυσκολότερη και αδιαμφισβήτητα λιγότερο ευχάριστη εμπειρία.

Στη συνέχεια, θα ήθελα να ευχαριστήσω τους παιδικούς μου φίλους που δεν έπαψαν να πιστεύουν σε μένα, τους συμφοιτητές μου -που πλέον γίναν φίλοι- και μου έδωσαν την απαραίτητη ώθηση για την αποπεράτωση των σπουδών μου, τους ανθρώπους που μέσω της σχολής απέκτησα τη δυνατότητα να γνωρίσω, αλλά και όλους όσους συμπορευθήκαμε για κάποιο διάστημα και έπειτα ο κύκλος μας έκλεισε.

Τέλος, χρωστώ αμέριστη ευγνωμοσύνη στην οικογένειά μου για την ανιδιοτελή αγάπη τους και τη στήριξή τους όλα αυτά τα χρόνια.

Acknowledgments

First and foremost, I would like to warmly thank my supervisor, Prof. Panagiotis Tsanakas, who trusted me with this diploma thesis. Furthermore, I owe a huge thanks to the always willing Mr. Savvas Sifnaios and Mr. George Arvanitakis who were present every week and helped me in my endeavor. Without them, completing my diploma thesis would have been a much more difficult and undeniably less enjoyable experience.

Afterwards, I would like to thank my childhood friends who never stopped believing in me, my fellow students - who have now become friends - and who gave me the necessary push to finish my studies, the people who I would not have met if it had not been for this school, but also all those with whom I journeyed together on the road of life, until our paths diverged.

Finally, I owe a debt of gratitude to my family for their selfless love and support throughout the years.

Contents

Περίληψη.....	5
Abstract.....	6
Ευχαριστίες.....	9
Acknowledgments.....	10
Contents.....	11
List of Figures.....	14
List of Tables.....	17
Εκτεταμένη Περίληψη.....	18
Κίνητρα.....	18
Διαχείριση Αποβλήτων.....	20
Ξύλινα Απορρίμματα.....	20
Κυκλική Οικονομία.....	21
Ξύλινα Απορρίμματα και Κυκλική Οικονομία.....	23
Κατηγορίες Ξύλινων Απορριμάτων.....	24
Πολυφασματική Απεικόνιση.....	25
Φασματική Απεικόνιση.....	25
Πώς λειτουργεί η Πολυφασματική Απεικόνιση;.....	27
Τα Πλεονεκτήματα της Πολυφασματικής Απεικόνισης.....	27
Πεδία εφαρμογών της Πολυφασματικής Απεικόνισης.....	27
Τεχνητή Νοημοσύνη.....	28
Συνελικτικά Νευρωνικά Δίκτυα (ΣΝΔ).....	30
Συνελικτικό Επίπεδο.....	31
Επίπεδο Ομαδοποίησης.....	32
Πλήρως Συνδεδεμένο (ΠΣ) Επίπεδο.....	32
Αρχιτεκτονική του Δικτύου ResNet50.....	32
Όραση Υπολογιστών.....	33
Εργαλεία και Μέθοδοι που Ακολουθήθηκαν.....	35
Προσεγγίσεις για την επιλογή των ζωνών σε περίπτωση ανεπαρκούς εξοπλισμού.....	38
Διάταξη Πειράματος.....	39
Μετρικές Αξιολόγησης.....	41
Πίνακας Σύγχυσης.....	41
Ακρίβεια.....	41
Απώλεια.....	43
Αποτελέσματα και Συζήτηση.....	44
Συνεισφορά.....	49
Σύννοψη Διπλωματικής.....	50

Μελλοντικές Επεκτάσεις.....	51
Chapter 1 Introduction	54
1.1. Motivation	54
1.2. Waste Management	55
1.3. Recycling.....	56
1.4. Wood Recycling	57
1.4.1. The Obstacles related to Wood Recycling.....	57
1.5. Re-use.....	58
1.5.1. Biological reprocessing.....	58
1.5.2. Energy recovery.....	58
1.5.3. Pyrolysis	58
1.5.4. Resource recovery.....	58
1.6. Applications of Waste Management.....	59
1.7. Wood Wastes	59
1.8. Sectors of Wood Wastes.....	60
1.9. Circular Economy	61
1.10. Wood Waste and Circular Economy	62
1.11. Waste Frequency of Appearance.....	63
1.12. Wood Waste Categories	64
Chapter 2 Theoretical Background	67
2.1. Hyperspectral Images	67
2.2. Multispectral Imaging	68
2.2.1. Spectral Imaging.....	68
2.2.2. How does multispectral imaging work?	69
2.2.3. The Advantages of Multispectral Imaging.....	70
2.2.4. Fields of Multispectral Imaging	70
Chapter 3 Related work	71
3.1. General Waste Sorting Methods hitherto	71
3.2. Wood Waste Sorting Methods hitherto.....	72
3.3. Microscopic Methods for Wood Waste Characterization.....	73
3.4. Complete Autonomous Waste Sorter.....	75
Chapter 4 Artificial Intelligence	77
4.1. Artificial Intelligence.....	77
4.2. Machine Learning.....	78
4.3. Deep Learning.....	79
4.3.1. Epoch.....	80
4.4. Applications of Deep Learning and CNNs.....	80

4.5. Neural Networks (NN)	81
4.5.1. How do neural networks work?.....	82
4.6. Convolutional Neural Networks (CNN).....	84
4.6.1. Convolutional Layer	85
4.6.2. Pooling Layer.....	86
4.6.3. Fully-Connected Layer.....	86
4.6.4. ResNet50 Architecture.....	86
4.7. Computer Vision.....	87
4.8. Multispectral Imaging in Computer Vision	88
4.8.1. Principal Component Analysis (PCA)	88
4.8.2. Applying PCA on Multispectral Images.....	89
Chapter 5 Tools and Methods.....	91
5.1. Data Generation and Training of the Model	91
5.1.1. Approaches to the selection of the bands in case of inadequate equipment.....	94
Chapter 6 Experiment	95
6.1. Experimental Setup	95
6.2. Evaluation Metrics	96
6.2.1. Confusion Matrix.....	96
6.2.2. Accuracy	97
6.2.3. Loss Function	98
6.3. Results and Discussion	99
6.4. Contributions	104
Chapter 7 Conclusions.....	105
7.1. Thesis Summary	105
7.2. Future work	106
Bibliography	108

List of Figures

Εικόνα 1. Προβλεπόμενη παραγωγή αποβλήτων ανά περιοχή (εκατομμύρια τόνοι/έτος)...	18
Εικόνα 2. Γραφική αναπαράσταση της Γραμμικής Οικονομίας σε σύγκριση με την Κυκλική.	22
Εικόνα 3. Η υπερφασματική απεικόνιση καταγράφει μήκη κύματος από 250 nm έως 15.000 nm και θερμικές υπέρυθρες.	26
Εικόνα 4. Προκειμένου να αντιγραφεί η λειτουργία της ανθρώπινης όρασης, δημιουργείται μια ψηφιακή φωτογραφία ενός φύλλου (επάνω) χρησιμοποιώντας τρεις ζώνες: κόκκινο, πράσινο και μπλε. Τα RGB δεδομένα είναι συγκρίσιμα με ένα φυλλάδιο τριών σελίδων. Αντίθετα, μια υπερφασματική εικόνα του ίδιου φύλλου (κάτω) καταγράφει μια φασματική απόκριση από 220 μήκη κύματος, με αποτέλεσμα μια πιο εκτεταμένη και λεπτομερή αναπαράσταση του αντικειμένου, παρόμοια με ένα βιβλίο 220 σελίδων.	26
Εικόνα 5. Μια ακολουθία ΣΝΔ για την ταξινόμηση χειρόγραφων χαρακτήρων.	30
Εικόνα 6. Ένα διάγραμμα αναπαράστασης προεκπαιδευμένης αρχιτεκτονικής Resnet-50.33	
Εικόνα 7. Επαύξηση Εικόνων: Ξεκινώντας από την επάνω αριστερή γωνία: Αρχική RGB εικόνα ενός τεμαχίου ξύλου της κατηγορίας καπλαμά δρυός, αρχική VIS εικόνα, αρχική NIR εικόνα, ακολουθούμενες από πέντε επαυξημένες εκδόσεις RGB του ίδιου κομματιού.	37
Εικόνα 8. Εργαστηριακή διάταξη των πειραμάτων μας - Αυτόνομο σύστημα διαλογής, το οποίο αποτελείται από ένα βιομηχανικό κάδο, έναν μηχανικό διαχωριστή, ένα σύστημα αναγνώρισης με πολλαπλές κάμερες, έναν ρομποτικό βραχίονα, κάδους εναπόθεσης και μια κεντρική μονάδα επεξεργασίας.	40
Εικόνα 9. Πίνακας Σύγχυσης.	41
Εικόνα 10. Πίνακας σύγχυσης για ξύλο και πλαστικό του μέσου ταξινομητή.	45
Εικόνα 11. Πίνακας σύγχυσης για ξύλο και πλαστικό του μέσου ταξινομητή (πλειοψηφία των εκδοχών με αντιστροφή της επαύξησης).	45
Εικόνα 12. Πίνακας σύγχυσης και των 7 κατηγοριών του μέσου ταξινομητή. Παρουσιάζονται οι ακόλουθες λανθασμένες ταξινομήσεις: 17 δείγματα MDF, 3 εικόνες LDPE, 6 PET και 70 PP, ενώ δεν έγιναν λάθη στις κατηγορίες MFC, Βελανιδιά (ΟΑΚ) ή HDPE.	46
Εικόνα 13. Πίνακας σύγχυσης και των 7 κατηγοριών του μέσου ταξινομητή (Αντιστροφή της επαύξησης). Μπορούμε να παρατηρήσουμε ότι μόνο 14 από 214 σύνολα των 6 εκδοχών κατατάσσονται λάθος, εξού και 93.457% ακρίβεια.	46
Εικόνα 14. Ιστόγραμμα λανθασμένων εκδοχών ανά εικόνα. Έξι λάθη ανά εικόνα σημαίνει ότι και οι έξι εκδόσεις κατατάχθηκαν στη λανθασμένη κατηγορία, ενώ το μηδέν σημαίνει ότι όλες κατατάχθηκαν σωστά.	47
Εικόνα 15. Διάγραμμα που παρουσιάζει την εξέλιξη της ακρίβειας των συνόλων τριών ζωνών της κάμερας NIR καθώς προχωρούν οι εποχές. Η κορυφή κάθε γραμμής απεικονίζεται με δεκαδικό αριθμό. Επεξήγηση γραμμών: "randomAvg" αναφέρεται στο μέσο όρο των συνόλων τυχαίας επιλογής, το "bestTotal" αντιπροσωπεύει τη δεύτερη τεχνική ("καλύτερη εν συνόλω"), ενώ τα "bestEach" και "worstEach" αντιπροσωπεύουν την τρίτη και την τέταρτη μέθοδο που εφαρμόσαμε, αντίστοιχα.	48

Εικόνα 16. Διάγραμμα που παρουσιάζει την εξέλιξη της ακρίβειας των συνόλων τριών ζωνών της κάμερας VIS καθώς προχωρούν οι εποχές. Η κορυφή κάθε γραμμής απεικονίζεται με δεκαδικό αριθμό. Επεξήγηση γραμμών: "randomAvg" αναφέρεται στο μέσο όρο των συνόλων τυχαίας επιλογής, το "bestTotal" αντιπροσωπεύει τη δεύτερη τεχνική ("καλύτερη εν συνόλω"), ενώ τα "bestEach" και "worstEach" αντιπροσωπεύουν την τρίτη και την τέταρτη μέθοδο που εφαρμόσαμε, αντίστοιχα.48

Figure 1. Projected waste generation, by region (millions of tonnes/year).	54
Figure 2. Graphic representation of Linear in contrast with Circular economy.	62
Figure 3. Percentage proportion of Municipal Solid Waste (MSW) types.	63
Figure 4. RGB vs HSI images.	67
Figure 5. Hyperspectral imaging captures wavelengths from 250 nm to 15,000 nm and thermal infrared.	69
Figure 6. In order to replicate human vision, a digital photograph of a leaf (top) is created using three bands: red, green, and blue. The RGB data is comparable to a three-page brochure. In contrast, a hyperspectral image of the same leaf (bottom) captures a spectral response from 220 wavelengths, resulting in a more extensive and detailed representation of the object, akin to a book with 220 pages.	69
Figure 7. Principal component analysis performed on four wood products-plastic composites (A) and loadings analysis (B).	73
Figure 8. Microscopic monitoring of filamentous fungus growth during a bioremediation process: episcopy microscopy (A) and scanning electron microscopy (B).	74
Figure 9. CLSM image showing the location of the resin (red) on wood fiber (green) before (A) after (B) decontamination treatment.	74
Figure 10. Setups of Autonomous General Waste Sorters.	75
Figure 11. Graph of a typical Neural Network.	81
Figure 13. A CNN sequence to classify handwritten digits.	84
Figure 14. A block diagram representation of pre-trained Resnet-50 architecture.	87
Figure 15. Plot of the eigenvalues of the NIR image of the LDPE product depicted below.	90
Figure 16. The NIR image of the LDPE product whose eigenvalues are shown above.	90
Figure 17. Image Data Augmentation: Starting from the top left corner: Original RGB Image of an Oak Veneer piece of wood, Original VIS image, Original NIR image, Five Augmented RGB versions of the same piece.	92
Figure 18. Laboratory Setup of our Experiments - Autonomous sorting system, which is made up of an industrial bucket, a mechanical separator, a multi-camera identification system, a robotic arm, bins, and a central processing unit.	96
Figure 19. Confusion Matrix.	97
Figure 20. Confusion Matrix of Wood and Plastic of the Average Classifier.	100
Figure 21. Confusion Matrix of Wood and Plastic of the Average Classifier (Reverse Augmentation).	100
Figure 22. Confusion Matrix of all 7 categories of the Average classifier. The following misclassifications are depicted: 17 samples of MDF, 3 images of LDPE, 6 of PET and 70 of PP, while no mistakes were made for MFC, OAK or HDPE class.	101

Figure 23. Confusion Matrix of all 7 categories of the Average classifier (Reverse Augmentation). We can observe that only 14 out of 214 sets of 6 versions are misclassified, hence the 93.457% accuracy. 101

Figure 24. Histogram of wrong editions per image. Six mistakes per image means that all of the six editions were classified in the wrong category while zero stands for all of them being classified right. 102

Figure 25. Diagram presenting the NIR camera three-banded sets' accuracy evolution as epochs progress. The peak of each line is depicted with a decimal number. Line explanation: "randomAvg,, stands for the average of the random selection, "bestTotal,, stands for the second technique ("best in total,,), while "bestEach,, and "worstEach,, stand for the third and fourth method applied, respectively. 103

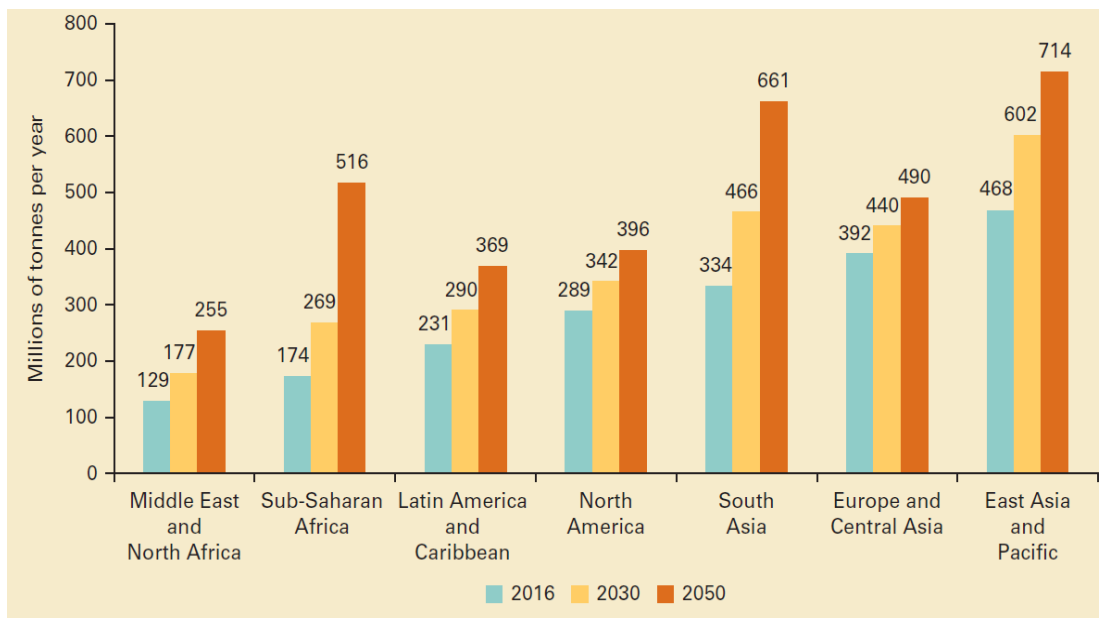
Figure 26. Diagram presenting the VIS camera three-banded sets' accuracy evolution as epochs progress. The peak of each line is depicted with a decimal number. Line explanation: "randomAvg,, stands for the average of the random selection, "bestTotal,, stands for the second technique ("best in total,,), while "bestEach,, and "worstEach,, stand for the third and fourth method applied, respectively. 103

List of Tables

Πίνακας 1. Συνεισφορά εικόνων κάθε υλικού στα διάφορα σύνολα.....	36
Πίνακας 2. Προδιαγραφές της βιομηχανικής κάμερας.	39
Πίνακας 3. Προδιαγραφές των πολυφασματικών αισθητήρων.	39
Πίνακας 4. Πίνακας Σύγκρισης ενός μοντέλου που κατατάσσει 100 όγκους ως καλοήθεις ή κακοήθεις.	42
Table 1. Wood Waste Categories: Markets and sources.....	64
Table 2. Wood Waste Categories: Materials, content before processing and notes.	65
Table 3. Contribution of each material image to the sets.....	92
Table 4. Specifications of the Industrial Camera.	95
Table 5. Specifications of the multispectral sensors.	95
Table 6. Confusion matrix of a model that classifies 100 tumors as either malignant or benign.	98

Εκτεταμένη Περίληψη

Κίνητρα



Εικόνα 1. Προβλεπόμενη παραγωγή αποβλήτων ανά περιοχή (εκατομμύρια τόνοι/έτος).

Στον εκσυγχρονισμό και την αστικοποίηση της καθημερινής ζωής, αναπόφευκτα παράγεται μεγάλη ποσότητα αστικών στερεών αποβλήτων (ΑΣΑ). Τα απόβλητα ξύλου αποτελούν την 6η μεγαλύτερη κατηγορία ΑΣΑ. Στην Ευρώπη μαζί με την Κεντρική Ασία, το 2016, παρήχθησαν 392 εκατομμύρια τόνοι ΑΣΑ, οι οποίοι μεταφράζονται σε περίπου 23,5 εκατομμύρια τόνους αποβλήτων ξύλου, δεδομένου ότι τα απορρίμματα ξύλου καταλαμβάνουν σχεδόν το 6% του συνολικού ποσοστού των ΑΣΑ [1].

Οι πεπερασμένοι πόροι και οι περιβαλλοντικοί κίνδυνοι έχουν καταστήσει αναγκαία τη μετάβαση στο μοντέλο της κυκλικής οικονομίας. Λόγω της ταχείας αύξησης που προβλέπεται στην παραγωγή ΑΣΑ παγκοσμίως και για λόγους βιωσιμότητας, η Ευρωπαϊκή Ένωση έχει θέσει αυστηρότερους στόχους για την ανακύκλωση και έχει αρχίσει να ενθαρρύνει τις κυβερνήσεις να ανακυκλώνουν τα στερεά απόβλητα. Ωστόσο, ο διαχωρισμός των απορριμμάτων ξύλου μέσω τεχνικών τεχνητής νοημοσύνης εξακολουθεί να παραμένει ένα ανεξερεύνητο πεδίο.

Ο παλιός παραδοσιακός τρόπος διαχωρισμού των απορριμμάτων είναι η μέθοδος της χειροδιαλογής, κατά την οποία απασχολούνται άνθρωποι για να διαχωρίσουν τις διάφορες κατηγορίες υλικών. Ωστόσο, λόγω της μακροχρόνιας έκθεσης σε τοξικά και παθογόνα περιβάλλοντα εργασίας, η χειρωνακτική διαλογή έχει πολλούς παράγοντες κινδύνου, υψηλό εργατικό κόστος και χαμηλή απόδοση διαλογής. Ο αποτελεσματικός διαχωρισμός των υλικών αποτελεί κρίσιμη προϋπόθεση για την επιτυχή ανακύκλωση, γι' αυτό και η ερευνητική κοινότητα

έχει στραφεί στις αυτοματοποιημένες τεχνικές διαλογής με σκοπό τον εκσυγχρονισμό και την ψηφιοποίηση της διαδικασίας ανακύκλωσης, όπου χρησιμοποιούνται αισθητήρες για την ανίχνευση ανακυκλώσιμων υλικών στα απόβλητα, προκειμένου να αναπτυχθούν αυτόνομα συστήματα διαχωρισμού.

Συνήθως, τα ΑΣΑ είναι ένα αρκετά σύνθετο μείγμα στερεών υλικών που περιλαμβάνει μέταλλα, πολυμερή-πλαστικά, ξύλο, χαρτί, γυαλί, υλικά κατασκευών και κατεδαφίσεων κ.λπ. Επιπλέον, τα απορρίμματα ξύλου έχουν μεγάλη ποικιλία, πολύπλοκη σύνθεση και ποσοστό ανακύκλωσης που εξαρτάται από την κατάσταση κάθε μεμονωμένου τεμαχίου.

Λαμβάνοντας όλα αυτά υπόψη, παρακινηθήκαμε να αναπτύξουμε ένα αυτοματοποιημένο σύστημα που είναι σε θέση να αναγνωρίζει τις διάφορες κατηγορίες απορριμμάτων ξύλου και να διαχωρίζει τα ξύλα από το πλαστικά. Αυτό το σύστημα δύναται να απαιτεί λιγότερο χρόνο για τη διαλογή των αποβλήτων και θα είναι πιο ακριβές στην ταξινόμηση από τον χειροκίνητο τρόπο. Το νέο σύστημα διαχωρισμού που προτείνεται στην παρούσα διπλωματική περιλαμβάνει οπτική διαλογή και διαλογή με βάση την πολυφασματική απεικόνιση. Συγκεκριμένα, η εκπαίδευση του μοντέλου θα βασίζεται σε δεδομένα που προέρχονται από δύο πολυφασματικούς αισθητήρες και μια βιομηχανική κάμερα. Με αυτόν τον τρόπο, η ταυτοποίηση, ο χαρακτηρισμός και η κατηγοριοποίηση των ξύλων θα γίνεται χωρίς κόπο με βάση τη σύστασή τους. Με την χρησιμοποίηση του συστήματος, τα διαχωρισμένα απόβλητα που βρίσκονται σε καλή κατάσταση θα μπορούν ακόμη να ανακυκλωθούν και να μετατραπούν σε ενέργεια ή καύσιμα για την ανάπτυξη της οικονομίας. Άλλωστε, η έννοια της κυκλικής οικονομίας συνδέεται στενά με τη διαχείριση των ξύλινων απορριμμάτων και μια τέτοια προσέγγιση περιλαμβάνει την εξεύρεση τρόπων επαναχρησιμοποίησης του ξύλου αντί της απλής απόρριψής του.

Επιπλέον, ένας ακόμη στόχος αυτής της έρευνας ήταν να εξετάσει το βαθμό στον οποίο ο υψηλής τεχνολογίας εξοπλισμός επηρεάζει την απόδοση του διαχωριστή. Στόχος μας ήταν να αξιολογήσουμε κατά πόσον η αξιοποίηση κορυφαίου εξοπλισμού θα απέφερε σημαντικές βελτιώσεις στη διαδικασία διαλογής. Αυτή η κατεύθυνση έρευνας ήταν ζωτικής σημασίας, διότι εάν ο εξοπλισμός υψηλής τεχνολογίας δεν παρουσίαζε σημαντική επιρροή στην απόδοση του διαχωριστή, αυτό θα σήμαινε ότι η διαχείριση των ξύλινων απορριμμάτων θα μπορούσε να γίνει προσιτή σε ένα ευρύτερο κοινό παγκοσμίως. Αυτό θα άνοιγε ευκαιρίες για περιοχές και κοινότητες με περιορισμένους πόρους να υιοθετήσουν παρόμοιες πρακτικές διαχείρισης αποβλήτων. Στόχος μας ήταν να αποκαλύψουμε τις δυνατότητες για οικονομικά αποδοτικές και επεκτάσιμες λύσεις στη διαχείριση αποβλήτων ξύλου, οι οποίες θα μπορούσαν να επηρεάσουν θετικά τις προσπάθειες περιβαλλοντικής βιωσιμότητας παγκοσμίως.

Το σύστημα που αναπτύχθηκε για τον διαχωρισμό των συσσωρευμένων αποβλήτων βασίζεται στην έννοια του συνελκτικού νευρωνικού δικτύου. Οι εξελίξεις στην τεχνητή νοημοσύνη, ιδίως στον τομέα της μηχανικής μάθησης και της όρασης υπολογιστών, έχουν βελτιώσει σημαντικά την ακρίβεια και την αποτελεσματικότητα της ταξινόμησης των απορριμμάτων, βοηθώντας έργα όπως αυτό. Με τη χρήση αλγορίθμων και νευρωνικών δικτύων που εκπαιδεύονται σε μεγάλα σύνολα δεδομένων από εικόνες αποβλήτων, οι μηχανές μπορούν πλέον να αναγνωρίζουν και να ταξινομούν διαφορετικούς τύπους υλικών που απορρίπτονται με μεγάλη ακρίβεια.

Διαχείριση Αποβλήτων

Η διαχείριση αποβλήτων περιλαμβάνει τις διαδικασίες και τις ενέργειες που είναι απαραίτητες για το χειρισμό των αποβλήτων καθ' όλη τη διάρκεια του κύκλου ζωής τους. Περιλαμβάνεται η συλλογή, η μεταφορά, η επεξεργασία και η απόθεση στη χωματερή των απορριμμάτων, καθώς και η εποπτεία και ο έλεγχος της διαδικασίας αυτής μαζί με οποιοδήποτε νομοθετικό, τεχνολογικό ή οικονομικό μηχανισμό σχετίζεται με τα απόβλητα.

Τα απόβλητα ενίοτε ενέχουν κίνδυνο για την ανθρώπινη υγεία. Τα θέματα υγείας μπορούν να συνδεθούν με ολόκληρη τη διαδικασία διαχείρισης απορριμμάτων. Ο χειρισμός των στερεών αποβλήτων μπορεί να προκαλέσει άμεσα προβλήματα υγείας, ενώ η κατανάλωση νερού και τροφίμων μπορεί να τα προκαλέσει έμμεσα. Η ανθρώπινη δραστηριότητα, όπως η εξόρυξη και η επεξεργασία βασικών πόρων, παράγει απόβλητα [2]. Η διαχείριση απορριμμάτων στοχεύει στη μείωση των αρνητικών συνεπειών των απορριμμάτων στο περιβάλλον, την ανθρώπινη υγεία, τους πόρους του πλανήτη και την αισθητική.

Τα αστικά στερεά απόβλητα, τα οποία παράγονται από βιομηχανική, εμπορική και οικιακή δραστηριότητα καταλαμβάνουν σημαντικό μέρος της διαχείρισης των απορριμμάτων. Υπάρχουν διαφορές στις τακτικές διαχείρισης απορριμμάτων μεταξύ ανεπτυγμένων και αναπτυσσόμενων χωρών, αστικών και αγροτικών περιοχών, οικιστικών και βιομηχανικών τομέων, ακόμη και εντός της ίδιας χώρας [3].

Αν και η αποτελεσματική διαχείριση των απορριμμάτων είναι ζωτικής σημασίας για τη δημιουργία βιώσιμων και βιώσιμων κοινοτήτων, πολλά αναπτυσσόμενα έθνη εξακολουθούν να δυσκολεύονται με αυτήν. Και ο λόγος πίσω από αυτό, σύμφωνα με μία έρευνα [4], είναι το κόστος της αποτελεσματικής διαχείρισης των απορριμμάτων που συνήθως αντιπροσωπεύει το 20% έως το 50% των δημοτικών προϋπολογισμών.

Οι πρώτες μονάδες αποτέφρωσης, ή «καταστροφείς», όπως έγιναν τότε γνωστοί, αναπτύχθηκαν ως απάντηση στην απότομη αύξηση των απορριμμάτων που έπρεπε να απορριφθούν. Ωστόσο, λόγω της τεράστιας ποσότητας στάχτης που δημιουργούσαν και του τρόπου με τον οποίο αυτή κυκλοφορούσε πάνω από τις γύρω περιοχές, εκείνες δέχθηκαν μεγάλη κριτική [5].

Ξύλινα Απορρίμματα

Ο όρος "ξύλινα απορρίμματα" αναφέρεται σε ένα σύνολο απορριμμάτων που περιλαμβάνει απορριπτόμενα προϊόντα ξύλου από διάφορες βιομηχανίες, συμπεριλαμβανομένων των τομέων συσκευασίας και επεξεργασίας ξύλου, κατεδάφισης και κατασκευής, καθώς και ιδιωτικών νοικοκυριών και κατασκευής σιδηροδρόμων. Η παραγωγή ενέργειας και τα νέα δυνητικά προϊόντα, συμπεριλαμβανομένων των χημικών, των βιοκαυσίμων και άλλων λιγνοκυτταρινικών υλικών, μπορούν επίσης να χρησιμοποιούν αυτά τα απορρίμματα ως δευτερεύουσα πηγή πρώτων υλών. Για παράδειγμα, λιπόφιλα και υδρόφιλα εκχυλίσματα που βρίσκονται στο φλοιό του ξύλου μπορούν να χρησιμοποιηθούν για τη δημιουργία προϊόντων υψηλής αξίας, όπως καλλυντικές ενώσεις ή φαρμακευτικά προϊόντα. Με παρόμοιο τρόπο, το βιοέλαιο που εξάγεται

από πόρους απορριμμάτων ξύλου είναι πολύ χρήσιμο ως επεκτατικό και τροποποιητικό για συνδετικά ασφάλτου, με βάση το πετρέλαιο, στη δημιουργία ασφαλτοστρωμένων δρόμων. Άλλες χρήσεις υψηλής προστιθέμενης αξίας για τα απορρίμματα ξύλου μπορεί να περιλαμβάνουν την παραγωγή βιοκαυσίμων και σύνθετων υλικών με βάση το ξύλο, τα οποία μπορούν να συμβάλουν στη μείωση των περιβαλλοντικών επιπτώσεων της βιομηχανίας της ξυλείας και στην κάλυψη της αυξανόμενης ζήτησης για ξύλο. Επομένως, η δασική βιομηχανία θα πρέπει να δώσει προτεραιότητα σε μέτρα που μειώνουν, ανακτούν και βελτιστοποιούν τη χρήση των απορριμμάτων ξύλου που παράγονται κατά τη συγκομιδή και την επεξεργασία.

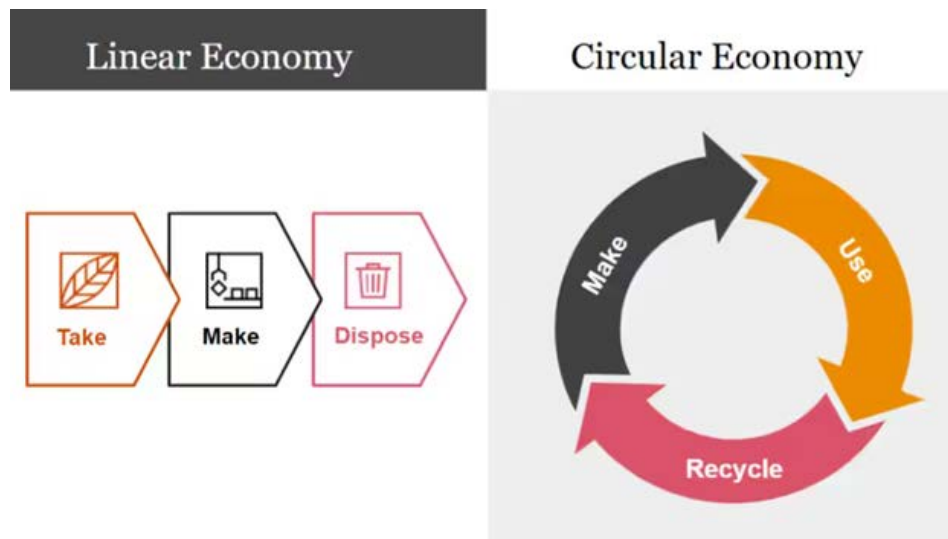
Τα απορρίμματα ξύλου και τα υποπροϊόντα που παράγονται από βιομηχανικές διεργασίες με βάση το ξύλο μπορούν να χρησιμοποιηθούν για τη δημιουργία ενός ευρέος φάσματος ωφέλιμων βιομηχανικών προϊόντων. Τα πριστήρια συγκεντρώνουν συνήθως 1 τόνο πριονίδι, ροκανίδια, χοντρές πλάκες, φέτες και γωνίες ανά 1000 πόδια ξυλείας που παράγονται. Περίπου το 75% αυτού του υλικού προς απόρριψη, αποτελείται από ξύλο και το υπόλοιπο 25% είναι φλοιός, ο οποίος μπορεί να μετατραπεί σε ενεργειακές ή μη εφαρμογές. Τα απόβλητα ξύλου μπορούν να είναι επωφελή όσον αφορά την παραγωγή ενέργειας μέσω πολλών μεθόδων όπως η καύση και η συμπαραγωγή. Από την άλλη πλευρά, τα απορρίμματα ξύλου μπορούν επίσης να χρησιμοποιηθούν για μη ενεργειακούς σκοπούς, όπως η κατασκευή σύνθετων σανίδων, προϊόντων κάλυψης επιφανειών, κομποστοποίησης και τσιμεντοσανίδων [6].

Η έρευνα έχει δείξει πολλούς τρόπους για την προστιθέμενη αξία, μετατρέποντας τα απορρίμματα ξύλου σε άλλα χρήσιμα προϊόντα. Για παράδειγμα, μια μελέτη που πραγματοποιήθηκε στη Φινλανδία έδειξε πολλές αυξανόμενες αγορές για προϊόντα με βάση το ξύλο, κυρίως υφάσματα, χημικά, βιοκαύσιμα και υποκατάστατα πλαστικών. Επιπλέον, μια μελέτη στη Ζιμπάμπουε έδειξε ότι η πλειονότητα των τεμαχίων και των ροκανιδιών από επιχειρήσεις με βάση το ξύλο χρησιμοποιούνται ως καυσόξυλα από τους ντόπιους και σε εμπορικά πριονιστήρια για τη δημιουργία ατμού για στεγνωτήρια κλιβάνου. Εν τω μεταξύ, ερευνητές στην Ιαπωνία αποκάλυψαν ότι οι εταιρείες επίπλων παρήγαγαν 15 εκατομμύρια κυβικά μέτρα απορριμμάτων ξύλου, από τα οποία πάνω από το 90% ανακυκλώθηκαν για την κατασκευή πάνελ και καυσίμων με βάση το ξύλο.

Κυκλική Οικονομία

Η κυκλική οικονομία είναι μια εναλλακτική λύση αντί μιας παραδοσιακής γραμμικής οικονομίας (κατασκευή, χρήση, απόρριψη) και επομένως ένα νέο μοντέλο παραγωγής και κατανάλωσης που προάγει τη βιώσιμη ανάπτυξη μακροπρόθεσμα. Υιοθετώντας αυτήν την προσέγγιση, μπορούμε να βελτιστοποιήσουμε τη χρήση των πόρων, να ελαχιστοποιήσουμε την κατανάλωση πρώτων υλών και να επαναχρησιμοποιήσουμε τα απόβλητα μέσω της ανακύκλωσης ή άλλων μέσων για τη δημιουργία νέων προϊόντων. Αν και αυτός ο όρος «Κυκλική Οικονομία» έχει πρόσφατα αποκτήσει μεγάλη δημοτικότητα, η έννοια υπάρχει από τη δεκαετία του 1970. Αρχικά προτάθηκε ως μέσο αντιμετώπισης ανησυχιών σχετικά με την ασφάλεια των πόρων και τη σπανιότητα σε τοπικό και εθνικό επίπεδο, καθώς και ως μέθοδος μείωσης των περιβαλλοντικών επιπτώσεων των πρακτικών παραγωγής και κατανάλωσης.

Ο στόχος της κυκλικής οικονομίας είναι να μεγιστοποιήσει τη χρησιμότητα των διαθέσιμων πόρων ακολουθώντας τρεις θεμελιώδεις αρχές: μείωση, επαναχρησιμοποίηση και ανακύκλωση. Αυτή η μεθοδολογία συμβάλλει στην παράταση του κύκλου ζωής των προϊόντων, στην ελαχιστοποίηση των απορριμμάτων και στην καθιέρωση μιας πιο αποτελεσματικής και οικολογικής μεθόδου παραγωγής με την πάροδο του χρόνου. Η κυκλική οικονομία εμπνέεται από τη φύση, όπου κάθε αγαθό είναι πολύτιμο και τα απόβλητα μετατρέπονται σε νέο πόρο, επιτυγχάνοντας έτσι μια ισορροπία μεταξύ προόδου και βιωσιμότητας.



Εικόνα 2. Γραφική αναπαράσταση της Γραμμικής Οικονομίας σε σύγκριση με την Κυκλική.

Η συσσώρευση μη ανακυκλώσιμων απορριμμάτων σε χώρους υγειονομικής ταφής σε όλο τον κόσμο αποτελεί σημαντική απειλή για τον τρόπο ζωής μας στο εγγύς μέλλον, εάν η γενιά μας δεν ενεργήσει γρήγορα για να την αποτρέψει. Αυτά τα απόβλητα μπορούν να αυξήσουν την εξάπλωση ασθενειών μέσω φορέων όπως τα έντομα και πιο συγκεκριμένα οι μύγες και τα κουνούπια. Επιπλέον, η συσσώρευση απορριμμάτων μπορεί να καταστρέψει τη φυσική ομορφιά των οικοτόπων, να οδηγήσει σε αποψίλωση των δασών και ρύπανση του εδάφους και των υδάτων λόγω της παρουσίας τοξικών χημικών ουσιών σε ακατάλληλα επεξεργασμένα υλικά. Αυτή η ρύπανση μπορεί να διαταράξει την τροφική αλυσίδα, οδηγώντας σε περισσότερες ασθένειες και προβλήματα υγείας για τους ανθρώπους και τα οικοσυστήματα παγκοσμίως.

Υπάρχουν τρεις βασικοί λόγοι για τους οποίους η συσσώρευση απορριμμάτων έχει γίνει πιο σοβαρή απειλή τα τελευταία 50 χρόνια. Ως πρώτος λόγος είναι η έλλειψη διαθεσιμότητας στην αγορά ανακυκλώσιμων ειδών, παρά τις προσπάθειες των εταιρειών να αναπτύξουν πιο βιώσιμα και φιλικά προς το περιβάλλον προϊόντα. Ο δεύτερος λόγος είναι ο υπερπληθυσμός, ο οποίος δείχνει την υλικοτεχνική πρόκληση κατά τη διαχείριση της παραγωγής απορριμμάτων που έχει ως αποτέλεσμα υψηλότερο ποσοστό αγαθών να καταλήγουν σε χωματερές ή ακόμα και στον ωκεανό, επηρεάζοντας τη ζωή των θαλάσσιων ειδών. Τέλος, ο τρίτος λόγος είναι η έλλειψη συμμετοχής της κοινωνίας στην αντιμετώπιση ζητημάτων όπως η κλιματική αλλαγή και η συσσώρευση αποβλήτων.

Για να φανεί το ζήτημα της συσσώρευσης αποβλήτων μέσα από δεδομένα, είναι σημαντικό να σημειωθεί ότι ο παγκόσμιος πληθυσμός παράγει μεταξύ 7 και 9 δισεκατομμυρίων τόνων απορριμμάτων ετησίως, εκ των οποίων το 70% δεν διαχειρίζεται σωστά και καταλήγει σε χώρους υγειονομικής ταφής, θέτοντας κίνδυνο ρύπανσης του φυσικού περιβάλλοντος και τη δημιουργία νέων κινδύνων για την υγεία, όπως τα μικροπλαστικά των ωκεανών. Στα δεδομένα αυτά περιλαμβάνονται όλα τα χρησιμοποιημένα, ανεπιθύμητα και απορριπτόμενα αντικείμενα που δημιουργούνται από ανθρώπους. Ωστόσο, υπάρχει διαφορά μεταξύ των συνολικών παραγόμενων αποβλήτων και των αστικών στερεών αποβλήτων (ΑΣΑ), τα οποία αποτελούνται μόνο από τα σκουπίδια που παράγονται σε αστικές περιοχές ή στις γύρω περιοχές. Περίπου 2 δισεκατομμύρια τόνοι ΑΣΑ παράγονται κάθε χρόνο, με περίπου το 33% αυτών να μην υφίσταται κατάλληλη επεξεργασία. Αυτό σημαίνει ότι κάθε άτομο παράγει από 0,1 έως 4,5 κιλά απορριμμάτων την ημέρα, με μέσο όρο 0,7 κιλά. Επιπλέον, αναμένεται ότι τα ΑΣΑ θα αυξηθούν σε 3,4 δισεκατομμύρια τόνους μέχρι το 2050 λόγω του ταχέως αυξανόμενου παγκόσμιου πληθυσμού και της ανάγκης να χρησιμοποιηθούν εκτενώς οι φυσικοί πόροι για τη βιομηχανική ανάπτυξη και τη διατήρηση του πολιτισμού μας.

Ξύλινα Απορρίμματα και Κυκλική Οικονομία

Το ξύλο είναι ένα υλικό φυσικό, ανανεώσιμο και βιοδιασπώμενο και έχει εξαιρετικές μηχανικές και θερμικές ιδιότητες. Οι περιβαλλοντικές επιπτώσεις κατά τα στάδια παραγωγής και τέλους ζωής των ξύλινων υλικών είναι γενικά λιγότερο επιβλαβείς από εκείνες των ισοδύναμων ανόργανων υλικών ή υλικών με βάση τα ορυκτά. Η έννοια της κυκλικής οικονομίας συνδέεται στενά με τη διαχείριση και τη χρήση απορριμμάτων ξύλου και μια προσέγγιση αυτού του είδους θα συνεπαγόταν την εύρεση τρόπων επαναχρησιμοποίησης ή ανακύκλωσης του ξύλου αντί απλώς της απόρριψής του [7].

Αυτό θα μπορούσε να περιλαμβάνει την επαναχρησιμοποίησή του ως πρώτη ύλη για νέα προϊόντα με βάση το ξύλο ή τη χρήση του ως πηγή καυσίμου, όπως για την παραγωγή ενέργειας από βιομάζα. Με την επαναχρησιμοποίηση ή την ανακύκλωση απορριμμάτων ξύλου, μπορούμε να μειώσουμε την ποσότητα των απορριμμάτων που καταλήγουν στις χωματερές και να ελαχιστοποιήσουμε την ανάγκη για νέες πρώτες ύλες, που συμβάλλουν στη διατήρηση των φυσικών πόρων και στη μείωση των περιβαλλοντικών επιπτώσεων της εξόρυξης και επεξεργασίας πόρων. Επιπλέον, η επαναχρησιμοποίηση και η ανακύκλωση απορριμμάτων ξύλου μπορεί επίσης να δημιουργήσει οικονομικές ευκαιρίες, όπως το άνοιγμα νέων θέσεων εργασίας στη βιομηχανία ανακύκλωσης και επαναχρησιμοποίησης. Λαμβάνοντας υπόψη τα παραπάνω, η προσέγγιση της κυκλικής οικονομίας είναι ένας σημαντικός τρόπος διαχείρισης των απορριμμάτων ξύλου με πιο βιώσιμο και προσεκτικό τρόπο.

Κατηγορίες Ξύλινων Απορριμάτων

Για τους σκοπούς της παρούσας διπλωματικής εργασίας θα περιορίσουμε τις τέσσερις επίσημες κατηγορίες στις οποίες χωρίζονται τα ξύλινα απορρίμματα από την Ένωση Ανακυκλωτών Ξύλου (Wood Recyclers' Associations – WRA)[8] σε τρεις, που είναι μια επίσης συχνή κατηγοριοποίηση των αποβλήτων ξύλου, δεδομένου ότι οι κατηγορίες Β και Γ έχουν πολλές ομοιότητες:

- Ως μία κατηγορία (κατηγορία με αριθμό 1), θα έχουμε το ακατέργαστο καθαρό μη επεξεργασμένο ξύλο.
- Ως μία κατηγορία (κατηγορία αριθμού 2), θα θεωρήσουμε το επεξεργασμένο μαζί με το μη επικίνδυνο ξύλο.
- Ως μία κατηγορία (κατηγορία με αριθμό 3) θα είναι τα επικίνδυνα απόβλητα ξύλου.

Τα πειράματά μας διεξήχθησαν με τρεις διαφορετικές κλάσεις απορριμάτων ξύλου, καθώς αυτές οι τρεις συγκεντρώνουν το μεγαλύτερο ποσοστό δημοτικότητας στην ελληνική αγορά ξύλου και έχουν επιπλέον ερευνητικό ενδιαφέρον λόγω της σύστασης τους:

- **Κλάση 1, μέσης πυκνότητας Ινοσανίδες (MDF):** Η ινοσανίδα μέσης πυκνότητας είναι ένα κατασκευασμένο προϊόν ξύλου που δημιουργείται από τη μετατροπή απορριμάτων σκληρής ή μαλακής ξυλείας σε ίνες ξύλου, συχνά σε μηχανή αποϊνωσης, που αναμιγνύονται με κεριά και συνδετικό υλικό ρητίνης και υποβάλλονται σε πάνελ με υψηλή θερμοκρασία και πίεση. Σε σύγκριση με το κόντρα πλακέ, το MDF έχει συνήθως υψηλότερη πυκνότητα. Αν και αποτελείται από διαχωρισμένες ίνες, μπορεί να χρησιμοποιηθεί σε κατασκευαστικές εφαρμογές παρόμοιες με το κόντρα πλακέ. Στην πραγματικότητα, το MDF είναι ισχυρότερο και ανθεκτικότερο από τη μοριοσανίδα [9].

- **Κλάση 2, Μοριοσανίδα με επένδυση μελαμίνης (MFC):** Οι μοριοσανίδες με επένδυση μελαμίνης, είναι ένας τύπος σύνθετου ξύλου που κατασκευάζεται με την ανάμειξη μικρών κομματιών επεξεργασμένου ξύλου με ισχυρή κόλλα (ρητίνη). Αυτός ο τύπος μοριοσανίδας είναι εξαιρετικά ευέλικτος και μπορεί να χρησιμοποιηθεί σε διάφορες εφαρμογές, όπως σε ντουλάπια μπάνιου και κουζίνας, έπιπλα γραφείου, σκάφη σε κανάλια, τροχόσπιτα και άλλες εγκαταστάσεις. Είναι αρκετά εύκολο να δουλέψει κανείς με αυτό και παρέχει ένα ελκυστικό φινιρίσμα στην επιφάνεια. Με τις σύγχρονες μεθόδους παραγωγής, οι μοριοσανίδες με επένδυση μελαμίνης είναι όλο και πιο δύσκολο να διακριθούν από το φυσικό ξύλο, καθώς η τεχνολογία επιτρέπει τη δημιουργία φύλλων που μιμούνται την εμφάνιση του πραγματικού ξύλου.

- **Κλάση 3, Καπλαμάς δρυός:** Ο καπλαμάς δρυός είναι μια λεπτή φέτα ξύλου δρυός που χρησιμοποιείται για να καλύψει άλλα υλικά, όπως κόντρα πλακέ ή μοριοσανίδες, για να δημιουργήσει την εμφάνιση συμπαγούς βελανιδιάς. Οι καπλαμάδες παράγονται συνήθως με τεμαχισμό ενός κορμού σε λεπτά φύλλα, τα οποία στη συνέχεια επικολλώνται σε ένα υλικό υποστρώματος με κόλλα. Ο καπλαμάς δρυός είναι δημοφιλής στην κατασκευή επίπλων και στη διακόσμηση εσωτερικών χώρων, καθώς προσφέρει την ομορφιά του ξύλου δρυός, δεδομένου ότι παρέχει μια συνεπή και ομοιόμορφη εμφάνιση, με χαμηλότερο κόστος από το μασίφ δρυ. Ο καπλαμάς δρυός διατίθεται σε διάφορες ποιότητες και πάχη και μπορεί να βαφτεί ή να φινιριστεί ώστε να ταιριάζει με ποικίλες σχεδιαστικές προτιμήσεις. Ωστόσο, δεν είναι τόσο ανθεκτικός ή μακράς διάρκειας όσο η συμπαγής βελανιδιά, καθώς το στρώμα καπλαμά είναι σχετικά λεπτό και μπορεί να είναι επιρρεπές σε ζημιές ή φθορά με την πάροδο του χρόνου.

Στην κατάταξη αυτών των κομματιών ξύλου στις τρεις προαναφερθείσες κατηγορίες, με βάση τον τρόπο παραγωγής τους, θα είχαμε τον καπλαμά δρυός να ανήκει στην κατηγορία 1, ενώ τα MDF και MFC να ανήκουν στην κατηγορία 2.

Πολυφασματική Απεικόνιση

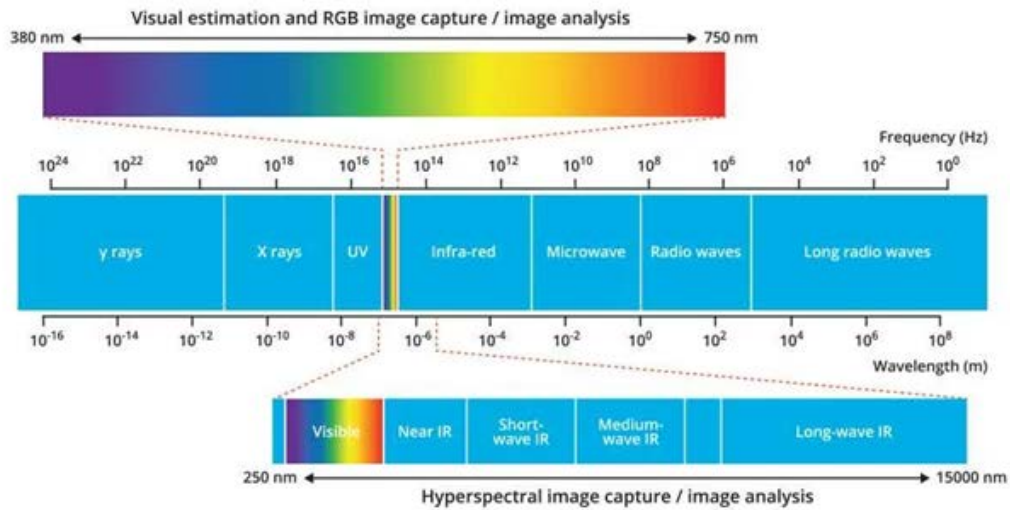
Η πολυφασματική απεικόνιση είναι μια προηγμένη τεχνολογία που συνδυάζει τη φασματοσκοπία με την απεικόνιση για τη συλλογή εξαντλητικών πληροφοριών σχετικά με τη σύνθεση και τα χαρακτηριστικά των υλικών και των επιφανειών. Αυτή η τεχνολογία είναι πιο ισχυρή από τα παραδοσιακά συστήματα απεικόνισης και παρέχει μη επεμβατικά και μη καταστροφικά μέσα αναγνώρισης και ποσοτικοποίησης αντικειμένων. Λόγω της ευελιξίας και της χρησιμότητάς της, η πολυφασματική απεικόνιση γίνεται όλο και πιο δημοφιλής σε μια ποικιλία βιομηχανιών και ερευνητικών έργων.

Η πολυφασματική απεικόνιση λειτουργεί συλλέγοντας και αναλύοντας πληροφορίες από όλο το ηλεκτρομαγνητικό φάσμα για να σχηματίσει το φάσμα - συγκρίσιμο με τα δακτυλικά αποτυπώματα - για κάθε εικονοστοιχείο (pixel) μιας εικόνας, επιτρέποντας την αναγνώριση αντικειμένων και υλικών από τις μοναδικές φασματικές υπογραφές τους, παρόμοιο με τον τρόπο με τον οποίο τα δακτυλικά αποτυπώματα χρησιμοποιούνται για την αναγνώριση ενός ατόμου. Η πολυφασματική απεικόνιση έχει πολλές πρακτικές εφαρμογές που συνίστανται στη βελτίωση της ποιότητας και της ασφάλειας των τροφίμων, στη διευκόλυνση της διαλογής και της ανακύκλωσης των απορριμμάτων και στην ενίσχυση του ελέγχου και της παρακολούθησης στη φαρμακευτική παραγωγή.

Το ηλεκτρομαγνητικό φάσμα περιλαμβάνει όλους τους τύπους φωτός, που κυμαίνονται από ραδιοκύματα έως ακτίνες γάμμα, οι περισσότεροι εκ των οποίων δεν μπορούν να ανιχνευθούν από το ανθρώπινο μάτι (Εικόνα 3).

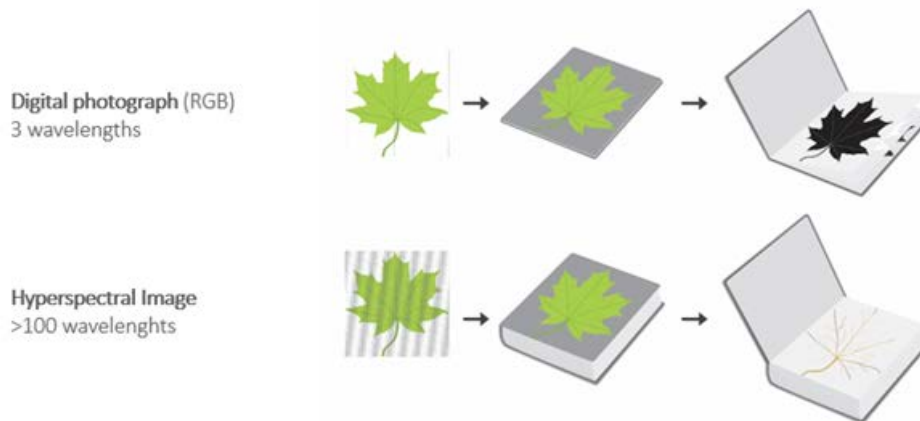
Φασματική Απεικόνιση

Η φασματική απεικόνιση χρησιμοποιεί διάφορες ζώνες (bands) σε όλο το ηλεκτρομαγνητικό φάσμα για την απεικόνιση. Σε αντίθεση με την κάμερα Κόκκινου-Πράσινου-Μπλε (RGB) που χρησιμοποιεί μόνο τις τρεις ζώνες ορατού φωτός (κόκκινο, πράσινο και μπλε), η υπερφασματική απεικόνιση μπορεί να εξερευνήσει πώς αλληλεπιδρούν τα αντικείμενα με ένα πολύ μεγαλύτερο εύρος ζωνών, που εκτείνονται από 250 nm έως 15.000 nm και επιπρόσθετα των θερμικών υπέρυθρων. Η έρευνα της αλληλεπίδρασης φωτός-ύλης είναι γνωστή ως φασματοσκοπία ή φασματική ανίχνευση.



Εικόνα 3. Η υπερφασματική απεικόνιση καταγράφει μήκη κύματος από 250 nm έως 15.000 nm και θερμικές υπέρυθρες.

Αυτά τα συστήματα είναι ικανά να συλλαμβάνουν πολλαπλά κανάλια ή ζώνες πληροφοριών σε όλο το ηλεκτρομαγνητικό φάσμα. Έτσι, αυτές οι πληροφορίες μπορούν να υποβληθούν σε επεξεργασία για να δημιουργήσουν μια χρωματικά κωδικοποιημένη αναπαράσταση των φασματικών δεδομένων, παρέχοντας πληροφορίες για τις χημικές και φυσικές ιδιότητες των αντικειμένων στην εικόνα.



Εικόνα 4. Προκειμένου να αντιγραφεί η λειτουργία της ανθρώπινης όρασης, δημιουργείται μια ψηφιακή φωτογραφία ενός φύλλου (επάνω) χρησιμοποιώντας τρεις ζώνες: κόκκινο, πράσινο και μπλε. Τα RGB δεδομένα είναι συγκρίσιμα με ένα φυλλάδιο τριών σελίδων. Αντίθετα, μια υπερφασματική εικόνα του ίδιου φύλλου (κάτω) καταγράφει μια φασματική απόκριση από 220 μήκη κύματος, με αποτέλεσμα μια πιο εκτεταμένη και λεπτομερή αναπαράσταση του αντικειμένου, παρόμοια με ένα βιβλίο 220 σελίδων.

Πώς λειτουργεί η Πολυφασματική Απεικόνιση;

Η πολυφασματική απεικόνιση κάνει χρήση ενός φασματομέτρου απεικόνισης, γνωστό και ως πολυφασματική κάμερα, για τη συλλογή φασματικών πληροφοριών. Αυτή η κάμερα συλλαμβάνει το φως μιας σκηνής χωρισμένο σε μεμονωμένα μήκη κύματος ή φασματικές ζώνες, δημιουργώντας μια δισδιάστατη εικόνα που καταγράφει επίσης τις φασματικές πληροφορίες κάθε εικονοστοιχείου. Το αποτέλεσμα είναι μια πολυφασματική εικόνα, με κάθε εικονοστοιχείο να αντιπροσωπεύει ένα μοναδικό φάσμα.

Η πολυφασματική απεικόνιση παρέχει τόσο χωρικές όσο και φασματικές πληροφορίες, επιτρέποντας την αναγνώριση και ταξινόμηση των υλικών, καθώς και τη χωρική κατανομή και τον τοπικό διαχωρισμό αυτών των υλικών.

Τα Πλεονεκτήματα της Πολυφασματικής Απεικόνισης

Το κύριο πλεονέκτημα της πολυφασματικής απεικόνισης έγκειται στην ικανότητά της να παρέχει ακριβή χαρακτηρισμό υλικού, χάρη στην υψηλή χωρική και φασματική της ανάλυση. Καταγράφοντας δεκάδες ή και εκατοντάδες φάσματα, μια πολυφασματική κάμερα μπορεί να κατασκευάσει έναν τεράστιο πολυφασματικό κύβο δεδομένων που περιέχει δεδομένα θέσης, μήκους κύματος και χρόνου. Σε σύγκριση με τη συμβατική απεικόνιση, η πολυφασματική παρέχει σημαντικότερο όγκο πληροφοριών, επιτρέποντας ακριβέστερη ανάλυση, ταυτοποίηση και διαφοροποίηση υλικών και ουσιών.

Με την πολυφασματική απεικόνιση, είναι δυνατό να διακρίνουμε υλικά που έχουν παρόμοιες φυσικές ή οπτικές ιδιότητες ή εκείνα που είναι αόρατα στο ανθρώπινο μάτι, όπως διάφορα ορυκτά.

Πεδία εφαρμογών της Πολυφασματικής Απεικόνισης

Η πολυφασματική απεικόνιση είναι μια τεχνική που γίνεται όλο και πιο διαδεδομένη στη βιομηχανία, την έρευνα και την τηλεπισκόπηση. Αυτή η μέθοδος παρέχει δεδομένα που μπορούν να χρησιμοποιηθούν για τον εντοπισμό, την κατηγοριοποίηση ή τη μέτρηση της συγκέντρωσης διαφόρων ουσιών που δεν είναι ανιχνεύσιμες με κανονικές κάμερες ή με γυμνό μάτι. Για παράδειγμα, ένα εν σειρά σύστημα ποιοτικού ελέγχου που περιλαμβάνει ένα σύστημα πολυφασματικής απεικόνισης μπορεί να ανιχνεύσει ξένα αντικείμενα, ρύπους, καθώς και τα επίπεδα λίπους, ζάχαρης ή υγρασίας στα προϊόντα. Επιπλέον, η πολυφασματική ή υπερφασματική απεικόνιση που λαμβάνεται μέσω τηλεπισκόπησης παρέχει πολύτιμα δεδομένα για τις επιφάνειες της Γης, συμπεριλαμβανομένων των ορυκτών και της βλάστησης.

Η πολυφασματική απεικόνιση έχει ένα ευρύ φάσμα εφαρμογών, όπως:

- **Περιβαλλοντική παρακολούθηση:** Παρατήρηση αλλαγών στη χρήση γης, την υγεία της βλάστησης και την ποιότητα του νερού με την πάροδο του χρόνου για τον εντοπισμό πρώιμων ενδείξεων οικολογικής υποβάθμισης και την παρακολούθηση των προσπάθειών συντήρησης.
- **Εξερεύνηση ορυκτών:** Χαρτογράφηση κοιτασμάτων ορυκτών, ανίχνευση της σύνθεσης και της ποιότητάς τους.
- **Ποιοτικός έλεγχος:** Αξιολόγηση προϊόντων διατροφής, ανίχνευση ρύπων και ελαττωμάτων σε βιομηχανικά προϊόντα με μη καταστροφικό τρόπο.

- **Διαχείριση απορριμμάτων:** Διαχωρισμός διαφορετικών υλικών με ακρίβεια και αποτελεσματικότητα για τη βελτιστοποίηση των διαδικασιών ανακύκλωσης και την ενίσχυση της αξίας των ανακυκλωμένων υλικών.
- **Γεωργία:** Εκτίμηση της υγείας και της απόδοσης των καλλιεργειών, παρακολούθηση της υγρασίας του εδάφους και της περιεκτικότητας σε θρεπτικά συστατικά για τη βελτίωση των πρακτικών διαχείρισης των καλλιεργειών και την ενίσχυση των αποδόσεων τους. Συνήθως, η παρακολούθηση της γεωργίας χρησιμοποιεί υπερφασματικές κάμερες τοποθετημένες σε μη επανδρωμένα αεροχήματα (drones).
- **Στρατιωτική παρακολούθηση:** Ανίχνευση επικίνδυνων υλικών.

Αυτές είναι μόνο μερικές από τις χρήσιμες πληροφορίες που μπορεί να παρέχει αυτό το είδος απεικόνισης για διάφορους σκοπούς. Συμπερασματικά, η πολυφασματική απεικόνιση είναι μια πολύτιμη τεχνολογία μηχανικής όρασης για πολλαπλές βιομηχανίες και έρευνες σχετικά με τη βελτίωση των διαδικασιών και της ποιότητας τους και τη μείωση των απορριμμάτων.

Τεχνητή Νοημοσύνη

Η έρευνα για την Τεχνητή Νοημοσύνη (TN) ξεκίνησε από ένα εργαστήριο που έλαβε χώρα στο Dartmouth College το 1956 [10], όπου οι συμμετέχοντες έγιναν τελικά οι πρωτοπόροι και οι ηγέτες του τομέα. Αυτοί, μαζί με τους μαθητές τους, ανέπτυξαν προγράμματα που θεωρήθηκαν «εκπληκτικά» από τα μέσα, καθώς οι υπολογιστές ήταν σε θέση να μάθουν στρατηγικές για σκάκι, να λύσουν προβλήματα λέξεων στην άλγεβρα, να αποδείξουν λογικά θεωρήματα, ακόμη και να μιλούν αγγλικά. Αυτό οδήγησε στην ίδρυση ερευνητικών εργαστηρίων τεχνητής νοημοσύνης παγκοσμίως μέχρι τα μέσα της δεκαετίας του 1960 [11].

Η πρόοδος στη μηχανική μάθηση και αντίληψη κατέστη δυνατή χάρη στους ταχύτερους υπολογιστές, τις βελτιώσεις στους αλγόριθμους και την πρόσβαση σε τεράστιες ποσότητες δεδομένων. Οι μέθοδοι βαθιάς μάθησης, οι οποίες βασίζονται σε μεγάλο βαθμό στα δεδομένα, άρχισαν να πληρούν στόχους ακρίβειας γύρω στο 2012. Το 2015 [12], η τεχνητή νοημοσύνη είχε μια σημαντική χρονιά, με τον αριθμό των έργων λογισμικού που χρησιμοποιούσαν TN στην Google να εκτοξεύεται από μια «σποραδική χρήση» το 2012 σε πάνω από 2.700 έργα. Αυτό το κύμα πιστώθηκε στην αυξημένη διαθεσιμότητα προσιτών νευρωνικών δικτύων, χάρη στον πολλαπλασιασμό της υποδομής υπολογιστικού νέφους και την επέκταση των ερευνητικών εργαλείων και συνόλων δεδομένων [13].

Η τεχνητή νοημοσύνη αναφέρεται στην αναπαραγωγή διαδικασιών ανθρώπινης νοημοσύνης από μηχανές, ιδιαίτερα συστήματα υπολογιστών. Η TN υλοποιείται μέσω διαφόρων εφαρμογών όπως αναγνώριση ομιλίας, επεξεργασία φυσικής γλώσσας και μηχανική όραση [14].

Καθώς ο ενθουσιασμός γύρω από την τεχνητή νοημοσύνη έχει ενταθεί, οι εταιρείες ήταν πρόθυμες να προωθήσουν τον τρόπο με τον οποίο τα προϊόντα και οι υπηρεσίες τους τη χρησιμοποιούν. Ωστόσο, αυτό που χαρακτηρίζουν ως TN, συχνά είναι απλώς ένα συστατικό της τεχνολογίας, όπως η μηχανική μάθηση. Η τεχνητή νοημοσύνη απαιτεί εξειδικευμένο υλικό και λογισμικό για την ανάπτυξη και την εκπαίδευση αλγορίθμων μηχανικής εκμάθησης. Αν και δεν

υπάρχει συγκεκριμένη γλώσσα προγραμματισμού που να αντιπροσωπεύει αποκλειστικά τη TN, οι Python, R, Java, C++ και Julia είναι δημοφιλείς επιλογές μεταξύ των προγραμματιστών.

Κυρίως, τα συστήματα TN λειτουργούν καταναλώνοντας σημαντικές ποσότητες δεδομένων εκπαίδευσης με επικέτα, αναλύοντας τα δεδομένα για μοτίβα και συσχετίσεις και χρησιμοποιώντας αυτά τα μοτίβα για να κάνουν προβλέψεις για μελλοντικά γεγονότα. Για παράδειγμα, ένα διαλογικό ρομπότ (chatbot) που είναι εκπαιδευμένο με διάφορα δείγματα κειμένου μπορεί να μάθει να δημιουργεί ρεαλιστικές συνομιλίες με ανθρώπους ή ένα εργαλείο αναγνώρισης εικόνων μπορεί να αναγνωρίσει και να περιγράψει αντικείμενα σε εικόνες εξετάζοντας εκατομμύρια παραδείγματα. Οι νέες τεχνικές δημιουργικής τεχνητής νοημοσύνης προχωρούν γρήγορα, επιτρέποντας τη παραγωγή ρεαλιστικού κειμένου, εικόνων, μουσικής και άλλων μέσων.

Ο προγραμματισμός με TN επικεντρώνεται σε γνωστικές δεξιότητες που περιλαμβάνουν τα ακόλουθα:

- **Μάθηση:** Ασχολείται με τη λήψη δεδομένων και τη θέσπιση κανόνων ή αλγορίθμων για την επεξεργασία τους σε πρακτικές πληροφορίες που πρέπει να ακολουθήσουν οι υπολογιστικές συσκευές.
- **Συλλογισμός:** Εστιάζει στην επιλογή του κατάλληλου αλγορίθμου για να φτάσει σε ένα συγκεκριμένο αποτέλεσμα.
- **Αυτοδιόρθωση:** Στοχεύει στη συνεχή βελτίωση των αλγορίθμων για την παροχή των πιο ακριβών αποτελεσμάτων.
- **Δημιουργικότητα:** Χρησιμοποιεί διαφορετικές τεχνικές τεχνητής νοημοσύνης, συμπεριλαμβανομένων νευρωνικών δικτύων, συστημάτων που βασίζονται σε κανόνες και στατιστικές μεθόδους, για την παραγωγή νέων εικόνων, κειμένου, μουσικής και ιδεών.

Η τεχνητή νοημοσύνη έχει τη δυνατότητα να επιφέρει σημαντικές αλλαγές στον τρόπο ζωής, την εργασία και τις ψυχαγωγικές μας δραστηριότητες. Στις επιχειρήσεις, η τεχνητή νοημοσύνη ήταν επιτυχής στην αυτοματοποίηση διαφόρων εργασιών που παραδοσιακά εκτελούνταν από ανθρώπους, όπως η εργασία εξυπηρέτησης πελατών, ο εντοπισμός απάτης και ο ποιοτικός έλεγχος. Η τεχνητή νοημοσύνη μπορεί να εκτελέσει ορισμένες εργασίες πιο αποτελεσματικά από τους ανθρώπους, ειδικά επαναλαμβανόμενες και προσανατολισμένες στη λεπτομέρεια, όπως η ανάλυση νομικών εγγράφων. Με τη δυνατότητα επεξεργασίας ογκωδών συνόλων δεδομένων, η τεχνητή νοημοσύνη μπορεί να παρέχει πολύτιμες πληροφορίες για τις επιχειρηματικές λειτουργίες. Η ταχεία επέκταση των εργαλείων παραγωγής τεχνητής νοημοσύνης αναμένεται επίσης να έχει σημαντικό αντίκτυπο σε τομείς όπως η εκπαίδευση, το μάρκετινγκ και ο σχεδιασμός προϊόντων.

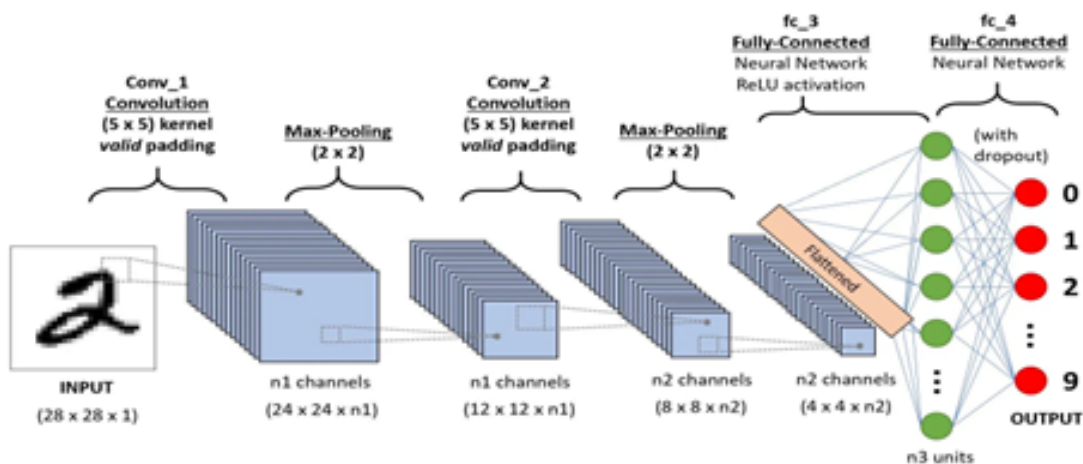
Η εφαρμογή τεχνικών τεχνητής νοημοσύνης όχι μόνο οδήγησε σε αυξημένη αποτελεσματικότητα, αλλά άνοιξε επίσης νέες επιχειρηματικές ευκαιρίες, όπως η χρήση της τεχνητής νοημοσύνης από τις εταιρείες συνεπιβατισμού και των ταξί, για τη σύνδεση των αναβατών με τα οχήματα τους [15].

Πολλές επιτυχημένες εταιρείες, συμπεριλαμβανομένων των Alphabet, Apple, Microsoft και Meta, έχουν υιοθετήσει τεχνολογίες TN για να βελτιώσουν τις δραστηριότητές τους και να αποκτήσουν ανταγωνιστικό πλεονέκτημα. Η Google, για παράδειγμα, χρησιμοποιεί τεχνητή νοημοσύνη στη μηχανή αναζήτησής της και τα αυτόνομα αυτοκίνητα της. Ενώ, επίσης, η τεχνητή

νοημοσύνη είναι υπεύθυνη για τις ανακαλύψεις της Google στην επεξεργασία της φυσικής γλώσσας.

Συνελικτικά Νευρωνικά Δίκτυα (ΣΝΔ)

Διαφορετικά είδη νευρωνικών δικτύων χρησιμοποιούνται για διαφορετικές περιπτώσεις και σύνολα δεδομένων. Πριν από την έλευση των Συνελικτικών Νευρωνικών Δικτύων (ΣΝΔ), χρησιμοποιούνταν χειροκίνητες, χρονοβόρες τεχνικές εξαγωγής χαρακτηριστικών για τον εντοπισμό αντικειμένων σε εικόνες. Τώρα, τα συνελικτικά νευρωνικά δίκτυα προσφέρουν μια πιο κλιμακούμενη προσέγγιση σε εφαρμογές ταξινόμησης εικόνων και αναγνώρισης αντικειμένων. Το επιτυγχάνουν αυτό αξιοποιώντας αρχές από τη γραμμική άλγεβρα, ιδίως τον πολλαπλασιασμό πινάκων, για τον εντοπισμό μοτίβων σε μια εικόνα. Παρά την αδιαμφισβήτητη χρησιμότητά τους, μπορεί να είναι πολύ απαιτητικά από υπολογιστική άποψη, απαιτώντας μονάδες γραφικής επεξεργασίας (GPU) για την εκπαίδευση των μοντέλων [16].



Εικόνα 5. Μια ακολουθία ΣΝΔ για την ταξινόμηση χειρόγραφων χαρακτήρων.

Ένα ΣΝΔ είναι ένας αλγόριθμος βαθιάς μάθησης που είναι ικανός να λαμβάνει μια εικόνα ως είσοδο, να αποδίδει σημαντικότητα (μαθησιακά βάρη και πολώσεις) σε διάφορα αντικείμενα/πτυχές της εικόνας και να διακρίνει το ένα από το άλλο. Οι απαιτήσεις προεργασίας ενός ΣΝΔ είναι πολύ χαμηλότερες σε σύγκριση με άλλους αλγορίθμους ταξινόμησης. Αν και τα φίλτρα είναι φτιαγμένα από τον άνθρωπο με πρωτόγονες μεθόδους, τα ΣΝΔ μπορούν να μάθουν αυτά τα φίλτρα/χαρακτηριστικά με επαρκή εκπαίδευση [17].

Η αρχιτεκτονική ενός ΣΝΔ είναι παρόμοια με το μοτίβο συνδεσιμότητας των νευρώνων στον ανθρώπινο εγκέφαλο και εμπνεύστηκε από την οργάνωση του οπτικού φλοιού. Στο οπτικό πεδίο, οι μεμονωμένοι νευρώνες αντιδρούν σε ερεθίσματα μόνο σε μια περιορισμένη περιοχή γνωστή ως πεδίο υποδοχής (receptive field). Ένα σύνολο τέτοιων πεδίων επικαλύπτεται-υπερδιπλώνεται για να καλύψει ολόκληρη την οπτική περιοχή.

Τα συνελικτικά νευρωνικά δίκτυα είναι ιδιαίτερα αναγνωρισμένα για την ανώτερη απόδοσή τους με εισόδους εικόνας, ομιλίας ή ηχητικού σήματος σε σύγκριση με άλλα νευρωνικά δίκτυα. Διαθέτουν τρεις κύριους τύπους επιπέδων, οι οποίοι είναι οι εξής:

- **Συνελικτικό επίπεδο**
- **Επίπεδο ομαδοποίησης**
- **Πλήρως συνδεδεμένο (ΠΣ) επίπεδο**

Το συνελικτικό επίπεδο είναι το αρχικό επίπεδο ενός συνελικτικού δικτύου, το οποίο μπορούν να διαδεχθούν άλλα συνελικτικά επίπεδα ή επίπεδα ομαδοποίησης. Αντίθετα, το πλήρως συνδεδεμένο επίπεδο είναι το τελικό. Καθώς προστίθεται κάθε επίπεδο του δικτύου, αυτό γίνεται πιο πολύπλοκο και μπορεί να αναγνωρίζει μεγαλύτερα τμήματα της εικόνας. Αρχικά, τα προηγούμενα επίπεδα ανιχνεύουν απλά χαρακτηριστικά όπως οι ακμές και τα χρώματα. Καθώς τα δεδομένα της εικόνας προχωρούν μέσω των επιπέδων-στρωμάτων του CNN, αρχίζει να αντιλαμβάνεται πιο ουσιαστικά χαρακτηριστικά ή μορφές του αντικείμενου, έως ότου τελικά αναγνωρίσει το αντικείμενο-στόχο.

Συνελικτικό Επίπεδο

Το κύριο στοιχείο ενός ΣΝΔ είναι το συνελικτικό επίπεδο, το οποίο εκτελεί το μεγαλύτερο μέρος των υπολογισμών. Χρειάζεται τρία στοιχεία: δεδομένα εισόδου, έναν χάρτη χαρακτηριστικών και ένα φίλτρο. Όταν δουλεύει με μια έγχρωμη εικόνα, τα δεδομένα εισόδου θα αποτελούνται από έναν τρισδιάστατο πίνακα εικονοστοιχείων που αντιπροσωπεύουν το ύψος, το πλάτος και το βάθος (που αντιστοιχεί στο RGB σε μια εικόνα). Ο ανιχνευτής χαρακτηριστικών, που ονομάζεται επίσης πυρήνας ή φίλτρο, κινείται στα πεδία υποδοχής της εικόνας για να ελέγξει την παρουσία ενός χαρακτηριστικού. Η διαδικασία αυτή ονομάζεται συνέλιξη.

Ο ανιχνευτής χαρακτηριστικών είναι ένας δισδιάστατος (2Δ) πίνακας βαρών που αντιπροσωπεύει ένα τμήμα της εικόνας. Αν και το μέγεθος του φίλτρου δεν είναι σταθερό, είναι συνήθως ένας πίνακας 3x3, ο οποίος καθορίζει επίσης το μέγεθος του πεδίου υποδοχής. Στη συνέχεια, το φίλτρο εφαρμόζεται σε ένα τμήμα της εικόνας και τα εικονοστοιχεία εισόδου μαζί με το φίλτρο δημιουργούν ένα εσωτερικό γινόμενο. Αυτό το εσωτερικό γινόμενο προωθείται στη συνέχεια σε έναν πίνακα εξόδου. Το φίλτρο κινείται κατά ένα βήμα και επαναλαμβάνει τη διαδικασία μέχρι να καλυφθεί ολόκληρη η εικόνα. Η έξοδος που προκύπτει από τα εσωτερικά γινόμενα της εισόδου και του φίλτρου ονομάζεται χάρτης χαρακτηριστικών ή χάρτης ενεργοποίησης.

Μετά από κάθε πράξη συνέλιξης, το ΣΝΔ εφαρμόζει έναν μετασχηματισμό ReLU (Rectified Linear Unit - Ανορθωμένης Γραμμικής Μονάδας) στον χάρτη χαρακτηριστικών, ο οποίος εισάγει μη γραμμικότητα στο μοντέλο.

Όπως αναφέρθηκε και προηγουμένως, ένα ακόμα συνελικτικό επίπεδο μπορεί να ακολουθεί το αρχικό συνελικτικό στρώμα, δημιουργώντας μια ιεραρχική δομή στο νευρωνικό δίκτυο. Έτσι, τα μεταγενέστερα στρώματα έχουν πρόσβαση στα εικονοστοιχεία των πεδίων υποδοχής προηγούμενων στρωμάτων. Για να το καταδείξουμε αυτό, μπορούμε να υποθέσουμε την αναγνώριση ενός αυτοκινήτου σε μια εικόνα. Το όχημα μπορεί να θεωρηθεί ως ένα σύνολο τμημάτων, όπως ανεμοθώρακας, παράθυρα, τροχοί, καθρέφτες, πόρτες κ.λπ. Κάθε μέρος του αυτοκινήτου αντιστοιχεί σε ένα μοτίβο χαμηλότερου επιπέδου στο νευρωνικό δίκτυο, ενώ ο

συνδυασμός όλων αυτών των μερών αντιστοιχεί σε ένα μοτίβο υψηλότερου επιπέδου, σχηματίζοντας έτσι μια ιεραρχία χαρακτηριστικών στο νευρωνικό.

Επίπεδο Ομαδοποίησης

Τα επίπεδα ομαδοποίησης, τα οποία αναφέρονται επίσης ως μειοδειγματοληψία, επιτελούν μείωση της διάστασης μειώνοντας τον αριθμό των παραμέτρων εισόδου. Η λειτουργία της ομαδοποίησης είναι παρόμοια με εκείνη του συνελκτικού επιπέδου, καθώς χρησιμοποιεί ένα φίλτρο που σαρώνει ολόκληρη την είσοδο. Ωστόσο, το φίλτρο στο στρώμα ομαδοποίησης δεν έχει βάρη. Αντίθετα, ο πυρήνας χρησιμοποιεί μια συνάρτηση συνάθροισης στις τιμές εντός του πεδίου υποδοχής, οι οποίες στη συνέχεια διαμορφώνουν τον πίνακα εξόδου. Υπάρχουν δύο βασικοί τύποι ομαδοποίησης:

- **Μέγιστη ομαδοποίηση (Max pooling):** Καθώς το φίλτρο κινείται στην είσοδο, επιλέγει το εικονοστοιχείο με τη μέγιστη τιμή για να το στείλει στη συστοιχία εξόδου. Σε αυτό το σημείο, πρέπει να αναφερθεί ότι αυτή η μέθοδος χρησιμοποιείται γενικά συχνότερα από την ομαδοποίηση μέσου όρου (average pooling).
- **Ομαδοποίηση Μέσου Όρου (Average pooling):** Καθώς το φίλτρο κινείται στην είσοδο, υπολογίζει τη μέση τιμή εντός του πεδίου υποδοχής για να την στείλει στη συστοιχία εξόδου.

Αν και χάνονται πολλές πληροφορίες στο στρώμα ομαδοποίησης, έχει επίσης αρκετά πλεονεκτήματα για το ΣΝΔ. Συγκεκριμένα, είναι επωφελής όσον αφορά τη μείωση της πολυπλοκότητας, την επίτευξη αποδοτικότητας και την ελαχιστοποίηση του κινδύνου υπερπροσαρμογής (overfitting).

Πλήρως Συνδεδεμένο (ΠΣ) Επίπεδο

Το πλήρως συνδεδεμένο επίπεδο έχει πάρει αυτήν την ονομασία επειδή συνδέει κάθε κόμβο στο στρώμα εξόδου απευθείας με έναν κόμβο στο προηγούμενο στρώμα. Αυτό έρχεται σε αντίθεση με τα μερικώς συνδεδεμένα επίπεδα, όπου οι τιμές των εικονοστοιχείων της εικόνας εισόδου δεν συνδέονται άμεσα με το στρώμα εξόδου.

Το επίπεδο είναι υπεύθυνο για την κατηγοριοποίηση χρησιμοποιώντας τα χαρακτηριστικά που προέκυψαν από τα προηγούμενα στρώματα και τα διακριτά φίλτρα τους. Οι συναρτήσεις ReLU χρησιμοποιούνται κυρίως στα συνελκτικά στρώματα και στρώματα ομαδοποίησης, αλλά τα ΠΣ επίπεδα εφαρμόζουν συνήθως μια συνάρτηση ενεργοποίησης «softmax» για να εξασφαλίσουν την κατάλληλη ταξινόμηση της εισόδου, δημιουργώντας ένα πιθανοτικό αποτέλεσμα μεταξύ 0 και 1.

Αρχιτεκτονική του Δικτύου ResNet50

Το ResNet σημαίνει Υπολειμματικό Δίκτυο (Residual Network) και είναι ένας ειδικός τύπος ΣΝΔ που παρουσιάστηκε το 2015 [18]. Το ResNet-50 είναι μια παραλλαγή αυτού του δικτύου που αποτελείται από 50 στρώματα, συμπεριλαμβανομένων 48 επιπέδων συνελκτικού τύπου, ενός στρώματος Μέγιστης Ομαδοποίησης και ενός στρώματος μέσου όρου. Τα υπολειμματικά

νευρωνικά δίκτυα είναι ένας τύπος νευρωνικού δικτύου που κατασκευάζει δίκτυα συσσωρεύοντας υπολειμματικά στοιχεία.

Το ResNet-50 έχει μια αρχιτεκτονική η οποία απεικονίζεται στο παρεχόμενο διάγραμμα μοντέλου. Το ResNet των 50 στρωμάτων χρησιμοποιεί ένα σχεδιασμό «bottleneck» για τα δομικά του στοιχεία. Αυτός ο σχεδιασμός ενσωματώνει 1×1 συνελίξεις, γνωστές ως «bottlenecks», οι οποίες μειώνουν τον αριθμό των παραμέτρων και των πολλαπλασιασμών πινάκων. Κατά συνέπεια, η προσέγγιση αυτή επιτρέπει ταχύτερη εκπαίδευση για κάθε επίπεδο. Σε αντίθεση με την τυπική δομή δύο στρωμάτων, το ResNet-50 χρησιμοποιεί μια στοιβα τριών στρωμάτων στα bottleneck υπολειμματικά του στοιχεία.

Η αρχιτεκτονική του ResNet των 50 στρωμάτων περιλαμβάνει τα ακόλουθα στοιχεία, όπως απεικονίζονται και στο παρακάτω διάγραμμα [19]:

- **Μια συνέλιξη πυρήνα 7×7** μαζί με 64 άλλους πυρήνες με βήμα μεγέθους 2.
- **Ένα στρώμα μέγιστης ομαδοποίησης** με βήμα μεγέθους 2.
- **9 πρόσθετα στρώματα συνελίξης 3×3** , 64 πυρήνων, ακολουθούμενα από ένα άλλο με 1×1 , 64 πυρήνων και ένα τρίτο με 1×1 , 256 πυρήνων. Αυτή η ακολουθία επαναλαμβάνεται 3 φορές.
- **12 ακόμη στρώματα με συνέλιξη 1×1** , 128 πυρήνων, 3×3 , 128 πυρήνων και 1×1 , 512 πυρήνων, που επαναλαμβάνονται 4 φορές.
- **18 ακόμη στρώματα με συνέλιξη 1×1** , 256 πυρήνων, 3×3 , 256 πυρήνων και 1×1 , 1024 πυρήνων, που επαναλαμβάνονται 6 φορές.
- **9 ακόμη στρώματα με 1×1 συνέλιξη**, 512 πυρήνων, 3×3 , 512 πυρήνων και 1×1 , 2048 πυρήνων, επαναλαμβανόμενα 3 φορές.

(Σε αυτό το σημείο, το δίκτυο έχει φτάσει τα 50 επίπεδα)

- **Επίπεδο ομαδοποίησης μέσου όρου**, ακολουθούμενο από ένα πλήρως συνδεδεμένο επίπεδο με 1000 κόμβους, που εφαρμόζει τη συνάρτηση ενεργοποίησης softmax.



Εικόνα 6. Ένα διάγραμμα αναπαράστασης προεκπαιδευμένης αρχιτεκτονικής Resnet-50.

Όραση Υπολογιστών

Η Όραση Υπολογιστών, γνωστή και ως Μηχανική Όραση, είναι ένα υποσύνολο της Τεχνητής Νοημοσύνης (TN) που χρησιμοποιεί ψηφιακές εικόνες, βίντεο και άλλα οπτικά δεδομένα για να επιτρέπει στους υπολογιστές και τα συστήματα να εξάγουν σημαντικές πληροφορίες και να παρέχουν προτάσεις ή να αναλαμβάνουν δράση με βάση αυτές τις πληροφορίες [20].

Ενώ η Τεχνητή Νοημοσύνη δίνει τη δυνατότητα στους υπολογιστές να σκέφτονται, ο σκοπός της όρασης υπολογιστών είναι να βοηθήσει τους υπολογιστές να βλέπουν, να εξετάζουν και να κατανοούν το οπτικό περιεχόμενο, παρόμοια με την ανθρώπινη όραση, αν και με διαφορετικά εργαλεία και χρόνο εκμάθησης. Η ανθρώπινη όραση έχει το πλεονέκτημα των εμπειριών μιας ολόκληρης ζωής για να εκπαιδευτεί στο πώς να διαφοροποιεί αντικείμενα, να καθορίζει την απόστασή τους, να ανιχνεύει την κίνηση και να εντοπίζει ζητήματα σε εικόνες.

Η όραση υπολογιστών χρησιμοποιεί κάμερες, δεδομένα και αλγόριθμους για να διδάξει στις μηχανές να εκτελούν τέτοιες λειτουργίες, χωρίς να βασίζεται σε βιολογικά συστήματα που μοιάζουν με τα ανθρώπινα, όπως ο αμφιβληστροειδής, τα οπτικά νεύρα και ο οπτικός φλοιός. Όταν οι μηχανές εκπαιδεύονται να εξετάζουν προϊόντα ή να επιβλέπουν τη λειτουργία ενός περιουσιακού στοιχείου παραγωγής, μπορούν να αναλύουν χιλιάδες προϊόντα ή διαδικασίες ανά λεπτό και να εντοπίζουν ακόμη και τις παραμικρές ατέλειες ή παρατυπίες. Αυτή η ικανότητα επιτρέπει στις μηχανές να ξεπερνούν γρήγορα τους ανθρώπους.

Η όραση υπολογιστών έχει γίνει πανταχού παρούσα σε μια ποικιλία βιομηχανιών, από την ενέργεια και τις υπηρεσίες κοινής ωφέλειας έως τη μεταποίηση και την αυτοκινητοβιομηχανία, και η αγορά για αυτή την τεχνολογία εξακολουθεί να επεκτείνεται [21].

Για την επίτευξη ακριβούς αναγνώρισης εικόνων, η όραση υπολογιστών απαιτεί μεγάλο όγκο δεδομένων. Επεξεργάζεται και αναλύει τα δεδομένα επανειλημμένα μέχρι να μπορέσει να ξεχωρίσει τα αντικείμενα και τελικά να τα αναγνωρίσει. Για παράδειγμα, προκειμένου ένας υπολογιστής να μάθει να αναγνωρίζει ελαστικά αυτοκινήτων, πρέπει να του δοθεί ένας τεράστιος όγκος εικόνων ελαστικών και αντικειμένων που σχετίζονται με ελαστικά, ώστε να μάθει τις διαφορές και να εντοπίζει ένα ελαστικό. Επιπρόσθετη προσοχή πρέπει να δοθεί σε αυτά που δεν παρουσιάζουν ατέλειες.

Για να επιτευχθεί αυτό, χρησιμοποιούνται δύο ζωτικής σημασίας τεχνολογίες: η βαθιά μάθηση και τα ΣΝΔ. Η μηχανική μάθηση περιλαμβάνει τη χρήση αλγοριθμικών μοντέλων που επιτρέπουν στους υπολογιστές να μαθαίνουν μόνοι τους για το περιεχόμενο των οπτικών δεδομένων. Όταν αρκετά δεδομένα τροφοδοτηθούν μέσω του μοντέλου, ο υπολογιστής "βλέπει" τα δεδομένα και μαθαίνει να διαφοροποιεί τη μία εικόνα από την άλλη. Οι αλγόριθμοι επιτρέπουν στη μηχανή να μαθαίνει μόνη της, χωρίς να προγραμματίζεται από άνθρωπο για να αναγνωρίσει μια συγκεκριμένη εικόνα.

Τα Συνελικτικά Νευρωνικά Δίκτυα βοηθούν τα μοντέλα μηχανικής και βαθιάς μάθησης να "βλέπουν" τις εικόνες, αναλύοντάς τις σε εικονοστοιχεία και αποδίδοντάς τους ετικέτες. Αυτές οι ετικέτες χρησιμοποιούνται στη συνέχεια για την εκτέλεση μαθηματικών πράξεων, που ονομάζονται συνελίξεις, ώστε να προβλεφθεί τι αντιπροσωπεύει η εικόνα. Το νευρωνικό δίκτυο πραγματοποιεί συζεύξεις και επαληθεύει την ακρίβεια των προβλέψεών του με μια σειρά επαναλήψεων έως ότου οι προβλέψεις φτάσουν στην πραγματική τιμή, επιτρέποντάς του να αναγνωρίζει ή να "βλέπει" εικόνες με τρόπο ανάλογο όπως ο άνθρωπος.

Όμοια με τον τρόπο με τον οποίο οι άνθρωποι αναγνωρίζουν μακρινά αντικείμενα, ένα ΣΝΔ ανιχνεύει αρχικά τραχιές ακμές και βασικά σχήματα και στη συνέχεια βελτιώνει την πρόβλεψή του

με κάθε επανάληψη για να αναγνωρίσει την εικόνα. Τα ΣΝΔ χρησιμοποιούνται για την κατανόηση μεμονωμένων εικόνων, ενώ υπάρχουν και άλλοι τύποι νευρωνικών, όπως τα Επαναλαμβανόμενα Νευρωνικά Δίκτυα, που χρησιμοποιούνται με παρόμοιο τρόπο, αλλά για εφαρμογές βίντεο ώστε να βοηθήσουν τους υπολογιστές να κατανοήσουν πώς οι εικόνες συνδέονται μεταξύ τους σε μια ακολουθία καρέ.

Εργαλεία και Μέθοδοι που Ακολουθήθηκαν

Σε αυτό το κεφάλαιο, θα εξεταστούν τα βήματα που ακολουθήθηκαν στην παρούσα έρευνα. Πρώτον, συλλέξαμε ένα μεγάλο σύνολο δεδομένων από αντικείμενα. Το σύνολο δεδομένων μας αποτελείται από δύο κύριες κλάσεις και επτά υποκλάσεις. Οι δύο κύριες κλάσεις μας είναι το Ξύλο και το Πλαστικό και οι επτά υποκλάσεις μας περιλαμβάνουν την Ινοσανίδα μέσης πυκνότητας (MDF), τη Μοριοσανίδα (Νοβοπάν) με επένδυση μελαμίνης (MFC), τον καπλαμά δρυός, οι οποίες ανήκουν στο Ξύλο, ενώ οι υπόλοιπες, Πολυαιθυλένιο υψηλής πυκνότητας (HDPE), Πολυαιθυλένιο χαμηλής πυκνότητας (LDPE), Τερεφθαλικό πολυαιθυλένιο (PET), Πολυπροπυλένιο (PP) ανήκουν στο Πλαστικό. Έτσι, συνολικά έχουμε τρία είδη ξύλου και τέσσερα είδη πλαστικού. Συγκεντρώθηκαν περίπου 100 τεμάχια από κάθε κατηγορία που κυμαίνονται σε διαφορετικούς τύπους και μεγέθη.

Στη συνέχεια, προχωρήσαμε με τη δημιουργία των συνόλων δεδομένων εικόνας. Η διαδικασία αυτή περιλάμβανε την τοποθέτηση κάθε αντικείμενου στην πλατφόρμα και τη λήψη φωτογραφιών με τη βοήθεια των φωτογραφικών μηχανών. Χρησιμοποιήσαμε τη βιομηχανική κάμερα RGB, την πολυφασματική κάμερα ορατού (VIS) και την εγγύς υπέρυθρου (NIR) πολυφασματική κάμερα. Επομένως, για κάθε αντικείμενο έχουμε τρεις διαφορετικούς τύπους φωτογραφιών που έχουν ληφθεί και οι οποίες τελικά θα αποτελέσουν τους τρεις ταξινομητές μας. Επιπλέον, αλλάξαμε τον προσανατολισμό των αντικειμένων και επαναλάβαμε την ίδια διαδικασία για δύο βασικούς λόγους. Ο ένας είναι για να αυξήσουμε τον αριθμό των εικόνων προκειμένου το μοντέλο να εκπαιδευτεί καλύτερα και ο άλλος είναι για να δοκιμάσουμε διαφορετικές σκιές, τοποθετήσεις και γωνίες των αντικειμένων.

Έπειτα, στάθηκε η προεπεξεργασία των πολυφασματικών δεδομένων. Περιλάμβανε την ανάγνωση των ακατέργαστων αρχείων και τη μετατροπή τους σε τρισδιάστατους υπερκύβους, οι οποίοι είχαν είτε 16 ζώνες (VIS) είτε 25 ζώνες (NIR). Επιπλέον, έπρεπε να εκτελεστεί μια ανάλυση PCA προκειμένου να ληφθούν οι τρεις κορυφαίες ιδιοτιμές με τις πιο χρήσιμες πληροφορίες.

Μετάπειτα, αφού μετατρέψαμε και τους τρεις τύπους εικόνων σε αρχεία με τα οποία μπορούμε να εργαστούμε, η αλλαγή μεγέθους και η κανονικοποίηση ήταν τα επόμενα βήματά μας. Η επαύξηση των δεδομένων ήρθε αμέσως μετά και έπαιξε σημαντικό ρόλο. Σε αυτό το βήμα, ο στόχος ήταν να δημιουργήσουμε ακόμα περισσότερα δεδομένα προκειμένου να ελέγξουμε αν τα μοντέλα μας θα έπαιρναν λανθασμένη απόφαση σε κάθε επαυξημένη έκδοση των δειγμάτων. Η τεχνική αυτή περιλάμβανε τη δημιουργία πέντε νέων (άρα συνολικά 6 εικόνων από κάθε αντικείμενο) εκδόσεων των εικόνων που καταγράφηκαν εφαρμόζοντας τυχαίο ζουμ (0 έως +20%), τυχαία περιστροφή (-36° έως +36°) και τυχαία οριζόντια αναστροφή.

Ως εκ τούτου, καταλήξαμε σε 6522 αρχεία εικόνων από κάθε τύπο κάμερας, τα οποία αποτελούν την είσοδο του ΣΝΔ μας. Αυτές οι εικόνες στη συνέχεια χωρίστηκαν σε σύνολα εκπαίδευσης και επικύρωσης, τα οποία δεν επικαλύπτονται μεταξύ τους και δεν μοιράζονται καμία έκδοση του ίδιου δείγματος. Ο κανόνας που ακολουθήθηκε για το διαχωρισμό αυτό είναι αυτός που χρησιμοποιείται συχνότερα [22], η αναλογία 80-20%, 80% για το σύνολο εκπαίδευσης και 20% για το σύνολο επικύρωσης.

Ως αποτέλεσμα, το σύνολο εκπαίδευσης περιέχει 5238 αρχεία, ενώ το σύνολο επικύρωσης περιέχει 1284, τα οποία αθροίζονται συνολικά σε 6522 αρχεία εικόνων. Αναλυτικότερα, κάθε κατηγορία υλικού των επτά συνεισέφερε όπως φαίνεται παρακάτω:

Κλάση	Αριθμός Εικόνων για Εκπαίδευση	Αριθμός Εικόνων για Επικύρωση	Συνολικός Αριθμός Εικόνων
MDF	456	108	564
MFC	552	132	684
Καπλαμάς Βελανιδιάς	576	144	720
HDPE	498	120	618
LDPE	492	120	612
PET	1398	348	1746
PP	1266	312	1578

Πίνακας 1. Συνεισφορά εικόνων κάθε υλικού στα διάφορα σύνολα.

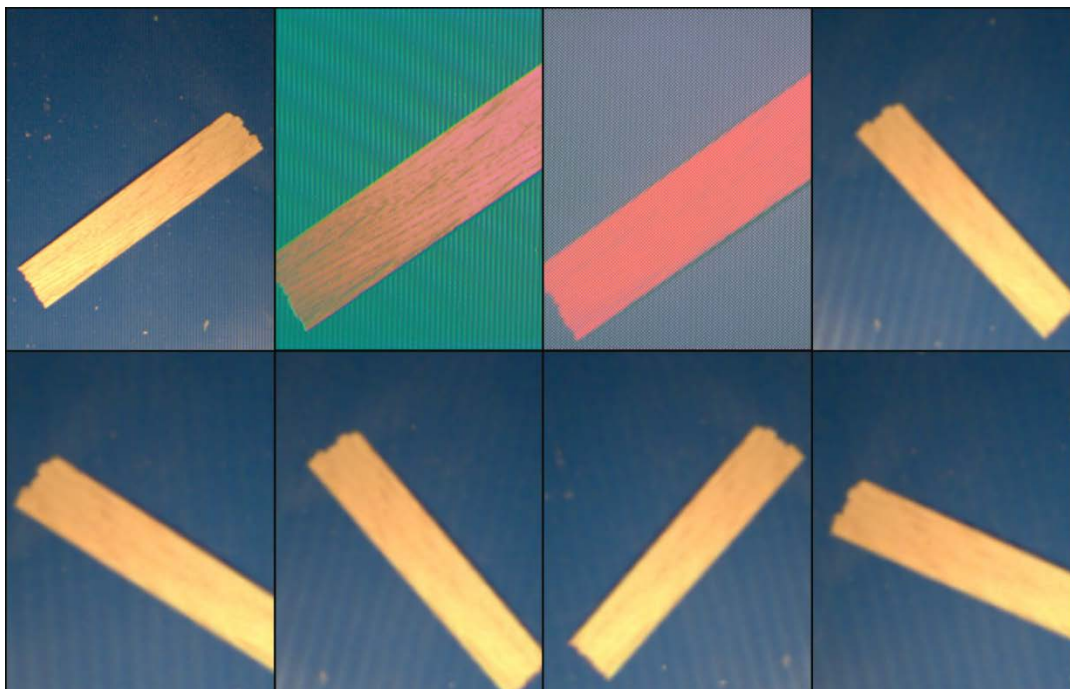
Μια σημαντική διαδικασία κατά τη διάρκεια αυτής της μελέτης ήταν η εκπαίδευση μοντέλων με διάφορες αρχιτεκτονικές και ο πειραματισμός με την αλλαγή των παραμέτρων τους μέχρι να βρεθεί εκείνη που είχε ικανοποιητική απόδοση.

Οι τελικές παράμετροι που επιλέχθηκαν για την εκπαίδευση του μοντέλου είναι:

- **Κριτήριο:** Διασταυρούμενη Απώλεια Εντροπίας (Cross Entropy Loss) μεταξύ των μη κανονικοποιημένων προβλέψεων (logits) εισόδου και του στόχου, καθώς αυτός ο τύπος απώλειας φτάνει στο μέγιστο των δυνατοτήτων του κατά την εκπαίδευση ενός προβλήματος ταξινόμησης με διάφορες κλάσεις, προσαρμόζοντας τα βάρη του μοντέλου. Ο μαθηματικός τύπος της απώλειας διασταυρούμενης εντροπίας είναι:

$$L_{CE} = - \sum_{i=1}^n t_i \log(p_i),$$

όπου n είναι ο αριθμός των κλάσεων, t_i είναι η ετικέτα αλήθειας και p_i είναι η πιθανότητα Softmax για την i^{th} κλάση.



Εικόνα 7. Επαύξηση Εικόνων: Ξεκινώντας από την επάνω αριστερή γωνία: Αρχική RGB εικόνα ενός τεμαχίου ξύλου της κατηγορίας καπλαμά δρυός, αρχική VIS εικόνα, αρχική NIR εικόνα, ακολουθούμενες από πέντε επαυξημένες εκδόσεις RGB του ίδιου κομματιού.

- **Βελτιστοποιητής:** Στοχαστική κλίση καθόδου (Stochastic Gradient Descent - SGD). Η βασική ιδέα πίσω από την SGD είναι η επαναληπτική ενημέρωση των παραμέτρων του μοντέλου με μικρά βήματα προς την κατεύθυνση της πιο απότομης καθόδου της συνάρτησης απώλειας. Σε κάθε επανάληψη, η SGD επιλέγει τυχαία ένα υποσύνολο παραδειγμάτων εκπαίδευσης (που ονομάζεται μίνι-πακέτο) και υπολογίζει τις κλίσεις με βάση αυτό το μίνι-πακέτο. Οι κλίσεις χρησιμοποιούνται στη συνέχεια για την ενημέρωση των παραμέτρων του μοντέλου. Αυτή η διαδικασία επαναλαμβάνεται για πολλές επαναλήψεις μέχρι να επιτευχθεί σύγκλιση ή ένα κριτήριο διακοπής. Απαιτεί προσεκτικό συντονισμό του ρυθμού μάθησης και άλλων υπερπαραμέτρων, και η σύγκλιση της μπορεί να είναι ευαίσθητη στην επιλογή του μεγέθους της μίνι-παρτίδας. Αυτός είναι ο λόγος πίσω από τις τέσσερις επόμενες επιλογές μας.
- **Ρυθμός μάθησης:** 0.001. Ο ρυθμός μάθησης είναι μια υπερπαραμέτρος που χρησιμοποιείται για να ρυθμίζει τον ρυθμό με τον οποίο ο αλγόριθμος προσαρμόζει ή μαθαίνει τις τιμές μιας εκτίμησης παραμέτρου, δηλαδή ρυθμίζει τα βάρη του ΣΝΔ μας όσον αφορά την κλίση απώλειας. Δείχνει τη συχνότητα με την οποία το νευρωνικό δίκτυο ενσωματώνει τη γνώση που έχει αποκτήσει, ανανεώνοντας τις έννοιες που έχει μάθει.
- **Μέγεθος παρτίδας:** 4. Το μέγεθος της παρτίδας καθορίζει τον αριθμό των δειγμάτων που θα διαδοθούν μέσω του δικτύου.
- **Γάμμα (Gamma):** 0.1, ένας πολλαπλασιαστικός παράγοντας με βάση τον οποίο θα επιδεινώνεται ο ρυθμός μάθησης. Για παράδειγμα, εάν ο ρυθμός μάθησης είναι 1000 και το γάμμα είναι 0.1, ο νέος ρυθμός μάθησης θα είναι $1000 \times 0.1 = 100$ κ.ο.κ.

- **Μέγεθος βήματος:** 7, μειώνει το ρυθμό μάθησης κάθε ομάδας παραμέτρων κατά γάμμα κάθε N εποχές. Το N -εδώ- ορίζεται ως 7.
- **Εποχές:** 50

Οι επιδόσεις τους υπολογίστηκαν μέσω των μετρικών ΣΝΔ, όπως η ακρίβεια και η απώλεια. Τα αποτελέσματα αυτής της προσπάθειας παρουσιάζονται σε επόμενη ενότητα. Έτσι, αξιοποιώντας την αρχιτεκτονική ΣΝΔ που επιλέξαμε, δημιουργήσαμε τρία μοντέλα πρόβλεψης εικόνας - ταξινομητές (RGB, VIS και NIR). Το έργο αυτών των ταξινομητών είναι να δέχονται ως είσοδο μια εικόνα του τύπου τους και να δίνουν ως έξοδο το διάνυσμα των πιθανοτήτων τους, δηλαδή το πόσο πιθανό είναι να εντάσσεται η δεδομένη είσοδος σε κάθε μία από τις επτά προκαθορισμένες κατηγορίες. Η κατηγορία με το υψηλότερο ποσοστό θεωρείται ως η τελική τους έξοδος-πρόβλεψη.

Σε αυτό το σημείο, ήταν απαραίτητο να συνδυαστούν αυτοί οι τρεις ταξινομητές και να συγκριθούν οι επιδόσεις τους μεμονωμένα αλλά και σε ζεύγη. Και πάλι, τα αποτελέσματα προβάλλονται σε επόμενη ενότητα.

Τέλος, η εκτεταμένη εξέταση των ζωνών απαιτούσε την επανάληψη της όλης διαδικασίας, τουλάχιστον για τις πολυφασματικές εικόνες. Σε αυτό το στάδιο, εξετάζουμε τον αντίκτυπο χειρότερου εξοπλισμού, εργαστηριακών συνθηκών και, κατά συνέπεια, εικόνων με λιγότερες ζώνες και, επομένως, τη σημασία του ρόλου που διαδραματίζουν οι ζώνες. Για το σκοπό αυτό, είναι θεμελιώδες να συγκρίνουμε τις επιδόσεις ΣΝΔ ενός καλού εξοπλισμού και ενός ανεπαρκούς εξοπλισμού, αυτού που καταγράφει πολυφασματικές εικόνες με μόνο τρεις ζώνες.

Προσεγγίσεις για την επιλογή των ζωνών σε περίπτωση ανεπαρκούς εξοπλισμού

Υπάρχουν τέσσερις διαφορετικές προσεγγίσεις που ακολουθήθηκαν για την επιλογή μόνο τριών ζωνών αντί του συνδυασμού τους που προβλέπει η μέθοδος PCA. Φυσικά, οι προσεγγίσεις αυτές αναφέρονται στις πολυφασματικές εικόνες λόγω της φύσης τους να έχουν περισσότερες από τρεις ζώνες σε σύγκριση με τις εικόνες RGB. Όπως έχουμε ήδη αναφέρει, εργαζόμαστε με επτά διαφορετικές κατηγορίες και δύο τύπους πολυφασματικών εικόνων, τις NIR και VIS.

Η πρώτη μέθοδος ήταν η τυχαία επιλογή, η οποία επαναλήφθηκε αρκετές φορές μέχρι να μην παρουσιάζεται πλέον σημαντική βελτίωση ή επιδείνωση στα μέσα αποτελέσματα. Η πρόθεση πίσω από την εφαρμογή αυτής της μεθόδου ήταν να λειτουργήσει ως σημείο αναφοράς για τις άλλες τρεις προσεγγίσεις.

Η δεύτερη τεχνική ήταν η επιλογή "καλύτερη εν συνόλω". Με την εφαρμογή της PCA στα δύο πολυφασματικά σύνολα δεδομένων μας καταλήξαμε στις ιδιοτιμές και τα ιδιοδιανύσματα κάθε εικόνας. Απομονώσαμε τις τρεις πρώτες ιδιοτιμές και ιδιοδιανύσματα κάθε εικόνας, καθώς οι υπόλοιπες ιδιοτιμές περιλαμβάνουν κυρίως θόρυβο. Στη συνέχεια, πολλαπλασιάζοντάς τις και επαναλαμβάνοντας την ίδια διαδικασία για κάθε εικόνα προέκυψε το άθροισμα του γινομένου όλων των εικόνων το οποίο έδειξε τις τρεις ζώνες που προσφέρουν τις περισσότερες πληροφορίες για κάθε τύπο φάσματος. Έτσι, για εικόνες τύπου NIR συμπεράναμε ότι οι πιο σημαντικές ζώνες είναι η 1η, η 12η και η 17η, ενώ για εικόνες τύπου VIS η 2η, η 15η και η 16η.

Η τρίτη μέθοδος που εφαρμόστηκε ήταν η "καλύτερη από κάθε κατηγορία" και στη συνέχεια η επιλογή των τριών πιο επαναλαμβανόμενων ζωνών. Χρησιμοποιήσαμε την ίδια διαδικασία με τη δεύτερη τεχνική. Παρ' όλα αυτά, την εφαρμόσαμε ξεχωριστά σε κάθε κλάση, σημειώνοντας τις τρεις κορυφαίες ζώνες για κάθε κλάση. Τελικά, εντοπίσαμε τις τρεις ζώνες που εμφανίζονταν συχνότερα. Έτσι, για τις εικόνες τύπου NIR καταλήξαμε στο συμπέρασμα ότι οι πιο επαναλαμβανόμενες ζώνες είναι η 5η, η 12η και η 18η, ενώ για τις εικόνες τύπου VIS η 2η, η 10η και η 14η.

Η τέταρτη προσέγγιση ήταν παρόμοια με την τρίτη, ωστόσο εφαρμόστηκε στις χειρότερες ζώνες κάθε κατηγορίας, ακολουθούμενη από την επιλογή των πιο συχνά εμφανιζόμενων ζωνών. Έτσι, για τις εικόνες τύπου NIR καταλήξαμε στο συμπέρασμα ότι οι πιο επαναλαμβανόμενες ζώνες είναι η 20ή, η 23η και η 24η, ενώ για τις εικόνες τύπου VIS η 3η, η 4η και η 7η.

Διάταξη Πειράματος

Η εργαστηριακή διάταξη για τα πειράματα ακολουθεί τις συνήθεις διατάξεις που βρίσκονται στη βιβλιογραφία [23], [24], [25], [26]. Στην Εικόνα 8 παρουσιάζεται η διάταξη που υλοποιήθηκε για τους σκοπούς της παρούσας μελέτης. Τα κύρια στοιχεία που απεικονίζονται περιλαμβάνουν, πρώτον, έναν ιμάντα μεταφοράς ο οποίος έχει τα ίδια χαρακτηριστικά με μια πραγματική γραμμή παραγωγής.

Δεύτερον, υπάρχει μια βιομηχανική κάμερα RGB, με τις προδιαγραφές της να παρουσιάζονται στη συνέχεια. Τρίτον, υπάρχουν δύο πολυφασματικοί αισθητήρες, ένας του ορατού, υπέρυθρου φωτός μικρού κύματος (αισθητήρας SWIR) και ένας του εγγύς υπέρυθρου φωτός (πολυφασματικός αισθητήρας) με τις προδιαγραφές τους να φαίνονται παρακάτω.

Χαρακτηριστικό	Τιμή	Χαρακτηριστικό	Τιμή
Προμηθευτής Αισθητήρων	e2v	Ανάλυση (Οριζ. x Καθ.)	1280 px x 1024 px
Αισθητήρας	EV76C560	Ανάλυση	1.3 MP
Κλείστρο	Σφαιρικό & Κυλιόμενο Κλείστρο	Μέγεθος Εικονοστοιχείου (Οριζ x Κάθ.)	5.3 μm x 5.3 μm
Μορφή Αισθητήρα	1/1.8"	Ρυθμός Καρέ	60 fps
Τύπος Αισθητήρα	CMOS	Μονόχρωμη/Πολύχρωμη	Πολύχρωμη
Μέγεθος Αισθητήρα	6.8 mm x 5.4 mm	Βάθος bit pixel	12 bits

Πίνακας 2. Προδιαγραφές της βιομηχανικής κάμερας.

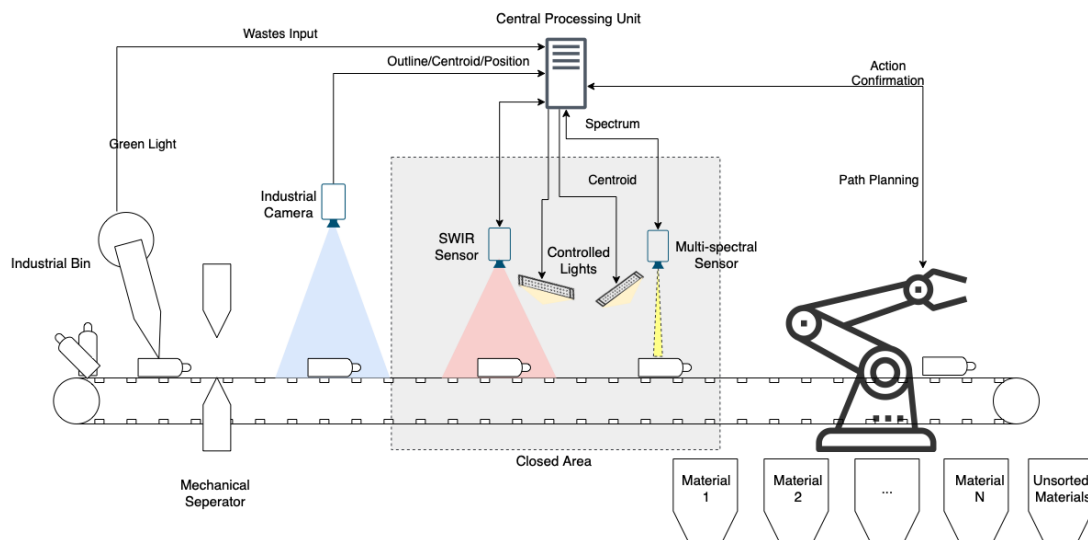
Μέθοδος Λήψης, Κεφαλές Κάμερας	Φάσμα
Στιγμιότυπο, Κεφαλή Αισθητήρα SWIR (VIS)	470 – 620 nm, 16 ζώνες
Στιγμιότυπο, Κεφαλή Πολυφασματικού Αισθητήρα (NIR)	600 – 975 nm, 25 ζώνες

Πίνακας 3. Προδιαγραφές των πολυφασματικών αισθητήρων.

Επιπλέον, η πηγή φωτός που χρησιμοποιείται προσφέρει ένα συνεχές φάσμα που είναι θεμελιώδες για τη λήψη των εικόνων. Τέλος, υπάρχει ένας ρομποτικός βραχίονας για την τοποθέτηση των δειγμάτων στο σωστό δοχείο μετά την πρόβλεψη της κατηγορίας τους.

Όλα αυτά τα εξαρτήματα συνδέονται μεταξύ τους μέσω μιας κεντρικής μονάδας επεξεργασίας (CPU), η οποία συγκεντρώνει τις παρεχόμενες πληροφορίες και τις επεξεργάζεται. Όχι μόνο αυτό, αλλά η CPU είναι επίσης υπεύθυνη για τις ενέργειες του ρομποτικού βραχίονα κατευθύνοντάς τον μέσα από εντολές.

Ο ιμάντας μεταφοράς μεταφέρει κάθε αντικείμενο μπροστά από την κάμερα RGB, όπου λαμβάνεται η πρώτη του φωτογραφία και στη συνέχεια ακολουθεί η ίδια διαδικασία για τις άλλες δύο πολυφασματικές κάμερες. Η πηγή του φωτισμού αποτελεί σημαντικό παράγοντα για τη λήψη των πολυφασματικών εικόνων, καθώς κάνει το δείγμα ευδιάκριτο και (όπως φαίνεται στην επόμενη ενότητα) επηρεάζει σε αρκετό βαθμό τα αναμενόμενα αποτελέσματα. Στη συνέχεια, ο μέσος όρος των πιθανοτήτων από τις προβλέψεις των τριών ταξινομητών παρέχει το τελικό αποτέλεσμα (εκτίμηση για το υλικό) και μια εντολή για τον ρομποτικό βραχίονα προκειμένου να τοποθετήσει το αντικείμενο στο σωστό δοχείο. Όσον αφορά το κομμάτι της όρασης των υπολογιστών, το πλαίσιο μηχανικής μάθησης και η αρχιτεκτονική ΣΝΔ που χρησιμοποιούνται είναι το PyTorch και το ResNet50 αντίστοιχα. Το μοντέλο έχει ρυθμιστεί να εκτελείται για 50 εποχές και να σταματά σε περίπτωση που η απώλεια του συνόλου δεδομένων εκπαίδευσης γίνει μικρότερη από 0.1 για να αποφευχθεί η υπερπροσαρμογή. Τέλος, η κατάσταση του μοντέλου στην εποχή με την καλύτερη ακρίβεια αποθηκεύεται αφού ολοκληρωθεί η εκπαίδευση του μοντέλου.



Εικόνα 8. Εργαστηριακή διάταξη των πειραμάτων μας - Αυτόνομο σύστημα διαλογής, το οποίο αποτελείται από ένα βιομηχανικό κάδο, έναν μηχανικό διαχωριστή, ένα σύστημα αναγνώρισης με πολλαπλές κάμερες, έναν ρομποτικό βραχίονα, κάδους εναπόθεσης και μια κεντρική μονάδα επεξεργασίας.

Ο κώδικας που ανέπτυξα για την υλοποίηση της εργασίας μπορεί να βρεθεί στο ακόλουθο αποθετήριο του GitHub ([27]).

Μετρικές Αξιολόγησης

Πίνακας Σύγχυσης

Ο πίνακας σύγχυσης είναι ένας πίνακας που αναπαριστά την απόδοση ενός αλγορίθμου ταξινόμησης. Παρέχει μια συνοπτική περίληψη της απόδοσης του μοντέλου όσον αφορά τον ορθό εντοπισμό και την εσφαλμένη ταξινόμηση περιπτώσεων διαφορετικών κλάσεων. Ένα παράδειγμα ενός πίνακα σύγχυσης απεικονίζεται στην παρακάτω εικόνα, όπου ο ιστός «N = καλοήθης» ταξινομείται ως υγιής και ο ιστός «P = κακοήθης» ταξινομείται ως καρκινικός [28].

		Actual class	
		P	N
Predicted class	P	TP	FP
	N	FN	TN

Εικόνα 9. Πίνακας Σύγχυσης.

Ο πίνακας σύγχυσης αποτελείται από τέσσερις κύριες τιμές που χρησιμοποιούνται για τον καθορισμό των μετρικών αξιολόγησης ενός ταξινομητή. Αυτές οι τέσσερις τιμές είναι οι εξής:

- **Αληθώς θετικά (True Positive, TP):** Αυτή αντιπροσωπεύει τον αριθμό των ασθενών που έχουν ταξινομηθεί σωστά ως έχοντες κακοήθεις κόμβους, δηλαδή ως έχοντες τη νόσο.
- **Αληθώς αρνητικά (True Negative, TN):** Αυτή αντιπροσωπεύει τον αριθμό των ασθενών που έχουν ταξινομηθεί σωστά ως υγιείς
- **Ψευδώς θετικά (False Positive, FP):** Αντιπροσωπεύει τον αριθμό των ασθενών που έχουν ταξινομηθεί εσφαλμένα ως έχοντες τη νόσο, αλλά στην πραγματικότητα είναι υγιείς. Το FP είναι επίσης γνωστό ως σφάλμα τύπου I.
- **Ψευδώς αρνητικά (False Negative, FN):** Αντιπροσωπεύει τον αριθμό των ασθενών που έχουν ταξινομηθεί εσφαλμένα ως υγιείς, ενώ στην πραγματικότητα έχουν τη νόσο. Το FN είναι επίσης γνωστό ως σφάλμα τύπου II.

Ακρίβεια

Η ακρίβεια είναι μια μετρική απόδοσης που χρησιμοποιείται για την αξιολόγηση της αποτελεσματικότητας των μοντέλων ταξινόμησης. Αναφέρεται στο ποσοστό των σωστών προβλέψεων που έχει κάνει το μοντέλο [29]. Πιο συγκεκριμένα, η ακρίβεια έχει τον ακόλουθο ορισμό:

$$\text{Ακρίβεια} = \frac{\text{Αριθμός σωστών προβλέψεων}}{\text{Συνολικός αριθμός προβλέψεων}}$$

Για δυαδική κατηγοριοποίηση, η ακρίβεια μπορεί επίσης να υπολογιστεί ως προς τα θετικά και τα αρνητικά ως εξής:

$$\text{Ακρίβεια} = \frac{TP+TN}{TP+TN+FP+FN} .$$

Όπου TP = Αληθώς θετικά, TN = Αληθώς αρνητικά, FP = Ψευδώς θετικά, και FN = Ψευδώς Αρνητικά, όπως ορίστηκαν προηγουμένως.

Θα εξετάσουμε τώρα την ακρίβεια ενός συγκεκριμένου μοντέλου χρησιμοποιώντας τον πίνακα σύγχυσης. Το εν λόγω μοντέλο έχει ταξινομήσει 100 όγκους είτε ως καλοήθεις (η θετική κατηγορία) είτε ως καλοήθεις (η αρνητική κατηγορία) με τα αποτελέσματα να παρουσιάζονται παρακάτω:

<p>Αληθώς θετικά (TP):</p> <p>Αληθινή τιμή: Κακοήθης Το μοντέλο προέβλεψε: Κακοήθης Αριθμός TP αποτελεσμάτων: 1</p>	<p>Ψευδώς θετικά (FP):</p> <p>Αληθινή τιμή: Καλοήθης Το μοντέλο προέβλεψε: Κακοήθης Αριθμός FP αποτελεσμάτων: 1</p>
<p>Ψευδώς Αρνητικά (FN):</p> <p>Αληθινή τιμή: Κακοήθης Το μοντέλο προέβλεψε: Καλοήθης Αριθμός FN αποτελεσμάτων: 8</p>	<p>Αληθώς αρνητικά (TN):</p> <p>Αληθινή τιμή: Καλοήθης Το μοντέλο προέβλεψε: Καλοήθης Αριθμός TN αποτελεσμάτων: 90</p>

Πίνακας 4. Πίνακας Σύγχυσης ενός μοντέλου που κατατάσσει 100 όγκους ως καλοήθεις ή κακοήθεις.

Η ακρίβεια του μοντέλου υπολογίζεται σε 0,91 ή 91%, που σημαίνει ότι έχει κάνει 91 σωστές προβλέψεις από το σύνολο των 100 παραδειγμάτων. Με μια πρώτη ματιά, αυτή η υψηλή ακρίβεια μπορεί να υποδηλώνει ότι ο ταξινομητής όγκων αναγνωρίζει αποτελεσματικά τις κακοήθειες. Ωστόσο, για να κατανοήσουμε καλύτερα την απόδοση του μοντέλου, είναι απαραίτητο να προβούμε σε λεπτομερέστερη ανάλυση των θετικών και αρνητικών αποτελεσμάτων.

Από τα 100 παραδείγματα όγκων, τα 91 από αυτά είναι καλοήθη (90 αληθής αρνητικά και 1 ψευδώς θετικό), ενώ τα υπόλοιπα 9 είναι κακοήθη (1 αληθώς θετικό και 8 ψευδώς αρνητικά). Από τους 91 όγκους που είναι καλοήθεις, το μοντέλο έχει ταξινομήσει με ακρίβεια τους 90 από αυτούς ως καλοήθεις, γεγονός που αποτελεί θετικό αποτέλεσμα. Ωστόσο, από τους 9 κακοήθεις όγκους, το μοντέλο αναγνώρισε σωστά μόνο 1 ως κακοήθη. Αυτό είναι ένα ανησυχητικό αποτέλεσμα, καθώς το μοντέλο έχει χάσει 8 από τους 9 κακοήθεις όγκους, υποδεικνύοντας ότι η ικανότητα του μοντέλου να εντοπίζει κακοήθειες είναι ανεπαρκής.

Μια ακρίβεια 91% μπορεί να φαίνεται εντυπωσιακή. Ωστόσο, αν εξετάσουμε ένα άλλο μοντέλο ταξινόμησης όγκων που προβλέπει πάντα καλοήθεις όγκους, θα επιτύχει ακριβώς την ίδια ακρίβεια, δηλαδή 91 στις 100 σωστές προβλέψεις. Αυτό σημαίνει ότι το τρέχον μοντέλο δεν είναι καλύτερο από ένα μοντέλο που δεν έχει καμία προβλεπτική ικανότητα στη διάκριση μεταξύ κακοήθων και καλοήθων όγκων. Όταν έχουμε να κάνουμε με ένα σύνολο δεδομένων με ανισορροπία μεταξύ των κατηγοριών του, όπως αυτό του παρόντος σεναρίου, όπου υπάρχει

σημαντική διαφορά μεταξύ του αριθμού των θετικών και των αρνητικών ετικετών, η ακρίβεια από μόνη της δεν μπορεί να περιγράψει πλήρως την κατάσταση.

Απώλεια

Μια συνάρτηση απώλειας είναι μια συνάρτηση που συγκρίνει τις προβλεπόμενες τιμές εξόδου ενός νευρωνικού δικτύου με τις αναμενόμενες τιμές εξόδου και υπολογίζει τη διαφορά μεταξύ τους. Η συνάρτηση απωλειών χρησιμοποιείται για να αξιολογηθεί πόσο καλά αποδίδει το μοντέλο στα δεδομένα εκπαίδευσης. Κατά τη διάρκεια της εκπαίδευσης, ο στόχος είναι η ελαχιστοποίηση της απώλειας μεταξύ των προβλεπόμενων και των αναμενόμενων εξόδων [30].

Προσαρμόζουμε τις υπερπαραμέτρους για να ελαχιστοποιήσουμε τη μέση απώλεια, η οποία περιλαμβάνει την εύρεση των τιμών των βαρών, w^T , και των πολώσεων, b , που οδηγούν στην ελάχιστη τιμή του J , που αντιπροσωπεύει τη μέση απώλεια:

$$J(w^T, b) = \frac{1}{m} \sum_{i=1}^m L(\hat{y}^{(i)}, y^{(i)}).$$

Σε αναλογία με την έννοια των υπολειμμάτων στη στατιστική, τα οποία ποσοτικοποιούν την απόκλιση των πραγματικών τιμών y από τη γραμμή παλινδρόμησης (προβλεπόμενες τιμές), στόχος μας είναι η ελαχιστοποίηση της καθαρής απόστασης.

Υπάρχουν δύο πρωταρχικοί τύποι συναρτήσεων απωλειών που χρησιμοποιούνται στην επιβλεπόμενη μάθηση, οι οποίοι αντιστοιχούν στους δύο κύριους τύπους νευρωνικών δικτύων: οι συναρτήσεις απωλειών παλινδρόμησης και οι συναρτήσεις απωλειών ταξινόμησης.

1. **Συναρτήσεις απωλειών παλινδρόμησης:** χρησιμοποιούνται στα νευρωνικά δίκτυα παλινδρόμησης, όπου το μοντέλο προβλέπει μια τιμή εξόδου που αντιστοιχεί σε μια τιμή εισόδου αντί για προκαθορισμένες ετικέτες. Παραδείγματα συναρτήσεων απωλειών παλινδρόμησης περιλαμβάνουν το μέσο τετραγωνικό σφάλμα και το μέσο απόλυτο σφάλμα.
2. **Συναρτήσεις απωλειών ταξινόμησης:** χρησιμοποιούνται στα νευρωνικά δίκτυα ταξινόμησης. Παράγουν ένα διάνυσμα πιθανοτήτων που δείχνει την πιθανότητα μια είσοδος να ανήκει σε διάφορες προκαθορισμένες κατηγορίες, επιτρέποντας στο δίκτυο να επιλέξει την κατηγορία με την υψηλότερη πιθανότητα. Παραδείγματα συναρτήσεων απωλειών ταξινόμησης περιλαμβάνουν τη δυαδική διασταυρούμενη εντροπία, την κατηγορική διασταυρούμενη εντροπία.

Σε αντίθεση με την ακρίβεια, η απώλεια δεν εκφράζεται ως ποσοστό αλλά ως άθροισμα των σφαλμάτων που έγιναν σε κάθε δείγμα στα σύνολα εκπαίδευσης ή επικύρωσης. Κατά τη διαδικασία εκπαίδευσης, η ελαχιστοποίηση της τιμής της συνάρτησης απώλειας είναι ο στόχος προκειμένου να καθοριστούν οι βέλτιστες τιμές παραμέτρων για το μοντέλο, όπως τα βάρη σε ένα νευρωνικό δίκτυο.

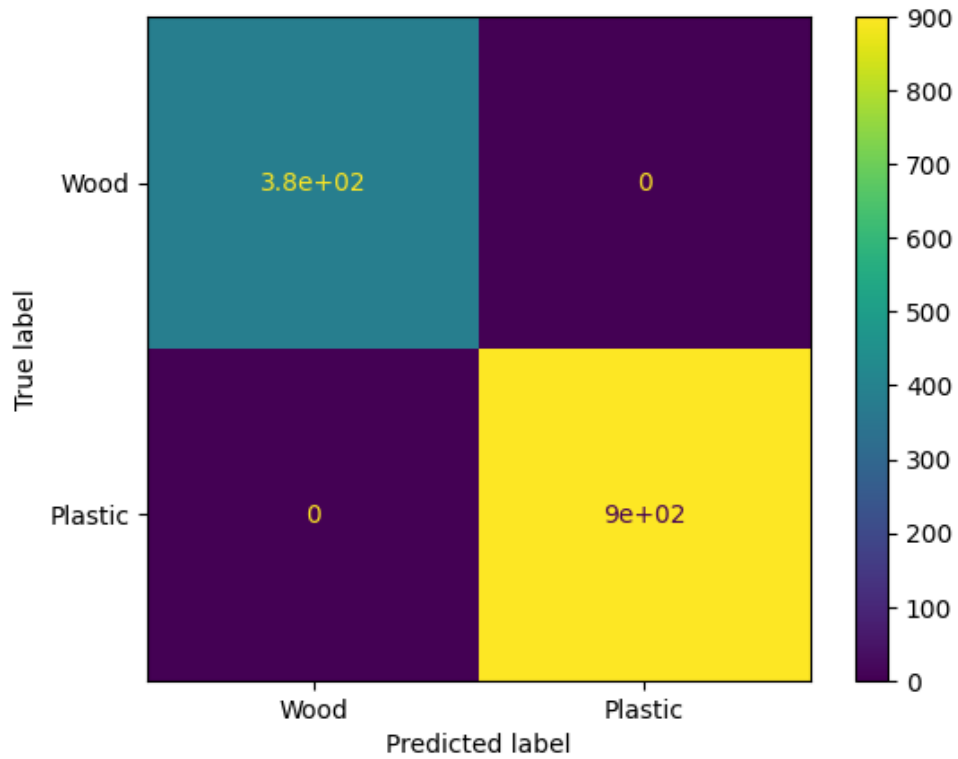
Αποτελέσματα και Συζήτηση

Το πρώτο αποτέλεσμα σχετίζεται με την αρχιτεκτονική του μοντέλου που χρησιμοποιήθηκε. Αφού δοκιμάσαμε διαφορετικά μοντέλα, εμείς καταλήξαμε στο συμπέρασμα ότι το ResNet50 είναι αυτό που ταιριάζει καλύτερα στους σκοπούς μας και το πιο βέλτιστο για την παρούσα έρευνα από την άποψη της χρήσης GPU και των διαθέσιμων πόρων, της ακρίβειας πρόβλεψης και τη διαχείριση ενός μεσαίου μεγέθους συνόλου δεδομένων εικόνων.

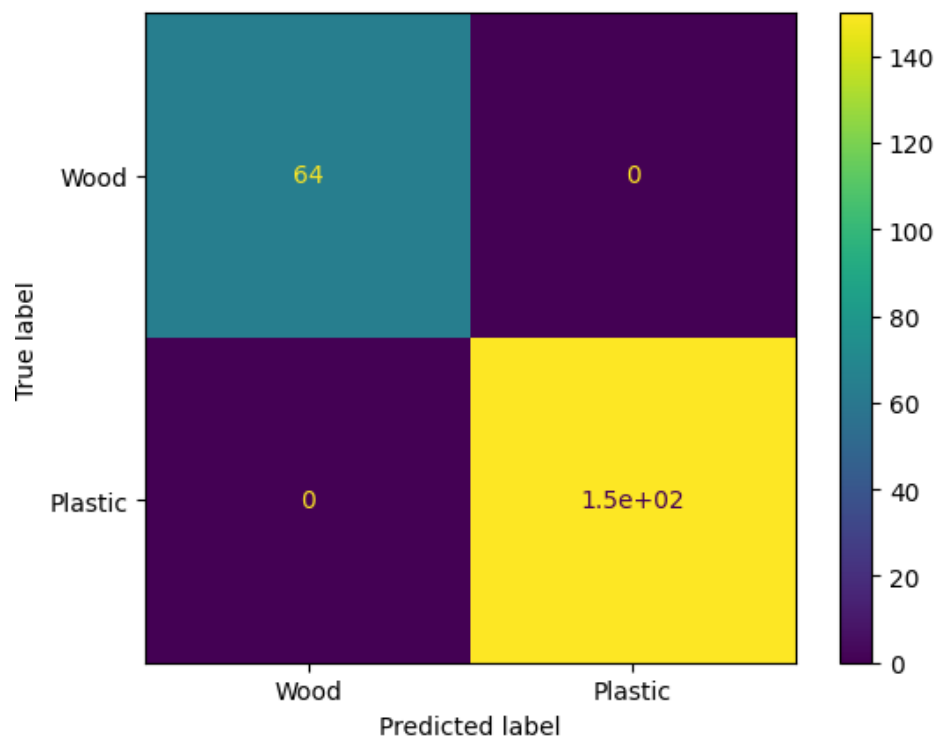
Το δεύτερο αποτέλεσμα αφορά τους ταξινομητές. Οι τρεις ταξινομητές που δημιουργήθηκαν, δίνουν καλύτερα αποτελέσματα όταν συνδυάζονται εν αντιθέσει με τον καθένα χωριστά. Επιπλέον, ο μέσος ταξινομητής είναι αυτός που παρέχει τα καλύτερα αποτελέσματα. Ο μέσος ταξινομητής δημιουργείται προσθέτοντας τις τιμές των ποσοστών πιθανότητας κάθε κλάσης των τριών άλλων ταξινομητών (RGB, NIR, VIS) και στη συνέχεια διαιρείται με το 3 για να προκύψει το μέσο ποσοστό. Πιο συγκεκριμένα, αξίζει να αναφερθεί ότι η ταξινόμηση στα νευρωνικά δίκτυα λειτουργεί με την εξαγωγή ενός διανύσματος πιθανοτήτων. Αυτές οι πιθανότητες αναφέρονται στο πόσο καλά το δεδομένο εισόδου ταιριάζει σε κάθε μία από τις προκαθορισμένες κατηγορίες και στη συνέχεια τα Νευρωνικά Δίκτυα Ταξινόμησης επιλέγουν την κατηγορία-κλάση με την υψηλότερη πιθανότητα ως τελική έξοδο.

Τα σημαντικότερα αποτελέσματα που παρέχονται από τον μέσο ταξινομητή είναι:

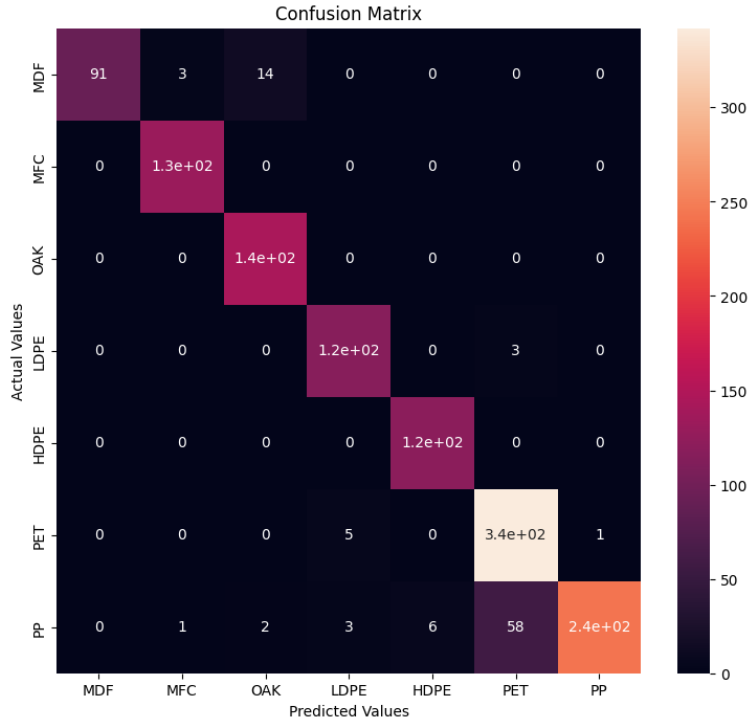
- **Εμπιστοσύνη** σε περίπτωση σωστής πρόβλεψης: 80.968%
- **Εμπιστοσύνη** σε περίπτωση λανθασμένης εκτίμησης: 52.705%
- **Ακρίβεια**: 92.523%
- **Ακρίβεια**: 93.457% με την αντίστροφη διαδικασία της επαύξησης.
- Ακολουθούν οι **πίνακες σύγχυσης** του μέσου ταξινομητή. Ο δεύτερος και ο τέταρτος πίνακας απεικονίζουν τα αποτελέσματα της αντίστροφης επαύξησης. Η αντίστροφη διαδικασία της επαύξησης είναι μια τεχνική παρόμοια με την κωδικοποίηση επανάληψης (repetition coding) στις τηλεπικοινωνίες. Πιο αναλυτικά, παρουσιάζουν ποιο είναι το αποτέλεσμα εάν λάβουμε υπόψη την πλειοψηφία των εκδοχών από κάθε εικόνα. Για παράδειγμα, εάν 4 από τις 6 εκδοχές μιας εικόνας επιλέγονται ως LDPE, τότε θεωρείται ως LDPE.



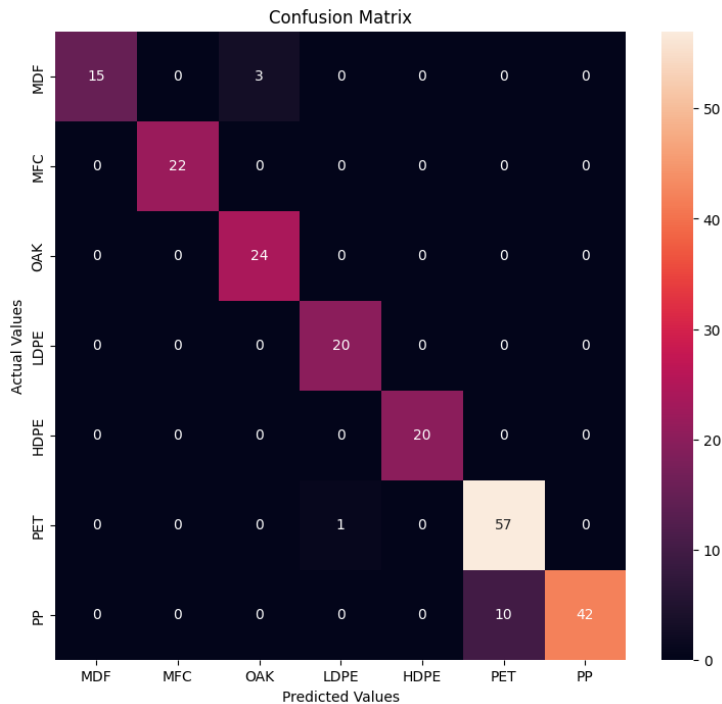
Εικόνα 10. Πίνακας σύγχυσης για ξύλο και πλαστικό του μέσου ταξινομητή.



Εικόνα 11. Πίνακας σύγχυσης για ξύλο και πλαστικό του μέσου ταξινομητή (πλειοψηφία των εκδοχών με αντιστροφή της επαύξεσης).

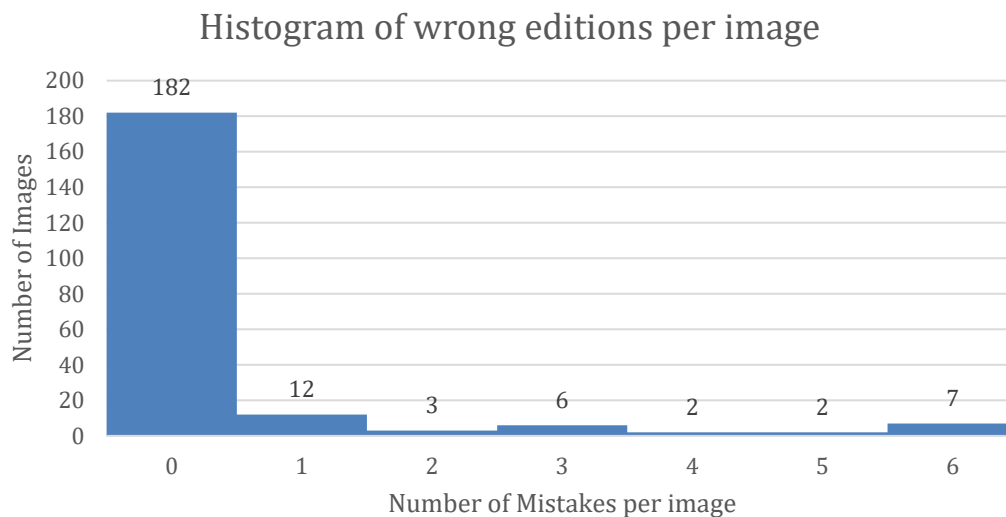


Εικόνα 12. Πίνακας σύγχυσης και των 7 κατηγοριών του μέσου ταξινομητή. Παρουσιάζονται οι ακόλουθες λανθασμένες ταξινομήσεις: 17 δείγματα MDF, 3 εικόνες LDPE, 6 PET και 70 PP, ενώ δεν έγιναν λάθη στις κατηγορίες MFC, Βελανιδιά (OAK) ή HDPE.



Εικόνα 13. Πίνακας σύγχυσης και των 7 κατηγοριών του μέσου ταξινομητή (Αντιστροφή της επαύξεσης). Μπορούμε να παρατηρήσουμε ότι μόνο 14 από 214 σύνολα των 6 εκδοχών κατατάσσονται λάθος, εξού και 93.457% ακρίβεια.

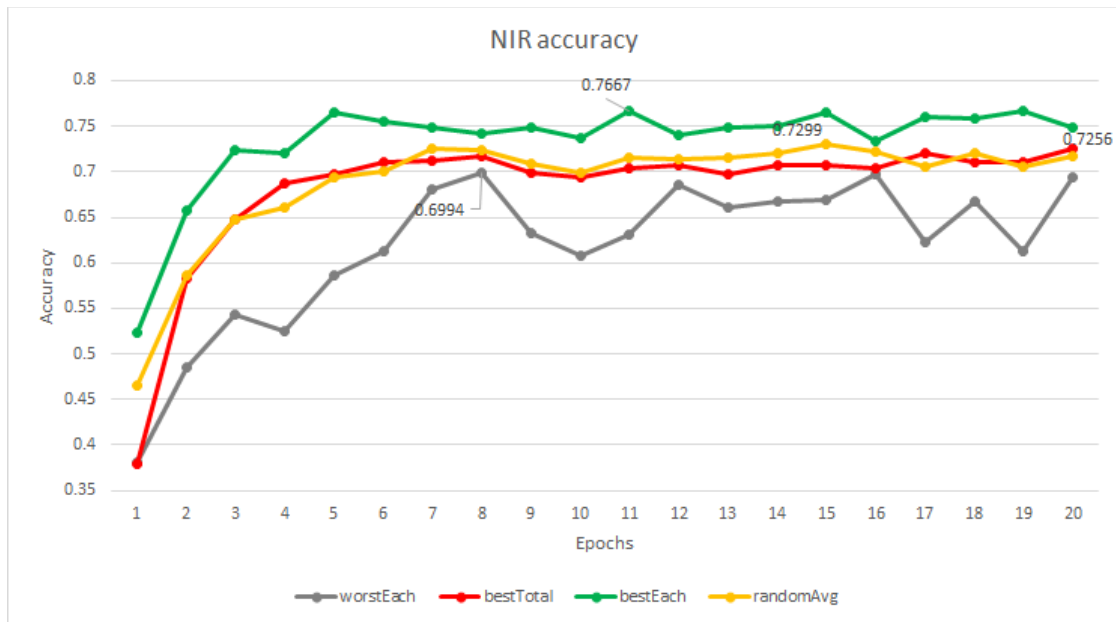
Ένα συμπληρωματικό αποτέλεσμα που εξάγεται από τον μέσο ταξινομητή και έχει τη δική του αξία είναι ο αριθμός των λανθασμένων εκδοχών ανά εικόνα. Για να το παρουσιάσουμε αυτό, χρησιμοποιούμε τη βοήθεια του κάτωθι ιστογράμματος. Στο ιστόγραμμα επιδεικνύεται ο αριθμός των εικόνων που έχουν κατανεμημένες ατυχώς από μηδέν έως και τις έξι τους εκδοχές. Για παράδειγμα, τέσσερα λάθη ανά εικόνα σημαίνει ότι τέσσερις από τις έξι εκδοχές κατατάχθηκαν σε λανθασμένη κατηγορία, ενώ το μηδέν σημαίνει ότι όλες κατατάχθηκαν σωστά.



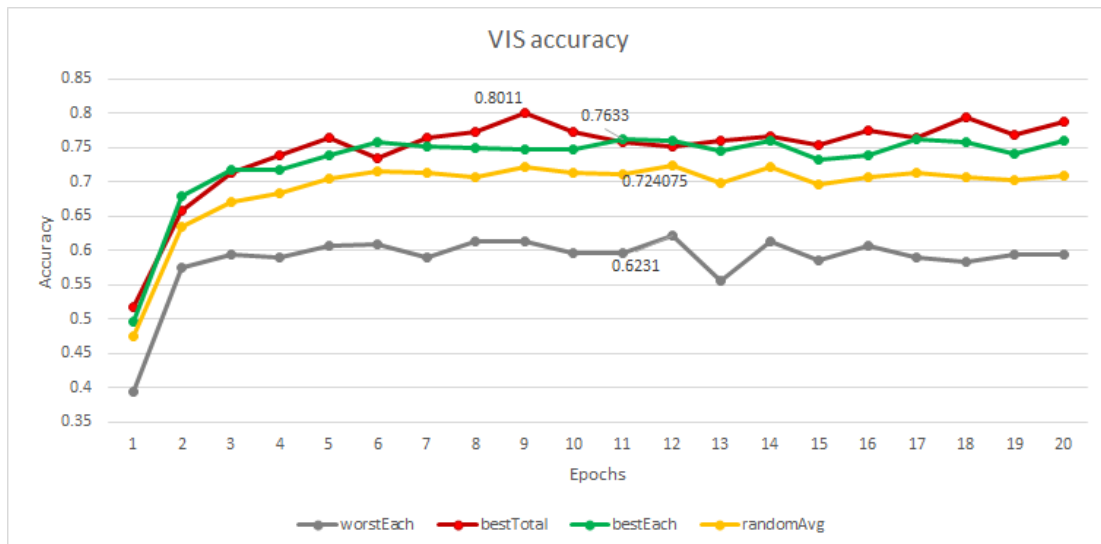
Εικόνα 14. Ιστόγραμμα λανθασμένων εκδοχών ανά εικόνα. Έξι λάθη ανά εικόνα σημαίνει ότι και οι έξι εκδόσεις κατατάχθηκαν στη λανθασμένη κατηγορία, ενώ το μηδέν σημαίνει ότι όλες κατατάχθηκαν σωστά.

Το τελευταίο κομμάτι των συμπερασμάτων ασχολείται με τον αντίκτυπο που θα είχε ένας ανεπαρκής εξοπλισμός στην ακρίβεια των ταξινομητών μαζί με μια μελέτη επάνω στην επιλογή ζωνών. Η επιρροή του χαμηλότερης ποιότητας εξοπλισμού καταδεικνύεται με την εξέταση της ακρίβειας των πολυφασματικών ταξινομητών. Η έρευνα δείχνει ότι οι κάμερες NIR παρουσιάζουν μόλις 2,07% έλλειμμα ακρίβειας, ενώ οι κάμερες VIS εμφανίζουν οριακή μείωση κατά 1,69% σε σύγκριση με την καλύτερη ακρίβεια που επιτυγχάνεται με κάμερες NIR (78,74%) και VIS (81,72%), αντίστοιχα. Τα επόμενα διαγράμματα απεικονίζουν την εξέλιξη της ακρίβειας κατά τη διάρκεια των εποχών έως ότου εμφανιστεί η υπερπροσαρμογή.

Επιπρόσθετα, πραγματοποιήσαμε πειράματα επάνω στη μείωση των διαστάσεων για πολυφασματικές κάμερες, παραλληλίζοντας την επιλογή χαρακτηριστικών με τη χρήση μόνο τριών διαφορετικών ζωνών και αξιολογώντας τις επιπτώσεις της. Ακόμη, επιλέγοντας διάφορα σύνολα τριών ζωνών από τις πολυφασματικές κάμερες, αποδείξαμε ότι η χωρική εστίαση και η εστίαση στο σχήμα απέχουν σημαντικά από το απόλυτο 100% της ακρίβειας, αφήνοντας περιθώρια βελτίωσης σε μία εστίαση στο φασματικό αποτύπωμα.



Εικόνα 15. Διάγραμμα που παρουσιάζει την εξέλιξη της ακρίβειας των συνόλων τριών ζωνών της κάμερας NIR καθώς προχωρούν οι εποχές. Η κορυφή κάθε γραμμής απεικονίζεται με δεκαδικό αριθμό. Επεξήγηση γραμμών: "randomAvg" αναφέρεται στο μέσο όρο των συνόλων τυχαίας επιλογής, το "bestTotal" αντιπροσωπεύει τη δεύτερη τεχνική ("καλύτερη εν συνόλω"), ενώ τα "bestEach" και "worstEach" αντιπροσωπεύουν την τρίτη και την τέταρτη μέθοδο που εφαρμόσαμε, αντίστοιχα.



Εικόνα 16. Διάγραμμα που παρουσιάζει την εξέλιξη της ακρίβειας των συνόλων τριών ζωνών της κάμερας VIS καθώς προχωρούν οι εποχές. Η κορυφή κάθε γραμμής απεικονίζεται με δεκαδικό αριθμό. Επεξήγηση γραμμών: "randomAvg" αναφέρεται στο μέσο όρο των συνόλων τυχαίας επιλογής, το "bestTotal" αντιπροσωπεύει τη δεύτερη τεχνική ("καλύτερη εν συνόλω"), ενώ τα "bestEach" και "worstEach" αντιπροσωπεύουν την τρίτη και την τέταρτη μέθοδο που εφαρμόσαμε, αντίστοιχα.

Αυτό υποστηρίζεται από το γεγονός ότι η μέση ακρίβεια των τυχαία επιλεγμένων συνόλων ζωνών διαφέρει κατά λιγότερο από 8% από την υψηλότερη ακρίβεια που επιτυγχάνεται από σύνολα τριών ζωνών και λιγότερο από 10% από την αξιοποίηση της PCA και τη συμπερίληψη

όλων των ζωνών. Προς υπεράσπιση των παραπάνω, έρχεται και το γεγονός πως όταν επιλέγονται οι τρεις χειρότερες ζώνες των καμερών NIR, η διαφορά στα αποτελέσματα σε σύγκριση με την καλύτερη επιλεγμένη τριάδα ζωνών είναι μικρότερη από 7%.

Η προτεινόμενη έρευνά μας αποδεικνύει ότι τα ΣΝΔ μπορούν να χρησιμοποιηθούν για την ακριβή και αποτελεσματική ταξινόμηση των απορριμμάτων ξύλου. Παρέχοντας μια γρήγορη και στοχευμένη μέθοδο διαλογής, τα ΣΝΔ έχουν τη δυνατότητα να βελτιώσουν την αποδοτικότητα και την αποτελεσματικότητα της διαχείρισης των ξύλινων απορριμμάτων, οδηγώντας σε εξοικονόμηση κόστους και μείωση των περιβαλλοντικών επιπτώσεων. Μέσω αυτών των ευρημάτων, προκύπτουν νέες δυνατότητες για την αξιοποίηση νευρωνικών δικτύων σε παρόμοιες πρακτικές διαχείρισης αποβλήτων και σε άλλες βιομηχανίες, όπως είναι τα πλαστικά, τα μέταλλα και τα απόβλητα κατασκευών και κατεδαφίσεων.

Συνεισφορά

Πρώτον, ως σημαντική συμβολή θεωρείται η δημιουργία του συνόλου δεδομένων εικόνας των προϊόντων από ξύλο και πλαστικό. Αυτό το σύνολο δεδομένων δημιουργήθηκε με τη λήψη εικόνων από τρεις διαφορετικούς τύπους καμερών (RGB, VIS, NIR). Στη συνέχεια ακολούθησε η επαύξηση των δεδομένων που περιλάμβανε μικρές αλλαγές στις εικόνες που καταγράφηκαν και η οποία μας επέτρεψε να καταλήξουμε σε 6522 δείγματα εικόνων για κάθε τύπο κάμερας, ώστε να έχουμε ασφαλέστερα τελικά αποτελέσματα.

Επιπλέον, μια άλλη συμβολή της παρούσας έρευνας είναι η δοκιμή του TensorFlow έναντι του PyTorch υπό τις ίδιες συνθήκες και η καθιέρωση της άποψης ότι το TensorFlow είναι ευκολότερο στη χρήση για σκοπούς μηχανικής μάθησης εισαγωγικού επιπέδου, ενώ το PyTorch μπορεί να αποδειχθεί πιο χρήσιμο λόγω των πολύπλοκων ρυθμίσεων που διαθέτει, αν και πιο περίπλοκο.

Τρίτον, η κύρια συνεισφορά μας είναι η ακρίβεια διαλογής που επιτυγχάνεται. Αναλυτικότερα, πετύχαμε ακρίβεια 100% στη διάκριση του ξύλου από το πλαστικό και 93.45% στην ταξινόμηση κάθε υλικού από τις επτά κλάσεις με την αντιστροφή της επαύξησης. Αυτό το αποτέλεσμα οφείλεται στην ανάπτυξη του Μέσου Ταξινομητή, ο οποίος είναι ένας συνδυασμός των άλλων τριών ταξινομητών και μπορούμε να συμπεράνουμε με ασφάλεια ότι ο Μέσος ταξινομητής είναι η καλύτερη διαθέσιμη επιλογή, αφού παράγει τα καλύτερα αποτελέσματα.

Επιπρόσθετα, η έρευνά μας περιλάμβανε πολλές δοκιμές σε διάφορα μοντέλα ΣΝΔ για σκοπούς ταξινόμησης εικόνων και, ως εκ τούτου, το συμπέρασμα που εξάγεται είναι ότι το Resnet50 μπορεί να χειριστεί με επιτυχία ένα μεσαίου μεγέθους σύνολο δεδομένων, να παρέχει αρκετά ακριβή τελικά αποτελέσματα και να συνεργαστεί αποτελεσματικά με υπολογιστικές μονάδες που δεν ανήκουν στην υψηλότερη κατηγορία.

Τέλος, ένα ουσιαστικό συμπέρασμα που προέκυψε αφορά τη χρήση εξοπλισμού μικρότερων δυνατοτήτων. Ως τελική πτυχή της έρευνάς μας, πραγματοποιήσαμε πειράματα σχετικά με τη μείωση των διαστάσεων για τις πολυφασματικές κάμερες, που χρησίμευσαν ως επίδειξη της επιλογής χαρακτηριστικών σε περιπτώσεις όπου ο εξοπλισμός μπορεί να μην πληροί τα βέλτιστα πρότυπα, και αξιολογήσαμε τις επιπτώσεις της. Είναι αξιοσημείωτο ότι τα αποτελέσματα έδειξαν

μια οριακή μείωση μικρότερη του 2.1% και για τους δύο τύπους καμερών. Το εύρημα αυτό αξιοποιείται ως πειστική απόδειξη ότι ακόμη και με μη βέλτιστο εξοπλισμό μπορούν να επιτευχθούν ικανοποιητικά αποτελέσματα. Επιπλέον, επιλέγοντας διαφορετικά σύνολα τριών ζωνών από τις πολυφασματικές κάμερες, αποδείξαμε ότι η χωρική εστίαση και η εστίαση στο σχήμα των αντικειμένων δεν παίζουν τόσο κυρίαρχο ρόλο σε σύγκριση με το φασματικό τους αποτύπωμα. Τον ισχυρισμό αυτό στηρίζει το γεγονός ότι η μέση ακρίβεια των συνόλων των τυχαία επιλεγμένων ζωνών διαφέρει κατά λιγότερο από 8% από την υψηλότερη ακρίβεια που επιτυγχάνεται από σύνολα τριών ζωνών και λιγότερο από 10% από τη χρήση της PCA και την αξιοποίηση όλων των ζωνών.

Σύνοψη Διπλωματικής

Συμπερασματικά, η παρούσα διπλωματική διερευνά τις δυνατότητες χρήσης Συνελκτικών Νευρωνικών Δικτύων (ΣΝΔ) για την ταξινόμηση των απορριμμάτων ξύλου. Η προτεινόμενη έρευνα καταδεικνύει τη σκοπιμότητα και την αποτελεσματικότητα της χρήσης ΣΝΔ και θα μπορούσε να συμβάλει στην καθιέρωση των ΣΝΔ σε περισσότερους τομείς διαχείρισης βιομηχανικών αποβλήτων. Ως αποτέλεσμα, η προτεινόμενη έρευνα έχει τη δυνατότητα να συμβάλει σημαντικά στον τομέα της διαχείρισης απορριμμάτων ξύλου και στην ανάπτυξη βιώσιμων βιομηχανιών με βάση το ξύλο, βοηθώντας περαιτέρω την κυκλική οικονομία.

Η βαθιά μάθηση έχει επιδείξει τεράστιες δυνατότητες σε πολλούς τομείς και ιδιαίτερα τα ΣΝΔ στην επεξεργασία εικόνας. Η παρούσα έρευνα βασίστηκε στο κορυφαίο μοντέλο ταξινόμησης εικόνας ResNet, το οποίο είχε εισήχθη για πρώτη φορά στον τομέα της γενικής διαλογής αποβλήτων και χρησιμοποιήθηκε ως ταξινομητής για τα απόβλητα σε ένα πολύπλοκο βιομηχανικό περιβάλλον. Προτάθηκε μια εφικτή μέθοδος διαχωρισμού με βάση το ResNet και περιγράφηκε λεπτομερώς ολόκληρη η διαδικασία, συμπεριλαμβανομένης της δημιουργίας εικόνων, της δημιουργίας του συνόλου εκπαίδευσης και του συνόλου επικύρωσης, της επαύξησης εικόνων, του συντονισμού του μοντέλου ResNet50, του διαχωρισμού αποβλήτων ξύλου, της διαλογής ξύλου από πλαστικά απορρίμματα και της βελτιστοποίησης αυτών των αποτελεσμάτων. Τα αποτελέσματα δείχνουν ότι η προτεινόμενη μέθοδος έχει τη δυνατότητα να διευκολύνει την αυτόματη ταξινόμηση των αποβλήτων ξύλου, η οποία είναι σημαντική για πολλές βιομηχανίες και δήμους στην προσπάθεια επίτευξης καλύτερης διαχείρισης του κύκλου ζωής των ξύλινων προϊόντων, τη μείωση της περιβαλλοντικής επιβάρυνσης και την αύξηση της βιωσιμότητας. Συγκεκριμένα, εξήχθησαν κάποια καίρια συμπεράσματα ως εξής.

Πρώτον, αξιολογήθηκαν διάφορες αρχιτεκτονικές μοντέλων για να προσδιοριστεί η καταλληλότερη για τις απαιτήσεις μας. Κατά συνέπεια, διαπιστώθηκε ότι το ResNet50 επέδειξε εξαιρετικές επιδόσεις στο χειρισμό συνόλων δεδομένων μεσαίου μεγέθους χωρίς να απαιτεί υπερβολικά ποσά υπολογιστικών πόρων. Δεύτερον, επιτύχαμε 93.45% ακρίβεια στο διαχωρισμό καθεμιάς από τις επτά κλάσεις και την τέλεια βαθμολογία 100% στο διαχωρισμό του ξύλου από το πλαστικό.

Τρίτον, πραγματοποιήσαμε πειράματα μείωσης των διαστάσεων για τις πολυφασματικές κάμερες, που χρησιμεύουν ως παράδειγμα επιλογής χαρακτηριστικών σε σενάρια όπου ο

εξοπλισμός μπορεί να είναι ανεπαρκής, και αξιολογήσαμε την επίδρασή της. Τα αποτελέσματα έδειξαν μείωση μόλις κατά 1.69% σε σύγκριση με τη χρήση όλων των ζωνών στο ορατό φάσμα (VIS) και μείωση κατά 2.07% στο φάσμα του εγγύς υπέρυθρου (NIR).

ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ

Η ταξινόμηση των ξύλινων απορριμμάτων και γενικά η ταξινόμηση των αποβλήτων είναι ακόμη ανοικτά πεδία που πρέπει να ανακαλυφθούν και έχουν μια σειρά από ανοικτά ζητήματα που πρέπει να μελετηθούν. Πρώτον, υπάρχει ανάγκη επέκτασης της μελέτης ώστε να συμπεριληφθούν υπερφασματικές κάμερες, οι οποίες θα ήταν επωφελείς για την εξέταση ενός ευρύτερου πεδίου κατηγοριών ξύλου. Η υπερφασματική απεικόνιση προσφέρει μια μοναδική ευκαιρία για τη λήψη λεπτομερών φασματικών πληροφοριών από τα υλικά ξύλου, επιτρέποντας μια πιο ολοκληρωμένη κατανόηση της σύνθεσης και των χαρακτηριστικών τους. Αυτή η επέκταση της τεχνολογίας μπορεί να βελτιώσει την αποτελεσματικότητα των διαδικασιών διαχωρισμού των απορριμμάτων ξύλου, οδηγώντας σε καλύτερη ακρίβεια διαλογής τους. Ακόμη, η χρησιμοποίηση υπερφασματικών καμερών θα μπορούσε να συμβάλει στην ανάπτυξη προηγμένων αλγορίθμων και τεχνικών μηχανικής μάθησης ειδικά προσαρμοσμένων στην ανάλυση αποβλήτων ξύλου, επιτρέποντας αυτοματοποιημένες μεθόδους διαχωρισμού.

Η βελτίωση της ακρίβειας της ταξινόμησης των ξύλινων απορριμμάτων είναι ένας κρίσιμος τομέας για μελλοντική έρευνα. Ενώ έχει σημειωθεί σημαντική πρόοδος στην παρούσα μελέτη, υπάρχουν ακόμη περιθώρια για την ενίσχυση της ακρίβειας και της αξιοπιστίας των αλγορίθμων ταξινόμησης. Οι ερευνητές μπορούν να διερευνήσουν προηγμένες τεχνικές μηχανικής μάθησης ή μοντέλα για την επίτευξη υψηλότερων ποσοστών ακρίβειας. Αυτό μπορεί να συμβεί με την αντιμετώπιση ζητημάτων όπως η μεταβλητότητα μεταξύ των κλάσεων και οι λεπτές φασματικές διαφορές τους.

Ο αντίκτυπος του φωτισμού στην πολυφασματική απεικόνιση είναι επίσης μια σημαντική πτυχή που πρέπει να διερευνηθεί περαιτέρω. Η κατανόηση του τρόπου με τον οποίο οι διαφορετικές συνθήκες φωτισμού επηρεάζουν τη φασματική απόκριση των υλικών ξύλου και η διερεύνηση τεχνικών αντιστάθμισης των διακυμάνσεων του φωτισμού θα οδηγήσει σε πιο συνεπή αποτελέσματα. Οι ερευνητές θα πρέπει να εξετάσουν παράγοντες όπως οι γωνίες φωτισμού, η ένταση και η επιρροή των περιβαλλοντικών συνθηκών για την ανάπτυξη ισχυρών αλγορίθμων που μπορούν να χειριστούν διαφορετικές καταστάσεις φωτισμού.

Η μελέτη του φασματικού αποτυπώματος αντί της αποκλειστικής στήριξης στο σχήμα των αντικειμένων αποτελεί μια ενδιαφέρουσα κατεύθυνση για μελλοντική έρευνα στην ταξινόμηση των απορριμμάτων ξύλου. Με την ανάλυση των μοναδικών φασματικών χαρακτηριστικών των διαφόρων υλικών ξύλου, οι ερευνητές μπορούν να αποκτήσουν πολύτιμες γνώσεις σχετικά με τη σύνθεση και τις ιδιότητές τους. Αυτή η μετατόπιση της εστίασης μπορεί να απαιτήσει την ανάπτυξη καινοτόμων τεχνικών επεξεργασίας δεδομένων που μπορούν να εξάγουν και να αναλύουν αποτελεσματικά τα φασματικά αποτυπώματα των αποβλήτων ξύλου. Με τη διερεύνηση της φασματικής διάστασης, μπορούν να αποκαλυφθούν νέες δυνατότητες για την αποτελεσματική διαλογή αποβλήτων ξύλου.

Επιπροσθέτως, οι μελλοντικές μελέτες θα πρέπει να εξετάσουν τη χρήση καλύτερων μονάδων επεξεργασίας, όπως οι κάρτες γραφικών υψηλής απόδοσης και η αυξημένη χωρητικότητα της μνήμης RAM, ώστε να βελτιώσουν την ταχύτητα και την αποτελεσματικότητα του διαχωρισμού των ξύλινων απορριμμάτων. Αυτοί οι προηγμένοι υπολογιστικοί πόροι μπορούν να επιταχύνουν σημαντικά την επεξεργασία μεγάλων υπερφασματικών συνόλων δεδομένων, επιτρέποντας δυνατότητες ταξινόμησης σε πραγματικό ή σχεδόν πραγματικό χρόνο.

Πέραν αυτού, η ενσωμάτωση διαφορετικών ειδών ξύλινων απορριμμάτων στο σύνολο δεδομένων και η εξασφάλιση ότι το μέγεθος και η ποικιλομορφία τους είναι αντιπροσωπευτικά των πραγματικών σεναρίων θα ενισχύσει περαιτέρω την ακρίβεια και τη δυνατότητα γενίκευσης των μοντέλων διαχωρισμού αποβλήτων ξύλου. Η έρευνά μας επικεντρώνεται πρωτίστως σε ένα περιορισμένο σύνολο τύπων ξύλου, ενώ η διερεύνηση πρόσθετων κατηγοριών, όπως μεγάλα κομμάτια ξύλου που δεν χωράνε εξ ολοκλήρου στην κάμερα ή πριονίδι που είναι μικροσκοπικό, θα προσφέρει μια πιο ολοκληρωμένη κατανόηση των προκλήσεων και των απαιτήσεων που αφορούν κάθε τύπο.

Επιπλέον, η παρούσα έρευνα ανοίγει το δρόμο για τη διερεύνηση των δυνατοτήτων χρήσης των ΣΝΔ σε διάφορους άλλους τομείς διαχείρισης βιομηχανικών αποβλήτων πέραν των απορριμμάτων ξύλου. Παρουσιάζοντας την αποτελεσματικότητα των ΣΝΔ στην ταξινόμηση αποβλήτων ξύλου, ανοίγονται ευκαιρίες για την αξιοποίηση παρόμοιων τεχνικών μηχανικής μάθησης σε τομείς διαφορετικών υλικών αποβλήτων, όπως τα πλαστικά, τα μέταλλα και τα απορρίμματα κατασκευών και κατεδαφίσεων. Αυτή η διεύρυνση του πεδίου εφαρμογής έχει τη δυνατότητα να φέρει επανάσταση στις πρακτικές διαχείρισης αποβλήτων σε πολλές βιομηχανίες, επιτρέποντας πιο αποτελεσματικές διαδικασίες διαλογής, ανακύκλωσης και ανάκτησης πόρων.

Παρά την πρόοδο που έχει σημειωθεί, παραμένει μεγάλη πρόκληση η εφαρμογή συστημάτων διαλογής ξύλινων απορριμμάτων σε πραγματικές συνθήκες, πέρα από τα όρια των προστατευμένων εργαστηριακών περιβαλλόντων. Παράγοντες όπως ο μεταβαλλόμενος φωτισμός, οι ανεξέλεγκτες και σκονισμένες περιβαλλοντικές συνθήκες και η ανάγκη προσαρμογής των αλγορίθμων και των μοντέλων για την αντιμετώπιση απρόβλεπτων σεναρίων δημιουργούν σημαντικές προκλήσεις που δεν εμφανίζονται σε μια προσομοίωση. Η μελλοντική έρευνα θα πρέπει να επικεντρωθεί στην αντιμετώπιση αυτών των πρακτικών εμποδίων, ώστε να διασφαλιστεί η επιτυχής ανάπτυξη και η ευρεία υιοθέτηση των τεχνολογιών διαχωρισμού αποβλήτων ξύλου σε πραγματικές συνθήκες.

Τέλος, είναι ζωτικής σημασίας τα σύνολα δεδομένων που αφορούν τα απορρίμματα ξύλου να καταστούν ελεύθερα προσβάσιμα σε όλους, διευκολύνοντας την πρόοδο της επιστημονικής γνώσης. Εξασφαλίζοντας τη διαθεσιμότητα και την εύκολη πρόσβαση σε αυτά τα σύνολα δεδομένων, οι ερευνητές μπορούν να συνεργάζονται αποτελεσματικότερα, να ανταλλάσσουν πολύτιμες γνώσεις και να αντιμετωπίζουν συλλογικά τις προκλήσεις που σχετίζονται με τη διαχείριση των αποβλήτων ξύλου. Τα ανοικτά σύνολα δεδομένων όχι μόνο ενθαρρύνουν τις διεπιστημονικές συνεργασίες αλλά και την καινοτομία, επιτρέποντας στους ερευνητές να αναπτύξουν και να αξιολογήσουν νέους αλγορίθμους, μοντέλα και τεχνικές, ενώ η ελεύθερη διαθεσιμότητα αυτών των συνόλων δεδομένων προωθεί τη διαφάνεια, την αναπαραγωγιμότητα και την υπευθυνότητα. Με τον εκδημοκρατισμό της πρόσβασης σε σύνολα δεδομένων για τα απορρίμματα, η επιστήμη μπορεί να προχωρήσει περαιτέρω, ανοίγοντας το δρόμο για βιώσιμες

πρακτικές διαχείρισης αποβλήτων και τεκμηριωμένη λήψη αποφάσεων που ωφελούν τόσο το περιβάλλον όσο και την κοινωνία στο σύνολό της.

Chapter 1

Introduction

1.1. Motivation

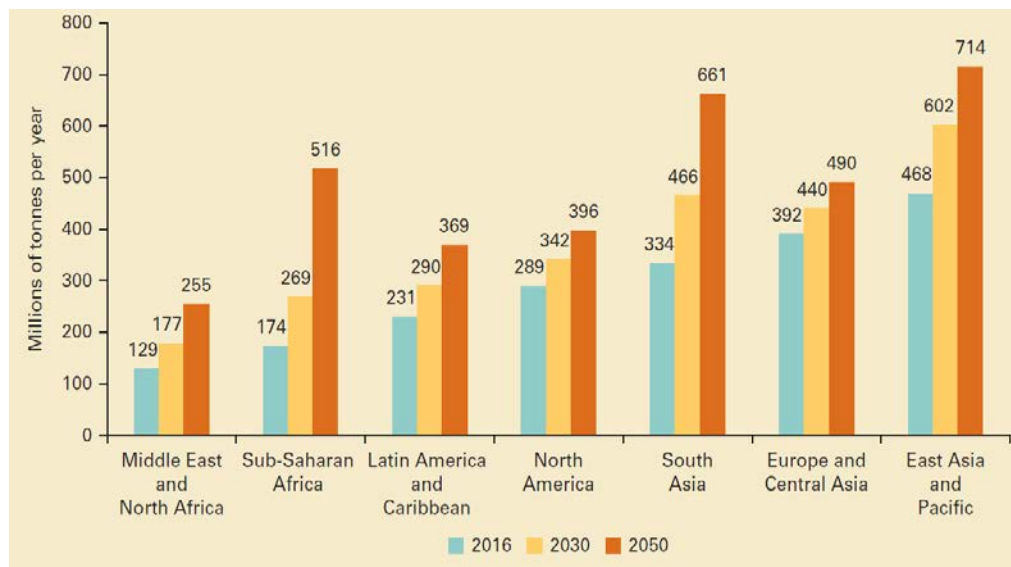


Figure 1. Projected waste generation, by region (millions of tonnes/year).

In the modernization and urbanization of daily life, a large amount of Municipal Solid Waste (MSW) is inevitably generated. Wood waste constitutes the 6th largest category of MSW. In Europe together with Central Asia, in 2016, 392 million tonnes of MSW were produced, which translates to roughly 23.5 millions of tonnes of Wood Waste, since wood waste takes up almost 6% of the total MSW proportion [1].

Finite resources and environmental hazards have necessitated the transition to the circular economy model. Due to the rapid increase projected in MSW production worldwide and for sustainability purposes, the European Union has set stricter targets for recycling and begun to encourage governments to recycle solid waste. However, wood waste segregation through artificial intelligence techniques still remains an unexplored field.

The old traditional way of separating waste is the hand-picking method, whereby humans are employed to separate out the different categories of materials. However, because of the long-term exposure to toxic and pathogenic work environments, manual sorting has many risk factors, high labor costs, and low sorting efficiency. The effective segregation of materials is a critical condition for successful recycling, therefore the research community has turned to automated sorting techniques in order to modernize and digitize the recycling process, where sensors are

used to detect recyclable materials in waste in order to develop autonomous systems for their sorting.

Typically, MSW is a fairly complex stream of solid materials that includes metals, polymer-plastics, wood, paper, glass, construction and demolition materials, etc. Moreover, wood waste has a wide variety, complex composition, and a recycling rate dependent on the condition of each individual piece.

Taking all this into account, it motivated us to develop an automated system which is able to identify the different categories of wood waste and segregate wood from plastic. This system can take less time to sort the waste, and it will be more accurate in classifying than the manual way. The new system of segregation recommended in this thesis includes optical sorting and hyperspectral imaging-based sorting. Specifically, the training of the model will be based on data coming from multispectral sensors and an industrial camera. In this way, the identification, characterization and categorization of wood will be made effortless based on its composition. With the system in place, the well-conditioned separated waste can still be recycled and converted to energy or fuel for the growth of the economy. Besides, the circular economy concept is closely linked to the management of wood waste and an approach of this kind involves finding ways to reuse wood instead of just disposing it.

Furthermore, we conducted this study to examine the extent to which high-end equipment influences the performance of the created sorter. Our aim was to assess whether the implementation of top-of-the-line equipment would yield significant improvements in the sorting process. This line of inquiry was crucial because if high-end equipment did not demonstrate a substantial influence on the sorter's performance, it would imply that wood waste management could be made accessible to a wider global audience. This would open opportunities for resource-constrained regions and communities to adopt similar waste management practices. Our aim was to shed light on the potential for cost-effective and scalable solutions in wood waste management, which could positively impact environmental sustainability efforts worldwide.

The system that is developed for the separation of the accumulated waste is based on the Convolutional Neural Network concept. Advancements in artificial technology, particularly in the field of machine learning and computer vision, have greatly improved the accuracy and efficiency of waste classification, assisting projects like this one. By utilizing algorithms and neural networks trained on large datasets of waste images, machines can now recognize and classify different types of waste materials with a high degree of accuracy.

1.2. Waste Management

Waste management encompasses the processes and actions necessary to manage waste throughout its entire life cycle. The collection, transport, treatment and disposal of waste are all included, as well as the supervision and control over the waste management process along with any waste-related legislation, technologies, or economic mechanisms.

Waste occasionally poses a risk to human health. Health issues can be linked with the entire process of waste management. The handling of solid waste can cause health problems directly,

while the consumption of water, soil, and food can cause them indirectly. Human activity, such as the extraction and processing of basic resources, produces waste [2]. Waste management aims to lessen the negative consequences of waste on the environment, human health, the resources of the planet, and aesthetics.

Municipal solid waste, which is produced by industrial, commercial, and household activity takes up a significant portion of waste management. There are differences in waste management tactics between developed and developing countries, urban and rural regions, residential and industrial sectors, and even within the same country [3].

Although effective waste management is crucial for creating sustainable and livable communities, many developing nations still struggle with it. And the reason behind that, according to a report [4], is the cost of efficient waste management which typically accounts for 20% to 50% of municipal budgets.

The earliest incineration plants, or "destructors," as they were then known, were developed in response to the sharp increase in trash that needed to be disposed of. Nevertheless, due to the huge sum of ash they created and the way it blew over the surrounding areas, these were met with great criticism [5].

1.3. Recycling

Recycling is a resource recovery tactic that involves the collecting and reuse of waste materials. This procedure entails dismantling and repurposing materials that would otherwise be discarded as trash. Recycling has many advantages, and since so many new technologies are making less materials not reusable, it is less demanding to live in a cleaner environment. Not only can our living conditions become better, but this can have a positive impact on the economy.

Utilizing specialized containers and collection vehicles, recycling materials can be collected apart from regular rubbish. Prior to the waste being collected, the owner of the waste may be required by certain communities to segregate the materials into different containers (for example, paper, plastics, and metals). While, in some other municipalities recyclables are collected all together in a single bin, and the sorting is done later at a central facility.

Different cities and nations accept different types of materials for recycling and each one of them has their own recycling programs to manage the different recyclable materials. Nonetheless, there are alterations on a product after being recycled, which means that its quality and its resale value usually drops. There are many types of materials that can be recycled. Of course, what interests us most in this thesis is the wood recycling, which is being elaborated further below.

1.4. Wood Recycling

Wood recycling is the process of converting wood waste into usable products. Recycling timber became popular in the early 1990s as environmental conditions (mostly deforestation and climate change) led timber producers and customers to seek a more environmentally friendly wood source. Recycling timber is the most sustainable method of producing wood and is really widespread among Oceanian countries where supplies of old wooden constructions are highly available. Wood can be chopped into wood chips and then be used to heat households or produce electric energy.

Most recycled timber originates from old buildings, bridges and piers, where demolition workers carefully segregate it. Simultaneously, any useful pieces of rock are also saved for reuse. After saving the usable timber, it is then sold by the wreckers to merchants who use metal detectors to scan the wood in order to remove nails and cut it to size. This re-milled timber is ordinarily offered to customers as wood beams or flooring and decking purposes.

Recycling wood has gained popularity as a sustainable product. Many consumers believe that by buying recycled timber, they can help reduce the demand for "green timber" and consequently help the environment. The use of recycled wood as a building material has been crucial in increasing awareness about deforestation among industrial sectors and households and encouraging timber mills to adopt more sustainable procedures.

1.4.1. The Obstacles related to Wood Recycling

Numerous challenges hinder the widespread acceptance of recycled timber as a construction material. Firstly, wall studs may require trimming at the ends to prevent decay and cracking, resulting in shorter pieces of wood that may not meet building codes. Although recycled wood may be cheaper than new wood, it needs a respectable amount of time and labor to separate the useful pieces, to de-nail and refinish them for reuse. Demolition must be done carefully to preserve the maximum amount of timber in a building, which takes more time than tearing it down carelessly. A trustworthy source of well-preserved recycled wood may be difficult to obtain for builders, due to the uneven development of the recycled wood market. There may also be a negative perception associated with using "used" or "cheap" wood, which is viewed as being of lower quality than "new" wood. Additionally, not all salvaged wood pieces may fit into a new building, and it might be less expensive and easier to get new wood, from the viewpoint of a constructor. It is a fact that these hurdles are not impossible to overcome. They are more logistical than structural, but a lot of builders prefer to just buy new uniform-sized wood due to its simplicity and faster application.

1.5. Re-use

1.5.1. Biological reprocessing

Recoverable organic materials, including plant materials, food leftovers, and paper goods, can be recovered through composting and digestion processes to decompose the organic matter. This organic material can be reused as mulch or compost for agricultural or landscaping purposes, and the waste gas generated during the process, such as methane, can be captured and used to produce electricity and heat (CHP/cogeneration), increasing efficiency [31]. The aim of biological processing in waste management is to regulate and speed up the natural process of organic matter decomposition.

1.5.2. Energy recovery

There are several practices, such as combustion, gasification, pyrolysis, anaerobic digestion, and landfill gas recovery, that can turn non-recyclable waste materials into usable heat, power or fuel. These practices belong to energy recovery, are included in the non-hazardous waste management group and are named more specifically as waste-to-energy. Through energy recovery, the generated renewable energy source can help reduce carbon emissions by replacing the demand for energy from fossil fuels and lower methane production from landfills. On a global scale, waste-to-energy makes up 16% of waste management [32].

1.5.3. Pyrolysis

Pyrolysis is commonly employed to transform various types of household and industrial residues into a recovered fuel. By placing different waste inputs, like plant waste, food waste, and tires, into the pyrolysis process, an alternative to fossil fuels can potentially be produced. Pyrolysis is a thermo-chemical decomposition of organic materials by heating them with limited oxygen availability, which results in the production of diverse hydrocarbon gasses [33].

Slow pyrolysis of wood waste produces solid charcoal [34]. Pyrolysis of wood waste and plastics is a promising method for fuel production. The solid residue after pyrolysis contains metals, glass, sand and pyrolysis coke which does not convert to gas. Some kinds of pyrolysis processes tend to release fewer harmful by-products that contain alkali metals, sulfur, and chlorine, compared to incineration; while other types of pyrolysis release gasses that are detrimental to the environment, such as HCl and SO₂ [35].

1.5.4. Resource recovery

Resource recovery is the deliberate diversion of waste, which was initially intended for disposal, for a particular future use. This process involves extracting or recovering materials and resources from recyclables or converting them into a source of energy. Resource recovery, in contrast with waste management, implements life cycle analysis (LCA) to propose alternatives. Several extensive studies have shown that for mixed Municipal Solid Waste (MSW), the preferred approach is administration, source separation, and collection, followed by reusing and recycling

of the non-organic fraction, as well as producing energy and compost/fertilizer from the organic material via anaerobic digestion.

An example of how resource recycling can be beneficial is recycling for profit discarded items, like circuit boards, which contain metals. Similarly, wood chippings from pallets and packaging can be repurposed for horticulture by creating useful goods, like path and walkway covers, or covering arena surfaces.

1.6. Applications of Waste Management

Waste Management methods can be applied in numerous fields, including:

1. **Economic:** The implementation of effective and consistent waste management practices can result in economic benefits such as the optimization of resource use, treatment, and disposal, as well as the creation of markets for recycled materials. This can lead to efficient production and consumption of goods, enabling the recovery of valuable materials and the potential creation of new positions for work and business opportunities.
2. **Social:** Proper waste management practices can have a positive impact on health, leading to more appealing and cleaner civil communities. This can result in social benefits and along with the aforementioned reason potentially help to reduce poverty, particularly in developing countries and cities.
3. **Environmental:** Decreasing or eradicating detrimental effects on the natural environment via reusing and recycling, reduction of greenhouse gas emissions and minimizing resource extraction can lead to a better quality of air and water
4. **Inter-generational Equity:** By adapting useful waste management practices, future generations can benefit from a more robust economy, a more inclusive and equitable society and a healthier environment [36].

1.7. Wood Wastes

The term "wood wastes" refers to a group of wastes that comprises discarded wood products from a variety of industries, including the wood packaging and processing sector, demolition and construction as well as private households and railway construction. The generation of energy and a variety of new potential products, including chemicals, biofuels, and other lignocellulosic materials, may both use this waste as a secondary source of raw materials. For instance, lipophilic and hydrophilic extractives found in wood bark can be used to create high-value goods like cosmetic compounds or pharmaceuticals. In a similar way, bio-oil extracted from wood waste resources works well as an extender and modifier for petroleum-based asphalt binders in the creation of asphalt roads. Other high added value uses for wood waste might include the production of biofuels and wood-based composites, which can help reduce the environmental

impact of the timber industry and meet the increasing demand for wood. The forest-based industry should therefore prioritize measures that reduce, recover, and optimize the use of wood waste generated during harvesting and processing.

Wood waste and byproducts generated from wood-based industrial processes can be utilized to create a broad spectrum of beneficial industrial products. Sawmills typically gather 1 ton of sawdust, shavings, slabs, and edgings for every 1000 board feet of lumber produced. Around 75% of this waste material consists of wood content, and the remaining 25% is bark, which can be transformed into energy or non-energy applications. Wood waste can be advantageous in terms of energy generation through several methods such as combustion and cogeneration. On the other hand, wood waste can also be used for non-energy purposes such as manufacturing composite boards, surfacing products, composting and cement boards [6].

Research has shown numerous ways to add value by converting wood waste into other useful products. For instance, a research that took place in Finland showed plenty of rising markets for wood-based products, mainly textiles, chemicals, biofuels and plastic substitutes. Moreover, a study in Zimbabwe indicated that the majority of offcuts and chips from wood-based businesses are employed as firewood by the locals and in commercial sawmills to create steam for kiln driers. Meanwhile, researchers in Japan revealed that furniture firms generated 15 million cubic meters of wood waste, of which over 90% was recycled to make wood-based panels and fuel.

1.8. Sectors of Wood Wastes

Discarded wood products are collected from various sources such as wood packaging, demolition and construction, wood processing industry, and others like private households and railway construction. Although current wood waste management strategies are mainly focused on landfill, energy and material recovery [37], these are not the only sectors related to wood waste. One of the main sectors involved is forestry since wood waste is an inevitable byproduct of forestry operations. When trees are harvested for timber, there are often leftover branches, treetops, and other types of wood waste generated. Another industry related to wood waste is the wood processing and manufacturing industry. This sector involves the conversion of wood into various products such as furniture, flooring, doors, windows, and other building materials. During this conversion, leftover wood pieces and scraps are generated. Moreover, another important use of wood waste in wood processing and manufacturing is as a raw material for new wood products. For example, sawdust and wood shavings can be used to make particleboard, medium-density fiberboard (MDF), and other composite wood products.

In bioenergy, wood waste is a common source of this type of energy, such as biomass, which can be burned to generate heat and electricity. One main form of bioenergy that is produced from wood waste is wood pellets. They are a convenient and efficient fuel source for heating homes and buildings, as well as for producing electrical energy. Additionally, wood waste can be used in agriculture as mulch, compost or soil amendment, which can help improve soil

quality and fertility, retain soil moisture, suppress weeds and reduce the need for synthetic fertilizers and herbicides.

Moreover, in construction and demolition industry, wood waste collected from these sites can be resold for furniture production, reused or recycled for other purposes, such as new construction projects and energy generation. The disposal of wood waste is an important aspect of waste management and common disposal methods include, preferably repurposing and composting, or landfilling as a last resort. As mentioned before, managing wood waste through proper disposal practices is vital. Landfills are a common destination for wood waste, but they also stand as an inefficient method of disposal since they require valuable space and decomposing wood waste can produce methane, a potent greenhouse gas. Finally, wood waste is of interest to governmental policy makers and sustainability advocates because it represents an opportunity to reduce waste, conserve resources, promote more sustainable practices and help build a brighter future. There are numerous policies and initiatives at the local and international levels that aim to encourage the responsible management of wood waste and promote circular economy principles.

1.9. Circular Economy

Circular economy is an alternative to a traditional linear economy (make, use, dispose) and therefore a new production and consumption model that promotes sustainable growth over the long term. By adopting this approach, we can optimize resource utilization, minimize the consumption of raw materials, and repurpose waste through recycling or other means to create new products. Although this term 'Circular Economy' has recently gained much popularity, the concept has been in existence since the 1970s. It was initially suggested as a means of addressing resource security and scarcity concerns at local and national levels, as well as a method of reducing the environmental impact of our production and consumption practices.

The objective of the circular economy is to maximize the usefulness of the available resources by following three fundamental principles: reducing, reusing, and recycling. This methodology helps to prolong the life cycle of products, minimize waste, and establish a more efficient and ecological method of production over time. The circular economy takes inspiration from nature, where every good is valuable and waste is converted into a new resource, thereby achieving a balance between progress and sustainability.

The accumulation of non-recyclable waste in landfills all over the world poses a significant threat to our way of life in the near future, if our generation does not act rapidly to prevent it. This waste can increase the spread of diseases through vectors such as insects like flies and mosquitoes. Additionally, waste accumulation can destroy the natural beauty of habitats, lead to deforestation and soil and water pollution due to the presence of toxic chemicals in improperly treated materials. This pollution can disrupt the food chain, leading to more diseases and health problems for humans and ecosystems globally.

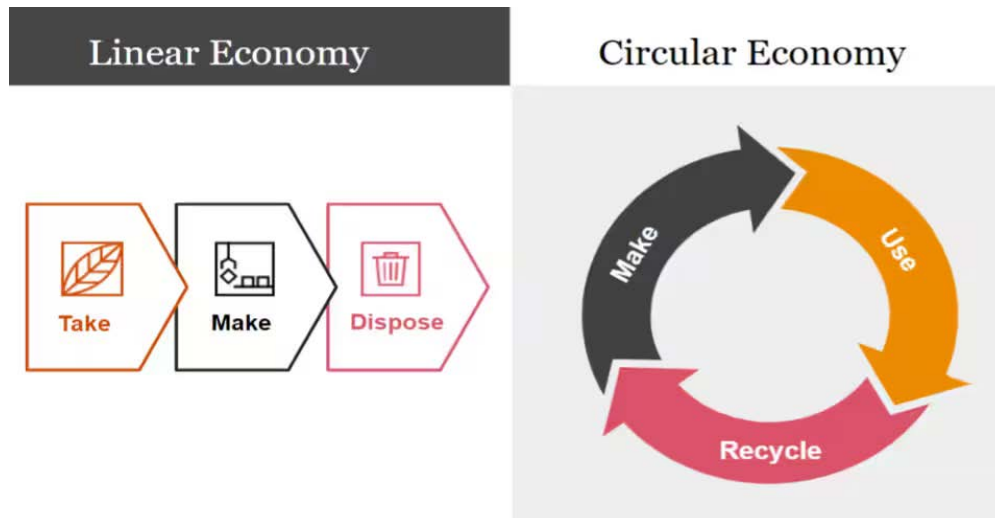


Figure 2. Graphic representation of Linear in contrast with Circular economy.

There are three basic reasons why waste accumulation has become a more serious threat over the last 50 years. As the first reason stands the lack of availability on the market of recyclable items, despite companies' efforts on developing more sustainable and environmentally friendly products. The second reason is overpopulation, which shows the logistic challenge when managing the generation of waste that results in a higher percentage of goods ending up in landfills or even the ocean, affecting the lives of marine species. Finally, the third reason is the lack of societal involvement in addressing issues such as climate change and waste accumulation.

To illustrate the issue of waste accumulation through data, it is important to note that the global population generates between 7 and 9 billion tons of waste annually, of which 70% is not treated properly and ends up in landfills, posing a risk of polluting natural environments and creating new health hazards such as ocean microplastics. This data encompasses all used, unwanted, and discarded items created by humans. However, there is a difference between the total waste produced and Municipal Solid Waste (MSW), which only consists of the garbage produced in urban areas or their surrounding regions. About 2 billion tons of MSW are generated each year, with approximately 33% of it not being appropriately treated. This implies that every person produces between 0.1 and 4.5 kilograms of waste per day, with an average of 0.7 kg. Furthermore, it is expected that MSW will increase to 3.4 billion tons by 2050 due to the rapidly increasing world population and the need to use natural resources extensively for industrial development and sustaining our civilization.

1.10. Wood Waste and Circular Economy

Wood is a material that is natural, renewable, and biodegradable, and it has exceptional mechanical and thermal properties. The environmental effects during the production and end-of-life stages of wood materials are generally less harmful than those of equivalent inorganic or

fossil-based materials. The circular economy concept is closely linked to the management and use of wood waste and an approach of this kind would involve finding ways to reuse or recycle wood instead of just disposing it [7].

This could include repurposing it as raw material for new wood-based goods or using it as a fuel source, such as in biomass energy generation. By reusing or recycling wood waste, we can decrease the amount of waste that ends up in landfills and minimize the need for new raw materials, which helps to conserve natural resources and reduce the environmental impacts of resource extraction and processing. Additionally, reusing and recycling wood waste can also create economic opportunities, such as opening new work positions in the recycling and repurposing industry. Taking everything into account, the circular economy approach is a significant way to manage wood waste in a more sustainable and valuable manner.

1.11. Waste Frequency of Appearance

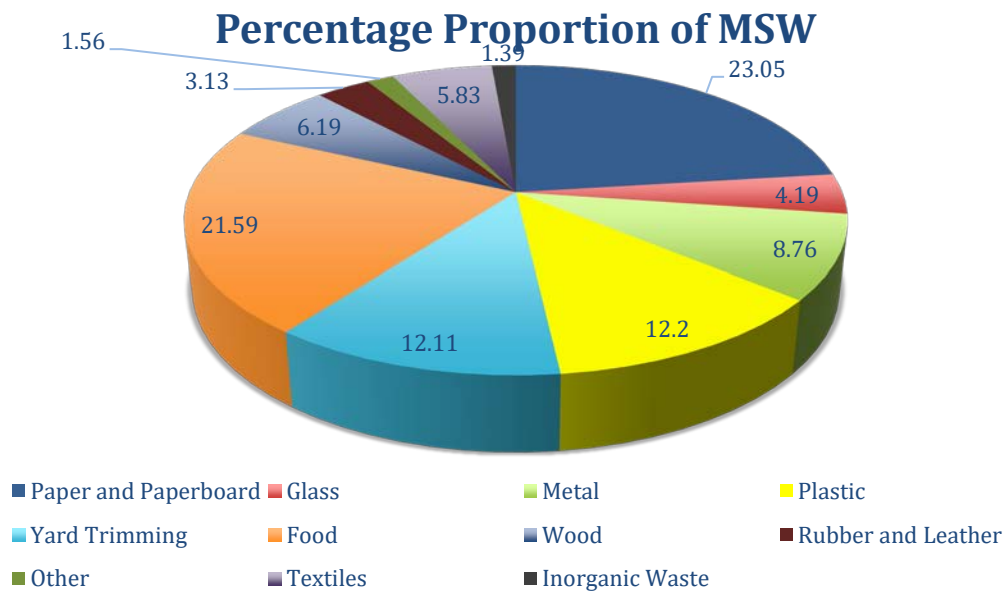


Figure 3. Percentage proportion of Municipal Solid Waste (MSW) types.

In the image attached, the percentage proportion of every major category in Municipal Solid Wastes (MSW) is shown. It can be observed the significance of Wood Waste among the eleven other Wastes, since it constitutes the 6th largest source of residues [38]. Although there are many studies and data analyses related to MSW and General Wastes, there has not been conducted any percentage research on the different grades or categories of Wood Wastes, which is another indicator of how unexplored the territory we are investigating is.

1.12. Wood Waste Categories

According to the official Wood Recyclers Association (WRA) [8], there are four different variations of wood waste:

Grade	Typical Markets	Typical sources of raw material for recycling
Grade A – Pre-Consumer Waste Wood (*) and untreated wooden packaging = Clean un-treated	A feedstock for the manufacture of professional and consumer products such as animal bedding, equine and landscaping surfacing. May also be used as a fuel in domestic biomass installations and for the manufacture of pellets and briquettes.	Wood product manufacturing, distribution, retailing, packaging and secondary manufacture, e.g. joinery and pallet reclamation.
Grade B – Industrial waste wood = Treated and Non-hazardous	This is the preferred feedstock for industrial wood processing operations such as the manufacture of panel board products. It can also be used for biomass.	As Grade A, plus construction and demolition operations, skip operators, transfer stations.
Grade C – Municipal waste wood = Treated and Non-hazardous	For use in the biomass installations and for panel board in controlled volumes.	All above plus municipal collections and transfer stations.
Grade D – Hazardous waste wood	Requires disposal at facilities licensed to accept hazardous waste.	All of the above plus Agricultural fencing, trackwork and transmission pole contractors.

Table 1. Wood Waste Categories: Markets and sources.

Clean/untreated waste wood is suitable for processing into animal bedding, panel board feedstock, landscaping or equestrian surfaces and biomass. Treated, but non-hazardous waste wood is suitable for processing as a feedstock for panel board or energy recovery in a compliant facility. Hazardous waste wood can only be disposed of in a facility licensed for this purpose.

*Pre-consumer waste wood is waste wood material created during the manufacturing process of virgin wood, not involving the application of treatments, e.g., offcuts or trimmings from virgin/sawn timber. It is also a waste wood material created during the manufacturing process of raw, untreated board products such as panel board, MDF and plywood (for clarity, this waste wood can only be used/burnt at source). Waste from joinery activity using these untreated wood materials is also included in this definition.

For the purposes of this thesis and its experiments we will narrow down these grades to three, which is another common distinguishing of wood waste, since Grade B and C share many similarities:

- **Grade A** will count as one category (Category number 1), namely Raw Clean untreated wood.
- **Grade B and C** will count as one category (Category number 2), namely Treated and Non-hazardous wood and
- **Grade D** will count as one category (Category number 3), Hazardous wood waste.

Grade	Typical Materials	Typical non-wood content prior to processing	Notes
A	Solid softwood and hardwood. Packaging waste, scrap pallets, packing cases and cable drums. Process off-cuts from the manufacture of virgin/sawn timber and untreated board products.	Nails and metal fixings. Minor amounts of non-hazardous surface coatings, such as water-soluble paint.	Is a waste for the requirements of Waste Management Regulations. Should not contain any treated or low-grade material.
B	May contain Grade A material as above plus building and demolition materials and domestic furniture made from solid wood.	Nails and metal fixings. Some paints, plastics, glass, grit, coatings, binders and glues. Limits on treated or coated materials as defined by end users.	This is mostly solid wood. Some feedstock specifications contain a 5% to 10% limit on former panel products such as chipboard, MDF and plywood. Is a waste for the requirements of Waste Management Regulations. Will require a compliant installation for biomass.
C	All of the above plus flat pack furniture made from board products and DIY materials.	Nails and metal fixings. Paints, coatings and glues, paper, plastics and rubber, glass, grit. Coated and treated timber (non CCA or creosote).	This is mostly board products. Mainly suitable for compliant biomass installations, but also suitable for panel board manufacture with correct processing and blending. Is a waste for Waste Management Regulations.
D	Agricultural fencing, transmission poles, railway sleepers, cooling towers.	Copper chrome arsenic (CCA) preservation treatments and creosote.	Is a waste for Waste Management Regulations. Requires disposal in a process regulated to take hazardous waste.

Table 2. Wood Waste Categories: Materials, content before processing and notes.

In addition, our experiments were conducted on three different classes of wood waste, since these three concentrate most of the popularity in the Greek wood market and they also have research interest due to their figure:

- Class 1, Medium-Density Fiberboard (MDF):** Medium-density fiberboard is an engineered wood product that is created by converting hardwood or softwood waste into wood fibers, often in a defibrator, mixed with wax and a resin binder, and subjected to panels with high temperature and pressure. Compared to plywood, MDF typically has a higher density. Although it is composed of separated fibers, it can be utilized in construction applications similar to plywood. In fact, MDF is stronger and more durable than particle board [9].

- **Class 2, Melamine Faced Chipboard (MFC):** Melamine faced chipboard, also known as Conti-board chipboard is a type of composite wood that is made by blending small pieces of machined wood with a powerful adhesive (resin). This type of chipboard is highly versatile and can be used in a variety of applications, including in bathroom and kitchen cabinetry, office furniture, canal boats, caravans, and other settings. It is pretty effortless to work with and provides an attractive surface finish. With modern production methods, melamine-faced chipboards are becoming more and more difficult to distinguish from natural wood, as technology allows for sheets that mimic the look of real wood.
- **Class 3, Oak Veneer:** Oak veneer is a thin slice of oak wood that is used to cover other materials, such as plywood or particleboard, to create the appearance of solid oak. Veneers are typically produced by slicing a log into thin sheets, which are then glued onto a substrate material using an adhesive. Oak veneer is popular in furniture making and interior design, as it offers the beauty of oak wood, since the veneer provides a consistent and uniform appearance, at a lower cost than solid oak. Oak veneer is available in a range of grades and thicknesses and can be stained or finished to match a variety of styles and design preferences. However, it is not as durable or long-lasting as solid oak, as the veneer layer is relatively thin and can be prone to damage or wear over time.

If we were to enlist the wood pieces of our three classes to the three categories mentioned earlier, based on their production method, we would have Oak Veneer belonging to Category 1, while MDF and MFC belonging to Category 2.

Chapter 2

Theoretical Background

2.1. Hyperspectral Images

As it is commonly accepted, the reason why we observe various colors such as red, green, and blue is because when light reflects off of objects, it falls into specific ranges of wavelengths within the visible spectrum of electromagnetic radiation. Specifically, long wavelengths around 564-580 nm appear red, medium-wavelengths around 534-545 nm appear green, and short-wavelengths near 420-440 nm appear blue. However, it's important to note that there are many other wavelengths beyond the visible spectrum that our eyes are unable to perceive and are therefore easily missed or invisible to us [39].

The colors or characteristics that are missed by humans may be visible to other animals since color perception is subjective and varies from species to species. As the saying goes, '*Color is in the eye of the beholder*' [40].

With the evolution in imaging technology, we are now able to gather information beyond the visible spectrum and interpret it. Spectral imaging refers to the process of capturing spatial and spectral information simultaneously in an image space and combining them. This technique is similar to an RGB-color image but contains numerous additional channels, making it difficult to visualize as a whole. Nonetheless, we possess all the necessary data for our computers to comprehend. To explain further, let's assume that we have an n band hyper-spectral image. This means that we have n grayscale images, with each band recording diverse light intensity data according to its wavelength values. These grayscale images are stacked on top of one another over an adjacent range of wavelengths, resulting in an image of dimensions n rows * m columns * n bands [41].

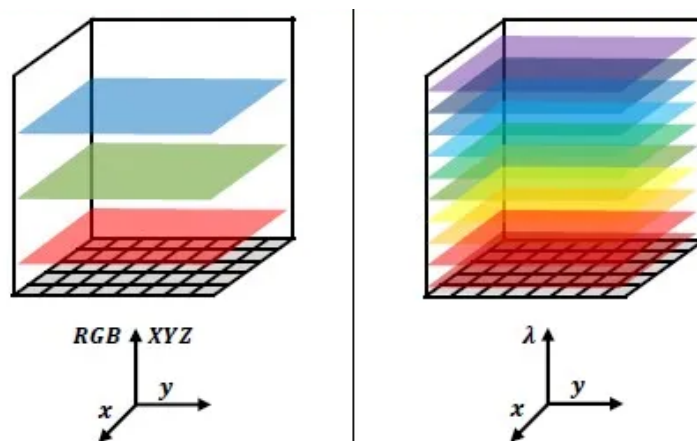


Figure 4. RGB vs HSI images.

Multi-spectral images, which are n band images, differ from hyperspectral images as they are not necessarily spread over an adjacent range of wavelengths and have fewer bands than an average hyperspectral image (HSI).

Images captured by HSI cameras are presented in the form of a hypercube, containing n -dimensional image data, which can be difficult to work with. As a result, it's common to apply dimensional-reduction techniques and various other approaches to process the hypercube. For instance, if we focus on a single hyper-pixel, we can plot a continuous spectrum for that specific pixel.

The applications of HSI are widespread, spanning across industries such as medical, agricultural and automotive. HSI enables us to explore uncharted territory by visualizing information that is invisible to the human eye and finding ways to comprehend it.

2.2. Multispectral Imaging

Multispectral imaging is an advanced technology that combines spectroscopy with imaging to gather exhaustive information about the composition and features of materials and surfaces. This technology is more powerful than traditional imaging systems and provides non-invasive and non-destructive means of recognizing and quantifying objects. Due to its versatility and usefulness, multispectral imaging has become increasingly popular in a variety of industries and research projects.

Multispectral imaging works by collecting and analyzing information across the electromagnetic spectrum to form the spectrum - comparable to fingerprints - for each pixel in an image, allowing for the identification of objects and materials by their unique spectral signatures, similar to how fingerprints are used to identify a person. Multispectral imaging has many practical applications which consist of improving food quality and safety, facilitating waste sorting and recycling, and enhancing control and monitoring in pharmaceutical production.

The electromagnetic spectrum encompasses all types of light, ranging from radio waves to gamma rays, most of which cannot be detected by the human eye (Figure 7).

2.2.1. Spectral Imaging

Spectral imaging uses various bands across the electromagnetic spectrum for imaging. Unlike the RGB camera that makes use of only the three visible light bands (red, green, and blue), hyperspectral imaging can explore how objects interact with a much wider range of bands, spanning from 250 nm to 15,000 nm and including thermal infrared. The research of light-matter interaction is known as spectroscopy or spectral sensing.

These systems are capable of capturing multiple channels or bands of information across the electromagnetic spectrum. Thus, this information can be processed to create a color-coded representation of the spectral data, providing insights into the chemical and physical properties of objects in the image.

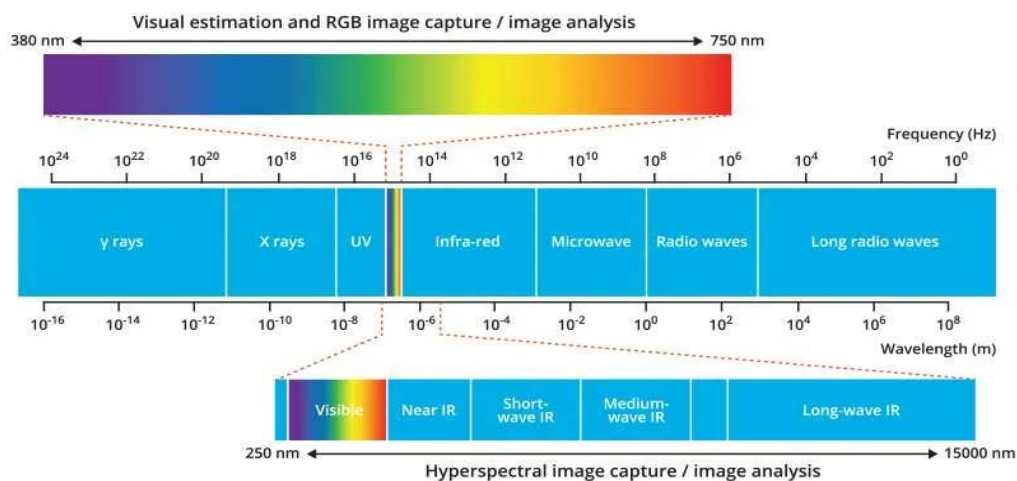


Figure 5. Hyperspectral imaging captures wavelengths from 250 nm to 15,000 nm and thermal infrared.

2.2.2. How does multispectral imaging work?

Multispectral imaging consists of using an imaging spectrometer, also known as a multispectral camera, to gather spectral information. This camera captures the light of a scene separated into individual wavelengths or spectral bands, generating a two-dimensional image that also records the spectral information of each pixel. The result is a multispectral image, with each pixel representing a unique spectrum.

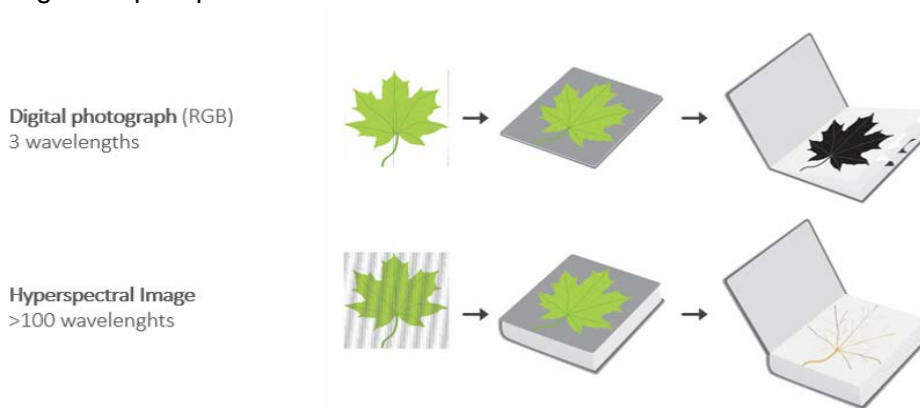


Figure 6. In order to replicate human vision, a digital photograph of a leaf (top) is created using three bands: red, green, and blue. The RGB data is comparable to a three-page brochure. In contrast, a hyperspectral image of the same leaf (bottom) captures a spectral response from 220 wavelengths, resulting in a more extensive and detailed representation of the object, akin to a book with 220 pages.

Multispectral imaging provides both spatial and spectral information, allowing for the identification and classification of materials, as well as the spatial distribution and areal separation of these materials.

2.2.3. The Advantages of Multispectral Imaging

Multispectral imaging's main benefit lies in its ability to provide precise material characterization, thanks to its high spatial and spectral resolution. By capturing tens or even hundreds of spectra, a multispectral camera can construct a massive multispectral data cube that contains position, wavelength, and time-related data. When compared to conventional imaging, multispectral imaging provides a more significant amount of information, enabling more precise analysis, identification, and differentiation of materials and substances.

With multispectral imaging, it is possible to distinguish materials that share similar physical or visual properties or those that are invisible to the human eye, such as various minerals.

2.2.4. Fields of Multispectral Imaging

Multispectral imaging is a technique that is becoming more prevalent in industry, research, and remote sensing. This method provides data that can be utilized for locating, categorizing, or measuring the concentration of various substances that are not detectable with regular cameras or the naked eye. For example, an in-line quality control system that includes a multispectral imaging system can detect foreign objects, contaminants, as well as the levels of fat, sugar, or moisture in products. Furthermore, multispectral or hyperspectral imaging obtained via remote sensing provides valuable data about Earth's surfaces, including minerals and vegetation.

Multispectral imaging has a wide range of applications, such as:

1. **Environmental monitoring:** Observing changes in land use, vegetation health, and water quality over time to identify early signs of ecological degradation and monitor conservation efforts.
2. **Mineral exploration:** Mapping mineral deposits, detecting the mineral composition and quality.
3. **Quality control:** Evaluating food products, detecting contaminants and defects in industrial products non-destructively.
4. **Waste management:** Separating different materials accurately and efficiently to optimize recycling processes and enhance the value of recycled materials.
5. **Agriculture:** Assessing the health and yield of crops, monitoring the soil moisture and nutrient content to enhance crop management practices and boost crop yields. Typically, agriculture monitoring employs hyperspectral cameras mounted on drones.
6. **Military surveillance:** Detecting hazardous materials.

These are just a few instances of the useful information that this kind of imaging can provide for various purposes. To conclude, multispectral imaging is a valuable machine vision technology for multiple industries and research to enhance their processes, improve quality, and reduce waste.

Chapter 3

Related work

3.1. General Waste Sorting Methods hitherto

There are several well-known methods [42] used for general waste classification. One method is hand sorting, which involves physically sorting through waste materials by hand to identify and separate different types of waste. Another method is magnetic separation: which utilizes magnets to separate ferrous materials, such as iron and steel, from non-ferrous materials. Eddy current separation is another technique that employs eddy currents to separate non-ferrous metals, such as aluminum and copper, from other materials. Density and flotation separation involves separating materials based on their density or their ability to float on a liquid surface, such as through the use of air classifiers or water-based/oil-based separation methods.

Centrifugation is a subset of density separation and involves the use of centrifugal force (high-speed rotation) to separate waste materials based on their physical properties, like size, shape, density, and viscosity. The denser waste particles move towards the bottom of the container, while the lighter waste particles move towards the top, allowing for effective separation of different types of waste. Chemical analysis involves analyzing waste materials to determine their chemical composition and properties. Thermal analysis subjects waste materials to high temperatures to determine their thermal properties, such as their melting and combustion points. X-ray fluorescence employs X-rays to identify and analyze the elemental composition of waste materials.

Hyperspectral imaging is a sophisticated method that uses state-of-the-art and complex imaging technology to capture and analyze the spectral properties of waste materials to identify their composition and properties. Optical sorting, also known as Digital sorting, uses sensors, RGB cameras and lasers to identify and sort waste materials based on their color, shape, size and other visual properties. Machine or deep learning algorithms can be combined with optical sorting and hyperspectral imaging to achieve better results in waste segregation. The cameras, either hyperspectral or RGB, can provide a wealth of data about the chemical and physical properties of waste materials which the algorithms can utilize in order to identify patterns and features that are not easily discernible by the human eye, and classify the waste materials into different categories.

There are also other less popular methods for general waste classification which can be more appropriate depending on the specific type of waste and the desired level of accuracy and efficiency.

3.2. Wood Waste Sorting Methods hitherto

There are traditional and non-automated ways of wood waste classification which are destructive or not efficient due to time or cost, like hand picking, centrifugation, flotation methods and chemical analysis, which have been already mentioned. Moreover, chemical analysis needs a particular mention since it involves Pyrolysis – Gas Chromatography – Mass Spectrometry in which the material is heated to decomposition in order to produce smaller molecules that are then separated by gas chromatography and detected using mass spectrometry.

Nonetheless, there are also some modern ways which include X-ray fluorescence, hyperspectral or digital imaging along with AI techniques, such as deep learning algorithms [43], [44]. Thus, the sorting system is trained to distinguish different types of materials, such as processed and unprocessed wood. This separation of wood by category is crucial in contemporary particleboard manufacturing and wood recycling plants since only the purest non-processed wood fractions are suitable for producing high-quality particleboards made of recycled content. In addition to that, in order to analyze the state of wood wastes, it is vital to employ microscopic methods after the segregation has been settled.

The use of microscopy in the monitoring and analysis of the mechanisms involved in a wood waste treatment process, in combination with other analytical methods, makes it possible to visualize and better understand the various transformations of the material as well as the physiological mechanisms developed by the microorganisms.

Until now, the vast majority of research was concentrated in wood separation from other materials and the process followed is analyzed further in more detail. This thesis is focused on segregation of different types of wood and segregation of wood waste from plastic.

Once collected, recoverable wood waste must be sorted to separate it from other types of waste such as stones, plastics or metals. Density and flotation sorting methods separate materials according to their density, and metals are separated by magnetic methods. However, these conventional methods have limitations and cannot effectively discard MDF and heavily contaminated wood. The wood treated with preservatives products containing heavy metals can be efficiently identified and sorted out, due to X-ray fluorescence spectroscopy (XRF). The clean wood chips can then be further sorted according to their material type based on the operator's requirements.

On the other hand, the most advanced technology for automated sorting utilizes a spectroscopic technique that employs a near infrared (NIR) sensor along with blowing nozzles to separate elements according to their composition. Wastes are exposed to radiation within a specific wavelength range, typically between 0.7 and 2.5 μm . Depending on their chemical composition, materials absorb some wavelengths, while others are captured and transmitted to a spectrometer. These residual wavelengths are then analyzed by a computer, both spectrally and spatially. To reduce the raw dataset to a limited set of significant and independent variables, a principal component analysis (PCA) is performed, which involves calculating the second derivatives of the spectra. The result is represented as a scatter plot on a graph, with one component being a combination of all the loadings of the different variables. The principal components loadings are then displayed on a plot to identify the relationships between the variables, allowing for the meaning and contribution of each variable to be determined [45].

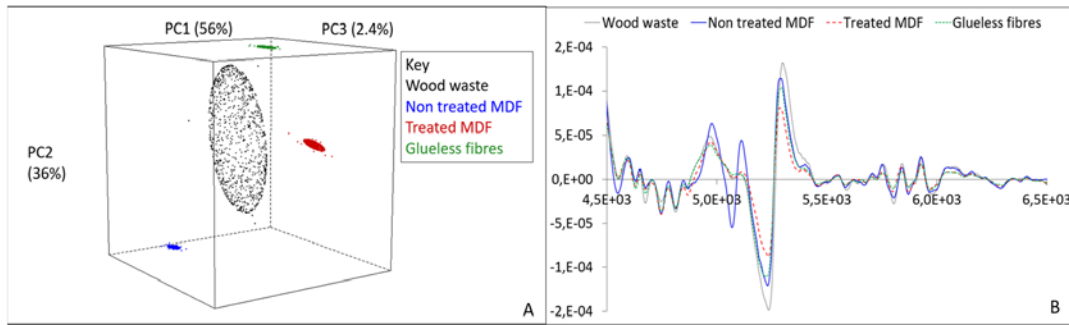


Figure 7. Principal component analysis performed on four wood products-plastic composites (A) and loadings analysis (B).

Near-infrared spectroscopy can be an important factor in enhancing the elimination of pollutants in recycled wood by characterizing the main components of wood. The signals of each characteristic functional group of wood components have been identified (in [46]), thus enabling the detection of wood compounds in the near infrared. A principal component analysis was conducted on four wood products with varying compositions, as shown in Figure 7A.

The principal component analysis yields four well-defined scatterplots, each representing a wood product. The analysis separates three different materials: the wood waste, the MDF, and the glueless fibres. Also, it separates the treated MDF from the untreated MDF. The wood waste scatterplot is the largest, thanks to its variability. The average NIR spectra of the wood products are shown in Figure 7B to justify the observed separations.

Each spectrum exhibits a distinct profile and the amplitude of a peak indicates the difference between the samples. Each peak corresponds to a range of wavelengths attributed to a chemical bond, and the differences between the peaks explain the differences in composition between the composites. As a result, NIR analysis can distinguish between different wood products based on their unique compositions.

3.3. Microscopic Methods for Wood Waste Characterization

Recent advancements in imaging and control of fermentation conditions offer new opportunities to better observe and characterize the changes in the wood material involved in recycling processes. It is really crucial to be able to monitor alterations in the material in the recycling processes. In addition to conventional transmission photon microscopy, more sophisticated techniques such as confocal laser scanning microscopy (CLSM) or scanning electron microscopy with microanalysis (EDS/WDS-SEM) allow the concurrent investigation of wood waste during processing. These methods have been widely used in previous studies and here some examples of laboratory-scale results will be provided. For instance, the study of copper detoxification by a filamentous fungus requires the sampling of wood particles, which can be analyzed using different imaging techniques, as shown in Figure 8.

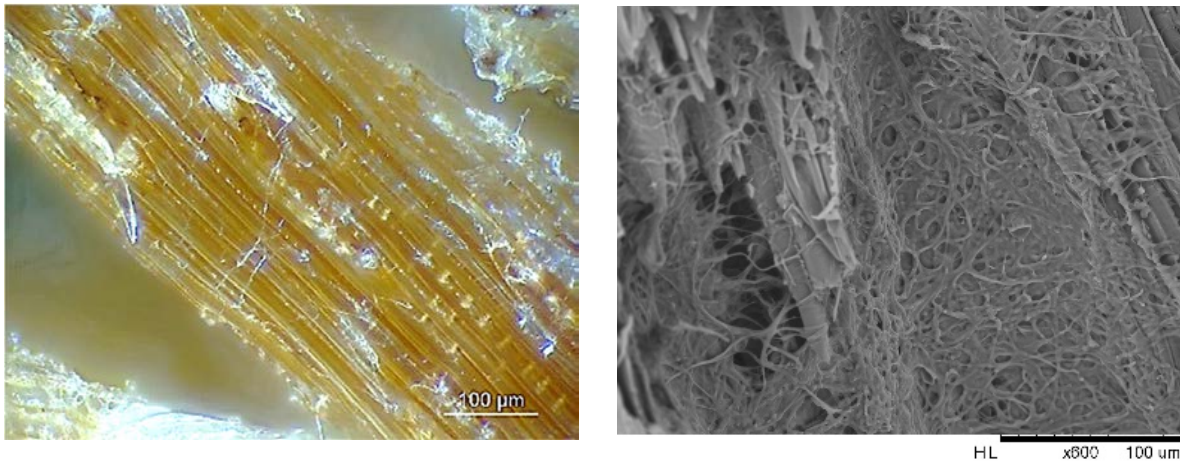


Figure 8. Microscopic monitoring of filamentous fungus growth during a bioremediation process: episcopy microscopy (A) and scanning electron microscopy (B).

Scanning electron microscopy (SEM) coupled with energy-dispersive X-ray spectroscopy (EDS) can be useful in bio-remediation processes involving wood contaminated with heavy metals like copper, as it can provide information about how the fungus mediates the relocation of the contaminants. Industrial wood that has been treated with a commercially formulated preservative product is ground into 2-5 mm particles. These particles, containing 2700 ppm of copper according to XRF quantification, are then incubated with a consortium of fungus and bacteria. In the case of organic contaminants such as urea formaldehyde resin, CLSM can be employed to observe the constituents of interest if they are fluorescent or can be stained with fluorescent dyes.

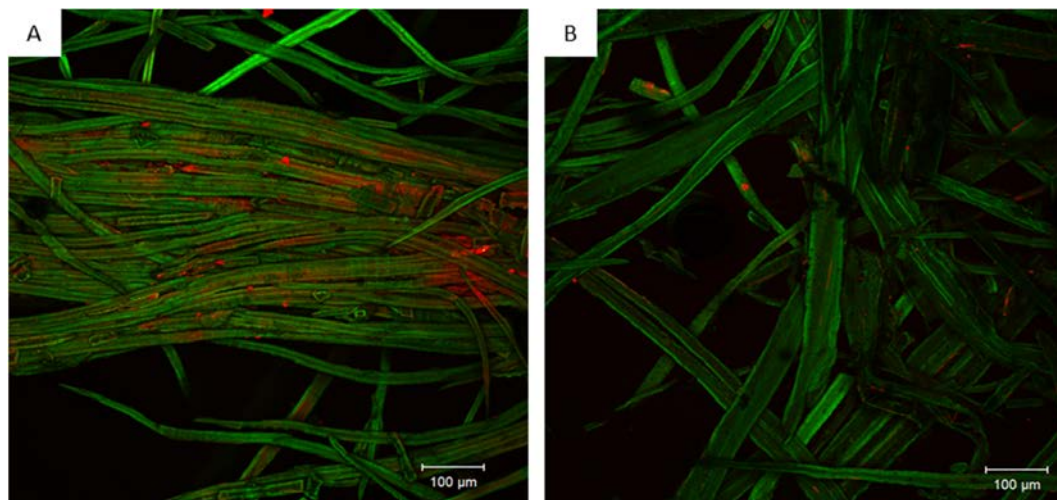


Figure 9. CLSM image showing the location of the resin (red) on wood fiber (green) before (A) after (B) decontamination treatment.

Urea-formaldehyde (UF) resin removal from wood fibers in MDF panels waste after steam explosion can be visualized by CLSM using spectral deconvolution. This type of imaging is commonly used in the field of botanical sciences [47],[48]. Once spectral calibration is completed,

hyperspectral imaging becomes possible, allowing for the evaluation of the UF removal efficiency of a process (as shown in Figure 9). Microscopic analysis of the sample can provide additional information that may not be easily obtained through traditional analytical chemistry techniques.

Microscopic inspection of the sample provides complementary information that might be hardly available with the use of analytical chemistry.

3.4. Complete Autonomous Waste Sorter

A complete autonomous sorter is yet to be found for distinguishing different types of wood waste. Nonetheless, there have been developed sorters for other wastes, such as Construction & Demolition waste, Plastic and Metal Waste [23], in which we will base our work. The Figures show the setup of these sorters and their common primary elements. Firstly, they all make use of a conveyor belt in accordance with an actual recycling production line. Secondly, there exist industrial, 3D and hyperspectral (visible and near-infrared light) cameras and a light source. Lastly, there is a robotic arm/manipulator/gripper to place the samples in the right container.

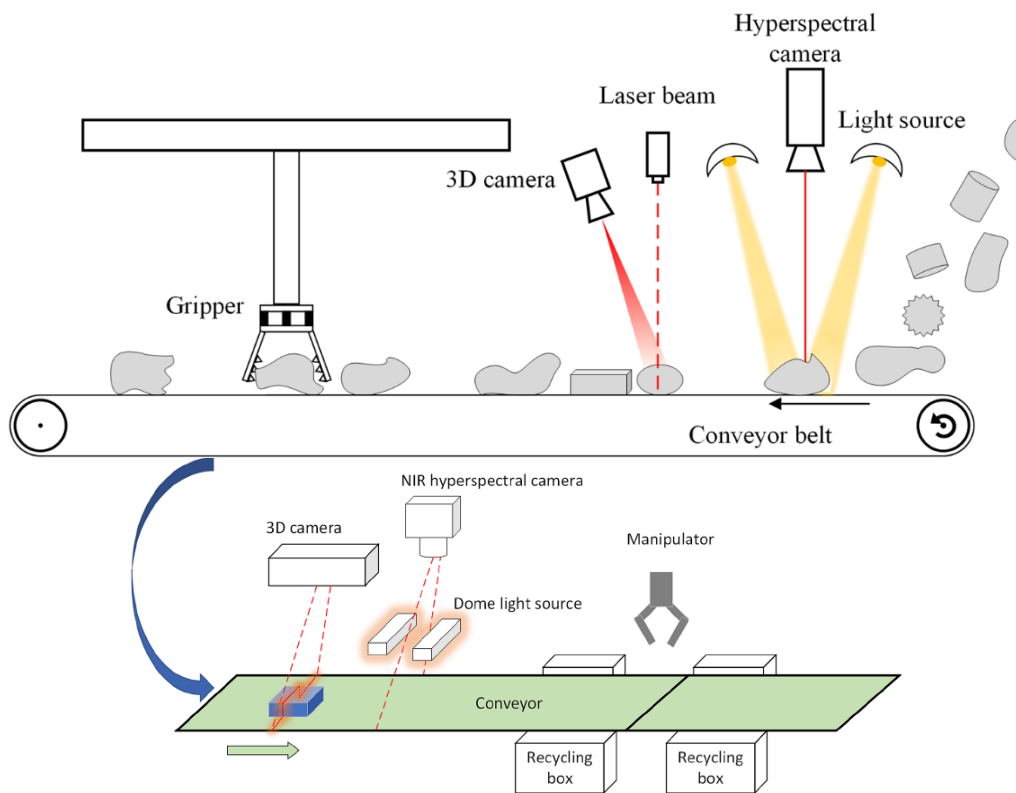


Figure 10. Setups of Autonomous General Waste Sorters.

The conveyor belt transfers the object at the front and the hyperspectral camera captures the curve of the object's reflectivity in the near-infrared spectrum. This curve is used as input to the classification model which outputs the material of the object, since objects made of different materials have different reflectance curves. The object is illuminated by an invisible infrared laser of a specific wavelength, and the industrial or 3D camera is tilted to capture the location or a stereoscopic image of the object. On the processing scope, the algorithms applied after the hyperspectral imaging are mostly deep learning combined with convolutional neural networks in order to enable computer vision.

Nevertheless, a complete autonomous sorter, even for other types of waste, is still far from real-condition usage. There are four main challenges behind that:

1. **Scale-up:** The performance of a sorter in the laboratory may not necessarily translate to a larger scale. The throughput, time-efficiency and accuracy of the prototype may be affected by changes in the size of the equipment and the amount of material being sorted.
2. **Variability:** The materials being sorted in the laboratory may not be representative of the variability that exists in the real world. The composition and characteristics of the material can vary widely, which can affect the accuracy of the sorter.
3. **Environmental conditions:** The conditions in a laboratory are typically controlled and stable, whereas real-world conditions can be unpredictable and harsh. Factors such as temperature, humidity, and dust can affect the performance of the sorter. The random lighting source existing in the outside world can play a significant role in the pictures capturing and dramatically affecting the results, as well as the background which is no longer the conveyor belt but soil or something new.
4. **Integration:** Sorting technologies are often just one component in a larger system or process and may need to be integrated with other equipment in a real-world setting, which can be complex and require specialized knowledge and expertise to ensure that all parts can work together effectively.

Chapter 4

Artificial Intelligence

4.1. Artificial Intelligence

AI research originated from a workshop that took place at Dartmouth College in 1956 [10], where the attendees eventually became the pioneers and leaders of the field. They, along with their students, developed programs that were deemed "astonishing" by the media, as computers were able to learn strategies for checkers, solve word problems in algebra, prove logical theorems, and even speak English. This led to the establishment of research labs worldwide by the mid-1960s. [11]

Advances in machine learning and perception were made possible by faster computers, improvements in algorithms, and access to vast amounts of data. Deep learning methods, which rely heavily on data, began to meet accuracy benchmarks around 2012. In 2015 [12], artificial intelligence had a breakthrough year, with the number of software projects that employed AI at Google skyrocketing from a "sporadic usage" in 2012 to over 2,700 projects. This surge was credited to the increased availability of affordable neural networks, thanks to the proliferation of cloud computing infrastructure and the expansion of research tools and datasets [13].

Artificial intelligence (AI) refers to the replication of human intelligence processes by machines, particularly computer systems. AI is implemented through various applications such as expert systems, speech recognition, natural language processing, and machine vision. [14]

As the excitement around AI has intensified, vendors have been eager to promote how their products and services use it. However, what they label as AI is often just a component of technology, like machine learning. AI requires specialized hardware and software to develop and train machine learning algorithms. While there is no specific programming language that solely represents AI, Python, R, Java, C++, and Julia are popular choices among AI developers.

Mainly, AI systems operate by consuming significant amounts of labeled training data, analyzing the data for patterns and correlations, and utilizing these patterns to make predictions about future events. For example, a chatbot that is trained with various text samples can learn to generate lifelike conversations with people, or an image recognition tool can identify and describe objects in images by reviewing millions of examples. New generative AI techniques are advancing quickly, allowing realistic text, images, music, and other media to be created.

The programming of AI concentrates on cognitive skills that involve the following:

- **Learning:** Deals with obtaining data and establishing rules or algorithms to process it into practical information for computing devices to follow.
- **Reasoning:** Focuses on selecting the appropriate algorithm to reach a specific outcome.
- **Self-correction:** Aims to continuously refine algorithms to deliver the most precise results.

- **Creativity:** Uses different AI techniques, including neural networks, rules-based systems, and statistical methods, to produce novel images, text, music, and ideas.

Artificial intelligence has the potential to bring about significant changes to our lifestyle, work, and recreational activities. In business, AI has been successful in automating various tasks that were traditionally done by humans, including customer service work, fraud detection, and quality control. AI can perform certain tasks more efficiently than humans, especially repetitive and detail-oriented ones such as analyzing legal documents. With the ability to process massive data sets, AI can provide valuable insights into business operations. The rapid expansion of generative AI tools is also expected to have a significant impact on fields such as education, marketing, and product design.

The application of AI techniques has not only resulted in increased efficiency but has also created new business opportunities, such as ridesharing companies' use of AI to connect riders with taxis [15].

Many successful companies, including Alphabet, Apple, Microsoft, and Meta, have adopted AI technologies to improve their operations and gain a competitive edge. Google, for example, uses AI in its search engine, self-driving cars and AI is responsible for Google's natural language processing breakthroughs.

4.2. Machine Learning

Machine learning is a subset of artificial intelligence and computer science that involves using algorithms and data to imitate human learning and improve accuracy over time. Recent advancements in processing power and data storage have allowed for innovative products based on machine learning, like Netflix's recommendation engine and self-driving cars. [49]

Machine learning is an important part of data science and involves training algorithms to make predictions, classifications, and uncover insights from data mining projects using statistical methods. These insights subsequently drive decision making in many applications and businesses, with the goal of reaching key growth metrics. As big data continues to expand and grow, the market demand for data scientists will increase. They will be required to help identify the most relevant business questions and the data to answer them.

Machine learning algorithms are typically built using frameworks such as TensorFlow[50] and PyTorch [51], which help accelerate solution development.

UC Berkeley has divided the learning mechanism of a machine learning algorithm into three primary components. [52]

1. **A Decision Process:** Here the algorithm utilizes input data, labeled or unlabeled, to make a prediction or classification about a pattern in the data.
2. **An Error Function:** An error function assesses the accuracy of the model by comparing it with known examples.
3. **A Model Optimization Process:** A model optimization process adjusts the weights to minimize the difference between the known example and the model estimate, enabling

the algorithm to update the weights independently until a certain level of accuracy is achieved.

Many types of machine learning models exist, which are categorized based on the degree of human intervention in the raw data, such as rewards, feedback, or labels. As per information provided by Nvidia [53], there are various machine learning models, including:

- **Supervised learning:** The dataset is labeled and classified by users to help the algorithm determine the accuracy of its performance.
- **Unsupervised learning:** The raw training dataset is unlabeled and an algorithm tries to detect patterns and relationships within the data offered without human help.
- **Semi-supervised learning:** The dataset contains structured and unstructured data, guiding the algorithm on its own way of reaching autonomous conclusions and labeling unlabeled data.
- **Reinforcement learning:** The dataset uses a “rewards and punishments” system, providing feedback to the algorithm to learn from its own mistakes through trial and error.

Finally, there is a newer area of machine learning called *deep learning*, which can automatically learn from datasets without human intervention. To improve the accuracy of predictions, this model requires huge amounts of raw data for processing.

4.3. Deep Learning

Deep learning is a subset of machine learning, which is primarily a neural network with three or more layers. These neural networks aim to mimic the behavior of the human brain, learning from massive amounts of data. While a neural network with a single layer can still make rough predictions, having additional hidden layers can optimize and refine accuracy. [54]

Deep learning technology is behind many AI applications and services that automate analytical and physical tasks without human intervention, such as digital assistants, voice-enabled TV remotes, and credit card fraud detection, as well as emerging technologies like self-driving cars. Deep learning/artificial neural networks aim to replicate the human brain's functioning by combining data inputs, weights, and biases to accurately identify, categorize, and describe objects within the data.

Deep neural networks include many layers of interconnected nodes that work together based on the previous layer to improve the accuracy of predictions or categorizations. The process of moving data forward through the network from the input to the output layer is known as forward propagation. The input and output layers are known as *visible* layers. In a deep learning model, the input layer is responsible for receiving the data for processing, while the output layer is responsible for producing the final prediction or classification.

Another process known as backpropagation utilizes algorithms, such as gradient descent, to compute prediction errors and then adapt the function's weights and biases by going in reverse through the layers, with the objective of training the model. Both of them, forward propagation and backpropagation enable a neural network to make forecasts and rectify any errors. As the epochs pass, the algorithm increases its accuracy.

The preceding paragraphs provide a basic explanation of the most straightforward form of deep neural network. Nevertheless, deep learning algorithms are highly intricate, and there exist different types of neural networks that cater to particular datasets or issues. For instance,

- **Convolutional neural networks (CNNs)**, are mainly used for computer vision and image classification purposes. They are capable of recognizing patterns and features within an image, thus enabling tasks such as object identification or detection. In 2015, for the first time, a CNN surpassed a human in an object recognition challenge.
- **Recurrent neural networks (RNNs)**, are commonly used in natural language and speech recognition tasks. They are advantageous in terms of processing sequential or times-series data.

4.3.1. Epoch

To avoid storage space limitations of a computer system, the training data is typically divided into smaller batches before being fed into a machine learning model for training. This approach, known as batch processing in machine learning, enables the model to be trained effectively with smaller amounts of data. When all of the batches have been processed through the model, this is referred to as an epoch [55].

In machine learning, an epoch refers to a cycle of training during which all of the available training data is used. Specifically, an epoch can be defined as the total number of iterations that are required to process all of the training data in one cycle. Alternatively, an epoch can also be understood as the number of times a training dataset is passed through an algorithm. In this context, each pass involves both forward and backward passes through the dataset.

The number of epochs is considered a hyperparameter that determines how many cycles of training are required to process the entire dataset through the learning algorithm. During each epoch, every sample within the training dataset has a chance to update the internal model parameters once, ensuring that the model has had exposure to all of the available data.

By plotting the number of epochs against the corresponding model performance metrics, a learning curve can be generated. The x-axis of this curve represents the number of training iterations, while the y-axis represents the skill of the model. Examining the shape of the curve can help determine whether the model is underfitting, overfitting, or performing well on the training dataset.

4.4. Applications of Deep Learning and CNNs

Deep learning finds extensive applications in different sectors. Here are a few instances:

1. **Computer vision:** Deep learning is widely used in computer vision applications such as object detection, image classification, face recognition, and autonomous vehicles.
2. **Natural language processing (NLP):** NLP tasks such as speech recognition, language translation, and sentiment analysis are now powered by deep learning.

3. **Robotics:** Deep learning is used in robotics applications for object recognition, obstacle avoidance, and path planning.
4. **Healthcare:** Deep learning is used in medical image analysis, disease diagnosis, drug discovery, and personalized medicine.
5. **Finance:** Deep learning is used in finance for fraud detection, risk assessment, and trading algorithms.

Convolutional Neural Networks (CNNs) have numerous applications in a wide range of fields, including:

1. **Image and video recognition:** CNNs have been highly successful in image and video recognition tasks such as object detection, face recognition, and gesture recognition.
2. **Natural language processing:** CNNs can be used for tasks such as text classification, sentiment analysis, and machine translation.
3. **Medical image analysis:** CNNs are used for tasks such as tumor detection, segmentation, and classification in medical images such as CT scans and MRI scans.

4.5. Neural Networks (NN)

Neural networks, also referred to as simulated neural networks (SNNs) or artificial neural networks (ANNs), are a subset of machine learning and serve as the foundation for deep learning algorithms. Their construction and name are modeled after the human brain, imitating the way biological neurons communicate with one another [56].

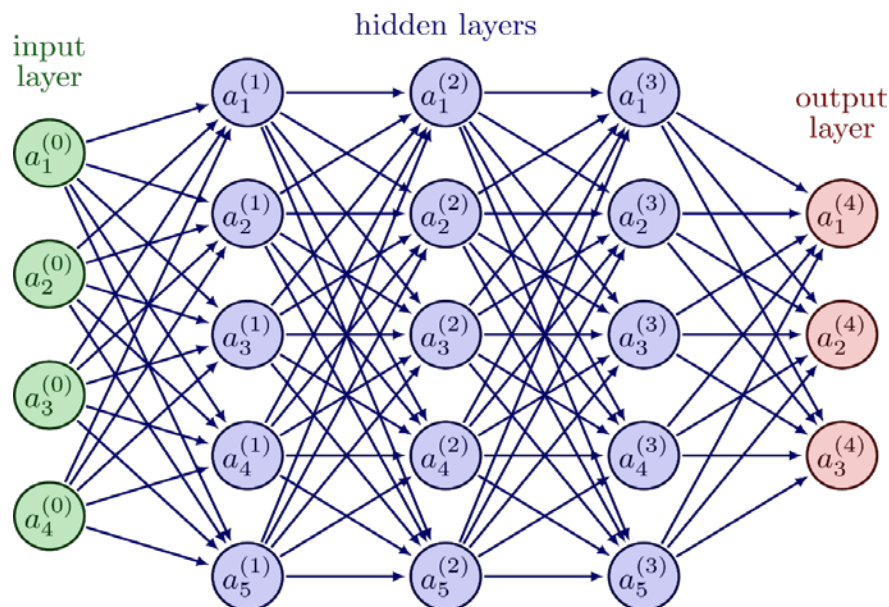


Figure 11. Graph of a typical Neural Network.

Artificial neural networks are constructed from layers of nodes, featuring an input layer, one or more hidden layers, and an output layer. Each node, or artificial neuron, is connected to another

and comes with a weight and threshold value. If the output of any individual node surpasses the given threshold value, that node activates, transmitting data to the next layer of the network. If not, no data is transferred to the next layer.

Neural networks utilize training data to learn and increase their accuracy over time. Nonetheless, after being fine-tuned for accuracy, these learning algorithms are potent tools in computer science and artificial intelligence, enabling us to classify and cluster data speedily. Tasks like speech or image recognition can take minutes compared to the hours required for human experts to identify them manually. Google's search algorithm and ChatGPT are two of the most renowned neural networks.

4.5.1. How do neural networks work?

Imagine each individual node to be an independent linear regression model that consists of input data, weights, a bias (or threshold), and an output. The formula would follow the below equations:

$$\sum w_i * x_i + bias = w_1x_1 + w_2x_2 + w_3x_3 + bias$$

$$output = f(x) = 1 \text{ if } \sum w_i * x_i + b \geq 0; 0 \text{ if } \sum w_i * x_i + b < 0$$

First, the input layer is established, and weights are assigned to determine the importance of each variable. Larger weights contribute more significantly to the output than smaller ones. All input values are multiplied by their corresponding weights and then added together. The output is then processed by an activation function that determines if the output meets a certain threshold. If it does, the node is activated and passes the data to the next layer of the network, where it becomes the input for the next node. This pattern of transmitting information from one layer to the following layer characterizes a feedforward neural network.

To better understand how a single node works using binary values, we can think of a tangible example of deciding whether to go shopping (Yes: 1, No: 0). The predicted outcome, or y-hat, is whether to go shopping or not. Suppose that three factors impact the decision:

1. Is the weather good? (Yes: 1, No: 0)
2. Are many people outside shopping? (Yes: 0, No: 1)
3. Do you have enough money to spend? (Yes: 1, No: 0)

Next, we can make the following assumption, which will provide us with the subsequent inputs:

- $X_1 = 1$, since it is a sunny day, but not too hot.
- $X_2 = 0$, since it is Christmas time and almost everyone is outside.
- $X_3 = 1$, since you have saved an amount of money to spend on gifts.

We must now assign weights to each input to determine their significance in the decision-making process. Higher weight values correspond to variables that are more critical in the decision or result.

- $W1 = 3$, since you prefer not going shopping on rainy or extremely warm days.
- $W2 = 2$, since you are used to the crowd, but you have a liking to not wait in large queues.
- $W3 = 5$, since it is the most important factor to have deposited money to spend and not use your credit card.

Finally, we will also assume a threshold value of 4, which is equivalent to a bias value of -4. Using the above variables and this value, we can start utilizing the formula to obtain the desired output.

$$\hat{y} = (1 * 3) + (0 * 2) + (1 * 5) - 4 = 4$$

By using the activation function described earlier in this section, we can conclude that the output of this node would be 1 because 4 is greater than 0. In this scenario, the decision would be to go shopping. However, if we alter the weights or the threshold, we can obtain different results from the model. The example above demonstrates how a neural network can make progressively complex decisions by building upon the outputs of previous decisions or layers.

In the preceding instance, we used perceptrons to illustrate the mathematical concepts involved in neural networks. However, in reality, neural networks employ sigmoid neurons that produce values between 0 and 1. As neural networks function similarly to decision trees by transferring data from one node to another, using x values between 0 and 1 diminishes the effect of any single variable on the output of an individual node and, consequently, the output of the entire neural network.

When we consider more pragmatic applications for neural networks, such as image recognition or classification, we will leverage supervised learning or labeled datasets to train the algorithm.

As we start to think about more practical use cases for neural networks, like image recognition or classification, we'll leverage supervised learning or labeled datasets, to train the algorithm. As we train the model, we will assess its accuracy using a loss function (or cost function), which is commonly known as the mean squared error (MSE). The equation for MSE is presented below, where:

- i stands for the index of the sample,
- \hat{y} (y -hat) is the predicted outcome,
- y is the actual value, and
- n is the number of samples.

Loss Function (Cost Function) that we chose is Mean Square Error:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 .$$

The primary objective is to minimize the loss function to ensure the best possible fit for any given observation. To achieve this, the model modifies its weights and bias using reinforcement learning and the loss function until it reaches a local minimum, also known as the point of

convergence. Gradient descent is used to adjust the weights of the model, allowing it to determine the best direction to minimize errors and reduce the cost function. The parameters of the model gradually adjust with each training example, allowing it to converge towards the minimum.

Most deep neural networks are feedforward, meaning they only proceed in one direction, from input to output. Nevertheless, it is also possible to train the model using backpropagation, which is the opposite direction from output to input. Backpropagation permits the calculation and assignment of the error associated with each neuron, allowing for appropriate adjustment and fitting of the model parameters.

4.6. Convolutional Neural Networks (CNN)

Different kinds of neural networks are utilized for different use cases and datasets. Before the advent of CNNs, manual, time-consuming feature extraction techniques were used to identify objects in images. Now, convolutional neural networks offer a more scalable approach to image classification and object recognition applications. They accomplish this by leveraging principles from linear algebra, particularly matrix multiplication, to identify patterns in an image. Despite their undisputed usefulness, they can be very computationally demanding, requiring graphical processing units (GPUs) to train models [16].

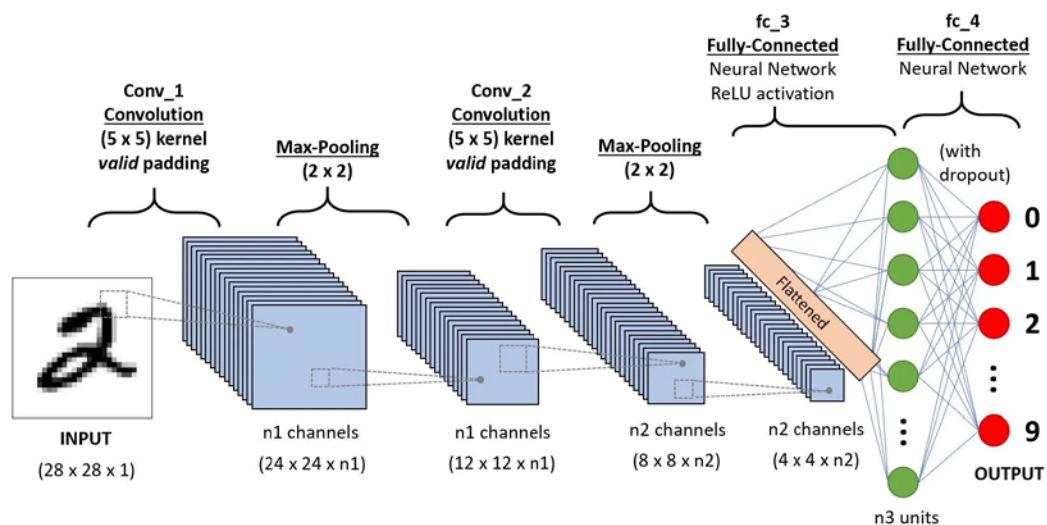


Figure 12. A CNN sequence to classify handwritten digits.

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm that is capable of receiving an input image, assigning importance (learnable weights and biases) to various objects/aspects in the image, and discerning one from another. The pre-processing requirements of a CNN are much lower compared to other classification algorithms. Although filters are hand-

engineered in primitive methods, CNNs can learn these filters/characteristics with sufficient training [17].

The architecture of a ConvNet is like the connectivity pattern of neurons in the human brain and was inspired by the organization of the Visual Cortex. In the visual field, Individual neurons react to stimuli only in a limited area known as the Receptive Field. A set of such fields overlap to cover the entire visual area.

Convolutional neural networks are highly recognized for their superior performance with image, speech, or audio signal inputs in comparison to other neural networks. They have three main types of layers, which are:

- **Convolutional layer**
- **Pooling layer**
- **Fully-connected (FC) layer**

The convolutional layer is the initial layer of a convolutional network, which can be succeeded by other convolutional or pooling layers. In contrast, the fully-connected layer is the ultimate layer. As each layer of the CNN is added, it becomes more complex and can identify larger parts of the image. At first, the earlier layers detect straightforward characteristics like edges and colors. As the image data advances through the CNN layers, it begins to perceive more substantial features or forms of the object until it ultimately recognizes the target object.

4.6.1. Convolutional Layer

The core block of a CNN is the convolutional layer, which performs most of the computation. It needs three components: input data, a feature map, and a filter. When working with a color image, the input data will consist of a 3D matrix of pixels representing height, width, and depth (corresponding to RGB in an image). The feature detector, also named as a kernel or filter, moves across the image receptive fields to check for the presence of a feature. This process is called convolution.

The feature detector is a two-dimensional (2-D) array of weights that represents a portion of the image. Although the filter size is not fixed, it is usually a 3x3 matrix, which also determines the receptive field size. Next, the filter is implemented to a section of the image, and the input pixels together with the filter generate a dot product. This dot product is then forwarded into an output array. The filter moves by a stride and repeats the process until the entire image is covered. The output obtained from the dot products of the input and filter is called a feature map, activation map.

Following every convolution operation, a CNN implements a Rectified Linear Unit (ReLU) transformation to the feature map, which introduces nonlinearity to the model.

As previously mentioned, a subsequent convolution layer may follow the initial convolution layer, creating a hierarchical structure in the CNN. Thus, the later layers have access to the pixels in the receptive fields of the prior layers. To illustrate this, we can consider identifying an automobile in an image. The vehicle can be viewed as a set of parts, such as a windshield, windows, wheels, mirrors, doors, etc. Each part of the car corresponds to a lower-level pattern in the neural network, while the combination of all these parts corresponds to a higher-level pattern, forming a feature hierarchy within the CNN.

4.6.2. Pooling Layer

Pooling layers, which are also referred to as downsampling, perform dimensionality reduction by decreasing the number of input parameters. The pooling operation is similar to that of the convolutional layer, as it employs a filter that is swept across the entire input. However, the filter in the pooling layer does not have weights. Rather, the kernel uses an aggregation function to the values within the receptive field, which are then formulating the output array. There are two primary types of pooling:

- **Max pooling:** As the filter moves across the input, it selects the pixel with the maximum value to send to the output array. At this point, it should be stated that this method is generally employed more frequently than the average pooling.
- **Average pooling:** As the filter moves across the input, it calculates the average value within the receptive field to send to the output array.

Although much information is lost in the pooling layer, it also has several advantages for the CNN. Specifically, they are beneficial in terms of reducing complexity, achieving efficiency, and minimizing the risk of overfitting.

4.6.3. Fully-Connected Layer

The fully-connected layer is appropriately named because it connects every node in the output layer directly to a node in the previous layer. This comes in contrast to partially connected layers, where the pixel values of the input image are not directly linked to the output layer.

The layer is responsible for classification by utilizing the resulted features from the previous layers and their distinct filters. ReLu functions are mainly used in convolutional and pooling layers, but FC layers typically implement a softmax activation function to ensure appropriate input classification, generating a probability score between 0 and 1.

4.6.4. ResNet50 Architecture

ResNet stands for Residual Network and is a specific type of CNN introduced in 2015 [18]. ResNet-50 is a variant of this network that consists of 50 layers, including 48 convolutional layers, one MaxPool layer, and one Average pool layer. Residual neural networks are a type of ANN that constructs networks by stacking residual blocks.

ResNet-50 has an architecture which is illustrated in the provided model diagram. The 50-layer ResNet employs a bottleneck design for its building blocks. This design incorporates 1×1 convolutions, known as "bottlenecks," that reduce the number of parameters and matrix multiplications. Consequently, this approach enables faster training for each layer. Unlike the typical two-layer structure, ResNet-50 utilizes a stack of three layers in its bottleneck residual blocks.

The architecture of the 50-layer ResNet includes the following components, as depicted in the diagram below [19]:

- **A 7×7 kernel convolution** alongside 64 other kernels with a 2-sized stride.
- **A max pooling layer** with a 2-sized stride.

- **9 additional layers** of 3x3, 64-kernel convolution, followed by another with 1x1, 64 kernels, and a third with 1x1, 256 kernels. This sequence is repeated 3 times.
- **12 more layers** with 1x1, 128 kernels, 3x3, 128 kernels, and 1x1, 512 kernels, iterated 4 times.
- **18 more layers** with 1x1, 256 kernels, 3x3, 256 kernels and 1x1,1024 kernels, repeated 6 times.
- **9 more layers** with 1x1, 512 kernels, 3x3, 512 kernels, and 1x1, 2048 kernels iterated 3 times.

(At this stage, the network has reached 50 layers)

- **Average pooling layer**, followed by a fully connected layer with 1000 nodes, implementing the softmax activation function.



Figure 13. A block diagram representation of pre-trained Resnet-50 architecture.

4.7. Computer Vision

Computer Vision, also known as Machine Vision, is a subfield of Artificial Intelligence (AI) that utilizes digital images, videos, and other visual data to enable computers and systems to extract significant information and provide suggestions or take action based on that information [20].

While AI empowers computers to think, the purpose of computer vision is to help computers see, scrutinize, and comprehend visual content, like human vision, albeit with different tools and time to learn. Human sight has the benefit of a lifetime of experiences to train itself on how to differentiate objects, determine their distance, detect motion and identify issues in images.

Computer Vision uses cameras, data, and algorithms to teach machines to perform such functions, without relying on human-like biological systems such as retinas, optic nerves, and visual cortex. When machines are trained to scrutinize products or oversee the operation of a production asset, they can analyze thousands of products or processes per minute and identify even the slightest imperfections or irregularities. This ability allows machines to outperform humans quickly. Computer vision has become ubiquitous in a variety of industries, from energy and utilities to manufacturing and automotive, and the market for this technology is still expanding [21].

To achieve accurate image recognition, computer vision requires a large amount of data. It processes and analyzes the data repeatedly until it can differentiate between objects and eventually recognize them. For instance, in order for a computer to learn to identify automobile tires, it needs to be given a massive amount of tire images and tire-related objects so it can learn the differences and detect a tire, additional caution has to be paid to the ones with no abnormalities.

To accomplish this, two vital technologies are utilized: deep learning and CNNs. Machine learning involves the use of algorithmic models that allow computers to teach themselves about visual data context. When enough data is fed through the model, the computer will "see" the data and learn to differentiate one image from another. Algorithms allow the machine to learn on its own, without being programmed by a human to recognize a particular image.

Convolutional Neural Networks assist machine learning and deep learning models to "view" images by breaking them down into pixels and assigning tags or labels to them. These labels are then used to perform mathematical operations, called convolutions, to predict what the image represents. The neural network conducts convolutions and verifies the accuracy of its predictions in a series of iterations until the predictions reach the real value, allowing it to recognize or "see" images in an analogous manner to humans.

Similar to how humans identify distant objects, a CNN initially detects hard edges and basic shapes and then refines its prediction with each iteration to recognize the image. CNNs are used to comprehend single images, while RNNs are used similarly but for video applications to assist computers in comprehending how images are interconnected in a sequence of frames.

4.8. Multispectral Imaging in Computer Vision

Multispectral images provide additional information to that of RGB images, which make them useful in increasing accuracy for our image classification model. However, the implementation of those images is not as effortless as for the RGB ones, since they have either 16 bands (VIS) or 25 bands (NIR). First, in order to read the raw binary files and convert them into the commonly established 3-dimensional hypercubes, a certain procedure is required. Then, the process followed to end up with 3-channel images and maintaining the most information possible is known as Principal Component Analysis. However, in this case, the 3 channels do not represent Red, Green and Blue colors but they are different shades of black and white.

4.8.1. Principal Component Analysis (PCA)

Principal component analysis, or PCA, is a dimensionality reduction method that is often used to reduce the dimensionality of large datasets, by transforming a large set of variables into a smaller one that retains most of the information in the initial large set [57].

PCA involves a mathematical process that calculates the covariance matrix of the original variables, which represents the relationships between each variable. It then computes the eigenvectors and eigenvalues of this matrix, which indicate the direction and magnitude of the

most significant patterns in the data, respectively. The eigenvectors are the principal components that form a new coordinate system for the data, and the eigenvalues represent the amount of variance in the data captured by each principal component.

The principal components are arranged in descending order of their eigenvalues, so that the first principal component accounts for the most significant pattern in the data, the second principal component accounts for the next most significant pattern, and so on. By selecting a subset of the principal components that account for the majority of the variance in the data, PCA can effectively reduce the dimensionality of the dataset without losing important information.

PCA is widely used in various fields, such as image and signal processing, data compression, and machine learning, where it is used as a preprocessing step to simplify and enhance the data before applying other algorithms. Although reducing the number of variables in a dataset can result in a loss of accuracy, the trade of dimensionality reduction is the simplicity increase, making it easier to explore and visualize smaller datasets. This, in turn, makes it easier and quicker for machine learning algorithms to analyze data points without extraneous variables to process. In essence, the goal of PCA is to decrease the number of variables in a dataset while maintaining as much information as possible.

PCA is commonly used in multispectral imaging. Multispectral imaging produces high-dimensional data, which can be challenging to analyze and interpret. PCA can be used to reduce the dimensionality of the data by identifying the most significant spectral features in the images and transforming them into a smaller set of principal components.

In multispectral imaging, the principal components represent the most significant spectral patterns in the data. For example, the first principal component might represent the overall brightness of the scene, while the second principal component might capture the difference between vegetation and non-vegetation areas. By selecting a subset of the principal components that capture most of the spectral variability in the data, PCA can effectively reduce the dimensionality of the data without losing much important information.

PCA can also be used to enhance the visual appearance of multispectral images by creating false-color composites that highlight specific features in the data. For example, by combining the first principal component with the third and fourth principal components, it is possible to create a false-color composite that enhances the contrast between different types of vegetation and other features in the scene.

4.8.2. Applying PCA on Multispectral Images

Dealing with multispectral data has a drawback, which is the large number of bands that need to be processed, and the challenge of storing such a huge amount of data. The time complexity also increases with the increase in data volume. Therefore, it is important to either reduce the data amount or select only the relevant bands. Nonetheless, it is important to ensure that the quality of classification does not decrease with the reduction in the number of bands. One way to address this issue is to preprocess the multispectral data using Principal Component Analysis (PCA) [58].

PCA relies on the fact that the adjacent bands of the Multispectral Image are highly correlated and convey the same information. This technique is based on the mathematical principle known

as Eigenvalue decomposition of the covariance matrix of the Hyperspectral bands that are being analyzed. The eigenvalues and eigenvectors are then arranged in a descending order to identify the ones that hold the most significant information, which will be placed at the top of the list [59].

Initially, we begin with Multispectral images containing 16 or 25 bands, which need to be narrowed down to just three bands. To achieve this, we should select the three eigenvectors that carry the most information, just as in RGB images that are composed of three channels. Thus, we extract the top three eigenvectors after the PCA is completed.

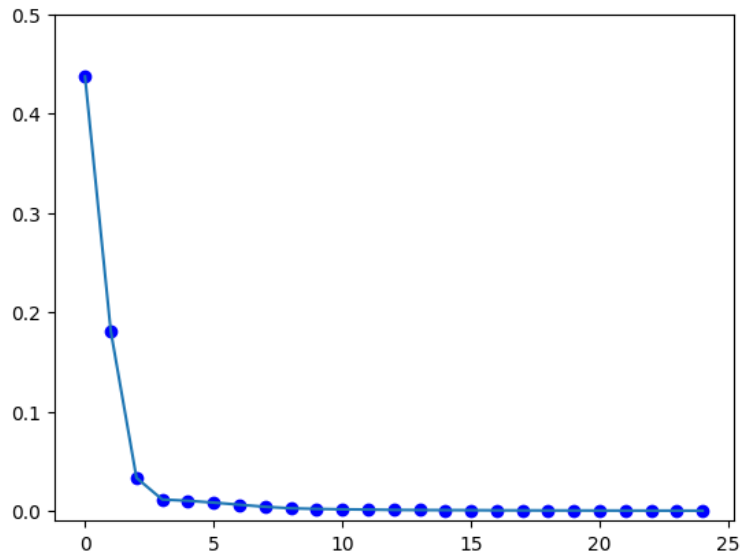


Figure 14. Plot of the eigenvalues of the NIR image of the LDPE product depicted below.



Figure 15. The NIR image of the LDPE product whose eigenvalues are shown above.

The above graph demonstrates that the three highest eigenvalues contain most of the relevant information, while the remaining ones primarily represent noise. Employing PCA not only decreases the time required for classification, but it also decreases the amount of data that must be processed. Lastly, this preprocessing technique produces classification outcomes that are reasonably precise and accurate.

Chapter 5

Tools and Methods

5.1. Data Generation and Training of the Model

In this chapter, the steps that were followed in this research will be addressed. Firstly, we collected a large dataset of objects. Our dataset consists of two main classes and seven subclasses. Our two main classes are Wood and Plastic and our seven subclasses include Medium-Density Fiberboard (MDF), Melamine-Faced Chipboard (MFC), Oak Veneer, which belong to the Wood, while the rest High Density PolyEthylene (HDPE), Low Density PolyEthylene (LDPE), Polyethylene terephthalate (PET), Polypropylene (PP) belong to Plastic. Thus, in total we have three kinds of wood and four kinds of plastic. Approximately 100 pieces of each category were accumulated that range in types and sizes.

Then, we continued with the generation of the image dataset. This process involved placing every object on the platform and capturing photographs with the help of the cameras. We used the RGB industrial camera, the visible (VIS) multispectral camera and the near-infrared (NIR) multispectral camera. Therefore, for every object we have three different types of images captured which will eventually form our three classifiers. In addition, we changed the orientation of the objects and repeated the same process for two main reasons. One of them is to increase the number of images in order for the model to train better and the other is to test different shadows, placements and angles of the object.

Next, stood the multispectral data preprocessing. It included reading the raw files and converting them to 3-dimensional hypercubes, which had either 16-bands (VIS) or 25-bands (NIR). Furthermore, a PCA had to be executed in order to obtain the top three eigenvalues with the most useful information.

Subsequently, after having converted all three types of images into files that we can work with, the resizing and normalization were our succeeding steps. Data Augmentation came just after and played a major role. At this step, the aim was to create more data in order to test whether our models would make a wrong decision in every augmented version of the samples. This technique included creating five new (so, in total 6 images of every object) versions of the captured images by applying random zoom (0 to +20%), random rotation (-36° to +36°) and random horizontal flip.

Hence, we ended up with 6522 image files from every camera type, which are the input of our CNN. These images were then split into training and validation sets, which do not overlap with each other and do not share any version of the same sample. The rule followed for this separation is the one that is the most commonly used [22], the 80-20% ratio, 80% for the training set and 20% for the validation set.

As a result, the training set contains 5238 files while the validation set contains 1284, which add up to the total of 6522 image files. Into more detail, each material class of the seven contributed as shown below:

Class	Number of Training Images	Number of Validation Images	Total Images
MDF	456	108	564
MFC	552	132	684
Oak Veneer	576	144	720
HDPE	498	120	618
LDPE	492	120	612
PET	1398	348	1746
PP	1266	312	1578

Table 3. Contribution of each material image to the sets.

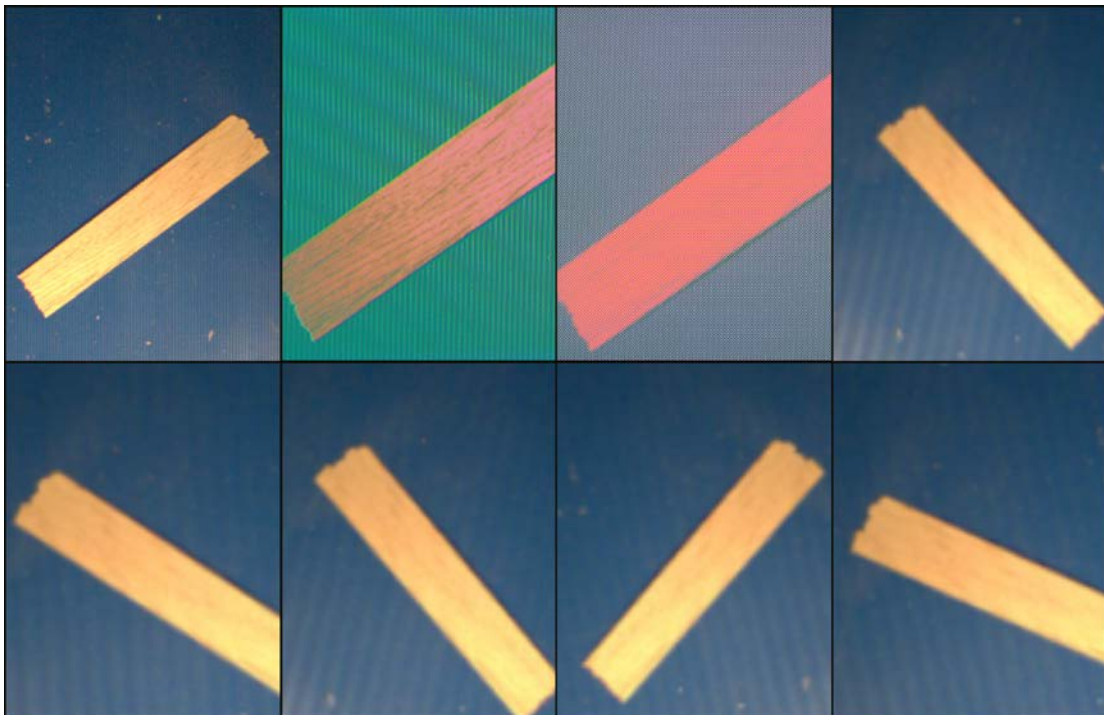


Figure 16. Image Data Augmentation: Starting from the top left corner: Original RGB Image of an Oak Veneer piece of wood, Original VIS image, Original NIR image, Five Augmented RGB versions of the same piece.

An important procedure during this study was training models with various architectures and experimenting by changing their parameters until finding the one that had satisfactory performance.

The ultimate parameters selected for the training of the model are:

- **Criterion:** Cross Entropy Loss between input logits and target, since this type of loss reaches its maximum potential when training a classification problem with various classes by adjusting the weights of the model. The mathematical formula of cross entropy loss is:

$$L_{CE} = -\sum_{i=1}^n t_i \log(p_i),$$

where n is the number of classes, t_i is the truth label and p_i is the Softmax probability for the i^{th} class.

- **Optimizer:** Stochastic Gradient Descent (SGD). The key idea behind SGD is to update the model parameters iteratively by taking small steps in the direction of the steepest descent of the loss function. In each iteration, SGD randomly selects a subset of training examples (called a mini-batch) and computes the gradients based on that mini-batch. The gradients are then used to update the model parameters. This process is repeated for multiple iterations until convergence, or a stopping criterion is reached. It requires careful tuning of learning rate and other hyperparameters, and its convergence can be sensitive to the choice of mini-batch size. That is the reason behind our four next choices.
- **Learning Rate:** 0.001. The learning rate is a hyperparameter used to govern the pace at which the algorithm adjusts or learns the values of a parameter estimate, meaning that it regulates the weights of our CNN regarding the loss gradient. It indicates the frequency at which the neural network incorporates the knowledge it has gained, refreshing its learned concepts.
- **Batch Size:** 4. The batch size defines the number of samples that will be propagated through the network.
- **Gamma:** 0.1, a multiplicative factor by which the learning rate will deteriorate. For example, if the learning rate is 1000 and gamma is 0.1, the new learning rate will be $1000 \times 0.1 = 100$ and so on.
- **Step size:** 7, decays the learning rate of each parameter group by gamma every N (N is defined as 7 here) epochs.
- **Epochs:** 50

Their performance was calculated through the CNN metrics, such as accuracy and loss. The outcomes of this attempt are presented in a subsequent section. Thus, by utilizing our selected CNN architecture, we generated three image prediction models - classifiers (RGB, VIS and NIR). The task of these classifiers is to accept an image of their type as an input and give as an output their vector of probabilities – the probability that the given input fits into each of the seven predetermined categories. The category with the highest percentage is considered as their final output-prediction.

At this point, it was necessary to combine those three classifiers and compare their performance individually and in pairs. Again, the outcomes are projected in a later section.

Finally, an excessive review upon bands required the repetition of the whole process, at least for the multispectral images. At this stage, we are examining the impact of worse equipment, laboratory conditions and consequently images with less bands and therefore the significance of the role that the bands play. For this purpose, it is fundamental to compare the CNN performances of the good equipment and of the inadequate equipment, the one that captures multispectral images with less bands.

5.1.1. Approaches to the selection of the bands in case of inadequate equipment

There are four different approaches that were followed to select only three bands instead of their combination provided by the method of PCA. Of course, these approaches are referring to the multispectral images due to their nature of having more than three bands in comparison to RGB images. As we have already mentioned, we are working with seven different classes and two types of multispectral images, NIR and VIS.

The first perspective was random selection, which was repeated enough times until no major improvement or deterioration was shown to the average results. The intention behind implementing this method was to act as a reference point for the other three approaches.

The second technique was “best in total” selection. By applying PCA on our two multispectral datasets we ended up with the eigenvalues and eigenvectors of every image. We isolated the first three eigenvalues and eigenvectors of every image since the rest eigenvalues include mostly noise. Then, by multiplying them and repeating the same process for every image we derived the sum of all the images’ product which indicated the three bands that offer the most information for every type of spectrum. Thus, for images of NIR type we concluded that the most important bands are 1st, 12th and 17th, and for images of VIS type 2nd, 15th and 16th.

The third method applied was “best of each class” and then selecting the 3 most repeated bands. We employed the same process as in the second technique. Nonetheless, we applied it separately to each class, noting down the top three bands for every class. Ultimately, we identified the three bands that appeared most frequently. Thus, for images of NIR type we concluded that the most repeated bands are 5th, 12th and 18th, and for images of VIS type 2nd, 10th and 14th.

The fourth approach was similar to the third one, however, implemented on the worst bands of each category, followed by the selection of the most frequently recurring bands. Thus, for images of NIR type we concluded that the most repeated bands are 20th, 23rd and 24th, and for images of VIS type 3rd, 4th and 7th.

Chapter 6

Experiment

6.1. Experimental Setup

The laboratory setup for the experiments follows the common setups found in bibliography [23], [24], [25], [26]. The Figure shows the setup implemented for the purposes of this study. The main elements portrayed include, firstly, a conveyor belt which shares the same characteristics of an actual production line.

Secondly, there is an industrial RGB camera, with the following specifications:

Characteristic	Value	Characteristic	Value
Sensor Vendor	e2v	Resolution (H x V)	1280 px x 1024 px
Sensor	EV76C560	Resolution	1.3 MP
Shutter	Global & Rolling Shutter	Pixel Size (H x V)	5.3 μm x 5.3 μm
Sensor Format	1/1.8"	Frame Rate	60 fps
Sensor Type	CMOS	Mono/Color	Color
Sensor Size	6.8 mm x 5.4 mm	Pixel Bit Depth	12 bits

Table 4. Specifications of the Industrial Camera.

Thirdly, there exist two multispectral sensors, one of visible, Short-Wave InfraRed light (SWIR Sensor) and one of near-infrared light (Multi-spectral Sensor) with their specifications shown below:

Method of Capturing, Camera Heads	Spectrum
Snapshot, SWIR Sensor Head (VIS)	470 – 620 nm, 16 bands
Snapshot, Multispectral Sensor Head (NIR)	600 – 975 nm, 25 bands

Table 5. Specifications of the multispectral sensors.

In addition, the light source that is utilized offers a continuous spectrum that is fundamental for the image capturing. Lastly, there is a robotic arm/manipulator to place the samples in the right container after predicting their classes.

All these components are interconnected through a Central Processing Unit (CPU), which gathers the information provided and processes it. Not only that, but the CPU is also responsible for the actions of the robotic arm by directing it.

The conveyor belt transfers the object to the front of the RGB camera where its first photograph is taken and then the same procedure is followed with the other two multispectral cameras. The light source is an important factor for the capture of the multispectral images since it illuminates the sample and (as it is shown in a following section) affects the expected results tremendously. Then, the average probability of the three classifiers' predictions provides the

ultimate result (prediction of the material) and an order for the robotic arm to place the object in the correct container. As far as the computer vision part is concerned, the machine learning framework and the CNN architecture employed are PyTorch and ResNet50 respectively. The model is set to run for 50 epochs and to stop in case the loss of the training dataset falls under 0.1 to avoid overfitting. Moreover, the state of the model at the epoch with the best accuracy is saved at the end.

The code developed to implement this thesis can be found in the following GitHub repository ([27]).

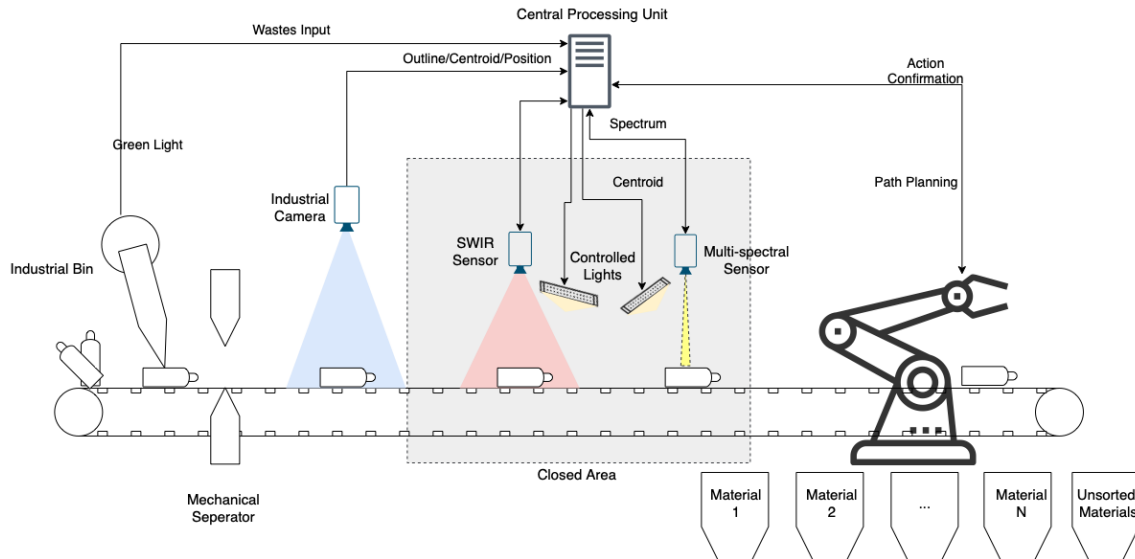


Figure 17. Laboratory Setup of our Experiments - Autonomous sorting system, which is made up of an industrial bucket, a mechanical separator, a multi-camera identification system, a robotic arm, bins, and a central processing unit.

6.2. Evaluation Metrics

6.2.1. Confusion Matrix

A confusion matrix is a table representing the performance of a classification algorithm. It provides a concise summary of the model's performance in terms of correctly identifying and misclassifying instances of different classes. An example of a confusion matrix is depicted at the previous figure, where "N = benign" tissue is classified as healthy and "P = malignant" tissue is classified as cancerous [28].

The confusion matrix consists of four main values that are utilized to define the evaluation metrics of a classifier. These four numbers are:

1. **True Positive (TP):** This represents the number of patients who have been correctly classified as having malignant nodes, meaning having the disease.

2. **True Negative (TN):** This represents the number of patients who have been correctly classified as healthy.
3. **False Positive (FP):** This represents the number of patients who have been misclassified as having the disease, but they are actually healthy. FP is also known as a *Type I error*.
4. **False Negative (FN):** This represents the number of patients who have been misclassified as healthy when they actually have the disease. FN is also known as a *Type II error*.

		Actual class	
		P	N
Predicted class	P	TP	FP
	N	FN	TN

Figure 18. Confusion Matrix.

6.2.2. Accuracy

Accuracy is a performance metric used to assess the effectiveness of classification models. It refers to the proportion of right predictions that the model has made [29]. More specifically, accuracy has the following definition:

$$Accuracy = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} .$$

For binary classification, accuracy can also be computed in terms of positives and negatives as follows:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} .$$

Where TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives, as defined earlier.

We will now examine the accuracy of a given model using its confusion matrix. The model in question has classified 100 tumors as either malignant (the positive class) or benign (the negative class) with the results shown in Table 6 below.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} = \frac{1+90}{1+90+1+8} = 0.91 .$$

The accuracy of the model is calculated to be 0.91 or 91%, which means that it has made 91 correct predictions out of the total of 100 examples. At first glance, this high accuracy may suggest that the tumor classifier is effectively identifying malignancies. However, to gain a more thorough

understanding of the model's performance, it is necessary to conduct a more detailed analysis of the positives and negatives.

Out of the 100 tumor examples, 91 of them are benign (90 true negatives and 1 false positive), while the remaining 9 are malignant (1 true positive and 8 false negatives). Out of the 91 tumors that are benign, the model has accurately classified 90 of them as benign which is a positive outcome. However, out of the 9 malignant tumors, the model has only correctly identified 1 as malignant. This is a concerning result, as the model has missed 8 out of 9 malignant tumors, indicating that the model's ability to detect malignancies is poor.

<p>True Positive (TP):</p> <p>Real value: Malignant ML model predicted: Malignant Number of TP results: 1</p>	<p>False Positive (FP):</p> <p>Real value: Benign ML model predicted: Malignant Number of FP results: 1</p>
<p>False Negative (FN):</p> <p>Real value: Malignant ML model predicted: Benign Number of FN results: 8</p>	<p>True Negative (TN):</p> <p>Real value: Benign ML model predicted: Benign Number of TN results: 90</p>

Table 6. Confusion matrix of a model that classifies 100 tumors as either malignant or benign.

An accuracy of 91% may seem impressive. However, if we consider another tumor-classifier model that always predicts benign, it would achieve the exact same accuracy of 91 out of 100 correct predictions. This means that the current model is no better than a model that has no predictive ability in distinguishing between malignant and benign tumors. When dealing with a class-imbalanced data set, such as the one in this scenario, where there is a significant difference between the number of positive and negative labels, accuracy alone cannot fully describe the situation.

6.2.3. Loss Function

A loss function is a function that compares the predicted output values of a neural network to the expected output values and calculates the difference between them. The loss function is used to evaluate how well the model is performing on the training data. During training, the goal is to minimize the loss between the predicted and target outputs [30].

We tune the hyperparameters to minimize the average loss, which involves finding the values of weights, w^T , and biases, b , that lead to the minimum value of J , representing the average loss:

$$J(w^T, b) = \frac{1}{m} \sum_{i=1}^m L(\hat{y}^{(i)}, y^{(i)}).$$

Analogous to the concept of residuals in statistics, which quantify the deviation of the actual y values from the regression line (predicted values), our aim is to minimize the net distance.

There are two primary types of loss functions used in supervised learning, which correspond to the two main types of neural networks: regression and classification loss functions.

1. **Regression Loss Functions:** are employed in regression neural networks, where the model predicts an output value corresponding to an input value instead of pre-defined labels. Examples of regression loss functions include Mean Squared Error and Mean Absolute Error.
2. **Classification Loss Functions:** are employed in classification neural networks. They produce a vector of probabilities that indicates the likelihood of an input belonging to various pre-set categories, allowing the network to select the category with the highest probability. Examples of classification loss functions include Binary Cross-Entropy and Categorical Cross-Entropy.

Unlike accuracy, loss is not expressed as a percentage but rather as a summation of errors made on each sample in the training or validation sets. In the training process, minimizing the value of the loss function is the objective in order to determine the optimal parameter values for the model, such as the weights in a neural network.

6.3. Results and Discussion

The first outcome is related to the model architecture used. After trying different models, we reached the conclusion that ResNet50 is the one that best fitted our purposes and the most optimal for this research in terms of GPU and available resources usage, expected accuracy provision and of handling a medium dataset of images.

The second result refers to the classifiers. The three classifiers created, give better results when combined instead of each one separately. Additionally, the Average Classifier is the one that provides the best results. The Average Classifier is created by adding every class's probability percentage of the three other classifiers (RGB, NIR, VIS) and then dividing by 3 to get the average percentage. For a better understanding, it is worth mentioning that classification neural networks work by outputting a vector of probabilities. These probabilities refer to how well the given input fits into each of the pre-set categories and then the classification NNs select the category-class with the highest probability as the final output.

Some of the outcomes provided by the average classifier are:

- **Confidence** in case of a right guess: 80.968%.
- **Confidence** in case of a wrong guess: 52.705%.
- **Accuracy:** 92.523%.
- **Accuracy:** 93.457%, by applying the reverse process of augmentation.
- Below stand the **confusion matrices** of the Average Classifier. The second and fourth matrix depict the results of the reverse augmentation. The reverse process of augmentation is a technique similar to repetition coding in telecommunications. Into more detail, they present what the result is if we take into consideration the majority of the

versions' guesses for every image. For example, if 4 out of 6 versions of an image are selected as LDPE then it is considered as an LDPE.

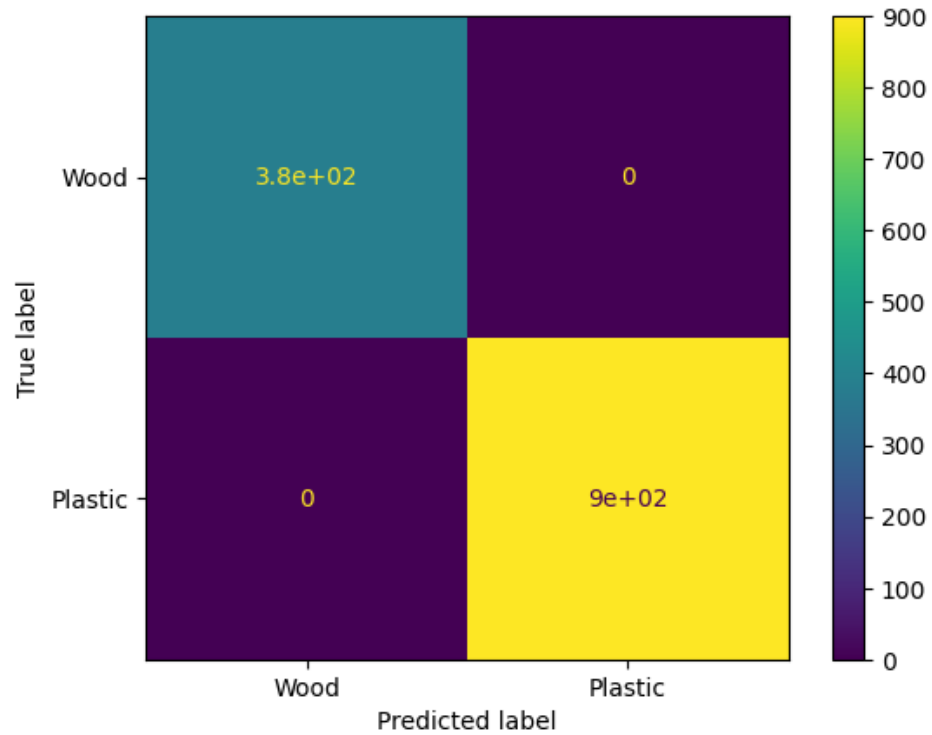


Figure 19. Confusion Matrix of Wood and Plastic of the Average Classifier.

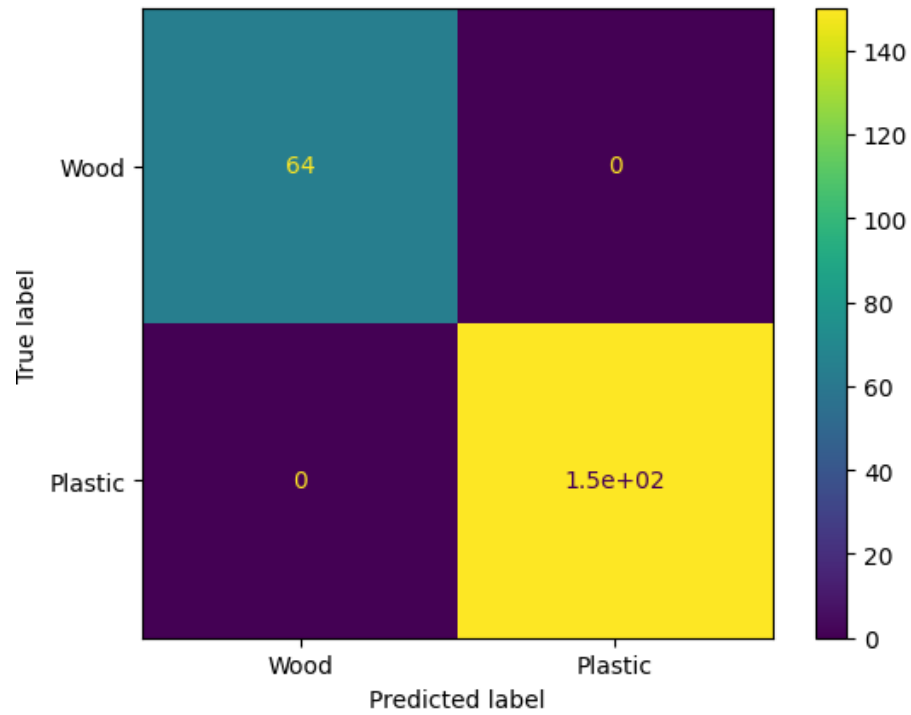


Figure 20. Confusion Matrix of Wood and Plastic of the Average Classifier (Reverse Augmentation).

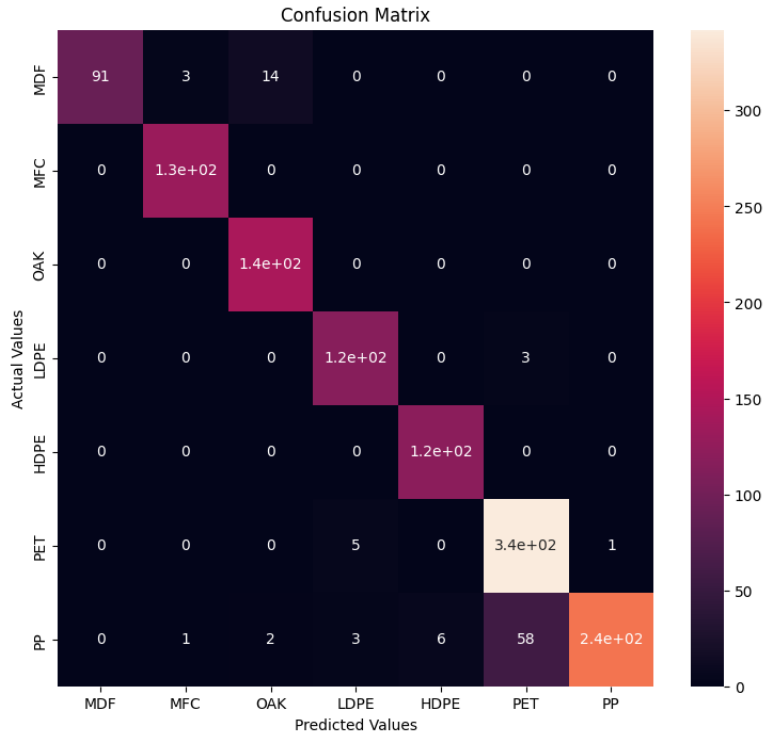


Figure 21. Confusion Matrix of all 7 categories of the Average classifier. The following misclassifications are depicted: 17 samples of MDF, 3 images of LDPE, 6 of PET and 70 of PP, while no mistakes were made for MFC, OAK or HDPE class.

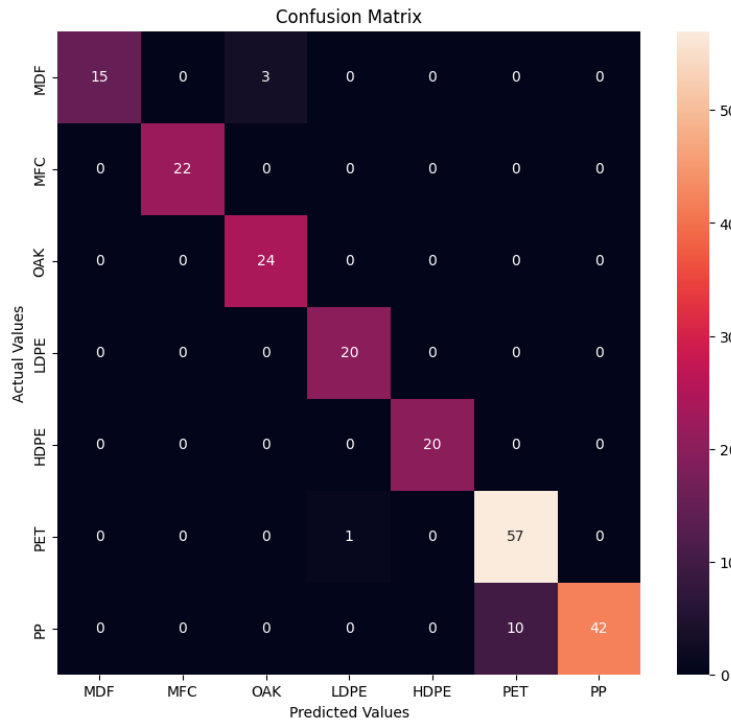


Figure 22. Confusion Matrix of all 7 categories of the Average classifier (Reverse Augmentation). We can observe that only 14 out of 214 sets of 6 versions are misclassified, hence the 93.457% accuracy.

Another outcome produced by the average classifier that has its own value is the number of wrong guessed versions per image. To present this, we use the help of histogram. The histogram displays the number of images that have had their versions -from zero to all six of them- distributed wrong. For instance, four errors per image means that four of the six versions were misclassified, while zero means that all were correctly classified.

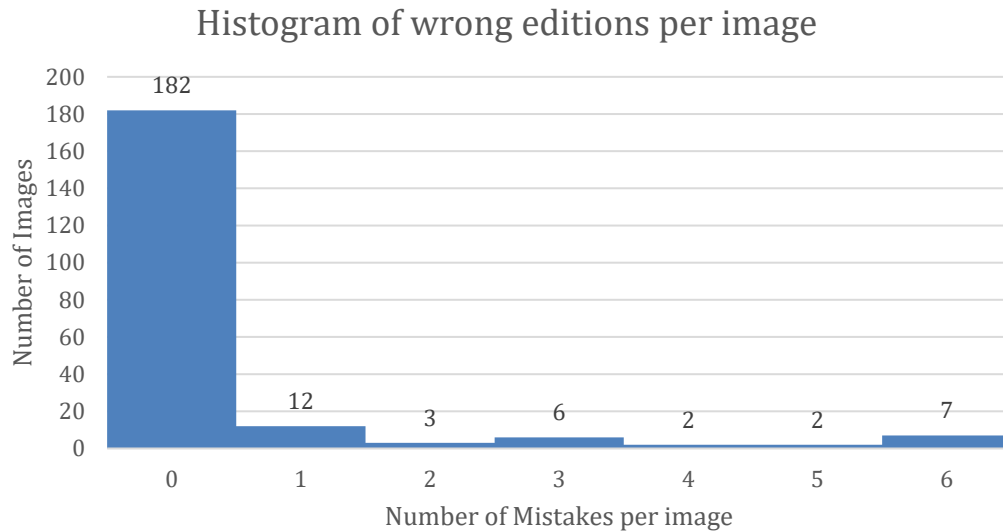


Figure 23. Histogram of wrong editions per image. Six mistakes per image means that all of the six editions were classified in the wrong category while zero stands for all of them being classified right.

The final section of the conclusions addresses the impact of inadequate equipment on classifiers' accuracy along with a study on band selection. The influence of lower-quality equipment is demonstrated by examining the accuracy of multispectral classifiers. The analysis reveals that NIR cameras exhibit a mere 2.07% accuracy deficit, whereas VIS cameras display a marginal 1.69% decrease when compared to the best accuracy achieved with NIR (78.74%) and VIS cameras (81.72%), respectively. The subsequent diagrams illustrate the accuracy progression during epochs until overfitting occurs. Moreover, we conducted experiments on dimensionality reduction for multispectral cameras, exemplifying feature selection using only three different bands and evaluating its implications. Furthermore, by selecting various sets of three bands from the multispectral cameras, we demonstrated that spatial and shape focus do not significantly outweigh spectral footprint. This is supported by the fact that the average accuracy of randomly selected band sets differs by less than 8% from the highest accuracy achieved by three-band sets and less than 10% from utilizing PCA and including all bands. Additionally, when utilizing the three worst bands of NIR cameras, the difference in outcomes compared to the best three-band selected is less than 7%.

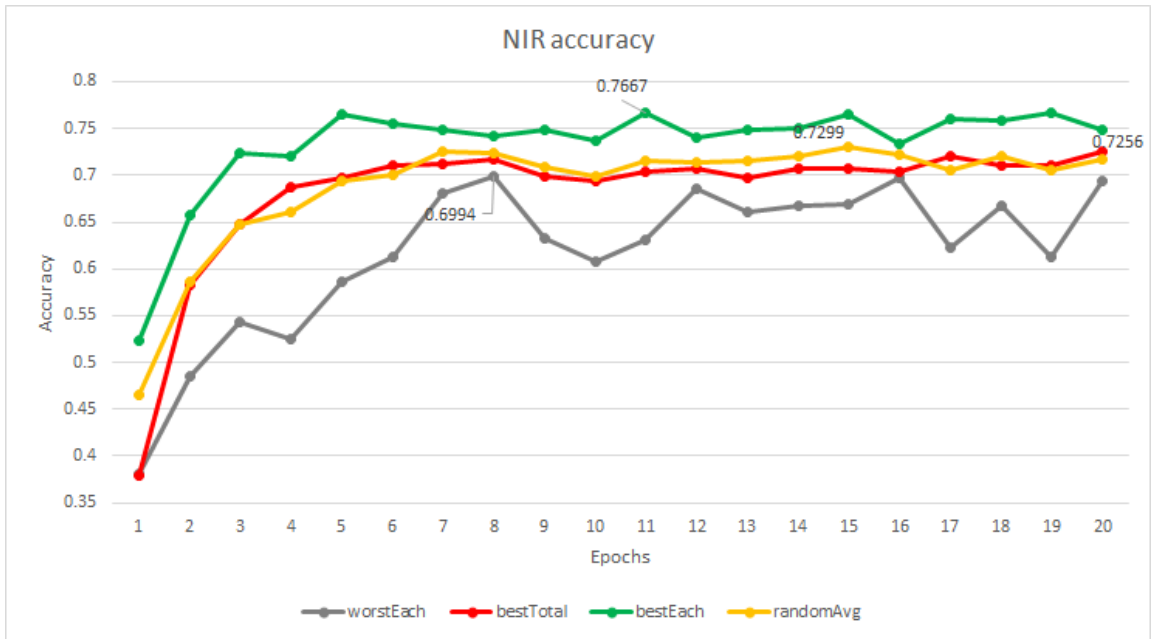


Figure 24. Diagram presenting the NIR camera three-banded sets' accuracy evolution as epochs progress. The peak of each line is depicted with a decimal number. Line explanation: "randomAvg,, stands for the average of the random selection, "bestTotal,, stands for the second technique ("best in total,,), while "bestEach,, and "worstEach,, stand for the third and fourth method applied, respectively.

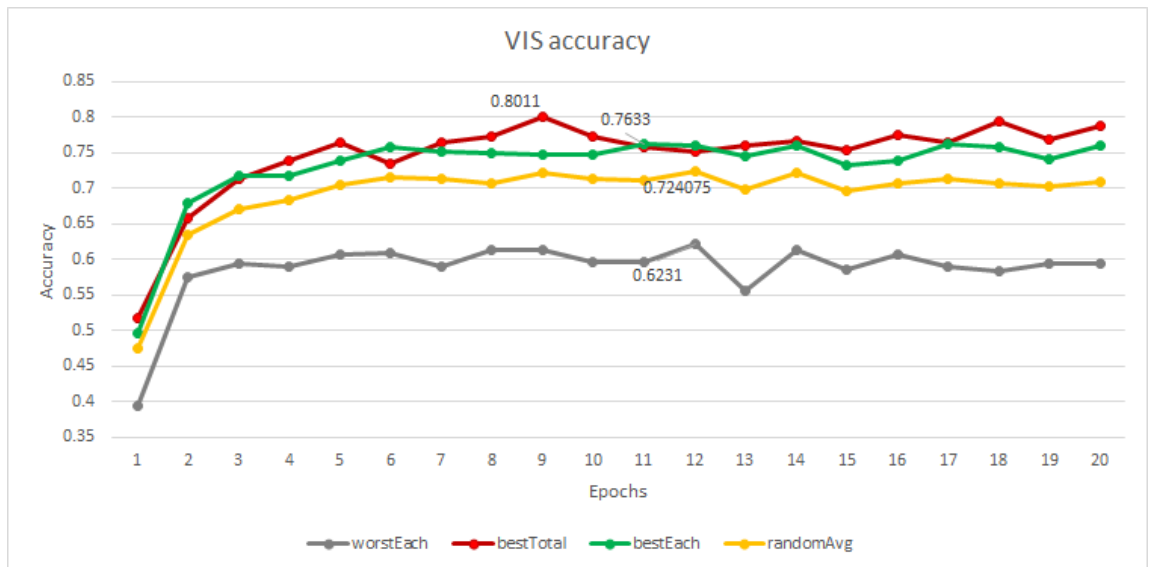


Figure 25. Diagram presenting the VIS camera three-banded sets' accuracy evolution as epochs progress. The peak of each line is depicted with a decimal number. Line explanation: "randomAvg,, stands for the average of the random selection, "bestTotal,, stands for the second technique ("best in total,,), while "bestEach,, and "worstEach,, stand for the third and fourth method applied, respectively.

The proposed research demonstrates that CNNs can be used for the accurate and efficient classification of wood waste. By providing a fast and accurate method of classification, CNNs have the potential to improve the efficiency and effectiveness of wood waste management,

leading to cost savings and reduced environmental impact. Through these findings, new possibilities emerge for leveraging CNNs to implement similar waste management practices across diverse waste industries such as Plastic, Metals, and Construction & Demolition Waste.

6.4. Contributions

Firstly, as a major contribution stands the generation of the image dataset of the wood and plastic goods. This dataset was created by capturing images with three different types of cameras (RGB, VIS, NIR). Then followed the data augmentation that included small changes in the images captured and which allowed us to end up with 6522 image samples for every type of camera in order to have safer final results.

Moreover, another contribution of this research is the testing of TensorFlow against PyTorch under the same conditions and the establishment of the opinion that TensorFlow is easier to use for entry level machine learning purposes, while PyTorch can prove to be more useful due to its complex settings, although more complicated.

Thirdly, our main contribution is the classification accuracy achieved. Into more detail, we accomplished a 100% accuracy into differentiating wood from plastic and a 93.45% into classifying every material of the seven classes by reversing the augmentation. This outcome was due to the implementation of the Average Classifier, which is a combination of the other three classifiers and we can safely conclude that the Average Classifier is the best option available since it produces the best results.

Furthermore, our research included many trials over different CNN models for image classification purposes and therefore the conclusion drawn is that Resnet50 can successfully handle a medium-sized dataset, provide accurate enough final outputs and collaborate efficiently with processing units that do not belong to the higher end.

Finally, an essential outcome pertains to the use of inadequate equipment. As a final aspect of our research, we conducted experiments on dimensionality reduction for multispectral cameras, serving as a demonstration of feature selection in situations where the equipment may not meet optimal standards, and assessed its implications. Remarkably, the results indicated a small marginal reduction of less than 2% for both types of cameras. This finding serves as compelling evidence that even with suboptimal equipment, satisfactory outcomes can still be achieved. Moreover, by selecting different sets of three bands from the multispectral cameras, we demonstrated that spatial and shape focus do not play a significantly dominant role in comparison to their spectral footprint. Supporting this assertion is the fact that the average accuracy of the sets of randomly selected bands differs by less than 8% from the highest accuracy achieved by sets of three bands and less than 10% from utilizing PCA and benefiting from all bands.

Chapter 7

Conclusions

7.1. Thesis Summary

In conclusion, this thesis explores the potential of using CNNs for the classification of wood waste. The proposed research demonstrates the feasibility and effectiveness of CNN usage and could help to establish the possibilities of CNNs in other industrial waste management areas. As a result, the proposed research has the potential to make a significant contribution to the field of wood waste management and to the development of sustainable wood-based industries, further helping circular economy.

Deep learning has shown enormous potential in many fields and especially CNNs in image processing. This research was based on the state-of-the-art image classification CNN model ResNet that was first introduced into the field of general waste sorting and used as a classifier for waste in a complex industrial environment. A feasible sorting method based on ResNet was proposed, and the whole flow was described in detail, including image generation, building of the training set and validation set, image augmentation, construction of the ResNet50 model, wood waste classification, wood from plastic waste classification and optimization of those results. The results indicate that the proposed method has the potential to facilitate the automatic classification of wood waste, which is important for many industries and municipalities to achieve better wooden products life cycle management, reduce environmental burdens and increase sustainability. Specifically, several conclusions were drawn as follows.

First, several model architectures were evaluated to determine the most suitable for our requirements. Consequently, it was discovered that ResNet50 demonstrated excellent performance in handling medium-sized datasets without demanding excessive amounts of computing resources. Second, we achieved a score of 93.45% Accuracy in segregating each of the seven classes and a perfect score of 100% in segregating wood from plastic.

Third, we conducted experiments on dimensionality reduction for the multispectral cameras, serving as an illustration of feature selection in scenarios where the equipment may be insufficient, and assessed its impact. The results showed a mere 1.69% reduction compared to utilizing all the bands in the visible spectrum (VIS), and a 2.07% reduction in the near-infrared (NIR) spectrum.

7.2. Future work

Wood waste classification and waste classification in general are still open fields that need to be discovered and they have a variety of open issues that need to be studied. Firstly, there is a need to expand the study to include hyperspectral cameras which would be beneficial in examining a wider range of wood categories. Hyperspectral imaging offers a unique opportunity to capture detailed spectral information from wood waste materials, enabling a more comprehensive understanding of their composition and characteristics. This expansion of technology can enhance the effectiveness of wood waste segregation processes, leading to better sorting accuracy of wood waste. Furthermore, the application of hyperspectral cameras can contribute to the development of advanced algorithms and machine learning techniques specifically tailored to wood waste analysis, allowing for automated and efficient segregation methods.

Improving the accuracy of wood waste classification is a crucial area for future research. While significant progress has been made in this study, there is still room for enhancing the precision and reliability of classification algorithms. Researchers can explore advanced machine learning techniques, or ensemble models to achieve higher accuracy rates. This could happen by addressing issues such as inter-class variability and subtle spectral differences.

The impact of light on multispectral imaging is also an important aspect that needs to be explored further. Understanding how different lighting conditions affect the spectral response of wood waste materials and investigating techniques to compensate for lighting variations will lead to more consistent results. Researchers should consider factors such as lighting angles, intensity, and the influence of environmental conditions to develop robust algorithms that can handle varying lighting situations.

Researching the spectral footprint instead of relying solely on the shape of objects presents an interesting avenue for future study in wood waste classification. By analyzing the unique spectral characteristics of different wood waste materials, researchers can gain valuable insights into their composition and properties and unlock new possibilities for efficient wood waste sorting. This shift in focus may require the development of innovative data processing techniques that can effectively extract and analyze the spectral footprints of wood waste.

Furthermore, future studies should consider the utilization of better processing units, such as high-performance GPUs and increased RAM capacity, to enhance the speed and efficiency of wood waste segregation. These advanced computing resources can significantly accelerate the processing of large hyperspectral datasets, enabling real-time or near-real-time classification capabilities.

Additionally, incorporating different kinds of wood waste into the dataset and ensuring its size and diversity are representative of real-world scenarios will further enhance the accuracy and generalizability of wood waste separation models. Our research primarily focuses on a limited set of wood types, while exploring additional categories, like large wood pieces that do not wholly fit in the camera or sawdust that is tiny, will provide a more comprehensive understanding of the challenges and requirements specific to each type.

Moreover, this research paves the way to explore the potential of utilizing Convolutional Neural Networks (CNNs) in various other industrial waste management domains beyond wood

waste. By showcasing the effectiveness of CNNs in wood waste classification, it opens up opportunities for leveraging similar machine learning techniques to address challenges in areas of distinct waste materials like plastic, metals, and construction & demolition waste. This broadening scope has the potential to revolutionize waste management practices across multiple industries, enabling more efficient recycling, and resource recovery processes.

Despite the progress made, a big challenge remains in implementing wood waste sorting systems in real-world conditions, beyond the confines of protected laboratory environments. Factors such as varying lighting, uncontrolled and dusty environmental conditions, and the need to adapt algorithms and models to handle unpredictable scenarios pose significant challenges which do not appear in a simulation. Future research should focus on addressing these practical obstacles to ensure the successful deployment and widespread adoption of wood waste separation technologies in real-world settings.

Finally, it is crucial that datasets pertaining to wood waste become openly accessible to everyone, facilitating the advancement of scientific knowledge. By ensuring the availability and ease of access to these datasets, researchers can collaborate more effectively, exchange valuable insights, and collectively address the challenges associated with wood waste management. Open datasets not only encourage interdisciplinary collaborations but also foster innovation by enabling researchers to develop and evaluate novel algorithms, models and techniques, while the availability of these datasets promotes transparency, reproducibility, and accountability. By democratizing access to wood waste datasets, science can move further, paving the way for sustainable waste management practices and informed decision-making that benefit both the environment and society as a whole.

Bibliography

- [1] “Trends in Solid Waste Management.” https://datatopics.worldbank.org/what-a-waste/trends_in_solid_waste_management.html (accessed May 12, 2023).
- [2] L. Giusti, “A review of waste management practices and their impact on human health,” *Waste Management*, vol. 29, no. 8, pp. 2227–2239, Aug. 2009, doi: 10.1016/j.wasman.2009.03.028.
- [3] “Waste Management Practices: Literature Review,” 2011.
- [4] “Solid Waste Management.” <https://www.worldbank.org/en/topic/urbandevelopment/brief/solid-waste-management> (accessed May 12, 2023).
- [5] Matthew. Gandy, *Recycling and the politics of urban waste*. 1994.
- [6] S. Pandey, “Wood waste utilization and associated product development from under-utilized low-quality wood and its prospects in Nepal,” *SN Appl Sci*, vol. 4, no. 6, p. 168, Jun. 2022, doi: 10.1007/s42452-022-05061-5.
- [7] I. Jahan, G. Zhang, M. Bhuiyan, and S. Navaratnam, “Circular Economy of Construction and Demolition Wood Waste—A Theoretical Framework Approach,” *Sustainability*, vol. 14, no. 17, p. 10478, Aug. 2022, doi: 10.3390/su141710478.
- [8] “WRA-Grades-of-Waste-Wood”.
- [9] W. P. (William P. Spence, “The home carpenter & woodworker’s repair manual,” p. 540, 2006.
- [10] “A PROPOSAL FOR THE DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE.” <https://web.archive.org/web/20070826230310/http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html> (accessed May 12, 2023).
- [11] “School of Informatics: History of Artificial Intelligence at Edinburgh.” <https://www.inf.ed.ac.uk/about/AIhistory.html> (accessed May 12, 2023).
- [12] “Why 2015 Was a Breakthrough Year in Artificial Intelligence - Bloomberg.” <https://www.bloomberg.com/news/articles/2015-12-08/why-2015-was-a-breakthrough-year-in-artificial-intelligence?leadSource=verify%20wall> (accessed May 12, 2023).
- [13] “Ask the AI experts: What’s driving today’s progress in AI? | McKinsey.” <https://www.mckinsey.com/capabilities/quantumblack/our-insights/ask-the-ai-experts-whats-driving-todays-progress-in-ai> (accessed May 12, 2023).
- [14] “What is artificial intelligence (AI)? - AI definition and how it works.” <https://www.techtarget.com/searchenterpriseai/definition/AI-Artificial-Intelligence> (accessed May 12, 2023).
- [15] “AI-Powered Ride-Hailing Apps can Help you in your Taxi Business.” <https://techcycled.com/apps-for-taxi-business/> (accessed May 12, 2023).
- [16] “What are Convolutional Neural Networks? | IBM.” <https://www.ibm.com/topics/convolutional-neural-networks> (accessed May 12, 2023).

- [17] “A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way | by Sumit Saha | Towards Data Science.” <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53> (accessed May 12, 2023).
- [18] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” Dec. 2015.
- [19] M. Talo, “Convolutional Neural Networks for Multi-class Histopathology Image Classification.”
- [20] “What is Computer Vision? | IBM.” <https://www.ibm.com/topics/computer-vision> (accessed May 12, 2023).
- [21] “7 Amazing Examples Of Computer And Machine Vision In Practice.” <https://www.forbes.com/sites/bernardmarr/2019/04/08/7-amazing-examples-of-computer-and-machine-vision-in-practice/#3dbb3f751018> (accessed May 12, 2023).
- [22] “What Is The 80/20 Rule In Machine Learning?” <https://www.techmediatoday.com/what-is-the-80-20-rule-in-machine-learning/> (accessed May 12, 2023).
- [23] Y. Ku, J. Yang, H. Fang, W. Xiao, and J. Zhuang, “Deep learning of grasping detection for a robot used in sorting construction and demolition waste,” *J Mater Cycles Waste Manag*, vol. 23, no. 1, pp. 84–95, Jan. 2021, doi: 10.1007/s10163-020-01098-z.
- [24] C. Bircanoglu, M. Atay, F. Beser, O. Genc, and M. A. Kizrak, “RecycleNet: Intelligent Waste Sorting Using Deep Neural Networks,” in *2018 Innovations in Intelligent Systems and Applications (INISTA)*, IEEE, Jul. 2018, pp. 1–7. doi: 10.1109/INISTA.2018.8466276.
- [25] O. Adedeji and Z. Wang, “Intelligent Waste Classification System Using Deep Learning Convolutional Neural Network,” *Procedia Manuf*, vol. 35, pp. 607–612, 2019, doi: 10.1016/j.promfg.2019.05.086.
- [26] J. Tuomas, T. Lukka, J. V Tossavainen, D. Kujala, and T. Raiko, “ZenRobotics Recycler-Robotic Sorting using Machine Learning,” 2014.
- [27] “alexveremis/WoodWasteThesis: Wood Waste Thesis | Alexandros Veremis under Prof. P. Tsanakas. All Rights Reserved 2023 ©.” <https://github.com/alexveremis/WoodWasteThesis> (accessed Jul. 23, 2023).
- [28] P. Singh, N. Singh, K. K. Singh, and A. Singh, “Diagnosing of disease using machine learning,” in *Machine Learning and the Internet of Medical Things in Healthcare*, Elsevier, 2021, pp. 89–111. doi: 10.1016/B978-0-12-821229-5.00003-3.
- [29] “Classification: Accuracy | Machine Learning | Google Developers” <https://developers.google.com/machine-learning/crash-course/classification/accuracy> (accessed May 12, 2023).
- [30] “Loss Functions and Their Use In Neural Networks | by Vishal Yathish | Towards Data Science.” <https://towardsdatascience.com/loss-functions-and-their-use-in-neural-networks-a470e703f1e9> (accessed May 12, 2023).
- [31] “Waste Management – Biological Reprocessing,Biological Reprocessing,Waste Management Resources.” <https://www.wastemanagement.in/biological-reprocessing.html> (accessed May 12, 2023).
- [32] “Waste Hierarchy | New Energy Corporation.” <https://web.archive.org/web/20140516094044/http://www.newenergycorp.com.au/what-we-do/waste-hierarchy> (accessed May 12, 2023).

- [33] D. Czajczyńska *et al.*, “Potential of pyrolysis processes in the waste management sector,” *Thermal Science and Engineering Progress*, vol. 3, pp. 171–197, Sep. 2017, doi: 10.1016/j.tsep.2017.06.003.
- [34] *Biomass Gasification, Pyrolysis and Torrefaction*. Elsevier, 2013. doi: 10.1016/C2011-0-07564-6.
- [35] D. Chen, L. Yin, H. Wang, and P. He, “Pyrolysis technologies for municipal solid waste: A review,” *Waste Management*, vol. 34, no. 12, pp. 2466–2486, Dec. 2014, doi: 10.1016/j.wasman.2014.08.004.
- [36] M. Hyman, B. Turner, A. Carpintero, United Nations Institute for Training and Research, Inter-Organization Programme for the Sound Management of Chemicals, and United Nations Environment Programme, *Guidelines for national waste management strategies : moving from challenges to opportunities*.
- [37] M. Zbieć, J. Franc-Dąbrowska, and N. Drejerska, “Wood Waste Management in Europe through the Lens of the Circular Bioeconomy,” *Energies (Basel)*, vol. 15, no. 12, p. 4352, Jun. 2022, doi: 10.3390/en15124352.
- [38] M. Malik *et al.*, “Waste Classification for Sustainable Development Using Image Recognition with Deep Learning Neural Network Models,” *Sustainability*, vol. 14, no. 12, p. 7222, Jun. 2022, doi: 10.3390/su14127222.
- [39] “What are Hyperspectral Images?. How are they different from the digital... | by Sarthak Panigrahi | Towards Data Science.” <https://towardsdatascience.com/what-are-hyperspectral-images-a5de5d9fa91> (accessed May 12, 2023).
- [40] “Color is in the eye of the beholder | Department of Organismic and Evolutionary Biology.” <https://oeb.harvard.edu/news/color-eye-beholder> (accessed May 12, 2023).
- [41] “What is hyperspectral Imaging?: A Comprehensive Guide - Specim Spectral Imaging.” <https://www.specim.com/technology/what-is-hyperspectral-imaging/> (accessed May 12, 2023).
- [42] W. Lu and J. Chen, “Computer vision for solid waste sorting: A critical review of academic research,” *Waste Management*, vol. 142, pp. 29–43, Apr. 2022, doi: 10.1016/j.wasman.2022.02.009.
- [43] I. Grigorev *et al.*, “Improving the quality of sorting wood chips by scanning and machine vision technology ,” *J For Sci (Prague)*, vol. 67, no. No. 5, pp. 212–218, May 2021, doi: 10.17221/10/2020-JFS.
- [44] M. Verheyen, W. Beckers, E. Claesen, G. Moonen, and E. Demeester, “Vision-based sorting of medium density fibreboard and grade A wood waste,” in *2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA)*, IEEE, Sep. 2016, pp. 1–6. doi: 10.1109/ETFA.2016.7733546.
- [45] A. Besserer, S. Troilo, P. Girods, Y. Rogauame, and N. Brosse, “Cascading Recycling of Wood Waste: A Review,” *Polymers (Basel)*, vol. 13, no. 11, p. 1752, May 2021, doi: 10.3390/polym13111752.
- [46] M. Schwanninger, J. C. Rodrigues, and K. Fackler, “A Review of Band Assignments in near Infrared Spectra of Wood and Wood Components,” *J Near Infrared Spectrosc*, vol. 19, no. 5, pp. 287–308, Oct. 2011, doi: 10.1255/jnirs.955.

- [47] L. Donaldson and N. Williams, "Imaging and Spectroscopy of Natural Fluorophores in Pine Needles," *Plants*, vol. 7, no. 1, p. 10, Feb. 2018, doi: 10.3390/plants7010010.
- [48] L. Donaldson, "Autofluorescence in Plants," *Molecules*, vol. 25, no. 10, p. 2393, May 2020, doi: 10.3390/molecules25102393.
- [49] "What is Machine Learning? | IBM." <https://www.ibm.com/topics/machine-learning> (accessed May 12, 2023).
- [50] "TensorFlow." <https://www.tensorflow.org/> (accessed May 12, 2023).
- [51] "PyTorch." <https://pytorch.org/> (accessed May 12, 2023).
- [52] "What Is Machine Learning (ML)? - I School Online." <https://ischoolonline.berkeley.edu/blog/what-is-machine-learning/> (accessed May 12, 2023).
- [53] "Difference Between Supervised, Unsupervised, & Reinforcement Learning | NVIDIA Blog." <https://blogs.nvidia.com/blog/2018/08/02/supervised-unsupervised-learning/> (accessed May 12, 2023).
- [54] "What is Deep Learning? | IBM." <https://www.ibm.com/topics/deep-learning> (accessed May 12, 2023).
- [55] "What is Epoch in Machine Learning? | Simplilearn." <https://www.simplilearn.com/tutorials/machine-learning-tutorial/what-is-epoch-in-machine-learning> (accessed May 12, 2023).
- [56] "What are Neural Networks? | IBM." <https://www.ibm.com/topics/neural-networks> (accessed May 12, 2023).
- [57] "Principal Component Analysis (PCA) Explained | Built In." <https://builtin.com/data-science/step-step-explanation-principal-component-analysis> (accessed May 12, 2023).
- [58] "PCA on HyperSpectral Data. A Beginner friendly tutorial on... | by Richa Dutt | Towards Data Science." <https://towardsdatascience.com/pca-on-hyperspectral-data-99c9c5178385> (accessed May 12, 2023).
- [59] K. Koonsanit and C. Jaruskulchai, "Band Selection for Hyperspectral Image Using Principal Components Analysis and Maxima-Minima Functional," 2011, pp. 103–112. doi: 10.1007/978-3-642-24788-0_10.