



Τεχνολογικό
Πανεπιστήμιο
Κύπρου

Σχολή Επικοινωνίας και
Μέσων Ενημέρωσης

Πτυχιακή εργασία

**ΜΕΛΕΤΗ ΤΩΝ ΧΑΡΤΩΝ ΘΕΡΜΟΤΗΤΑΣ (HEATMAPS)
ΠΟΥ ΔΗΜΙΟΥΡΓΟΥΝΤΑΙ ΚΑΤΑ ΤΟΝ ΕΝΤΟΠΙΣΜΟ
ΕΜΠΟΔΙΩΝ ΠΟΥ ΘΕΤΟΥΝ ΤΗΝ ΖΩΗ ΤΩΝ ΠΕΖΩΝ ΣΕ
ΚΙΝΔΥΝΟ: ΔΙΑΦΟΡΕΣ ΜΕΤΑΞΥ ΑΝΘΡΩΠΟΥ ΚΑΙ
ΜΗΧΑΝΗΣ**

Αικατερίνη Καρακασίδη

Λεμεσός, Μάιος 2023

ΤΕΧΝΟΛΟΓΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ
ΣΧΟΛΗ ΕΠΙΚΟΙΝΩΝΙΑΣ ΚΑΙ ΜΕΣΩΝ ΕΝΗΜΕΡΩΣΗΣ
ΤΜΗΜΑ ΕΠΙΚΟΙΝΩΝΙΑΣ ΚΑΙ ΣΠΟΥΔΩΝ ΔΙΑΔΙΚΤΥΟΥ

Πτυχιακή εργασία

ΜΕΛΕΤΗ ΤΩΝ ΧΑΡΤΩΝ ΘΕΡΜΟΤΗΤΑΣ (HEATMAPS) ΠΟΥ
ΔΗΜΙΟΥΡΓΟΥΝΤΑΙ ΚΑΤΑ ΤΟΝ ΕΝΤΟΠΙΣΜΟ ΕΜΠΟΔΙΩΝ
ΠΟΥ ΘΕΤΟΥΝ ΤΗΝ ΖΩΗ ΤΩΝ ΠΕΖΩΝ ΣΕ ΚΙΝΔΥΝΟ:
ΔΙΑΦΟΡΕΣ ΜΕΤΑΞΥ ΑΝΘΡΩΠΟΥ ΚΑΙ ΜΗΧΑΝΗΣ

της

Αικατερίνης Καρακασίδη

Επιβλέπων Καθηγητής

Δρ. Ζήνωνας Θεοδοσίου

Λεμεσός, Μάιος 2023

Πνευματικά δικαιώματα

Copyright © Αικατερίνη Καρακασίδη, 2023

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Η έγκριση της πτυχιακής εργασίας από το Τμήμα Επικοινωνίας και Σπουδών Διαδικτύου του Τεχνολογικού Πανεπιστημίου Κύπρου δεν υποδηλώνει απαραίτητως και αποδοχή των απόψεων του συγγραφέα εκ μέρους του Τμήματος.

Θα ήθελα να ευχαριστήσω ιδιαίτερα τον επιβλέποντα καθηγητή μου, Δρ. Ζήνωνα Θεοδοσίου, του τμήματος Επικοινωνίας και Σπουδών Διαδικτύου, που δίχως την πολύτιμη βοήθεια και υποστήριξη του η πραγματοποίηση αυτής της πτυχιακής εργασίας θα ήταν αδύνατη. Ακόμη, θα ήθελα να τον ευχαριστήσω για τον πολύτιμο χρόνο που διέθεσε για να με κατευθύνει και να συμβουλευτεί, με αποτέλεσμα την ολοκλήρωση αυτής της εργασίας, παρέχοντας τις κατάλληλες κατευθυντήριες οδηγίες. Επίσης θα ήθελα να ευχαριστήσω τον Περικλή Περικλέους που δίχως την πολύτιμη βοήθειά του δεν θα μπορούσα να πραγματοποιήσω τα πειράματα που έχουν διεξαχθεί. Τον ευχαριστώ πολύ για την βοήθεια του αλλά και τις συμβουλές του που μου έδωσε στα πλαίσια αυτής της εργασίας. Τέλος, θα ήθελα να ευχαριστήσω τους καθηγητές και τους συμφοιτητές μου στο τμήμα Επικοινωνίας και Σπουδών Διαδικτύου οι οποίοι, ήταν πάντα στο πλάι μου, υποστηρίζοντας και βοηθώντας με κατά την διάρκεια της πτυχιακής μου εργασίας.

ΠΕΡΙΛΗΨΗ

Οι σύγχρονες πόλεις έχουν κατακερματιστεί από έναν αυξανόμενο αριθμό εμποδίων και ζημιών στις υποδομές που δημιουργούν διάφορα προβλήματα που θέτουν σε κίνδυνο τους πεζούς πολίτες. Η ανάπτυξη αυτοματοποιημένων μεθόδων για τον εντοπισμό και την αναγνώριση ανθρώπων, εμποδίων και ζημιών σε οπτικά δεδομένα για τη δημιουργία ασφαλών αστικών περιβαλλόντων απασχολεί ιδιαίτερα την ερευνητική κοινότητα τα τελευταία χρόνια. Η χρήση αλγορίθμων βαθιάς μάθησης αποτελεί πλέον την κυρίαρχη προσέγγιση στην ανάλυση οπτικών δεδομένων με πολύ καλά αποτελέσματα σε ένα ευρύ φάσμα εφαρμογών, συμπεριλαμβανομένου του εντοπισμού εμποδίων. Ωστόσο, η εξήγηση της λειτουργίας των μοντέλων μάθησης παραμένει μια βασική πρόκληση για την απόκτηση σημαντικών γνώσεων σχετικά με τον τρόπο εκπαίδευσης των αλγορίθμων. Οι χάρτες θερμότητας που δημιουργούνται στο στάδιο εκπαίδευσης των αλγορίθμων βαθιάς μάθησης αναδεικνύουν τα σημεία εστίασης στην εικόνα εισόδου που βοήθησαν το παραγόμενο μοντέλο να προβλέψει την έξοδο και θα μπορούσαν να αποτελέσουν μια καλή αρχή προς αυτή την κατεύθυνση. Η παρούσα εργασία παρουσιάζει τις ομοιότητες μεταξύ των χαρτών θερμότητας που δημιουργούνται από αλγορίθμους βαθιάς μάθησης που εκπαιδεύτηκαν να εντοπίζουν εμπόδια σε πεζοδρόμια σε εικόνες που συλλέχθηκαν μέσω κάμερας κινητού τηλεφώνου, και των χαρτών θερμότητας που δημιουργούνται από ανθρώπους καθώς εντοπίζουν τα αντίστοιχα εμπόδια στα ίδια δεδομένα. Αρχικά δημιουργήθηκαν οι χάρτες θερμότητας για 20 εικόνες που αφορούσαν 10 διαφορετικά εμπόδια για δύο αλγόριθμους αναγνώρισης εικόνας που βασίζονται στις αρχιτεκτονικές των συνελικτικών δικτύων και οπτικών μετασχηματιστών, αντίστοιχα. Οι εικόνες δόθηκαν στη συνέχεια σε 35 χρήστες από τους οποίους ζητήθηκε να εντοπίσουν τα εμπόδια δημιουργώντας τους χάρτες θερμότητας με τη βοήθεια ενός eye tracker. Οι χάρτες θερμότητας συγκρίθηκαν οπτικά μεταξύ τους οδηγώντας σε ενδιαφέροντα συμπεράσματα σχετικά με την αντίληψη εμποδίων από τον άνθρωπο και τη μηχανή, καθώς και τις ομοιότητες μεταξύ των σημείων εστίασης τους όταν εντοπίζουν διαφορετικά εμπόδια που θέτουν σε κίνδυνο την ασφάλεια των πεζών. Τα αποτελέσματα δείχνουν ότι οι χάρτες θερμότητας που δημιουργούνται από ανθρώπους μοιάζουν περισσότερο με την αρχιτεκτονική οπτικών μετασχηματιστών.

Λέξεις κλειδιά: χάρτες θερμότητας, eye tracker, αλγόριθμοι βαθιάς μάθησης, επεξηγησιμότητα μοντέλων

ABSTRACT

Contemporary cities have been fragmented by a growing number of construction barriers and infrastructure damages that generate several problems that set pedestrian citizens at risk. The development of automated methods for detecting and recognizing people, barriers, and damages in visual data to create safe urban environments has been of particular concern to the research community in recent years. The use of deep learning algorithms is now the dominant approach in visual data analysis with very good results in a wide range of applications including obstacle detection. However, the explainability of the deep learning models remains a key challenge in gaining significant knowledge about how algorithms are trained. Post-hoc explainability for deep learning models using heatmaps highlight the focal points in the input image that helped the generated model predict the output and could be a good start in that direction. This work presents the similarities between heatmaps generated by deep learning algorithms trained while training models to detect obstacles on sidewalks in images collected via a smartphone camera, and heatmaps generated by humans as they detect the corresponding obstacles on the same data. Heatmaps of 20 images related to 10 different obstacles were first generated by two state-of-the-art image recognition algorithms based on convolutional neural network and vision transform architectures, respectively. The image data set was then given to 35 users who were asked to locate the obstacles by generating the heatmaps with the help of an eye-tracker. The heatmaps were visually compared using a multi-grid step approach, yielding interesting insights into human and machine obstacle perception, as well as the similarities between their focus points when detecting different obstacles that put pedestrian safety at risk. The results indicate that heatmaps created by humans resemble more closely to the vision transformer architecture.

Keywords: heatmaps, eye tracker, deep learning algorithms, model explainability

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΠΕΡΙΛΗΨΗ.....	5
ABSTRACT.....	7
ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ	8
ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ	10
ΚΑΤΑΛΟΓΟΣ ΔΙΑΓΡΑΜΜΑΤΩΝ	12
ΣΥΝΤΟΜΟΓΡΑΦΙΕΣ	13
ΑΠΟΔΟΣΗ ΟΡΩΝ	14
1 Εισαγωγή	15
1.1 Αναγκαιότητα Μελέτης	15
1.2 Βιβλιογραφική Ανασκόπηση	15
1.2.1 Μεγάλα Δεδομένα	15
1.2.2 Μηχανική Μάθηση	17
1.2.3 Βαθιά Μάθηση.....	21
1.2.4 Συνελκτικά Νευρωνικά Δίκτυα	22
1.2.5 Μηχανική Όραση.....	25
1.2.6 Επεξηγησιμότητα των Μοντέλων Βαθιάς Μάθησης.....	27
1.2.7 Χάρτες Θερμότητας.....	28
1.2.8 Eye Tracker	30
1.3 Ερευνητικά Ερωτήματα	30
1.3.1.1 Μεταβλητές	31
2 Μεθοδολογία Έρευνας.....	32
2.1 Δεδομένα.....	32
2.2 Συσκευή eye-tracker.....	32
2.3 Σχεδιασμός Πειράματος.....	33

2.4	Δημιουργία Χαρτών Θερμότητας	35
3	Ανάλυση	38
3.1	Χάρτες Θερμότητας Μηχανής	38
3.2	Χάρτες Θερμότητας Ανθρώπου	38
3.3	Ανάλυση Δεδομένων.....	42
3.4	Σύγκριση διαφορών.....	47
4	Αποτελέσματα.....	52
5	Συζήτηση	54
5.1	Μελλοντική Εργασία	55
	ΒΙΒΛΙΟΓΡΑΦΙΑ	56
	ΠΑΡΑΡΤΗΜΑ Ι	61

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

Πίνακας 1: Εικόνες που παρουσιάστηκαν στο Part A του πειράματος	33
Πίνακας 2: Εικόνες που παρουσιάστηκαν στο Part B του πειράματος	34
Πίνακας 3: Οι χάρτες θερμότητας που συλλέχθηκαν από το πρώτο μέρος των πειραμάτων	40
Πίνακας 4: Οι χάρτες θερμότητας που συλλέχθηκαν από το δεύτερο μέρος των πειραμάτων	40
Πίνακας 5: Οι χάρτες θερμότητας που συλλέχθηκαν από το πρώτο και δεύτερο μέρος των πειραμάτων πάνω στις εικόνες των εμποδίων σε έγχρωμη μορφή και οι χάρτες θερμότητας που δημιουργήθηκαν από τον κάθε αλγόριθμο.....	41
Πίνακας 6: Μέσος όρος των χαρτών όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο MobileNetV2 και των χαρτών όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο ViT-B16.....	48
Πίνακας 7: T-test ανάλυση για τους χάρτες όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο MobileNetV2 και των χαρτών όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο ViT-B16	48
Πίνακας 8: Μέσος όρος των χαρτών θερμότητας των αντρών σε σύγκριση με τον Αλγόριθμο MobileNetV2 και των χαρτών των αντρών σε σύγκριση με τον Αλγόριθμο ViT-B16	48
Πίνακας 9: T-test ανάλυση για τους χάρτες των αντρών σε σύγκριση με τον Αλγόριθμο MobileNetV2 και των χαρτών των αντρών σε σύγκριση με τον Αλγόριθμο ViT-B16..	48
Πίνακας 10: Μέσος όρος των χαρτών θερμότητας των γυναικών σε σύγκριση με τον Αλγόριθμο MobileNetV2 και των χαρτών των γυναικών σε σύγκριση με τον Αλγόριθμο ViT-B16	49
Πίνακας 11: T-test ανάλυση για τους χάρτες των γυναικών σε σύγκριση με τον Αλγόριθμο MobileNetV2 και των χαρτών των γυναικών σε σύγκριση με τον Αλγόριθμο ViT-B16	49
Πίνακας 12: Μέσος όρος των χαρτών θερμότητας του πρώτου μέρους όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο MobileNetV2 και των χαρτών	

του δεύτερου μέρους όλων των συμμετεχόντων σε σύγκριση με τον Αλγόριθμο MobileNetV2	49
Πίνακας 13: T-test μεταξύ των χαρτών θερμότητας του πρώτου μέρους όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο MobileNetV2 και των χαρτών του δεύτερου μέρους όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο MobileNetV2	49
Πίνακας 14: Μέσος όρος των χαρτών θερμότητας του πρώτου μέρους όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο ViT-B16 και των χαρτών του δεύτερου μέρους όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο ViT-B16.....	50
Πίνακας 15: T-test μεταξύ των χαρτών θερμότητας του πρώτου μέρους όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο ViT-B16 και των χαρτών του δεύτερου μέρους όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο ViT-B16.....	50

ΚΑΤΑΛΟΓΟΣ ΔΙΑΓΡΑΜΜΑΤΩΝ

Διάγραμμα 1: Ο όγκος των δεδομένων παγκοσμίως που δημιουργείται , συλλέγεται, αντιγράφεται και καταναλώνεται από το 2010 μέχρι το 2020, με προβλέψεις από το 2021 μέχρι το 2025 (σε zettabytes). Πηγή: Statista: Worldwide Data Created	16
Διάγραμμα 2: Οι προσεγγίσεις της μηχανικής μάθησης. Πηγή: C A Padmanabba Reddy et al, 2028: 82	20
Διάγραμμα 3: Η εξέλιξη των μεθόδων μηχανικής όρασης.....	26
Διάγραμμα 4: Σύγκριση των heatmaps όλων των συμμετεχόντων/ουσών με τα heatmaps του αλγορίθμου MobileNetV2. Μπάρα μπλε: Part A του πειράματος, μπάρα πράσινη: Part B του πειράματος	43
Διάγραμμα 5: Σύγκριση των heatmaps όλων των αντρών που συμμετείχαν με τα heatmaps του αλγορίθμου MobileNetV2. Μπάρα μπλε: Part A του πειράματος, μπάρα πράσινη: Part B του πειράματος	43
Διάγραμμα 6: Σύγκριση των heatmaps όλων των γυναικών που συμμετείχαν με τα heatmaps του αλγορίθμου MobileNetV2. Μπάρα μπλε: Part A του πειράματος, μπάρα πράσινη: Part B του πειράματος	43
Διάγραμμα 7: Σύγκριση των heatmaps όλων των συμμετεχόντων/ουσών με τα heatmaps του αλγορίθμου ViT-B16. Μπάρα μπλε: Part A του πειράματος, μπάρα πράσινη: Part B του πειράματος.....	44
Διάγραμμα 8: Σύγκριση των heatmaps όλων των αντρών που συμμετείχαν με τα heatmaps του αλγορίθμου ViT-B16. Μπάρα μπλε: Part A του πειράματος, μπάρα πράσινη: Part B του πειράματος	44
Διάγραμμα 9: Σύγκριση των heatmaps όλων των γυναικών που συμμετείχαν με τα heatmaps του αλγορίθμου ViT-B16. Μπάρα μπλε: Part A του πειράματος, μπάρα πράσινη: Part B του πειράματος	45
Διάγραμμα 10: Σύγκριση των heatmaps των αντρών με των γυναικών. Μπάρα μπλε: Part A του πειράματος, μπάρα πράσινη: Part B του πειράματος.....	45
Διάγραμμα 11: Σύγκριση των δύο αλγορίθμων (MobileNetV2 και ViT-B16). Μπάρα μπλε: Part A του πειράματος, μπάρα πράσινη: Part B του πειράματος.....	46

ΣΥΝΤΟΜΟΓΡΑΦΙΕΣ

AI:	Artificial Intelligence
ANN:	Artificial Neural Networks
BD:	Big data
CNN:	Convolutional Neural Networks
CRM:	Customer Relationship Management
DR:	Dimensionality Reduction
ERP:	Enterprise Resource Planning
GCN:	Graph Convolutional Neural Networks
ML:	Machine Learning
OCR:	Optical Character Recognition
ROI:	Return On Investment
SVM:	Support Vector Machines

ΑΠΟΔΟΣΗ ΟΡΩΝ

Big Data	Μεγάλα Δεδομένα
Black Box	Μαύρο Κουτί/ Επεξηγησιμότητα των Μοντέλων Βαθιάς Μάθησης
Convolutional Neural Networks	Συνελικτικά Νευρωνικά Δίκτυα
Heatmap	Χάρτης θερμότητας/ προεξοχής
Pattern Recognition	Αναγνώριση Προτύπων

1 Εισαγωγή

1.1 Αναγκαιότητα Μελέτης

Η ασφάλεια των πεζών έχει απασχολήσει ιδιαίτερα την ερευνητική κοινότητα λόγω του μεγάλου αριθμού πολιτών που τραυματίζονται ή πεθαίνουν στους δρόμους. Ένας μεγάλος αριθμός εμποδίων φράσσει την ελεύθερη διακίνηση των πεζών στη πόλη μεταξύ αυτών διάφορες ζημιές ή τρύπες στα πεζοδρόμια, σταθμευμένα αυτοκίνητα, δέντρα, κάδοι σκουπιδιών, κτλ. Η χρήση αυτοματοποιημένων μεθόδων αναγνώρισης των εμποδίων σε οπτικά δεδομένα θα μπορούσε να βοηθήσει τις τοπικές αρχές να περιορίσουν τα εμπόδια και να αυξήσουν την ασφάλεια των πολιτών.

Η βαθιά μάθηση σήμερα αποτελεί την κύρια μέθοδο ανάλυσης οπτικών δεδομένων και απ αριθμεί μεγάλα επιτεύγματα τα τελευταία χρόνια. Παρόλα αυτά, η δυσκολία επεξήγησης των παραγόμενων μοντέλων αποτελεί βασικό πρόβλημα για να γίνει αντιληπτό το πως το παραγόμενο μοντέλο παράγει την επιθυμητή έξοδο για δοθείσα είσοδο.

Με στόχο την επεξηγησιμότητα των μοντέλων βαθιάς μάθησης, η παρούσα πτυχιακή αφορά την μελέτη των χαρτών θερμότητας (heatmaps) που δημιουργούνται από τους αλγόριθμους βαθιάς μάθησης και του ανθρώπους κατά τον εντοπισμό εμποδίων που θέτουν την ζωή των πεζών σε κίνδυνο. Συγκεκριμένα, ο σκοπός της πτυχιακής εργασίας είναι η σύγκριση των χαρτών θερμότητας δύο αλγορίθμων βαθιάς μάθησης με τους χάρτες θερμότητας που δημιουργούνται από τους ανθρώπους όταν εντοπίζουν εμπόδια στα πεζοδρόμια μέσα σε εικόνες που συλλέχθηκαν από κάμερα κινητού τηλεφώνου.

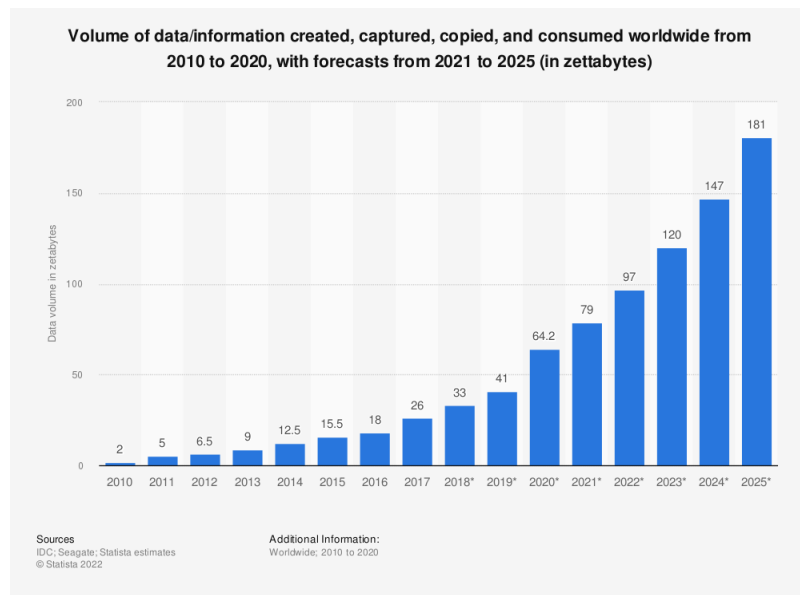
1.2 Βιβλιογραφική Ανασκόπηση

Σε αυτή την ενότητα ακολουθεί βιβλιογραφική ανασκόπηση του θέματος που μελετά η παρούσα πτυχιακή εργασία.

1.2.1 Μεγάλα Δεδομένα

Η ανάγκη για την ανάπτυξη αποτελεσματικών διαχείρισης του μεγάλου όγκου δεδομένων που δημιουργούνται καθημερινά οδήγησε την ερευνητική κοινότητα στην

δημιουργία του όρου «Μεγάλα Δεδομένα» (Big Data). Αυτή η ανάπτυξη απεικονίζεται στο Διάγραμμα 1, το οποίο παρουσιάζει την αύξηση των δεδομένων από το 2010 μέχρι και το 2025.



Διάγραμμα 1: Ο όγκος των δεδομένων παγκοσμίως που δημιουργείται, συλλέγεται, αντιγράφεται και καταναλώνεται από το 2010 μέχρι το 2020, με προβλέψεις από το 2021 μέχρι το 2025 (σε zettabytes). Πηγή: Statista: *Worldwide Data Created*

Ο όρος «Μεγάλα Δεδομένα» χρησιμοποιείται για την περιγραφή της συλλογής, ανάλυσης, επεξεργασίας και αναπαράστασης μεγάλων συνόλων δεδομένων (Emmanuel & Stanier, 2016: 1). Σύμφωνα με τους Emmanuel & Stanier (2016) χρησιμοποιείται για να καλύψει ένα εύρος δεδομένων, τεχνολογιών και εφαρμογών.

Οι George et al. (2014b: 1) αναφέρουν ότι τα Μεγάλα Δεδομένα έχουν αναδειχθεί από πηγές δεδομένων των οποίων ο πληθυσμός τους έχει αυξηθεί απότομα. Αυτές οι πηγές είναι τα διαδικτυακά clicks, οι κινητές συναλλαγές, το περιεχόμενο που δημιουργείται από τους χρήστες ή από επιχειρηματικές συναλλαγές. Τα δεδομένα που συλλέγονται αφορούν βίντεο, εικόνες, κείμενο, ήχο, αρχεία καταγραφής δεδομένων, emails, tweets, κτλ. (Mohanty et al., 2015).

Σύμφωνα με το επιστημονικό άρθρο «BigData: A Review» (Sagiroglu & Sinanc, 2013), το περιεχόμενο των δεδομένων κατατάσσεται σε τρεις κατηγορίες: δομημένο, ημι-δομημένο και αδόμητο. Δομημένο είναι το περιεχόμενο το οποίο χαρακτηρίζεται από

ετικέτες για την ταξινόμηση και την μετέπειτα κατηγοριοποίηση αυτών των δεδομένων σε αντίθεση με το αδόμητο περιεχόμενο το οποίο δύσκολα αναλύεται. Ημι-δομημένο περιεχόμενο έχουν τα δεδομένα τα οποία διαθέτουν ετικέτες για την κατηγοριοποίηση τους αλλά δεν μπορούν να ταξινομηθούν σε όλα τα πεδία μιας βάσης δεδομένων.

Τα δεδομένα κατηγοριοποιούνται ανάλογα με την ανάλυση που θα πρέπει να γίνει σε αυτά. Με βάση την αναφορά των Mohanty et al. (2015) στο βιβλίο τους «BigData» τα δεδομένα χαρακτηρίζονται από: πυκνότητα τιμής/πληροφορίας (value density/information), ευκολία ανάλυσης (analytic agility), συχνότητα και ταυτόχρονη πρόσβαση (frequency and concurrency of access).

Σε αντίθεση με τις παραδοσιακές βάσεις δεδομένων, τα μεγάλα δεδομένα, όπως αναφέρει και η NIST (National Institute of Standards and Technology) χαρακτηρίζονται από τα 3Vs: όγκος (Volume), ποικιλία (Variety) και ταχύτητα παραγωγής δεδομένων (Velocity) (Sagiroglu & Sinanc, 2013; Mohanty et al., 2015; Emmanuel & Stanier, 2016). Πιο αναλυτικά, το χαρακτηριστικό του όγκου των μεγάλων δεδομένων αναφέρεται στα δεδομένα τα οποία είναι από τη μία τεράστια και από την άλλη διατηρούν και επεκτείνουν ένα συγκεκριμένο μοτίβο για να δείξουν την σειρά αλληλουχίας που υπάρχει μέσα σε αυτά (Mohanty et al., 2015).

Ο καθορισμός των χαρακτηριστικών των μεγάλων δεδομένων γίνεται όλο και πιο περίπλοκος λόγω της συνεχούς αύξησης τους. Κάποιοι επιστήμονες θεωρούν πως υπάρχουν 4 χαρακτηριστικά και άλλοι από 5 μέχρι 7 χαρακτηριστικά. Τα πιο βασικά, όμως, είναι αυτά που έχουν αναφερθεί προηγουμένως: όγκος, ποικιλία και ταχύτητα παραγωγής δεδομένων (George et al., 2014b).

1.2.2 Μηχανική Μάθηση

Η μηχανική μάθηση είναι περιοχή της Τεχνητής Νοημοσύνης (Artificial Intelligence - AI) η οποία δίνει την ικανότητα σε ένα υπολογιστικό σύστημα να αντιλαμβάνεται, επεξεργάζεται και να αλληλεπιδρά με το περιβάλλον στο οποίο βρίσκεται, όπως και ο άνθρωπος. Προσπαθεί με λίγα λόγια να μιμηθεί την ανθρώπινη συμπεριφορά και ευφυΐα μαθαίνοντας από την αλληλεπίδραση του με το περιβάλλον (Issam El Naqa et al., 2015). Με την μεταφορά στην επόμενη εποχή της τεχνολογίας, οι υψηλές απαιτήσεις και προβλήματα που παρουσιάστηκαν κατά την επεξεργασία δεδομένων

επιλύθηκαν με τους αλγόριθμους της μηχανικής μάθησης. Η υιοθέτηση της μηχανικής μάθησης στα στάδια της επεξεργασίας και ανάλυσης έχει οδηγήσει σε ευκολότερο και αποτελεσματικότερο χειρισμό των δεδομένων (Mahesh, 2020). Η επεξεργασία των μεγάλων δεδομένων με την βοήθεια της μηχανικής μάθησης επιτρέπει την αποκάλυψη μοτίβων που είναι πιο λεπτομερείς ως προς τις πιο έγκυρες και ακριβείς προβλέψεις (Zhou et al., 2017).

Οι αλγόριθμοι μηχανικής μάθησης κατηγοριοποιούνται σε 3 κατηγορίες: μάθηση με επίβλεψη (supervised), μάθηση χωρίς επίβλεψη (unsupervised) και ημι-επιβλεπόμενη μάθηση (semi-supervised). Ανάλογα σε ποια κατηγορία μάθησης ανήκει ο κάθε αλγόριθμος οδηγεί και στο ανάλογο αποτέλεσμα συμπεριλαμβανομένων της ταξινόμησης, ομαδοποίησης, παρεμβολής, συσχέτισης, κτλ.. Οι βασικοί τύποι αλγορίθμων μηχανικής μάθησης βασίζονται σε δέντρα αποφάσεων (Decision Trees), μηχανές διανυσμάτων υποστήριξης (SVM), κανόνες συσχέτισης (Associations Rules), τεχνητά νευρωνικά δίκτυα (Neural Networks), δίκτυα Bayes (Bayes Networks), ομαδοποίηση (Clustering), κτλ. (Wang et al., 2015). Κάθε αλγόριθμος περνάει από δύο φάσεις: την φάση της εκπαίδευσης και την φάση της δοκιμής (Taygi & G, 2019). Στη φάση της εκπαίδευσης ο αλγόριθμος δέχεται ένα σύνολο δεδομένων έτσι ώστε να μπορέσει να εκπαιδευτεί και να μπορέσει να περάσει στη φάση της δοκιμής όπου θα αξιολογηθεί για τα αποτελέσματα που εξάγει.

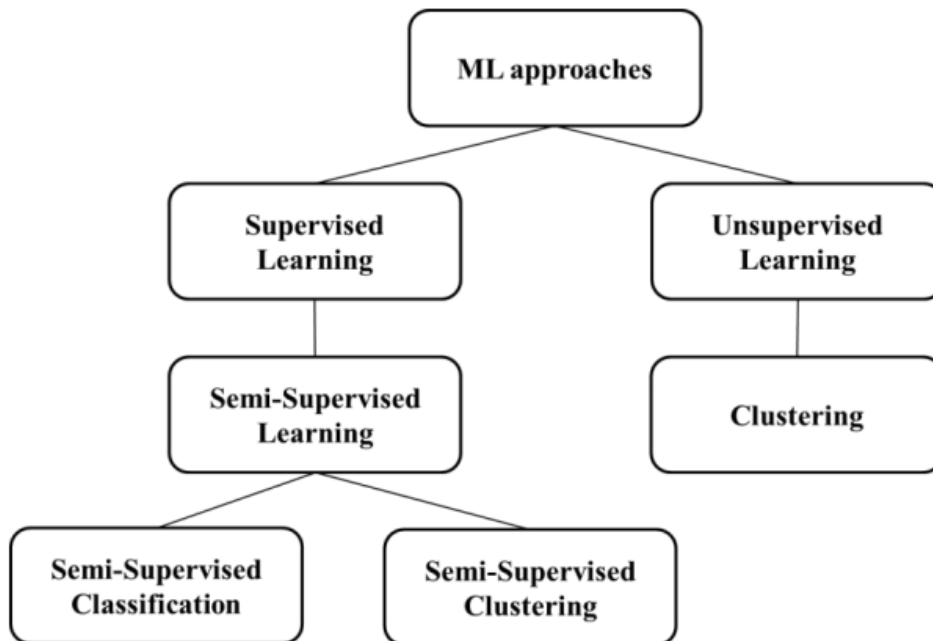
Πιο αναλυτικά, σύμφωνα με τον Mahesh στο επιστημονικό του άρθρο «Machine Learning Algorithms» (2020: 381), ένας αλγόριθμος που ανήκει στην κατηγορία μάθησης με επίβλεψη μαθαίνει διαμέσου ενός συνόλου δεδομένων που συμπεριλαμβάνει παραδείγματα εισόδου και εξόδου ούτως ώστε να κατηγοριοποιήσει σωστά τα δεδομένα. Κατά την εκπαίδευση του αλγορίθμου, εξάγονται πρότυπα μέσω των οποίων μαθαίνει ο αλγόριθμος να κατηγοριοποιεί τα δεδομένα του και στη συνέχεια τα εφαρμόζει για να κατηγοριοποιήσει άγνωστα δεδομένα. Από τα παραδείγματα αλγορίθμων που έχουν δοθεί νωρίτερα, τα δέντρα αποφάσεων, οι μηχανές διανυσμάτων υποστήριξης, και ο αφελής Bayes αποτελούν παραδείγματα αλγορίθμων μάθησης με επίβλεψη. Στην περίπτωση του δέντρου απόφασης ο αλγόριθμος έχει την μορφή δέντρου το οποίο αναπαριστά τις επιλογές που έχει να κάνει ο αλγόριθμος για να κατηγοριοποιήσει τα δεδομένα. Ο αφελής Bayes στηρίζεται στις πιθανότητες να ανήκουν τα δεδομένα σε κάποια κατηγορία με βάση τα χαρακτηριστικά

τους. Ο αλγόριθμος μηχανής διανυσμάτων υποστήριξης μεταφέρει τα δεδομένα σε ένα πολυεπίπεδο που μπορούν να είναι πλέον διαχωρίσιμα στις κατηγορίες στις οποίες ανήκουν (Mahesh, 2020).

Στην περίπτωση ενός αλγορίθμου χωρίς επίβλεψη δεν του δίνεται ο τρόπος με τον οποίο θα πρέπει να ταξινομηθούν τα δεδομένα που λαμβάνει στην είσοδο. Ο αλγόριθμος αναλύει τα δεδομένα εισόδου και τα ομαδοποιεί με βάση τα κοινά τους χαρακτηριστικά. Χρησιμοποιείται κυρίως για εφαρμογές ομαδοποίησης δεδομένων ή μείωσης των χαρακτηριστικών των δεδομένων (Mahesh, 2020).

Η ημι-επιβλεπόμενη μάθηση είναι συνδυασμός της μάθησης με επίβλεψη και της μάθησης χωρίς επίβλεψη. Ο πρωταρχικός της στόχος είναι να ξεπεράσει τα μειονεκτήματα των 2 κατηγοριών μάθησης. Η μάθηση με επίβλεψη απαιτεί ικανοποιητική ποσότητα επισημειωμένων δεδομένων (συνοδεύονται με την ετικέτα της κατηγορίας στην οποία ανήκουν) εκπαίδευσης προκειμένου να κατηγοριοποιηθούν σωστά, η οποία καταλήγει να είναι μη αποδοτική και χρονοβόρα διαδικασία. Από την άλλη, η μάθηση χωρίς επίβλεψη δεν απαιτεί επισημειωμένα δεδομένα εκπαίδευσης γιατί έχει την ικανότητα να ομαδοποιεί τα δεδομένα με βάση την ομοιότητα των χαρακτηριστικών. Το βασικό μειονέκτημα της μάθησης χωρίς επίβλεψη είναι ότι δεν μπορεί να ομαδοποιήσει με ακρίβεια άγνωστα δεδομένα. Έτσι, η ερευνητική κοινότητα εισήγαγε την ημι-επιβλεπόμενη μάθηση η οποία μπορεί να μάθει να κατηγοριοποιεί άγνωστα δεδομένα χρησιμοποιώντας μια μικρή ποσότητα δεδομένων εκπαίδευσης. Η ημι-επιβλεπόμενη μάθηση δημιουργεί ένα μοντέλο με λίγα επισημειωμένα δεδομένα ως δεδομένα εκπαίδευσης και αντιμετωπίζει τα υπόλοιπα ως δεδομένα δοκιμής. Σύμφωνα με το Διάγραμμα 2, η ημι-επιβλεπόμενη μάθηση διαχωρίζεται σε δύο μέρη: την ημι-επιβλεπόμενη ταξινόμηση (semi-supervised classification) και την ημι-επιβλεπόμενη ομαδοποίηση (semi-supervised clustering). Η ημι-επιβλεπόμενη ταξινόμηση είναι παρόμοια με την επιβλεπόμενη ταξινόμηση με την μόνη διαφορά ότι η δεύτερη απαιτεί περισσότερα δεδομένα εκπαίδευσης για την ταξινόμηση δεδομένων δοκιμής. Στην περίπτωση της ημι-επιβλεπόμενης ταξινόμησης χρησιμοποιώντας λιγότερα δεδομένα εκπαίδευσης είναι εφικτό να ταξινομηθεί μεγάλος όγκος δεδομένων δοκιμής. Από την άλλη πλευρά η ημι-επιβλεπόμενη ομαδοποίηση αποτελεί μια ειδική περίπτωση ομαδοποίησης. Ενώ στην περίπτωση της ομαδοποίησης χρησιμοποιούνται μη επισημειωμένα δεδομένα στην είσοδο, στην ημι-επιβλεπόμενη ομαδοποίηση

χρησιμοποιούνται τόσο επισημειωμένα όσο και μη επισημειωμένα δεδομένα (Padmanabha Reddy et al., 2018).



Διάγραμμα 2: Οι προσεγγίσεις της μηχανικής μάθησης. Πηγή: C A Padmanabha Reddy et al, 2028: 82

Παλαιότερα, όταν η επεξεργασία γινόταν με μικρή ποσότητα δεδομένων, τα συστήματα μηχανικής μάθησης ήταν σχεδιασμένα με την παραδοχή ότι όλα τα δεδομένα θα φορτώνονταν στην μνήμη για επεξεργασία. Με την ραγδαία αύξηση των διαθέσιμων δεδομένων αυτή η προσέγγιση έπρεπε να αναπτυχθεί καταλλήλως ούτως ώστε να καλυφθούν οι υψηλές απαιτήσεις που σχετίζονται με την επεξεργασία μεγάλου όγκου δεδομένων (Qiu et al., 2016). Κατ' επέκταση αναπτύχθηκαν πιο εξειδικευμένες τεχνικές μηχανικής μάθησης. Οι τεχνικές αυτές αφορούν τη μάθηση αναπαράστασης (Representation Learning), τη βαθιά μάθηση (Deep Learning), τη κατανεμημένη και παράλληλη μάθηση (Distributed and parallel learning), τη μεταφορά μάθησης (Transfer learning), την ενεργητική μάθηση (Active learning) και τη μάθηση με βάση τον πυρήνα (Kernel-based learning) (Qiu et al., 2016).

Σημαντικό βήμα για την εκπαίδευση ενός αλγορίθμου μηχανικής μάθησης είναι επίσης η μείωση διαστάσεων (Dimensionality Reduction). Η μείωση διαστάσεων είναι μια μέθοδος μείωσης του αριθμού διάστασης των χαρακτηριστικών των δεδομένων

διατηρώντας παράλληλα τις ιδιότητές τους. Πιο αναλυτικά, οι τεχνικές μείωσης διαστάσεων προβάλλουν δεδομένα από τον αρχικό πολυδιάστατο χώρο Rn σε ένα νέο χώρο χαμηλότερης διάστασης Rk . Μεταξύ άλλων η μείωση των διαστάσεων: (α) οδηγεί σε ταχύτερο υπολογισμό των αποστάσεων μεταξύ των δεδομένων, (β) μειώνει το μέγεθος του συνόλου δεδομένων και αναδεικνύει την δομή τους, (γ) βελτιώνεται η αποδοτικότητα των τεχνικών εξόρυξης δεδομένων, κτλ. (Murty & Devi, 2015; Περιγραφή: Μελέτη Μεθόδων Μείωσης Διαστατικότητας, 2020).

Είναι σημαντική η συμβολή της μηχανικής μάθησης ως προς την επίλυση προβλημάτων που απασχολούν την κοινωνία και σχετίζονται με την όραση (vision), αναγνώριση ομιλίας (speech recognition), ρομποτική (robotics), εξόρυξη δεδομένων και πληροφοριών (informatics and data mining), αυτόνομα συστήματα ελέγχου (autonomous control systems), συστήματα συστάσεων (recommendation systems) και συστήματα αναγνώρισης (recognition systems) (Aplaydin, 2020; Qiu et al., 2016). Οι προκλήσεις που αντιμετωπίζει η μηχανική μάθηση σχετίζονται με τα μεγάλα δεδομένα όπως αναφέρεται στο επιστημονικό άρθρο των L'Heureux et al. (2017) καθώς και του Zhou (2017).

1.2.3 Βαθιά Μάθηση

Η βαθιά μάθηση είναι το πεδίο της μηχανικής μάθησης που επικεντρώνεται στην ανάπτυξη σημαντικών μοντέλων νευρωνικών δικτύων ικανών να παράγουν ακριβείς αποφάσεις βάσει δεδομένων. Η βαθιά μάθηση είναι η καταλληλότερη για σενάρια με περίπλοκα δεδομένα και μεγάλα σύνολα δεδομένων. Η βαθιά μάθηση χρησιμοποιείται πλέον από τη συντριπτική πλειονότητα των επιχειρήσεων του διαδικτύου και των καταναλωτικών τεχνολογιών υψηλής τεχνολογίας (Janiesch et al., 2021: 687; Kelleher, 2019: 14). Στοιχείο που κάνει τη βαθιά μάθηση ξεχωριστή είναι τα πολλαπλά κρυφά επίπεδα επεξεργασίας των δεδομένων τα οποία προέρχονται από την αρχιτεκτονική των συνελκτικών νευρωνικών δικτύων (Convolutional Neural Networks - CNN) (O'Shea & Nash, 2015: 2).

Η βαθιά μάθηση επιτρέπει την καινοτομία και τον μετασχηματισμό σε πολλές πτυχές της σύγχρονης ζωής και κρατά τα ηνία στις πλείστες ανακαλύψεις της τεχνητής νοημοσύνης (Kelleher, 2019: 10). Τα τελευταία χρόνια, η βαθιά μάθηση έχει οδηγήσει

σε πολύ καλές αποδόσεις σε διάφορα αντικείμενα, όπως η οπτική αναγνώριση, η αναγνώριση ομιλίας και η επεξεργασία φυσικής γλώσσας (Gu et al., 2018: 354) και κατ' επέκταση, πολλές έρευνες να έχουν χρησιμοποιήσει τη βαθιά μάθηση για την αντιμετώπιση διαφόρων δυσκολιών που εμφανίζονται σε άλλους τομείς της επιστήμης όπως είναι η ηχημεία (Mater & Coote, 2019), η οικολογία (Christin et al., 2019), κτλ. Εμπνευσμένο από τις πιο πρόσφατες εξελίξεις στα νευρωνικά αυτοπαλινδρομούμενα παραγωγικά μοντέλα που μοντελοποιούν σύνθετες κατανομές, το WaveNet χρησιμοποιεί το παραγωγικό μοντέλο του CNN για την αναπαράσταση των δεδομένων (Gu et al., 2018: 370).

Μια εργασία στην οποία η βαθιά μάθηση μπορεί να ξεπεράσει σημαντικά τη συμβατική μηχανική μάθηση είναι ο σχεδιασμός χαρακτηριστικών (feature design). Ο σχεδιασμός χαρακτηριστικών στη συμβατική μηχανική μάθηση συχνά απαιτεί σημαντική ανθρώπινη προσπάθεια. Προσπαθώντας να εντοπίσει αυτόματα τα χαρακτηριστικά που είναι πιο ωφέλιμα για την εργασία από τα ακατέργαστα δεδομένα, η βαθιά μάθηση υιοθετεί μια διαφορετική προσέγγιση στη δημιουργία χαρακτηριστικών (Kelleher, 2019 :45). Στην πραγματικότητα, τα μοντέλα βαθιάς μάθησης έχουν περισσότερη ακρίβεια από πολλά άλλα μοντέλα μηχανικής μάθησης που χρησιμοποιούν χειροποίητα χαρακτηριστικά, δεδομένου ότι η βαθιά μάθηση έχει αποδειχθεί τόσο καλή στην εκμάθηση χαρακτηριστικών δεδομένων σε μεγάλα σύνολα δεδομένων. Γι' αυτό και ο πρωταρχικός παράγοντας που καθορίζει την βαθιά μάθηση ως επιτυχής είναι η γνώση του τι πρέπει να μετρηθεί και πώς να μετρηθεί (Kelleher, 2019 :43).

1.2.4 Συνελκτικά Νευρωνικά Δίκτυα

Τα συνελκτικά νευρωνικά δίκτυα (Convolutional Neural Networks - CNN) είναι μία πολύ διαδεδομένη μορφή αρχιτεκτονικής βαθιάς μάθησης, η οποία είναι εμπνευσμένη από την φυσική διαδικασία αντίληψης του οπτικού φάσματος από έναν ζωντανό οργανισμό (Gu et al., 2018: 1). Τα CNN αποδείχθηκαν επιτυχημένα για την επεξεργασία φυσικής γλώσσας και ομιλίας, με εξαιρετικά αποτελέσματα στην ταξινόμηση προτάσεων, στην μοντελοποίηση προτάσεων και στην αναγνώριση ομιλίας. Εφαρμόστηκαν με μεγάλη επιτυχία στην ενισχυτική μάθηση του μοντέλου AlphaGo που αναπτύχθηκε από την Google το 2015 και κατόρθωσε να νικήσει τους

μεγαλύτερους παίκτες στο European Champion στο αρχαίο κινέζικο παιχνίδι Go (Lopez Pinaya et al., 2020: 2 ; Zhang, 2018: 434).

Τα συνελκτικά νευρωνικά δίκτυα είναι βασισμένα στα τεχνητά νευρωνικά δίκτυα τα οποία είναι εμπνευσμένα από τη λειτουργία οργανικών νευρικών συστημάτων (όπως ο ανθρώπινος εγκέφαλος). Τα νευρωνικά δίκτυα αποτελούνται κυρίως από μεγάλο αριθμό διασυνδεδεμένων υπολογιστικών κόμβων (που αναφέρονται ως νευρώνες), οι οποίοι συνεργάζονται με καταναμημένη μέθοδο για να μάθουν από την είσοδο και να μεγιστοποιήσουν την τελική τους έξοδο. Η δομή ενός τεχνητού νευρωνικού δικτύου είναι ως εξής: αποτελείται από ένα επίπεδο εισόδου, ένα επίπεδο εξόδου και ενδιάμεσα υπάρχουν πολλά κρυφά επίπεδα. Στο επίπεδο εισόδου δίνονται τα δεδομένα τα οποία διανέμονται στα κρυφά επίπεδα μέχρι να φτάσουν στο επίπεδο εξόδου. Το επίπεδο εξόδου μπορεί να αποτελεί την έξοδο για ένα συγκεκριμένο τμήμα του δικτύου, όμως για κάποιο άλλο τμήμα μπορεί να αποτελεί την είσοδο. Με αυτό τον τρόπο η βαθιά μάθηση ορίζεται συνήθως και ως η ύπαρξη πολλών κρυφών ενδιάμεσων επιπέδων που συσσωρεύονται το ένα πάνω στο άλλο (O'Shea & Nash, 2015).

Η ιδέα ανάπτυξης των CNN προήλθε από ένα πείραμα σχετικά με τον οπτικό φλοιό που έγινε το 1962 (Zhang, 2018: 434). Στο πείραμα αποδείχτηκε ότι μεμονωμένοι νευρώνες στον εγκέφαλο ανταποκρίνονται μόνο όταν βρίσκονται στις παρυφές μιας συγκεκριμένης κατεύθυνσης. Ορισμένοι νευρώνες, για παράδειγμα, ενεργοποιούνται μόνο από κάθετες ακμές, ενώ άλλοι ενεργοποιούνται από οριζόντιες ή διαγώνιες ακμές. Το πείραμα ανακάλυψε ότι όλοι αυτοί οι νευρώνες ομαδοποιούνται και συνεργάζονται για να παράγουν την οπτική αντίληψη. Η παραδοχή ότι ένα συστατικό ενός τέτοιου συστήματος έχει ένα συγκεκριμένο σκοπό (π.χ. οι νευρώνες στον οπτικό φλοιό αναζητούν συγκεκριμένα χαρακτηριστικά) είναι εξίσου σημαντική στις μηχανές και αποτελεί την βάση των CNN. Ο Fukushima πρότεινε το Neocognitron (Ballester, 2016 όπως γίνεται αναφ. στο Zhang, 2018: 435) για πρώτη φορά το 1980, με βάση την έννοια των δεκτικών πεδίων (receptive fields) οι οποίες αφορούν συγκεκριμένες περιοχές του αισθητηριακού χώρου όπου ένα κατάλληλο ερέθισμα μπορεί να οδηγήσει μια ηλεκτρική απόκριση σε έναν αισθητηριακό νευρώνα (Gu et al., 2018: 1). Το Neocognitron είναι ένα αυτοοργανωτικό μοντέλο νευρωνικού δικτύου πολλαπλών επιπέδων. Το τοπικό αισθητήριο πεδίο του ανώτερου επιπέδου καθορίζει την απόκριση κάθε επιπέδου. Η αναγνώριση του μοντέλου δεν επηρεάζεται από τη θέση, τις

μικροσκοπικές αλλαγές στο σχήμα ή το μέγεθος της κλίμακας. Η μάθηση χωρίς επίβλεψη (unsupervised learning) με τη χρήση του Neocognitron ήταν επίσης η κυρίαρχη προσέγγιση μάθησης στην πρώιμη έρευνα των CNN.

Από το 2006, με την παράλληλη εξέλιξη των CNNs, έχουν αναπτυχθεί και εφαρμοστεί πολλές στρατηγικές ως προς την επίλυση των προβλημάτων που αντιμετωπίζονται με την εκπαίδευση πιο εξελιγμένων δικτύων. Πιο συγκεκριμένα, οι Hinton et al. (2012 όπως γίνεται αναφ. στο Gu et al., 2018: 354) ανέπτυξαν έναν παραδοσιακό σχεδιασμό CNN και κατέδειξαν την υπεροχή του στην ταξινόμηση εικόνων σε σχέση με προηγούμενες προσεγγίσεις. Η συνολική αρχιτεκτονική της μεθόδου τους, που ονομάστηκε AlexNet, είναι παρόμοια με εκείνη του LeNet-5, αλλά με πιο λεπτομερή δομή. Μετά την επιτυχία του AlexNet προτάθηκαν διάφορα σχέδια για την αύξηση της απόδοσής του μεταξύ αυτών και οι αξιοσημείωτες αρχιτεκτονικές: ZFNet, VGGNet, GoogleNet και ResNet (Gu et al., 2018:1). Σύμφωνα με τους Zeiler and Fergus (2014: 1), παρά τις θετικές εξελίξεις που έγιναν μέχρι το 2012 με την εξέλιξη των CNN, εξακολουθούσε να είναι δύσκολη η κατανόηση του τρόπου με τον οποίο συμπεριφέρονται αυτά τα περίπλοκα μοντέλα στο εσωτερικό τους ή του τρόπου με τον οποίο καταφέρνουν να το αποδώσουν με την πολυπλοκότητά τους. Λόγω αυτού, η ερευνητική κοινότητα πειραματίστηκε περαιτέρω έτσι ώστε να αναπτύξει καλύτερα μοντέλα με καλύτερη κατανόηση για το πώς λειτουργούν.

Μεταξύ των CNNs που χρησιμοποιούνται ευρέως για την ταξινόμηση εικόνων είναι το LeNet-5, το οποίο αναπτύχθηκε αρχικά για την αναγνώριση χαρακτήρων γραφής. Το LeNet-5 δημιουργήθηκε από τους Le Cun et al. (1989 όπως γίνεται αναφ. στο Gu et al., 2018: 354) για την αναγνώριση εικόνων 2D. Αυτό το δίκτυο εξαλείφει την απαίτηση για την εξαγωγή χαρακτηριστικών στο στάδιο της προεπεξεργασίας. Στην πραγματικότητα, το δίκτυο μπορεί να αναγνωρίσει πρότυπα απευθείας σε εικόνες εισόδου. Το LeNet-5 παράγει μια αυτό-μελετώμενη επεξεργασία μέσω της οπισθοδρομικής διάδοσης (back-propagation - BP) για συνεχή βελτιστοποίηση συνδυάζοντας την εξαγωγή χαρακτηριστικών και την αναγνώριση προτύπων. Μέσω της διαδικασίας αυτό-μελέτης, το δίκτυο ανακαλύπτει τα βέλτιστα χαρακτηριστικά για την ταξινόμηση (Li et al., 2019: 110). Όσα CNNs ακολούθησαν μετά το LeNet-5, βασίστηκαν στην αρχιτεκτονική του με κάποιες διαφορές (Ajit et al., 2020: 1; Kaye et al., 2020: 238).

Μία από τις εξελίξεις των συνελκτικών δικτύων η οποία επηρέασε θετικά την μηχανική όραση είναι η ανάπτυξη του γραφικού συνελκτικού δικτύου (Graphical Convolutional Network - GCN). Αποτελεί νέα δομή συνελκτικού δικτύου το οποίο επιτρέπει την συνέλιξη που χρησιμοποιείται σε αλγορίθμους βαθιάς μάθησης σε δεδομένα γραφημάτων (Cao et al., 2022: 13388). Το GCN κατάφερε να επιλύσει διάφορα προβλήματα της μηχανικής όρασης βελτιώνοντας τον εντοπισμό και την αναγνώριση αντικειμένων, την παρακολούθηση αντικειμένου, κτλ. (Cao et al., 2022: 13401).

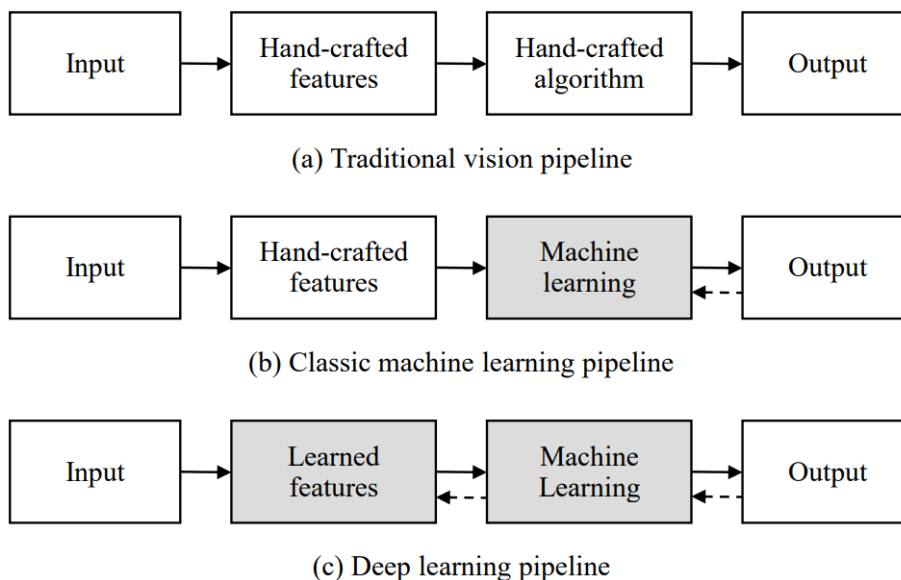
Ένα άλλο μοντέλο που έχει αναπτυχθεί για να αντιμετωπιστούν διάφορα προβλήματα και δυσκολίες που σχετίζονται με το χρόνο εξαγωγής συμπερασμάτων και την ακρίβεια, είναι το MobileNet. Το MobileNet, χρησιμοποιεί συνέλιξεις που απαιτούν μόνο το ένα όγδοο του χρόνου υπολογισμού (Sinha & El-Sharkawy, 2019: 280). Αυτό έχει ως αποτέλεσμα την δημιουργία μοντέλων που μπορούν να χρησιμοποιηθούν σε εφαρμογές κινητών συσκευών λόγω του μικρού τους μεγέθους και του λιγότερου χρόνου που χρειάζονται για να δώσουν το αποτέλεσμα στην έξοδο.

1.2.5 Μηχανική Όραση

Όσο εύκολη είναι η αναγνώριση των τριών διαστάσεων από ένα ανθρώπινο μάτι, τόσο δύσκολη γίνεται από μία μηχανή. Παράδειγμα σε αυτή τη δυσκολία είναι η αναγνώριση προσώπων σε μια εικόνα, όπως επίσης και η αναγνώριση των συναισθημάτων αυτών των ατόμων. Μια μηχανή από την άλλη πλευρά δεν μπορεί με τόση μεγάλη ευκολία να αναγνωρίσει διαστάσεις, πρόσωπα, συναισθήματα κτλ., όπως ο άνθρωπος και τα ζώα. Λύση σε αυτό το πρόβλημα έδωσε η μηχανική όραση, μέσω της οποίας αναπτύχθηκαν μαθηματικές τεχνικές, οι οποίες αποσκοπούσαν στην ανάκτηση του τρισδιάστατου σχήματος και εμφάνισης των αντικειμένων σε οπτικά δεδομένα (Szeliski, 2022: 25). Η μηχανική όραση έχει εξελιχθεί μέσα από τα πεδία της τεχνητής νοημοσύνης, της ψηφιακής επεξεργασίας εικόνας (digital image processing) και της αναγνώρισης προτύπων (pattern recognition). Η ανάπτυξη της μηχανικής όρασης έχει δώσει ώθηση να αναπτυχθούν και άλλοι τομείς της τεχνολογίας και να βοηθήσουν σε διάφορες εφαρμογές τόσο σε επιστημονικό όσο και σε βιομηχανικό επίπεδο (Szeliski, 2022: 189). Η μηχανική όραση χρησιμοποιείται σε πραγματικές εφαρμογές όπως είναι

αυτοοδηγούμενα αυτοκίνητα, η ιατρική απεικόνιση, η οπτική αναγνώριση χαρακτήρων, κτλ. (Szeliski, 2022: 25, 28).

Η εξέλιξη της μηχανικής όρασης, όπως απεικονίζεται και στο Διάγραμμα 3, είναι άμεσα συνδεδεμένη με τις εξελίξεις στον τομέα της μηχανικής μάθησης ξεκινώντας από τις παραδοσιακές μεθόδους και καταλήγοντας σήμερα στις μεθόδους που αξιοποιούν βαθιά μάθηση. Οι τεχνικές βαθιάς μάθησης έδωσαν λύσεις σε πολλά προβλήματα που εμφανίστηκαν κατά καιρούς και η χρήση των συνελκτικών νευρωνικών δικτύων είναι αρκετά σημαντική σε πολλές εφαρμογές της μηχανικής όρασης (Voulodimos et al., 2018). Σύμφωνα με την έρευνα των Zhou et al. (2019: 2), είναι πλέον εφικτή η ανάλυση του περιεχομένου σκηνών με πολλές εφαρμογές που έχουν σχέση με την δράση όπως η πλοήγηση στην πόλη.



Διάγραμμα 3: Η εξέλιξη των μεθόδων μηχανικής όρασης.

Σημαντικό ρόλο στην μηχανική όραση διαδραματίζει η προ-εκπαίδευση (pre-training) του μοντέλου. Παρόλο που το ImageNet (Deng et al., 2009) έχει χρησιμοποιηθεί αρκετά από την επιστημονική κοινότητα δεν κατόρθωσε να φέρει τις επιθυμητές επιδόσεις, ανάλογες των επιδόσεων, που κατόρθωσε η καινούρια κατηγορία αλγορίθμων βαθιάς μάθησης γνωστή και ως μετασχηματιστές όρασης (Vision Transformers - ViT) (Liang et al., 2022: 1). Στην περίπτωση της ταξινόμησης εικόνας,

το ViT είναι ένα μοντέλο που ταξινομεί τις εικόνες σύμφωνα με ένα μηχανισμό αυτοργάνωσης (self attention) (Xu et al., 2023). Το μοντέλο διαιρεί την εικόνα σε μια σειρά διατεταγμένων επιφανειών που στη συνέχεια επεξεργάζονται μέσω ενός κωδικοποιητή-μετασχηματιστή (Yue et al., 2021: 387).

1.2.6 Επεξηγησιμότητα των Μοντέλων Βαθιάς Μάθησης

Οι αλγόριθμοι βαθιάς μάθησης είναι πολλές φορές δυσνόητοι ως προς την επεξήγηση των βημάτων που εκτελούνται μεταξύ εισόδου και εξόδου και για αυτό τα παραγόμενα μοντέλα χαρακτηρίζονται από πολλούς ως «μαύρα κουτιά» (black boxes). Οι αλγόριθμοι συνήθως επιλέγονται βάσει της απόδοσης τους και όχι απαραίτητα με βάση την ερμηνεία του μοντέλου που αναπτύσσουν. Πολλές μέθοδοι, συμπεριλαμβανομένης της διαδραστικής εκπαίδευσης και της διαδραστικής επιλογής χαρακτηριστικών, έχουν αναπτυχθεί πρόσφατα για να εμπλέξουν τους ανθρώπους κατά την διαδικασία δημιουργίας των μοντέλων. Ωστόσο, εξακολουθούν να χρειάζονται διαδραστικά συστήματα που υπερβαίνουν τα απλά μέτρα ακρίβειας για την αξιολόγηση και την εκτίμηση της ερμηνευσιμότητας ενός μοντέλου (Krause et al., 2016: 5686).

Στην επιστημονική κοινότητα της μηχανικής μάθησης χρησιμοποιούνται οι έννοιες «Black boxes» και «white boxes» για να χαρακτηρίσουν τα μοντέλα που αναπτύσσονται. Ως «Black box» ονομάζονται τα μοντέλα που: (α) περιέχουν πολύπλοκες μαθηματικές συναρτήσεις, όπως για παράδειγμα τα μοντέλα που δημιουργούνται από μηχανές διανυσμάτων υποστήριξης και νευρωνικά δίκτυα, και (β) χρειάζονται επεξήγηση για την συνάρτηση που χρησιμοποιούν σε σχέση με τον χώρο που αναπαρίσταται (όπως για παράδειγμα στην περίπτωση του ταξινομητή του πλησιέστερου γείτονα k-Nearest Neighbor). Αυτού του είδους μοντέλα δύσκολα μπορούν εξηγηθούν και να κατανοηθούν από τους ειδικούς σε πρακτική εφαρμογή. Από την άλλη πλευρά, τα μοντέλα που ονομάζονται ως «white boxes» μπορούν συχνά να γίνουν κατανοητά από τους ειδικούς όπως στην περίπτωση των μοντέλων που βασίζονται σε πρότυπα, κανόνες, δέντρα αποφάσεων (Loyola-Gonzalez, 2019: 154096) ή λογιστική παλινδρόμηση (logistic regression) (Krause et al., 2016: 5687).

Η πολυπλοκότητα των αλγορίθμων σχετίζεται με την βασική ιδέα που ακολουθεί ο αλγόριθμος για να εκτελέσει τις εντολές που του έχουν δοθεί. Δηλαδή, ό,τι εισάγεται

στον αλγόριθμο και ό,τι εξάγεται από αυτόν δεν είναι απαραίτητως επεξηγήσιμα (Poon & Sung, 2021: 581). Η επεξηγησιμότητα των μοντέλων μηχανικής μάθησης απασχόλησε ιδιαίτερα την ερευνητική κοινότητα η οποία προσπάθησε να φωτίσει τα «Black boxes» όπως για παράδειγμα οι έρευνες που παρουσιάστηκαν από τους Krause et al. (2016), Bourguignon & Sundberg (2007), Tiedje et al. (1999) και Poon & Sung, (2021).

1.2.7 Χάρτες Θερμότητας

Οι χάρτες θερμότητας είναι μια θεμελιώδης τεχνική οπτικοποίησης που χρησιμοποιείται συχνά για την αποκάλυψη κρυμμένων μοτίβων σε γονιδιωματικά δεδομένα (Gu et al., 2016: 2847). Πιο αναλυτικά, οι χάρτες θερμότητας είναι διδιάστατες γραφικές απεικονίσεις δεδομένων στις οποίες οι τιμές μιας μεταβλητής εμφανίζονται με διαφορετικά χρώματα (Bojko, 2009: 30). Δηλαδή, είναι απεικονίσεις της οπτικής συμπεριφοράς και της κατανομής της προσοχής του ανθρώπου (Bojko, 2009: 37). Οι χάρτες θερμότητας βασίζονται κυρίως στην κωδικοποίηση χρωμάτων και στην ουσιαστική αναδιάταξη γραμμών και στηλών οι οποίες καθορίζουν και την ποιότητα της απεικόνισης (Gehlenborg & Wong, 2012). Οι χάρτες θερμότητας είναι χρήσιμοι για δύο λόγους. Πρώτον, ο διαισθητικός χαρακτήρας της χρωματικής κλίμακας σε σχέση με τη θερμότητα μειώνει τον όγκο της μελέτης που απαιτείται για την κατανόησή της. Δεύτερον, οι χάρτες θερμότητας εμφανίζουν τα δεδομένα ακριβώς πάνω στο ερέθισμα. Επειδή τα δεδομένα βρίσκονται τόσο κοντά στις πτυχές στις οποίες αφορούν, η ανάγνωση ενός χάρτη θερμότητας απαιτεί μικρή πνευματική προσπάθεια (Bojko, 2009: 30).

Οι χάρτες θερμότητας μπορούν να είναι χρήσιμα σε έγγραφα, εκθέσεις και παρουσιάσεις, επειδή συνοψίζουν τεράστιες ποσότητες δεδομένων που θα ήταν δύσκολο να κατανοηθούν αν παρουσιάζονταν αριθμητικά. Οι χάρτες επιτρέπουν την εξέταση και ανάλυση πιθανών μοτίβων ή τάσεων και μπορούν να χρησιμοποιηθούν για την παραγωγή δεδομένων σχετικά με την εμπειρία του χρήστη, όπως για παράδειγμα η χρήση, η ακρίβεια ή η οπτική προσοχή (Bojko, 2009: 30). Έχουν συγκεκριμένους προορισμούς που καθορίζουν την ερμηνεία τους ανάλογα με το είδος τους.

Οι χάρτες του ανθρώπου χρησιμοποιούνται συνήθως για την κατανόηση της λειτουργίας του εγκεφάλου κατά τη διάρκεια διαφορετικών δραστηριοτήτων, όπως η γλωσσική επεξεργασία, η κίνηση, η αντίληψη και η συναίσθηση. Μπορούν να χρησιμοποιηθούν για τη βελτίωση της θεραπείας ή της διάγνωσης κάποιων νοσημάτων που σχετίζονται με τον εγκέφαλο (Islam & Zhang, 2019: 1). Επιπρόσθετα μπορούν να αξιοποιηθούν για την αναγνώριση της προσοχής και του ενδιαφέροντος των χρηστών σε συγκεκριμένα στοιχεία, όπως εικόνες, διαφημίσεις ή προϊόντα με μεγάλο αντίκτυπο στο σχεδιασμό ιστοσελίδων και διαφημίσεων (Lettner & Holzmann, 2012). Τέλος, οι ανθρώπινοι χάρτες μπορούν να αξιοποιηθούν για την ανάπτυξη εφαρμογών που σχετίζονται με την όραση μετρώντας αρκετές επιτυχίες στον τομέα της μηχανικής όρασης (Kamel et al., 2021: 1330).

Υπάρχουν τέσσερα διαφορετικά είδη χαρτών θερμότητας, συμπεριλαμβανομένων των: (i) Fixation Count Heatmap, (ii) Absolute Gaze Duration Heatmap, (iii) Relative Gaze Duration Heatmap, και (iv) Participant Percentage Heatmap (Bojko, 2009: 31-33). Για την παραγωγή των χαρτών θερμότητας χρησιμοποιούνται διάφορα εργαλεία όπως είναι η R, Tobii's ClearView, Tobii Studio, SMI's BeGaze, NYAN, EyeTools, κτλ. (Gu et al., 2016: 2847; Bojko, 2009: 30, 31).

Οι χάρτες θερμότητας μπορούν επίσης να χρησιμοποιηθούν και στην περίπτωση των αλγορίθμων μηχανικής μάθησης υποδεικνύοντας ποια τμήματα στα δεδομένα εισόδου ήταν τα πιο σημαντικά και βοήθησαν το παραγόμενο μοντέλο να προβλέψει την έξοδο (Zhou et al., 2018).

Οι χάρτες των αλγορίθμων μπορούν να βοηθήσουν στην επεξήγηση των μοντέλων μηχανικής μάθησης και στην ανάλυση μεγάλων συνόλων δεδομένων. Μέσω των χαρτών μπορεί να γίνει ο εντοπισμός των στοιχείων που έχουν μεγαλύτερη συσχέτιση με τα επιθυμητά αποτελέσματα ούτως ώστε να βελτιωθεί η ακρίβεια και η απόδοση των μοντέλων. Μπορούν να συμβάλλουν στη βελτίωση διαφόρων εφαρμογών της μηχανικής μάθησης όπως είναι η αναγνώριση προσώπων, η ανίχνευση και πρόληψη ατυχημάτων σε βιομηχανικές εφαρμογές (Wang & Mei, 2022).

1.2.8 Eye Tracker

Ο eye tracker είναι μια συσκευή που χρησιμοποιείται για να μετρήσει την κίνηση του ματιού και να καταγράψει τη θέαση του χρήστη. Μπορεί να χρησιμοποιηθεί σε πολλούς τομείς, όπως στην επιστήμη της ψυχολογίας, της κοινωνιολογίας, του μάρκετινγκ, του σχεδιασμού και της ανάπτυξης λογισμικού.

Όταν χρησιμοποιείται στη δημιουργία χαρτών θερμότητας, ο eye tracker καταγράφει τη θέαση του χρήστη καθώς αλληλεπιδρά με μια διεπαφή ή με μια ιστοσελίδα. Τα δεδομένα από τον eye tracker μπορούν να αναλυθούν για να δημιουργήσουν ένα χάρτη θερμότητας, που είναι μια οπτική αναπαράσταση των σημείων που έχει επικεντρωθεί ο χρήστης κατά τη διάρκεια της χρήσης της διεπαφής ή της ιστοσελίδας. Εκτός από τη δημιουργία των χαρτών θερμότητας, ο eye tracker μπορεί να χρησιμοποιηθεί και σε άλλους τομείς που αφορούν τις αντιδράσεις των ανθρώπων σε διαφορετικά περιβάλλοντα. Ανάλογα με τον σκοπό και τον τομέα εφαρμογής, τα δεδομένα από τον eye tracker μπορούν να χρησιμοποιηθούν για:

- Ανάλυση της κοινωνικής συμπεριφοράς: ο eye tracker μπορεί να χρησιμοποιηθεί για να μελετηθεί η κοινωνική συμπεριφορά των ανθρώπων σε διάφορες καταστάσεις, όπως σε μια συνομιλία ή σε έναν κοινωνικό χώρο.
- Ανάλυση της συμπεριφοράς του καταναλωτή: ο eye tracker μπορεί να χρησιμοποιηθεί για να μελετηθεί η συμπεριφορά των καταναλωτών κατά την αγορά προϊόντων ή υπηρεσιών.
- Αξιολόγηση της απόδοσης των ανθρώπων: ο eye tracker μπορεί να χρησιμοποιηθεί για να αξιολογηθεί η απόδοση των ανθρώπων σε διάφορες εργασιακές ή εκπαιδευτικές καταστάσεις.

1.3 Ερευνητικά Ερωτήματα

Τα ερευνητικά ερωτήματα που θέτονται στην παρούσα πτυχιακή εργασία είναι τα εξής:

1. Πόσο όμοιοι οπτικά είναι οι χάρτες θερμότητας που δημιουργούνται από τον άνθρωπο όταν εντοπίζει εμπόδια σε εικόνες που θέτουν την ζωή των πεζών σε κίνδυνο με τους χάρτες θερμότητας που δημιουργούνται από τους αλγόριθμους

συνελικτικών δικτύων που έχουν εκπαιδευτεί να εντοπίζουν τα εμπόδια στα ίδια δεδομένα;

2. Πόσο όμοιοι οπτικά είναι οι χάρτες θερμότητας που δημιουργούνται από τον άνθρωπο όταν εντοπίζει εμπόδια σε εικόνες που θέτουν την ζωή των πεζών σε κίνδυνο με τους χάρτες θερμότητας που δημιουργούνται από τους αλγόριθμους οπτικών μετατροπέων που έχουν εκπαιδευτεί να εντοπίζουν τα εμπόδια στα ίδια δεδομένα;
3. Πόσο όμοιοι οπτικά είναι οι χάρτες θερμότητας που δημιουργούνται από τις γυναίκες με τους χάρτες που δημιουργούνται από τους άντρες όταν εντοπίζουν εμπόδια σε εικόνες που θέτουν την ζωή των πεζών σε κίνδυνο;
4. Πόσο όμοιοι οπτικά είναι οι χάρτες θερμότητας που δημιουργούνται από τους αλγόριθμους οπτικών μετατροπέων που έχουν εκπαιδευτεί να εντοπίζουν εμπόδια σε εικόνες που θέτουν την ζωή των πεζών σε κίνδυνο με τους χάρτες θερμότητας που δημιουργούνται από τους αλγόριθμους συνελικτικών δικτύων που έχουν εκπαιδευτεί να εντοπίζουν τα ίδια εμπόδια στα ίδια δεδομένα;

1.3.1.1 Μεταβλητές

Εξαρτημένη μεταβλητή είναι η διαφορά μεταξύ χαρτών θερμότητας. Ο/η κάθε συμμετέχοντας/ουσα στο πείραμα καθοδηγείται να εντοπίσει συγκεκριμένα εμπόδια σε εικόνες που εμφανίζονται στην οθόνη. Με την χρήση του eye-tracker γίνεται εξαγωγή των χαρτών θερμότητας, οι οποίοι στη συνέχεια συγκρίνονται είτε μεταξύ τους και εξάγεται η οπτική διαφορά, είτε μεταξύ των χαρτών θερμότητας που δημιουργήθηκαν από τους αλγορίθμους βαθιάς μάθησης MobileNetV2 και ViT-B16 οι οποίοι χρησιμοποιήθηκαν για την δημιουργία μοντέλων εντοπισμού των ίδιων εμποδίων στις ίδιες εικόνες. Ανεξάρτητες μεταβλητές αποτελούν το φύλο των συμμετεχόντων/ουσών, τα εμπόδια, η διαδικασία εντοπισμού των εμποδίων και ο αλγόριθμος βαθιάς μάθησης.

2 Μεθοδολογία Έρευνας

Στα πλαίσια της παρούσας πτυχιακής εργασίας εκπονήθηκε πειραματική έρευνα η οποία αφορούσε: (1) την δημιουργία ανθρώπινων χαρτών θερμότητας χρησιμοποιώντας συσκευή eye-tracker, και (2) την οπτική τους σύγκριση με τους αντίστοιχους χάρτες θερμότητας οποίοι εξήχθησαν μετά την εκπαίδευση 2 αλγορίθμων βαθιάς μάθησης.

Για την διεξαγωγή των πειραμάτων χρησιμοποιήθηκε το Usability Room του Cyprus Interaction Lab του Τμήματος Πολυμέσων και Γραφικών Τεχνών του Τεχνολογικού Πανεπιστημίου Κύπρου το οποίο είναι εξοπλισμένο με το Tobii Eye-tracker Nano. Για το πείραμα χρησιμοποιήθηκαν εικόνες που απεικονίζουν εμπόδια που θέτουν την ζωή των πεζών σε κίνδυνο για τις οποίες ζητήθηκε από τους συμμετέχοντες του πειράματος να τα εντοπίσουν. Στην συνέχεια για κάθε εικόνα εξήχθησαν οι χάρτες θερμότητας των περιοχών που εστίασαν για περισσότερο χρόνο οι χρήστες για να εντοπίσουν το κάθε εμπόδιο.

2.1 Δεδομένα

Τα δεδομένα που χρησιμοποιήθηκαν για την συλλογή των χαρτών θερμότητας αφορούν εικόνες οι οποίες συλλέχθηκαν με κάμερα κινητού τηλεφώνου στους δρόμους της πόλης της Λευκωσίας. Συγκεκριμένα χρησιμοποιήθηκαν 20 εικόνες που απεικονίζουν εμπόδια που επηρεάζουν την προστασία των πεζών στο πεζοδρόμιο. Οι εικόνες ήταν κατηγοριοποιημένες στις ακόλουθες 10 κατηγορίες: δίκυκλο όχημα, αυτοκίνητο, παγκάκι, συνωστισμένο πεζοδρόμιο, τρύπα, παρκόμετρο, εμπόδιο στάθμευσης, σπασμένο πλακόστρωτο, κώνος τροχαίας και δέντρο. Επιπρόσθετα χρησιμοποιήθηκαν ακόμα 5 φωτογραφίες οι οποίες δεν αφορούσαν κάποιο εμπόδιο.

2.2 Συσκευή eye-tracker

Η συλλογή των χαρτών θερμότητας έγινε με την χρήση του eye tracker οθόνης Tobii Eye tracker Nano. Το λογισμικό Pro Lab που συνοδεύει τον eye tracker προσφέρει μια ολοκληρωμένη πλατφόρμα για την καταγραφή και την ανάλυση δεδομένων του βλέμματος, η οποία βοηθά στην ερμηνεία της ανθρώπινης συμπεριφοράς. Μέσω του

λογισμικού, τα δεδομένα παρακολούθησης των ματιών επεξεργάζονται εύκολα για σύγκριση, ερμηνεία και παρουσίαση συνδυάζοντας απλές διαδικασίες με σύνθετα εργαλεία οπτικοποίησης και ανάλυσης.

2.3 Σχεδιασμός Πειράματος

Το πείραμα που χρησιμοποιήθηκε για την δημιουργία των χαρτών θερμότητας του ανθρώπου αποτελείται από 2 μέρη (Part A και Part B). Από τις 25 εικόνες, οι 15 χρησιμοποιήθηκαν στο Part A και οι 10 εικόνες στο Part B. Το Part A (Πίνακας 1) αποτελείται από: (α) 1 εικόνα από κάθε κατηγορία εμποδίου (σύνολο 10 εικόνες), και (β) 5 εικόνες που δεν αφορούσαν κανένα εμπόδιο οι οποίες προστέθηκαν για να μην δημιουργηθεί η προκατάληψη από τους συμμετέχοντες/ουσες ότι όλες οι εικόνες που θα τους παρουσιαστούν περιέχουν κατ' ανάγκη κάποιο εμπόδιο.

Πίνακας 1: Εικόνες που παρουσιάστηκαν στο Part A του πειράματος

Part A				
				
Εικόνα χωρίς εμπόδιο 1	Εικόνα χωρίς εμπόδιο 2	Εικόνα χωρίς εμπόδιο 3	Εικόνα χωρίς εμπόδιο 4	Εικόνα χωρίς εμπόδιο 5
				
Δίκυκλο όχημα	Τετράτροχο όχημα	Παγκάκι	Συνωσθημένο Πεζοδρόμιο	Τρύπα
				
Παρκόμετρο	Εμπόδιο στάθμευσης	Σπασμένο πλακόστρωτο	Κώνος τροχαίας	Δέντρο

Το Part B του πειράματος αποτελείται από τις υπόλοιπες 10 εικόνες, 1 εικόνα από κάθε κατηγορία εμποδίου (Πίνακας 2).

Πίνακας 2: Εικόνες που παρουσιάστηκαν στο Part B του πειράματος

Part B				
				
Δίκυκλο όχημα	Τετράτροχο όχημα	Παγκάκι	Συνωσθημένο Πεζοδρόμιο	Τρύπα
				
Παρκόμετρο	Εμπόδιο στάθμευσης	Σπασμένο πλακόστρωτο	Κώνος τροχαίας	Δέντρο

Κατά την διάρκεια του πειράματος, τόσο στο Part A όσο και στο Part B, μετά την προβολή κάθε εικόνας ακολουθούσε η παρουσίαση των κατηγοριών των εμποδίων σε αλφαβητική σειρά (Εικόνα 1) από τις οποίες οι συμμετέχοντες/ουσες θα έπρεπε να επιλέξουν σε ποια κατηγορία ανήκει το εμπόδιο που εντόπισαν.

> Δέντρο	> Σπασμένο πλακόστρωτο
> Δίκυκλο όχημα	> Στάση λεωφορείου
> Εμπόδιο στάθμευσης	> Στενό πεζοδρόμιο
> Επικίνδυνη διασταύρωση	> Συνωστισμένο πεζοδρόμιο
> Κάλαθος	> Τετρακίνητο όχημα
> Κώνος τροχαίας	> Τρύπα
> Παγκάκι	
> Παρκόμετρο	
> Πινακίδα	

Εικόνα 1: Κατηγορίες εμποδίων με αλφαβητική σειρά

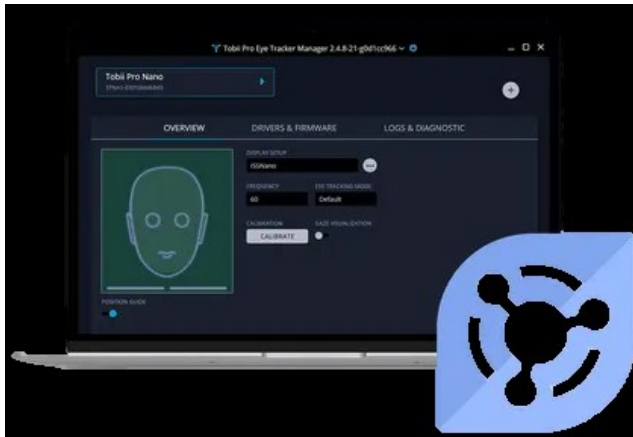
Οι χρήστες μπορούσαν να επιλέξουν την κατάλληλη κατηγορία εμποδίου αναμεσα σε 15 διαφορετικές επιλογές. Εκτός από τις 10 αρχικές κατηγορίες εμποδίων, προστέθηκαν ακόμα 5 συμπεριλαμβανομένου της στάσης λεωφορείου, κάλαθος, επικίνδυνη

διασταύρωση, πινακίδα και στενό πεζοδρόμιο. Όλοι οι συνδυασμοί εικόνας και λίστας εμποδίων εμφανίζονταν με τυχαία σειρά και στα δύο μέρη του πειράματος.

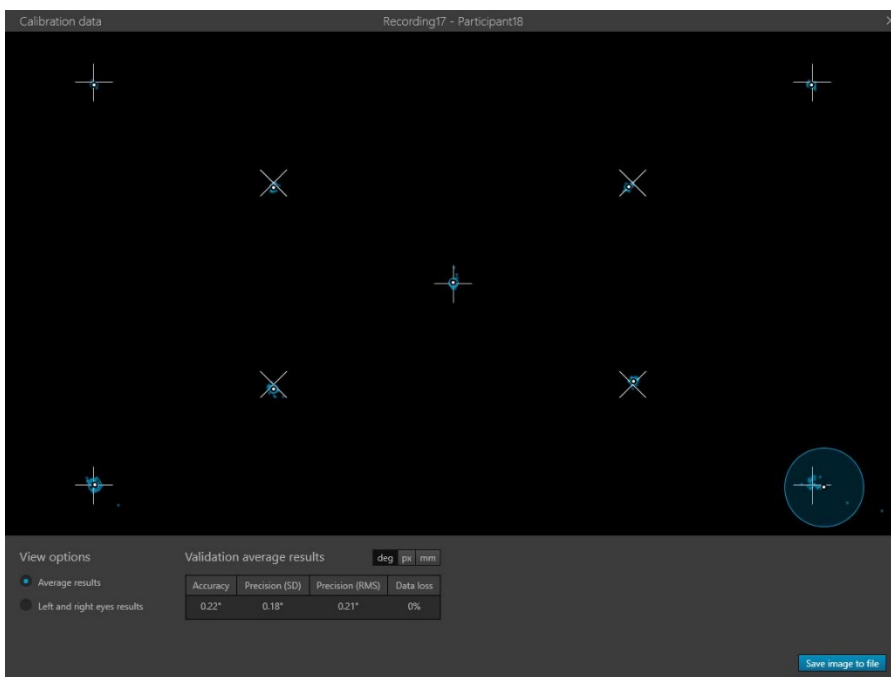
Πριν ξεκινήσει το πείραμα για την δημιουργία των χαρτών, έγινε δοκιμαστικά με την συμμετοχή 3 ατόμων για να επικυρωθεί η διαδικασία και να λυθούν τυχόν προβλήματα.

2.4 Δημιουργία Χαρτών Θερμότητας

Η συλλογή των συμμετεχόντων/ουσών για να συμμετάσχουν στο πείραμα έγινε μέσω του διαδικτύου όπου είχαν την δυνατότητα να δηλώσουν συμμετοχή και να επιλέξουν την ημέρα και την ώρα της αρεσκείας τους. Με την συμπλήρωση των 35 ατόμων που ήταν ο προαπαιτούμενος αριθμός συμμετοχής ξεκίνησε η διεξαγωγή των πειραμάτων. Τα πειράματα διεξήχθησαν τον Μάρτιο του 2023 μεταξύ των ωρών 9:00 π.μ. και 6:00 μ.μ λόγω των περιορισμών του φωτισμού στην αίθουσα του πειράματος. Κάθε πείραμα είχε διάρκεια περίπου 15 λεπτά μαζί με τις οδηγίες. Το πείραμα ξεκινούσε με την ενημέρωση των συμμετεχόντων/ουσών για την διαδικασία (Παράρτημα Ι). Το πρώτο βήμα μετά την ενημέρωση ήταν η προσθήκη του/της συμμετέχοντα/ουσας και των χαρακτηριστικών του/της μεταβλητών στο λογισμικό του eye-tracker. Συγκεκριμένα δόθηκε η ηλικία και το φύλο όπου το δεύτερο χρησιμοποιήθηκε για την ανάλυση των αποτελεσμάτων. Στην συνέχεια ακολούθησε η πρώτη βαθμονόμηση (calibration) του eye-tracker μέσω του “Eye Tracker Manager” (Εικόνα 2). Για την πρώτη βαθμονόμηση, ο/η συμμετέχοντας/ουσα έπρεπε να έχει την κατάλληλη απόσταση και ύψος από τον eye-tracker έτσι ώστε η συσκευή να μπορεί να εντοπίσει το μάτι του/της. Ο συμμετέχοντας/ουσα στην συνέχεια όφειλε να ακολουθήσει τις 5 κουκίδες που εμφανιζόντουσαν στην οθόνη. Μετά ακολούθησε και δεύτερη βαθμονόμηση μέσω του “Tobii Pro Lab” (Εικόνα 3), για ανατροφοδότηση και επικύρωση της βαθμονόμησης, όπου αυτή την φορά ο/η συμμετέχοντας/ουσα όφειλε να ακολουθήσει τις 9 κουκίδες που εμφανιζόντουσαν στην οθόνη.



Εικόνα 2: Αρχική ρύθμιση του eye tracker στο “Eye Tracker Manager”. *Πηγή:* https://connect.tobii.com/s/article/how-to-calibrate-and-validate-in-tobii-pro-lab?language=en_US



Εικόνα 3: Δεύτερη ρύθμιση και επικύρωση των τιμών του eye tracker στο “Tobii Pro Lab”. *Πηγή:* https://connect.tobii.com/s/article/how-to-calibrate-and-validate-in-tobii-pro-lab?language=en_US

Μετά την ρύθμιση της βαθμονόμησης ξεκινούσε το πείραμα όπου στο Part A, παρουσιάζονταν μια προς μια οι εικόνες στον/στην συμμετέχοντα/ουσα και καλείτο να απαντήσει εάν παρατηρούσε κάποιο εμπόδιο που εμπόδιζε την διέλευση των πεζών σε

κάθε εικόνα ξεχωριστά. Εάν η απάντηση ήταν θετική ακολουθούσε η ερώτηση εάν εντόπιζε ένα ή περισσότερα εμπόδια και στην συνέχεια ακολουθούσε η οδηγία για να εστιάσει περισσότερο στο εμπόδιο που θα τον/την ενοχλούσε περισσότερο. Μετά την διευκρινιστική ερώτηση ότι όντως το εντόπισε ακολουθούσε η λίστα με τις κατηγορίες εμποδίων για να επιλέξει σε ποια ανήκει το εμπόδιο που εντόπισε. Στην περίπτωση που η απάντηση στην αρχική ερώτηση εάν παρατηρούσε κάποιο εμπόδιο ήταν αρνητική, ο/η συμμετέχοντας/ουσα συνέχιζε κατευθείαν με την επόμενη εικόνα του πειράματος. Στο Part B ο/η συμμετέχοντας/ουσα έβλεπε μια προς μια τις εικόνες των εμποδίων και καλείτο να απαντήσει εάν εντόπιζε συγκεκριμένα κάποιο εμπόδιο. Μετά την διευκρινιστική ερώτηση ότι όντως το εντόπισε συνέχιζε με την επόμενη εικόνα του πειράματος.

3 Ανάλυση

3.1 Χάρτες Θερμότητας Μηχανής

Οι χάρτες θερμότητας των αλγορίθμων δημιουργήθηκαν χρησιμοποιώντας τα δεδομένα που συλλέχθηκαν με κάμερα κινητού τηλεφώνου στους δρόμους της πόλης της Λευκωσίας και αφορούσαν 29 εμπόδια που θέτουν την ζωή των πεζών σε κίνδυνο. Η συλλογή αφορούσε 2 βίντεο για κάθε εμπόδιο τα οποία επεξεργάστηκαν μέσω της εξαγωγής μεμονωμένων καρτέ (frames) για να δημιουργηθεί το σύνολο δεδομένων εκπαίδευσης. Η εξαγωγή των καρτέ έγινε σε τακτά διαστήματα καθ' όλη τη διάρκεια των βίντεο. Συνολικά εξήχθησαν 500 καρτέ από τα βίντεο για κάθε τύπο εμποδίου, με αποτέλεσμα ένα σύνολο δεδομένων με $29 \times 500 = 14500$ εικόνες. Με βάση το σύνολο δεδομένων που δημιουργήθηκε, εκπαιδεύτηκαν 2 μοντέλα νευρωνικών δικτύων, το ViT-B16 και το MobileNetV2, χρησιμοποιώντας την τεχνική εκπαίδευσης transfer learning (μεταφορά μάθησης). Συγκεκριμένα, τα μοντέλα που χρησιμοποιήθηκαν ήταν προ-εκπαιδευμένα στο ImageNet, και η μεταφορά μάθησης που έγινε με το σύνολο δεδομένων τα κατέστησε ικανά να αναγνωρίζουν με μεγάλη ακρίβεια τους 29 τύπους εμποδίων. Τα τελικά μοντέλα είχαν μια επικύρωση ακρίβειας (validation accuracy) κοντά στο 100%.

Για τις ανάγκες της παρούσας πτυχιακής εργασίας έγινε εξαγωγή των χαρτών θερμότητας των αλγορίθμων για 10 εμπόδια χρησιμοποιώντας τη βιβλιοθήκη PytorchGradCAM στην Python. Για κάθε εμπόδιο έγινε η εξαγωγή χαρτών θερμότητας από 2 διαφορετικές εικόνες. Συνολικά εξήχθησαν 20 χάρτες (10 τύποι εμποδίων \times 2 χάρτες θερμότητας). Για την επιλογή των χαρτών, έγινε ο απαραίτητος έλεγχος ούτως ώστε να διασφαλιστεί ότι οι χάρτες θερμότητας που χρησιμοποιήθηκαν δεν αλλοιώθηκαν από το φαινόμενο της υπερεκπαίδευσης.

3.2 Χάρτες Θερμότητας Ανθρώπου

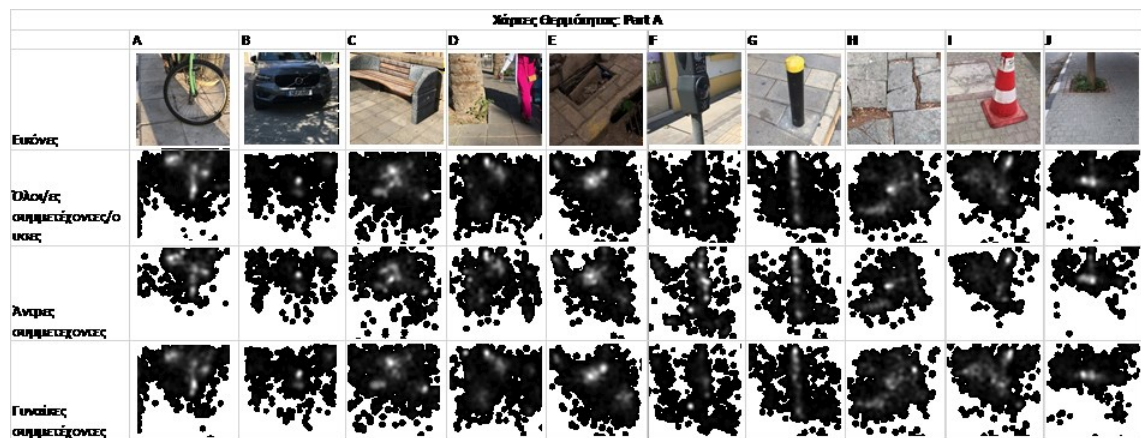
Μέσω του πειράματος που περιγράφηκε στο προηγούμενο κεφάλαιο δημιουργήθηκαν 60 ανθρώπινοι χάρτες θερμότητας οι οποίοι συμπεριλαμβάνουν:

1. τον μέσο χάρτη θερμότητας όλων των συμμετεχόντων/ουσών για κάθε εμπόδιο που δημιουργήθηκε στο Part A του πειράματος (συνολικά δημιουργήθηκαν 10 χάρτες θερμότητας, ένας για κάθε εικόνα)
2. τον μέσο χάρτη θερμότητας όλων των συμμετεχόντων/ουσών για κάθε εμπόδιο που δημιουργήθηκε στο Part B του πειράματος (συνολικά δημιουργήθηκαν 10 χάρτες θερμότητας, ένας για κάθε εικόνα)
3. τον μέσο χάρτη θερμότητας των γυναικών που συμμετείχαν στο πείραμα για κάθε εμπόδιο που δημιουργήθηκε στο Part A του πειράματος (συνολικά δημιουργήθηκαν 10 χάρτες θερμότητας, ένας για κάθε εικόνα)
4. τον μέσο χάρτη θερμότητας των αντρών που συμμετείχαν στο πείραμα για κάθε εμπόδιο που δημιουργήθηκε στο Part A του πειράματος (συνολικά δημιουργήθηκαν 10 χάρτες θερμότητας, ένας για κάθε εικόνα)
5. τον μέσο χάρτη θερμότητας των γυναικών που συμμετείχαν στο πείραμα για κάθε εμπόδιο που δημιουργήθηκε στο Part B του πειράματος (συνολικά δημιουργήθηκαν 10 χάρτες θερμότητας, ένας για κάθε εικόνα)
6. τον μέσο χάρτη θερμότητας των αντρών που συμμετείχαν στο πείραμα για κάθε εμπόδιο που δημιουργήθηκε στο Part B του πειράματος (συνολικά δημιουργήθηκαν 10 χάρτες θερμότητας, ένας για κάθε εικόνα)

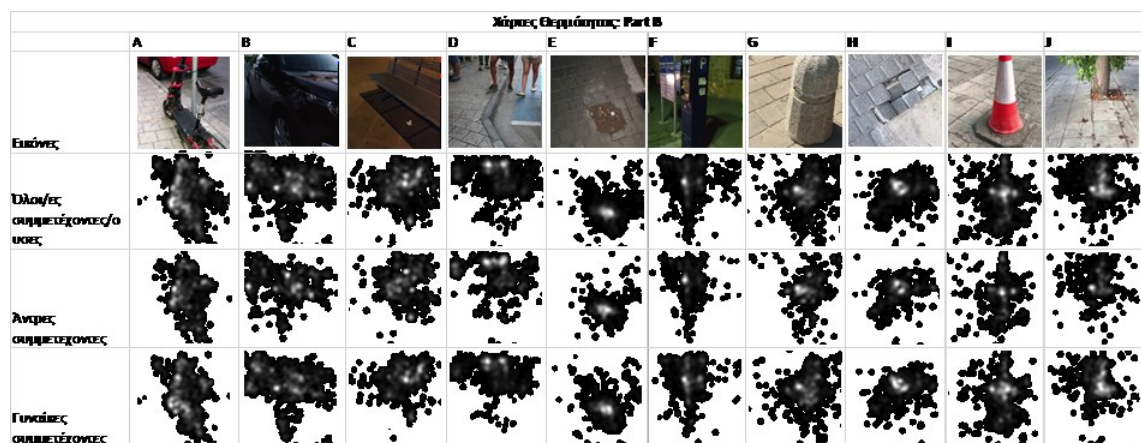
Όσον αφορά το Part A του πειράματος δεν δημιουργήθηκαν οι χάρτες θερμότητας για τις εικόνες που δεν περιείχαν κάποιο εμπόδιο. Επομένως από τις 15 εικόνες, δημιουργήθηκαν χάρτες θερμότητας μόνο για τις 10 εικόνες. Ο κάθε χάρτης θερμότητας εξήχθη σε μορφή εικόνας (.png) με διαστάσεις 1080x1080. Οι πληροφορίες στους χάρτες θερμότητας αναπαρίστανται σε αποχρώσεις του γκρι από 0 έως 255. Στους πίνακες 3 και 4 παρουσιάζονται οι χάρτες θερμότητας για το Part A και Part B, όπου στην συγκεκριμένη περίπτωση το χρώμα ξεκινά από λευκό για τα σημεία που δεν έχουν παρατηρηθεί καθόλου και καταλήγει σε μαύρο για τα σημεία που έχουν παρατηρηθεί τις περισσότερες φορές. Ο πίνακας 5 παρουσιάζει τους χάρτες θερμότητας που δημιουργήθηκαν στα Part A και Part B μαζί με τους χάρτες θερμότητας που δημιουργήθηκαν από τους αλγόριθμους MobileNetV2 και ViT-B16 σε έγχρωμη μορφή. Τα χρώματα που εμφανίζονται στους χάρτες θερμότητας του Πίνακα 5 είναι το κόκκινο, κίτρινο, πράσινο και μπλε για τους χάρτες των αλγορίθμων MobileNetV2 και ViT-B16. Το κάθε χρώμα υποδηλώνει και την προβολή που υπάρχει στο κάθε σημείο

της εικόνας. Τα σημεία στο χάρτη θερμότητας που είναι με κόκκινο χρώμα είναι όσα σημεία έχουν προβληθεί περισσότερες φορές από τους συμμετέχοντες/ουσες, δηλαδή είχε δοθεί μεγαλύτερη έμφαση σε αυτά. Τα σημεία με πράσινο ή και μπλε χρώμα είναι τα σημεία τα οποία έχουν προβληθεί ελάχιστα από τους συμμετέχοντες/ουσες ή τους αλγορίθμους. Ενδιάμεσα του κόκκινου χρώματος και του πράσινου είναι το κίτρινο χρώμα που έχει μία σχετικά μέτρια προβολή των σημείων αυτών από τους συμμετέχοντες/ουσες.

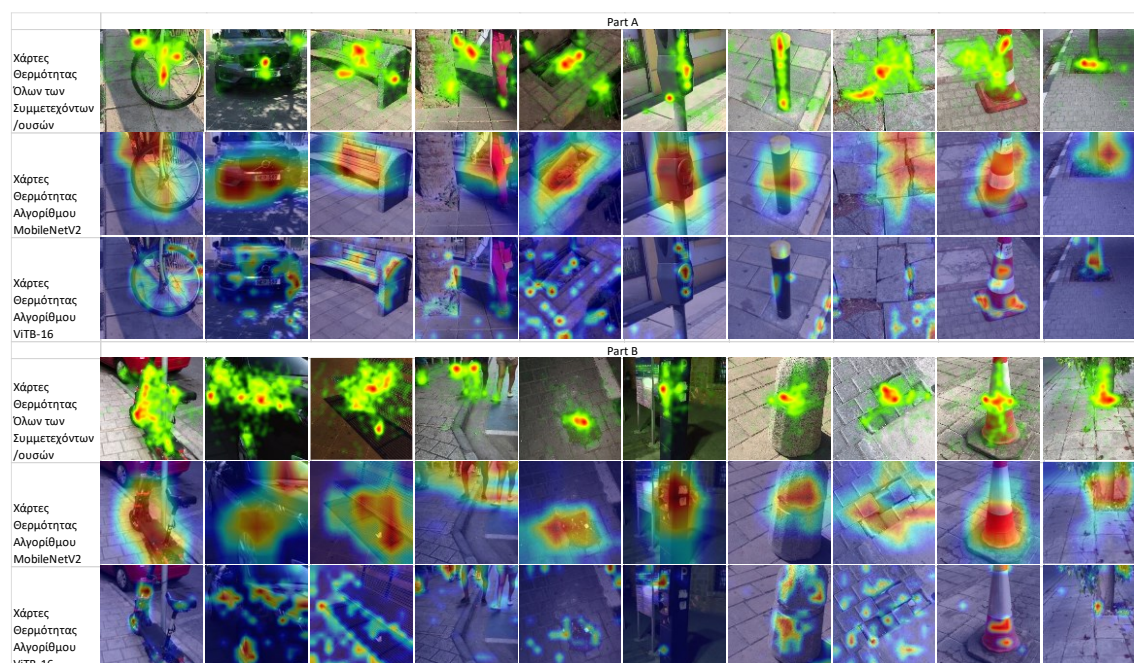
Πίνακας 3: Οι χάρτες θερμότητας που συλλέχθηκαν από το πρώτο μέρος των πειραμάτων



Πίνακας 4: Οι χάρτες θερμότητας που συλλέχθηκαν από το δεύτερο μέρος των πειραμάτων



Πίνακας 5: Οι χάρτες θερμότητας που συλλέχθηκαν από το πρώτο και δεύτερο μέρος των πειραμάτων πάνω στις εικόνες των εμποδίων σε έγχρωμη μορφή και οι χάρτες θερμότητας που δημιουργήθηκαν από τον κάθε αλγόριθμο



Για την σύγκριση της οπτικής διαφοράς μεταξύ κάθε χάρτη θερμότητας ανθρώπου και μηχανής ακολουθήθηκαν τα επόμενα βήματα:

1. Αλλαγή μεγέθους του ανθρώπινου χάρτη θερμότητας από 1080x1080 σε 224x224 τιμές για να έχει τις ίδιες διαστάσεις με τον χάρτη της μηχανής
2. Διαχωρισμός του χάρτη θερμότητας του ανθρώπου σε 196 τμήματα μεγέθους 16x16 το καθένα
3. Διαχωρισμός χάρτη θερμότητας της μηχανής σε 196 τμήματα μεγέθους 16x16 το καθένα
4. Υπολογισμός του μέσου όρου των τιμών της φωτεινότητας (οι τιμές της φωτεινότητας είναι μεταξύ 0 και 255) σε κάθε ένα από τα 196 τμήματα σε κάθε χάρτη θερμότητας
5. Υπολογισμός της διαφοράς της τιμής μεταξύ χάρτη θερμότητας ανθρώπου και μηχανής για κάθε ένα από τα 196 τμήματα
6. Άθροισμα των τιμών της διαφοράς για όλα τα τμήματα

Η υλοποίηση των βημάτων έγινε στην πλατφόρμα ταχείας πρωτοτυποποίησης MATLAB.

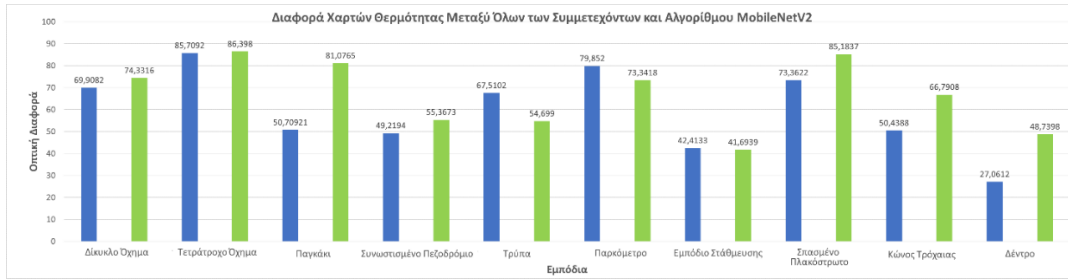
3.3 Ανάλυση Δεδομένων

Ερευνητικό Ερώτημα 1: Πόσο όμοιοι οπτικά είναι οι χάρτες θερμότητας που δημιουργούνται από τον άνθρωπο όταν εντοπίζει εμπόδια σε εικόνες που θέτουν την ζωή των πεζών σε κίνδυνο με τους χάρτες θερμότητας που δημιουργούνται από τους αλγόριθμους συνελκτικών δικτύων που έχουν εκπαιδευτεί να εντοπίζουν τα εμπόδια στα ίδια δεδομένα;

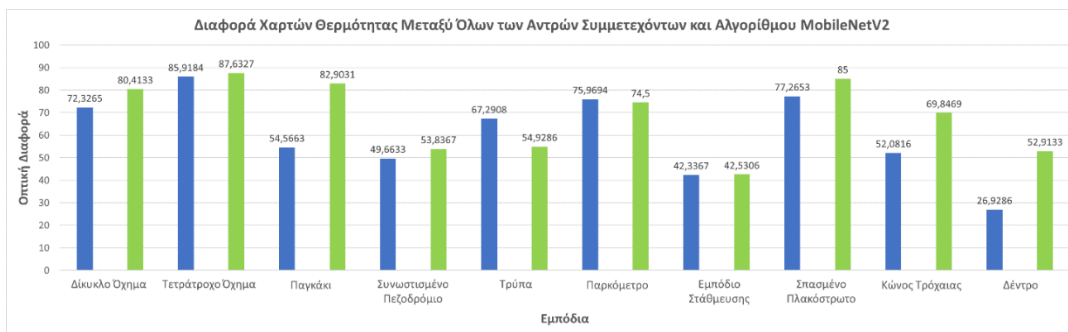
Για το πρώτο ερευνητικό ερώτημα δημιουργήθηκαν 3 γραφήματα τα οποία αναπαριστούν την:

1. Οπτική διαφορά των χαρτών θερμότητας όλων των συμμετεχόντων/ουσών με τους χάρτες θερμότητας του Αλγορίθμου MobileNetV2 (Διάγραμμα 4)
2. Οπτική διαφορά των χαρτών θερμότητας των αντρών με τους χάρτες θερμότητας του Αλγορίθμου MobileNetV2 (Διάγραμμα 5)
3. Οπτική διαφορά των χαρτών θερμότητας των γυναικών με τους χάρτες θερμότητας του Αλγορίθμου MobileNetV2 (Διάγραμμα 6)

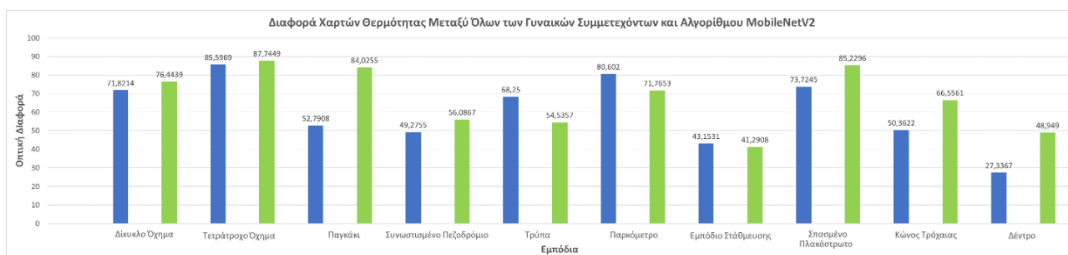
Όπως μπορεί να παρατηρηθεί από τα 3 αυτά γραφήματα οι διαφορές μεταξύ των χαρτών θερμότητας του αλγορίθμου MobileNetV2 και των χαρτών θερμότητας όλων των συμμετεχόντων/ουσών, των ανδρών και των γυναικών είναι μεγάλη. Οι πλείστες τιμές περικλείονται στο διάστημα 40-90 και αυτό υποδηλώνει πως οι χάρτες θερμότητας που δημιουργούνται από τον άνθρωπο όταν εντοπίζει εμπόδια σε εικόνες που θέτουν την ζωή των πεζών σε κίνδυνο δεν είναι όμοιοι με τους χάρτες θερμότητας που δημιουργούνται από τους αλγόριθμους συνελκτικών δικτύων που έχουν εκπαιδευτεί να εντοπίζουν τα εμπόδια στα ίδια δεδομένα.



Διάγραμμα 4: Σύγκριση των heatmaps όλων των συμμετεχόντων/ουσών με τα heatmaps του αλγορίθμου MobileNetV2. Μπάρα μπλε: Part A του πειράματος, μπάρα πράσινη: Part B του πειράματος



Διάγραμμα 5: Σύγκριση των heatmaps όλων των αντρών που συμμετείχαν με τα heatmaps του αλγορίθμου MobileNetV2. Μπάρα μπλε: Part A του πειράματος, μπάρα πράσινη: Part B του πειράματος



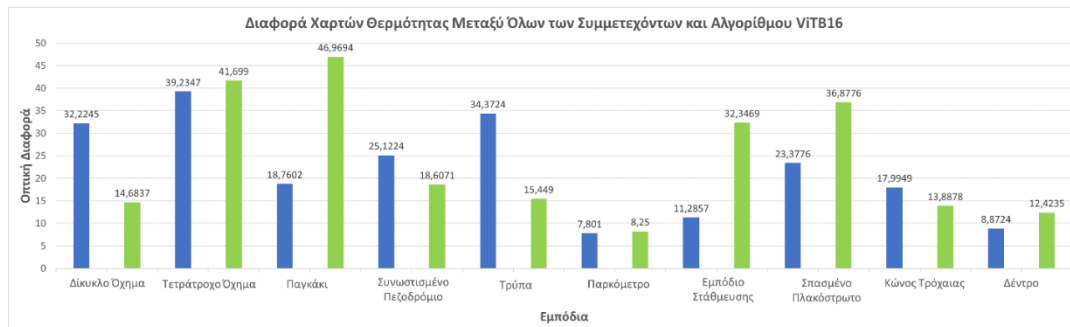
Διάγραμμα 6: Σύγκριση των heatmaps όλων των γυναικών που συμμετείχαν με τα heatmaps του αλγορίθμου MobileNetV2. Μπάρα μπλε: Part A του πειράματος, μπάρα πράσινη: Part B του πειράματος

Ερευνητικό Ερώτημα 2: Πόσο όμοιοι οπτικά είναι οι χάρτες θερμότητας που δημιουργούνται από τον άνθρωπο όταν εντοπίζει εμπόδια σε εικόνες που θέτουν την ζωή των πεζών σε κίνδυνο με τους χάρτες θερμότητας που δημιουργούνται από τους

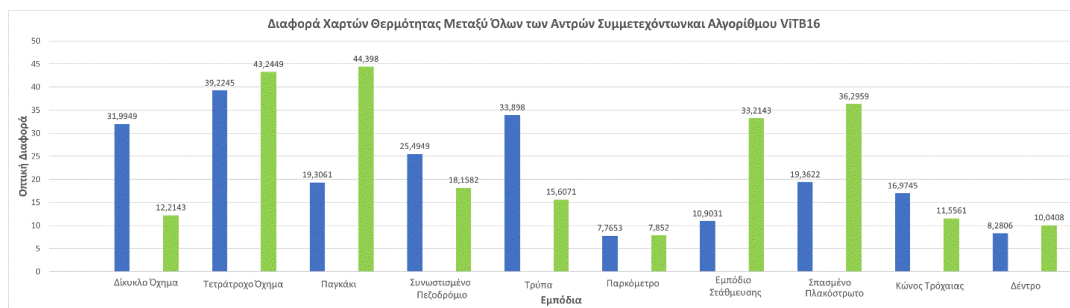
αλγόριθμους οπτικών μετατροπών που έχουν εκπαιδευτεί να εντοπίζουν τα εμπόδια στα ίδια δεδομένα;

Προκειμένου να μπορέσει να απαντηθεί το δεύτερο ερευνητικό ερώτημα δημιουργήθηκαν επίσης 3 γραφήματα τα οποία αναπαριστούν την:

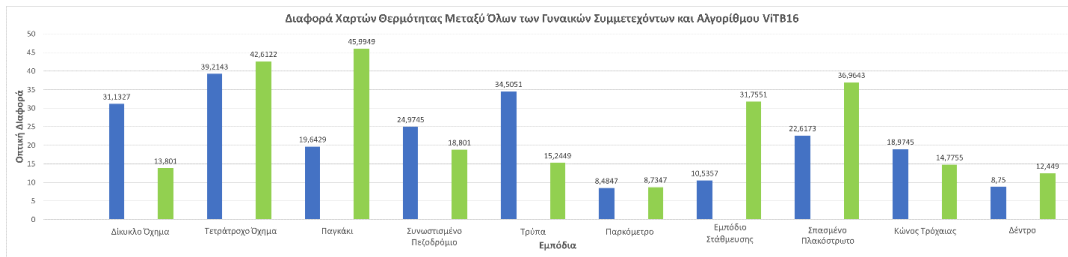
1. Οπτική διαφορά των χαρτών θερμότητας όλων των συμμετεχόντων με τους χάρτες θερμότητας του Αλγορίθμου ViT-B16 (Διάγραμμα 7)
2. Οπτική διαφορά των χαρτών θερμότητας των αντρών με τους χάρτες θερμότητας του Αλγορίθμου ViT-B16 (Διάγραμμα 8)
3. Οπτική διαφορά των χαρτών θερμότητας των γυναικών με τους χάρτες θερμότητας του Αλγορίθμου ViT-B16 (Διάγραμμα 9)



Διάγραμμα 7: Σύγκριση των heatmaps όλων των συμμετεχόντων/ουσών με τα heatmaps του αλγόριθμου ViT-B16. Μπάρα μπλε: Part A του πειράματος, μπάρα πράσινη: Part B του πειράματος



Διάγραμμα 8: Σύγκριση των heatmaps όλων των αντρών που συμμετείχαν με τα heatmaps του αλγορίθμου ViT-B16. Μπάρα μπλε: Part A του πειράματος, μπάρα πράσινη: Part B του πειράματος

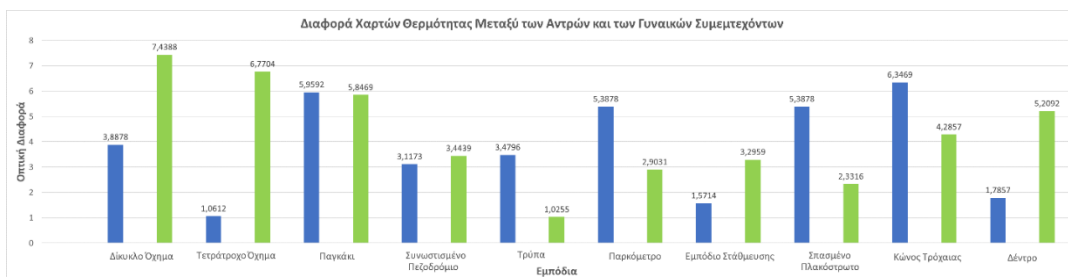


Διάγραμμα 9: Σύγκριση των heatmaps όλων των γυναικών που συμμετείχαν με τα heatmaps του αλγορίθμου ViT-B16. Μπάρτα μπλε: Part A του πειράματος, μπάρτα πράσινη: Part B του πειράματος

Στα γραφήματα 7, 8 και 9 η οπτική διαφορά, στους πλείστους χάρτες θερμότητας, κυμαίνεται μεταξύ των τιμών 7 και 25. Το συμπέρασμα που εξάγεται από αυτά τα γραφήματα είναι το εξής: οι χάρτες θερμότητας που δημιουργούνται από τον άνθρωπο όταν εντοπίζει εμπόδια σε εικόνες που θέτουν την ζωή των πεζών σε κίνδυνο είναι όμοιοι με τους χάρτες θερμότητας που δημιουργούνται από τους αλγόριθμους οπτικών μετατροπέων που έχουν εκπαιδευτεί να εντοπίζουν τα εμπόδια στα ίδια δεδομένα.

Ερευνητικό Ερώτημα 3: Πόσο όμοιοι είναι οι χάρτες θερμότητας που δημιουργούνται από τις γυναίκες με τους χάρτες που δημιουργούνται από τους άντρες όταν εντοπίζουν εμπόδια σε εικόνες που θέτουν την ζωή των πεζών σε κίνδυνο;

Για το τρίτο ερευνητικό ερώτημα δημιουργήθηκε ένα γράφημα το οποίο αναπαριστά την οπτική διαφορά μεταξύ των χαρτών θερμότητας των αντρών με τους χάρτες θερμότητας των γυναικών. Αυτή η οπτική διαφορά αναπαρίσταται στο Διάγραμμα 10.



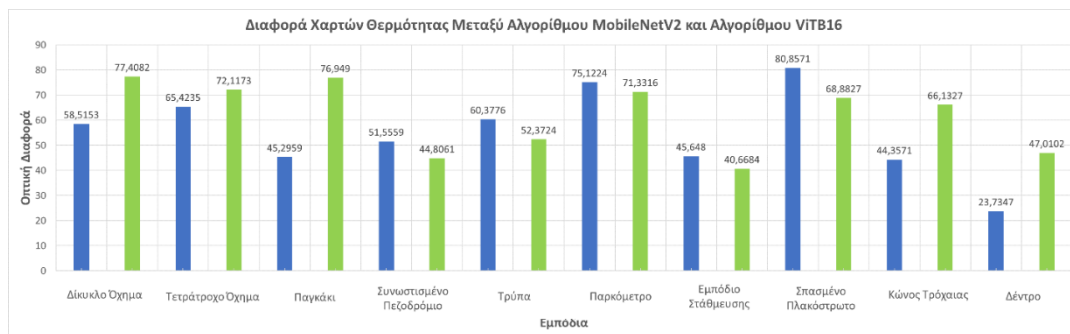
Διάγραμμα 10: Σύγκριση των heatmaps των αντρών με των γυναικών. Μπάρτα μπλε: Part A του πειράματος, μπάρτα πράσινη: Part B του πειράματος

Οι οπτικές διαφορές των χαρτών θερμότητας που διαγράφονται στο Διάγραμμα 10 είναι μεταξύ των τιμών 1 και του 10. Αυτές οι τιμές υποδηλώνουν ότι οι χάρτες θερμότητας που δημιουργούνται από τις γυναίκες είναι όμοιοι με τους χάρτες που δημιουργούνται από τους άντρες όταν εντοπίζουν εμπόδια σε εικόνες που θέτουν την ζωή των πεζών σε κίνδυνο.

Ερευνητικό Ερώτημα 4: Πόσο όμοιοι είναι οι χάρτες θερμότητας που δημιουργούνται από τους αλγόριθμους οπτικών μετατροπέων που έχουν εκπαιδευτεί να εντοπίζουν εμπόδια σε εικόνες που θέτουν την ζωή των πεζών σε κίνδυνο με τους χάρτες θερμότητας που δημιουργούνται από τους αλγόριθμους συνελκτικών δικτύων που έχουν εκπαιδευτεί να εντοπίζουν τα ίδια εμπόδια στα ίδια δεδομένα;

Για το τέταρτο ερευνητικό ερώτημα δημιουργήθηκε το γράφημα οπτικής διαφοράς μεταξύ των χαρτών θερμότητας που δημιουργήθηκαν από τον Αλγόριθμο MobileNetV2 και των χαρτών θερμότητας που δημιουργήθηκαν από τον Αλγόριθμο ViT-B16 (Διάγραμμα 11).

Με βάση το Διάγραμμα 11 παρατηρείται ότι οι τιμές της οπτικής διαφοράς μεταξύ των χαρτών θερμότητας του αλγορίθμου MobileNetV2 και των χαρτών θερμότητας του αλγορίθμου ViT-B16 είναι αρκετά ψηλές. Αυτό υποδεικνύει πως ο αλγόριθμος MobileNetV2 και ο αλγόριθμος ViT-B16 εστιάζουν σε διαφορετικά σημεία για να εντοπίσουν τα εμπόδια.



Διάγραμμα 11: Σύγκριση των δύο αλγορίθμων (MobileNetV2 και ViT-B16). Μπάρα μπλε: Part A του πειράματος, μπάρα πράσινη: Part B του πειράματος

3.4 Σύγκριση διαφορών

Μελετήθηκαν οι διαφορές μεταξύ των χαρτών θερμότητας που υπολογίστηκαν στην προηγούμενη ενότητα ως προς το εάν υπάρχει σημαντική διαφορά μεταξύ τους.

Συγκεκριμένα υλοποιήθηκαν οι ακόλουθες αναλύσεις με την βοήθεια του στατιστικού εργαλείου IBM SPSS (Levesque & SPSS Inc, 2006).

1. Ανάλυση t-test μεταξύ της διαφοράς των χαρτών θερμότητας που δημιουργήθηκαν από όλους/ες τους/ις συμμετέχοντες/ουσες σε σύγκριση με τους χάρτες του Αλγορίθμου MobileNetV2 και των χαρτών που δημιουργήθηκαν από όλους/ες τους/ις συμμετέχοντες/ουσες σε σύγκριση με τους χάρτες του Αλγορίθμου ViT-B16 (Πίνακες 6, 7).
2. Ανάλυση t-test μεταξύ της διαφοράς των χαρτών θερμότητας που δημιουργήθηκαν από τους άντρες σε σύγκριση με τους χάρτες του Αλγορίθμου MobileNetV2 και των χαρτών που δημιουργήθηκαν από τους άντρες σε σύγκριση με τους χάρτες του Αλγορίθμου ViT-B16 (Πίνακες 8, 9).
3. Ανάλυση t-test μεταξύ της διαφοράς των χαρτών θερμότητας που δημιουργήθηκαν από γυναίκες σε σύγκριση με τους χάρτες του Αλγορίθμου MobileNetV2 και των χαρτών που δημιουργήθηκαν από γυναίκες σε σύγκριση με τον Αλγόριθμο ViT-B16 (Πίνακες 10, 11)
4. Ανάλυση t-test μεταξύ της διαφοράς των χαρτών θερμότητας του Part A που δημιουργήθηκαν από όλους/ες τους/ις συμμετέχοντες/ουσες σε σύγκριση με τον Αλγόριθμο MobileNetV2 και των χαρτών του Part B που δημιουργήθηκαν από όλους/ες τους/ις συμμετέχοντες/ουσες σε σύγκριση με τον Αλγόριθμο MobileNetV2 (Πίνακες 12, 13)
5. Ανάλυση t-test μεταξύ των χαρτών θερμότητας του Part A μέρους που δημιουργήθηκαν από όλους/ες τους/ις συμμετέχοντες/ουσες σε σύγκριση με τον Αλγόριθμο ViT-B16 και των χαρτών του Part B που δημιουργήθηκαν από όλους/ες τους/ις συμμετέχοντες/ουσες σε σύγκριση με τον Αλγόριθμο ViT-B16 (Πίνακες 14, 15)

Πίνακας 6: Μέσος όρος των χαρτών όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο MobileNetV2 και των χαρτών όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο ViT-B16

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	all_users vs MobileNetV2	63,1903055	20	17,15120725	3,83512653
	all_users vs ViTB16	23,011990	20	12,2593749	2,7412796

Πίνακας 7: t-test ανάλυση για τους χάρτες όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο MobileNetV2 και των χαρτών όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο ViT-B16

		N	Correlation	Significance	
				One-Sided p	Two-Sided p
Pair 1	all_users vs MobileNetV2 & all_users vs ViTB16	20	,503	,012	,024

Πίνακας 8: Μέσος όρος των χαρτών θερμότητας των αντρών σε σύγκριση με τον Αλγόριθμο MobileNetV2 και των χαρτών των αντρών σε σύγκριση με τον Αλγόριθμο ViT-B16

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	male vs MobileNetV2	64,442605	20	17,3146751	3,8716791
	male vs ViTB16	22,289285	20	12,5424466	2,8045763

Πίνακας 9: t-test ανάλυση για τους χάρτες των αντρών σε σύγκριση με τον Αλγόριθμο MobileNetV2 και των χαρτών των αντρών σε σύγκριση με τον Αλγόριθμο ViT-B16

		N	Correlation	Significance	
				One-Sided p	Two-Sided p
Pair 1	male vs MobileNetV2 & male vs ViTB16	20	,465	,019	,039

Πίνακας 10: Μέσος όρος των χαρτών θερμότητας των γυναικών σε σύγκριση με τον Αλγόριθμο MobileNetV2 και των χαρτών θερμότητας των γυναικών σε σύγκριση με τον Αλγόριθμο ViT-B16

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	female vs MobileNetV2	63,777030	20	17,4019857	3,8912023
	female vs ViTB16	22,998215	20	12,1221058	2,7105853

Πίνακας 11: t-test ανάλυση για τους χάρτες των γυναικών σε σύγκριση με τον Αλγόριθμο MobileNetV2 και των χαρτών των γυναικών σε σύγκριση με τον Αλγόριθμο ViT-B16

		N	Correlation	Significance	
				One-Sided p	Two-Sided p
Pair 1	female vs MobileNetV2 & female vs ViTB16	20	,526	,009	,017

Πίνακας 12: Μέσος όρος των χαρτών θερμότητας του Part A όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο MobileNetV2 και των χαρτών του Part B όλων των συμμετεχόντων σε σύγκριση με τον Αλγόριθμο MobileNetV2

	part	N	Mean	Std. Deviation	Std. Error Mean
all_users vs MobileNetV2	A	10	59,6183710	18,48406402	5,84517427
	B	10	66,7622400	15,84295976	5,00998377

Πίνακας 13: t-test μεταξύ των χαρτών θερμότητας του Part A όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο MobileNetV2 και των χαρτών του Part B όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο MobileNetV2

		Levene's Test for Equality of Variances		t-test for Equality of Means							
		F	Sig.	t	df	Significance		Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						One-Sided p	Two-Sided p			Lower	Upper
all_users vs MobileNetV2	Equal variances assumed	,443	,514	-,928	18	,183	,366	-7,14386900	7,69844138	-23,31769417	9,02995617
	Equal variances not assumed			-,928	17,588	,183	,366	-7,14386900	7,69844138	-23,34486627	9,05712827

Πίνακας 14: Μέσος όρος των χαρτών θερμότητας του Part A όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο ViT-B16 και των χαρτών του Part B όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο ViT-B16

Group Statistics					
	part	N	Mean	Std. Deviation	Std. Error Mean
all_users vs ViTB16	A	10	21,904580	10,9619874	3,4664848
	B	10	24,119400	13,9424992	4,4090054

Πίνακας 15: t-test μεταξύ των χαρτών θερμότητας του Part A όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο ViT-B16 και των χαρτών του Part B όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο ViT-B16

Independent Samples Test											
		Levene's Test for Equality of Variances				t-test for Equality of Means					
		F	Sig.	t	df	Significance		Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						One-Sided p	Two-Sided p			Lower	Upper
all_users vs ViTB16	Equal variances assumed	1,912	,184	-,395	18	,349	,698	-2,2148200	5,6085511	-13,9979486	9,5683086
	Equal variances not assumed			-,395	17,051	,349	,698	-2,2148200	5,6085511	-14,0451567	9,6155167

Παρατηρώντας τα αποτελέσματα των αναλύσεων t-test αποδεικνύεται στατιστικά πως:

1. Η οπτική διαφορά μεταξύ των χαρτών θερμότητας όλων των συμμετεχόντων/ουσών σε σύγκριση με τον αλγόριθμο MobileNetV2 και των χαρτών όλων των συμμετεχόντων/ουσών σε σύγκριση με τον αλγόριθμο ViT-B16 είναι στατιστικά σημαντική γιατί η τιμή p διπλής όψης είναι ίση με 0,024 (< 0,05).
2. Η διαφορά μεταξύ των χαρτών θερμότητας των αντρών σε σύγκριση με τον αλγόριθμο MobileNetV2 και των χαρτών των αντρών σε σύγκριση με τον αλγόριθμο ViT-B16 είναι στατιστικά σημαντική γιατί η τιμή p διπλής όψης είναι ίση με 0,039 (<0.05).
3. Η διαφορά μεταξύ των χαρτών θερμότητας των γυναικών σε σύγκριση με τον αλγόριθμο MobileNetV2 και των χαρτών των γυναικών σε σύγκριση με τον αλγόριθμο ViT-B16 είναι στατιστικά σημαντική γιατί η τιμή p διπλής όψης είναι ίση με 0,017 (<0.05).
4. Η διαφορά μεταξύ των χαρτών θερμότητας του Part A όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο MobileNetV2 και των χαρτών του Part B όλων των συμμετεχόντων/ουσών σε σύγκριση με τον

Αλγόριθμο MobileNetV2 δεν είναι στατιστικά σημαντική γιατί η τιμή p διπλής όψης είναι ίση με 0,366 ($> 0,05$).

5. Η διαφορά μεταξύ των χαρτών θερμότητας του Part A όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο ViT-B16 και των χαρτών του Part B όλων των συμμετεχόντων/ουσών σε σύγκριση με τον Αλγόριθμο ViT-B16 δεν είναι στατιστικά σημαντική γιατί η τιμή p διπλής όψης είναι ίση με 0,349 ($>0,05$).

4 Αποτελέσματα

Όσον αφορά το πρώτο ερευνητικό ερώτημα, παρατηρώντας τα γραφήματα (Διάγραμμα 4-6) εξάγεται το συμπέρασμα πως ο αλγόριθμος MobileNetV2 και κατ' επέκταση η αρχιτεκτονική των συνελκτικών δικτύων δεν βασίζεται στα ίδια σημεία της εικόνας που βασίζεται ο άνθρωπος για να εντοπίσει κάποιο εμπόδιο. Αυτό αποδεικνύεται κυρίως από τις μεγάλες τιμές οπτικής διαφοράς που παρατηρούνται κατά την σύγκριση των χαρτών θερμότητας του αλγορίθμου MobileNetV2 με τους χάρτες θερμότητας όλων των συμμετεχόντων/ουσών ανεξαιρέτως του φύλου για όλα τα εμπόδια που συμπεριλήφθηκαν στην πείραμα.

Οι οπτικές διαφορές των χαρτών που δημιουργήθηκαν στο Part A φαίνεται να είναι κοντά με τις οπτικές διαφορές των χαρτών του Part B οι οποίες διαφοροποιούνται ανάλογα με τον τύπο του εμποδίου. Στην περίπτωση σύγκρισης των χαρτών όλων των συμμετεχόντων/ουσών με τον χάρτες του MobileNetV2 (Διάγραμμα 4) οι μεγαλύτερες οπτικές διαφορές μεταξύ Part A και Part B παρατηρούνται στα εμπόδια «Παγκάκι» και «Δέντρο». Παρόλα αυτά δεν υπάρχει στατιστική διαφορά μεταξύ των οπτικών διαφορών του Part A και Part B όταν συγκρίνονται με τον αλγόριθμο MobileNetV2 (p value >0.05).

Αντίστοιχα για το δεύτερο ερευνητικό ερώτημα, μελετώντας τα γραφήματα (Διάγραμμα 7-9) των οπτικών διαφορών των χαρτών όλων των συμμετεχόντων/ουσών, γυναικών και αντρών, με τους χάρτες του αλγορίθμου ViT-B16 παρατηρούνται μικρές διαφορές. Αυτό οδηγεί στο συμπέρασμα ότι οι άνθρωποι βασίστηκαν σε σχεδόν ίδια σημεία με τα σημεία που βασίστηκε ο αλγόριθμος της αρχιτεκτονικής οπτικών μετασχηματιστών για να εντοπίσουν τα εμπόδια στις εικόνες. Οι οπτικές διαφορές μεταξύ χαρτών ανθρώπου και μηχανής (Διάγραμμα 9) είναι πολύ μικρές για όλα τα εμπόδια με την μικρότερη να παρατηρείται στην περίπτωση του εμποδίου «Παρκόμετρο» και η μεγαλύτερη στην περίπτωση του εμποδίου «Παγκάκι». Αντίστοιχα, ενώ δεν έχει αποδειχθεί να υπάρχει στατιστικά σημαντική διαφορά μεταξύ των οπτικών διαφορών των χαρτών του Part A και του Part B όταν συγκρίνονται με τους χάρτες του αλγορίθμου οπτικών μετασχηματιστών (p value >0.05), παρατηρούνται μεγάλες διαφορές στην περίπτωση κάποιων εμποδίων. Για παράδειγμα στην περίπτωση του εμποδίου «Παγκάκι», η διαφορά της οπτικής διαφοράς μεταξύ του Part A και του Part B είναι ίση με 28.2.

Όσον αφορά το τρίτο ερευνητικό ερώτημα, εξάγεται το συμπέρασμα (Διάγραμμα 10) ότι οι οπτικές διαφορές μεταξύ των χαρτών των αντρών και των γυναικών είναι πολύ μικρές και κατ' επέκταση ότι το φύλο δεν παίζει ρόλο στον τρόπο που οι άνθρωποι εντοπίζουν τα εμπόδια στις εικόνες. Η τιμή της διαφοράς κυμαίνεται από 1 στην περίπτωση του εμποδίου «Τετράτροχο Όχημα» στο Part A μέχρι 7.4 στην περίπτωση του εμποδίου «Δίκυκλο Όχημα» στο Part B. Επιπρόσθετα, οι οπτικές διαφορές μεταξύ Part A και Part B είναι εξίσου αμελητέες στο διάστημα 0.1 (εμπόδιο «Παγκάκι») - 5.7 (εμποδίου «Τετράτροχο Όχημα»).

Τέλος για το τέταρτο ερευνητικό ερώτημα, το γράφημα (Διάγραμμα 11) της οπτικής διαφοράς μεταξύ των χαρτών των αλγορίθμων MobileNetV2 και ViT-B16 επιβεβαιώνει ότι οι 2 αλγόριθμοι εστίασαν κυρίως σε διαφορετικά σημεία στην εικόνα για να εντοπίσουν τα αντίστοιχα εμπόδια. Η οπτική διαφορά των χαρτών κυμαίνεται στο εύρος από 24 μέχρι 80. Η μικρότερη διαφορά σημειώθηκε για το εμπόδιο «Δέντρο» και η μεγαλύτερη για το εμπόδιο «Σπασμένο Πλακόστρωτο» στο Part A. Όπως έχει αποδειχθεί από την ανάλυση t-test δεν υπάρχει στατιστική διαφορά μεταξύ των οπτικών διαφορών των χαρτών που δημιουργήθηκαν στο Part A και τους χάρτες που δημιουργήθηκαν στο Part B και στις 2 περιπτώσεις των αλγορίθμων MobileNetV2 και ViT-B16.

Με βάση τα αποτελέσματα εξάγεται το τελικό συμπέρασμα ότι οι χάρτες του ανθρώπου είναι πιο κοντά με τους χάρτες του αλγορίθμου της αρχιτεκτονικής οπτικών μετασχηματιστών (ViT-B16) παρά με τους χάρτες του αλγορίθμου της αρχιτεκτονικής των συνελκτικών δικτύων (MobileNetV2). Επίσης αποδείχτηκε να υπάρχει σημαντική στατιστική διαφορά ($p < 0.05$) μεταξύ των οπτικών διαφορών των χαρτών θερμότητας του ανθρώπου σε σύγκριση με τον αλγόριθμο MobileNetV2 και των χαρτών του ανθρώπου σε σύγκριση με τον αλγόριθμο ViT-B16.

5 Συζήτηση

Τα διάφορα προβλήματα που αντιμετωπίζουν οι πολίτες κατά την διάρκεια του περπατήματος στο κέντρο της πόλης θα μπορούσαν να μειωθούν με την ανάπτυξη αυτοματοποιημένων μεθόδων εντοπισμού των εμποδίων και ζημιών που παρατηρούνται στα πεζοδρόμια. Με την βοήθεια καμερών κινητών συσκευών και μεθόδων μηχανικής όρασης θα μπορούσε να γίνεται ο εντοπισμός σε πραγματικό χρόνο ούτως ώστε να ενημερώνονται έγκαιρα οι πολίτες για εμπόδια που θέτουν την ζωή τους σε κίνδυνο καθώς και οι αρμόδιες αρχές για την άμεση επίλυση των προβλημάτων. Η ανάπτυξη των αλγορίθμων βαθιάς μάθησης βοήθησε αρκετά προς την επίτευξη αυτού του στόχου επιτυγχάνοντας υψηλά ποσοστά αυτόματης αναγνώρισης αντικειμένων σε οπτικά δεδομένα. Η δυσκολία όμως στην επεξηγηματικότητα των παραγόμενων μοντέλων αναγνώρισης δημιουργεί αρκετά προβλήματα στην κατανόηση του πως λειτουργούν οι αλγόριθμοι βαθιάς μάθησης.

Η παρούσα πτυχιακή εργασία μελετά τις οπτικές ομοιότητες των χαρτών θερμότητας που δημιουργούνται από τους αλγόριθμους βαθιάς μάθησης με τους χάρτες θερμότητας που δημιουργούνται από τους ανθρώπους κατά τον εντοπισμό εμποδίων σε οπτικά δεδομένα που συλλέχθηκαν από κάμερα κινητού τηλεφώνου. Συγκεκριμένα, οι χάρτες θερμότητας αλγορίθμων από δυο διαφορετικές αρχιτεκτονικές βαθιάς μάθησης, συμπεριλαμβανομένων των συνελκτικών νευρωνικών δικτύων και οπτικών μετασχηματιστών, συγκρίθηκαν με τους χάρτες που δημιουργήθηκαν μέσω eye-tracker για 35 συμμετέχοντες/ουσες.

Η σύγκριση των χαρτών θερμότητας οδήγησε σε ενδιαφέροντα αποτελέσματα σχετικά με την αντίληψη των εμποδίων από τον άνθρωπο και την μηχανή αλλά και τις ομοιότητες μεταξύ των σημείων εστίασης τους όταν εντοπίζουν διάφορα εμπόδια που θέτουν την ασφάλεια των πεζών σε κίνδυνο. Οι 2 αλγόριθμοι που χρησιμοποιήθηκαν στην παρούσα εργασία εστιάζουν σε διαφορετικά σημεία της εικόνας για να εντοπίσουν τα εμπόδια, με τα σημεία εστίασης του αλγορίθμου από την αρχιτεκτονική των οπτικών μετασχηματιστών να είναι πιο κοντά με τα σημεία εστίασης του ανθρώπου. Το συγκεκριμένο εύρημα μπορεί να βοηθήσει προς την επίλυση του προβλήματος του μαύρου κουτιού και να οδηγήσει προς την καλύτερη κατανόηση των μοντέλων που δημιουργούνται από τους αλγόριθμους βαθιάς μάθησης.

5.1 Μελλοντική Εργασία

Η μελέτη που εκπονήθηκε στα πλαίσια της παρούσας πτυχιακής εργασίας έχει αρκετές προοπτικές για επέκταση με απώτερο σκοπό την επεξηγηματικότητα των μοντέλων βαθιάς μάθησης. Η μελλοντική έρευνα θα μπορούσε να εστιάσει σε περισσότερα και διαφορετικά εμπόδια, σε μεγαλύτερο αριθμό οπτικών δεδομένων, στην δημιουργία χαρτών θερμότητας με την βοήθεια περισσότερων συμμετεχόντων και στην σύγκριση διαφορετικών αλγορίθμων βαθιάς μάθησης. Τέλος θα μπορούσε να χρησιμοποιηθούν διαφορετικοί τρόποι μέτρησης της οπτικής διαφοράς μεταξύ των χαρτών θερμότητας.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- Ajit, A., Acharya, K., & Samanta, A. (2020). A Review of Convolutional Neural Networks. 2020 International Conference on Emerging Trends in Information Technology and Engineering (Ic-ETITE). <https://doi.org/10.1109/ic-etite47903.2020.049>
- Alpaydin, E. (2020). Introduction To Machine Learning. Mit Press.
- Bishop, C. (2006a). Pattern recognition and machine learning. Springer Verlag.
- Bishop, C. (2006b). Pattern recognition and machine learning. Springer Verlag.
- Bojko, A. (2009). Informative or Misleading? Heatmaps Deconstructed. Human-Computer Interaction. New Trends, 30–39. https://doi.org/10.1007/978-3-642-02574-7_4
- Bourguignon, F., & Sundberg, M. (2007). Aid Effectiveness—Opening the Black Box. *American Economic Review*, 97(2), 316–321. <https://doi.org/10.1257/aer.97.2.316>
- C A Padmanabha Reddy, Y., Viswanath, P., & Eswara Reddy, B. (2018). Semi-supervised learning: a brief review. *International Journal of Engineering & Technology*, 7(1.8), 81. <https://doi.org/10.14419/ijet.v7i1.8.9977>
- Cao, P., Zhu, Z., Wang, Z., Zhu, Y., & Niu, Q. (2022). Applications of graph convolutional networks in computer vision. *Neural Computing and Applications*. <https://doi.org/10.1007/s00521-022-07368-1>
- Christin, S., Hervet, É., & Lecomte, N. (2019). Applications for deep learning in ecology. *Methods in Ecology and Evolution*. <https://doi.org/10.1111/2041-210x.13256>
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. 2009 IEEE Conference on Computer Vision and Pattern Recognition. <https://doi.org/10.1109/cvpr.2009.5206848>
- Emmanuel, I., & Stanier, C. (2016). Defining Big Data. *Proceedings of the International Conference on Big Data and Advanced Wireless Technologies - BDAW '16*. <https://doi.org/10.1145/3010089.3010090>
- Gehlenborg, N., & Wong, B. (2012). Heat maps. *Nature Methods*, 9(3), 213–213. <https://doi.org/10.1038/nmeth.1902>
- George, G., Haas, M. R., & Pentland, A. (2014a). Big Data and Management. *Academy of Management Journal*, 57(2), 321-326. <https://doi.org/10.5465/amj.2014.4002>
- George, G., Haas, M. R., & Pentland, A. (2014b). Big Data and Management. *Academy of Management Journal*, 57(2), 321-326. <https://doi.org/10.5465/amj.2014.4002>
- Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., & Chen, T. (2018). Recent advances in convolutional neural networks. *Pattern Recognition*, 77, 354–377. <https://doi.org/10.1016/j.patcog.2017.10.013>

- Gu, Z., Eils, R., & Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics (Oxford, England)*, 32(18), 2847–2849. <https://doi.org/10.1093/bioinformatics/btw313>
- Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. R. (2012). Improving neural networks by preventing co-adaptation of feature detectors. Cornell University. <https://arxiv.org/abs/1207.0580>
- Islam, J., & Zhang, Y.-Q. (2019). Understanding 3D CNN Behavior for Alzheimer’s Disease Diagnosis from Brain PET Scan. *ArXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.1912.04563>
- Issam El Naqa, Li, R., & Murphy, M. J. (2015). *Machine Learning in Radiation Oncology : Theory and Applications*. Springer International Publishing.
- Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31, 687-695. Springer. <https://doi.org/10.1007/s12525-021-00475-2>
- Kamel, A., Sheng, B., Li, P., Kim, J., & David Dagan Feng. (2021). Hybrid Refinement-Correction Heatmaps for Human Pose Estimation. *IEEE Transactions on Multimedia*, 23, 1330–1342. <https://doi.org/10.1109/tmm.2020.2999181>
- Kayed, M., Anter, A., & Mohamed, H. (2020, February 1). Classification of Garments from Fashion MNIST Dataset Using CNN LeNet-5 Architecture. *IEEE Xplore*. <https://doi.org/10.1109/ITCE48509.2020.9047776>
- Kelleher, J. D. (2019). *Deep learning*. Mit Press.
- Krause, J., Perer, A., & Ng, K. (2016). Interacting with Predictions. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 5686–5697. ACM Digital Library. <https://doi.org/10.1145/2858036.2858529>
- L’Heureux, A., Grolinger, K., Elyamany, H. F., & Capretz, M. A. M. (2017). Machine Learning With Big Data: Challenges and Approaches. *IEEE Access*, 5, 7776–7797. <https://doi.org/10.1109/access.2017.2696365>
- Lettner, F., & Holzmann, C. (2012). Heat maps as a usability tool for multi-touch interaction in mobile applications. *Mobile and Ubiquitous Multimedia*. <https://doi.org/10.1145/2406367.2406427>
- Levesque, R., & SPSS Inc. (2006). *SPSS Programming and Data Management (2nd ed.)*. https://www.spsstools.net/static/spss-programming-book/SPSS_Programming_and_Data_Management_2nd_edition.pdf
- Li, T., Jin, D., Du, C., Cao, X., Chen, H., Yan, J., Chen, N., Chen, Z., Feng, Z., & Liu, S. (2019). The image-based analysis and classification of urine sediments using a LeNet-5 neural network. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 8(1), 109–114. <https://doi.org/10.1080/21681163.2019.1608307>

- Liang, F., Li, Y., & Marculescu, D. (2022). SupMAE: Supervised Masked Autoencoders Are Efficient Vision Learners. ArXiv (Cornell University), 2. <https://doi.org/10.48550/arxiv.2205.14540>
- Lopez Pinaya, W. H., Vieira, S., Garcia-Dias, R., & Mechelli, A. (2020). Convolutional neural networks. *Machine Learning*, 173–191. <https://doi.org/10.1016/B978-0-12-815739-8.00010-9>
- Loyola-Gonzalez, O. (2019). Black-Box vs. White-Box: Understanding Their Advantages and Weaknesses From a Practical Point of View. *IEEE Access*, 7, 154096–154113. <https://doi.org/10.1109/access.2019.2949286>
- M Narasimha Murty, & V Susheela Devi. (2015a). *Introduction to Pattern Recognition and Machine Learning*. World Scientific.
- M Narasimha Murty, & V Susheela Devi. (2015b). *Introduction to Pattern Recognition and Machine Learning (Vol. 5)*. World Scientific.
- Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR)*. [Internet], 9, 381-386
- Mater, A. C., & Coote, M. L. (2019). Deep Learning in Chemistry. *Journal of Chemical Information and Modeling*, 59(6), 2545–2559. <https://doi.org/10.1021/acs.jcim.9b00266>
- Mohanty, H., Bhuyan, P., & Chenthati, D. (2015). *Big Data: A Primer (Studies in Big Data, 11) (2015th ed.)*. Springer.
- O’Shea, K., & Nash, R. (2015). *An Introduction to Convolutional Neural Networks*. <https://arxiv.org/pdf/1511.08458.pdf>
- Poon, A. I. F., & Sung, J. J. Y. (2021). Opening the black box of AI-Medicine. *Journal of Gastroenterology and Hepatology*, 36(3), 581–584. <https://doi.org/10.1111/jgh.15384>
- Qiu, J., Wu, Q., Ding, G., Xu, Y., & Feng, S. (2016). A survey of machine learning for big data processing. *EURASIP Journal on Advances in Signal Processing*, 2016(1). <https://doi.org/10.1186/s13634-016-0355-x>
- Sagiroglu, S., & Sinanc, D. (2013). Big data: a Review. *2013 International Conference on Collaboration Technologies and Systems (CTS)*, 42–47. <https://doi.org/10.1109/cts.2013.6567202>
- Sinha, D., & El-Sharkawy, M. (2019). Thin MobileNet: An Enhanced MobileNet Architecture. *2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, 280–285. <https://doi.org/10.1109/uemcon47517.2019.8993089>
- Statista. (2021, June 7). *Data Created Worldwide 2010-2025 | Statista*. Statista; Statista. <https://www.statista.com/statistics/871513/worldwide-data-created/>
- Szeliski, R. (2022). *Computer Vision: Algorithms and Applications (2nd ed.)*. Springer.

- Tiedje, J. M., Asuming-Brempong, S., Nüsslein, K., Marsh, T. L., & Flynn, S. J. (1999). Opening the black box of soil microbial diversity. *Applied Soil Ecology*, 13(2), 109–122. [https://doi.org/10.1016/S0929-1393\(99\)00026-8](https://doi.org/10.1016/S0929-1393(99)00026-8)
- Tyagi, A. K., & G, R. (2019). Machine Learning with Big Data. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3356269>
- Ulisses Braga-Neto. (2020). *Fundamentals of Pattern Recognition and Machine Learning*. Springer Nature.
- Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). Deep Learning for Computer Vision: A Brief Review. *Computational Intelligence and Neuroscience*, 2018, 1–13. <https://doi.org/10.1155/2018/7068349>
- Wang, L., Wang, G., & Alexander, C. A. (2015). Machine Learning in Big Data [Review of Machine Learning in Big Data]. *International Journal of Advances in Applied Sciences (IJAAS)*, 4, 117–123. <https://download.garuda.kemdikbud.go.id/article.php?article=1494662&val=158&title=Machine%20Learning%20in%20Big%20Data>
- Wang, Q., & Mei, J. (2022). SHDM-NET: Heat Map Detail Guidance with Image Matting for Industrial Weld Semantic Segmentation Network. *ArXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2207.04297>
- Xu, Y., Chen, Y., Wu, J., Pan, J., Liao, C., & Su, H. (2023). The utility of Vision Transformer in preoperatively predicting microvascular invasion status of hepatocellular carcinoma. *HPB*. <https://doi.org/10.1016/j.hpb.2023.01.015>
- Yue, X., Sun, S., Kuang, Z., Wei, M., Philip, Zhang, W. W., & Lin, D. (2021). Vision Transformer with Progressive Sampling. 2021 IEEE/CVF International Conference on Computer Vision (ICCV), 387-396. <https://doi.org/10.1109/iccv48922.2021.00044>
- Zeiler, M. D., & Fergus, R. (2014). Visualizing and Understanding Convolutional Networks. *Computer Vision – ECCV 2014*, 818–833. https://doi.org/10.1007/978-3-319-10590-1_53
- Zhang, Q. (2018). Convolutional Neural Networks (pp. 434–439). In 3rd International Conference on Electromechanical Control Technology and Transportation. <https://pdfs.semanticscholar.org/6b96/b7bba3b1accbcb0361ecd1bdb58999cd76ac.pdf>
- Zhou, B., Krähenbühl, P., & Koltun, V. (2019). Does computer vision matter for action? *ScienceRobotics*, 4(30). <https://doi.org/10.1126/scirobotics.aaw6661>
- Zhou, B., Sun, Y., Bau, D., & Torralba, A. (2018). Interpretable Basis Decomposition for Visual Explanation. 119–134. *ECCV 2018 papers*. https://openaccess.thecvf.com/content_ECCV_2018/html/Antonio_Torralba_Interpretable_Basis_Decomposition_ECCV_2018_paper.html
- Zhou, L., Pan, S., Wang, J., & Vasilakos, A. V. (2017). Machine learning on big data: Opportunities and challenges. *Neurocomputing*, 237, 350–361. <https://doi.org/10.1016/j.neucom.2017.01.026>

Ευαγγέλου, Γ., & Ρούσσου, Ν. (2020). Αρχιτεκτονικές και Υλοποιήσεις του Νευρωνικού Δικτύου LeNet-5 σε FPGAs. <https://nemertes.library.upatras.gr/server/api/core/bitstreams/31b578be-5cea-41f0-a560-363188706d73/content>

Περιγραφή: Μελέτη Μεθόδων Μείωσης Διαστατικότητας. (2020). Auth.gr. <https://search.lib.auth.gr/Record/ikee-123819/Description>

Φίτσιος, Α. (2017). Εφαρμογές Συνελικτικών Νευρωνικών Δικτύων με χρήση του Zynq-7000 SoC. https://ikee.lib.auth.gr/record/292355/files/thesis_Fitsios_7547.pdf

ΠΑΡΑΡΤΗΜΑ Ι

Παράρτημα Ι: Οδηγίες που δόθηκαν στους συμμετέχοντες/ουσες πριν την έναρξη της πρώτης βαθμονόμησης.

«Το πείραμα αφορά εμπόδια στο πεζοδρόμιο τα οποία εμποδίζουν την διέλευση των πεζών. Το πείραμα αποτελείται από δύο μέρη. Στο πρώτο μέρος θα προβάλλονται κάποιες φωτογραφίες και θα πρέπει να μας πεις εάν εντοπίζεις κάποιο εμπόδιο με ένα ναι ή όχι και αναλόγως εάν εντοπίζεις θα σου εμφανιστεί στη συνέχεια μια διαφάνεια με πιθανές κατηγορίες και θα μας πεις σε ποια κατηγορία πιστεύεις πως μπορεί να ανήκει το εμπόδιο που εντόπισες. Οι κατηγορίες θα σου δοθούν σε ένα χαρτί στη συνέχεια για να τις δεις για να ξέρεις πάνω κάτω ποιες είναι οι κατηγορίες, αν έχεις κάποια απορία σχετικά με κάποια από αυτές ευχαρίστως να την λύσουμε. Το δεύτερο μέρος είναι το αντίστροφο του πρώτου. Δηλαδή, θα σου αναφέρουμε πως πιθανότατα να υπάρχει εμπόδιο συγκεκριμένης κατηγορίας στην φωτογραφία που θα ακολουθεί και εσύ θα μας πεις εάν το εντοπίζεις ή όχι. Η διάρκεια του πειράματος είναι 10 λεπτά. Δεν υπάρχει ούτε σωστό ούτε λάθος. Δεν υπάρχει λόγος άγχους. Εμείς την δική σου απάντηση θα λάβουμε υπόψη και αν σου κάνουμε κάποια ερώτηση δεν θα είναι για να σε κάνουμε να σκεφτείς καλύτερα την απάντηση που έδωσες και να δώσεις μία που πιστεύεις πως αυτή θέλουμε να ακούσουμε. Λόγω του ότι για να πραγματοποιηθεί αυτό το πείραμα θα χρησιμοποιηθεί ένας eye tracker ο οποίος βρίσκεται λίγο κάτω από την οθόνη που βρίσκεται μπροστά σου. Για να τον χρησιμοποιήσουμε σωστά, θα πρέπει να γίνουν δύο ρυθμίσεις στις οποίες ο eye tracker θα πιάσει την πορεία του βλέμματος σου και θα το αποθηκεύσει για να μπορέσει να κάνει τα ανάλογα στατιστικά στη συνέχεια. Καθ' όλη τη διάρκεια του πειράματος θα πρέπει να μείνεις ακίνητος/η. Μόνο τα μάτια σου θα μπορούν να κουνιούνται. Ξεκινάμε με τις ρυθμίσεις.»