



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ – ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ**

**Πρόγραμμα Μεταπτυχιακών Σπουδών**

**«Κυβερνοασφάλεια και Επιστήμη των Δεδομένων»**

**Μεταπτυχιακή Διατριβή**

Τίτλος Διατριβής	<b>Σύστημα πρόβλεψης για την ταξινόμηση διαφορετικών ειδών και ποικιλιών σπόρων δημητριακών</b> <b>Prediction system for the classification of different species and varieties of cereal seeds</b>
Όνοματεπώνυμο Φοιτητή	<b>Λεκαράκος Χρήστος</b>
Πατρώνυμο	<b>Φίλιππος</b>
Αριθμός Μητρώου	<b>ΜΠΚΕΔ21027</b>
Επιβλέπων	<b>Δημήτριος Αποστόλου, Καθηγητής</b>

Ημερομηνία Παράδοσης **Φεβρουάριος 2025**

---

**Τριμελής Εξεταστική Επιτροπή**

Δημήτριος Αποστόλου  
Καθηγητής

Διονύσιος Σωτηρόπουλος  
Επίκουρος Καθηγητής

Γρηγόριος Κορωνάκος  
Διδάσκων ΠΜΣ

## **Ευχαριστίες**

Θα ήθελα να ευχαριστήσω τον επιβλέπων καθηγητή Δημήτριο Αποστόλου για την συνεργασία μας και την ευκαιρία που μου έδωσε, να ασχοληθώ με ένα τόσο ενδιαφέρον θέμα.

Την Σίσσυ Ευθυμιάδου και τον Χριστόφορο Κασιμάτη για την πολύτιμη βοήθεια τους στην εκπόνηση αυτής της μεταπτυχιακής διατριβής. Η γνώση που μου μετέφεραν για τον γεωργικό τομέα, η συνεισφορά τους σε τεχνικό και υλικό επίπεδο ήταν κρίσιμη για την πραγματοποίηση του έργου αυτού. Τον οργανισμό ΕΛΓΟ-Δήμητρα για την παραχώρηση των σπόρων προς φωτογράφιση.

Φυσικά, δεν θα μπορούσα να παραλείψω να ευχαριστήσω την οικογένεια μου για την υποστήριξη και την βοήθεια καθ' όλη την διάρκεια της πορείας μου, αλλά και τους φίλους μου για την διαρκή τους ενθάρρυνση.

## Περίληψη

Για χιλιάδες χρονιά τα δημητριακά αποτελούν μια σταθερή πηγή τροφής για τους ανθρώπους και τα ζώα. Η ποιότητα των προϊόντων δημητριακών είναι άρρηκτα συνδεδεμένη με την ποιότητα των σπορών τους καθώς υψηλής ποιότητας σπόροι είναι πιο πιθανό να δημιουργήσουν υγιή φυτά με υψηλές αποδόσεις. Επί του παρόντος η αναγνώριση σπόρων είναι μια πολύπλοκη και χρονοβόρα διαδικασία η οποία διεξάγεται μέσω χειρωνακτικού ελέγχου. Ωστόσο με την βοήθεια των νευρωνικών δικτύων η διαδικασία της ταξινόμησης μπορεί να αυτοματοποιηθεί οδηγώντας έτσι σε ταχύτερα και πιο ακριβή αποτελέσματα. Στην εργασία αυτή προτείνεται ένα μοντέλο πρόβλεψης βαθιάς μάθησης ικανό να ταξινομήσει σπόρους δημητριακών από διαφορετικά είδη και ποικιλίες (Κριθάρι Τριπτόλεμος, Δήμητρα, βρόμη Φλεγα, βρόμη Πήγασος, σκληρό σιτάρι Σέλλας και σιτάρι Λεβάντε). Για την ανάπτυξη του μοντέλου πρόβλεψης δοκιμάστηκαν τρεις παραλλαγές από κοινές αρχιτεκτονικές, συνελικτικών νευρωνικών δικτύων (CNN). Για την αξιολόγηση της αποτελεσματικότητας των προτεινόμενων μοντέλων, χρησιμοποιήθηκε ένα σύνολο δεδομένων που περιλαμβάνει 12.720 μεμονωμένες εικόνες σπόρων, που καταγράφηκαν με χρήση ψηφιακής κάμερας. Τα αποτελέσματα έδειξαν ακρίβεια ταξινόμησης που κυμαίνονταν από 65% έως 97%. Οι υψηλότερες ακρίβειες δοκιμών επιτεύχθηκαν χρησιμοποιώντας τις παραλλαγές των αρχιτεκτονικών ResNet50 και Xception που ανέρχονται στο 97%. Κατά συνέπεια, τα αποτελέσματα της προτεινόμενης προσέγγισης είναι ακριβή και αξιόπιστα, ενθαρρύνοντας την πρακτική εφαρμογή της.

**Λέξεις κλειδιά:** Συνελικτικά Νευρωνικά Δίκτυα, Βαθιά μάθηση, Μεταφορά γνώσης, Ταξινόμηση Σπόρων, μοντέλο VGG16, μοντέλο ResNet50, μοντέλο Xception

## Abstract

Cereals play a crucial role in food supply globally. The quality of cereal seeds is of major importance within the grain chain industry, particularly in seed testing and certification procedures. Currently, this recognition and verification process relies on manual visual analysis of grains. However, an automated approach leveraging computer vision and machine learning classification has emerged as a faster and more efficient alternative method. Nevertheless, classifying cereals from different species accurately remains a complex and demanding task, especially at the varietal level. This study introduces a deep learning-based method designed to achieve precise classification for cereals of different species and varieties. Specifically, Convolutional Neural Networks (CNNs) were employed to classify wheat grain images into six different varieties (Krithari Triptolemos, Dimitra, Vromi Flega, Vromi Pigasos, Skliro Sitari Sellas and Levante). Additionally, three variations of commonly used CNN architectures were employed, namely VGG16, ResNet50 and Xception. To evaluate the effectiveness of the proposed models, a dataset comprising of 12.720 individual grain seed images, captured using a digital camera, was utilized. The results demonstrated test accuracies ranging from 65% to 97% for classification. The highest test accuracies were achieved using the ResNet50 and Xception architectures standing at 97%. Consequently, the outcomes of the proposed approach are both accurate and dependable, encouraging its practical implementation.

**Keywords:** Convolutional Neural Networks, Deep Learning, Transfer Learning, Seed classification, VGG16 model, ResNet50 model, Xception model

**Πίνακας περιεχομένων**

Περίληψη.....	4
1. Εισαγωγή.....	9
1.1 Ερευνητικοί Στόχοι .....	10
2. Θεωρητικό Υπόβαθρο .....	11
2.1 Συνελκτικά Νευρωνικά Δίκτυα .....	12
2.2 Εκμάθηση με μεταφορά γνώσης .....	13
3. Βιβλιογραφική Ανασκόπηση του πεδίου .....	15
4. Υλικά και μέθοδοι .....	20
4.1 Το σύνολο δεδομένων.....	20
4.2 Κανονικοποίηση των δεδομένων .....	23
4.3 Τα μοντέλα που δοκιμάστηκαν .....	25
4.3.1 Το μοντέλο VGG16 .....	25
4.3.2 Το μοντέλο ResNet50 .....	26
4.3.3 Το μοντέλο Xception .....	27
4.4 Εκπαίδευση των νευρωνικών δικτύων .....	29
4.5 Μετρικές αξιολόγησης.....	33
5. Αποτελέσματα .....	34
5.1 Ακρίβεια και απώλεια .....	35
5.2 Η απόδοση των μοντέλων στην ταξινόμηση.....	39
5.3 Σύγκριση μετρικών.....	41
6. Συζήτηση .....	42
7. Μελλοντική Έρευνα .....	45
8. Συμπέρασμα.....	47
Βιβλιογραφία .....	50

**Κατάλογος Πινάκων**

Πίνακας 1: Κατανομή του συνόλου δεδομένων ανά είδος και ποικιλία. ....	22
Πίνακας 2: Γενικές πληροφορίες για τις παραλλαγές των τριών μοντέλων που δοκιμάστηκαν .....	33
Πίνακας 3: Συγκεντρωτικά αποτελέσματα της ακρίβειας των μοντέλων για κάθε ποικιλία στο σύνολο ελέγχου.....	34
Πίνακας 4: Το classification report του μοντέλου VGG16.....	40
Πίνακας 5: Το classification report του μοντέλου ResNet50.....	41
Πίνακας 6: Το classification report του μοντέλου Xception.....	42

## Κατάλογος Εικόνων

Εικόνα 1: Η απεικόνιση ενός τεχνητού νευρωνικού δικτύου.....	11
Εικόνα 2: Απεικόνιση της αρχιτεκτονικής ενός απλου συνελκτικού νευρωνικού δικτύου .....	13
Εικόνα 3: Σχηματική απεικόνιση της διαδικασίας μάθησης με μεταφορά γνώσης. ....	15
Εικόνα 4: Η διαδικασία επισήμανσης των εικόνων του συνόλου δεδομένων.....	21
Εικόνα 5: Δοκιμαστικές λήψεις σπόρων .....	23
Εικόνα 6: Οι κατανομές πλάτους και ύψους των εικόνων του συνόλου δεδομένων σε ρixel .....	25
Εικόνα 7: Η βασική αρχιτεκτονική του μοντέλου VGG16. ....	26
Εικόνα 8: Η βασική αρχιτεκτονική του μοντέλου ResNet50.....	27
Εικόνα 9: Η αρχιτεκτονική του μοντέλου Xception. ....	28
Εικόνα 10: Η παραλλαγή του μοντέλου VGG16. ....	29
Εικόνα 11: Η παραλλαγή του μοντέλου ResNet50. ....	30
Εικόνα 12: Η παραλλαγή του μοντέλου Xception. ....	31
Εικόνα 13: Μαθηματικοί τύποι των μετρικών αξιολόγησης .....	34
Εικόνα 14: Διάγραμμα ροής της διαδικασίας ανάπτυξης της προτεινόμενης εφαρμογής 35	
Εικόνα 15: Η μέγιστη και ελάχιστη ακρίβεια κατά την διάρκεια της εκπαίδευσης του μοντέλου VGG16.....	36
Εικόνα 16: Η μέγιστη και ελάχιστη ακρίβεια κατά την διάρκεια της εκπαίδευσης του μοντέλου Xception. ....	37
Εικόνα 17: Η μέγιστη και ελάχιστη ακρίβεια κατά την διάρκεια της εκπαίδευσης του μοντέλου ResNet50. ....	37
Εικόνα 18: Μέγιστη και ελάχιστη απώλεια στο σύνολο επικύρωσης για το μοντέλο ResNet50. ....	38
Εικόνα 19: Μέγιστη και ελάχιστη απώλεια στο σύνολο επικύρωσης για το μοντέλο Xception. ....	39
Εικόνα 20: Μέγιστη και ελάχιστη απώλεια στο σύνολο επικύρωσης για το μοντέλο VGG16.....	39

**Συντομογραφίες και ακρωνύμια**

CNN(s)	Convolutional Neural Network(s)
ANN	Artificial Neural Network
ReLU	Rectified Linear Unit
NLP	Natural Language Processing
VGG	Visual Geometry Group
ResNet	Residual Network
SVM	Support Vector Machines
SSD	Single Shot Detector
ML	Machine Learning
MLP	Multilayer Perceptron
JPG	Joint Photographic Experts Group
KTR	Krithari Triptolemos
DMT	Krithari Dimitra
SSS	Skliri Sitari Sellas
LVT	Sitari Levante
VRF	Vromi Flega
VPG	Vromi Pigasos
ILSVRC	ImageNet Large Scale Visual Recognition Challenge
CFIA	Canadian Food Inspection Agency
ISTA	International Seed Testing Association
AOSA	Association of Official Seed Analysts



## 1. Εισαγωγή

Τα δημητριακά αποτελούν τα πιο μαζικά παραγόμενα και υψηλής κατανάλωσης προϊόντα καλλιέργειας παγκοσμίως [1]. Για χιλιάδες χρονιά οι άνθρωποι στηρίζονται στην καλλιέργεια δημητριακών ως πηγή τροφής για τους ίδιους αλλά και για την κτηνοτροφία. Είναι ζωτικής σημασίας για την ανθρώπινη επιβίωση καθώς αποτελούν κύριο συστατικό σε βασικά τρόφιμα όπως το ψωμί και τα ζυμαρικά τα οποία συμβάλλουν πάνω από 50% στην ημερήσια θερμιδική πρόσληψη [2]. Η ποιότητα των δημητριακών μπορεί να ποικίλλει ανάλογα με το είδος και την ποικιλία που ανήκουν καθορίζοντας έτσι την ποιότητα των προϊόντων που παράγονται από την αντίστοιχη ποικιλία δημητριακών [3]. Για αυτό ο εντοπισμός και η επαλήθευση των διαφορετικών ειδών και ποικιλιών των δημητριακών είναι μεγάλης σημασίας για την εκτίμηση κόστους, πρωτεΐνης και θερμιδικής αξίας, παράγοντες που καθορίζουν σε μεγάλο βαθμό την ποιότητα του σπόρου [4]. Επομένως, η ταξινόμηση των σπορών ανάλογα με το είδος και την ποικιλία αποτελεί ένα πρόβλημα που χρειάζεται την παρέμβαση ειδικών του γεωργικού τομέα, ένα γεγονός που δεν αποκλείει το ανθρώπινο λάθος. Η μηχανική μάθηση έχει ευρέως χρησιμοποιηθεί για προβλήματα ταξινόμησης σε διάφορους κλάδους με εντυπωσιακά αποτελέσματα και θα μπορούσε να αποτελέσει μια λύση για την ταξινόμηση σπορών δημητριακών που ανήκουν σε διαφορετικά είδη και ποικιλίες [5].

Η εξέλιξη των ψηφιακών τεχνολογιών δημιουργεί το πλαίσιο για ολοκληρωμένες εφαρμογές της μηχανικής μάθησης στον γεωργικό τομέα. Σε συνδυασμό με την άνθιση του πεδίου των μεγάλων δεδομένων, δημιουργούνται νέες προοπτικές μετασχηματίζοντας την γεωργική βιομηχανία [6]. Μεγάλο ρολό έχει επίσης διαδραματίσει η ευρεία χρήση ψηφιακών μέσων όπως αισθητήρων, drones και τεχνολογίας GPS. Η χρήση ψηφιακών μέσων έχει επιτρέψει την συλλογή μεγάλων ποσοτήτων δεδομένων και κατ' επέκταση όλο και περισσότερες εφαρμογές από τα πεδία της μηχανικής μάθησης και των μεγάλων δεδομένων εφαρμόζονται στον γεωργικό τομέα [7]. Μοντέλα μηχανικής μάθησης ενσωματωμένα σε ολοκληρωμένα συστήματα υποστήριξης αποφάσεων έχουν χρησιμοποιηθεί για παράδειγμα στην παρακολούθηση και την ακριβή διαχείριση της ανάπτυξης των καλλιεργειών [8], των συνθηκών του εδάφους και της χρήσης του νερού [9].

Η τεχνολογία ανάλυσης εικόνας έχει τη δυνατότητα να προσφέρει ουσιαστικά στον αγροτικό τομέα, δίνοντας τη δυνατότητα στους αγρότες να παρακολουθούν και να διαχειρίζονται τις καλλιέργειες με μεγαλύτερη ακρίβεια και αποτελεσματικότητα. Με την ανάλυση εικόνων που λαμβάνονται από κάμερες, drones ή άλλες συσκευές, η μηχανική όραση μπορεί, για παράδειγμα να ανιχνεύσει πρώιμα σημάδια φυτικών ασθενειών, προσβολές από παράσιτα και ελλείψεις θρεπτικών συστατικών, επιτρέποντας στους αγρότες να αναλάβουν γρήγορα μέτρα και να αποτρέψουν τις απώλειες των καλλιεργειών [10]. Επιπλέον, η μηχανική όραση μπορεί να αυτοματοποιήσει πολλές εργασίες, όπως η αναγνώριση και η ταξινόμηση των προϊόντων, η μείωση του κόστους εργασίας και η αύξηση της αποδοτικότητας.

Ένα κοινό πρόβλημα που αντιμετωπίζουν οι αγρότες είναι η αναγνώριση και η ταξινόμηση σπορών δημητριακών ανάλογα με το είδος και την ποικιλία τους. Μερικοί σπόροι δημητριακών, παρόλο που ανήκουν σε διαφορετικά είδη ή ποικιλίες, μπορεί να έχουν παρόμοια σχήματα, μεγέθη ή χρώματα, γεγονός που καθιστά δύσκολη την ακριβή αναγνώριση και ταξινόμησή τους, ακόμα και για τους ειδικούς του γεωργικού τομέα [11]. Επιπλέον, οι σπόροι μπορεί να διαφέρουν σε σημεία, όπως η υφή της επιφάνειας ή η εσωτερική δομή, που δεν είναι εύκολα ορατά στο ανθρώπινο μάτι. Οι παραδοσιακές μέθοδοι αναγνώρισης σπόρων δημητριακών βασίζονται στον χειρωνακτικό έλεγχο, ο οποίος είναι χρονοβόρος και συχνά υποκειμενικός, οδηγώντας σε σφάλματα και αντιφάσεις. Σε αυτές τις περιπτώσεις, η τεχνολογία μηχανικής όρασης μπορεί να είναι

ιδιαίτερα χρήσιμη για την ταξινόμηση των σπόρων, καθώς μπορεί να αναλύσει εικόνες των σπόρων και να εντοπίσει ακόμη και μικρές διαφορές που μπορεί να είναι ανεπαίσθητες για τον άνθρωπο. Οι αλγόριθμοι μηχανικής όρασης μπορούν επίσης να εκπαιδευτούν, ώστε να αναγνωρίζουν διαφορετικές ποικιλίες σπόρων με βάση χαρακτηριστικά όπως το μέγεθος, το σχήμα, την υφή και το χρώμα, παρέχοντας ακριβείς και αξιόπιστες δυνατότητες ταξινόμησης και αναγνώρισης [12].

Στη γεωργία, ο ποιοτικός έλεγχος των σπορών είναι μια διαδικασία ζωτικής σημασίας, καθώς εμπεριέχει πληροφορίες απαραίτητες για την επιλογή καλλιέργειας όπως για παράδειγμα η καθαρότητα, οι συνθήκες βλάστησης ή η ύπαρξη επιβλαβών ζιζανίων. Η επιλογή της καλλιέργειας εξαρτάται άμεσα από το περιβάλλον στο οποίο αυτή πραγματοποιείται, με τους αγρότες να επιλέγουν είδη δημητριακών που είναι κατάλληλα για τις κλιματικές και εδαφολογικές συνθήκες της περιοχής τους. Κάθε είδος αναπτύσσεται σε διαφορετικές κλιματικές και εδαφολογικές συνθήκες, με αποτέλεσμα να απαιτούνται διαφορετικές ποσότητες νερού ή λιπάσματος για την ανάπτυξη του φυτού. Επομένως, η ακριβής αναγνώριση και ταξινόμηση των σπόρων επιτρέπει στους αγρότες να επιλέγουν υψηλής ποιότητας σπόρους από τα είδη τα οποία είναι καταλληλά για το περιβάλλον παραγωγής τους. Επιπλέον, μέσω αυτής της διαδικασίας ελέγχου και ταξινόμησης προλαμβάνονται διάφορα προβλήματα όπως το ενδεχόμενο πρόσμιξης διαφορετικών ποικιλιών, της εμφάνισης άρρωστων σπόρων ή ζιζανίων που μπορεί να δημιουργήσουν πρόβλημα στη σοδειά. Με τον εντοπισμό και την αφαίρεση κατεστραμμένων και άρρωστων σπόρων, οι αγρότες μπορούν να μειώσουν τον κίνδυνο αποτυχίας της καλλιέργειας, ενώ παράλληλα προστατεύουν το περιβάλλον, καθώς αν τα φυτά είναι υγιή μειώνεται η ανάγκη για χρήση λιπασμάτων ή εντομοκτόνων.

Η ταξινόμηση των σπόρων είναι μια χρονοβόρα και εντατική εργασία και η χρήση της τεχνολογίας μηχανικής όρασης μπορεί να αυτοματοποιήσει τη διαδικασία, μειώνοντας την ανάγκη για χειρωνακτική εργασία και βελτιώνοντας την αποτελεσματικότητα [13]. Με την αυτοματοποίηση της διαδικασίας ταξινόμησης σπόρων, οι αγρότες μπορούν να αισθάνονται πολύ πιο σίγουροι για τα είδη φυτών που επιλέγουν, να εξοικονομήσουν χρόνο και πόρους, ενώ ταυτόχρονα εξασφαλίζονται η ακρίβεια και η συνέπεια των διαδικασιών διαλογής. Καταληκτικά, η ταξινόμηση των σπόρων διαφορετικών ειδών δημητριακών ανάλογα με το είδος και την ποικιλία τους είναι ένα κρίσιμο στοιχείο για τις σύγχρονες πρακτικές καλλιέργειας, δίνοντας τη δυνατότητα στους αγρότες να βελτιστοποιήσουν τις στρατηγικές φύτευσης και τις αποδόσεις των καλλιεργειών τους, ενώ παράλληλα μειώνουν τα απόβλητα και αυξάνουν την αποδοτικότητα.

## 1.1 Ερευνητικοί Στόχοι

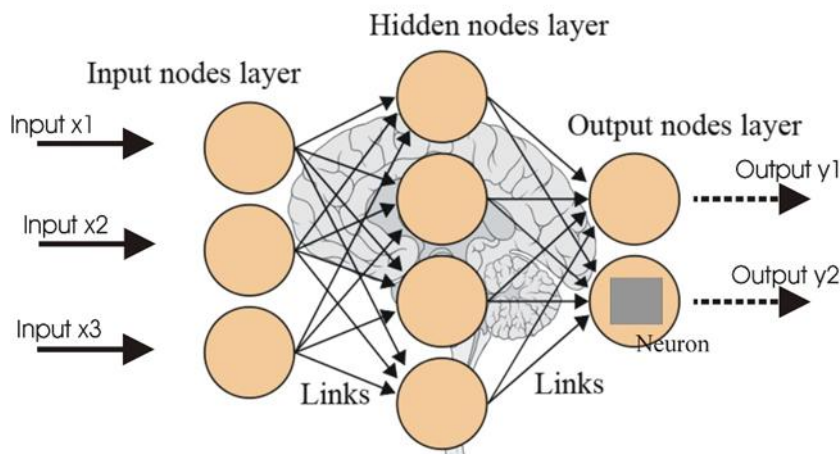
Η παρούσα διπλωματική εργασία έχει ως κύριο στόχο την ανάπτυξη ενός μοντέλου πρόβλεψης για τους αγρότες, που θα βασίζεται σε μοντέλα μηχανικής μάθησης, ικανό να ταξινομήσει σπόρους δημητριακών από τρία διαφορετικά είδη και έξι διαφορετικές ποικιλίες. Πιο συγκεκριμένα, θα δημιουργηθεί ένα σύνολο δεδομένων αποτελούμενο από εικόνες της κάθε ποικιλίας οι οποίες είναι: Κριθάρι Τριπτόλεμος, κριθάρι Δήμητρα, Σκληρό Σιτάρι Σέλλας, βρόμη Φλεγα, βρόμη Πήγασος και σιτάρι Λεβάντε. Για την επίτευξη του στόχου θα γίνει χρήση τεχνικών επεξεργασίας εικόνας σε συνδυασμό με αλγορίθμους βαθιάς μάθησης όπως τα συνελκτικά νευρωνικά δίκτυα. (Convolutional Neural Networks – CNNs). Επιπλέον, θα συγκριθούν τρία CNN τα οποία εκπαιδεύτηκαν με μεταφορά γνώσης και θα αξιολογηθούν για την ακρίβεια και την ταχύτητα τους στην ταξινόμηση των σπορών των δημητριακών. Τέλος, μέσω της αξιολόγησης της

αποτελεσματικότητας του συστήματος θα προσδιοριστεί η προσφορά του στην επίλυση ενός δύσκολου προβλήματος του γεωργικού τομέα.

## 2. Θεωρητικό Υπόβαθρο

Η μηχανική μάθηση ανήκει στο πεδίο της επιστήμης της πληροφορικής και η ανάπτυξη της οφείλεται στη συνεχή εξέλιξη του πεδίου της αναγνώρισης προτύπων [14]. Στηρίζεται σε μεγάλο βαθμό στη στατιστική καθώς ένα υπολογιστικό σύστημα χρειάζεται δεδομένα ώστε να εκπαιδευτεί και να καταφέρει να πραγματοποιήσει την εργασία που του ανατέθηκε. Οι πιο συχνές εργασίες που επιλύουν τα μοντέλα Μηχανικής μάθησης είναι η πρόβλεψη τιμών και η ταξινόμηση διακριτών τάξεων. Για παράδειγμα, ένα μοντέλο μπορεί να προβλέψει την τιμή μιας μετοχής ή να ταξινομήσει τις διαφορετικές ομάδες αίματος. Η εξέλιξη της τεχνολογίας έχει επιτρέψει την δημιουργία εξαιρετικά πολύπλοκων μοντέλων και την εφαρμογή διάφορων μηχανισμών του πραγματικού κόσμου [15].

Σήμερα υπάρχουν μοντέλα που μιμούνται την οπτική λειτουργία του εγκέφαλου, ειδικότερα τον τρόπο με τον οποίο οπτικές πληροφορίες επεξεργάζονται στον πρωτογενή οπτικό φλοιό όπως τα Συνελικτικά Νευρωνικά Δίκτυα (CNN). Τέτοιου είδους δίκτυα αποτελούν τη βάση της σύγχρονης τεχνολογίας για την αναγνώριση εικόνων, λόγω της ικανότητάς τους να ανιχνεύουν και να εξάγουν πολύπλοκα χαρακτηριστικά από οπτικά δεδομένα. Μέσω των συνελικτικών επιπέδων, τα CNNs εντοπίζουν χαρακτηριστικά όπως άκρα, υφές και σχήματα, τα οποία στη συνέχεια συνδυάζονται για την αναγνώριση αντικειμένων ή μοτίβων σε μια εικόνα. Αυτή η ιδιότητα τα καθιστά ιδιαίτερα αποτελεσματικά σε εφαρμογές όπως η αναγνώριση προσώπων, η κατηγοριοποίηση αντικειμένων και η ανάλυση δορυφορικών εικόνων, επιτυγχάνοντας εξαιρετικά αποτελέσματα ακόμα και σε πολύπλοκα σύνολα δεδομένων.



**Εικόνα 1: Η απεικόνιση ενός τεχνητού νευρωνικού δικτύου**  
<https://www.analyticsvidhya.com/blog/2016/08/evolution-core-concepts-deep-learning-neural-networks/>

## 2.1 Συνελικτικά Νευρωνικά Δίκτυα

Τα συνελικτικά νευρωνικά δίκτυα (CNN) είναι ένας τύπος τεχνικής βαθιάς μάθησης που χρησιμοποιείται συνήθως για εργασίες αναγνώρισης και ταξινόμησης εικόνων. Η βαθιά μάθηση είναι ένα υποσύνολο της μηχανικής μάθησης που περιλαμβάνει την εκπαίδευση τεχνητών νευρωνικών δικτύων με πολλαπλά επίπεδα για την εκμάθηση σύνθετων αναπαραστάσεων δεδομένων [16].

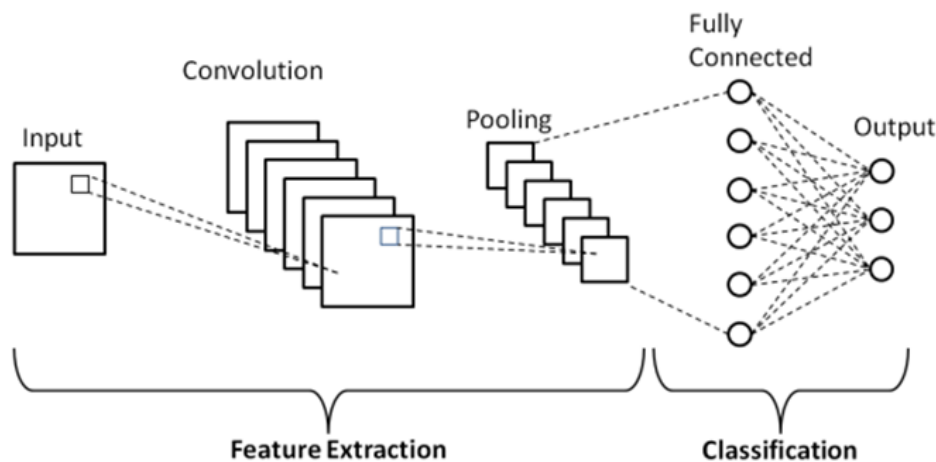
Τα CNN εμπνεύστηκαν από τη δομή του οπτικού φλοιού στον εγκέφαλο και έχουν σχεδιαστεί για να μιμούνται τον τρόπο με τον οποίο ο οπτικός φλοιός επεξεργάζεται τις οπτικές πληροφορίες. Η βασική αρχιτεκτονική ενός CNN αποτελείται από ένα επίπεδο εισόδου, μια σειρά από συνελικτικά επίπεδα, επίπεδα ενεργοποίησης, επίπεδα συγκέντρωσης και πλήρως συνδεδεμένα επίπεδα [17].

Το επίπεδο εισόδου ενός CNN είναι το σημείο όπου η εικόνα εισόδου τροφοδοτείται στο δίκτυο. Η εικόνα εισόδου είναι συνήθως μια μήτρα τιμών  $pxq$  που αντιπροσωπεύει την εικόνα. Τα συνελικτικά στρώματα ενός CNN είναι υπεύθυνα για την εξαγωγή χαρακτηριστικών από την εικόνα εισόδου. Κάθε συνελικτικό επίπεδο αποτελείται από ένα σύνολο φίλτρων (ή πυρήνων) που συνενώνονται με την εικόνα εισόδου για την παραγωγή ενός συνόλου χαρτών χαρακτηριστικών. Αυτοί οι χάρτες χαρακτηριστικών αντιπροσωπεύουν την παρουσία ή την απουσία ορισμένων χαρακτηριστικών (όπως ακμές ή υφές) στην εικόνα εισόδου.

Τα επίπεδα ενεργοποίησης ενός CNN εισάγουν την μη γραμμικότητα στο δίκτυο εφαρμόζοντας μια μη γραμμική συνάρτηση ενεργοποίησης στην έξοδο των συνελικτικών στρωμάτων. Η πιο συχνά χρησιμοποιούμενη συνάρτηση ενεργοποίησης στα CNN είναι η συνάρτηση διορθωμένης γραμμικής μονάδας (ReLU), η οποία θέτει τις αρνητικές τιμές στο μηδέν και αφήνει αμετάβλητες τις θετικές τιμές.

Τα στρώματα συγκέντρωσης ενός CNN μειώνουν τη δειγματοληψία της εξόδου των συνελικτικών στρωμάτων λαμβάνοντας τη μέγιστη, μέση ή άλλη λειτουργία συγκέντρωσης σε μια μικρή χωρική περιοχή. Αυτό μειώνει τη χωρική διάσταση των χαρτών χαρακτηριστικών και βοηθά στο να γίνει το δίκτυο πιο υπολογιστικά αποδοτικό. Τα πλήρως συνδεδεμένα στρώματα ενός CNN εκτελούν το τελικό βήμα ταξινόμησης ή παλινδρόμησης. Παίρνουν την έξοδο των προηγούμενων επιπέδων (που είναι συνήθως ένα σύνολο χαρακτηριστικών υψηλού επιπέδου) και παράγουν ένα σύνολο τιμών εξόδου που αντιπροσωπεύουν την προβλεπόμενη τιμή κλάσης ή παλινδρόμησης. Ο αριθμός των κόμβων στο επίπεδο εξόδου εξαρτάται από τον αριθμό των κλάσεων ή των τιμών παλινδρόμησης που έχει σχεδιαστεί να προβλέπει το δίκτυο.

Σε σύγκριση με τα παραδοσιακά τεχνητά νευρωνικά δίκτυα (ANN), τα CNN έχουν σχεδιαστεί ειδικά για εργασίες επεξεργασίας εικόνας και χρησιμοποιούν συνελικτικά επίπεδα, τα οποία δεν υπάρχουν στα παραδοσιακά μοντέλα μηχανικής μάθησης. Χρησιμοποιούν επίσης συνήθως στρώματα συγκέντρωσης για να μειώσουν τη χωρική διάσταση των χαρτών χαρακτηριστικών, κάτι που δεν γίνεται στα παραδοσιακά. Ωστόσο, τα CNN είναι συνήθως πιο ακριβά υπολογιστικά από τα παραδοσιακά ANN λόγω του μεγάλου αριθμού των παραμέτρων που εμπλέκονται. Συνολικά, τα CNN έχουν επιδείξει μεγάλη επιτυχία στις εργασίες αναγνώρισης και ταξινόμησης εικόνων και χρησιμοποιούνται ευρέως σε εφαρμογές όρασης υπολογιστών στον τομέα της βαθιάς μάθησης [18].



**Εικόνα 2: Απεικόνιση της αρχιτεκτονικής ενός απλού συνελκτικού νευρωνικού δικτύου**

[https://www.researchgate.net/figure/Schematic-diagram-of-a-basic-convolutional-neural-network-CNN-architecture-26\\_fig1\\_336805909](https://www.researchgate.net/figure/Schematic-diagram-of-a-basic-convolutional-neural-network-CNN-architecture-26_fig1_336805909)

## 2.2 Εκμάθηση Με Μεταφορά Γνώσης

Η διαδικασία εκπαίδευσης είναι κρίσιμη ώστε τα τεχνητά νευρωνικά δίκτυα να κάνουν ακριβείς προβλέψεις. Τα μοντέλα μπορούν να εκπαιδευτούν με διάφορες τεχνικές, όπως για παράδειγμα με επιβλεπόμενη μάθηση που περιλαμβάνει ένα επισημασμένο σύνολο δεδομένων ή την μη-επιβλεπόμενη μάθηση που δεν χρειάζονται δεδομένα με επισημασμένες ετικέτες [19]. Υπάρχουν και περιπτώσεις όπου το σύνολο δεδομένων μπορεί να αποτελείται και από τα δυο είδη, δηλαδή μια ποσόστωση επισημασμένων και μη-επισημασμένων δεδομένων. Για αυτό τον λόγο έχουν διερευνηθεί μικτές προσεγγίσεις που οδήγησαν στην εξέλιξη τεχνικών ημι-επιβλεπόμενης μάθησης, με σκοπό να δημιουργηθούν πιο αποτελεσματικοί ταξινομητές [20]. Ωστόσο στις περισσότερες προσεγγίσεις αυτού του είδους γίνεται η υπόθεση ότι τα επισημασμένα και μη-επισημασμένα δεδομένα ακολουθούν την ίδια κατανομή [21].

Στην εκμάθηση με μεταφορά γνώσης δεν χρειάζεται αυτή η υπόθεση. Ένα μοντέλο που εκπαιδεύεται με αυτήν την τεχνική αξιοποιεί την γνώση που απέκτησε κατά την διαδικασία εκπαίδευσης σε μια εργασία ώστε να βελτιώσει την απόδοσή του σε μια άλλη [22]. Αυτή η προσέγγιση έχει εξελιχθεί σημαντικά της τελευταίες δεκαετίες με την άνθιση της βαθιάς μάθησης και έχει υιοθετηθεί ευρέως.

Οι πρώτες προσπάθειες ξεκίνησαν την δεκαετία του 1990 σε σχετικές έννοιες όπως η επαγωγική μάθηση και η εκπαίδευση πολλαπλών παράλληλων εργασιών. Η θεμελίωση αυτών των εννοιών έθεσαν την βάση για περαιτέρω διερεύνηση. Μία από τις θεμελιώδεις συνεισφορές στον τομέα της εκπαίδευσης πολλαπλών εργασιών ήταν από τον Caruana, ο οποίος έδειξε ότι η εκμάθηση πολλαπλών εργασιών ταυτόχρονα με κοινές αναπαραστάσεις θα μπορούσε να αποφέρει ανώτερα αποτελέσματα σε σύγκριση με τις

μαθησιακές εργασίες μεμονωμένα, αποδεικνύοντας ότι η επαγωγική μάθηση έχει πρακτική εφαρμογή [23].

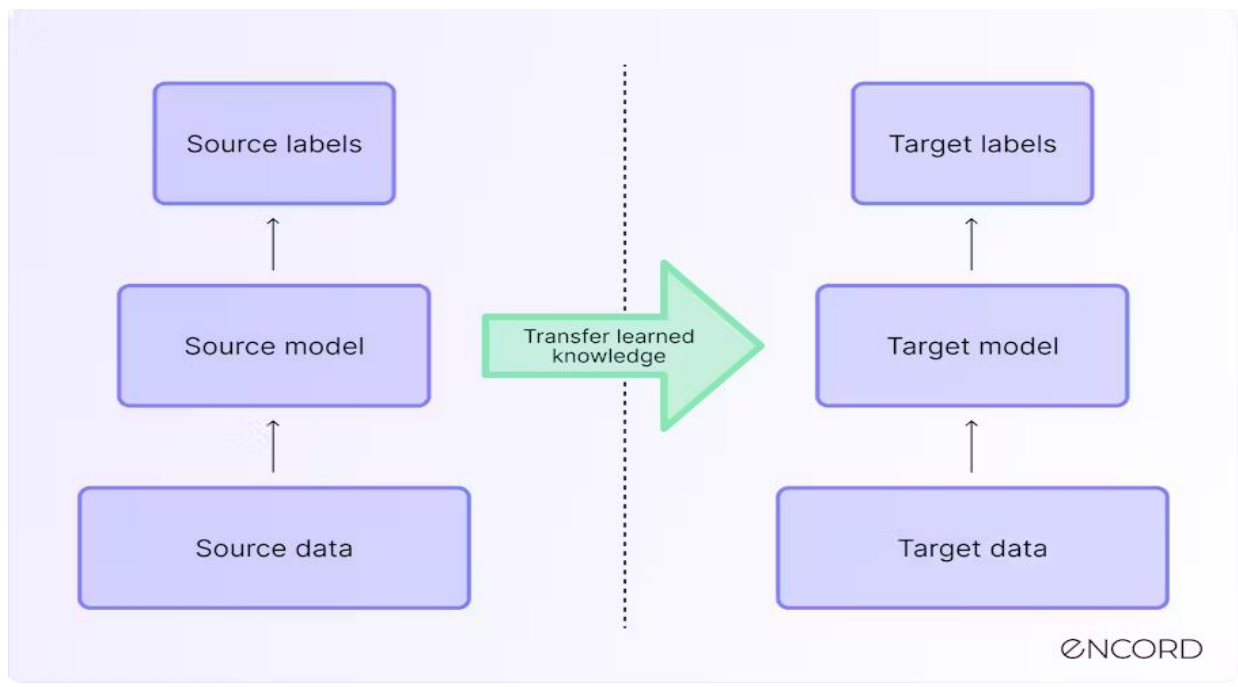
Παρόλα αυτά οι υπολογιστικοί περιορισμοί της εποχής δεν μπορούσαν να επιτρέψουν την εφαρμογή της τεχνικής μεταφοράς γνώσης σε ευρεία κλίμακα. Οι παραδοσιακές μέθοδοι μηχανικής μάθησης απαιτούν εκτεταμένη επεξεργασία ώστε να παραχθούν χαρακτηριστικά και τις περισσότερες φορές είναι συγκεκριμένα για κάθε εργασία, καθιστώντας τη δυνατότητα μεταφοράς της μαθημένης γνώσης δύσκολη.

Η έλευση της βαθιάς μάθησης τη δεκαετία του 2010 σηματοδότησε ένα σημείο καμπής για τη μεταφορά μάθησης. Τα βαθιά νευρωνικά δίκτυα, ιδιαίτερα τα συνελκτικικά νευρωνικά δίκτυα (CNN), αποδείχθηκαν εξαιρετικά αποτελεσματικά στην εκμάθηση ιεραρχικών αναπαραστάσεων δεδομένων. Αυτές οι αναπαραστάσεις, που αντλήθηκαν από μεγάλα σύνολα δεδομένων, θα μπορούσαν να βελτιστοποιηθούν ώστε να αποδίδουν καλά σε μικρότερα σύνολα δεδομένων για συγκεκριμένες εργασίες. Το AlexNet (2012) [17] και οι επόμενες αρχιτεκτονικές όπως το VGGNet (2014) [24] και το ResNet (2015) κατέδειξαν τη χρησιμότητα των προ εκπαιδευμένων μοντέλων σε εργασίες αναγνώρισης εικόνων [25].

Ταυτόχρονα, ο τομέας της επεξεργασίας φυσικής γλώσσας (NLP) γνώρισε καινοτομίες με την ανάπτυξη ενσωματώσεων λέξεων όπως το Word2Vec (2013) [26] και το GloVe (2014) [27], οι οποίες κατέγραψαν σημασιολογικές σχέσεις μεταξύ των λέξεων. Η τεχνική της εκμάθησης με μεταφορά γνώσης εξελίχθηκε ακόμα περισσότερο με την εισαγωγή αρχιτεκτονικών που βασίζονται σε μετασχηματιστές (transformers), όπως το BERT (2018) [28] και το GPT (2018) [29], οι οποίες θα μπορούσαν να γενικεύσουν σε διάφορες εργασίες επεξεργασίας φυσικής γλώσσας με ελάχιστες προσαρμογές για συγκεκριμένες εργασίες.

Σήμερα, η μάθηση με μεταφορά γνώσης αποτελεί ακρογωνιαίο λίθο της τεχνητής νοημοσύνης, δίνοντας τη δυνατότητα σε ερευνητές και επαγγελματίες να αντιμετωπίσουν πολύπλοκα προβλήματα σε όλους τους τομείς. Οι εφαρμογές της μεθόδου έχουν υιοθετηθεί ευρέως σε τομείς όπως η ιατρική [30], η οικονομία [31] αλλά και τον αγροτικό τομέα που είναι σχετικός με την παρούσα μεταπτυχιακή διατριβή. Μειώνοντας την ανάγκη για εκτεταμένα δεδομένα με ετικέτα και υπολογιστικούς πόρους, η μάθηση με μεταφορά γνώσης έχει επιτρέψει την πρόσβαση σε ισχυρές δυνατότητες μηχανικής εκμάθησης.

Συμπερασματικά, το πεδίο της εκπαίδευσης με μεταφορά γνώσης έχει εξελιχθεί από μια θεωρητική έννοια σε ένα μετασχηματιστικό εργαλείο που στηρίζει τα σύγχρονα συστήματα τεχνητής νοημοσύνης. Η ικανότητά του να αξιοποιεί την προηγούμενη γνώση και να την προσαρμόζει σε νέες εργασίες έχει επαναπροσδιορίσει τα όρια του τι μπορεί να επιτύχει η τεχνητή νοημοσύνη [32]. Καθώς η υπολογιστική ισχύς και η διαθεσιμότητα δεδομένων συνεχίζουν να αυξάνονται, η μάθηση με μεταφορά πιθανότατα θα διαδραματίσει ακόμη μεγαλύτερο ρόλο στη διαμόρφωση του μέλλοντος της τεχνολογίας.



**Εικόνα 3: Σχηματική απεικόνιση της διαδικασίας μάθησης με μεταφορά γνώσης.**

<https://encord.com/blog/transfer-learning/>

### 3. Βιβλιογραφική Ανασκόπηση του πεδίου

Η τεχνητή νοημοσύνη έχει συνεισφέρει σημαντικά στον τομέα της γεωργίας, προσφέροντας ευέλικτες, υψηλής απόδοσης, ακριβείς και οικονομικά αποδοτικές λύσεις σε διάφορα προβλήματα [33]. Στην βιβλιογραφία εντοπίζονται πολλές έρευνες που βασίζονται σε τεχνικές βαθιάς μάθησης και συνελκτικά νευρωνικά δίκτυα CNN, όμως κάθε μια από αυτές γίνεται σε διαφορετικό πλαίσιο. Έτσι, η βιβλιογραφία θα μπορούσε να χωριστεί σε τρεις μεγάλες κατηγορίες ανάλογα με το πλαίσιο που διεξάγονται οι έρευνες και ποιο πρόβλημά προσπαθούν να επιλύσουν. Οι πιο κοινές περιπτώσεις είναι τρεις. Πιο συγκεκριμένα, εντοπίζονται έρευνες στην βιβλιογραφία που στοχεύουν στην αναγνώριση και την ταξινόμηση φύλλων διάφορων φυτών, στην αναγνώριση ασθενειών μέσω εικόνων των φύλλων τους και στην ταξινόμηση και αναγνώριση σπόρων. Είναι λοιπόν προφανές πως η επεξεργασία εικόνας και η ανάλυση της έχουν ιδιαίτερο ενδιαφέρον για τους ερευνητές στη γεωργία, ειδικά για τη διαλογή και την ταξινόμηση των σπόρων για τη βελτίωση της ποιότητας των καλλιεργειών.

Στο πλαίσιο της αναγνώρισης και ταξινόμησης φύλλων διάφορων φυτών, αρκετές μέθοδοι έχουν προταθεί. Για παράδειγμα ο Hall [34], πρότεινε ένα συνεκτικό νευρωνικό δίκτυο ώστε να πραγματοποιηθεί ταξινόμηση σε ένα σύνολο δεδομένων που αποτελούνταν από 1900 εικόνες φύλλων 32 διαφορετικών ειδών φυτών. Η μέση ακρίβεια του συγκεκριμένου μοντέλου ήταν 97.3% και η συνδυαστική προσέγγιση που ακολουθήθηκε ώστε να πραγματοποιηθεί η ταξινόμηση ξεπερνά τις επιδόσεις των παραδοσιακών τεχνικών. Παρόλα αυτά, οι εικόνες που χρησιμοποιήθηκαν είχαν ληφθεί

σε ιδανικές συνθήκες κάτι που σπάνια συναντιέται στον φυσικό κόσμο. Στο σενάριο όπου η λήψη πραγματοποιηθεί σε εξωτερικό χώρο στο φυσικό περιβάλλον δηλαδή που αναπτύσσεται το φυτό, τότε οι καιρικές συνθήκες θα μπορούσαν να επηρεάσουν το φως και κατ' επέκταση την ποιότητα της εικόνας την στιγμή που λήφθηκε η φωτογραφία.

Σε παρόμοιο πλαίσιο κινείται η ερευνά της Di Ruberto [35], όπου χρησιμοποιήθηκαν εικόνες από το ίδιο σύνολο δεδομένων σε τεχνητό φόντο. Ο στόχος των συγγραφέων ήταν να προτείνουν ένα σύστημα ταξινόμησης φύλλων το οποίο βασίζεται σε Support Vector Machines και στην εξαγωγή ενός καινούριου και ευρύτερου συνόλου χαρακτηριστικών (features). Το σύνολο αυτό αποτελείται από 138 διαφορετικά χαρακτηριστικά και ενσωματώνει το χρώμα, το σχήμα και την υφή του φύλλου ώστε να διασφαλιστεί η μεγαλύτερη δυνατή ακρίβεια στον καλύτερο δυνατό χρόνο. Σημειώνεται ότι τα αποτελέσματα ήταν εντυπωσιακά επιτυγχάνοντας ακρίβεια 99% και δίνοντας την δυνατότητα ταξινόμησης ενός αγνώστου είδους φυτού σε περίπου 1 δευτερόλεπτο.

Στην κατεύθυνση της αναγνώρισης ασθενειών μέσω εικόνων των φύλλων ενός φυτού, χρησιμοποιούνται παρόμοιες τεχνικές σε αρκετά διαφορετικά είδη με αξιολογία αποτελέσματα. Σε μια έρευνα που αφορούσε την αναγνώριση δυο διαφορετικών ασθενειών που εμφανίζονται στις μπανάνες χρησιμοποιήθηκε η αρχιτεκτονική LeNet σε ένα σύνολο δεδομένων 3700 εικόνων. Ο σκοπός των ερευνητών ήταν να παρουσιάσουν ένα εύκολο στην χρήση και χαμηλό σε κόστος εργαλείο που θα μπορέσει να βοηθήσει τους παράγωγους να δράσουν εγκαίρως ώστε να διατηρηθούν υγιή τα φυτά τους. Οι ερευνητές για να προσομοιώσουν την μεταβλητή φύση της φωτεινότητας του πραγματικού κόσμου επεξεργάστηκαν τις φωτογραφίες δημιουργώντας δυο υποσύνολα ένα έγχρωμο και ένα σε αποχρώσεις γκρι κλίμακας (grayscale images). Στα πειράματα που διεξήχθησαν, το προτεινόμενο μοντέλο επέτυχε μέγιστη ακρίβεια 99% για τις έγχρωμες φωτογραφίες, ενώ για τις εικόνες γκρι κλίμακας παρουσίασε μέγιστη ακρίβεια ίση με 97% [36].

Μια άλλη αξιολογία πρόταση που εφαρμόστηκε σε μεγάλη κλίμακα αποτελεί το CaffeNet [37], ένα συνελκτικό νευρωνικό δίκτυο (CNN) το οποίο χρησιμοποιήθηκε για ταξινόμηση σε μια βάση δεδομένων, η οποία περιέχει περισσότερες από 33.000 εικόνες. Η ακρίβεια του συγκεκριμένου μοντέλου ήταν μεταξύ 91-98% και έχει την ικανότητα να διακρίνει τα φύλλα των φυτών στο φυσικό τους περιβάλλον, διαπιστώνοντας αν υπάρχουν οπτικές διάφορες σε σχέση με υγιή φυτά, επιτρέποντας του να αναγνωρίζει επιτυχώς 13 διαφορετικούς τύπους ασθενειών.

Εξίσου αξιολογία αποτελέσματα παρουσιάζει το έργο του Mohanty [38], όπου δοκιμάστηκαν δυο αρχιτεκτονικές βαθιάς μάθησης σε ένα σύνολο δεδομένων που αποτελούταν από 54.000 δημόσια διαθέσιμες εικόνες φυτών με σκοπό την αναγνώριση ασθενειών μέσω των εικόνων των φύλλων τους. Οι αρχιτεκτονικές που δοκιμάστηκαν ήταν βασισμένες στις αρχιτεκτονικές των μοντέλων AlexNet και GoogleNet χρησιμοποιώντας δυο τρόπους εκπαίδευσης. Εφαρμόστηκε εκπαίδευση με μεταφορά γνώσης (Transfer Learning) και η παραδοσιακή μέθοδος εκπαίδευσης χωρίζοντας το σύνολο δεδομένων σε σύνολο εκπαίδευσης και σύνολο ελέγχου. Το τελικό μοντέλο είχε ακρίβεια 99.35% στο σύνολο ελέγχου παρουσιάζοντας εξαιρετικά αποτελέσματα για την αναγνώριση ασθενειών σε διαφορετικά είδη φυτών.

Μια πιο πρόσφατη ερευνά σε σύγκριση με τις δυο προηγούμενες η οποία κινείται και αυτή στο πλαίσιο αναγνώρισης ασθενειών προτείνει ένα ολοκληρωμένο πλαίσιο εργασίας που βασίζεται στα συνελκτικά νευρωνικά δίκτυα. Το προτεινόμενο πλαίσιο ακολουθεί μια υβριδική προσέγγιση και αποτελεί συνδυασμό δυο διαφορετικών μοντέλων ενός Single Shot Detector (SSD) MobileNet και ενός CNN, επιτρέποντας την αναγνώριση ασθενειών με ακρίβεια 96.88%. Ένα αξιοσημείωτο γεγονός είναι πως το



μοντέλο αναγνωρίζει τα φύλλα της καλλιέργειας και έχει την ικανότητα να ταξινομεί τον τύπο της ασθένειας ακόμα και αν υπάρχει σε μεμονωμένα φύλλα. Το πλήρες σύστημα αποτελείται από μια κάμερα η οποία λαμβάνει στιγμιότυπα σε πραγματικό χρόνο και ένα ενσωματωμένο υλικό το οποίο επιτρέπει την εύκολη μεταφορά του. Έτσι, μπορεί να γίνει αναγνώριση ασθενειών σε πραγματικό χρόνο κατευθείαν στην τοποθεσία του χωραφιού γρηγορά και αποτελεσματικά, χωρίς να είναι απαραίτητη η παρουσία κάποιου ειδικού [39].

Στην τελευταία κατηγορία, ανήκουν οι έρευνες που έχουν παρόμοιο αντικείμενο με την παρούσα διπλωματική και αφορούν την αναγνώριση και ταξινόμηση σπορών. Από τη μελέτη της βιβλιογραφίας προκύπτει πως οι ερευνητές έχουν χρησιμοποιήσει επιτυχώς τεχνικές μηχανικής μάθησης (ML) όπως ταξινομητές Naïve Bayes, ταξινομητές δέντρων αποφάσεων (Decision Tree Classifier) και MLP για να εκπαιδεύσουν μοντέλα ώστε να πραγματοποιήσουν ταξινόμηση σπόρων. Αυτές οι τεχνικές χρησιμοποιήθηκαν ώστε να εκπαιδευτεί ένα μοντέλο σε ένα δείγμα 900 σπόρων βαμβακιού και αναφέρθηκε ποσοστό ακρίβειας 98.7% για τον ταξινομητή δέντρου απόφασης και MLP και 94.22% για τον ταξινομητή Naïve Bayes. Ο ταξινομητής Naïve Bayes είχε το υψηλότερο ποσοστό σφάλματος, κάνοντας 52 λανθασμένες ταξινομήσεις, ενώ ο ταξινομητής δέντρου αποφάσεων και ο MLP έκαναν μόνο 11 το καθένα. Σε μια παρόμοια ερευνά με την προηγούμενη, η οποία όμως έχει ως στόχο να βελτιώσει την διαλογή σπόρων πιπεριού, χρησιμοποιήθηκε λογισμικό αναγνώρισης εικόνων. Αρχικά, με την βοήθεια του λογισμικού εντοπίστηκαν τα φυσικά χαρακτηριστικά που υποδεικνύουν την ευρωστία των σπόρων. Στη συνέχεια, με την βοήθεια ενός μοντέλου MLP έγινε πρόβλεψη της βλάστησης των σπόρων χρησιμοποιώντας 15 μεταβλητές οι οποίες προήλθαν από τα φυσικά τους χαρακτηριστικά επιτυγχάνοντας ακρίβεια διαλογής 90% [40].

Η εξαγωγή χαρακτηριστικών στις παραδοσιακές τεχνικές ML βασίζεται σε χαρακτηριστικά που καθορίζονται από τον χρήστη, το γεγονός αυτό μπορεί να οδηγήσει σε απώλεια σημαντικών πληροφοριών επηρεάζοντας τα αποτελέσματα σε σημαντικό βαθμό. Οι τεχνικές βαθιάς μάθησης μπορούν να βοηθήσουν στην αντιμετώπιση αυτού του προβλήματος, αφού τα χαρακτηριστικά της εικόνας μπορούν να προσδιοριστούν σε διαφορετικά επίπεδα. Οι Rozman και Stajniko πρότειναν ένα σύστημα όρασης υπολογιστή που χρησιμοποιούσε ένα φίλτρο Gaussian, τμηματοποίηση και περιοχή ενδιαφέροντος (ROI) για την εξαγωγή χαρακτηριστικών σε ένα δείγμα 700 σπόρων τομάτας [41]. Εφάρμοσαν διαφορετικούς αλγόριθμους ταξινόμησης ML, συμπεριλαμβανομένων ταξινομητών Naïve Bayes, k-πλησιέστερων γειτόνων (K-NN), ταξινομητών δέντρων αποφάσεων, μηχανών υποστήριξης διανυσμάτων και τεχνητών νευρωνικών δικτύων (ANN) με αρχιτεκτονική MLP και ανέφεραν ποσοστό ακρίβειας 95,44% για το ANN. Τα ποσοστά ακρίβειας για τους άλλους αλγόριθμους ήταν 87,89% για το NBC, 91,66% για το k-NN, 93,66% για το DT και 93,09% για το SVM, επομένως είναι εμφανής η υπεροχή της απόδοσης του νευρωνικού δικτύου MLP σε σχέση με τους υπόλοιπους αλγόριθμους.

Σε μια άλλη έρευνα που χρησιμοποιήθηκαν παραδοσιακές τεχνικές μάθησης και αφορούσε την ταξινόμηση σπόρων τριών διαφορετικών ποικιλιών Ιρανικού ρυζιού χρησιμοποιήθηκαν 666 φωτογραφίες συνολικά οι οποίες παρουσιάζουν σταθερές συνθήκες φωτισμού. Μετά από ανάλυση των κύριων συνιστωσών (Principal component analysis) οι ερευνητές καταλήγουν στα πιο σημαντικά χαρακτηριστικά του συνόλου δεδομένων και με χρήση ενός νευρωνικού δικτύου MLP κάνουν ταξινόμηση των σπόρων επιτυγχάνοντας μέση ακρίβεια 90% [42]. Μια άλλη προσέγγιση για την ταξινόμηση σπόρων ρυζιού στηρίζεται στην εξαγωγή των χαρακτηριστικών που πηγάζουν από τις φυσικές ιδιότητες των σπόρων με σκοπό την αποτελεσματική ταξινόμηση 14 διαφορετικών ποικιλιών ώστε να εξαχθούν τα χαρακτηριστικά τους. Στη συνέχεια,

δοκιμάστηκαν και συγκρίθηκαν τέσσερις στατιστικές μέθοδοι μηχανικής μάθησης (LR, LDA, k-NN και SVM) και πέντε προεκπαιδευμένα μοντέλα βαθιάς μάθησης. (VGG16, VGG19, Xception, InceptionV3 και InceptionResNetV2) [43]. Τα χαρακτηριστικά εξήχθησαν από ένα σύνολο δεδομένων 50.000 εικόνων σπόρων ρυζιού και αφορούσαν το χρώμα, την ύφη και το σχήμα. Στις στατιστικές μεθόδους η μεγαλύτερη ακρίβεια επιτεύχθηκε από το μοντέλο SVM με 83.9% ενώ από τα μοντέλα βαθιάς μάθησης την καλύτερη επίδοση σημείωσε το InceptionResNetV2 με ακρίβεια 95.15%. Μια ακόμη έρευνα που αφορά σπόρους ρυζιού σύγκρινε τους αλγορίθμους K-NN και SVM με μοντέλα CNN [44]. Ο σκοπός της συγκεκριμένης έρευνας είναι η αναγνώριση της ποικιλίας ρυζιού που ανήκει ο σπόρος. Το σύνολο δεδομένων αποτελούνταν από υπερφασματικές εικόνες και τα αποτελέσματά της δείχνουν την ανωτερότητα του CNN σε σχέση με τους άλλους δυο αλγορίθμους επιτυγχάνοντας ακρίβεια 89.6 % στο σύνολο εκπαίδευσης και 87% στο σύνολο ελέγχου. Η μέθοδος με τις υπερφασματικές εικόνες μπορεί να αποτελέσει μια λύση για ποικιλίες που σπόρων που εμφανίζουν τεράστια οπτική ομοιότητα.

Σε πιο πρόσφατες έρευνες οι ερευνητές προτείνουν νέες παραλλαγές μοντέλων βαθιάς μάθησης ώστε να επιτύχουν ακόμα καλύτερες επιδόσεις, ειδικότερα σε περιπτώσεις όπου οι οπτικές διαφορές μπορεί να είναι ανεπαίσθητες. Ενώ τα CNN έχουν επικρατήσει ως η πρώτη επιλογή για την αναγνώριση και ταξινόμηση διάφορων αντικειμένων παρουσιάζουν μερικά μειονεκτήματα. Χρειάζονται πολύ μεγάλη υπολογιστική ισχύ ενώ τα χαρακτηριστικά που εξάγονται από τα συνελκτικά στρώματα δεν διερευνώνται σε βάθος διότι έχουν πολύ μεγάλο χώρο διαστάσεων .

Έτσι, σε μια έρευνα που διεξήχθη σε σπόρους ηλιοτροπίων [45], οι συγγραφείς προτείνουν ένα CNN που αποτελείται από 8 συνελκτικά στρώματα και συνοδεύεται από έναν μηχανισμό προσοχής ώστε να αποφευχθούν τα παραπάνω μειονεκτήματα. Το προτεινόμενο μοντέλο έχει μικρότερο μέγεθος από ένα κλασικό CNN και με την βοήθεια του μηχανισμού προσοχής έχει την ικανότητα να ενισχύει τα χαρακτηριστικά της εικόνας που θα ταξινομηθεί οδηγώντας σε μεγαλύτερη ακρίβεια η οποία αγγίζει .

Μια άλλη παραλλαγή ενός CNN που χρησιμοποιεί τεχνικές μεταφοράς μάθησης (transfer learning) είναι το P-ResNet, το οποίο προτάθηκε σε μια έρευνα που έγινε για την ταξινόμηση σπόρων αραβοσίτου [46]. Το σύνολο δεδομένων αποτελούνταν από περισσότερες από 8000 εικόνες σπόρων και το προτεινόμενο μοντέλο αποτελείται από 5 συνελκτικά στρώματα και 1 πλήρως συνδεδεμένο στρώμα. Τα συγκριτικά πειράματα ανέδειξαν την ανωτερότητα του P-ResNet σε σχέση με άλλα κλασικά CNN μοντέλα όπως το AlexNet και το GoogleNet. Η ακρίβεια που επιτεύχθηκε ήταν κατά μέσο όρο 99.5% και για τις 5 ποικιλίες αραβοσίτου που δοκιμάστηκαν.

Από την μελέτη της σχετικής βιβλιογραφίας προκύπτουν έρευνες που είναι συνυφασμένες με το αντικείμενο της παρούσας διατριβής και αφορούν την εφαρμογή τεχνικών βαθιάς μάθησης για την ταξινόμηση σπόρων διάφορων φυτών συμπεριλαμβανομένου του σιταριού αλλά και άλλων ειδών δημητριακών. Αρκετοί ερευνητές φαίνεται πως κατευθύνονται προς την λύση της εφαρμογής τεχνικών βαθιάς μάθησης για την ταξινόμηση σπόρων και πειραματίζονται με διαφορετικές αρχιτεκτονικές ώστε να προτείνουν καλύτερα μοντέλα ταξινόμησης.

Στο πλαίσιο αυτό κινείται και η ερευνά του Yasar, με σκοπό την αξιολόγηση διαφορετικών μοντέλων CNN για την ταξινόμηση πέντε ποικιλιών σίτου [11]. Αρχικά δημιουργήθηκε ένα σύνολο δεδομένων με 8354 εικόνες σπόρων σιταριού το οποίο χρησιμοποιήθηκε για την εκπαίδευση και την αξιολόγηση των μοντέλων. Τα τρία μοντέλα που συγκρίθηκαν ήταν το Inception-V3, το Mobilenet-V2 και το Resnet18, τα οποία ταξινόμησαν σπόρους σιταριού στην σωστή ποικιλία με ακρίβεια 97.37%, 97.07%, και 97.67%, αντίστοιχα.

Εξίσου μεγάλο ενδιαφέρον παρουσιάζει μια ερευνά στην οποία χρησιμοποιήθηκε ένα υβριδικό μοντέλο CNN-SVM για την ταξινόμηση σπόρων σιταριού [47]. Η διαφορά που εντοπίζεται σε σχέση με άλλα έργα του πεδίου είναι ότι σε αυτή την περίπτωση το μοντέλο βαθιάς μάθησης χρησιμοποιείται για την εξαγωγή χαρακτηριστικών και όχι για την ταξινόμηση του σύνολο δεδομένων. Είναι γεγονός πως για την αποτελεσματική εκπαίδευση ενός CNN πρέπει να υπάρχει μεγάλη ποσότητα δεδομένων, για αυτό οι ερευνητές επέλεξαν εικόνες υψηλής ανάλυσης διαφορετικών ποικιλιών σποριασμένου σίτου οι οποίες επεξεργάστηκαν και έγιναν περικοπή σε μικρότερες. Με αυτό τον τρόπο δημιουργήθηκε ένα δείγμα 800 εικόνων από τέσσερις διαφορετικές ποικιλίες. Στη συνέχεια, επτά διαφορετικά προ-εκπαιδευμένα μοντέλα CNN χρησιμοποιήθηκαν για την εξαγωγή χαρακτηριστικών ενώ η ταξινόμηση πραγματοποιήθηκε από ένα ταξινομητή SVM επιτυγχάνοντας ακρίβεια 98.1%.

Μια ακόμα λύση για την ανάπτυξη αποδοτικών μοντέλων μπορεί να αποτελέσει η μέθοδος της μεταφοράς γνώσης. Η ανάπτυξη των τεχνικών μεταφοράς γνώσης (Transfer Learning) έχει ανοίξει νέους ορίζοντες εξερεύνησης και μπορεί να επιτρέψει την ταξινόμηση σπόρων σε ευρεία κλίμακα. Τα μοντέλα που εκπαιδεύονται με μεταφορά γνώσης χρειάζονται λιγότερη υπολογιστική ισχύ αλλά και λιγότερο χρόνο εκπαίδευσης σε σχέση με την εκπαίδευση ενός παραδοσιακού μοντέλου βαθιάς μάθησης [22].

Σε αυτή την κατεύθυνση μια από τις έρευνες εξετάζει και συγκρίνει αυτές δοκιμάστηκαν πέντε διαφορετικές αρχιτεκτονικές CNN με τεχνικές μεταφοράς γνώσης (Transfer Learning) ώστε να γίνει ταξινόμηση των σπόρων σιταριού σε τέσσερις διαφορετικές ποικιλίες [48]. Οι ερευνητές χρησιμοποίησαν ένα σύνολο δεδομένων που αποτελούταν από περισσότερες από 31.000 εικόνες σπόρων σιταριού από διάφορες περιοχές της Αλγερίας. Στο σύνολο δεδομένων ελέγχου τα τρία καλύτερα μοντέλα ήταν με αρχιτεκτονική DensNet201, Inception V3 και MobileNet επιτυγχάνοντας ακρίβεια 95.68%, 95.62% και 95.49% αντίστοιχα.

Παρόμοια κινείται και μια έρευνα που χρησιμοποιεί την μέθοδο της μεταφοράς γνώσης ώστε να πραγματοποιηθεί ταξινόμηση σε εικόνες σπόρου καλαμποκιού. Οι ερευνητές προσπαθούν να αναδείξουν τα μοντέλα βαθιάς μάθησης ως ένα χρήσιμο εργαλείο για την αναγνώριση και διάκριση σπόρων μεταξύ απλοειδούς και διπλοειδούς καλαμποκιού. Χρησιμοποιώντας ένα μικρό σύνολο δεδομένων που αποτελούταν συνολικά από περίπου 3000 εικόνες των δυο ειδών καλαμποκιού και εφαρμόζοντας την μέθοδο της μεταφοράς γνώσης επέτυχαν ακρίβεια μεγαλύτερη από 93% [49]. Το γεγονός πως με την μεταφορά γνώσης επιτευχθήκαν τόσο καλά αποτελέσματα σε ένα μικρό σύνολο δεδομένων μπορεί να αποτελέσει πυλώνα προόδου σε ένα πεδίο που η απουσία μεγάλης ποσότητας δεδομένων είναι συχνή.

Βέβαια το ερευνητικό ενδιαφέρον δεν περιορίζεται μόνο στις κατηγορίες που αναφέρθηκαν σε αυτήν την ενότητα. Αρκετοί ερευνητές έχουν χρησιμοποιήσει την βαθιά μάθηση και σε άλλα υπο-πεδία της γεωργίας όπως τον ποιοτικό έλεγχο των γεωργικών προϊόντων. Για παράδειγμα, ένας συνδυασμός ενός CNN για την εξαγωγή χαρακτηριστικών μαζί με ένα μοντέλο SVM χρησιμοποιήθηκε σε εικόνες καρότων ώστε να διασφαλιστεί η οπτική τους ποιότητα [50]. Ενώ σε μια άλλη έρευνα μια βελτιωμένη παραλλαγή του μοντέλου YOLO χρησιμοποιήθηκε σε ένα σύνολο δεδομένων αποτελούμενο από εικόνες που έχουν ληφθεί από drone [51]. Η έρευνα αφορούσε τους καρπούς των φοινικόδεντρων και επιτεύχθηκε αναγνώριση των φρούτων σε λιγότερο από 35 ms με ακρίβεια 99.76%

Συμπερασματικά, από την μελέτη της βιβλιογραφίας φαίνεται πως μοντέλα CNN αλλά και προηγμένες τεχνικές βαθιάς μάθησης έχουν υιοθετηθεί ευρέως στον γεωργικό τομέα για διαφορετικών ειδών εργασίες. Εξίσου σημαντικό είναι το γεγονός πως αρκετές έρευνες έχουν αποδείξει ότι η τεχνητή νοημοσύνη μπορεί να εφαρμοστεί αποτελεσματικά

και σε μεγάλη κλίμακα για την επίλυση χρονοβόρων και πολύπλοκων προβλημάτων. Το μόνο σίγουρο είναι ότι όσο εξελίσσεται ο τομέας της μηχανικής μάθησης όλο και περισσότερες εφαρμογές θα προκύπτουν οδηγώντας σε νέες καινοτομίες και ουσιαστική βελτίωση και εξέλιξη των υπαρχόντων πρακτικών.

#### **4. Υλικά και μέθοδοι**

Στην ενότητα αυτή θα παρουσιαστούν τα υλικά και οι μέθοδοι που χρησιμοποιήθηκαν για την ανάπτυξη του συστήματος ταξινόμησης σπορών. Το προτεινόμενο σύστημα πρόβλεψης μπορεί να χρησιμοποιηθεί ως βάση για ένα ολοκληρωμένο σύστημα υποστήριξης αποφάσεων για αγρότες αφού επιτρέπει γρήγορη και με μεγάλη ακρίβεια ταξινόμηση.

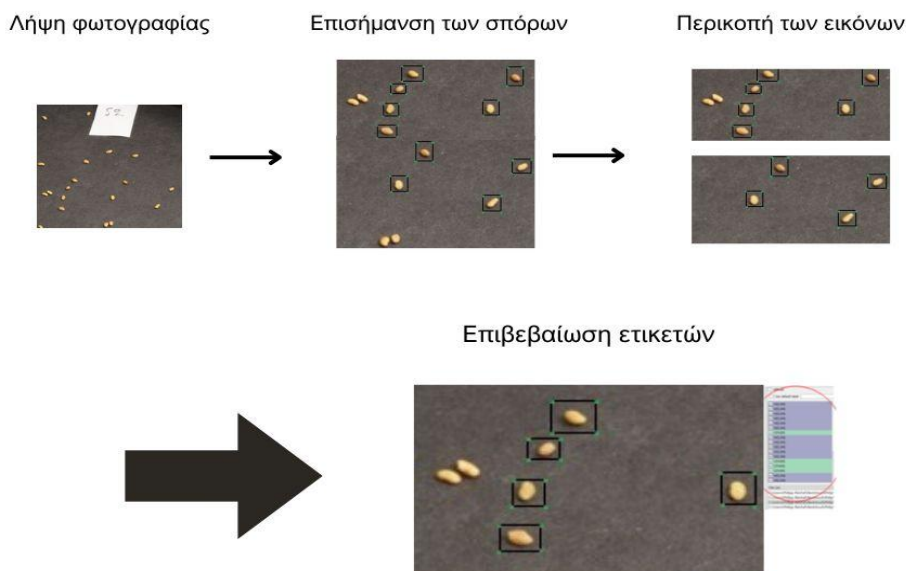
Τα δεδομένα που χρησιμοποιήθηκαν δεν έχουν αντληθεί από κάποια δημοσιάζουσα βάση δεδομένων αλλά έχουν δημιουργηθεί από την αρχή χρησιμοποιώντας μια φωτογραφική μηχανή με αποσπώμενο φλας για να ληφθούν οι εικόνες των σπορών. Στη συνέχεια επεξεργάστηκαν με τη βοήθεια διάφορων υπολογιστικών εργαλείων τα οποία και περιγράφονται λεπτομερώς στις επόμενες υπο-ενοότητες. Τέλος, θα αναφερθούν λεπτομερώς οι μέθοδοι, οι τεχνικές και οι αρχιτεκτονικές που χρησιμοποιήθηκαν.

##### **4.1 Το Σύνολο Δεδομένων**

Για την δημιουργία του συνόλου δεδομένων χρησιμοποιήθηκε η φωτογραφική μηχανή SONY ILCE-6000 με φακό SONY E 18-55mm και χρήση φλας. Οι φωτογραφίες αποθηκεύτηκαν σε μορφή JPG με ανάλυση 6000x4000. Οι εικόνες των σπόρων προέρχονται από τρία διαφορετικά είδη και έξι διαφορετικές ποικιλίες και διαφέρουν σε μέγεθος, υφή και χρώμα. Για την λήψη των φωτογραφιών, η κάμερα τοποθετήθηκε σε σταθερό τρίποδο ώστε να διασφαλίζεται η σταθερότητα του φωτογραφικού φακού και να αποφευχθούν τυχόν παραμορφώσεις στην εικόνα λόγω αστάθειας. Η απόσταση των σπόρων από την κάμερα ήταν 1.5 μέτρο και τοποθετήθηκαν τυχαία σε διάφορες θέσεις πάνω σε ένα μαύρο χαρτόνι το οποίο καλυπτόταν περιμετρικά από άλλα χαρτόνια δημιουργώντας στην ουσία ένα κουτί, πάνω από το οποίο τοποθετήθηκε το φλας ώστε να εξασφαλίζονται σταθερές συνθήκες φωτισμού. Η επιλογή μαύρου φόντου για την λήψη των φωτογραφιών έγινε ώστε να περιοριστεί ο «θόρυβος» στα δεδομένα της εικόνας.

Με αυτόν τον τρόπο, ελήφθησαν πολλές διαφορετικές φωτογραφίες από κάθε κατηγορία σπόρου οι οποίες, στη συνέχεια επεξεργάστηκαν με την βοήθεια του εργαλείου Labellmage [52] ώστε να γίνει επισήμανση των σπόρων που υπήρχαν σε κάθε εικόνα. Η επισήμανση γίνεται ως εξής: Το εργαλείο προσφέρει στον χρήστη ένα τετράγωνο πλαίσιο επισήμανσης στο οποίο μπορεί να περικλείει το αντικείμενο ενδιαφέροντος του. Στην συνέχεια ο χρήστης μπορεί να επισημάνει με την ετικέτα που επιθυμεί το αντικείμενο και να το αποθηκεύσει σε διάφορες μορφές.

### Σχηματική απεικόνιση της διαδικασίας επισήμανσης



**Εικόνα 4: Η διαδικασία επισήμανσης των εικόνων του συνόλου δεδομένων.**

Για την επισήμανση των εικόνων με τις κατάλληλες ετικέτες χρησιμοποιήθηκαν κάποιοι κανόνες. Αρχικά απορρίφθηκαν περικομμένα δείγματα των οποίων το συνολικό μέγεθος εικόνας ήταν λιγότερο από 400 pixels. Το γεγονός αυτό σε συνδυασμό με την επιλογή μαύρου φόντου κατά την διαδικασία λήψης των φωτογραφιών, εξασφαλίζει την ύπαρξη όσο δυνατόν λιγότερου «θορύβου» στο σύνολο δεδομένων. Το πλαίσιο επισήμανσης επεκτάθηκε συμμετρικά όπου χρειάστηκε ώστε ο σπόρος να βρίσκεται στο κέντρο του και το ελάχιστο μέγεθος του πλαισίου ήταν 28x28 pixels. Όμως τα μεγέθη των σπόρων διαφέρουν ελαφρώς από είδος σε είδος. Επομένως σε περίπτωση που κάποιοι σπόροι είχαν μεγαλύτερα μεγέθη τότε το πλαίσιο επισήμανσης επεκτεινόταν ώστε να περικλείεται ομοιόμορφα και συμμετρικά η μορφή του σπόρου μέσα σε αυτό. Σε όσες εικόνες παρατηρήθηκε ότι το μέγεθος 28x28 δεν ήταν αρκετό εφαρμόστηκε ίση επέκταση 4 pixels σε όλες τις κατευθύνσεις.

Επιπλέον για να είναι απολύτως ξεκάθαρες οι απεικονίσεις κάθε μορφής των διαθέσιμων ειδών σπόρων, τα πλαίσια επισήμανσης δεν επιτρεπόταν να καλύπτουν σε μεγάλη έκταση άλλα επισημασμένα πλαίσια. Ο στόχος ήταν να διατηρηθεί η συνοχή του συνόλου δεδομένων ώστε σε κάθε φωτογραφία να αντιστοιχούν ξεκάθαρες απεικονίσεις των διαφορετικών ειδών σπόρων. Σε διαφορετική περίπτωση, δηλαδή αν υπήρχε μεγάλη επικάλυψη μεταξύ των πλαισίων επισήμανσης, το σχήμα του σπόρου θα γινόταν διαφορετικά αντιληπτό από το μοντέλο και επομένως θα επηρέαζε την ικανότητα του να γενικεύσει. Με αυτό τον τρόπο εξασφαλίζεται ότι τα περικομμένα δείγματα θα αποτελούνται από εικόνες σπόρων οι οποίες απεικονίζουν με απολυτή σαφήνεια όλα τα χαρακτηριστικά τους.

Η θέσπιση των κανόνων επισήμανσης έγινε για την εξασφάλιση της ποιότητας των δεδομένων. Άλλωστε η απόδοση ενός μοντέλου σε μια εργασία είναι άρρηκτα συνδεδεμένη με την ποιότητα και την ακρίβεια του συνόλου δεδομένων με το οποίο

εκπαιδεύτηκε. Για παράδειγμα η ύπαρξη «θορύβου» ή λανθασμένων τιμών σε κάποιες μεταβλητές μπορεί να οδηγήσει το μοντέλο στην εκμάθηση μη εγκύρων μοτίβων έχοντας επομένως κακή απόδοση. Διατηρώντας μια ενιαία προσέγγιση και δομημένη στρατηγική κατά την διαδικασία επισήμανσης εξασφαλίζεται ότι θα παραχθεί ένα υψηλής ποιότητας σύνολο δεδομένων το οποίο θα περιέχει όλες τις απαραίτητες πληροφορίες που θα βοηθήσουν το μοντέλο να εκπαιδευτεί αποτελεσματικά.

Μετά την διαδικασία επισήμανσης και χρησιμοποιώντας την βοήθεια ενός ειδικού στην γεωργία έγινε επιβεβαίωση των ετικετών μέσω του ίδιου εργαλείου ώστε να εξασφαλιστεί ότι οι ετικέτες που έχουν αποδοθεί στις εικόνες του συνόλου δεδομένων είναι ακριβείς. Στη συνέχεια, οι εικόνες έγιναν περικοπή σε μικρότερες οι οποίες περιέχουν ένα μικρό αλλά ξεκάθαρα επισημασμένο πλήθος σπόρων όπως εμφανίζεται στην εικόνα 4. Οι εικόνες που προέκυψαν, αποθηκευτήκαν σε ένα αρχείο με πληροφορίες που τις αφορούν όπως ημερομηνία λήψης, είδος και ποικιλία σπόρου, μέγεθος φωτογραφίας στην μορφή πλάτος επί ύψος και το συνολικό μέγεθος ριχεί που τις αποτελούν. Έτσι, προέκυψαν 12.720 διαφορετικά δείγματα από τα τρία είδη και τις έξι διαφορετικές ποικιλίες σπόρων, τα οποία χωρίστηκαν σε τρία υποσύνολα δεδομένων. Πιο αναλυτικά, στο σύνολο των 12.720 φωτογραφιών εντοπίζονται 812 φωτογραφίες της ποικιλίας κριθαριού Δήμητρα και 1.137 φωτογραφίες της ποικιλίας κριθαριού Τριπτόλεμος. Από την οικογένεια σιταρίων υπάρχουν 6.512 φωτογραφίες της ποικιλίας Λεβάντε και 1.891 φωτογραφίες σκληρού σιταριού Σέλλας. Ενώ από τις ποικιλίες βρόμης χρησιμοποιήθηκαν 1.178 φωτογραφίες βρόμη Πήγασος και 1.190 από την ποικιλία Φλέγα. Αναλυτική κατανομή του διαθέσιμου πλήθους φωτογραφιών από κάθε είδος και ποικιλία βρίσκεται στον πίνακα 1.

<b>Ποικιλία</b>	<b>Σύνολο Εικόνων</b>
Κριθάρι Τριπτόλεμος (KTR)	1.137
Κριθάρι Δήμητρα (DMT)	812
Σκληρό Σιτάρι Σέλλας (SSS)	1.891
Σιτάρι Λεβάντες (LVT)	6.512
Βρόμη Φλέγα (VRF)	1.190
Βρόμη Πήγασος (VPG)	1.178
<b>Σύνολο</b>	<b>12.720</b>

**Πίνακας 1: Η κατανομή των εικόνων του συνόλου δεδομένων**

Ένα υποσύνολο χρησιμοποιήθηκε για την εκπαίδευση των μοντέλων (training dataset), ένα άλλο υποσύνολο χρησιμοποιήθηκε για επικύρωση (validation dataset) ενώ το τρίτο υποσύνολο χρησιμοποιήθηκε για δοκιμές (test dataset) και αποτελούνταν από εικόνες που δεν έχουν προσπελαστεί καθόλου από τα μοντέλα που αναπτυχθήκαν. Το τελευταίο υποσύνολο χρησιμοποιήθηκε μετά την ολοκλήρωση της εκπαίδευσης ώστε να αξιολογηθεί η ικανότητα ταξινόμησης του μοντέλου σε καινούρια δεδομένα που δεν έχουν ξανά εμφανιστεί σε καμία φάση ανάπτυξης της εφαρμογής.



**Εικόνα 5: Δοκιμαστικές λήψεις σπόρων**

## 4.2 Κανονικοποίηση Των Δεδομένων

Όπως αναφέρθηκε και στην ενότητα 4.1 δημιουργήθηκαν τρία υποσύνολα δεδομένων. Από τις 12.720 εικόνες που προέκυψαν συνολικά και πριν υποστούν οποιαδήποτε επεξεργασία, για κάθε κατηγορία ετικετών διαχωρίστηκε εκ των προτέρων περίπου το 70% του συνόλου των εικόνων και χρησιμοποιήθηκε για την εκπαίδευση των μοντέλων (Training dataset), ένα ποσοστό 15% για την επικύρωση τους (Validation dataset) ενώ το υπόλοιπο 15% για την αξιολόγηση τους και την εξέταση των αποτελεσμάτων (Testing Dataset). Επομένως σε απολυτούς αριθμούς προέκυψαν 8.550 φωτογραφίες για το σύνολο εκπαίδευσης και από 2085 για το σύνολο επικύρωσης και το σύνολο δοκίμων αντίστοιχα.

Ο διαχωρισμός του συνόλου δεδομένων έγινε χρησιμοποιώντας έναν αλγόριθμο βασιζόμενο σε δυο κλάσεις ο οποίος καλείται μέσω ενός `rython dictionary`. Ο αλγόριθμος δέχεται ως παραμέτρους το μονοπάτι του φακέλου όπου βρίσκονται οι διαθέσιμες φωτογραφίες προς διαχωρισμό, το ποσοστό που θα καταλαμβάνουν τα υποσύνολα δεδομένων εκπαίδευσης, επικύρωσης και δοκιμής και τα μονοπάτια των φακέλων όπου θα αποθηκευτεί το κάθε υποσύνολο. Αυτή η προσέγγιση θα βοηθήσει στην αποτελεσματική προετοιμασία και διαχείριση του συνόλου δεδομένων ώστε να διευκολυνθεί η εκπαίδευση των μοντέλων ταξινόμησης.

Εξίσου σημαντικό σε προβλήματα ταξινόμησης είναι να υπάρχει ισορροπία στα δεδομένα κάθε ομάδας προκειμένου τα μοντέλα να εκπαιδευτούν αποτελεσματικά και να γίνει σαφής αξιολόγηση τους. Αν κάποια από τις ομάδες δεν εκπροσωπείται επαρκώς δηλαδή υπάρχουν λίγα δείγματα διαθέσιμα ή κάποια από τις ομάδες αποτελεί την πλειοψηφία εκπροσωπώντας μια μεγάλη ποσότητα του συνόλου δεδομένου, τότε το σύνολο δεδομένων δεν μπορεί να είναι αμερόληπτο. Το γεγονός αυτό έχει ως αποτέλεσμα να επηρεάζεται η ικανότητα του μοντέλου να γενικεύσει την συμπεριφορά τέτοιου είδους τάξεων [53]. Το σύνολο δεδομένων που δημιουργήθηκε είχε ανισορροπία αφού 6512 φωτογραφίες από το σύνολο 12.720 αποτελούσαν από την ποικιλία *σιταριού λεβάντε* δηλαδή καταλαμβάνει το 51% του συνολικού διαθέσιμου πλήθους φωτογραφιών. Για να εξομαλυνθεί η ανισορροπία που υπήρχε στο σύνολο δεδομένων χρησιμοποιήθηκε υποδειγματοληψία σε συνδυασμό με την εφαρμογή μιας απλής τεχνικής επαύξησης δεδομένων.

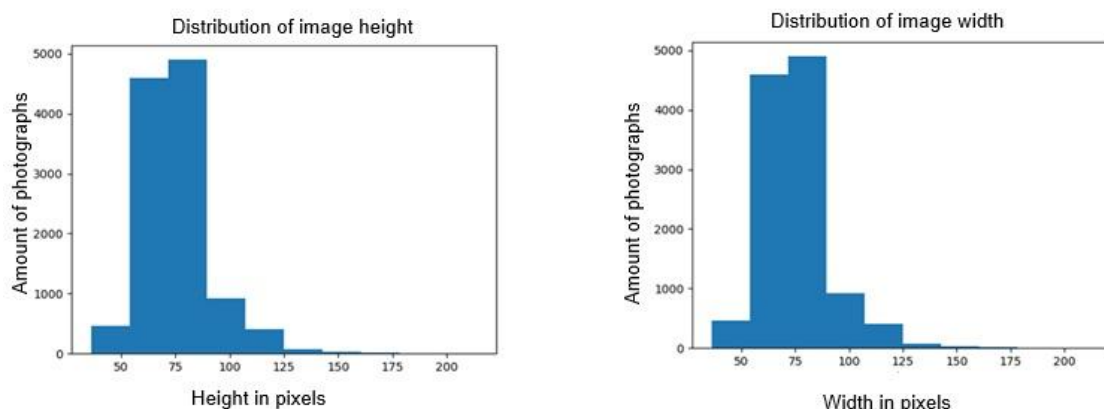
Η υποδειγματοληψία αφορά την επιλογή ενός αντιπροσωπευτικού υποσυνόλου δεδομένων από ένα μεγαλύτερο. Στην πράξη, από κάθε κατηγορία ετικετών επιλέγεται

ένα πλήθος παρατηρήσεων ίσο με το πλήθος παρατηρήσεων που βρίσκεται στην μειοψηφική τάξη. Αν δηλαδή για παράδειγμα, στην μειοψηφική τάξη εντοπίζονται 500 παρατηρήσεις τότε από κάθε κατηγορία επιλέγεται αντίστοιχος αριθμός παρατηρήσεων ώστε να εξαλειφθεί η ύπαρξη προκατάληψης και να επέλθει ισορροπία μεταξύ των διαφορετικών τάξεων. Στην περίπτωση μας το κριθάρι Δήμητρα και Τριπτόλεμος είχαν τα λιγότερα δείγματα με 812 και 1137 αντίστοιχα, ενώ ακόμα και αυτές οι ποικιλίες που έχουν αρκετά μεγαλύτερο αριθμό δειγμάτων σε σχέση με τις δυο ποικιλίες κριθαριών εξακολουθούν να έχουν μεγάλη απόκλιση σε σχέση με την πιο πλειοψηφική τάξη του συνόλου δεδομένων. Η εξαιρετικά μεγάλη απόκλιση των παρατηρήσεων μεταξύ της μικρότερης και μεγαλύτερης τάξης δημιουργεί πολύ μεγάλη δυσκολία στο να προσπελάσει το μοντέλο έναν ικανοποιητικό αριθμό εικόνων ως παραδείγματα. Αρχικά η προσέγγιση ήταν να κατασκευαστεί ένα υποσύνολο, διαλέγοντας ως πλήθος από κάθε τάξη το συνολικό πλήθος της πιο μειοψηφικής τάξης η οποία αποτελούταν από 812 φωτογραφίες. Όμως με αυτό τον τρόπο τα δείγματα των μειοψηφικών τάξεων θα εξαντιλούνταν γρηγορά χωρίς να έχει ολοκληρωθεί η προσπέλαση όλων των εικόνων της πλειοψηφικής τάξης, γεγονός που κάνει πολύ πιθανή την ύπαρξη υπερπροσαρμογής. Έτσι αποφασίστηκε να προστεθεί ένας αριθμός δειγμάτων σε όλες τις μειοψηφικές τάξεις επιλέγοντας τελικά 1000 φωτογραφίες από κάθε τάξη. Σε περίπτωση που κάποια τάξη δεν είχε τόσο μεγάλο πλήθος διαθέσιμων φωτογραφιών εφαρμόστηκε η τεχνική επαύξησης δεδομένων.

Συνοπτικά η τεχνική της επαύξησης δεδομένων η οποία εφαρμόστηκε δυναμικά κατά την διάρκεια της εκπαίδευσης, βοήθησε να δημιουργηθούν τεχνητά δείγματα για τις μειοψηφικές τάξεις ώστε να εξομαλυνθεί η άνιση κατανομή των διαθέσιμων δεδομένων. Η διαδικασία επαύξησης και οι μετατροπές που πραγματοποιήθηκαν ώστε να παραχθούν τα τεχνητά δείγματα, αναφέρονται αναλυτικά παρακάτω στο κεφάλαιο 4.4 που αφορά την εκπαίδευση των νευρωνικών δικτύων. Η διαδικασία επεξεργασίας της εικόνας ξεκινά με την δημιουργία εικόνων για κάθε κατηγορία σπόρων και επισήμανση των σπορών που υπάρχουν σε αυτές, οι οποίες στην συνέχεια εξετάστηκαν διεξοδικά από έναν ειδικό στην γεωργία ώστε να διασφαλιστεί η ακρίβεια των ετικετών.

Στη συνέχεια, οι εικόνες επεξεργάζονται ώστε να μετασχηματιστούν στις κατάλληλες διαστάσεις εισόδου που χρειάζεται το κάθε μοντέλο. Πιο συγκεκριμένα για τα μοντέλα ResNet50 και VGG16 οι εικόνες μετασχηματίζονται σε μέγεθος 224x224 ενώ για το Xception σε 299x299, ενώ πάντοτε διατηρείται η αρχική αναλογία των πραγματικών τους διαστάσεων. Τέλος για την αξιολόγηση της ταξινόμησης χρησιμοποιήθηκαν όλες οι εικόνες του υποσυνόλου δοκίμων, χωρίς να ισορροπηθούν ώστε τα αποτελέσματα να αξιολογηθούν με βάση τις συνθήκες του πραγματικού κόσμου όπου η ανισορροπία των συνόλων δεδομένων είναι συχνή [54].





**Εικόνα 6: Οι κατανομές πλάτους και ύψους των εικόνων του συνόλου δεδομένων σε pixel**

### 4.3 Τα Μοντέλα Που Δοκιμάστηκαν

Για την επίλυση του προβλήματος ταξινόμησης των διαφορετικών ειδών και ποικιλιών σπόρων δημητριακών, δοκιμάστηκαν τρία διαφορετικά μοντέλα. Αυτά είναι: Το VGG (Visual Geometry Group), το ResNet50 και το Xception και θα παρουσιαστούν αναλυτικά παρακάτω.

#### 4.3.1 Το μοντέλο VGG16

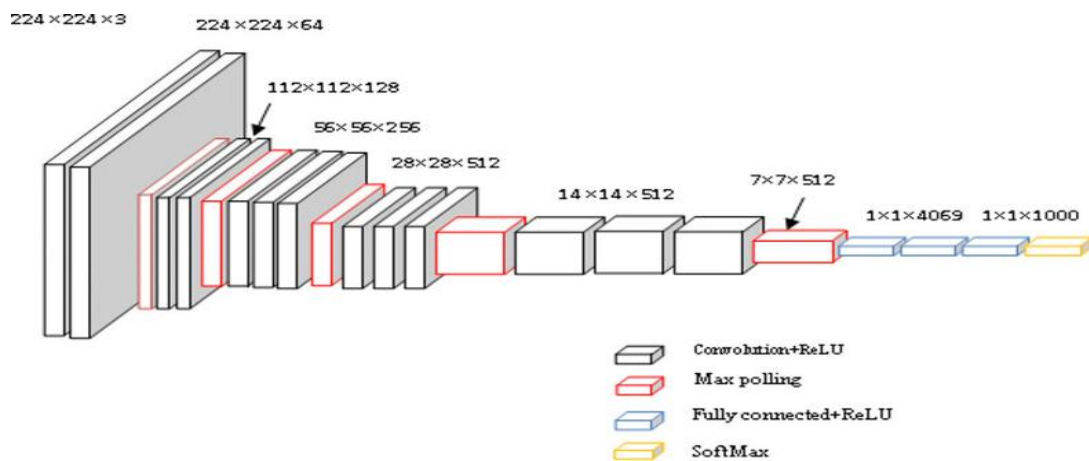
Το Συνελκτικό Νευρωνικό Δίκτυο (Convolutional Neural Network - CNN) της Ομάδας Οπτικής Γεωμετρίας (Visual Geometry Group - VGG) είναι ένας αλγόριθμος βαθιάς μάθησης που αναπτύχθηκε από την Ομάδα Οπτικής Γεωμετρίας του Πανεπιστημίου της Οξφόρδης. Αυτό το μοντέλο βαθιάς μάθησης έγινε διάσημο αφού επέτυχε εξαιρετικά αποτελέσματα στο ImageNet Large Scale Visual Recognition Challenge (ILSVRC) του 2014 [24].

Η βασική αρχιτεκτονική του VGG χρησιμοποιεί μικρά αντιδραστικά πεδία (φίλτρα συνέλιξης  $3 \times 3$ ), επιτρέποντάς του να μάθει πιο πολύπλοκα μοτίβα με λιγότερες παραμέτρους. Τοποθετώντας σε στήβες αυτά τα μικρά συνελεκτικά στρώματα, το VGG μπορεί να προσεγγίσει αποτελεσματικά μεγαλύτερα αντιδραστικά πεδία (όπως  $5 \times 5$  ή  $7 \times 7$ ), αλλά με λιγότερες παραμέτρους και βαθύτερες δομές.

Υπάρχουν διάφορες παραλλαγές του VGG, τα πιο συνηθισμένα είναι τα VGG16 και VGG19, τα οποία διακρίνονται από τον αριθμό των στρωμάτων στις αρχιτεκτονικές τους. Παρά το γεγονός ότι έχουν περισσότερο βάθος από προηγούμενα μοντέλα όπως το AlexNet [37], τα μοντέλα VGG παραμένουν σχετικά απλά και ομοιόμορφα, αποτελούμενα κυρίως από στοιβαγμένα στρώματα συνέλιξης  $3 \times 3$  ακολουθούμενα από μια λειτουργία max pooling, όλα τα οποία χρησιμοποιούν την συνάρτηση ενεργοποίησης ReLU (Rectified Linear Unit).

Ένα από τα κύρια οφέλη του μοντέλου VGG είναι οι εξαιρετικές δυνατότητες μεταφοράς μάθησης. Τα προ-εκπαιδευμένα μοντέλα VGG χρησιμοποιούνται συχνά ως

εξαγωγείς χαρακτηριστικών σε διάφορες εφαρμογές όρασης λόγω της ικανότητάς τους να γενικεύουν καλά σε άλλες εργασίες. Ωστόσο, αξίζει να σημειωθεί ότι τα μοντέλα VGG έρχονται με υψηλό υπολογιστικό κόστος. Είναι μνήμης-εντατικά λόγω του αριθμού των παραμέτρων που εμπλέκονται, ιδιαίτερα στα πλήρως συνδεδεμένα στρώματα προς το τέλος του δικτύου. Αυτό το καθιστά λιγότερο εφικτό για ανάπτυξη σε εφαρμογές σε πραγματικό χρόνο ή σε συσκευές με περιορισμένους υπολογιστικά πόρους.



**Εικόνα 7: Η βασική αρχιτεκτονική του μοντέλου VGG16.**

[https://www.researchgate.net/figure/VGG-Neural-Network-Architecture-Source\\_fig5\\_369334711](https://www.researchgate.net/figure/VGG-Neural-Network-Architecture-Source_fig5_369334711)

#### 4.3.2 Το μοντέλο ResNet50

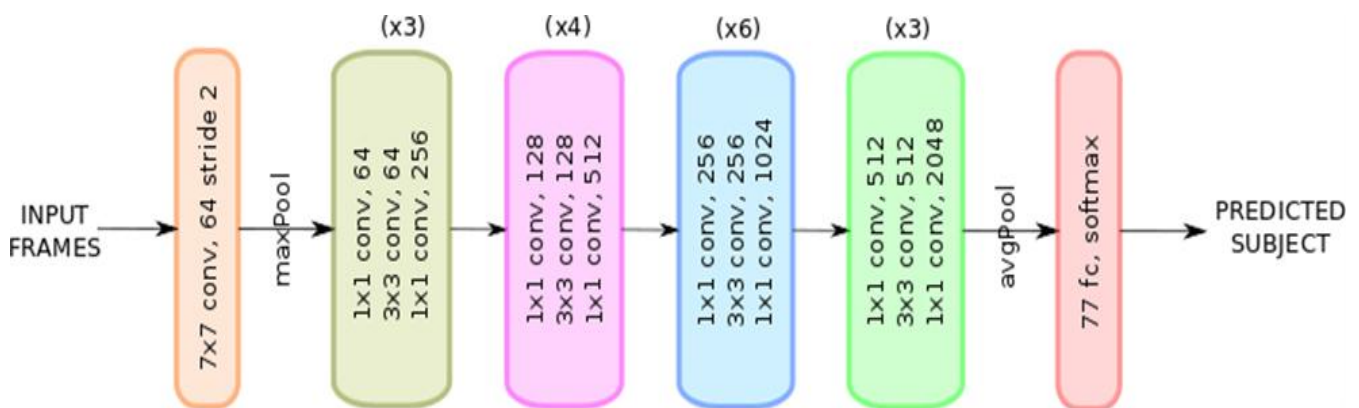
Το Residual Network (ResNet) είναι ένα Συνελικτικό Νευρωνικό Δίκτυο (CNN) που αναπτύχθηκε από την Microsoft Research. Ειδικότερα, το ResNet50 αναφέρεται στην παραλλαγή της αρχιτεκτονικής ResNet που έχει 50 στρώσεις. Το ResNet κέρδισε διεθνή αναγνώριση για την επίδοσή του στο ImageNet Large Scale Visual Recognition Challenge (ILSVRC) του 2015 [25].

Το μοναδικό χαρακτηριστικό της αρχιτεκτονικής ResNet είναι η χρήση των "ενωτικών διασυνδέσεων" ή "διασυνδέσεων συντομεύσεων", οι οποίες επιτρέπουν την άμεση αντίστροφη διάδοση του διανύσματος κλίσης στα προηγούμενα στρώματα. Αυτές οι συνδέσεις προστίθενται στην έξοδο των στρωμάτων συνέλιξης, βοηθώντας να αντιμετωπιστεί το πρόβλημα των εξαφανιζόμενων κλίσεων (vanishing gradients) που συχνά συμβαίνει σε βαθιά δίκτυα. Αυτό, με τη σειρά του, καθιστά δυνατή την εκπαίδευση πολύ βαθιών δικτύων, με εκδόσεις ResNet που φτάνουν έως και τα 152 στρώματα.

Όσον αφορά τη δομή, ένα δίκτυο ResNet50 περιλαμβάνει συνήθως ένα στρώμα συνέλιξης και ένα στρώμα μέγιστης συγκέντρωσης, ακολουθούμενο από τέσσερα στάδια που αποτελούνται από διάφορες δομές bottleneck (ένα τριστρωματικό block που περιλαμβάνει συνέλιξεις 1x1, 3x3 και 1x1). Το δίκτυο τελειώνει με έναν παγκόσμιο μέσο υπολογισμό, ένα πλήρως συνδεδεμένο στρώμα και μια συνάρτηση softmax [25].

Το κύριο πλεονέκτημα του ResNet50 είναι η ικανότητά του να γενικεύει καλά σε άλλες εργασίες, καθιστώντας το δημοφιλές επιλογή για εφαρμογές μεταφοράς μάθησης. Παρά το βάθος τους, τα μοντέλα ResNet μπορούν να εκπαιδευτούν σε λογικό χρονικό διάστημα λόγω των υπολειμματικών τους μπλοκ και της αποδοτικής αντίστροφης διάδοσης. Η εισαγωγή της υπολειμματικής μάθησης στην οποία στηρίχτηκε η δημιουργία του μοντέλου ήταν καθοριστική για την προώθηση του τομέα της βαθιάς μάθησης, επιτρέποντας την εκπαίδευση ολοένα και βαθύτερων δικτύων και βελτιώνοντας την ακρίβεια σε ποικίλες εργασίες.

Παρόλα αυτά, όπως και τα περισσότερα μοντέλα βαθιάς μάθησης, το ResNet50 μπορεί να είναι υπολογιστικά έντονο και απαιτεί σημαντική ποσότητα μνήμης, ειδικά για εισόδους υψηλής ανάλυσης. Αυτό θα μπορούσε να περιορίσει τη χρήση του σε περιβάλλοντα με περιορισμένους πόρους ή σε εφαρμογές πραγματικού χρόνου.



**Εικόνα 8: Η βασική αρχιτεκτονική του μοντέλου ResNet50.**

[https://www.researchgate.net/figure/ResNet-50-neural-network-architecture-56\\_fig4\\_336805103](https://www.researchgate.net/figure/ResNet-50-neural-network-architecture-56_fig4_336805103)

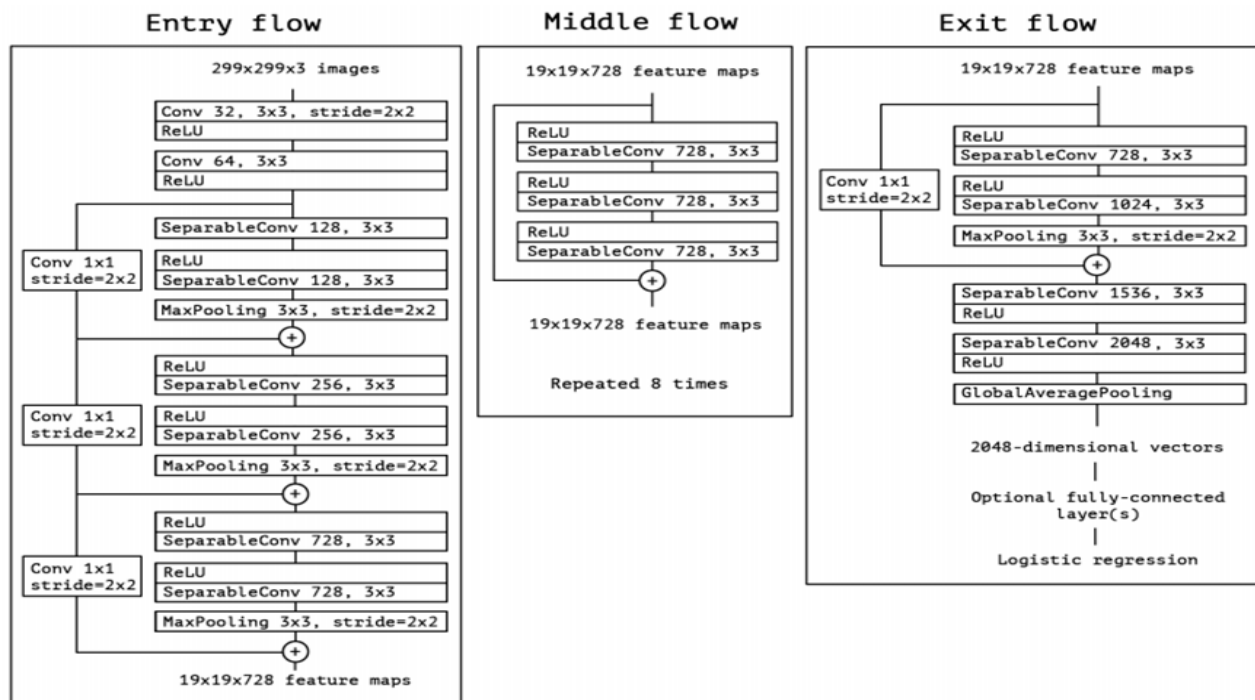
#### 4.3.3 Το μοντέλο Xception

Το μοντέλο Xception είναι μια προσαρμογή του μοντέλου Inception-v3, μια αρχιτεκτονική Convolutional Neural Network (CNN) που προτάθηκε από ερευνητές της Google [55]. Αυτή η αρχιτεκτονική, βασισμένη στο προγενέστερο μοντέλο Inception (γνωστό και ως GoogLeNet), έχει επιτύχει εντυπωσιακή επίδοση σε εργασίες ταξινόμησης εικόνων, συμπεριλαμβανομένης της ImageNet Large Scale Visual Recognition Challenge (ILSVRC).

Οι αρχιτεκτονικές τύπου Inception είναι σχεδιασμένες γύρω από την έννοια ενός "Inception module" και ακολουθούν μια διαφορετική προσέγγιση, η οποία επιτρέπει την αύξηση του «πλάτους» του δικτύου. Το "Inception module" προσπαθεί να υπολογίσει πολλαπλά διαφορετικά επίπεδα τα οποία παράλληλα, βρίσκονται στον ίδιο «χάρτη εισόδου». Με αυτό τον τρόπο δημιουργείται μια ποικιλία διαφορετικών λειτουργιών συνέλιξης σε κάθε στρώμα, με την έξοδο αυτών των λειτουργιών να συνενώνεται μαζί. Επομένως, προσφέρει το πλεονέκτημα της αύξησης των διαστάσεων του δικτύου χωρίς αντίστοιχη αύξηση της υπολογιστικής πολυπλοκότητας. Αυτή η προσέγγιση αναφέρεται συχνά ως σχεδιασμός "δικτύου-μέσα-σε-δίκτυο".

Η ειδοποιός διαφορά μεταξύ Xception και Inception βρίσκεται στην αντικατάσταση των modules σε διαχωριζόμενες παράλληλες συνελίξεις. Πιο συγκεκριμένα, η διαδικασία

Ξεκινά με τον υπολογισμό των χωρικών συσχετίσεων σε κάθε κανάλι ανεξάρτητα από τα αλλά, η οποία όμως τελικά καταλήγει σε μια συνέλιξη  $1 \times 1$  ώστε να αποτυπωθεί με ακρίβεια η διακαναλική συσχέτιση. Στη συνέχεια, αυτές χρησιμοποιούνται όχι μόνο για την εκτέλεση μείωσης των χωρικών διαστάσεων (βελτιώνοντας έτσι την υπολογιστική αποδοτικότητα), αλλά και για την εισαγωγή μη γραμμικότητας στο μοντέλο. Αυτό συμβάλλει στην εκφραστική δύναμη του μοντέλου χωρίς να αυξάνει σημαντικά το υπολογιστικό κόστος. Επιπλέον έχει την δυνατότητα να υλοποιεί την εξομάλυνση των ετικετών, μια τεχνική που βοηθάει στην αποφυγή της υπερβολικής βεβαιότητας του μοντέλου στις προβλέψεις του, πράγμα που μπορεί να οδηγήσει σε υπερπροσαρμογής (over-fitting).



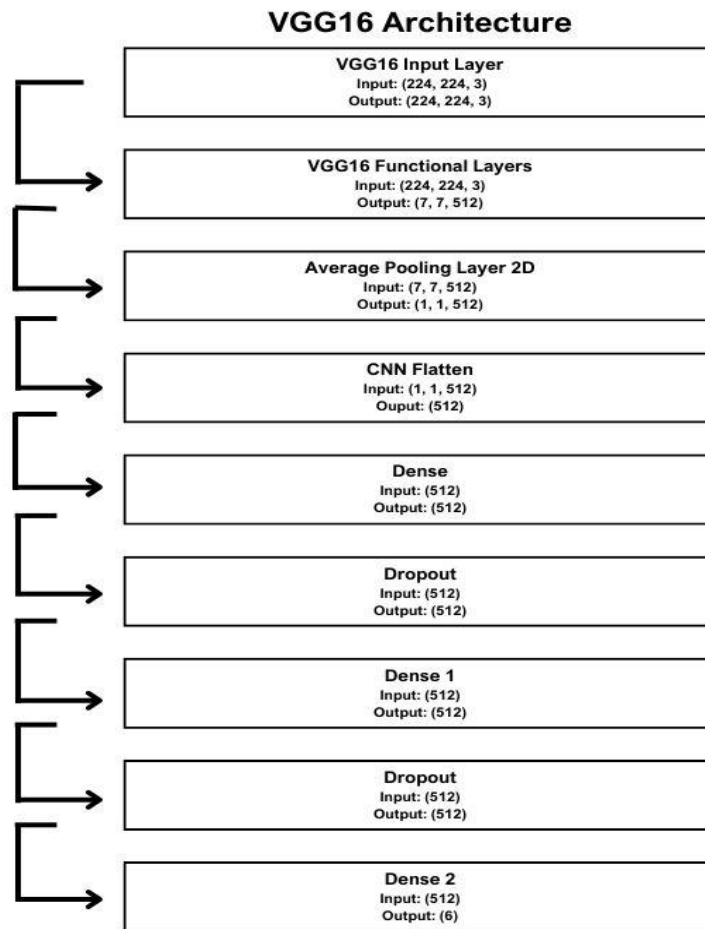
**Εικόνα 9: Η αρχιτεκτονική του μοντέλου Xception.**

<https://towardsdatascience.com/xception-from-scratch-using-tensorflow-even-better-than-inception-940fb231ced9>

#### 4.4 Εκπαίδευση Των Νευρωνικών Δικτύων

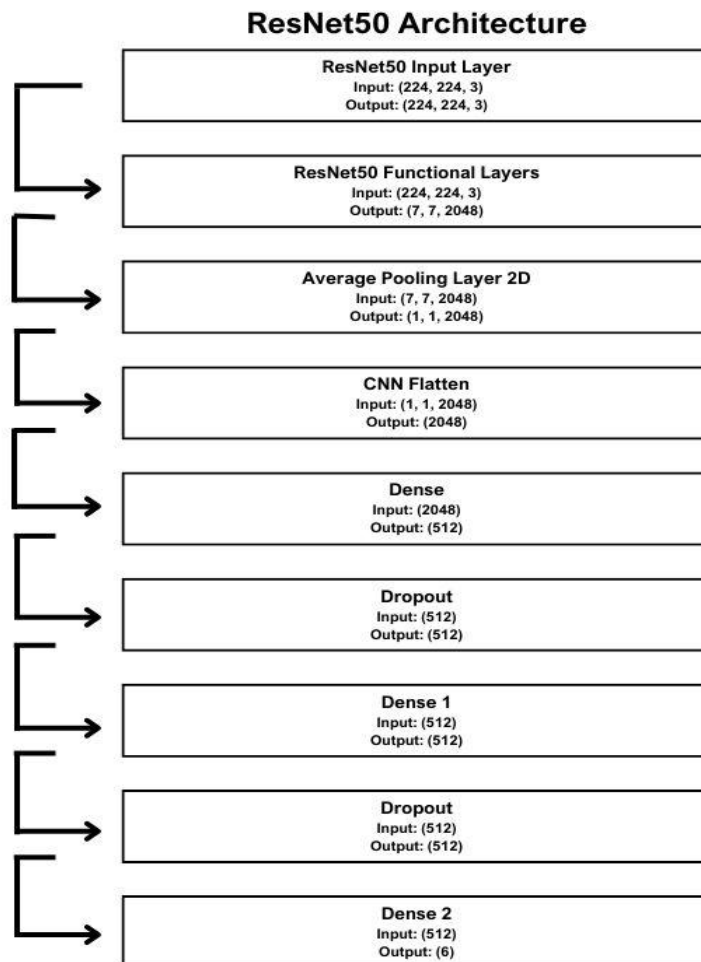
Η εκπαίδευση των νευρωνικών δικτύων πραγματοποιήθηκε με την γλώσσα προγραμματισμού Python version 3.11.0 χρησιμοποιώντας την βιβλιοθήκη Keras 2.11.0 και Tensorflow 2.11.0 ως backend. Τα δίκτυα που δοκιμάστηκαν ήταν προ εκπαιδευμένα χρησιμοποιώντας τα βάρη του συνόλου δεδομένων ImageNet χωρίς όμως να συμπεριλαμβάνουν τα τελευταία επίπεδα ταξινόμησης. Στα υπόλοιπα στρωματά των

τριών διαφορετικών μοντέλων βαθιάς μάθησης που δοκιμάστηκαν, έγινε μία ακόμα προσαρμογή λόγω ότι το κάθε δίκτυο έχει διαφορετικό μέγεθος εξόδου πριν τα τελικά στρώματά ταξινόμησης. Για το VGG16 η διάσταση εξόδου πριν τα στρώματά ταξινόμησης είναι (7, 7, 512) για το Resnet50 (7, 7, 2048) ενώ για το Xception (10, 10, 2048). Επομένως τοποθετήθηκαν δυο ενδιάμεσα στρώματά ένα Average Pooling 2D που μετατρέπει τις διαστάσεις σε (1, 1, 512) για το VGG16 και (1, 1, 2048) για το ResNet50 και το Xception και ένα στρώμα Flatten CNN που η τελική του έξοδος είναι ένας μονοδιάστατος πίνακας για κάθε δείγμα, που τροφοδοτείται στα πλήρως συνδεδεμένα στρώματα ταξινόμησης.



**Εικόνα 10: Η παραλλαγή του μοντέλου VGG16.**

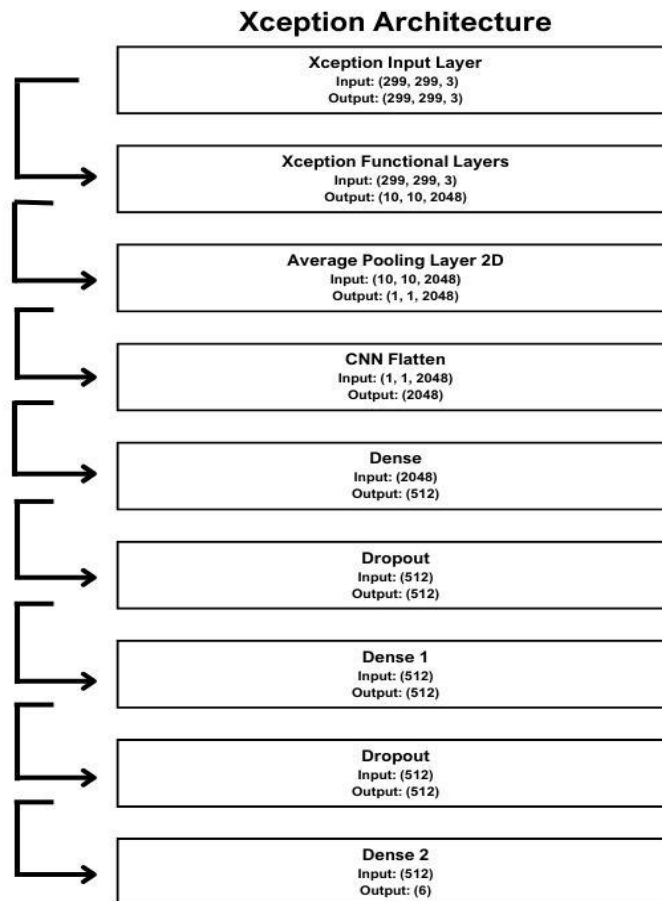
Η ειδοποιός διαφορά εντοπίζεται στα δυο τελευταία επίπεδα των προ εκπαιδευμένων μοντέλων όπου δεν χρησιμοποιήθηκαν τα στρώματα από τα δυο τελευταία επίπεδα ταξινόμησης για το σύνολο δεδομένων ImageNet. Τα τελευταία επίπεδα και στις τρεις προ εκπαιδευμένες παραλλαγές των μοντέλων αποτελούνται από δυο πλήρως συνδεδεμένα στρώματα τα οποία απαρτίζονται από 512 νευρώνες το καθένα. Για τα επίπεδα αυτά επιλέχθηκε η συνάρτηση ενεργοποίησης ReLU (Rectified Linear Unit) και προστέθηκαν 2 Dropout layers της τάξεως του 50% ώστε να αποφευχθεί το ενδεχόμενο υπερπροσαρμογής (overfitting). Τα δίκτυα εκπαιδεύτηκαν σε ένα laptop με κάρτα γραφικών Ge-Force GTX 1660-ti η οποία έχει μνήμη (VRAM) 6GB. Ως αλγόριθμος βελτιστοποίησης (optimizer) επιλέχθηκε ο αλγόριθμος Adam διότι ένα από τα βασικά του χαρακτηριστικά είναι ότι επιτρέπει την προσαρμογή του ρυθμού μάθησης των παραμέτρων του μοντέλου, γεγονός που μπορεί να οδηγήσει σε ταχύτερο αποτέλεσμα ταξινόμησης, ενώ για συνάρτηση απώλειας επιλέχθηκε η categorical cross entropy.



**Εικόνα 11: Η παραλλαγή του μοντέλου ResNet50.**

Όπως αναφέρθηκε και στην ενότητα 4.2 για να αντιμετωπιστεί η μεγάλη ανισορροπία του συνόλου δεδομένου και να ενισχυθεί η ποικιλομορφία, εφαρμοστήκαν μεταμορφώσεις στις εικόνες δημιουργώντας δεδομένα τεχνητά. Για την επαύξηση δεδομένων χρησιμοποιήθηκε η κλάση ImageDataGenerator της βιβλιοθήκης Keras. Σε πρώτη φάση μια εικόνα φορτώνεται στον αλγόριθμο και μετατρέπεται σε πίνακα numpy. Η διαδικασία της επαύξησης ξεκινά εφαρμόζοντας προσαρμογές στην φωτεινότητα των εικόνων εντός του εύρους [0.5, 1.2]. Τροποποιώντας τη φωτεινότητα των εικόνων γίνεται μια απόπειρα να προσομοιωθούν οι πραγματικές συνθήκες, όπου συνήθως εμφανίζονται μεταβολές στις συνθήκες φωτισμού.

Στη συνέχεια, χρησιμοποιώντας έναν επαναληπτικό βρόγχο δημιουργούνται τεχνητά δείγματα για τις μειωθηφικές κατηγορίες εκτός από τις μεταμορφώσεις της φωτεινότητας, εφαρμόζονται επιπλέον, περιστροφή των εικόνων, και μεγέθυνση ή σμίκρυνση ώστε να υπάρξει ποικιλομορφία και να ενισχυθεί η ποιότητα του συνόλου δεδομένων και να επέλθει ισορροπία κατά την διαδικασία της εκπαίδευσης. Πιο συγκεκριμένα, ο επαναληπτικός βρόγχος εκτελείται 9 φορές εφαρμόζοντας περιστροφή έως 120 μοίρες και μεγέθυνση ή σμίκρυνση έως 20% του πραγματικού μεγέθους της εικόνας.



**Εικόνα 12: Η παραλλαγή του μοντέλου Xception.**

Στο τελικό στάδιο της διαδικασίας το μέγεθος όλων των εικόνων που ανήκουν στο σύνολο εκπαίδευσης, μετατρέπεται στο κατάλληλο μέγεθος pixel ώστε να ξεκινήσει η διαδικασία της εκπαίδευσης.

Για το μοντέλο VGG και ResNet50 οι εικόνες εισόδου μετασχηματίζονται στο μέγεθος 224x224 ενώ για το μοντέλο Xception: 299x299. Μετά την ολοκλήρωση της διαδικασίας της επαύξησης δεδομένων, οι εικόνες αποθηκεύονται στους κατάλληλους φακέλους ανάλογα με το είδος και την ποικιλία κάθε σπόρου, ώστε να είναι έτοιμες για χρήση. Η μετατροπή του μεγέθους των εικόνων γίνεται με την βοήθεια της βιβλιοθήκης OpenCV, οι οποίες στη συνέχεια τροφοδοτούνται στο νευρωνικό δίκτυο ώστε να εκπαιδευτεί.

Για τις παραλλαγές των μοντέλων Xception και ResNet50 επιλέχθηκε ρυθμός μάθησης ίσος με  $1 \times 10^{-3}$ , ωστόσο για την παραλλαγή του μοντέλου VGG επιλέχθηκε ρυθμός μάθησης ίσος με  $1 \times 10^{-4}$ . Όλες οι παραλλαγές των μοντέλων εκπαιδεύτηκαν χρησιμοποιώντας batch size 32, δηλαδή το κάθε μοντέλο εκπαιδεύτηκε λαμβάνοντας κάθε φορά ως αριθμό δειγμάτων 32 εικόνες μέχρι να εξαντληθούν τα διαθέσιμα δείγματα.

Προκειμένου να γίνει ενδελεχής αξιολόγηση της ικανότητας των μοντέλων να ταξινομήσουν με μεγάλη ακρίβεια τους σπόρους δημητριακών, κάθε μοντέλο εκπαιδεύτηκε χρησιμοποιώντας την μέθοδο της διασταυρουμένης επικύρωσης 5 φορές (5-fold cross-validation). Αυτή η τεχνική περιλαμβάνει τη διαίρεση του συνόλου δεδομένων σε πέντε ίσα υποσύνολα, ή "πτυχές". Σε κάθε επανάληψη, μία πτυχή χρησιμοποιείται ως σύνολο επικύρωσης, ενώ οι υπόλοιπες τέσσερις πτυχές χρησιμοποιούνται για την εκπαίδευση.

Στην παρούσα διατριβή το σύνολο δεδομένων εκπαίδευσης χωρίζεται κάθε φορά σε 5 υποσύνολα όπου το καθένα περιέχει περίπου 1000 φωτογραφίες συμπεριλαμβανομένων των τεχνητών δειγμάτων που δημιουργήθηκαν από την επαύξηση δεδομένων. Λόγω του πολύ μεγάλου πλήθους της ποικιλίας Λεβαντε και τις 5 φορές που εκπαιδεύτηκε το μοντέλο, υπήρχαν περισσότερες εικόνες από την συγκεκριμένη κατηγορία. Η προσπάθεια εξισορρόπησης των δεδομένων επικεντρώθηκε περισσότερο στο να μην υπάρχει πολύ μεγάλη ανισορροπία μεταξύ της κατηγορίας Λεβάντε και των άλλων ποικιλιών. Επομένως, σε αρκετές περιπτώσεις εκπαίδευσης η κατηγορία Λεβαντε περιείχε περίπου 250-350 εικόνες παραπάνω. Σε αντίθεση με την παραδοσιακή μέθοδο, στην παρούσα εφαρμογή, το σύνολο επικύρωσης διαχωρίστηκε εκ των προτέρων διατηρώντας ένα σταθερό πλήθος 2085 φωτογραφιών διαχωρισμένα σε 5 υποσύνολα. Η απόδοση του μοντέλου στη συνέχεια υπολογίζεται ως ο μέσος όρος των αποτελεσμάτων και των πέντε επαναλήψεων, παρέχοντας μια πιο αξιόπιστη εκτίμηση της ακρίβειάς του σε δεδομένα που δεν παρουσιάζουν ισορροπία ούτε έχουν χρησιμοποιηθεί ξανά στην εκπαιδευτική διαδικασία. Αυτή η παραλλαγμένη προσέγγιση της παραδοσιακής μεθόδου διασταυρουμένης επικύρωσης, εξασφαλίζει ότι η αξιολόγηση πραγματοποιείται σε διάφορα υποσύνολα των δεδομένων, παρέχοντας έτσι μια ολοκληρωμένη κατανόηση της αποτελεσματικότητας του μοντέλου σε συνθήκες παρόμοιες με τον πραγματικό κόσμο.

Η διαδικασία εκπαίδευσης κάθε επανάληψης πραγματοποιήθηκε μέχρι η ακρίβεια του συνόλου επικύρωσης (validation dataset) να μην δείξει καμία βελτίωση για 100 συνεχόμενες εποχές. Με αυτό τον τρόπο εξασφαλίζεται ότι κατά την διάρκεια της εκπαίδευσης, κάθε φωτογραφία θα έχει χρησιμοποιηθεί τουλάχιστον μια φορά σε αυτές τις εποχές και ότι τα δίκτυα όντως έχουν πραγματοποιήσει την μέγιστη δυνατή σύγκλιση. Οι γενικές πληροφορίες για τα μοντέλα όπως οι παράμετροι της αρχιτεκτονικής, ο χρόνος εκπαίδευσης ανά εποχή και το μέγεθος εισόδου συνοψίζονται στον πίνακα 2.



	<b>VGG16</b>	<b>ResNet50</b>	<b>Xception</b>
Mean Time Per Epoch (Seconds)	144	148	229
Minimum Epochs Used	389	502	419
Maximum Epochs Used	769	967	789
Total Network Parameters	27.829.068	24.905.612	22.179.380
Network Depth - Layers	16	50	71
Input Image Size in pixels	224x224	224x224	299x299
Batch Size	32	32	32

**Πίνακας 2: Γενικές πληροφορίες για τις παραλλαγές των τριών μοντέλων που δοκιμάστηκαν**

#### 4.5 Μετρικές Αξιολόγησης

Για την αξιολόγηση των αποτελεσμάτων ταξινόμησης του κάθε μοντέλου χρησιμοποιήθηκαν οι τρεις πιο κοινές και ευρέως υιοθετημένες μετρικές, οι οποίες είναι η ακρίβεια (precision), η ανάκληση (recall) και το f1-score [56]. Η ακρίβεια ορίζεται ως ο λόγος των αληθών θετικά τιμών ως προς το άθροισμα των αληθών θετικά τιμών και των ψευδώς θετικών τιμών. Οι αληθείς θετικές τιμές αντιπροσωπεύουν σπόρους οι οποίοι ταξινομήθηκαν σε ένα συγκεκριμένο είδος και όντως ανήκουν σε αυτό. Ενώ, οι ψευδείς θετικές τιμές ως οι σπόροι που δεν ανήκουν στο συγκεκριμένο είδος, αλλά το μοντέλο λανθασμένα της ταξινόμησε σε αυτό.

Η ανάκληση χρησιμοποιείται για την αξιολόγηση της ευαισθησίας του κάθε μοντέλου. Όπως και στην μετρική της ακρίβειας, η ανάκληση είναι ένας λόγος, όπου ο αριθμητής αναπαριστά τις αληθώς θετικά τιμές, ενώ ο παρονομαστής τις ψευδώς αρνητικές τιμές. Σε αντίθεση με τις ψευδώς θετικές τιμές, οι ψευδώς αρνητικές τιμές αναπαριστούν τους σπόρους οι οποίοι ανήκουν στο είδος αλλά δεν ταξινομήθηκαν σε αυτό. Πρακτικά, η ανάκληση υποδηλώνει ποσό καλά το μοντέλο μπορεί να ταξινομήσει τις εικόνες σπορών που όντως ανήκουν στην ποικιλία.

Τέλος, το f1-score είναι ένας αρμονικός μέσος ο οποίος κυμαίνεται από το 0 έως το 1 και απεικονίζει την αναλογία μεταξύ ακρίβειας και ανάκλησης. Το f1-score είναι ιδιαίτερα χρήσιμο σε προβλήματα ταξινόμησης πολλών τάξεων καθώς υποδηλώνει την ισορροπία μεταξύ ψευδών θετικών και ψευδώς αρνητικών τιμών. Όσο πιο κοντά στο 1 βρίσκεται η τιμή του f1 score τόσο πιο καλή και ισορροπημένη είναι η αναλογία μεταξύ ακρίβειας και ανάκλησης.

$$\text{Precision} \quad \text{Precision} = \frac{tp}{tp + fp}$$

$$\text{Recall} \quad \text{Recall} = \frac{tp}{tp + fn}$$

$$\text{F1-score} \quad \text{F1-score} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

**Εικόνα 13: Μαθηματικοί τύποι των μετρικών αξιολόγησης**

<https://www.kaggle.com/discussions/getting-started/351680>

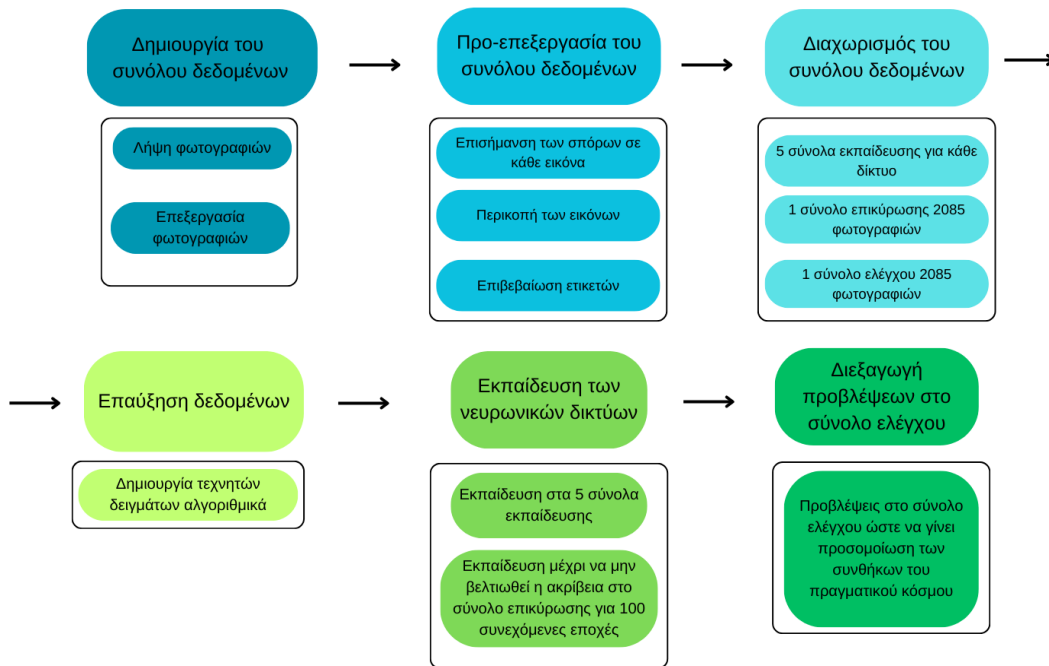
## 5. Αποτελέσματα

Στην ενότητα αυτή παρουσιάζονται τα αποτελέσματα της εφαρμογής των τριών μοντέλων τα οποία χρησιμοποιήθηκαν για την ταξινόμηση έξι διαφορετικών ποικιλιών από τρία διαφορετικά δημητριακά, οδηγώντας σε έξι διαφορετικές κατηγορίες προς ταξινόμηση. Στο πλαίσιο αυτής της έρευνας, δόθηκε έμφαση στην ανάλυση της αποτελεσματικότητας και της ακρίβειας των μοντέλων κατά την ταξινόμηση των ειδών στο σύνολο ελέγχου με βάση τα χαρακτηριστικά τους. Άλλωστε η ανάλυση των αποτελεσμάτων δίνοντας ιδιαίτερη βαρύτητα στο σύνολο ελέγχου προσφέρει χρήσιμα συμπεράσματα για την απόδοση των μοντέλων σε πραγματικές συνθήκες.

Variety	Model		
	ResNet50	VGG	Xception
KTR	0.91	0.50	0.88
DMT	0.92	0.28	0.88
VRF	0.97	0.45	0.97
VPG	1.00	0.21	0.98
SSS	0.95	0.42	0.94
LVT	1.00	0.82	1.00

**Πίνακας 3: Συγκεντρωτικά αποτελέσματα της ακρίβειας των μοντέλων για κάθε ποικιλία στο σύνολο ελέγχου.**

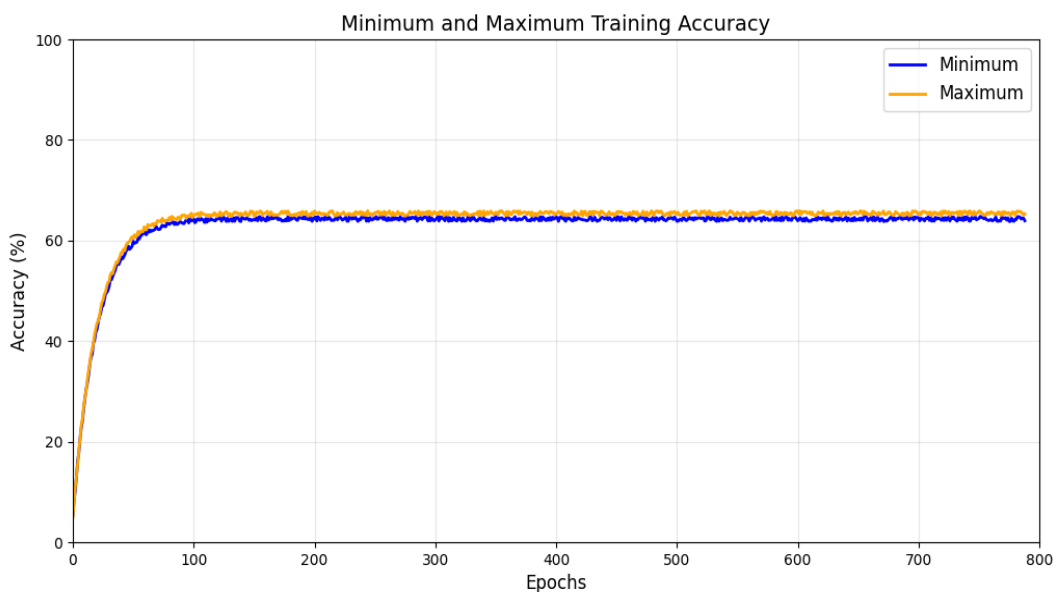
Διάγραμμα Ροής της διαδικασίας ανάπτυξης της εφαρμογής



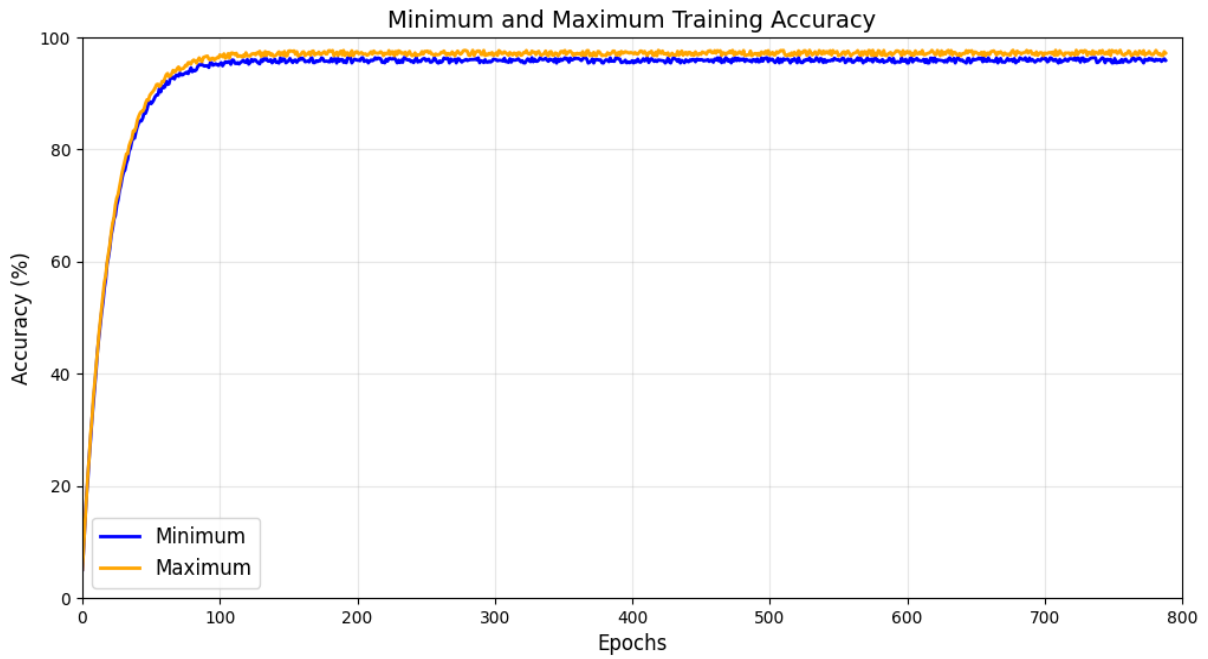
**Εικόνα 14: Διάγραμμα ροής της διαδικασίας ανάπτυξης της προτεινόμενης εφαρμογής**

## 5.1 Ακρίβεια Και Απώλεια

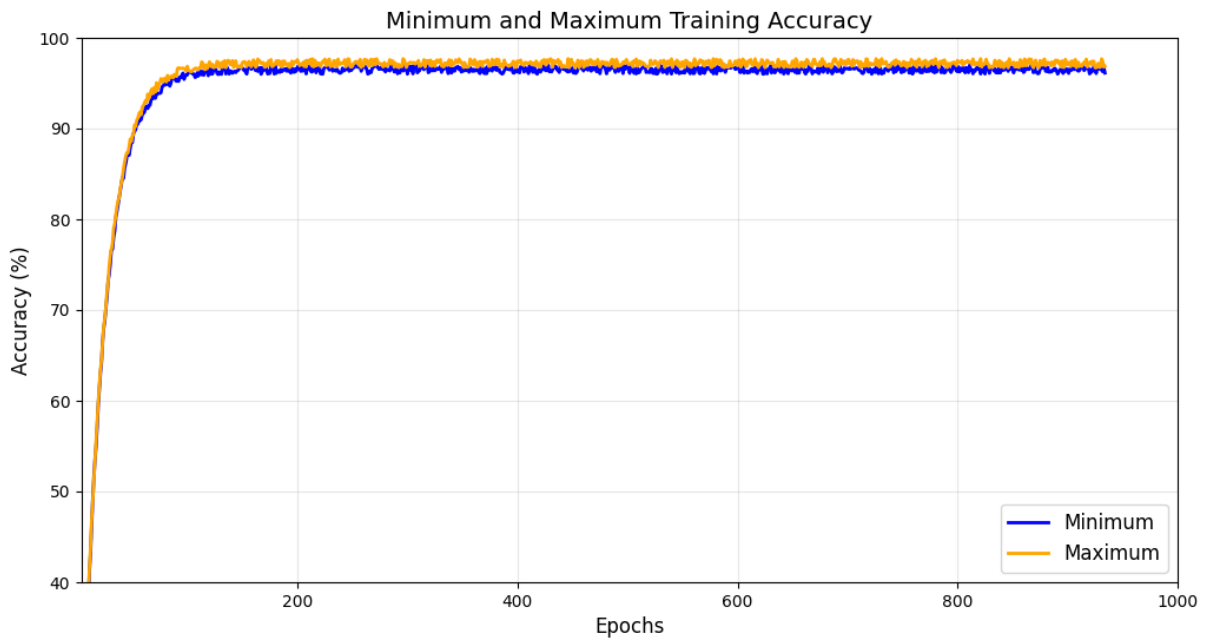
Η ακρίβεια των μοντέλων κατά την διάρκεια της εκπαίδευσης αρχικά αυξάνεται ραγδαία, ειδικότερα τις πρώτες 100 εποχές και εν συνέχεια, συγκλίνει σταδιακά προς το βέλτιστο. Η ταχύτητα σύγκλισης προς το βέλτιστο παρουσιάζει μικρές διαφορές ανάλογα με το μοντέλο που δοκιμάζεται, χωρίς να παρουσιάζονται διαφορές στην ακρίβεια. Οι παραλλαγές των μοντέλων ResNet50 και Xception σημείωσαν ελάχιστα παραπάνω από 97% ακρίβεια στο σύνολο ελέγχου ενώ η παραλλαγή του μοντέλου VGG16 παρουσίασε ακρίβεια 65% στο ίδιο υποσύνολο.



**Εικόνα 15: Η μέγιστη και ελάχιστη ακρίβεια κατά την διάρκεια της εκπαίδευσης του μοντέλου VGG16.**



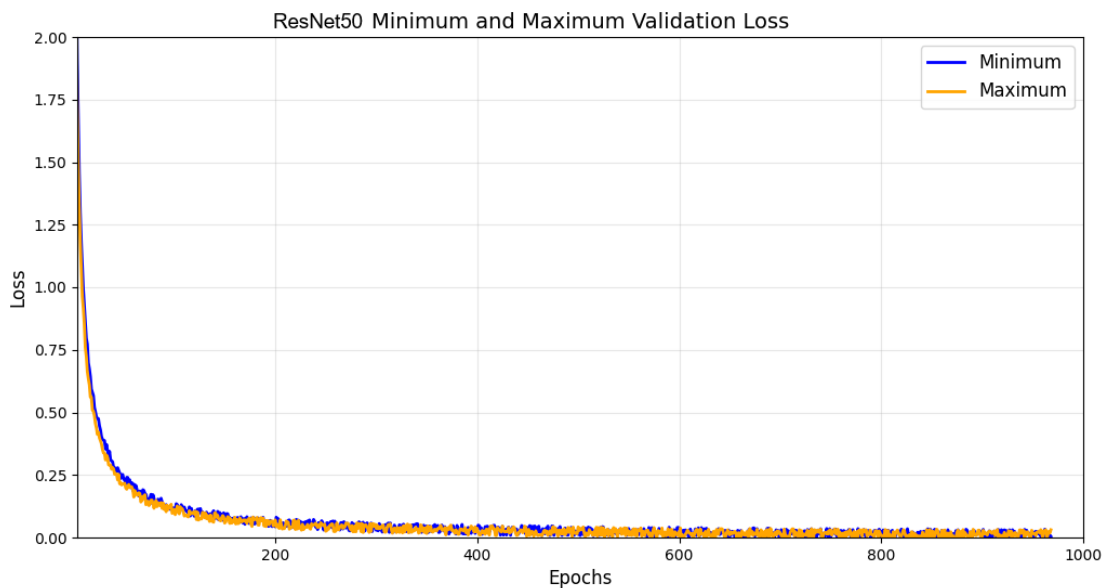
**Εικόνα 16: Η μέγιστη και ελάχιστη ακρίβεια κατά την διάρκεια της εκπαίδευσης του μοντέλου Xception.**



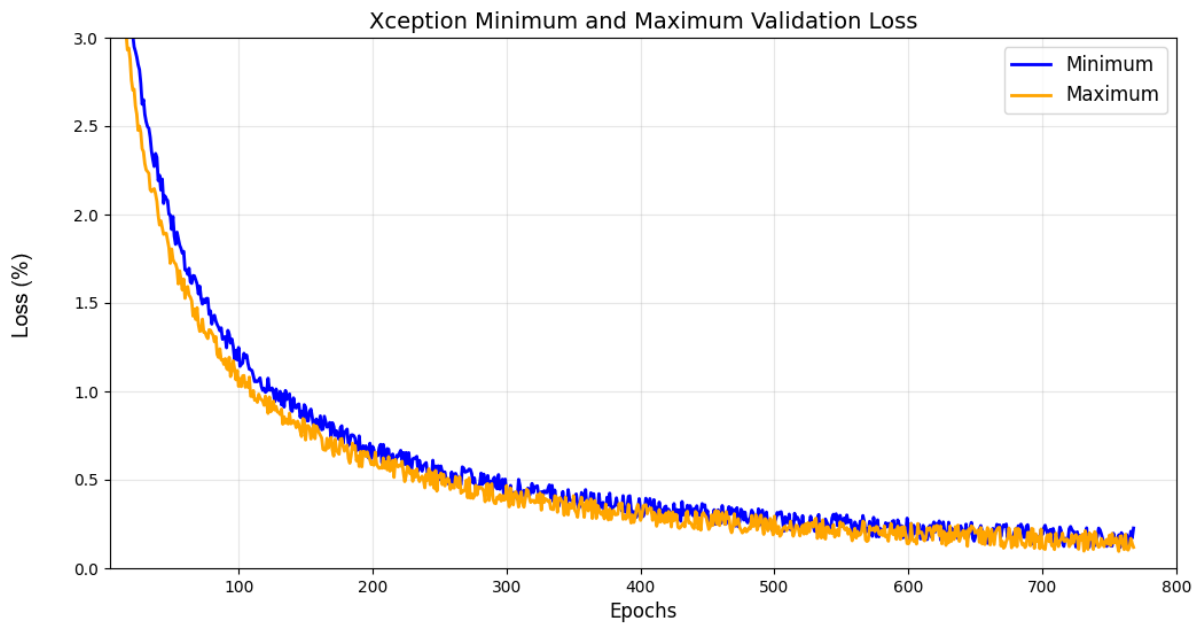
**Εικόνα 17: Η μέγιστη και ελάχιστη ακρίβεια κατά την διάρκεια της εκπαίδευσης του μοντέλου ResNet50.**

Στις πέντε επαναλήψεις που εκτελέστηκαν κατά την διάρκεια της εκπαίδευσης εφαρμόζοντας διασταυρούμενη επικύρωση 5 πτυχών, τα αποτελέσματα στα σύνολα επικύρωσης ήταν παρόμοια με μικρές αποκλίσεις. Τα μοντέλα ResNet50 και Xception έχουν απόδοση μέσης ακρίβειας 95% ενώ το VGG16 είχε μέση ακρίβεια 65% στα αντίστοιχα υποσύνολα επικύρωσης.

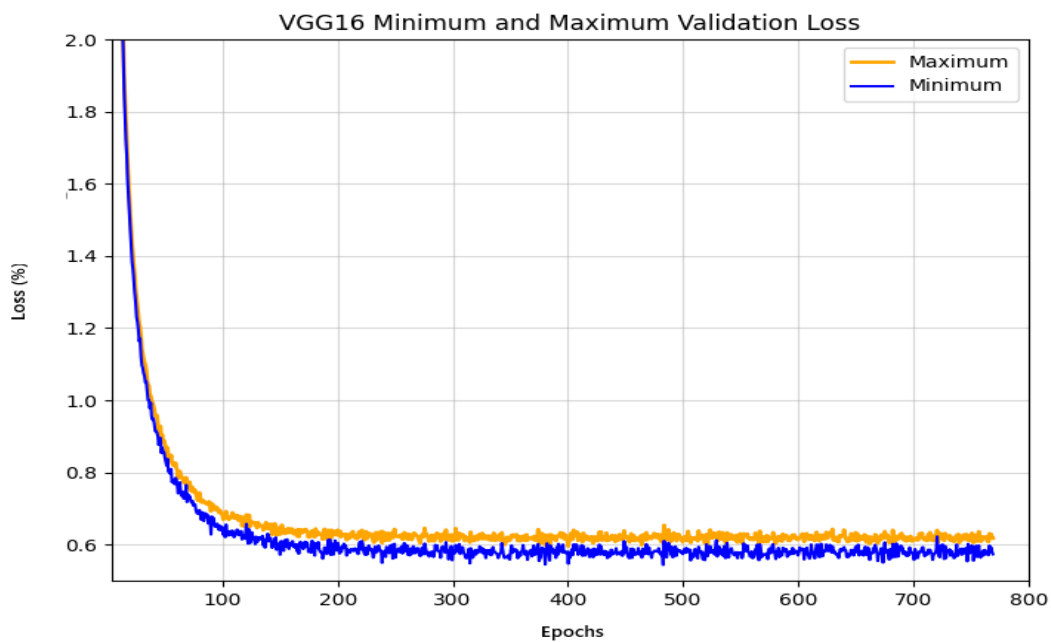
Εξίσου σημαντικό είναι να εξεταστεί η απώλεια (loss) που παρουσιάζουν τα τρία νευρωνικά δίκτυα στο σύνολο εκπαίδευσης. Οι συναρτήσεις κόστους ή αλλιώς συναρτήσεις απώλειας προσπαθούν να κάνουν μια αντιστοίχιση των παραμέτρων του νευρωνικού δικτύου σε μια τιμή. Η τιμή που λαμβάνει η συνάρτηση κόστους μετά την εκτέλεση κάθε επανάληψης, δηλαδή μετά από κάθε εποχή, είναι μια ένδειξη για το ποσό καλά αποδίδει ένα νευρωνικό δίκτυο. Και τα τρία δίκτυα παρουσίασαν παρόμοια συμπεριφορά όσον αφορά την απώλεια στο σύνολο δεδομένων ελέγχου. Το ίδιο συμβαίνει και στο σύνολο εκπαίδευσης ενώ μικρές διακυμάνσεις εντοπίζονται στην απώλεια του συνόλου επικύρωσης, όπου η απώλεια μειώνεται με σταθερό ρυθμό σε παρόμοιο αριθμό εποχών και για τα τρία μοντέλα. Τις μεγαλύτερες διακυμάνσεις στα δεδομένα επικύρωσης παρουσίασε η παραλλαγή του μοντέλου Xception.



**Εικόνα 18: Μέγιστη και ελάχιστη απώλεια στο σύνολο επικύρωσης για το μοντέλο ResNet50.**



**Εικόνα 19: Μέγιστη και ελάχιστη απώλεια στο σύνολο επικύρωσης για το μοντέλο Xception.**



**Εικόνα 20: Μέγιστη και ελάχιστη απώλεια στο σύνολο επικύρωσης για το μοντέλο VGG16**

## 5.2 Η Απόδοση Των Μοντέλων Στην Ταξινόμηση

Σε αυτήν την ενότητα παρουσιάζονται τα αποτελέσματα ταξινόμησης των νευρωνικών δικτύων στο σύνολο δοκιμών. Όπως αναφέρθηκε και στην ενότητα 4.1 το σύνολο δεδομένων αποτελείται από έξι διαφορετικές τάξεις, οι οποίες αντιπροσωπεύουν τις διαφορετικές ποικιλίες σπόρων δημητριακών. Οι φωτογραφίες που αποθηκεύτηκαν στο σύνολο δοκιμών δόθηκαν ως είσοδος και στα τρία νευρωνικά δίκτυα, ακολούθησε η δημιουργία των πινάκων σύγχυσης (confusion matrices) για κάθε δίκτυο και υπολογισμός των μετρικών που αναφέρθηκαν στην ενότητα 4.5.

Η παραλλαγή του μοντέλου VGG είχε την χειρότερη απόδοση σε όλες τις τάξεις με μέση ακρίβεια ταξινόμησης 65%. Η καλύτερη επίδοση ταξινόμησης του VGG ήταν για την ποικιλία Levante με ακρίβεια 82%, ενώ η χειρότερη επίδοση σημειώθηκε στην ποικιλία βρόμης Πήγασος με επίδοση ακρίβειας 21%. Η δεύτερη χειρότερη επίδοση σημειώθηκε στην ποικιλία Κριθάρι-Τριπτόλεμος (KTR) με 28% ακρίβεια, αλλά με τιμή 0.92 στην ανάκληση (recall) η οποία θεωρείται υψηλή και υποδηλώνει πως το μοντέλο ταξινομεί σωστά και με σχετικά μεγάλη ακρίβεια τις αληθώς θετικά τιμές, αλλά σε πολλές περιπτώσεις επισημαίνει λανθασμένα σπόρους οι οποίοι ανήκουν σε άλλες τάξεις ως κριθάρι-Τριπτόλεμος (KTR). Γενικά, η παραλλαγή του μοντέλου VGG16 δεν κατάφερε αξιολογικά αποτελέσματα σε καμία άλλη κατηγορία πέρα της ποικιλίας Levante, η οποία είναι η κυρίαρχη τάξη αφού περιέχει τον μεγαλύτερο αριθμό φωτογραφιών, δηλαδή τα περισσότερα δείγματα σε σχέση με τις υπόλοιπες. Παρόλη την προσπάθεια εξισορρόπησης του συνόλου δεδομένων χρησιμοποιώντας περίπου 1000 φωτογραφίες από κάθε τάξη σε κάθε επανάληψη εκπαίδευσης στο συγκεκριμένο μοντέλο φαίνεται να μην απέδωσε. Τα αποτελέσματα ταξινόμησης του μοντέλου VGG16 παρουσιάζονται αναλυτικά στον πίνακα 4.

### VGG16 Classification Report

	Precision	Recall	F1-Score	Support
<b>DMT</b>	0.28	0.92	0.43	106
<b>KTR</b>	0.50	0.03	0.06	184
<b>LVT</b>	0.82	0.96	0.89	1086
<b>SSS</b>	0.42	0.08	0.14	322
<b>VPG</b>	0.21	0.02	0.03	193
<b>VRF</b>	0.45	0.89	0.60	194
<b>accuracy</b>			0.65	2085
<b>macro avg</b>	0.45	0.48	0.36	2085
<b>weighted avg</b>	0.61	0.65	0.57	2085

**Πίνακας 4: Το classification report του μοντέλου VGG16.**

Οι παραλλαγές των άλλων δυο μοντέλων, ResNet και Xception επιτυγχάνουν πολύ καλύτερα αποτελέσματα από την παραλλαγή του μοντέλου VGG με ακρίβεια ταξινόμησης 97%. Η χειρότερη επίδοση της παραλλαγής ResNet σημειώνεται στην ποικιλία Κριθάρι-Θεσσαλονίκης με ακρίβεια 91% ενώ τις καλύτερες επιδόσεις σημείωσε



στις ποικιλίες Δήμητρα και Levante με ακρίβεια 99.2%. Αξιοσημείωτο είναι το γεγονός πως τα αποτελέσματά είναι ενθαρρυντικά ακόμα και για τις τάξεις που υποεκπροσωπούνται, όπως η ποικιλία κριθάρι Τριπτόλεμος και Δήμητρα, όπου στο σύνολο δοκιμών υπήρχαν 106 και 184 φωτογραφίες αντίστοιχα. Συμπερασματικά οι επιδόσεις του μοντέλου ResNet50 στις μειοψηφικές τάξεις, είναι μια ένδειξη ότι έχει επιτευχθεί η βέλτιστη δυνατή σύγκλιση χωρίς να παρουσιάζονται στοιχεία υπέρ-προσαρμογής του δικτύου (overfitting) στα δεδομένα εκπαίδευσης, επιτρέποντας του να πραγματοποιήσει μια πολύ καλή γενίκευση του προβλήματος. Επίσης, αναδεικνύεται η ορθότητα της προσέγγισης στην διαδικασία εκπαίδευσης χρησιμοποιώντας τα τεχνητά δείγματα σε συνδυασμό με την μέθοδο της υποδειγματοληψίας αφού το μοντέλο αναγνώρισε με πάρα πολύ μεγάλη ακρίβεια τα περισσότερα δείγματα του συνόλου δοκιμών.

### ResNet50 Classification Report

	Precision	Recall	F1-Score	Support
<b>DMT</b>	0.92	0.99	0.95	106
<b>KTR</b>	0.91	0.99	0.95	184
<b>LVT</b>	1.00	0.97	0.98	1086
<b>SSS</b>	0.95	0.99	0.97	322
<b>VPG</b>	1.00	0.98	0.98	193
<b>VRF</b>	0.97	0.97	0.97	194
<b>accuracy</b>			0.97	2085
<b>macro avg</b>	0.96	0.98	0.97	2085
<b>weighted avg</b>	0.97	0.97	0.97	2085

**Πίνακας 5: To classification report του μοντέλου ResNet50**

Η παραλλαγή του μοντέλου Xception, επιτυγχάνει 97% ακρίβεια ταξινόμησης. Όπως και με την παραλλαγή του μοντέλου ResNet50, τα καλύτερα αποτελέσματα επιτυγχάνονται στην ποικιλία Levante με ακρίβεια που ξεπερνά το 99%. Η χειρότερη επίδοση επιτυγχάνεται στην ποικιλία κριθάρι-Τριπτόλεμος με 88%, ενώ για όλες τις υπόλοιπες ποικιλίες η ακρίβεια ξεπερνά το 94%. Η επίδοση του μοντέλου Xception είναι εξίσου εντυπωσιακή παρόλο που επέδειξε την μεγαλύτερη μεταβλητότητα στην απώλεια των δεδομένων επικύρωσης. Η απόδοση του στις μειοψηφικές τάξεις είναι αρκετά καλή όσο της παραλλαγής ResNet50, παρουσιάζοντας ωστόσο μια πολύ μικρή διαφορά της τάξης του 3%. Επιπλέον, παρόλο που η επίδοση του μοντέλου στην ποικιλία κριθάρι-Τριπτόλεμος είναι η χειρότερη, το μοντέλο κατάφερε να ταξινομήσει σωστά τα αληθώς θετικά δείγματα με ακρίβεια 99%. Ωστόσο, υπήρχαν περιπτώσεις όπου σπόροι από άλλες τάξεις ταξινομήθηκαν εσφαλμένα ως κριθάρι-Τριπτόλεμος ανεβάζοντας έτσι τον αριθμό ψευδώς θετικών δειγμάτων. Καταληκτικά, ενώ το μοντέλο επιδεικνύει υψηλό βαθμό ακρίβειας, παραμένει ένα οριακό περιθώριο βελτίωσης σχετικά με την μείωση των ψευδώς θετικών δειγμάτων.

**Xception Classification report**

	Precision	Recall	F1-Score	Support
<b>DMT</b>	0.88	0.99	0.93	106
<b>KTR</b>	0.96	0.97	0.96	184
<b>LVT</b>	1.00	0.97	0.98	1086
<b>SSS</b>	0.94	0.98	0.96	322
<b>VPG</b>	0.98	0.98	0.98	193
<b>VRF</b>	0.97	0.98	0.97	194
<b>accuracy</b>			0.97	2085
<b>macro avg</b>	0.95	0.98	0.97	2085
<b>weighted avg</b>	0.97	0.97	0.97	2085

**Πίνακας 6: To classification report του μοντέλου Xception****5.3 Σύγκριση Μετρικών**

Η ακρίβεια της ταξινόμησης (accuracy) αποτελεί μόνο μια γενική επισκόπηση της απόδοσης ενός νευρωνικού δικτύου. Αντίθετα οι μετρικές όπως αυτές ορίστηκαν στην ενότητα 4.5, εμβαθύνουν επιτρέποντας έτσι μια πιο λεπτομερή και προσεκτική αξιολόγηση. Βοηθούν ουσιαστικά διασφαλίζοντας ότι οι προβλέψεις του μοντέλου ευθυγραμμίζονται με τις απαιτήσεις και τους περιορισμούς του προβλήματος προς επίλυση.

Το μοντέλο VGG16 παρουσιάζει τις χειρότερες μετρικές σε σχέση με τα άλλα δυο δίκτυα τα οποία επιτυγχάνουν παρόμοιες επιδόσεις. Η ακρίβεια και η ανάκληση του μοντέλου VGG16 παρουσιάζουν μεγάλες διακυμάνσεις μεταξύ των διαφορετικών τάξεων, δηλαδή των διαφορετικών ποικιλιών. Πιο αναλυτικά, μόνο για τρεις τάξεις εμφανίζονται αξιόλογες επιδόσεις στην ανάκληση, για το κριθαρι-Τριππολεμου, την ποικιλία Levante και την βρώμη Φλεγα, με 0.92, 0.96 και 0.89 αντίστοιχα. Και στην ακρίβεια εντοπίζονται μεγάλες διακυμάνσεις στις τάξεις αυτές, όπου σημειώνονται επιδόσεις ίσες με 0.28, 0.82 και 0.45. Σε τέτοιες περιπτώσεις όπου εμφανίζεται μεγάλη διαφορά μεταξύ ακρίβειας και ανάκλησης, το μοντέλο αποκτά την τάση να υπερ-προβλέπει αυτές τις κατηγορίες, συλλαμβάνοντας περισσότερες από τις πραγματικές περιπτώσεις, με αποτέλεσμα αυξημένες ψευδώς θετικές επισημάνσεις. Εν ολίγοις, ενώ το μοντέλο επιτυγχάνει συνολική ακρίβεια 0.65, μια βαθύτερη κατάδυση στις μετρικές για συγκεκριμένες κατηγορίες αποκαλύπτει σημαντικές διαφορές στην απόδοση. Για παράδειγμα, το μοντέλο υπερέχει στον προσδιορισμό ορισμένων κατηγοριών, όπως την ποικιλία Λεβάντε, αλλά αντιμετωπίζει προκλήσεις με άλλες, όπως η βρόμη Πήγασος (VPG). Ως εκ τούτου, είναι εμφανές πως χρειάζονται περαιτέρω ενέργειες για την βελτίωση των δυνατοτήτων ταξινόμησης σε όλες τις κατηγορίες.

Σε αντίθεση με την επίδοση του μοντέλου VGG, η παραλλαγή του ResNet50 παρουσιάζει υποδειγματική επάρκεια στις προβλέψεις του. Οι επιδόσεις στην αναγνώριση των ποικιλιών Λεβάντε και βρόμη Πήγασος ξεχωρίζουν με ακρίβεια 1.00, υποδεικνύοντας ότι όταν το μοντέλο προβλέπει ότι ένας σπόρος ανήκει σε οποιαδήποτε από αυτές τις κατηγορίες, είναι σχεδόν πάντα σωστό. Η υψηλή ανάκληση για την ποικιλία

Λεβάντε (LVT) υπογραμμίζει περαιτέρω την ισχυρή του απόδοση, ενώ η ελαφρώς χαμηλότερη ανάκληση για το βρόμη Πήγασος (VPG) (0.95) υποδηλώνει ότι μερικές περιπτώσεις αυτής της κατηγορίας μπορεί να έχουν χαθεί. Και οι υπόλοιπες κατηγορίες παρουσιάζουν εντυπωσιακές τιμές ακριβείας με 0.92, 0.91, 0.95, 0.97 αντίστοιχα. Οι οποίες, συνδυάζονται με υψηλές τιμές ανάκλησης που κυμαίνονται μεταξύ 0.97 και 0.99, καθιστώντας σαφές ότι το μοντέλο είναι ικανό να εντοπίζει και να ταξινομεί σωστά περιπτώσεις όλων των κατηγοριών με ελάχιστα σφάλματα. Οι μετρήσεις ανά κατηγορία τονίζουν περαιτέρω τη συνέπεια και την ακρίβεια του μοντέλου στις προβλέψεις του. Αυτή η απόδοση υποδηλώνει ότι το μοντέλο είναι καλά συντονισμένο και έχει εκπαιδευτεί αποτελεσματικά για να διακρίνει και να ταξινομεί περιπτώσεις στις διαφορετικές κατηγορίες που παρουσιάζονται.

Στο ίδιο μοτίβο κινούνται και οι μετρικές του μοντέλου Xception. Η ταξινόμηση της ποικιλίας LVT, με ακρίβεια 1.00, αποτελεί απόδειξη της άψογης ικανότητας του μοντέλου να διακρίνει. Επομένως όσον αφορά τις προβλέψεις για αυτήν την ποικιλία φαίνεται να υπάρχει μια σχεδόν άψογη απόδοση. Η ελαφρώς χαμηλότερη ανάκληση του 0.97 υποδηλώνει ότι ενώ οι προβλέψεις του μοντέλου είναι εξαιρετικά ακριβείς, μερικές περιπτώσεις μπορεί να έχουν παραλειφθεί. Στην ποικιλία KTR εμφανίζεται η χαμηλότερη ακρίβεια με 0.88, η οποία, όμως συνδυάζεται με υψηλή τιμή ανάκλησης της τάξης του 0.99, υποδεικνύοντας έτσι, ότι το μοντέλο όχι μόνο κάνει πολύ ακριβείς προβλέψεις για αυτήν την κατηγορία, αλλά καταγράφει επίσης σχεδόν όλες τις παρουσίες της. Για τις υπόλοιπες κατηγορίες, το μοντέλο δείχνει σταθερές τιμές ακριβείας που κυμαίνονται από 0.94 έως 0.98. Σε συνδυασμό με τις τιμές ανάκλησης οι οποίες κινούνται στα ίδια υψηλά επίπεδα, είναι προφανές ότι το μοντέλο επιτυγχάνει μια αρμονική ισορροπία μεταξύ ακρίβειας και ευαισθησίας για αυτές τις κατηγορίες, εξασφαλίζοντας ελάχιστες εσφαλμένες ταξινομήσεις και παραλείψεις. Καταληκτικά, το μοντέλο παρουσιάζει εντυπωσιακή συνολική ακρίβεια η οποία αντικατοπτρίζει την ικανότητα του για αξιόπιστες προβλέψεις. Αν και οι μετρικές είναι λίγο πιο μειωμένες σε σχέση με το ResNet50 η διαφορά είναι πολύ μικρή. Το γεγονός αυτό υποδηλώνει ότι η παραλλαγή του μοντέλου Xception έχει εκπαιδευτεί εξίσου αποτελεσματικά ώστε να μπορεί να διακρίνει πολύπλοκες περιπτώσεις μεταξύ των διαφορετικών ποικιλιών.

## 6. Συζήτηση

Η αναγνώριση διαφορετικών ειδών σπορών χρησιμοποιώντας τεχνικές βαθιάς μάθησης είναι μια σχετικά καινούρια προσέγγιση δεδομένου ότι οι περισσότερες δημοσιεύσεις αφορούν το χρονικό διάστημα της τελευταίας δεκαετίας [8]. Σε σχέση με αλλά παρόμοια έργα που εντοπίστηκαν κατά την βιβλιογραφική ανασκόπηση του πεδίου, επιτυγχάνονται παρόμοια αποτελέσματα. Στις περισσότερες περιπτώσεις τα έργα αυτά ασχολήθηκαν με σύνολα δεδομένων που αποτελούταν από ένα μικρό πλήθος διαφορετικών κατηγοριών ο οποίος ήταν μεταξύ τεσσάρων και πέντε διαφορετικών τάξεων [11][45][47]. Η ακρίβεια των αποτελεσμάτων στα έργα αυτά κυμαίνεται από 95% έως 97% τα οποία είναι αρκετά κοντά σε σχέση με τα αποτελέσματα της παρούσας εργασίας.

Η ύπαρξη ενός μικρού αριθμού διαφορετικών κατηγοριών δεν συμβαδίζει με τις συνθήκες στον πραγματικό κόσμο, αφού αρκετά είδη μπορεί να διαχωρίζονται σε πολλές διαφορετικές ποικιλίες. Όπως είδαμε και στην βιβλιογραφική ανασκόπηση σε μια έρευνα που αφορούσε το ρύζι εξετάστηκαν 14 διαφορετικές ποικιλίες [43]. Αντίστοιχα αυτό συμβαίνει και σε άλλα είδη σπόρων αυξάνοντας έτσι πάρα πολύ το σύνολο των διαφορετικών κατηγοριών που θα πρέπει το μοντέλο να ταξινομήσει. Όταν η ταξινόμηση συμβαίνει μεταξύ διαφορετικών ειδών και ποικιλιών οι διαφορετικές κατηγορίες που πρέπει να ταξινομηθούν αυξάνονται ακόμα περισσότερο, με αποτέλεσμα να αυξάνεται

κατά πολύ η πολυπλοκότητα του προβλήματος. Άλλωστε, η ύπαρξη ανισορροπίας πολλές φορές οδηγεί σε προκατάληψη υπέρ της πλειοψηφικής κατηγορίας επηρεάζοντας σημαντικά την ικανότητα του μοντέλου να γενικεύσει οδηγώντας σε λάθος ταξινομήσεις [57].

Όμως αν υπάρχει ένα πολύ μεγάλο σύνολο δεδομένων και όλες οι διαφορετικές κατηγορίες μπορούν να εκπροσωπηθούν επαρκώς τότε δεν επηρεάζεται τόσο η ικανότητα του μοντέλου να γενικεύσει, γιατί υπάρχει ένας πολύ μεγάλος αριθμός διαθέσιμων παραδειγμάτων για εκπαίδευση. Πράγματι σε όσα έργα υπήρχε διαθέσιμο ένα μεγάλο σύνολο δεδομένων ακόμα και αν ο αριθμός των τάξεων ήταν μεγάλος τα αποτελέσματά ήταν πολύ ενθαρρυντικά [48][42][11].

Συγκριτικά με αλλά έργα που μελετήθηκαν η παρούσα εισάγει μια διαφορετική προσέγγιση. Σε έναν ικανοποιητικό αριθμό περιπτώσεων από την εξεταζόμενη βιβλιογραφία, η προσέγγιση που ακολουθήθηκε σχετικά με την ταξινόμηση αφορούσε ποικιλίες του ίδιου είδους. Για παράδειγμα σε δυο μελέτες προσπάθησαν να ταξινομηθούν πέντε διαφορετικές ποικιλίες σίτου [11][47], ενώ σε κάποιες άλλες πραγματοποιούνταν ταξινόμηση σε διαφορετικές ποικιλίες ρυζιού [46][43]. Η κατεύθυνση τους ήταν στραμμένη προς την αναγνώριση μεταξύ διαφορετικών ποικιλιών του ίδιου είδους.

Αντιθέτως, στην παρούσα η προσέγγιση αφορά ταξινόμηση σε δυο επίπεδα, του είδους και της ποικιλίας, ένα γεγονός που μπορεί να προσφέρει ευελιξία και να αντιμετωπιστούν ταξινομικές προκλήσεις που προκύπτουν όταν ταυτόχρονα εξετάζονται διαφορετικά είδη και ποικιλίες. Επιπλέον, η ενσωμάτωση πολλαπλών ποικιλιών και ειδών στην ταξινόμηση καθιστά το σύστημα πιο ανθεκτικό και λιγότερο επιρρεπές σε υπερεκπαίδευση πάνω σε συγκεκριμένα χαρακτηριστικά που μπορεί να εμφανίζει κάθε ποικιλία. Με αυτόν τον τρόπο, το σύστημα αποκτά μεγαλύτερη ικανότητα προσαρμογής σε ευρύτερα σύνολα δεδομένων, προσφέροντας πιο αξιόπιστα αποτελέσματα σε διαφορετικά περιβάλλοντα και εφαρμογές. Αυτή η προσέγγιση εξασφαλίζει ότι η ταξινόμηση δεν περιορίζεται από τα στενά όρια μιας μόνο ποικιλίας, αλλά μπορεί να χρησιμοποιηθεί σε ευρύτερες και πιο ποικιλόμορφες συνθήκες.

Επιπλέον, η έρευνα σε επίπεδο μεταξύ ειδών και οικογενειών παραμένει σχετικά λιγότερο εξερευνημένη σε σύγκριση με τον μεγάλο αριθμό μελετών που εστιάζουν στην ταξινόμηση σε επίπεδο ποικιλιών είτε αυτές αφορούν εικόνες σπόρων ή φυτών. Και ακόμα και εκείνα τα έργα που έχουν ως στόχο την ταξινόμηση διαφορετικών ειδών με πολλές διαφορετικές κατηγορίες δεν επιχειρούν να ταξινομήσουν διαφορετικές ποικιλίες ίδιων ειδών [33][36]. Αυτή η διαφορά αναδεικνύει ένα κενό στην βιβλιογραφία καθώς τα υψηλότερα ταξινομικά επίπεδα συχνά περιλαμβάνουν ευρύτερα και πιο πολύπλοκα πρότυπα ποικιλομορφίας. Εστιάζοντας σε αυτό το κενό μπορεί να αποκτηθεί περισσότερη γνώση για τα διακριτικά χαρακτηριστικά των διαφορετικών ειδών και ποικιλιών αλλά και πιθανά μοτίβα που δεν είχαν προηγουμένως παρατηρηθεί. Τέτοιου είδους προσεγγίσεις ενισχύουν την επιστημονική κατανόηση και μπορούν να αποτελέσουν μια σταθερή βάση για μελλοντική ερευνά.

Μέχρι σήμερα οι διαδικασίες ταξινόμησης διεξάγονται χειροκίνητα από πιστοποιημένους αναλυτές, οι οποίοι στηρίζονται κατά βάση στον μορφολογικό έλεγχο κάθε σπόρου. Οι διαδικασίες αυτές είναι κατά βάση αποτελεσματικές και οι μονές έγκυρες που διεξάγονται σήμερα. Παρόλα αυτά είναι αρκετά χρονοβόρες διαδικασίες και μπορούν να μην είναι αντικειμενικές καθώς δεν αποκλείεται το ενδεχόμενο ανθρώπινου λάθους [58].

Εφόσον η αναγνώριση και ταξινόμηση σπόρων στηρίζεται στον ανθρώπινο έλεγχο αποτελεί μια επίκτητη ικανότητα. Ένας πιστοποιημένος αναλυτής εκπαιδεύεται για

περισσότερες από 1500 ώρες ώστε να είναι ικανός να αναγνωρίσει μεταξύ 200-300 ειδών διαφορετικών σπόρων (Canada Food Inspection Agency - CFIA 2015) [59][60] ενώ σύμφωνα με τον οργανισμό International Seed Testing Association (ISTA) [61] χρειάζονται τουλάχιστον δυο χρόνια για την συμπλήρωση του ανώτατου επιπέδου πιστοποίησης. Επιπλέον, χρειάζεται συστηματική μελέτη για θέματα που δεν είναι απαραίτητα για την εκπαίδευση αλλά απαιτείται να εξετάζονται ταυτόχρονα με την ανάλυση, ώστε να είναι αποτελεσματική η διαδικασία αναγνώρισης. Τέτοια θέματα περιλαμβάνουν την εθνική νομοθεσία, και τα πρότυπα πιστοποίησης ώστε να εξασφαλίζεται η ποιότητα [59]. Επομένως είναι εμφανές πως ακόμα και ένα πολύ έμπειρος αναλυτής με βαθιά γνώση του θέματος μπορεί να υποπέσει σε σφάλμα.

Προφανώς για τους αγρότες τα πράγματα είναι ακόμα πιο περιπλοκά αφού δεν έχουν λάβει καμία εκπαίδευση ούτε έχουν ειδική γνώση. Η ικανότητα να ταξινομήσουν σπόρους στηρίζεται αποκλειστικά στις προσωπικές τους εμπειρίες δηλαδή η γνώση τους είναι γύρω από τους σπόρους των φυτών που καλλιεργούν. Για τους αγρότες είναι αρκετά σημαντικό να γνωρίζουν με αρκετά μεγάλη βεβαιότητα το είδος και την ποικιλία της καλλιέργειας. Η ύπαρξη διαφορετικών ποικιλιών στο ίδιο χωράφι μπορεί να δημιουργήσει σοβαρά προβλήματα που θα επηρεάσουν την συγκομιδή. Για παράδειγμα, ποικιλίες οι οποίες έχουν διαφορετικό χρόνο ωρίμανσης σημαίνει πως θα έχουν και διαφορετικό χρόνο συγκομιδής κάνοντας πολύπλοκη την διαδικασία οδηγώντας ενδεχομένως σε απώλειες.

Επομένως, αν κάποιος αγρότης έχει αμφιβολία για τον σπόρο που βλέπει δεν υπάρχει κάποιο εργαλείο που θα μπορούσε να χρησιμοποιήσει να τον βοηθήσει να πάρει μια απόφαση. Η μοναδική λύση είναι να αποταθεί σε έναν ειδικό το οποίο είναι αρκετά χρονοβόρο. Η παρούσα διατριβή στοχεύει να θεμελιώσει αυτό το χάσμα μεταξύ εργαστήριου και αγρού, καλύπτοντας έτσι ένα σημαντικό ερευνητικό κενό. Η εστίαση σε δυο επίπεδα, του είδους και της ποικιλίας, δίνει την δυνατότητα ανάπτυξης αυτοματοποιημένων εργαλείων ικανά να διακρίνουν τους σπόρους αποτελεσματικά και με μεγάλη ακρίβεια ανταποκρινόμενα στην ποικιλομορφία που συναντάται στον αγρό.

Σε ένα γενικότερο πλαίσιο η αυτοματοποίηση της διαδικασίας ταξινόμησης σπόρων μπορεί να προσφέρει πολλαπλά οφέλη και σε άλλα υποπεδία της γεωργίας. Αρχικά, έχει την δυνατότητα να συμβάλλει στην βελτίωση των υπαρχουσών πρακτικών πιστοποίησης, οδηγώντας έτσι στην αναβάθμιση των προϊόντων μέσω συστηματικού ποιοτικού ελέγχου αφού η διαδικασία δεν θα είναι πια χρονοβόρα. Εξασφαλίζοντας ότι οι σπόροι ταξινομούνται με ακρίβεια μπορούν να βελτιωθούν οι πρακτικές και να δημιουργηθούν νέες εφαρμογές στην γεωργία ακρίβειας. Η ομοιομορφία της σποράς στους αγρούς εξασφαλίζει ομοιόμορφες συνθήκες άρδευσης και συγκομιδής μειώνοντας την υπερβολική άρδευση και την εκτεταμένη χρήση λιπασμάτων συμβάλλοντας έτσι στην βιώσιμη γεωργική ανάπτυξη.

Παρόλο που οι εφαρμογές βαθιάς μάθησης στη γεωργία αναπτύσσουν σημαντικές δυνατότητες, η πρόοδος στον τομέα παραμένει αργή, καθώς αντιμετωπίζει πολλούς περιορισμούς που οφείλονται σε διάφορους παράγοντες. Ένας κοινός περιορισμός που παρουσιάστηκε μεταξύ της παρούσης και άλλων έργων του πεδίου είναι η απουσία δημοσίων συνόλων δεδομένων με ελεύθερη πρόσβαση. Στα περισσότερα έργα οι εικόνες αποκτήθηκαν μέσω φωτογράφισης, των σπόρων οι οποίοι δωρίστηκαν από διάφορους γεωργικούς οργανισμούς. Δεν υπάρχουν ψηφιακά αρχεία με ακριβείς πληροφορίες, επομένως γίνεται επιτακτική η ανάγκη για έλεγχο επιβεβαίωσης των αρχείων αυτών από ειδικούς του πεδίου με αποτέλεσμα να επιμηκύνονται οι διαδικασίες. Επιπλέον, η δημιουργία μεγάλων συνόλων δεδομένων που θα αποτελέσουν την βάση για να εκπαιδευτούν βαθιά νευρωνικά δίκτυα απαιτούν σημαντικές δαπάνες πόρων και είναι μια εξαιρετικά χρονοβόρα διαδικασία.

Ένας άλλος περιορισμός που συναντήθηκε στα ερευνητικά έργα που επιχειρούν να ταξινομήσουν σπόρους με χρήση τεχνικών βαθιάς μάθησης πηγάζει από την διαδικασία εκπαίδευσης των νευρωνικών δικτύων. Τα Συνελικτικά νευρωνικά δίκτυα είναι ακριβά υπολογιστικά συστήματα που απαιτούν αρκετά μεγάλη υπολογιστική ισχύ. Η διαδικασία εκπαίδευσης γίνεται ακόμα πιο πολύπλοκη και χρονοβόρα λόγω της έλλειψης δεδομένων καθιστώντας έτσι την ευρεία εφαρμογή αυτής της τεχνολογίας πολύ δύσκολη σε πραγματικές συνθήκες. Σε αρκετά έργα όπως και στην παρούσα οι φωτογραφίες αποκτήθηκαν σε συγκεκριμένες συνθήκες όπου υπήρχε τεχνητό φόντο που βοηθά στην διαδικασία της εκπαίδευσης, όμως δεν προσομοιώνει τις συνθήκες του αγρού. Η ύπαρξη φωτός είναι απαραίτητη προϋπόθεση για να αποκτηθεί η φωτογραφία ώστε να λειτουργήσει ένα τέτοιο σύστημα και αυτές διαφέρουν αρκετά κατά την διάρκεια της ημέρας, αλλά και από εποχή σε εποχή. Συνεπώς, τα περισσότερα έργα αποτελούν απόδειξη εφικτότητας έχοντας όμως μεγάλα περιθώρια ερευνάς, λόγω του ότι αποτελεί ένα πολύ πρόσφατο ερευνητικό θέμα [62].

Κατά την εισαγωγή νέων τεχνικών, είναι απαραίτητη η ενδεδειγμένη αξιολόγησή τους και να πληρούν συγκεκριμένες απαιτήσεις διαπίστευσης. Μέχρι σήμερα στη ταξινόμηση σπόρων υπάρχουν καθιερωμένες μέθοδοι και κανόνες δοκιμών που πρέπει να ακολουθούνται, όπως αυτοί που ορίζονται από τους οργανισμούς Association of Official Seed Analysts (AOSA 2021) [63] και ISTA (2022) [64]. Επομένως, οι οργανισμοί τυποποίησης διαδραματίζουν καθοριστικό ρόλο στην προώθηση της υιοθέτησης και της μελλοντικής προσαρμογής αυτών των νέων τεχνολογιών. Όμως διάφορες μέθοδοι που έχουν αναπτυχθεί δεν έχουν ακόμα αξιολογηθεί μέσω τυποποιημένων ή ευρέως αποδεκτών προσεγγίσεων, ούτε έχουν δοθεί γενικές κατευθύνσεις από εργαστήρια ή γεωργικά ινστιτούτα. Με άλλα λόγια, η εφαρμογή τους παραμένει ανεπίσημη ή μη επαληθευμένη. Παράλληλα λόγω της έλλειψης κατεύθυνσης και της απουσίας επίσημων πρωτοκόλλων οι προγραμματιστές δυσκολεύονται να προσαρμόσουν τα μοντέλα τους ώστε να συμμορφώνονται με τα καθιερωμένα πρότυπα ή κανόνες ταξινόμησης. Η έλλειψη προτύπων και επίσημων διαδικασιών με την συμβολή εξειδικευμένων οργανισμών επιβραδύνει την υιοθέτηση προηγμένων τεχνολογιών [59].

Καταληκτικά, οι εφαρμογές τεχνητής νοημοσύνης έχουν ευρέως δοκιμαστεί σε διαφορά προβλήματα του αγροτικού τομέα επιδεικνύοντάς εντυπωσιακά αποτελέσματα [8]. Στο πλαίσιο της ταξινόμησης σπόρων τα αποτελέσματα πολλών έργων δείχνουν μεγάλη ακρίβεια αναγνώρισης αποδεικνύοντας ότι η υιοθέτηση προηγμένων τεχνολογιών μπορεί να αποτελέσει μια βιώσιμη λύση. Για να ξεπεραστούν όμως οι υπάρχοντες ερευνητικοί περιορισμοί, είναι απαραίτητος ο συντονισμός μεταξύ ερευνητών, βιομηχανίας και εξειδικευμένων οργανισμών για την δημιουργία αξιόπιστων και ευρέως αποδεκτών μεθόδων. Έτσι, το μέλλον της γεωργίας μπορεί να γίνει πιο ακριβές και αποτελεσματικό.

## 7. Μελλοντική Ερευνά

Η συνεχής εξέλιξη της τεχνολογίας των νευρωνικών δικτύων και των επιστημονικών προσεγγίσεων ανοίγει νέους δρόμους για την αντιμετώπιση των περιορισμών που υπάρχουν σήμερα στην ταξινόμηση σπόρων. Οι μελλοντικές έρευνες μπορούν να επικεντρωθούν τόσο στη βελτίωση των τεχνικών υπολογιστικής όρασης που εφαρμόζονται στον γεωργικό τομέα, όσο και στην ενσωμάτωση αυτών σε καθιερωμένα πρότυπα και διαδικασίες. Στο κεφάλαιο αυτό, εξετάζονται οι πιθανές κατευθύνσεις της έρευνας που μπορούν να οδηγήσουν σε βιώσιμες και πρακτικές λύσεις για το μέλλον.

Ο μεγαλύτερος περιορισμός που χρειάζεται να αντιμετωπιστεί είναι η έλλειψη διαθέσιμων πηγών δεδομένων με ελεύθερη πρόσβαση. Η ύπαρξη δεδομένων είναι η

βάση για την εκπαίδευση μοντέλων βαθιάς μάθησης. Λόγω του βάθους τέτοιων μοντέλων είναι εξίσου απαραίτητο τα σύνολα δεδομένων να περιέχουν ένα μεγάλο αριθμό παρατηρήσεων. Με αυτό τον τρόπο τα μοντέλα μπορούν να προσπελάσουν ένα μεγάλο αριθμό παραδειγμάτων και να επιτύχουν αποτελέσματα με πολύ υψηλή ακρίβεια.

Επιπλέον, λόγω της έλλειψης μιας μεγάλης ενιαίας βάσης δεδομένων οι έρευνες που διεξάγονται στο πλαίσιο ταξινόμησης σπόρων με χρήση τεχνικών βαθιάς μάθησης, μπορούν να εστιάσουν μόνο σε ένα μικρό αριθμό ειδών ή ποικιλιών. Ένα θετικό γεγονός είναι ότι, η κατασκευή ενός μεγάλου και αξιόπιστου συνόλου δεδομένων για την ταξινόμηση σπόρων μπορεί να μην χρειάζεται απαραίτητα ειδικό εξοπλισμό, χρειάζεται όμως επεξεργασία των εικόνων και επικύρωση από ειδικούς του πεδίου για την διαβεβαίωση των ετικετών τους. Αντίθετα, τα έργα τα οποία επικεντρώνονται σε αλλά μέρη του φυτού όπως τα φύλλα ή τα άνθη, δεν αντιμετωπίζουν τις ίδιες προκλήσεις διότι η πρόσβαση σε μεγάλα και αξιόπιστα σύνολα δεδομένων είναι πιο εύκολη.

Με την συμβολή ειδικών του πεδίου και εξειδικευμένων οργανισμών μπορούν να δημιουργηθούν σαφή πρότυπα σχετικά με τα χαρακτηριστικά ταξινόμησης κάθε είδους και ποικιλίας, βοηθώντας έτσι ουσιαστικά τους ερευνητές. Συνεπώς, άρτια εκπαιδευμένοι αναλυτές και επιστήμονες χρειάζονται καθ' όλη την διάρκεια κατασκευής των συνόλων δεδομένων. Με την γνώση τους και την εμπειρία τους μπορούν να βοηθήσουν στην απόκτηση των σημαντικών χαρακτηριστικών κάθε σπόρου και του είδους του και να επιβεβαιώσουν την ορθότητα των επισημάνσεων. Εξίσου απαραίτητη είναι η συνεργασία μεταξύ των εργαστήριων, της βιομηχανίας και αναλυτών σπόρων με προγραμματιστές τεχνητής νοημοσύνης. Μια τέτοια συνεργασία μπορεί να οδηγήσει στην δημιουργία μιας αξιόπιστης βάσης δεδομένων μεγάλης κλίμακας που θα αποτελείται από εικόνες σπόρων, συμβάλλοντας έτσι σημαντικά στην προώθηση της εφαρμογής της τεχνητής νοημοσύνης στην αναγνώριση σπόρων.

Ο δεύτερος περιορισμός που αναφέρθηκε αφορά την φύση των συστημάτων βαθιάς μάθησης τα οποία απαιτούν μεγάλη υπολογιστική ισχύ και της δύσκολης εφαρμογής τους σε πραγματικές συνθήκες. Επιπλέον, λόγω του γεγονότος ότι οι έρευνες που αφορούν εφαρμογές βαθιάς μάθησης είναι πολύ πρόσφατες, δεν έχουν γίνει εκτενείς συγκρίσεις με άλλες παραδοσιακές τεχνικές μηχανικής μάθησης, ούτε έχουν δοκιμαστεί εκτενώς αν και έχουν γίνει κάποιες προσπάθειες από κάποιους ερευνητές ενδεδειγμένης εξερεύνησης και αξιολόγησης του πεδίου [8][58]. Παρόλα αυτά χρειάζονται περισσότερες έρευνες ώστε να εξαχθούν ουσιαστικά συμπεράσματα σχετικά με το ποσό ικανά είναι αυτά τα συστήματα να επιλύσουν προβλήματα όπως την ταξινόμηση διάφορων ειδών σπόρων σε πραγματικές συνθήκες. Έτσι θα μπορέσει να αναδειχθεί μια ξεκάθαρη επιστημονική κατεύθυνση που θα μπορέσει να επιταχύνει τις σχετικές έρευνες στο πεδίο.

Σχετικά με την αποδοτική εκπαίδευση βαθιών νευρωνικών δικτύων που αφορά το αντικείμενο της παρούσης, φαίνεται πως αναδύεται μια ξεκάθαρη τάση της χρήσης μεταφοράς γνώσης. Αρκετές έρευνες παρουσίασαν εντυπωσιακά αποτελέσματα χρησιμοποιώντας τις συγκεκριμένες τεχνικές [45][47][58]. Εξίσου εντυπωσιακά αποτελέσματα παρουσιάζουν έργα τα οποία διατηρούν μια μικτή προσέγγιση. Σε εκείνα οι ερευνητές επιχειρούν ένα συνδυασμό παραδοσιακών τεχνικών μηχανικής μάθησης και συνελκτικών νευρωνικών δικτύων. Επομένως η εφαρμογή μικτών προσεγγίσεων και οι τεχνικές μεταφοράς γνώσης μπορούν να αποτελέσουν μια λύση για την ανάπτυξη μοντέλων ικανών να εκπαιδευτούν σε μικρότερα σύνολα δεδομένων [46][58].

Σε περίπτωση έλλειψης δεδομένων και άλλες τεχνικές μπορούν να χρησιμοποιηθούν και ήδη έχουν γίνει κάποιες πρώτες προσπάθειες διευθέτησης του προβλήματος από κάποιους ερευνητές. Ένας τρόπος αποτελεί η τεχνική της επαύξησης δεδομένων εφαρμόζοντας δηλαδή μεταμορφώσεις σε υπάρχουσες εικόνες ώστε να δημιουργηθούν νέα δείγματα, τεχνική που χρησιμοποιήθηκε και στην παρούσα διατριβή.

Άλλα έργα χρησιμοποίησαν πλήρως συνθετικά σύνολα δεδομένων κάνοντας όμως ταξινόμηση σε πραγματικά δείγματα με αρκετά μεγάλη ακρίβεια, προσφέροντας έτσι μια ακόμα λύση στην έλλειψη επαρκών δεδομένων για εκπαίδευση [40]. Μελλοντικές έρευνες θα μπορούσαν επίσης να στραφούν στην λύση των γενετικών ανταγωνιστικών δικτύων (Generative Adversarial Networks) τα οποία έχουν χρησιμοποιηθεί με μεγάλη επιτυχία στην δημιουργία τεχνητών δειγμάτων σε άλλα επιστημονικά προβλήματα που υπάρχουν μεγάλη έλλειψη δεδομένων [65] [66].

Στο πλαίσιο αντιμετώπισης της έλλειψης πρότυπων διαδικασιών και επίσημων πρωτοκόλλων, χρειάζεται περισσότερη εμπλοκή από επιστημονικούς φορείς και καινοτόμους του γεωργικού τομέα. Θα πρέπει να υπάρξει μια ευρεία συνεργασία με κατεύθυνση την σαφή αξιολόγηση της αποτελεσματικότητας μεταξύ των νέων τεχνολογιών και των υπαρχόντων. Για παράδειγμα, μελλοντικές έρευνες θα μπορούσαν να εστιάσουν στην ανάπτυξη μιας τυποποιημένης διαδικασίας επαλήθευσης ή επικύρωσης για τη σύγκριση των συστημάτων που βασίζονται στην τεχνητή νοημοσύνη με τα υπάρχοντα πρότυπα. Έτσι θα υπάρξει πλήρης επιβεβαίωση και ξεκάθαρος καθορισμός των προϋποθέσεων που πρέπει να τηρούνται ώστε τα συστήματα που στηρίζονται στις τεχνικές βαθιάς μάθησης να μπορούν να υιοθετηθούν ευρέως.

Καταληκτικά, χρειάζονται συντονισμένες προσπάθειες και αρκετές μελλοντικές έρευνες για την πρακτική εφαρμογή και την ευρεία υιοθέτηση συστημάτων τεχνητής νοημοσύνης. Η δημιουργία μεγάλων συνόλων δεδομένων είναι μια χρονοβόρα διαδικασία που απαιτεί πολλούς πόρους, γεγονός που περιορίζει την πρόοδο, για αυτό η εφαρμογή τεχνικών βαθιάς μάθησης μπορεί να γίνει σταδιακά. Σε πρώτο στάδιο χρειάζεται ένα πλαίσιο προτυποποίησης, στη συνέχεια τα συστήματα αυτά θα μπορούν να αποτελέσουν εργαλεία υποστήριξης απόφασης για πάρα πολύ δύσκολες περιπτώσεις ταξινόμησης σπορών. Στα επόμενα στάδια θα πρέπει να εξερευνηθεί η ταξινόμηση σπορών με όσο το δυνατόν λιγότερη ανθρώπινη παρέμβαση. Με αυτό το τρόπο θα υπάρξει μια ομαλή μετάβαση, που θα επιτρέψει μέσω μελλοντικής έρευνας την πλήρη αυτοματοποίηση της διαδικασίας ταξινόμησης δίνοντας λύση σε ένα πολύπλοκο πρόβλημα που αντιμετωπίζει ο γεωργικός τομέας.

## 8. Συμπέρασμα

Σκοπός της παρούσας ήταν η δημιουργία ενός μοντέλου πρόβλεψης για τους αγρότες, το οποίο μέσω φωτογραφιών υψηλής ανάλυσης θα είναι ικανό να ταξινομήσει με μεγάλη ακρίβεια του σπόρους στα αντίστοιχα είδη και ποικιλίες που ανήκουν. Για την επίτευξη αυτού του σκοπού ήταν επιτακτική η δημιουργία ενός μεγάλου συνόλου δεδομένων, αφού ήταν αδύνατο να βρεθεί ένα ικανοποιητικό πλήθος φωτογραφιών ελεύθερα στο διαδίκτυο. Έτσι δημιουργήθηκε ένα εκτενές σύνολο δεδομένων το οποίο αποτελούταν από 12.720 φωτογραφίες υψηλής ανάλυσης, που υποβλήθηκαν σε ανάλυση, επεξεργασία και σήμανση, αποτελώντας τον πυρήνα της εκπαίδευσης των μοντέλων και της επίτευξης των αποτελεσμάτων.

Τα αποτελέσματα ανέδειξαν την ικανότητα των νευρωνικών δικτύων να εντοπίζουν ανεπαίσθητες στον ανθρώπινο οφθαλμό, οπτικές διαφορές. Μεταξύ των τριών μοντέλων, τα μοντέλα ResNet και Xception επέτυχαν τις καλύτερες επιδόσεις αγγίζοντας και τα δυο ακρίβεια ίση με 97%. Σε αντίθεση με το VGG16 κανένα από τα δυο αυτά μοντέλα δεν επηρεάστηκε από την ελαφριά ανισορροπία μεταξύ των κατηγοριών του συνόλου δεδομένων. Λόγω της μεγάλης ομοιότητας των αποτελεσμάτων τους είναι δύσκολο να διακριθεί εάν κάποιο μοντέλο ήταν σαφώς καλύτερο από το άλλο. Στο πλαίσιο της



απόδοσης του μοντέλου για κάθε κατηγορία το ResNet50 είχε ελαφρώς καλύτερη απόδοση για τις LVT,VPG,VRF και SSS καθώς η μετρική του f1-score σε εκείνες τις κατηγορίες ήταν 0.98 για τις δυο πρώτες και 0.97 για τις δυο επόμενες. Παρομοίως το Xception επιτυγχάνει ακριβώς ίδια απόδοση βάσει των μετρικών στις κατηγορίες LVT και VPG με f1-score 0.98 αλλά και στην κατηγορία VRF με 0.97. Επομένως, από τα αποτελέσματα φαίνεται πως τα μοντέλα ResNet και Xception αποτελούν μια εξαιρετική επιλογή για την δημιουργία ενός προβλεπτικού μοντέλου ταξινόμησης διαφορετικών ειδών και ποικιλιών σπόρων επιτυγχάνοντας υψηλού επιπέδου απόδοση.

Αντίθετα, το VGG16 δείχνει να έχει τη χειρότερη απόδοση σε σχέση με τα άλλα δύο επιτυγχάνοντας μόνο 65%. Πάρα την προσπάθεια που έγινε για την εξισορρόπηση του συνόλου δεδομένων δημιουργώντας περισσότερα δείγματα από τις μειοψηφικές τάξεις και εφαρμόζοντας δειγματοληψία, αυτή δεν επιτεύχθηκε, καθώς σε 2 από τις 6 κατηγορίες υπήρχαν αρκετά περισσότερες φωτογραφίες σε σχέση με τις υπόλοιπες. Όπως αναφέρθηκε και στην ενότητα 4.4 σε αρκετές περιπτώσεις εκπαίδευσης η κατηγορία Λεβαντε περιείχε περίπου 250-350 εικόνες παραπάνω και για τα τρία μοντέλα που δοκιμάστηκαν. Αυτό το γεγονός ίσως επηρέασε την διαδικασία εκπαίδευσης του συγκεκριμένου μοντέλου. Ενώ το VGG16 παρουσίασε δυσκολίες στην ακριβή ταξινόμηση, τα άλλα δύο μοντέλα κατάφεραν να επιτύχουν καλύτερα αποτελέσματα. Αυτό υποδηλώνει ότι, παρόλο που το VGG16 είναι ένα ευρέως γνωστό και χρησιμοποιημένο μοντέλο σε πολλές εφαρμογές βαθιάς μάθησης, μπορεί να μην είναι ιδανικό για το πρόβλημα που εξετάσαμε.

Στη συνέχεια έγινε μια εκτενής εξερεύνηση και σύγκριση των αποτελεσμάτων με άλλα έργα του πεδίου και παρουσιάστηκε μια ενδελεχής ανάλυση όλων των περιορισμών που αντιμετωπίζει η εφαρμογή της βαθιάς μάθησης στην αναγνώριση και ταξινόμηση σπορών. Συγκριτικά με άλλα έργα η καινοτομία της παρούσης είναι η διάκριση σε ένα ανώτερο ταξινομικό επίπεδο μεταξύ διαφορετικών ειδών και ποικιλιών το οποίο επιχειρείται σε ελάχιστες άλλες έρευνες, ενώ ταυτόχρονα καθιστά το σύστημα πιο ανθεκτικό και λιγότερο επιρρεπές σε υπερεκπαίδευση. Διαπιστώθηκε η ανάγκη καθιέρωσης επίσημων διαδικασιών, δεδομένου ότι η προτεινόμενη τεχνολογία δεν έχει ακόμη επαρκώς επιβεβαιωθεί. Επιπλέον, αναδεικνύεται η σημασία της εξέλιξης και εφαρμογής καινοτόμων μεθόδων που θα μπορούν να ανταποκριθούν στις προκλήσεις της περιορισμένης διαθεσιμότητας δεδομένων και του υψηλού κόστους που συνεπάγεται η χρήση συνελκτικών νευρωνικών δικτύων με μεγάλο βάθος.

Η ακριβής και αυτοματοποιημένη ταξινόμηση των ειδών σπόρων σε δυο επίπεδα του είδους και της ποικιλίας, χρησιμοποιώντας προηγμένες τεχνικές βαθιάς μηχανικής μάθησης, μπορεί να βοηθήσει σημαντικά τον γεωργικό τομέα. Η ταξινόμηση σπόρων είναι μια χρονοβόρα διαδικασία που απαιτεί πόρους και την συμβολή ειδικών του είδους ώστε να ελεγχθεί η ορθότητα της προέλευσης τους. Η αυτοματοποίηση αυτής της διαδικασίας δίνει την δυνατότητα στους αγρότες να εξοικονομήσουν χρόνο και πόρους, διασφαλίζοντας ταυτόχρονα την ακρίβεια και τη συνέπεια της διαδικασίας της ταξινόμησης των σπόρων. Το προτεινόμενο σύστημα πρόβλεψης θα μπορούσε επίσης να ενσωματωθεί σε μια εφαρμογή για smartphones προσφέροντας ένα χρήσιμο και φιλικό προς τον χρήστη εργαλείο που θα μπορεί να μεταφερθεί οπουδήποτε. Με την ενσωμάτωση εφαρμογών τεχνητής νοημοσύνης στον γεωργικό τομέα δημιουργούνται νέες δυνατότητες. Οι αγρότες μπορούν να επωφεληθούν λόγω της ακρίβειας και της αποτελεσματικότητάς που προσφέρουν αυτά τα συστήματα, ανοίγοντας νέους δρόμους για την βελτίωση της παραγωγικότητας και της βιωσιμότητας στην γεωργία.

## Βιβλιογραφία

- [1] “Geographic distribution of major crops across the world - Leff - 2004 - Global Biogeochemical Cycles - <https://doi.org/10.1029/2003GB002108>
- [2] G. Igrejas and G. Branlard, “The Importance of Wheat,” in *Wheat Quality For Improving Processing And Human Health*, G. Igrejas, T. M. Ikeda, and C. Guzmán, Eds., Cham: Springer International Publishing, 2020, pp. 1–7. doi: 10.1007/978-3-030-34163-3\_1.
- [3] J. M. Awika, “Major Cereal Grains Production and Use around the World,” in *Advances in Cereal Science: Implications to Food Processing and Health Promotion*, vol. 1089, 0 vols., in ACS Symposium Series, no. 1089, vol. 1089. , American Chemical Society, 2011, pp. 1–13. doi: 10.1021/bk-2011-1089.ch001.
- [4] D. V. Vinogradov, M. V. Evsenina, and A. V. Novikova, “Improving the conditioning of wheat grain when preparing it for grinding into graded flour,” *IOP Conf. Ser.: Earth Environ. Sci.*, vol. 723, no. 2, Art. no. 2, Nov. 2021, doi: 10.1088/1755-1315/723/2/022081.
- [5] “Computer Vision and Image Processing: A Paper Review | Wiley | International Journal of Artificial Intelligence Research.” <https://doi.org/10.29099/ijair.v2i1.42>
- [6] K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, “Machine Learning in Agriculture: A Review,” *Sensors*, vol. 18, no. 8, Art. no. 8, Aug. 2018, doi: 10.3390/s18082674.
- [7] A. Kamilaris, A. Kartakoullis, and F. X. Prenafeta-Boldú, “A review on the practice of big data analysis in agriculture,” *Computers and Electronics in Agriculture*, vol. 143, pp. 23–37, Dec. 2017, doi: 10.1016/j.compag.2017.09.037.
- [8] A. Kamilaris and F. X. Prenafeta-Boldú, “Deep learning in agriculture: A survey,” *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, Apr. 2018, doi: 10.1016/j.compag.2018.02.016.
- [9] M. Rinaldi and Z. He, “Chapter Six - Decision Support Systems to Manage Irrigation in Agriculture\*Present address: Consiglio per la Ricerca e la Sperimentazione in Agricoltura, Cereal Research Centre, S.S. 673km 25,200, 71122 Foggia, Italy,” in *Advances in Agronomy*, vol. 123, D. L. Sparks, Ed., Academic Press, 2014, pp. 229–279. doi: 10.1016/B978-0-12-420225-2.00006-6.
- [10] K. P. Ferentinos, “Deep learning models for plant disease detection and diagnosis,” *Computers and Electronics in Agriculture*, vol. 145, pp. 311–318, Feb. 2018, doi: 10.1016/j.compag.2018.01.009.
- [11] A. Yasar, “Benchmarking analysis of CNN models for bread wheat varieties,” *Eur Food Res Technol*, vol. 249, no. 3, Art. no. 3, Mar. 2023, doi: 10.1007/s00217-022-04172-y.
- [12] A. Loddo and C. Di Ruberto, “On the Efficacy of Handcrafted and Deep Features for Seed Image Classification,” *Journal of Imaging*, vol. 7, no. 9, Art. no. 9, Sep. 2021, doi: 10.3390/jimaging7090171.
- [13] A. Golcuk and A. Yasar, “Classification of bread wheat genotypes by machine learning algorithms,” *Journal of Food Composition and Analysis*, vol. 119, p. 105253, Jun. 2023, doi: 10.1016/j.jfca.2023.105253.
- [14] *Pattern Recognition and Machine Learning*. 2023.. Available: <https://link.springer.com/book/9780387310732>
- [15] W. Fang, X. Li, M. Zhang, and M. Hu, “Nature-Inspired Algorithms for Real-World Optimization Problems,” *Journal of Applied Mathematics*, vol. 2015, pp. 1–2, 2015, doi: 10.1155/2015/359203.
- [16] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, Art. no. 7553, May 2015, doi: 10.1038/nature14539.

- [17] M. Z. Alom *et al.*, “The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches,” Sep. 12, 2018, *arXiv*: arXiv:1803.01164. doi: 10.48550/arXiv.1803.01164.
- [18] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, “Deep Learning for Computer Vision: A Brief Review,” *Computational Intelligence and Neuroscience*, vol. 2018, p. e7068349, Feb. 2018, doi: 10.1155/2018/7068349.
- [19] S. Ray, “A Quick Review of Machine Learning Algorithms,” in *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, Oct. 2019, pp. 35–39. doi: 10.1109/COMITCon.2019.8862451.
- [20] X. (Jerry) Zhu, “Semi-Supervised Learning Literature Survey,” University of Wisconsin-Madison Department of Computer Sciences, Technical Report, 2005. Available: <https://minds.wisconsin.edu/handle/1793/60444>
- [21] S. J. Pan and Q. Yang, “A Survey on Transfer Learning,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, Jul. 2010, doi: 10.1109/TKDE.2009.191.
- [22] M. Iman, H. R. Arabnia, and K. Rasheed, “A Review of Deep Transfer Learning and Recent Advancements,” *Technologies*, vol. 11, no. 2, Art. no. 2, Apr. 2023, doi: 10.3390/technologies11020040.
- [23] Caruana, Rich. (1996). Algorithms and Applications for Multitask Learning. [https://www.researchgate.net/publication/2758125\\_Algorithms\\_and\\_Applications\\_for\\_Multitask\\_Learning](https://www.researchgate.net/publication/2758125_Algorithms_and_Applications_for_Multitask_Learning)
- [24] K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman, “Return of the Devil in the Details: Delving Deep into Convolutional Nets,” Nov. 05, 2014, *arXiv*: arXiv:1405.3531. doi: 10.48550/arXiv.1405.3531.
- [25] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” Dec. 10, 2015, *arXiv*: arXiv:1512.03385. Available: <http://arxiv.org/abs/1512.03385>
- [26] K. W. Church, “Word2Vec,” *Nat. Lang. Eng.*, vol. 23, no. 1, pp. 155–162, Jan. 2017, doi: 10.1017/S1351324916000334.
- [27] J. Pennington, R. Socher, and C. Manning, “GloVe: Global Vectors for Word Representation,” in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, A. Moschitti, B. Pang, and W. Daelemans, Eds., Doha, Qatar: Association for Computational Linguistics, Jul. 2014, pp. 1532–1543. doi: 10.3115/v1/D14-1162.
- [28] M. V. Koroteev, “BERT: A Review of Applications in Natural Language Processing and Understanding,” Mar. 22, 2021, *arXiv*: arXiv:2103.11943. doi: 10.48550/arXiv.2103.11943.
- [29] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, “Improving Language Understanding by Generative Pre-Training”. <https://www.semanticscholar.org/paper/Improving-Language-Understanding-by-Generative-Radford-Narasimhan/cd18800a0fe0b668a1cc19f2ec95b5003d0a5035#citing-papers>
- [30] H. E. Kim, A. Cosa-Linan, N. Santhanam, M. Jannesari, M. E. Maros, and T. Ganslandt, “Transfer learning for medical image classification: a literature review,” *BMC Med Imaging*, vol. 22, no. 1, p. 69, Apr. 2022, doi: 10.1186/s12880-022-00793-7.
- [31] A. Koshiyama, S. B. Blumberg, N. Firoozye, P. Treleaven, and S. Flennerhag, “QuantNet: transferring learning across trading strategies,” *Quantitative Finance*, vol. 22, no. 6, pp. 1071–1090, Jun. 2022, doi: 10.1080/14697688.2021.1999487.
- [32] K. Weiss, T. M. Khoshgoftaar, and D. Wang, “A survey of transfer learning,” *J Big Data*, vol. 3, no. 1, p. 9, May 2016, doi: 10.1186/s40537-016-0043-6.

- [33] N. C. Eli-Chukwu, "Applications of Artificial Intelligence in Agriculture: A Review," *Eng. Technol. Appl. Sci. Res.*, vol. 9, no. 4, Art. no. 4, Aug. 2019, doi: 10.48084/etasr.2756.
- [34] D. Hall, C. McCool, F. Dayoub, N. Sunderhauf, and B. Upcroft, "Evaluation of Features for Leaf Classification in Challenging Conditions," in *2015 IEEE Winter Conference on Applications of Computer Vision*, Jan. 2015, pp. 797–804. doi: 10.1109/WACV.2015.111.
- [35] C. Di Ruberto and L. Putzu, "A fast leaf recognition algorithm based on SVM classifier and high dimensional feature vector," in *2014 International Conference on Computer Vision Theory and Applications (VISAPP)*, Jan. 2014, pp. 601–609.
- [36] J. Amara, B. Bouaziz, and A. Algergawy, "A Deep Learning-based Approach for Banana Leaf Diseases Classification," presented at the *Datenbanksysteme für Business, Technologie und Web (BTW 2017) - Workshopband*, Gesellschaft für Informatik e.V., 2017, pp. 79–88, [https://www.researchgate.net/publication/316561053\\_A\\_Deep\\_Learning-based\\_Approach\\_for\\_Banana\\_Leaf\\_Diseases\\_Classification](https://www.researchgate.net/publication/316561053_A_Deep_Learning-based_Approach_for_Banana_Leaf_Diseases_Classification)
- [37] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification," *Computational Intelligence and Neuroscience*, vol. 2016, p. e3289801, Jun. 2016, doi: 10.1155/2016/3289801.
- [38] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using Deep Learning for Image-Based Plant Disease Detection," *Frontiers in Plant Science*, vol. 7, 2016. <https://doi.org/10.3389/fpls.2016.01419>
- [39] R. Gajjar, N. Gajjar, V. J. Thakor, N. P. Patel, and S. Ruparelia, "Real-time detection and identification of plant leaf diseases using convolutional neural networks on an embedded platform," *Vis Comput*, vol. 38, no. 8, Art. no. 8, Aug. 2022, doi: 10.1007/s00371-021-02164-9.
- [40] K. S. Jamuna, S. Karpagavalli, M. S. Vijaya, P. Revathi, S. Gokilavani, and E. Madhiya, "Classification of Seed Cotton Yield Based on the Growth Stages of Cotton Crop Using Machine Learning Techniques," in *2010 International Conference on Advances in Computer Engineering*, Jun. 2010, pp. 312–315. doi: 10.1109/ACE.2010.71.
- [41] U. Škrubej, Č. Rozman, and D. Stajnko, "Assessment of germination rate of the tomato seeds using image processing and machine learning," *European Journal of Horticultural Science*, vol. 80, pp. 68–75, Apr. 2015, doi: 10.17660/eJHS.2015/80.2.4.
- [42] S. Fayyazi, M. H. Abbaspour-Fard, A. Rohani, S. A. Monadjemi, and H. Sadrnia, "Identification and Classification of Three Iranian Rice Varieties in Mixed Bulks Using Image Processing and MLP Neural Network," *International Journal of Food Engineering*, vol. 13, no. 5, May 2017, doi: 10.1515/ijfe-2016-0121.
- [43] K. Kiratiratanapruk *et al.*, "Development of Paddy Rice Seed Classification Process using Machine Learning Techniques for Automatic Grading Machine," *Journal of Sensors*, vol. 2020, p. e7041310, Jul. 2020, doi: 10.1155/2020/7041310.
- [44] Z. Qiu, J. Chen, Y. Zhao, S. Zhu, Y. He, and C. Zhang, "Variety Identification of Single Rice Seed Using Hyperspectral Imaging Combined with Convolutional Neural Network," *Applied Sciences*, vol. 8, no. 2, Art. no. 2, Feb. 2018, doi: 10.3390/app8020212.
- [45] Z. Luan, C. Li, S. Ding, M. Wei, and Y. Yang, "Sunflower seed sorting based on Convolutional Neural Network," in *Eleventh International Conference on Graphics and Image Processing (ICGIP 2019)*, SPIE, Jan. 2020, pp. 428–434. doi: 10.1117/12.2557789.

- [46] P. Xu, Q. Tan, Y. Zhang, X. Zha, S. Yang, and R. Yang, "Research on Maize Seed Classification and Recognition Based on Machine Vision and Deep Learning," *Agriculture*, vol. 12, no. 2, Art. no. 2, Feb. 2022, doi: 10.3390/agriculture12020232.
- [47] M. F. Unlarsen *et al.*, "CNN–SVM hybrid model for varietal classification of wheat based on bulk samples," *Eur Food Res Technol*, vol. 248, no. 8, Art. no. 8, Aug. 2022, doi: 10.1007/s00217-022-04029-4.
- [48] K. Laabassi, M. A. Belarbi, S. Mahmoudi, S. A. Mahmoudi, and K. Ferhat, "Wheat varieties identification based on a deep learning approach," *Journal of the Saudi Society of Agricultural Sciences*, vol. 20, no. 5, Art. no. 5, Jul. 2021, doi: 10.1016/j.jssas.2021.02.008.
- [49] J. Lin, J. Li, W. Li, H. Qin, and S. Chen, "A data transfer method for improving seed identification of maize (*Zea mays*) haploid breeding based on genetic similarity," *Plant Breeding*, vol. 138, no. 6, pp. 790–801, 2019, doi: 10.1111/pbr.12746.
- [50] H. Zhu, L. Yang, J. Fei, L. Zhao, and Z. Han, "Recognition of carrot appearance quality based on deep feature and support vector machine," *Computers and Electronics in Agriculture*, vol. 186, p. 106185, Jul. 2021, doi: 10.1016/j.compag.2021.106185.
- [51] M. H. Junos, A. S. Mohd Khairuddin, S. Thannirmalai, and M. Dahari, "Automatic detection of oil palm fruits from UAV images using an improved YOLO model," *Vis Comput*, vol. 38, no. 7, Art. no. 7, Jul. 2022, doi: 10.1007/s00371-021-02116-3.
- [52] *HumanSignal/labelImg*. Python. HumanSignal. Available: <https://github.com/HumanSignal/labelImg>
- [53] A. D. Pozzolo, O. Caelen, R. A. Johnson, and G. Bontempi, "Calibrating Probability with Undersampling for Unbalanced Classification," in *2015 IEEE Symposium Series on Computational Intelligence*, Sep. 2015, pp. 159–166. doi: 10.1109/SSCI.2015.33.
- [54] E. Burnaev, P. Erofeev, and A. Papanov, "Influence of resampling on accuracy of imbalanced classification," in *Eighth International Conference on Machine Vision (ICMV 2015)*, SPIE, Dec. 2015, pp. 423–427. doi: 10.1117/12.2228523.
- [55] C. Szegedy *et al.*, "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA: IEEE, Jun. 2015, pp. 1–9. doi: 10.1109/CVPR.2015.7298594.
- [56] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing & Management*, vol. 45, no. 4, Art. no. 4, Jul. 2009, doi: 10.1016/j.ipm.2009.03.002.
- [57] D. D. Ramyachitra and P. Manikandan, "IMBALANCED DATASET CLASSIFICATION AND SOLUTIONS: A REVIEW," *International Journal of Computing and Business Research*, vol. 5, no. 4, 2014.
- [58] A. Dell'Aquila, "Towards new computer imaging techniques applied to seed quality testing and sorting," *Seed Science and Technology*, vol. 35, no. 3, pp. 519–538, Oct. 2007, doi: 10.15258/sst.2007.35.3.01.
- [59] Invited Review: Automated seed identification with computer vision: challenges and opportunities," *ResearchGate*, Oct. 2024, doi: 10.15258/sst.2022.50.1. s.05.
- [60] "CFIA\_ACIA-#3468880-v11-Candidates\_Guide\_to\_Seed\_Analyst\_Accreditation.pdf." Available: [https://csi-ics.com/sites/default/files/CFIA\\_ACIA-%233468880-v11-Candidates\\_Guide\\_to\\_Seed\\_Analyst\\_Accreditation.pdf](https://csi-ics.com/sites/default/files/CFIA_ACIA-%233468880-v11-Candidates_Guide_to_Seed_Analyst_Accreditation.pdf)
- [61] I. S. T. Association, "Seed Analyst Training," International Seed Testing Association. Available: <https://www.seedtest.org/en/technical-committees/seed-analyst-training.html>

- [62] X. Yi, M. Eramian, R. Wang, J. Neudorf, A. Salzl, and J. Maruschak, "Seed Identification Using Computer Vision — A Proof of Concept Study," *Seed Technology*, vol. 37, no. 1, pp. 93–100, 2016.
- [63] "Publications – Analyzeseeds." Available: <https://analyzeseeds.com/publications/>
- [64] I. S. T. Association, "How to use the ISTA Rules," International Seed Testing Association. Available: <https://www.seedtest.org/en/publications/international-rules-seed-testing/how-use-ista-rules.html>
- [65] A. Tsirikoglou, G. Eilertsen, J. Unger, "A Survey of Image Synthesis Methods for Visual Machine Learning," *ResearchGate*, Oct. 2024, doi: 10.1111/cgf.14047.
- [66] K. Man and J. Chahl, "A Review of Synthetic Image Data and Its Use in Computer Vision," *Journal of Imaging*, vol. 8, no. 11, Art. no. 11, Nov. 2022, doi: 10.3390/jimaging8110310.