



ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ  
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ  
ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

«Ανάπτυξη εφαρμογής για κινητά τηλέφωνα με βάση  
την τεχνολογία για την τουριστική καθοδήγηση στην  
Θεσσαλονίκη»

Του φοιτητή  
Σαλαμπάση Νικόλαου  
Αρ. Μητρώου: ait262020

Επιβλέπων  
Διαμαντάρας Κωνσταντίνος  
Καθηγητής

Ημερομηνία 23-6-2023

Ανάπτυξη εφαρμογής κινητής συσκευής για την αναγνώριση των μνημείων της Θεσσαλονίκης

21270

Σαλαμπάσης Νικόλαος

Διαμαντάρας Κωνσταντίνος

26-10-2021

23-6-2023

*Βεβαιώνω ότι είμαι ο συγγραφέας αυτής της εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, έχω καταγράψει τις όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών, εικόνων και κειμένου, είτε αυτές αναφέρονται ακριβώς είτε παραφρασμένες. Επιπλέον, βεβαιώνω ότι αυτή η εργασία προετοιμάστηκε από εμένα προσωπικά, ειδικά ως διπλωματική εργασία, στο Τμήμα Μηχανικών Πληροφορικής και Ηλεκτρονικών Συστημάτων του ΔΙ.ΠΑ.Ε.*

*Η παρούσα εργασία αποτελεί πνευματική ιδιοκτησία του φοιτητή Σαλαμπάση Νικόλαου που την εκπόνησε/αν. Στο πλαίσιο της πολιτικής ανοικτής πρόσβασης, ο συγγραφέας/δημιουργός εκχωρεί στο Διεθνές Πανεπιστήμιο της Ελλάδος άδεια χρήσης του δικαιώματος αναπαραγωγής, δανεισμού, παρουσίασης στο κοινό και ψηφιακής διάχυσης της εργασίας διεθνώς, σε ηλεκτρονική μορφή και σε οποιοδήποτε μέσο, για διδακτικούς και ερευνητικούς σκοπούς, άνευ ανταλλάγματος. Η ανοικτή πρόσβαση στο πλήρες κείμενο της εργασίας, δεν σημαίνει καθ' οιονδήποτε τρόπο παραχώρηση δικαιωμάτων διανοητικής ιδιοκτησίας του συγγραφέα/δημιουργού, ούτε επιτρέπει την αναπαραγωγή, αναδημοσίευση, αντιγραφή, πώληση, εμπορική χρήση, διανομή, έκδοση, μεταφόρτωση (downloading), ανάρτηση (uploading), μετάφραση, τροποποίηση με οποιονδήποτε τρόπο, τμηματικά ή περιληπτικά της εργασίας, χωρίς τη ρητή προηγούμενη έγγραφη συναίνεση του συγγραφέα/δημιουργού.*

Η έγκριση της διπλωματικής εργασίας από το Τμήμα Μηχανικών Πληροφορικής και Ηλεκτρονικών Συστημάτων του Διεθνούς Πανεπιστημίου της Ελλάδος, δεν υποδηλώνει απαραίτητως και αποδοχή των απόψεων του συγγραφέα, εκ μέρους του Τμήματος.



## Πρόλογος

Η Διπλωματική εργασία αυτή επιλέχτηκε με σκοπό την δημιουργία ενός συστήματος ψηφιακής ξενάγησης μέσω κινητού τηλεφώνου για την αναγνώριση των 18 μνημείων πολιτιστικής κληρονομιάς της UNESCO στην Θεσσαλονίκη και την υποβοήθηση της ξενάγησης σε κάθε ένα από αυτά. Στόχος είναι το σύστημα αυτό να μπορέσει να αξιοποιηθεί από τον Οργανισμό Τουρισμού Θεσσαλονίκης ως μια λύση για τους επισκέπτες της πόλης μας και να χρησιμοποιηθεί από όσο τον δυνατό περισσότερους χρήστες. Η Διπλωματική εργασία αυτή μου έδωσε την δυνατότητα να ασχοληθώ και να ειδικευτώ σε τεχνολογίες αιχμής στο κομμάτι της μηχανικής μάθησης και συγκεκριμένα στην μηχανική όραση (Computer Vision), στην ανάπτυξη εφαρμογών για κινητές συσκευές με όλες τις ιδιαιτερότητες και τις ευκαιρίες που αυτή περικλείει και στις βέλτιστες πρακτικές σχετικά με τον σχεδιασμό ενός ελκυστικού και ενδιαφέροντος περιβάλλον χρήστη. Τέλος, αποσκοπεί στην δημιουργία ενός ολοκληρωμένου προϊόντος το οποίο θα τεθεί στην παραγωγή και θα μπορέσει να δημιουργήσει αξία για το κομμάτι του τουρισμού της πόλης μας μεταμορφώνοντας τον τρόπο με τον οποίο κάποιος μπορεί να ξεναγηθεί σε αυτήν.

## Περίληψη

Η βελτίωση της τουριστικής εμπειρίας μέσω της μηχανικής όρασης και της ανάπτυξης εφαρμογών για κινητά τηλέφωνα είναι το κύριο αντικείμενο αυτής της διπλωματικής εργασίας, η οποία διερευνά τη σύγκλιση της τεχνολογίας και του τουρισμού. Η διπλωματική εργασία αποτελείται από έξι κεφάλαια, καθένα από τα οποία πραγματεύεται ένα ξεχωριστό θέμα. Ο ρόλος των διαδραστικών τεχνολογιών στην τουριστική καθοδήγηση και ο αντίκτυπος του ψηφιακού μετασχηματισμού στην τουριστική βιομηχανία παρουσιάζονται στο πρώτο κεφάλαιο με μια ανάλυση της σχέσης μεταξύ τουρισμού και τεχνολογίας. Οι τεχνολογίες ψηφιακής καθοδήγησης στις έξυπνες πόλεις διερευνώνται επόμενο κεφάλαιο, το οποίο εμβαθύνει στη λειτουργία της τεχνολογίας στις τουριστικές περιηγήσεις και εξετάζει τις τεχνολογίες που προσανατολίζονται προς τους χρήστες και τις επιχειρήσεις.

Στη συνέχεια ακολουθούν τα πιο τεχνικά κομμάτια της εργασίας, με επίκεντρο αρχικά την μηχανική όραση (Computer Vision), εστιάζοντας σε πτυχές όπως η αναγνώριση αντικειμένων. Γίνεται ανάλυση του ρόλου της μηχανικής μάθησης στην την μηχανική όραση, καθώς και του τρόπου με τον οποίο η τελευταία λέξη της τεχνολογίας ενσωματώνεται στον τομέα αυτό. Στη συνέχεια, γίνεται λεπτομερής αναφορά στις διάφορες εκδόσεις του YOLO και στα διάφορα αρχιτεκτονικά στοιχεία που υπάρχουν στους αλγορίθμους ανίχνευσης αντικειμένων που βασίζονται σε συνελκτικά νευρωνικά δίκτυα, τα οποία συνθέτουν το μοντέλο YOLO.

Ένα από τα κύρια θέματα αποτελεί η ανάπτυξη εφαρμογών για κινητά τηλέφωνα, με ιδιαίτερη έμφαση στην πλατφόρμα iOS. Καλύπτονται θέματα όπως το λειτουργικό σύστημα iOS και το iOS SDK, η δομή ενός τυπικού project, η γλώσσα προγραμματισμού Swift και τα UI Frameworks τα οποία χρησιμοποιούνται. Το κεφάλαιο αυτό διευκρινίζει τις βέλτιστες πρακτικές στην αρχιτεκτονική εφαρμογών, τις αρχές του αντικειμενοστραφούς προγραμματισμού (OOP) και τις αρχιτεκτονικές σύνθετων συστημάτων.

Η πρακτική ενσωμάτωση των μοντέλων μηχανικής όρασης σε μια εφαρμογή για κινητά, καθώς και η σχεδίαση του UI αποτελούν το επόμενο θέμα. Διερευνάται η εκπαίδευση του μοντέλου, η εξαγωγή του μοντέλου σε μορφή ώστε να αξιοποιηθεί από την πλατφόρμα και η ενσωμάτωσή του στην εφαρμογή. Επιπλέον, το κεφάλαιο καλύπτει την ανάπτυξη του UI που φιλοξενεί τη χρήση του μοντέλου και την ανάκτηση σχετικών δεδομένων από τον ιστό σχετικά με τα μνημεία.

Τέλος εξετάζουμε την επίδοση της εφαρμογής, παρέχοντας μια ανάλυση της αποδοτικότητας και της αποτελεσματικότητάς της στο πλαίσιο του τουρισμού.

Συνολικά, η διπλωματική εργασία αυτή ασχολείται με το πώς οι εξελίξεις στις αναδυόμενες τεχνολογίες, όπως η χρήση εφαρμογών για κινητά τηλέφωνα σε συνδυασμό με συστήματα μηχανικής όρασης και γεωεντοπισμού, μπορούν να βελτιώσουν τις τουριστικές εμπειρίες.

# «Development of a mobile application for technology driven tourist guidance»

«Salampasis Nikolaos»

## **Abstract**

Improving the tourism experience through computer vision and the development of mobile phone applications is the main focus of this thesis, which explores the convergence of technology and tourism. The thesis consists of six chapters, each of which addresses a separate topic. The role of interactive technologies in tourism guidance and the impact of digital transformation on the tourism industry are presented in the first chapter with an analysis of the relationship between tourism and technology. Digital guidance technologies in smart cities are explored in the following chapter, which delves into the function of technology in tourism tours and examines user- and business-oriented technologies.

This is followed by the more technical parts of the thesis, focusing initially on computer vision, focusing on aspects such as object detection. An analysis is made of the role of machine learning in machine vision, and how the latest technology is being integrated in this area. A detailed discussion is then given of the different versions of YOLO and the various architectural elements present in the object detection algorithms based on convolutional neural networks that make up the YOLO model.

One of the main topics covered in this chapter is the development of mobile applications, with a special emphasis on iOS. This involves exploring the iOS SDK and operating system, as well as the structure of application development projects, the use of UI Frameworks, and the Swift programming language. You'll also discover insights on object-oriented programming (OOP) principles, the best practices for application architecture, and complex system architectures.

The practical integration of computer vision models into a mobile application is the main topic of Chapter 5. Training the model, exporting the model to a format usable by the application, and integrating it into the application are explored. In addition, the chapter covers the development of the UI that hosts the model and the retrieval of relevant data from the web about the monuments.

Finally, we examine the performance of the application, providing an analysis of its efficiency and effectiveness in the context of tourism.

Overall, this thesis deals with how developments in emerging technologies, such as the usage of mobile applications in combination with computer vision and geolocation systems, can enhance tourism experiences.



# Περιεχόμενα

Πρόλογος.....	iv
Περίληψη .....	v
Abstract .....	vi
Περιεχόμενα .....	viii
Κατάλογος Σχημάτων .....	x
Κατάλογος Πινάκων .....	x
Συντομογραφίες.....	xii
Κεφάλαιο 1ο: Εισαγωγή.....	1
1.1 Τουρισμός και τεχνολογία .....	1
1.2 Ψηφιακή μεταμόρφωση του τουρισμού.....	1
1.3 Διαδραστικές τεχνολογίες στην τουριστική ξενάγησή.....	2
1.4 Ο ρόλος των smartphones (Εξυπνες κινητές συσκευές) και των φορητών συσκευών στην εξέλιξη της τεχνολογίας και το αντίκτυπο στον τουρισμό .....	4
1.5 Στόχοι και σκοποί.....	5
Κεφάλαιο 2ο: Τουριστική ξενάγηση και τεχνολογία .....	6
2.1 Εισαγωγή στις τεχνολογίες που χρησιμοποιούνται στην τουριστική βιομηχανία.....	6
2.2 Τεχνολογίες προσδιορισμένες προς τον χρήστη .....	7
2.3 Τεχνολογίες προσδιορισμένες προς τις επιχειρήσεις.....	8
2.4 Συστήματα καθοδήγησης/ξενάγησης .....	9
2.5 Τεχνολογίες ψηφιακής καθοδήγησης έξυπνων πόλεων.....	11
Κεφάλαιο 3ο: Computer Vision – Αναγνώριση εικόνας και αντικειμένων.....	14
3.1 Εισαγωγή στην χρήση της μηχανικής μάθησης και το Computer Vision.....	14
3.2 State of the art τεχνολογίες για την μηχανική όραση .....	16
3.3 Αλγόριθμοι ανίχνευσης αντικειμένων με βάση τα CNN .....	19
3.3.1 Μετρικές που χρησιμοποιούνται στους αλγορίθμους ανίχνευσης αντικειμένων .....	19
3.3.2 Το μοντέλο YOLO.....	23
3.3.3 Αρχιτεκτονική των μοντέλων ανίχνευσης αντικειμένων – Αρχιτεκτονική του YOLOv2, YOLOv5 και YOLOv7 .....	23
Κεφάλαιο 4ο: Ανάπτυξη εφαρμογής για κινητά τηλέφωνα.....	29
4.1 Εισαγωγή στην ανάπτυξη εφαρμογών για κινητά τηλέφωνα.....	29
4.2 Εισαγωγή στην ανάπτυξη εφαρμογών για την πλατφόρμα του iOS .....	30
4.2.1 Το λειτουργικό σύστημα iOS και το iOS SDK .....	30



4.2.2	Τυπική δομή ενός project.....	31
4.2.3	Η Γλώσσα προγραμματισμού Swift.....	32
4.2.4	Οδηγίες ανθρώπινης διεπαφής (human interface guidelines HIG).....	35
4.2.5	Βασικά Frameworks που χρησιμοποιούνται στην ανάπτυξη εφαρμογών.....	35
4.2.6	Η Εφαρμογή και το περιβάλλον .....	36
4.2.7	UI Components – UIKit.....	38
4.3	Αρχιτεκτονική εφαρμογής – βέλτιστες πρακτικές.....	40
4.3.1	Γενικές αρχές του αντικειμενοστραφή προγραμματισμού ( object oriented programming OOP) 41	
4.3.2	Αρχιτεκτονικές και αρχές πολύπλοκων συστημάτων .....	41
4.3.3	Σχεδιαστικά πρότυπα MVC, MVVM, VIPER .....	44
4.3.4	Αρχιτεκτονικές που υιοθετεί η Swift .....	47
4.3.5	Εφαρμογές Computer Vision στο iOS .....	48
Κεφάλαιο 5ο:	Ενσωμάτωση – Integration.....	49
5.1	Εισαγωγή .....	49
5.2	Εκπαίδευση του μοντέλου .....	50
5.2.1	Προετοιμασία δεδομένων.....	50
5.2.2	Εκπαίδευση .....	51
5.2.3	Εκπαίδευση με το CreateML.....	55
5.3	Εξαγωγή του μοντέλου για χρήση από την εφαρμογή .....	56
5.4	Ενσωμάτωση στην εφαρμογή .....	57
5.5	Ανάπτυξη του UI της εφαρμογής για την αξιοποίηση του μοντέλου .....	58
Κεφάλαιο 6ο:	Αξιολόγηση .....	64
6.1	Περιπτώσεις Χρήσης– Use Cases .....	64
6.2	Αποτελέσματα.....	64
Κεφάλαιο 7ο:	Συμπεράσματα και προτάσεις βελτίωσης.....	65
7.1	Το μέλλον της τεχνολογίας στον τουρισμό.....	65
7.2	Πρόσθετη μελέτη, προτάσεις για το μέλλον .....	65
ΒΙΒΛΙΟΓΡΑΦΙΑ.....		66
ΠΑΡΑΡΤΗΜΑ Α : Αρχιτεκτονική του μοντέλου .....		70
ΠΑΡΑΡΤΗΜΑ Β : Υπερπαράμετροι του Yolov5.....		71

## Κατάλογος Σχημάτων

Σχήμα 1.1: Η σημασία του ανθρωποκεντρικού σχεδιασμού στη διαμόρφωση των επιπτώσεων των τεχνολογιών του Τουρισμού 4.0 στις τουριστικές εμπειρίες.....	3
Σχήμα 2.1: Τα επίπεδα εμπάθυνσης στα οποία κινούνται οι τεχνολογίες AR VR και MR .....	8
Σχήμα 2.2: Οι βασικοί παράγοντες μιας έξυπνης πόλης.....	11
Σχήμα 2.3: Η ροή της πληροφορίας σε μια ιδανική έξυπνη πόλη .....	12
Σχήμα 3.1: Τυπική διαδικασία εκπαίδευσης και ανάπτυξης ενός μοντέλου μηχανικής μάθησης.....	14
Σχήμα 3.2: Διάγραμμα συσχέτισης των εννοιών AI, ML και Deep Learning .....	15
Σχήμα 3.3: Τα στρώματα του VGG16 .....	18
Σχήμα 3.2: Παράδειγμα λειτουργίας ενός CNN.....	19
Σχήμα 3.5: Αναπαράσταση του IoU .....	20
Σχήμα 3.6: Ο πίνακας σύγχυσης (confusion matrix) .....	21
Σχήμα 3.7: Η εξέλιξη και οι εκδόσεις της οικογένειας YOLO .....	24
Σχήμα 3.8: Η Αρχιτεκτονική των Object Detectors σύμφωνα με το [25].....	24
Σχήμα 3.9: Σχηματική αναπαράσταση των backbone, neck , head του YOLO v5.....	27
Σχήμα 3.10: Η αρχιτεκτονική του YOLOv7 σύμφωνα με [29].....	28
Σχήμα 4.1: Η κατάταξη της ταχύτητας των πιο διαδεδομένων γλωσσών προγραμματισμού σύμφωνα με το [33] .....	32
Σχήμα 4.2: Ο κύκλος ζωής της εφαρμογής .....	36
Σχήμα 4.3: Τα Views σε μια τυπική οθόνη της εφαρμογής ημερολογίου .....	37
Σχήμα 4.4: Βασική δομή που ακολουθούν όλες οι οθόνες σε μια εφαρμογή iOS .....	38
Σχήμα 4.5 Προσανατολισμός και τύποι προβολών σε διαφορετικές διατάξεις .....	39
Σχήμα 4.6: MVC Patern στην swift.....	43
Σχήμα 4.7: MVVM Patern στην swift .....	44
Σχήμα 4.8: Η επικοινωνία μεταξύ των στρωμάτων στην αρχιτεκτονική MVVM.....	45
Σχήμα 4.9: Το διάγραμμα ροής πληροφορίας της αρχιτεκτονικής VIPER.....	45
Σχήμα 5.1 Σχεδιασμός της εφαρμογής και περιγραφή εργασιών .....	49
Σχήμα 5.2 Ένα παράδειγμα εικόνας με πλαίσιο οριοθέτησης από το σύνολο δεδομένων COCO [41] και η ερμηνεία της σε διάφορες μορφές αναπαράστασης.....	50
Σχήμα 5.3: Τα διαφορετικά μεγέθη του YOLOv5 και η απόδοση τους στο COCO dataset.....	52
Σχήμα 5.3: Αποτελέσματα inference στο test set .....	55
Σχήμα 5.5: Διάγραμμα UML με την ροή πληροφορίας.....	57

## Κατάλογος Πινάκων

Πίνακας 1: Στρώματα του YOLO v2.....	25
Πίνακας 2: Αποτελέσματα εκπαίδευσης - Confusion Matrix.....	53
Πίνακας 3: Αποτελέσματα εκπαίδευσης - F1.....	53
Πίνακας 4: Αποτελέσματα εκπαίδευσης - Χρονική Εξέλιξη Μετρικών.....	54
Πίνακας 5 Σύγκριση των μοντέλων YOLO v2 και v5 .....	54



## Συντομογραφίες

AR	Επαυξημένη Πραγματικότητα, Augmented Reality
VR	Εικονική Πραγματικότητα, Virtual Reality
AI	Τεχνητή Νοημοσύνη, Artificial Intelligence
MR	Μικτή Πραγματικότητα Mixed Reality
UI	Διεπαφή χρήστη, User Interface
UX	Εμπειρία χρήστη. User Experience
ΤΠΕ	Τεχνολογίες Πληροφοριών και Επικοινωνιών
GPS	Παγκόσμιο Σύστημα Στιγματοθέτησης, Global Positioning System
ML	Μηχανική Μάθηση, Machine Learning
CNN	Συνελκτικά νευρωνικά δίκτυα (convolutional neural networks)
IoU	Τομή επί της ένωσης (Intersection over Union,)
HIG	Human Interface Guidelines

# Κεφάλαιο 1ο: Εισαγωγή

## 1.1 Τουρισμός και τεχνολογία

Η τεχνολογική επανάσταση είχε βαθύτατο αντίκτυπο στην τουριστική βιομηχανία. Από τη μία πλευρά, η τεχνολογία διευκόλυνε τους ανθρώπους να προγραμματίζουν και να κάνουν κρατήσεις για τα ταξίδια τους, οδηγώντας σε αύξηση του αριθμού των ανθρώπων που ταξιδεύουν τόσο στο εσωτερικό όσο και στο εξωτερικό. Αυτό οφείλεται στο γεγονός ότι η τεχνολογία έχει καταστήσει τη διαδικασία εύρεσης πληροφοριών σχετικά με τους προορισμούς, τη σύγκριση τιμών και την πραγματοποίηση κρατήσεων πολύ πιο εύκολη και προσιτή στο ευρύτερο κοινό.

Από την άλλη πλευρά, η ευρεία χρήση της τεχνολογίας έχει επίσης αλλάξει τον τρόπο με τον οποίο οι άνθρωποι ταξιδεύουν. Για παράδειγμα, οι άνθρωποι χρησιμοποιούν πλέον τα smartphones και ηλεκτρονικές συσκευές για να σχεδιάζουν τα δρομολόγιά τους, να βρίσκουν εστιατόρια και να πλοηγούνται σε άγνωστες πόλεις, μειώνοντας την ανάγκη για παραδοσιακούς ταξιδιωτικούς οδηγούς και χάρτες. Επιπλέον, τα μέσα κοινωνικής δικτύωσης και οι ιστότοποι ταξιδιωτικών κριτικών έχουν δώσει στους ταξιδιώτες πρόσβαση σε πληθώρα πληροφοριών και απόψεων σχετικά με προορισμούς και καταλύματα, επιτρέποντάς τους να λαμβάνουν πιο ενημερωμένες και συνειδητοποιημένες αποφάσεις.

Ωστόσο, η τεχνολογική επανάσταση έφερε και κάποιες προκλήσεις για την τουριστική βιομηχανία. Για παράδειγμα, η άνοδος των διαδικτυακών πλατφορμών ταξιδιωτικών κρατήσεων και η παρακμή των παραδοσιακών ταξιδιωτικών πρακτορείων έχει δυσκολέψει τον ανταγωνισμό των μικρότερων, ανεξάρτητων τουριστικών επιχειρήσεων, καθώς συχνά δεν διαθέτουν τους πόρους και την τεχνολογία για την αποτελεσματική διαδικτυακή προώθησή τους.

Συνολικά, η τεχνολογική επανάσταση έχει επηρεάσει θετικά και αρνητικά την τουριστική βιομηχανία, δημιουργώντας νέες ευκαιρίες και προκλήσεις τόσο για τους ταξιδιώτες όσο και για τους επαγγελματίες του κλάδου.

## 1.2 Ψηφιακή μεταμόρφωση του τουρισμού

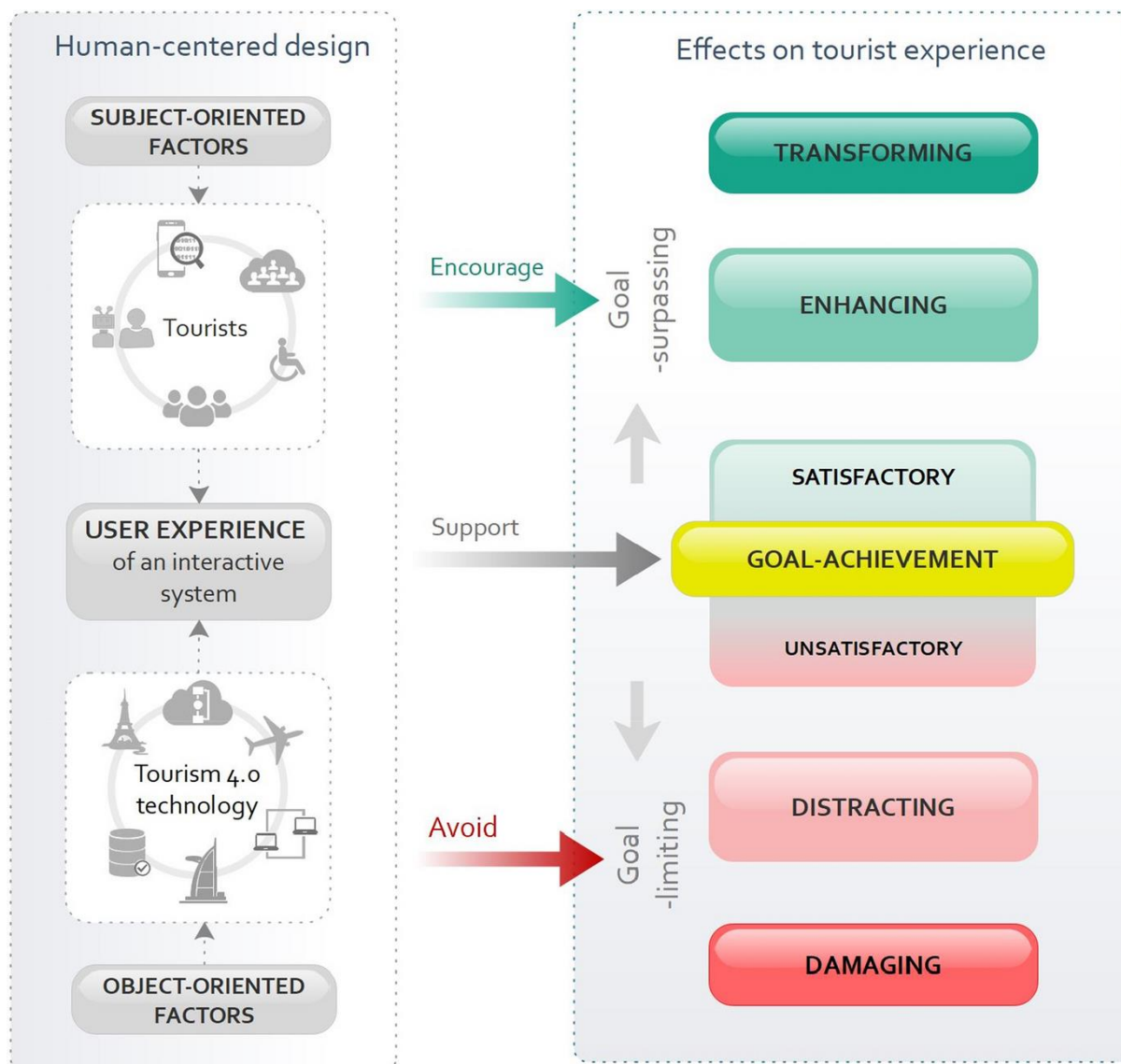
Η πρόοδος της τεχνολογίας έχει μεταμορφώσει βαθιά την τουριστική βιομηχανία με πολλαπλούς τρόπους. Από την έλευση των online κρατήσεων μέχρι την άνοδο των εικονικών περιηγήσεων και των εφαρμογών ηλεκτρονικής ξενάγησης, η ενσωμάτωση της τεχνολογίας στον τουρισμό έχει αλλάξει δραστικά τον τρόπο με τον οποίο οι άνθρωποι ταξιδεύουν και βιώνουν διαφορετικές κουλτούρες και τοποθεσίες. Συγκεκριμένα μερικοί τομείς οι οποίοι έχουν άμεσο αντίκτυπο είναι οι παρακάτω:

1. Τεχνολογία κινητών τηλεφώνων: Η έλευση της κινητής τεχνολογίας παρείχε στους ταξιδιώτες πρόσβαση στις απαραίτητες ταξιδιωτικές χωρίς χρονικούς και χωρικούς περιορισμούς σε πληροφορίες και υπηρεσίες. Από εφαρμογές πλοήγησης με βάση το GPS έως εργαλεία μετάφρασης και πρόσβαση στο διαδίκτυο μέσω κινητού τηλεφώνου, οι εφαρμογές για κινητά έχουν γίνει απαραίτητα βοηθήματα για το ταξίδι.
2. Ψηφιακές κρατήσεις και προγραμματισμός ταξιδιών: Οι διαδικτυακές πλατφόρμες κρατήσεων, όπως η Booking.com και η Airbnb, έχουν αλλάξει ριζικά τον τρόπο με τον οποίο οι ταξιδιώτες κλείνουν τη διαμονή τους [1]. Τα διαδικτυακά ταξιδιωτικά γραφεία, όπως το Booking.com, το Airbnb, το Expedia και το Skyscanner, δίνουν τη δυνατότητα στους χρήστες να συγκρίνουν και να κλείνουν πτήσεις, καταλύματα και ενοικίαση αυτοκινήτων από την άνεση του σπιτιού τους.

3. Κοινωνικά δίκτυα: Οι πλατφόρμες κοινωνικής δικτύωσης, όπως το Facebook, το Instagram και το TripAdvisor, είχαν σημαντικό αντίκτυπο στην τουριστική βιομηχανία. Έχουν γίνει πλατφόρμες έμπνευσης, ανταλλαγής πληροφοριών και κριτικών, επηρεάζοντας έτσι την τουριστική συμπεριφορά [2] των χρηστών τους.
4. Επαυξημένη Πραγματικότητα (Augmented Reality, AR) και Εικονική Πραγματικότητα (Virtual Reality, VR): AR και VR είναι αναδυόμενες τεχνολογίες που υιοθετούνται στον τουριστικό κλάδο, παρέχοντας καθηλωτικές εμπειρίες που μπορούν είτε να ενισχύσουν την εμπειρία του χρήστη είτε να προσομοιάσουν προορισμούς [3].
5. Διαδίκτυο των πραγμάτων (IoT): Το διαδίκτυο των πραγμάτων έχει εφαρμογές σε έξυπνα ξενοδοχεία, παρακολούθηση αποσκευών και την δημιουργία εξατομικευμένων τουριστικών εμπειριών με την χρήση δεδομένων σε πραγματικό χρόνο (real time data) [4].
6. Big Data: Η ανάλυση μεγάλων δεδομένων έχει επιτρέψει στις τουριστικές εταιρείες να αποκτήσουν πληροφορίες σχετικά με τη συμπεριφορά και τις προτιμήσεις των ταξιδιωτών, επιτρέποντάς τους έτσι να παρέχουν εξατομικευμένες προσφορές και υπηρεσίες.
7. Τεχνητή Νοημοσύνη (Artificial Intelligence , AI) : Τα chatbots με τεχνητή νοημοσύνη, τα συστήματα συστάσεων και η αυτοματοποίηση στην εξυπηρέτηση πελατών έχουν βελτιώσει την αποτελεσματικότητα και εξατομικεύσει την δημιουργία προτάσεων που βοηθούν στην τουριστική εμπειρία.

### **1.3 Διαδραστικές τεχνολογίες στην τουριστική ξενάγησή**

Οι τεχνολογίες αυτές μπορούν να βελτιώσουν, να εμπλουτίσουν ή να μεταμορφώσουν τις τουριστικές εμπειρίες, αλλά και να θέσουν προκλήσεις και κινδύνους για τη βιωσιμότητα του τουρισμού. Θα πρέπει να τονιστεί η σημασία του ανθρωποκεντρικού σχεδιασμού για τη δημιουργία διαδραστικών συστημάτων που ωφελούν τόσο τους τουρίστες [5].



Σχήμα 1.1 Η σημασία του ανθρωποκεντρικού σχεδιασμού στη διαμόρφωση των επιπτώσεων των τεχνολογιών του Τουρισμού 4.0 στις τουριστικές εμπειρίες [5]

Από την πλευρά των επιχειρήσεων εμφανίζονται ορισμένες προκλήσεις και ευκαιρίες στην ψηφιακή εποχή, όπως η ανάγκη προσαρμογής, καινοτομίας, και διαφοροποίησης. Οι επιχειρήσεις θα πρέπει να εναγκαλιστούν την ψηφιακή τεχνολογία ως εργαλείο για να βελτιώσουν τις υπηρεσίες τους και να δημιουργήσουν αξία για τους τουρίστες.

Ορισμένα βασικά χαρακτηριστικά των διαδραστικών τεχνολογιών στην τουριστική καθοδήγηση είναι:

- Ανθρωποκεντρική σχεδίαση που εστιάζει στις ανάγκες και τις προτιμήσεις των τουριστών, στο πλαίσιο και τους στόχους που θέλουν να επιτύχουν και των εμπειριών που θέλουν να βιώσουν.
- Η Χρήση τεχνολογιών VR, AR, η μικτή πραγματικότητα, όπου αποτελεί τον συνδυασμό των 2 προηγούμενων, (Mixed Reality, MR) ή το βίντεο 360 μοιρών για να δημιουργήσουν ρεαλιστικές και ελκυστικές προσομοιώσεις τουριστικών προορισμών ή δραστηριοτήτων.

- Η ενσωμάτωση με άλλες τεχνολογίες, όπως το διαδίκτυο των πραγμάτων, η ανάλυση μεγάλων δεδομένων, η τεχνητή νοημοσύνη, το Blockchain ή οι υπηρεσίες που βασίζονται στην χρήση τεχνολογιών προσδιορισμού τοποθεσίας (GPS), για να προσφέρουν εξατομικευμένες, συμφωνά με τις ανάγκες τους και «έξυπνες» υπηρεσίες στους τουρίστες

Εκτός από τις βελτιώσεις στις ευκαιρίες που προσφέρει στις επιχειρήσεις και τις εμπειρίες των χρηστών θα πρέπει να αναλογιστούμε και τις δυσκολίες οι οποίες προκύπτουν για την δημιουργία των εφαρμογών αυτών καθώς αποτελούν συνεχώς μεταβαλλόμενο πεδίο με την εμφάνιση νέων τεχνολογιών και εμποδίων. Επίσης οι εφαρμογές αυτές θα πρέπει να κρατούν μια ισορροπία της διαδραστικότητας μεταξύ της εξερεύνησης του πραγματικού αντικειμένου και της χρήσης της εφαρμογής, λαμβάνοντας υπόψιν την εμπειρία του χρήστη και την ελκυστική παρουσίαση τους.

### **1.4 Ο ρόλος των smartphones (Εξυπνες κινητές συσκευές) και των φορητών συσκευών στην εξέλιξη της τεχνολογίας και το αντίκτυπο στον τουρισμό**

Τα έξυπνα κινητά τηλέφωνα έχουν διαδραματίσει καθοριστικό ρόλο στην εξέλιξη της τεχνολογίας, όχι μόνο όσον αφορά το υλικό (hardware), αλλά και στην προώθηση νέου λογισμικού και υπηρεσιών. Επιτρέποντας την εφαρμογή του διάχυτου υπολογισμού (ubiquitous computing), τα πολύπλοκα υπολογιστικά συστήματα έχουν μεταμορφωθεί με τέτοιο τρόπο ώστε να χωράνε στις τσέπες μας, καθιστώντας το διαδίκτυο, τις εφαρμογές και τις ψηφιακές υπηρεσίες προσβάσιμες από οπουδήποτε. Ειδικότερα κάνοντας χρήση των ασύρματων τεχνολογιών, όπως το 3G, το 4G και τώρα το 5G, καθώς και τεχνολογίες για τη βελτιστοποίηση της μετάδοσης δεδομένων έχουν δώσει την δυνατότητα στους χρήστες να συνδέονται στο διαδίκτυο με πραγματικά φορητό τρόπο. Με την δυνατότητα αυτή και για να υπερβούν τους περιορισμούς σε πόρους υπολογισμού και αποθήκευσης οι φορητές συσκευές έχουν οδηγήσει στην υιοθέτηση και την εξέλιξη των τεχνολογιών cloud (Cloud Computing). Σε συνδυασμό με την πρόοδο και την ενσωμάτωση συστημάτων όπως το GPS, τα επιταχυνσιόμετρα, τα γυροσκόπια, η βιομετρική ασφάλεια (αναγνώριση δακτυλικών αποτυπωμάτων και προσώπου) και το NFC στις συσκευές αυτές έχει οδηγήσει περαιτέρω στην ανάπτυξη εφαρμογών όπως οι υπηρεσίες γεωγραφικού εντοπισμού, οι εφαρμογές υγείας και γυμναστικής, τα συστήματα ανέπαφων πληρωμών και οι εμπειρίες επανξιμένης πραγματικότητας (AR).

Οι δυνατότητες αυτές δημιούργησαν μια καινούργια οικονομία, την οικονομία των εφαρμογών για κινητές συσκευές, το Mobile App Economy. Ο περιορισμένος χώρος της οθόνης των έξυπνων κινητών τηλεφώνων έχουν καταστήσει αναγκαία την πρόοδο στο σχεδιασμό user interface (UI) και user experience (UX), οδηγώντας στην ανάπτυξη πιο φιλικών προς το χρήστη ψηφιακών διεπαφών. Αυτές οι “compact” εκδόσεις των κλασικών εφαρμογών λογισμικού ανταποκρίνονται σε διάφορες ανάγκες των χρηστών, οδηγώντας σε εξελίξεις σε τομείς όπως τα κινητά παιχνίδια, οι οικονομικές συναλλαγές, η υγεία και τα μέσα κοινωνικής δικτύωσης.

Οι κινητές συσκευές διαδραματίζουν μεγάλο ρόλο επίσης έχουν στην ανάπτυξη της τεχνητής νοημοσύνης και της μηχανικής μάθησης. Με την ενσωμάτωση της τεχνητής νοημοσύνης στα smartphones, έχουν σημειωθεί πρόοδοι σε τομείς όπως η αναγνώριση φωνής (π.χ. Siri, Google Assistant), η αναγνώριση εικόνας και αντικειμένων, οι εξατομικευμένες συστάσεις και η ψηφιακή βοήθεια.

Εκτός από το κομμάτι του λογισμικού, οι αυξημένες ανάγκες για μικρότερα, ταχύτερα και πιο ενεργειακά αποδοτικά εξαρτήματα στο κομμάτι του hardware, έχει συμβάλει ώστε να ωθήσει τις εξελίξεις στην τεχνολογία με γρηγορότερους ρυθμούς. Για παράδειγμα, οι βελτιώσεις στους



επεξεργαστές, τις οθόνες, τη μνήμη, τις μπαταρίες και τις κάμερες οδηγούνται συχνά από τη βιομηχανία των smartphones.

Δεν θα ήταν υπερβολή να ισχυριστεί κάποιος ότι ένα κινητό τηλέφωνο είναι πλέον αρκετό στον μέσο χρήστη για να πραγματοποιήσει τις περισσότερες του καθημερινές ανάγκες οπότε παλαιότερα θα χρειαζόταν πρόσβαση σε ηλεκτρονικό υπολογιστή. Σύμφωνα με έρευνες ο αριθμός των παγκόσμιων χρηστών smartphone εκτιμάται σε 6,8 δισεκατομμύρια, σημειώνοντας ετήσια αύξηση 4,2% [6].

Συγκεκριμένα το ποσοστό των χρηστών που χρησιμοποιούν κάποιου είδους εφαρμογή σχετική με τον τουρισμό φτάνει το 25%, με τις εφαρμογές αυτές να κατατάσσονται στην 7<sup>η</sup> θέση μεταξύ των πιο συχνά κατεβασμένων εφαρμογών από τα app stores. Για να προγραμματίσουν ένα ταξίδι εκ των προτέρων, περίπου το 15% των χρηστών χρησιμοποιούν κινητές συσκευές. Τέλος στον ελεύθερο χρόνο τους, περίπου το 85% χρησιμοποιεί smartphones για να προγραμματίσει το επόμενο ταξίδι του [7].

## 1.5 Στόχοι και σκοποί

### Στόχος:

Ο κύριος στόχος αυτής της διπλωματικής εργασίας είναι να περιγράψει την διαδικασία για την σχεδίαση και να ανάπτυξη μιας εφαρμογής πλοήγησης και ανίχνευσης των 18 μνημείων της UNESCO για κινητά με την χρήση μηχανικής μάθησης που θα βελτιώνει την εμπειρία των κατοίκων και των τουριστών της Θεσσαλονίκης οι οποίοι επιθυμούν να τα επισκεφτούν παρέχοντας πληροφορίες σε πραγματικό χρόνο για τα ορόσημα και τα σημεία ενδιαφέροντος καθώς και οδηγίες για το πως να φτάσουν σε αυτά.

### Σκοπός:

- **Βιβλιογραφική ανασκόπηση:** Διεξαγωγή ενδελεχούς ανασκόπησης της υπάρχουσας βιβλιογραφίας σχετικά με τους αλγορίθμους ανίχνευσης αντικειμένων, την ανάπτυξη κινητών εφαρμογών, τη χρήση παρόμοιων τεχνολογιών σε αστικά περιβάλλοντα και τις επιπτώσεις της τεχνολογίας στην βιομηχανία του τουρισμού.
- **Σχεδιασμός συστήματος:** Ανάπτυξη αρχιτεκτονικής για την εφαρμογή ανίχνευσης αντικειμένων που λαμβάνει υπόψη τις απαιτήσεις των χρηστών, και το κατά πόσο είναι εφικτή η χρήση μιας κινητής συσκευής.
- **Ανάπτυξη και βελτιστοποίηση αλγορίθμων:** Σχεδιασμός και βελτιστοποίηση ενός αλγορίθμου ανίχνευσης αντικειμένων, ικανού να αναγνωρίζει με ακρίβεια μια ποικιλία μνημείων, ορόσημων και σημείων ενδιαφέροντος σε διαφορετικές συνθήκες φωτισμού και από διαφορετικές γωνίες.
- **Ανάπτυξη εφαρμογής:** Δημιουργία μιας φιλικής προς το χρήστη εφαρμογής για κινητά τηλέφωνα που εφαρμόζει αποτελεσματικά τον αλγόριθμο ανίχνευσης αντικειμένων και παρέχει στους χρήστες πολύτιμες πληροφορίες σχετικά με τα αντικείμενα που ανιχνεύονται σε πραγματικό χρόνο.
- **Αξιολόγηση:** Δοκιμή της απόδοσης και της ακρίβειας της εφαρμογής σε πραγματικές συνθήκες στη Θεσσαλονίκη.
- **Ανάλυση επιπτώσεων:** Αξιολογήστε τις πιθανές επιπτώσεις της εφαρμογής στον τουρισμό, τον, την προσβασιμότητα και άλλους συναφείς τομείς στη Θεσσαλονίκη.

## Κεφάλαιο 2ο: Τουριστική ξενάγηση και τεχνολογία

### 2.1 Εισαγωγή στις τεχνολογίες που χρησιμοποιούνται στην τουριστική βιομηχανία

Η τεχνολογία και οι Τεχνολογίες Πληροφοριών και Επικοινωνιών (ΤΠΕ) αναφέρονται στη γενική χρήση ψηφιακών συσκευών, λογισμικού, πλατφορμών και δικτύων για την πρόσβαση, την επεξεργασία και την επικοινωνία πληροφοριών στον τουρισμό. Η τεχνολογία έχει φέρει επανάσταση στον τρόπο λειτουργίας των επιχειρήσεων και στον τρόπο αλληλεπίδρασής τους με τους πελάτες τους, οδηγώντας σε αυξημένη αποτελεσματικότητα, βελτίωση στην εμπειρία του χρήστη και νέες επιχειρηματικές ευκαιρίες. Με την έλευση των διαδικτυακών πλατφορμών κρατήσεων και συγκρίσεων, οι καταναλωτές μπορούν πλέον να έχουν εύκολη πρόσβαση σε ένα ευρύ φάσμα ταξιδιωτικών υπηρεσιών, από πτήσεις και καταλύματα έως τοπικές εμπειρίες. Αυτό έχει καταστήσει τα ταξίδια πιο προσιτά και βολικά από ποτέ.

Μπορούμε συνοπτικά να ταξινομήσουμε το επίπεδο υιοθέτησης της τεχνολογίας και των ΤΠΕ σε τέσσερα στάδια στον τουρισμό: προοπτική, αφομοίωση, και οικειοποίηση.

- Η προοπτική αναφέρεται στο αρχικό στάδιο όπου η τεχνολογία και οι ΤΠΕ προσφέρουν νέες δυνατότητες για την ανάπτυξη και την καινοτομία στον τουρισμό. Για παράδειγμα, τα διαδικτυακά ταξιδιωτικά γραφεία (online travel agencies OTA) επιτρέπουν στους τουρίστες να αναζητούν και να κάνουν κρατήσεις ταξιδιωτικών προϊόντων ή υπηρεσιών μέσω διαδικτύου. Οι OTA επιτρέπουν επίσης στους προμηθευτές να προσεγγίσουν μια ευρύτερη αγορά και να μειώσουν το κόστος διανομής. Ένα άλλο παράδειγμα είναι τα μέσα κοινωνικής δικτύωσης, τα οποία επιτρέπουν στους τουρίστες να μοιράζονται τις ταξιδιωτικές τους εμπειρίες και απόψεις με άλλους χρήστες.
- Η διατάραξη αναφέρεται στο στάδιο κατά το οποίο η τεχνολογία και οι ΤΠΕ προκαλούν τις υπάρχουσες δομές και πρακτικές του τουρισμού. Για παράδειγμα, οι πλατφόρμες peer-to-peer, όπως η Airbnb ή η Uber, επιτρέπουν στους τουρίστες να έχουν πρόσβαση σε εναλλακτικές επιλογές διαμονής ή μεταφοράς που ανταγωνίζονται τους παραδοσιακούς παρόχους. Ένα άλλο παράδειγμα είναι οι κινητές συσκευές, οι οποίες επιτρέπουν στους τουρίστες να έχουν πρόσβαση σε πληροφορίες και υπηρεσίες οποτεδήποτε και οπουδήποτε, αλλάζοντας τα ταξιδιωτικά πρότυπα και τις συμπεριφορές τους.
- Η αφομοίωση αναφέρεται στο στάδιο κατά το οποίο η τεχνολογία και οι ΤΠΕ ενσωματώνονται στην τουριστική εμπειρία, δημιουργώντας νέους τρόπους αλληλεπίδρασης και εμπλοκής. Για παράδειγμα, η επαυξημένη πραγματικότητα (AR) ή η εικονική πραγματικότητα (VR) επιτρέπουν στους τουρίστες να βιώσουν προσομοιωμένα περιβάλλοντα που μπορούν να εμπλουτίσουν την ψυχαγωγία τους ή να συμβάλουν στο να μάθουν μια καινούργια πληροφορία με διαδραστικό τρόπο.
- Η οικειοποίηση αναφέρεται στο στάδιο κατά το οποίο η τεχνολογία και οι ΤΠΕ αντικαθιστούν ή μετασχηματίζουν την τουριστική εμπειρία, δημιουργώντας νέους τύπους αξίας και περιεχομένου. Ένα παράδειγμα θα μπορούσε να είναι τα κρυπτονόμισμα, τα οποία επιτρέπουν στους τουρίστες να χρησιμοποιούν αποκεντρωμένες ή ασφαλείς συναλλαγές που μπορούν να μειώσουν τους μεσάζοντες ή τις αμοιβές.

Τεχνολογίες όπως η τεχνητή νοημοσύνη, η μηχανική μάθηση και η ανάλυση μεγάλων δεδομένων επιτρέπουν στις επιχειρήσεις να κατανοούν και να προβλέπουν τη συμπεριφορά και τις προτιμήσεις των πελατών, επιτρέποντάς τους έτσι να παρέχουν εξατομικευμένες προσφορές και βελτιωμένες εμπειρίες στους πελάτες τους. Τεχνολογίες όπως η τεχνητή νοημοσύνη (AI), η ανάλυση μεγάλων δεδομένων, διαδίκτυο των πραγμάτων (IoT), blockchain, ρομποτική και AR/VR ορίζουν την 4ή γενιά τουρισμού (Tourism 4.0 [8]), μπορούν να βοηθήσουν τις επιχειρήσεις να εκσυγχρονίσουν τις διεργασίες τους, να μειώσουν το κόστος και να αυξήσουν την αποδοτικότητα με την οποία λειτουργούν. Αυτό μπορεί να

κυμαίνεται από τη βελτίωση της διαχείρισης της αλυσίδας εφοδιασμού και τη μείωση του διοικητικού φόρτου έως την αυτοματοποίηση της εξυπηρέτησης πελατών και των στρατηγικών μάρκετινγκ. Αυτό μπορεί να επιτευχθεί με πλατφόρμες επιχειρηματικής ευφυΐας (Business Intelligence), οι εταιρείες μπορούν να συγκεντρώνουν, να αναλύουν και να απεικονίζουν μεγάλες ποσότητες δεδομένων, οδηγώντας σε πιο τεκμηριωμένη λήψη αποφάσεων και διαμόρφωση στρατηγικής.

Τέλος θα πρέπει να γίνει και αναφορά στις πλατφόρμες μέσω κοινωνικής οι οποίες έχουν προσφέρει στις επιχειρήσεις νέους τρόπους επικοινωνίας με τους πελάτες, από το μάρκετινγκ και την εξυπηρέτηση πελατών έως την ανταλλαγή περιεχομένου που δημιουργείται από τους χρήστες και τη συλλογή ανατροφοδότησης με σκοπό την συνεχή βελτίωση των υπηρεσιών τους.

Συνοψίζοντας, η τεχνολογία έχει επιφέρει σημαντικές αλλαγές στον κλάδο του τουρισμού, οδηγώντας σε έναν πιο προσιτό, εξατομικευμένο, αποτελεσματικό και καθοδηγούμενο από δεδομένα τομέα (Data driven).

## 2.2 Τεχνολογίες προσδιορισμένες προς τον χρήστη

Στην εποχή της ραγδαίας τεχνολογικής προόδου, διάφοροι κλάδοι υφίστανται σημαντικούς μετασχηματισμούς, και ο τουρισμός δεν αποτελεί εξαίρεση. Οι τεχνολογίες με επίκεντρο τον χρήστη παίζουν καθοριστικό ρόλο στην αναδιαμόρφωση της τουριστικής βιομηχανίας, αλλάζοντας τον τρόπο παροχής και κατανάλωσης των υπηρεσιών από τους χρήστες. Από το αρχικό στάδιο του σχεδιασμού ενός ταξιδιού έως το τελικό στάδιο της ανταλλαγής εμπειριών, η τεχνολογία έχει καταστήσει την όλη διαδικασία πιο βολική, εξατομικευμένη και ελκυστική προς τους ανθρώπους. Σημαντικό είναι να τονίσουμε ότι οι τεχνολογίες αυτές θα πρέπει να ακολουθούν την ανθρωποκεντρική σχεδίαση.

Σύμφωνα με το NIST(National Institute of Standards and Technology) [9], ο ανθρωποκεντρικός σχεδιασμός είναι μια προσέγγιση των διαδραστικών συστημάτων που αποσκοπεί στο να καταστήσει τα συστήματα εύχρηστα και χρήσιμα, εστιάζοντας στους χρήστες, τις ανάγκες και τις απαιτήσεις τους εφαρμόζοντας τις τεχνικές ευχρηστίας στον σχεδιασμό τους. Ο ανθρωποκεντρικός σχεδιασμός είναι μια προσέγγιση επίλυσης προβλημάτων που εμπλέκει την ανθρώπινη οπτική σε όλα τα βήματα της διαδικασίας. Χρησιμοποιείται για το σχεδιασμό προϊόντων και υπηρεσιών που βασίζονται στα προβλήματα, τις ανάγκες και τις προτιμήσεις των τελικών χρηστών.

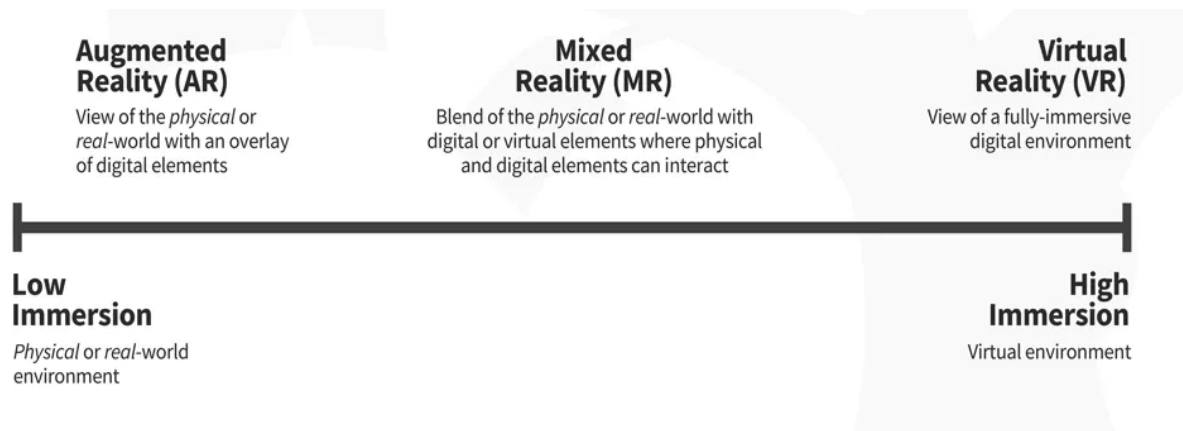
Μερικά παραδείγματα μπορούν να χωριστούν στις παρακάτω κατηγορίες:

**Διαδικτυακές πλατφόρμες κρατήσεων και συγκρίσεων:** Οι ιστότοποι και οι εφαρμογές όπως η Expedia, η Airbnb και η Booking.com έχουν διευκολύνει τους τουρίστες να βρίσκουν, να συγκρίνουν και να κλείνουν καταλύματα, πτήσεις και εμπειρίες από τις προσωπικές τους συσκευές. Τα διαδικτυακά ταξιδιωτικά γραφεία έχουν φέρει επανάσταση στον τρόπο με τον οποίο γίνονται οι κρατήσεις, παρέχοντας πληθώρα επιλογών για τους τελικούς χρήστες.

**Εφαρμογές για κινητά:** Από αεροπορικές εταιρείες έως τοπικούς ξεναγούς, πολλές επιχειρήσεις που σχετίζονται με τον τουρισμό διαθέτουν εφαρμογές για κινητά. Παρέχοντας στον χρήστη υπηρεσίες όπως ηλεκτρονικά εισιτήρια, ειδοποιήσεις σε πραγματικό χρόνο, υπηρεσίες βάσει τοποθεσίας και εξατομικευμένες συστάσεις.

**Επauξημένη και εικονική πραγματικότητα (AR/VR):** Αυτές οι τεχνολογίες επιτρέπουν στους χρήστες να εξερευνήσουν προορισμούς ή να βιώσουν δραστηριότητες πριν αποφασίσουν να τους επισκεφθούν. Μουσεία και ιστορικοί χώροι, για παράδειγμα, χρησιμοποιούν την AR για να βελτιώσουν

την εμπειρία των επισκεπτών παρέχοντας πρόσθετες ψηφιακές πληροφορίες ή αναπαραστάσεις που δεν θα ήταν διαθέσιμες στον χρήστη με διαφορετικό τρόπο.



Σχήμα 2.1 Τα επίπεδα εμβάθυνσης στα οποία κινούνται οι τεχνολογίες AR VR και MR [10]

**Τεχνητή νοημοσύνη (AI) και μηχανική μάθηση (ML):** Η τεχνητή νοημοσύνη μπορεί να βοηθήσει στην πρόβλεψη των ταξιδιωτικών τάσεων, στην εξατομίκευση των εμπειριών των χρηστών. Οι προσωπικοί ηλεκτρονικοί βοηθοί που υποστηρίζονται από AI χρησιμοποιούνται συνήθως για την οργάνωση και την παροχή πληροφοριών σε πραγματικό χρόνο.

**Κοινωνικά δίκτυα:** Πλατφόρμες όπως το Facebook, το Instagram και το YouTube έχουν μεταμορφώσει τον τρόπο με τον οποίο οι τουρίστες αναζητούν και μοιράζονται τις ταξιδιωτικές τους εμπειρίες. Το περιεχόμενο που δημιουργείται από τους χρήστες επηρεάζει σημαντικά τις αποφάσεις των ταξιδιωτών. Το περιεχόμενο που δημιουργείται από τους χρήστες μπορεί να παρέχει πληροφορίες και έμπνευση στους τουρίστες για να σχεδιάσουν τα ταξίδια τους, να επιλέξουν τους προορισμούς τους και να κάνουν κρατήσεις για τα ταξιδιωτικά προϊόντα ή τις υπηρεσίες τους. Επίσης συχνά δημιουργεί προσδοκίες και δίνει μια πιο προσωπική εικόνα για τους προορισμούς πριν από το ταξίδι [11].

**Τεχνολογία γεωγραφικού εντοπισμού:** επιτρέπει στις επιχειρήσεις να προσφέρουν υπηρεσίες που αφορούν την τοποθεσία, όπως εξατομικευμένες συστάσεις για εστιατόρια ή αξιοθέατα σε κοντινή απόσταση. Καθώς τα συστήματα αυτά πλέον είναι ενσωματωμένα στις περισσότερες έξυπνες συσκευές αποτελούν μαζί με τις εφαρμογές χαρτών όπως οι Χάρτες Google ή οι Χάρτες Apple, απαραίτητα εργαλεία ώστε οι χρήστες να μπορούν να περιηγηθούν σε άγνωστους προορισμούς ή ακόμα και μεγάλα κτήρια ή αεροδρόμια. Αξίζει να σημειωθεί, εκτός από την πλοήγηση οι υπηρεσίες γεωγραφικού εντοπισμού μπορεί να χρησιμοποιηθούν για την παροχή πληροφοριών ασφάλειας, όπως κοντινά νοσοκομεία ή υπηρεσίες έκτακτης ανάγκης. Μπορεί επίσης να χρησιμοποιηθεί σε εφαρμογές που επιτρέπουν στους ταξιδιώτες να μοιράζονται την τοποθεσία τους με φίλους ή συγγενείς.

### 2.3 Τεχνολογίες προσδιορισμένες προς τις επιχειρήσεις

Βέβαια, ενώ ένα σημαντικό μέρος της τεχνολογίας που χρησιμοποιείται στον τουριστικό κλάδο είναι όντως επικεντρωμένο στον χρήστη, υπάρχουν επίσης αρκετές τεχνολογίες και πλατφόρμες Business-to-Business (B2B) οι οποίες έχουν μεταμορφώσει τον κλάδο. Οι τεχνολογίες αυτές βοηθούν τις επιχειρήσεις να βελτιώσουν τις υπηρεσίες τους, τις δραστηριότητές τους και να ενισχύσουν τις στρατηγικές τους, οδηγώντας σε μια πιο αποτελεσματική και κερδοφόρα τουριστική βιομηχανία.

Χαρακτηριστικό παράδειγμα τις εξέλιξης αυτής αποτελούν τα παγκόσμια συστήματα διανομής. Τα συστήματα αυτά επιτρέπουν τις συναλλαγές μεταξύ των παροχών υπηρεσιών της ταξιδιωτικής βιομηχανίας και των ταξιδιωτικών γραφείων. Χρησιμοποιούνται κυρίως για κρατήσεις αεροπορικών εισιτηρίων, ενοικιάσεις αυτοκινήτων και δωμάτια ξενοδοχείων. Για την διαχείριση του όγκου της πληροφορίας η οποία είναι διαθέσιμη στις επιχειρήσεις έχουν δημιουργηθεί πλατφόρμες επιχειρηματικής ευφυΐας. Αυτά τα εργαλεία βοηθούν τις επιχειρήσεις να συγκεντρώνουν, να αναλύουν και να απεικονίζουν μεγάλα δεδομένα σε εύπεπτη μορφή. Βοηθούν στη λήψη στρατηγικών αποφάσεων, στην πρόβλεψη τάσεων, στη διαχείριση εσόδων και στην ανάλυση του ανταγωνισμού.

Για την διαχείριση των εσόδων και των πληρωμών χρησιμοποιούνται συστήματα διαχείρισης εσόδων (RMS): Τα RMS χρησιμοποιούνται για την πρόβλεψη της ζήτησης των καταναλωτών με σκοπό τη βελτιστοποίηση της διαθεσιμότητας αποθεμάτων και τιμών. Τα παραδείγματα περιλαμβάνουν τις IDeaS, Duetto και Rainmaker. Σε συνδυασμό με τις πλατφόρμες διαχείρισης καναλιών, τα συστήματα επιτρέπουν στους παρόχους καταλυμάτων να διαχειρίζονται πολλαπλά διαδικτυακά κανάλια διανομής από μια ενιαία πλατφόρμα, αποφεύγοντας έτσι τις υπερκρατήσεις και βελτιστοποιώντας την ροή εσόδων. Παραδείγματα είναι οι υπηρεσίες SiteMinder, eZee και STAAH.

Συστήματα διαχείρισης πελατειακών σχέσεων (Customer Relationship Management CRM): Τα CRM βοηθούν τις επιχειρήσεις να διαχειρίζονται τις σχέσεις τους με τους υφιστάμενους και δυνητικούς πελάτες. Αναλύουν το ιστορικό των πελατών, βελτιώνουν τις επιχειρηματικές σχέσεις, εστιάζοντας ειδικά στη διατήρηση των πελατών και τελικά στην αύξηση των πωλήσεων. Παραδείγματα περιλαμβάνουν Salesforce, Zoho και Hubspot.

Συστήματα διαχείρισης ακινήτων (Property Management Systems PMS): Πρόκειται για πλατφόρμες που επιτρέπουν στις επιχειρήσεις φιλοξενίας να διαχειρίζονται τις καθημερινές τους λειτουργίες. Οι λειτουργίες μπορεί να περιλαμβάνουν διαχείριση κρατήσεων, ανάθεση δωματίων, διαχείριση συντήρησης, τιμολόγηση όπως τα Oracle Hospitality, RMS Cloud και eZee Frontdesk.

Λύσεις πληρωμών B2B: Παρέχουν ασφαλείς μεθόδους (VivaWallet, PayPal) για την αποστολή και λήψη πληρωμών από τις επιχειρήσεις σε παγκόσμιο επίπεδο, που χρησιμοποιούνται συχνά από ταξιδιωτικά γραφεία και επιχειρήσεις φιλοξενίας για συναλλαγές με προμηθευτές και συνεργάτες.

## 2.4 Συστήματα καθοδήγησης/ξενάγησης

Τα συστήματα καθοδήγησης και ξενάγησης διαδραματίζουν κρίσιμο ρόλο στην τουριστική βιομηχανία, καθιστώντας τα ταξίδια πιο προσιτά, αποτελεσματικά και ευχάριστα για τους τουρίστες. Στην εποχή των smartphones και της διάχυτης συνδεσιμότητας, τα συστήματα καθοδήγησης και δρομολόγησης έχουν γίνει αναπόσπαστο κομμάτι των ταξιδιών. Περιλαμβάνουν ένα φάσμα τεχνολογιών και εφαρμογών που παρέχουν βοήθεια πλοήγησης στους τουρίστες, από χάρτες με GPS έως περιηγήσεις επαυξημένης πραγματικότητας (AR).

Τα πιο ευρέως χρησιμοποιούμενα εργαλεία καθοδήγησης και δρομολόγησης είναι εφαρμογές που βασίζονται στο GPS, όπως οι χάρτες Google Maps, Apple Maps. Αυτές οι εφαρμογές όχι μόνο προσφέρουν οδηγίες βήμα βήμα για διάφορους τρόπους μεταφοράς, όπως περπάτημα, ποδηλασία, οδήγηση και δημόσια συγκοινωνία, αλλά παρέχουν επίσης ενημερώσεις για την κυκλοφορία σε πραγματικό χρόνο, εκτιμώμενους χρόνους άφιξης και προτάσεις εναλλακτικών διαδρομών σε περίπτωση εμποδίων ή κυκλοφορικής συμφόρησης. Αυτή η δυνατότητα είναι απαραίτητη για τους τουρίστες που πλοηγούνται σε άγνωστους προορισμούς.

Στον τομέα του μάρκετινγκ και της διαφήμισης, οι επιχειρήσεις αξιοποιούν τα δεδομένα του GPS για να παρέχουν στοχευμένες προωθητικές ενέργειες σε δυνητικούς πελάτες σε συγκεκριμένες περιοχές, ενισχύοντας έτσι την εμβέλεια και τα ποσοστά μετατροπής τους. Η τεχνολογία αυτή έχει επίσης μεταμορφώσει τα αξιοθέατα και τις ξεναγήσεις, με πολλά μουσεία και πάρκα να προσφέρουν πλέον ξεναγήσεις με βάση τον γεωεντοπισμό, διευκολύνοντας πιο προσιτές και πολύγλωσσες εμπειρίες.

Τα δεδομένα γεωγραφικού εντοπισμού αποτελούν επίσης πολύτιμο πόρο για την ανάλυση δεδομένων ταξιδιών. Οι τουριστικές αρχές και οι επιχειρήσεις μπορούν να αντλήσουν πληροφορίες από αυτά τα δεδομένα για να κατανοήσουν τις δημοφιλείς διαδρομές, τα αξιοθέατα και τις συμπεριφορές των τουριστών, ενημερώνοντας έτσι τις αποφάσεις στρατηγικής, μάρκετινγκ και υποδομών τους ώστε να παρέχουν την καλύτερη δυνατή εμπειρία.

Τα ευφυή συστήματα δρομολόγησης για έναν εξατομικευμένο ηλεκτρονικό τουριστικό οδηγό σύμφωνα με τον P. Garcia [12] είναι ικανά επιλύσουν τις ανάγκες των τουριστών σε διάφορες πτυχές: i) θα πρέπει να περιλαμβάνει τις δημόσιες συγκοινωνίες, ii) λαμβάνει υπόψη τους διαφορετικούς χρόνους ταξιδιού, προσαρμόζόμενο σε πραγματικές συνθήκες όπως οι ώρες αιχμής, iii) υπολογίζει τις διαδρομές σε πραγματικό χρόνο για να αντιδρά σε απροσδόκητα γεγονότα, iv) εφαρμόζει ευρετικές μεθόδους για την αποτελεσματική δημιουργία διαδρομών, ακόμη και σε προορισμούς με μεγάλο αριθμό σημείων ενδιαφέροντος και πυκνό δίκτυο δημόσιων συγκοινωνιών.

Επιπλέον, αυτές οι εφαρμογές συχνά ενσωματώνουν σημεία ενδιαφέροντος, επιτρέποντας στους τουρίστες να βρίσκουν και να πλοηγούνται σε κοντινά αξιοθέατα, εστιατόρια ή δημόσιες υπηρεσίες (αστυνομία, νοσοκομεία, πρεσβείες κλπ). Μπορούν επίσης να προσφέρουν λειτουργίες σχεδιασμού διαδρομής, επιτρέποντας στους τουρίστες να σχεδιάζουν αποτελεσματικά το δρομολόγιο και τη σειρά των προορισμών τους με βάση την βέλτιστη διαδρομή ή τις προτιμήσεις τους.

Παράλληλα, οι τεχνολογίες AR ενσωματώνονται σε συστήματα καθοδήγησης και ξεναγήσης για τη δημιουργία πιο διαδραστικών και ελκυστικών ταξιδιωτικών εμπειριών. Η AR μπορεί να ενσωματώσει με την χρήση smartphone ψηφιακές πληροφορίες, όπως ιστορικές λεπτομέρειες ή κριτικές χρηστών, πάνω στο φυσικό περιβάλλον. Για παράδειγμα, ένας τουρίστας που στρέφει την κάμερα του smartphone του σε ένα ιστορικό μνημείο μπορεί να δει πληροφορίες σχετικά με την ιστορία και τη σημασία του. Ορισμένες πόλεις προσφέρουν ακόμη και περιπατητικές περιηγήσεις AR, παρέχοντας μια καθοδηγούμενη, εκπαιδευτική εμπειρία που συνδυάζει τον ψηφιακό και τον φυσικό κόσμο.

Τα συστήματα δημόσιων μεταφορών επωφελούνται επίσης από τις τεχνολογίες ψηφιακής καθοδήγησης. Πολλές πόλεις προσφέρουν εφαρμογές που παρέχουν πληροφορίες για τα μέσα μεταφοράς σε πραγματικό χρόνο, όπως δρομολόγια λεωφορείων ή τρένων, χάρτες διαδρομών και διακοπές υπηρεσιών. Αυτό καθιστά τη χρήση των δημόσιων συγκοινωνιών πιο προσιτή και ελκυστική για τους τουρίστες. Συχνά αυτά τα μέσα συνδυάζονται με τουριστικές ξεναγήσεις στα αξιοθέατα τις εκάστοτε πόλης.

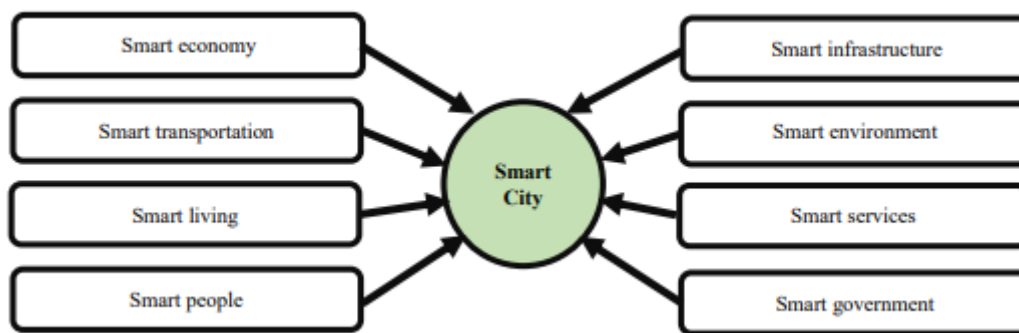
Τέλος, η κοινωνική πτυχή των ταξιδιών έχει ενισχυθεί με την τεχνολογία γεωεντοπισμού. Οι πλατφόρμες κοινωνικής δικτύωσης επιτρέπουν στους τουρίστες να κάνουν "check in" σε τοποθεσίες, προωθώντας τοπικές επιχειρήσεις και αξιοθέατα και επιτρέποντας στους ταξιδιώτες να μοιράζονται τις εμπειρίες τους σε πραγματικό χρόνο.

Εν κατακλείδι, τα συστήματα καθοδήγησης και δρομολόγησης έχουν φέρει επανάσταση στην τουριστική βιομηχανία, διευκολύνοντας την πλοήγηση, εμπλουτίζοντας τις εμπειρίες επίσκεψης στα αξιοθέατα και βοηθώντας τους τουρίστες να αξιοποιήσουν στο έπακρο τις επισκέψεις τους. Είναι σημαντικό να σημειωθεί ότι, ενώ η τεχνολογία γεωεντοπισμού προσφέρει πολλά οφέλη, πρέπει να

λαμβάνονται προσεκτικά υπόψη οι επιπτώσεις στην προστασία της ιδιωτικής ζωής, διασφαλίζοντας ότι οι χρήστες διατηρούν τον έλεγχο των προσωπικών τους δεδομένων. Καθώς η τεχνολογία συνεχίζει να εξελίσσεται, μπορούμε να περιμένουμε ότι αυτά τα εργαλεία θα γίνουν ακόμη πιο εξελιγμένα και αποτελέσουν αναπόσπαστο κομμάτι της τουριστικής εμπειρίας.

## 2.5 Τεχνολογίες ψηφιακής καθοδήγησης έξυπνων πόλεων

Μια έξυπνη πόλη είναι μια αστική περιοχή που χρησιμοποιεί ψηφιακές τεχνολογίες για να βελτιώσει τις υποδομές, να εκσυγχρονίσει τις κυβερνητικές υπηρεσίες, να βελτιώσει την προσβασιμότητα, να προωθήσει τη βιωσιμότητα και να επιταχύνει την οικονομική ανάπτυξη της [13]. Πρόκειται για μια τεχνολογικά σύγχρονη αστική περιοχή που χρησιμοποιεί διάφορες ψηφιακές μεθόδους και αισθητήρες για τη συλλογή συγκεκριμένων δεδομένων, τα οποία χρησιμοποιούνται για την αποτελεσματική διαχείριση υποδομών, των πόρων και των υπηρεσιών.



Σχήμα 2.2: Οι βασικοί παράγοντες μιας έξυπνης πόλης [13]

Οι έξυπνες πόλεις έχουν διάφορα οφέλη για τους κατοίκους, τις επιχειρήσεις και τις κυβερνητικές υπηρεσίες τους. Μερικά από τα χαρακτηριστικά οφέλη σύμφωνα με το [14] είναι:

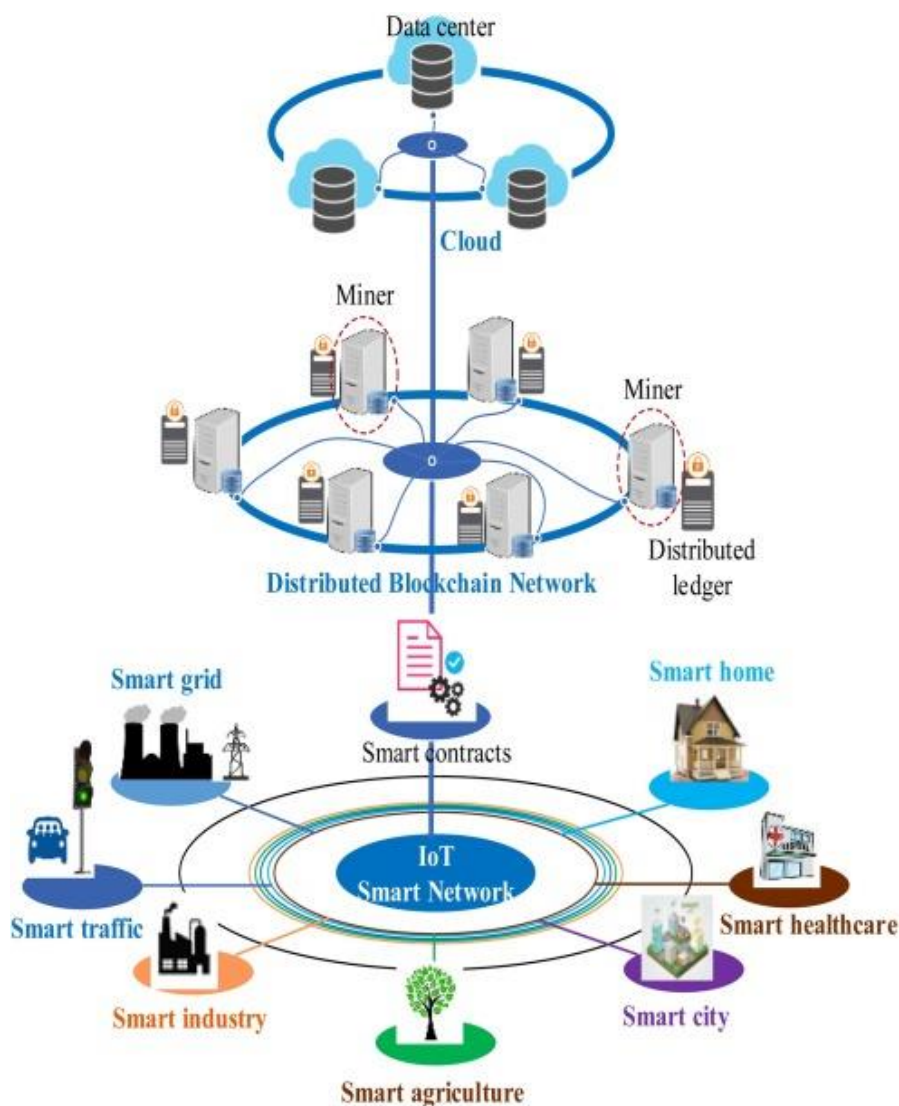
**Αποτελεσματικότερη και βασισόμενη σε δεδομένα λήψη αποφάσεων:** Οι έξυπνες πόλεις μπορούν να έχουν πρόσβαση και να αναλύουν μεγάλο όγκο δεδομένων από διάφορες πηγές και να τα χρησιμοποιούν για τη βελτίωση των λειτουργιών και των υπηρεσιών τους. Για παράδειγμα, οι έξυπνες πόλεις μπορούν να χρησιμοποιούν δεδομένα για τη βελτιστοποίηση της ροής της κυκλοφορίας, τη μείωση της εγκληματικότητας, τον προγραμματισμό της μελλοντικής ανάπτυξης και την αντιμετώπιση καταστάσεων έκτακτης ανάγκης.

**Ενισχυμένη συμμετοχή των πολιτών και της κυβέρνησης:** Οι έξυπνες πόλεις μπορούν να παρέχουν πιο εύχρηστες και βολικές ψηφιακές υπηρεσίες για τους πολίτες τους, όπως διαδικτυακές πύλες, εφαρμογές για κινητά τηλέφωνα, περίπτερα αυτοεξυπηρέτησης και chatbots. Οι υπηρεσίες αυτές μπορούν να αυξήσουν την ικανοποίηση, τη συμμετοχή των πολιτών, καθώς και να βελτιώσουν τη διαφάνεια των κυβερνητικών υπηρεσιών και να δημιουργήσουν μια ροή ανατροφοδότησης για την βελτίωση τους.

**Βελτιωμένες δημόσιες υπηρεσίες και υποδομές:** Οι έξυπνες πόλεις μπορούν να αξιοποιήσουν τεχνολογίες όπως Internet of Things, AI, cloud και blockchain για να παρακολουθούν και να διαχειρίζονται τις δημόσιες υποδομές και τους πόρους τους πιο αποτελεσματικά και αποδοτικά. Για παράδειγμα, οι έξυπνες πόλεις μπορούν να χρησιμοποιούν αισθητήρες για να εντοπίζουν διαρροές

νερού, διακοπές φωτισμού στους δρόμους, επίπεδα απορριμμάτων, ποιότητα αέρα και ηχορύπανση και να χρησιμοποιούν αυτοματισμούς για να τα διορθώνουν ή να τα αναφέρουν. Οι έξυπνες πόλεις μπορούν επίσης να χρησιμοποιούν έξυπνα δίκτυα για την εξισορρόπηση της προσφοράς και της ζήτησης ενέργειας, τη μείωση των διακοπών ρεύματος και την ενσωμάτωση ανανεώσιμων πηγών ενέργειας.

**Ισχυρότερη οικονομική ανάπτυξη και καινοτομία:** Οι έξυπνες πόλεις μπορούν να χρησιμοποιήσουν τις τεχνολογίες που αναφέρθηκαν παραπάνω για να προσελκύσουν περισσότερες επιχειρήσεις, επενδύσεις και τουρίστες, καθώς και να προωθήσουν μια κουλτούρα καινοτομίας και επιχειρηματικότητας. Για παράδειγμα, οι έξυπνες πόλεις μπορούν να χρησιμοποιούν δεδομένα και αναλύσεις για να εντοπίζουν νέες ευκαιρίες για οικονομική ανάπτυξη, να δημιουργούν ένα πιο φιλικό προς τις επιχειρήσεις περιβάλλον, να υποστηρίζουν νεοφυείς επιχειρήσεις και ΜΜΕ και να προωθούν την ψηφιακή ένταξη και την ανάπτυξη δεξιοτήτων. Οι έξυπνες πόλεις μπορούν επίσης να χρησιμοποιούν ψηφιακές πλατφόρμες για να προβάλλουν την πολιτιστική τους κληρονομιά, τις εκδηλώσεις, τα αξιοθέατα και τις υπηρεσίες τους σε δυνητικούς επισκέπτες.



Σχήμα 2.3 Η ροή της πληροφορίας σε μια ιδανική έξυπνη πόλη σύμφωνα με το [14]



**Οι Ψηφιακές Ξεναγήσεις στις έξυπνες πόλεις** μπορεί να γίνουν είτε με ψηφιακό (απομακρυσμένο) ή φυσικό (με την παρουσία του χρήστη) τρόπο που βοηθούν τους επισκέπτες να εξερευνήσουν και να απολαύσουν τα αξιοθέατα και πολιτιστικές εκδηλώσεις με βολικό και εξατομικευμένο τρόπο. Χρησιμοποιούν έξυπνες τεχνολογίες και δεδομένα για την παροχή πληροφοριών, συστάσεων, πλοήγησης, κρατήσεων, πληρωμών και κοινωνικής αλληλεπίδρασης για τους τουρίστες .

Σύμφωνα με το [15] οι εφαρμογές κινητών τηλεφώνων που παρέχουν υπηρεσίες για τους τουρίστες στις έξυπνες πόλεις, όπως πληροφορίες, συστάσεις, πλοήγηση, κρατήσεις, πληρωμές και κοινωνική αλληλεπίδραση. Μπορούν να βελτιώσουν την τουριστική εμπειρία στις έξυπνες πόλεις παρέχοντας λειτουργίες όπως η πρόσβαση σε δεδομένα σε πραγματικό χρόνο σχετικά με τον καιρό, την κυκλοφορία, τις δημόσιες συγκοινωνίες, τις εκδηλώσεις ή τα αξιοθέατα, η εύρεση κοντινών σημείων ενδιαφέροντος ή δραστηριοτήτων με βάση τις προτιμήσεις ή την τοποθεσία, η ηλεκτρονική κράτηση καταλύματος, μεταφορών, εισιτηρίων ή εστιατορίων, η πληρωμή υπηρεσιών με τη χρήση κινητών πορτοφολιών ή κωδικών QR και η ανταλλαγή εμπειριών ή κριτικών με άλλους ταξιδιώτες.

Επιπλέον με τη χρήση της επαυξημένης πραγματικότητας (AR) και εικονικής πραγματικότητας δίνεται η δυνατότητα να δημιουργηθούν καθηλωτικές και διαδραστικές εμπειρίες με την επικάλυψη ψηφιακών πληροφοριών ή εικόνων στον πραγματικό κόσμο (AR) ή την προσομοίωση ενός εντελώς τεχνητού περιβάλλοντος (VR). Μπορούν να βελτιώσουν την τουριστική εμπειρία στις έξυπνες πόλεις παρέχοντας λειτουργίες όπως ο εμπλουτισμός ιστορικών και πολιτιστικών χώρων με πρόσθετες πληροφορίες ή ιστορίες, η δημιουργία εικονικών περιηγήσεων ή προσομοιώσεων αξιοθέατων ή εκδηλώσεων, η δυνατότητα απομακρυσμένων ή εικονικών ταξιδιών και η παιχνιδιοποίηση της αστικής εξερεύνησης [16].

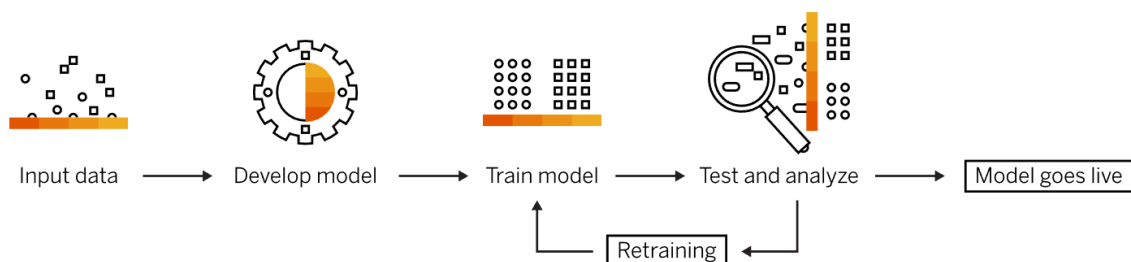
Τα συστήματα ψηφιακής καθοδήγησης, είτε πρόκειται για εφαρμογές για κινητά είτε για διαδραστικά περίπτερα, έχουν γίνει αναπόσπαστα εργαλεία για την πλοήγηση και τη διάδοση πληροφοριών, ενισχύοντας τις τουριστικές εμπειρίες παρέχοντας άμεσα, συναφή και εξειδικευμένα δεδομένα. Σε ευρύτερη κλίμακα, οι έξυπνες πόλεις σε όλο τον κόσμο πρωτοπορούν στη χρήση αυτών των τεχνολογιών, δημιουργώντας διασυνδεδεμένα και έξυπνα αστικά περιβάλλοντα που διευκολύνουν τον τουρισμό, εξασφαλίζοντας παράλληλα τη βιωσιμότητα των τουριστικών επιχειρήσεων και την ενσωμάτωση τους στην ψηφιακή εποχή.

## Κεφάλαιο 3ο: Computer Vision – Αναγνώριση εικόνας και αντικειμένων

### 3.1 Εισαγωγή στην χρήση της μηχανικής μάθησης και το Computer Vision

Σύμφωνα με την IBM [17] το MIT [18] Η μηχανική μάθηση είναι ένας κλάδος της τεχνητής νοημοσύνης (AI) και της επιστήμης των υπολογιστών που επικεντρώνεται στη χρήση δεδομένων και αλγορίθμων για τη μίμηση του τρόπου με τον οποίο μαθαίνει ο άνθρωπος, βελτιώνοντας σταδιακά την ακρίβειά του.

Η μηχανική μάθηση μπορεί να χρησιμοποιηθεί σε διάφορες εφαρμογές, όπως όπως η ταξινόμηση ("classification"), η παλινδρόμηση ("regression"), η ομαδοποίηση ("clustering"), η μείωση των διαστάσεων ("dimensionality reduction"), η ανίχνευση ανωμαλιών ("anomaly detection"), η ενισχυτική μάθηση ("reinforcement learning") και άλλα. Η μηχανική μάθηση μπορεί επίσης να κατηγοριοποιηθεί σε διάφορους τύπους, όπως η μάθηση με επίβλεψη (supervised learning), η μάθηση χωρίς επίβλεψη (unsupervised learning), η μάθηση με ημιεπίβλεψη (semi-supervised learning) και η βαθιά μάθηση (deep learning). Η μάθηση με επίβλεψη είναι όταν ο υπολογιστής λαμβάνει δεδομένα με ετικέτες και μαθαίνει να κάνει προβλέψεις με βάση τις ετικέτες. Μάθηση χωρίς επίβλεψη είναι όταν ο υπολογιστής λαμβάνει δεδομένα χωρίς ετικέτες και μαθαίνει να εξάγει την δομή και χαρακτηριστικά από τα δεδομένα χωρίς καμία καθοδήγηση. Η μάθηση με ημι-επίβλεψη είναι όταν το μοντέλο λαμβάνει ένα μείγμα δεδομένων με και χωρίς ετικέτες και μαθαίνει να αξιοποιεί και τους δύο τύπους πληροφοριών για να βελτιώσει την απόδοσή του.



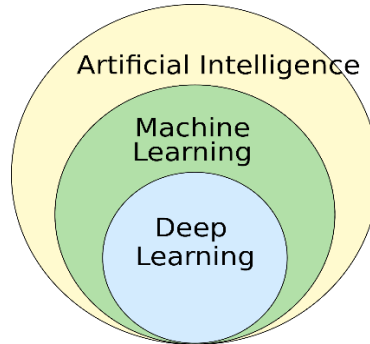
Σχήμα 3.1 Τυπική διαδικασία εκπαίδευσης και ανάπτυξης ενός μοντέλου μηχανικής μάθησης

Η λειτουργία ενός συστήματος μηχανικής μάθησης μπορεί να είναι περιγραφική, δηλαδή το σύστημα χρησιμοποιεί τα δεδομένα για να εξηγήσει τι συνέβη, προγνωστική, δηλαδή το σύστημα χρησιμοποιεί τα δεδομένα για να προβλέψει τι θα συμβεί, ή ρυθμιστική, δηλαδή το σύστημα χρησιμοποιεί τα δεδομένα για να κάνει προτάσεις σχετικά με τις ενέργειες που πρέπει να γίνουν.

Μερικές από τις βασικές έννοιες οι οποίες θα πρέπει να αναφέρουμε είναι τα νευρωνικά δίκτυα και η βαθιά μάθηση (Deep Learning), καθώς και έννοιες σχετικά με τα κομμάτια που αποτελούν ένα ολοκληρωμένο σύστημα μηχανικής μάθησης.

Τα νευρωνικά δίκτυα είναι ένας τύπος αλγορίθμου μηχανικής μάθησης που μιμείται τη δομή και τη λειτουργία του ανθρώπινου εγκεφάλου. Αποτελούνται από διασυνδεδεμένους κόμβους ή νευρώνες που επεξεργάζονται και μεταδίδουν πληροφορίες μεταξύ των επιπέδων. Τα νευρωνικά δίκτυα μπορούν να μαθαίνουν από δεδομένα και να εκτελούν εργασίες όπως ταξινόμηση, παλινδρόμηση, ομαδοποίηση κ.λπ. [19] [20].

Η βαθιά μάθηση είναι επίσης ένα υποσύνολο της μηχανικής μάθησης που χρησιμοποιεί νευρωνικά δίκτυα με πολλαπλά κρυφά στρώματα. Αυτά τα στρώματα μπορούν να εξάγουν σύνθετα χαρακτηριστικά από τα δεδομένα εισόδου και να εκτελούν πιο προηγμένες εργασίες όπως η επεξεργασία φυσικής γλώσσας, η όραση υπολογιστών, η αναγνώριση ομιλίας κ.λπ. Η βαθιά μάθηση απαιτεί περισσότερα δεδομένα και υπολογιστική ισχύ από τα νευρωνικά δίκτυα [19] [20].



Σχήμα 3.2 Διάγραμμα συσχέτισης των εννοιών AI, ML και Deep Learning

Η κύρια διαφορά μεταξύ των νευρωνικών δικτύων και της βαθιάς μάθησης είναι ο αριθμός και η πολυπλοκότητα των στρωμάτων. Τα νευρωνικά δίκτυα έχουν συνήθως ένα ή λίγα κρυφά στρώματα, ενώ τα δίκτυα βαθιάς μάθησης έχουν πολλά στρώματα που μπορούν να μάθουν από διαφορετικά επίπεδα αφαιρέσης. Τα δίκτυα βαθιάς μάθησης είναι επίσης πιο ευέλικτα και ανοιχτά σε scaling από τα νευρωνικά δίκτυα.

Τα θεμελιώδη κομμάτια ενός συστήματος μηχανικής μάθησης για την ολοκληρωμένη και επιτυχή λειτουργία του αποτελούνται από την συλλογή την επεξεργασία δεδομένων και την εξαγωγή χαρακτηριστικών για την εκπαίδευση του μοντέλου, την επιλογή και την υλοποίηση του μοντέλου και τέλος την εκπαίδευση και την αξιολόγηση των αποτελεσμάτων.

**Δεδομένα:** Τα δεδομένα είναι η πρώτη ύλη που χρησιμοποιούν οι αλγόριθμοι μηχανικής μάθησης για να μάθουν και να κάνουν προβλέψεις. Τα δεδομένα μπορεί να είναι δομημένα (οργανωμένα σε πίνακες ή πίνακες) ή μη δομημένα (όπως κείμενο, εικόνες ή ήχος). Τα δεδομένα μπορούν επίσης να είναι επισημασμένα (με προκαθορισμένες κατηγορίες ή τιμές) ή μη επισημασμένα. Η ποιότητα και η ποσότητα των δεδομένων είναι σημαντικοί παράγοντες που επηρεάζουν την απόδοση και την ακρίβεια των αλγορίθμων μηχανικής μάθησης (IBM SOURCE)

**Εξαγωγή Χαρακτηριστικών:** είναι τα χαρακτηριστικά ή οι ιδιότητες των δεδομένων που χρησιμοποιούνται ως είσοδος για τους αλγορίθμους μηχανικής μάθησης. Τα χαρακτηριστικά μπορούν να εξαχθούν από τα ακατέργαστα δεδομένα χρησιμοποιώντας διάφορες μεθόδους, όπως η χειροκίνητη επιλογή και μετατροπή των σχετικών χαρακτηριστικών ή η εξαγωγή χαρακτηριστικών (αυτόματη ανακάλυψη και εκμάθηση χαρακτηριστικών από τα δεδομένα). Τα χαρακτηριστικά γνωρίσματα μπορούν επίσης να είναι αριθμητικά (όπως το ύψος, το βάρος ή η ηλικία) ή κατηγορικά (όπως το χρώμα, το φύλο ή ο τύπος). Η επιλογή χαρακτηριστικών και η μείωση του προβλήματος σε μικρότερες διαστάσεις είναι τεχνικές που αποσκοπούν στη μείωση του αριθμού των χαρακτηριστικών και στην αφαίρεση των μη σχετικών με το πρόβλημα το οποίο καλείται να αντιμετωπίσει ο αλγόριθμος.

**Μοντέλα μηχανικής μάθησης:** Τα μοντέλα είναι οι μαθηματικές αναπαραστάσεις ή συναρτήσεις που χρησιμοποιούν οι αλγόριθμοι μηχανικής μάθησης για να αντιστοιχίσουν τα χαρακτηριστικά εισόδου στις τιμές εξόδου. Τα μοντέλα μπορεί να είναι παραμετρικά (όπως η γραμμική παλινδρόμηση ή η λογιστική παλινδρόμηση) ή μη παραμετρικά (όπως τα δέντρα αποφάσεων ή οι k-nearest neighbors). Τα μοντέλα μπορεί επίσης να είναι απλά (όπως perceptron ή Naïve Bayes) ή σύνθετα (όπως νευρωνικά δίκτυα ή μηχανές διανυσμάτων υποστήριξης). Η επιλογή και η αξιολόγηση μοντέλων είναι τεχνικές που αποσκοπούν στην επιλογή του καλύτερου μοντέλου για το δεδομένο πρόβλημα και στη μέτρηση της απόδοσής του χρησιμοποιώντας μετρικές όπως η ακρίβεια, η ανάκληση ή το F1-score [21] [22].

**Η εκπαίδευση** είναι η διαδικασία προσαρμογής των παραμέτρων ή των βαρών του μοντέλου με βάση τα δεδομένα και την ανατροφοδότηση. Η εκπαίδευση μπορεί να είναι εποπτευόμενη (χρησιμοποιώντας επισημασμένα δεδομένα για την καθοδήγηση της μάθησης), μη εποπτευόμενη (χρησιμοποιώντας μη επισημασμένα δεδομένα για την ανακάλυψη μοτίβων ή δομών) ή ημι-εποπτευόμενη (χρησιμοποιώντας έναν συνδυασμό επισημασμένων και μη επισημασμένων δεδομένων). Η μάθηση μπορεί επίσης να είναι online (συνεχής ενημέρωση του μοντέλου με νέα δεδομένα) ή offline (περιοδική ενημέρωση του μοντέλου με δέσμες δεδομένων). Η μάθηση μπορεί επίσης να είναι ενεργητική (αναζήτηση νέων δεδομένων που είναι πιο κατατοπιστικά) ή παθητική (χρήση υφιστάμενων δεδομένων χωρίς αναζήτηση [17]).

Σύμφωνα με την IBM [17], η Μηχανική όραση ή Computer Vision, είναι ένας κλάδος της τεχνητής νοημοσύνης ο οποίος επιτρέπει στα ψηφιακά συστήματα να εξάγουν συμπεράσματα και πληροφορίες από ψηφιακές εικόνες, βίντεο ή άλλες μεθόδους εισαγωγής οπτικών ροών πληροφορίας. Αξιοποιώντας και αναλύοντας τα δεδομένα αυτά τα συστήματα είναι ικανά να παρέχουν συστάσεις ή να πραγματοποιούν ενέργειες. Η μηχανική μάθηση και συγκεκριμένα τα νευρωνικά δίκτυα δίνουν την δυνατότητα στους υπολογιστές να λαμβάνουν πληροφορίες και να παρακολουθούν τον φυσικό κόσμο. Για να γίνει εφικτό αυτό, είναι αναγκαία η ανάπτυξη θεωρητικών και αλγοριθμικών βάσεων για την επίτευξη της αυτόματης ερμηνείας των οπτικών ερεθισμάτων. Μερικές από τις εφαρμογές οι οποίες είναι εφικτές είναι η ανίχνευση αντικειμένων, η ανίχνευση εκτάκτων συμβάντων, η τρισδιάστατη εκτίμηση πόζας, η εκτίμηση κίνησης, η τρισδιάστατη μοντελοποίηση σκηνών και η αποκατάσταση εικόνων.

Μερικοί τομείς στους οποίους υπάρχουν εφαρμογές Computer Vision, είναι ο τουρισμός, η ρομποτική, η ασφάλεια, η ιατρική, η ψυχαγωγία, η εκπαίδευση, η γεωργία, οι κατασκευές. Η Μηχανική όραση μπορεί να συμβάλλει στη βελτίωση της αποδοτικότητας, της ακρίβειας και της ασφάλειας διάφορων διαδικασιών και συστημάτων που βασίζονται σε οπτικά δεδομένα περιορίζοντας την ανθρώπινη επέμβαση, ελαχιστοποιώντας τους παράγοντες σφάλματος.

### 3.2 State of the art τεχνολογίες για την μηχανική όραση

Για τις ανάγκες της διπλωματικής εργασίας και την δημιουργία ενός συστήματος αναγνώρισης μνημείων χρησιμοποιήθηκαν 2 διαφορετικά μοντέλα (YOLOv2, YOLOv7), τα οποία αποτελούν την αιχμή της τεχνολογίας αναγνώρισης αντικειμένων στον τομέα της μηχανικής όρασης τα οποία είναι βασισμένα στα συνελκτικά νευρωνικά δίκτυα (convolutional neural networks ή CNN). Οι μετρικές οι οποίες μας ενδιαφέρουν είναι η μέση ακρίβεια (average precision) και ο χρόνος εκτέλεσης μιας πρόβλεψης του μοντέλου (inference time). Παρακάτω θα αναλύσουμε τις τεχνολογίες οι οποίες είναι στην αιχμή της τεχνολογίας και μπορούν να χρησιμοποιηθούν για τις ανάγκες της εφαρμογής.

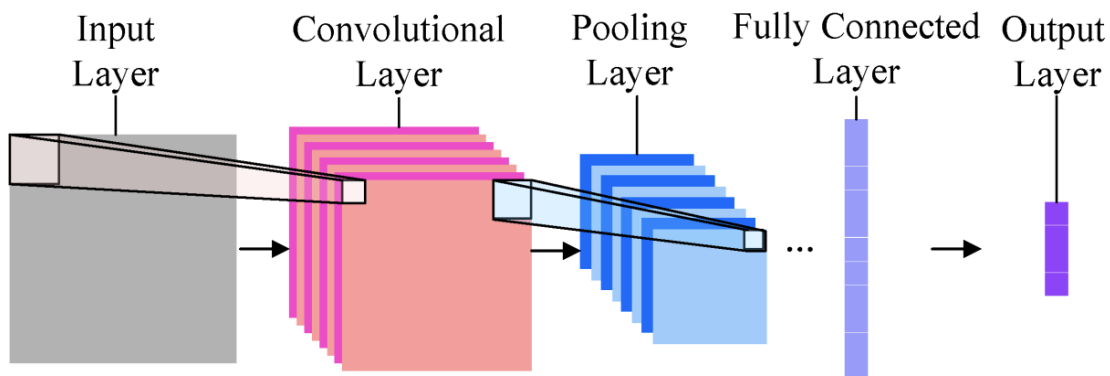
Τα CNN αποτελούνται από διάφορα βασικά συστατικά και λειτουργίες. Η πρωταρχική λειτουργία είναι η συνέλιξη, η οποία εκτελείται από τα στρώματα συνέλιξης. Αυτά τα στρώματα εφαρμόζουν ένα σύνολο

μαθησιακών φίλτρων (kernels) πάνω στην εικόνα εισόδου, εξάγοντας τοπικά χαρακτηριστικά και μοτίβα μέσω πολλαπλασιασμών και αθροισμάτων στα στοιχεία του διανύσματος της εικόνας εισόδου. Αυτή η διαδικασία αποτυπώνει χωρικές ιεραρχίες και επιτρέπει στο δίκτυο να μαθαίνει αναπαραστάσεις αυξανόμενης πολυπλοκότητας και χαρακτηριστικών.

Μετά τη συνέλιξη, μια συνάρτηση ενεργοποίησης, συνήθως ReLU (Rectified Linear Unit), εφαρμόζεται για κάθε φίλτρο. Αυτό εισάγει μη γραμμικότητα στο δίκτυο, επιτρέποντάς του να μάθει πιο σύνθετες αναπαραστάσεις και καθιστώντας το δίκτυο πιο ευπροσάρμοστο.

Για να μειωθούν οι χωρικές διαστάσεις και να διατηρηθούν σημαντικές πληροφορίες, χρησιμοποιούνται επίπεδα δειγματοληψίας (pooling layers). Αυτά τα στρώματα μειώνουν τη δειγματοληψία των χαρτών χαρακτηριστικών (feature maps) που παράγονται από τα στρώματα συνέλιξης, χρησιμοποιώντας συχνά τη μέγιστη συγκέντρωση για την επιλογή της μέγιστης τιμής από μια τοπική γειτνίαση. Αυτή η διαδικασία υποδειγματοληψίας μειώνει τον αριθμό των παραμέτρων και των απαιτούμενων υπολογισμών του μοντέλου. Αφού εφαρμοστούν πολλά convolution layers και pooling layers, συνήθως προστίθενται πλήρως συνδεδεμένα επίπεδα (Fully Connected layers). Αυτά τα στρώματα μοιάζουν με εκείνα που συναντώνται στα παραδοσιακά νευρωνικά δίκτυα και συνδέουν κάθε νευρώνα με κάθε νευρώνα στο προηγούμενο και στα επόμενα στρώματα. Καταγράφουν αναπαραστάσεις υψηλότερου επιπέδου και επιτρέπουν την πραγματοποίηση προβλέψεων με βάση τα εξαγόμενα χαρακτηριστικά.

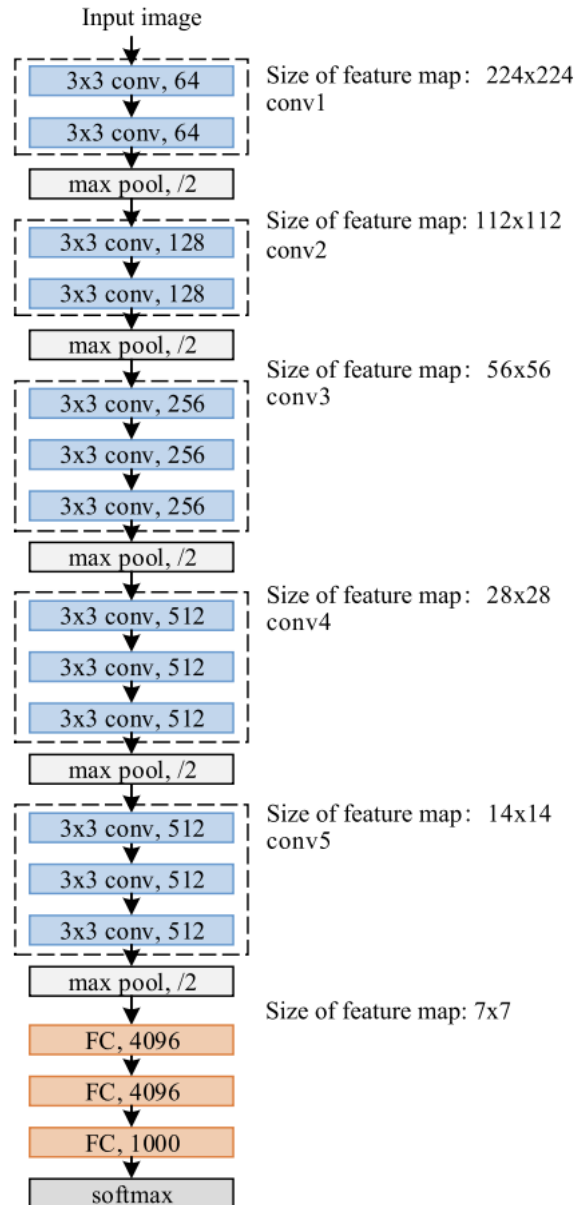
Το συνελκτικό στρώμα και το στρώμα συγκέντρωσης αποτελούν τον εξαγωγέα χαρακτηριστικών, ενώ το πλήρως συνδεδεμένο στρώμα είναι ο ταξινομητής. Το πλήρως συνδεδεμένο στρώμα χαρτογραφεί τον χάρτη χαρακτηριστικών υψηλών διαστάσεων που λαμβάνεται από την εξαγωγή χαρακτηριστικών σε ένα διάνυσμα χαρακτηριστικών 1 διάστασης. Αυτό το διάνυσμα περιέχει όλες τις πληροφορίες χαρακτηριστικών και μπορεί να μετατραπεί σε πιθανότητα κάθε κατηγορίας. Για παράδειγμα, μια εικόνα εισόδου  $224 \times 224$  μπορεί να μετατραπεί σε 4096 χάρτες χαρακτηριστικών  $1 \times 1$  μετά από πολλαπλά επίπεδα συνέλιξης και συγκέντρωσης. Με βάση αυτά τα 4096 χαρακτηριστικά, ένα πλήρως συνδεδεμένο επίπεδο μπορεί να χρησιμοποιηθεί για να καθορίσει ποια από αυτά είναι γουρούνια, σκύλοι, γάτες, αγελάδες ή πρόβατα. [23]



Κατά τη διάρκεια της εκπαίδευσης, τα CNN χρησιμοποιούν μια συνάρτηση απωλειών για να μετρήσουν τη διαφορά μεταξύ της προβλεπόμενης εξόδου και της θεμελιώδους αλήθειας. Οι συνήθεις συναρτήσεις απωλειών για εργασίες ταξινόμησης περιλαμβάνουν συναρτήσεις όπως categorical cross-entropy και softmax. Οι παράμετροι του δικτύου ενημερώνονται χρησιμοποιώντας gradient based μεθόδους βελτιστοποίησης, όπως η backpropagation, οι οποίες ελαχιστοποιούν επαναληπτικά την απώλεια και βελτιώνουν την απόδοση του δικτύου.

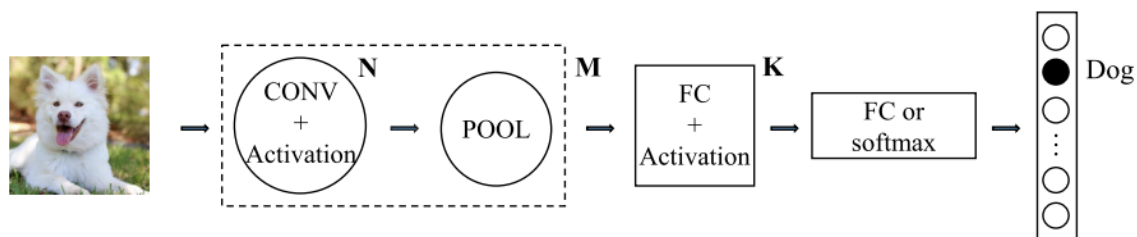
Διαφορετικά CNN κατασκευάζονται με διάφορους συνδυασμούς επιπέδων συνελκτικής σύνδεσης, πλήρως συνδεδεμένων επιπέδων και επιπέδων συγκέντρωσης. Το δίκτυο VGG16 στο Σχήμα 3.3

χρησιμοποιεί στρώματα συνελκτικού τύπου και στρώματα συγκέντρωσης για την εξαγωγή χαρακτηριστικών και μειώνει το μέγεθος του χάρτη χαρακτηριστικών από  $224 \times 224$  σε  $112 \times 112$ , σε  $56 \times 56$ , έως  $7 \times 7$ .



Σχήμα 3.3 Τα στρώματα του VGG16

Στην ταξινόμησή του, οι  $512 \times 7 \times 7$  χάρτες συνελκτικών χαρακτηριστικών περνούν από το πρώτο πλήρως συνδεδεμένο στρώμα για την έξοδο 4096 χαρακτηριστικών, τα χαρακτηριστικά εξόδου περνούν από το δεύτερο πλήρως συνδεδεμένο στρώμα χωρίς αλλαγή του μεγέθους των χαρακτηριστικών και στη συνέχεια τα χαρακτηριστικά που προκύπτουν τροφοδοτούνται στο τρίτο Deep learning πλήρως συνδεδεμένο στρώμα για να προκύψουν 1000 χαρακτηριστικά. Τέλος, αυτά τα 1000 χαρακτηριστικά περνούν μέσα από το στρώμα softmax για να παραχθεί η τελική έξοδος του νευρωνικού δικτύου, δηλαδή σε ποια κλάση από τις 1000 ανήκει το αντικείμενο εισόδου.



Σχήμα 3.4 Παράδειγμα λειτουργίας ενός CNN

Στο Σχήμα 3.4 παρουσιάζονται ορισμένοι συνήθεις συνδυασμοί επιπέδων στα CNN. Η συνέλιξη και τα στρώματα συγκέντρωσης εμφανίζονται συνήθως εναλλάξ, με ένα στρώμα συνέλιξης να ακολουθείται από ένα στρώμα συγκέντρωσης. Μερικές φορές, ένα στρώμα συνέλιξης μπορεί να συνδέεται με δύο τρία ακόμη στρώματα συνέλιξης, ακολουθούμενα από ένα στρώμα συγκέντρωσης, δηλαδή  $N$  στρώματα συνέλιξης στρώματα ακολουθούμενα από ένα στρώμα συγκέντρωσης. Μετά την επανάληψη αυτού του μοτίβου  $M$  φορές, όλα τα χαρακτηριστικά εξάγονται, και στη συνέχεια αυτά τα χαρακτηριστικά αντιστοιχίζονται σε  $O$  χαρακτηριστικά εξόδου μέσω  $K$  πλήρως συνδεδεμένα στρώματα. Τέλος, ένα πλήρως συνδεδεμένο στρώμα, ένα softmax ή μερικές φορές και τα δύο χρησιμοποιούνται για τον προσδιορισμό της κλάσης εξόδου.

### 3.3 Αλγόριθμοι ανίχνευσης αντικειμένων με βάση τα CNN

Σε εφαρμογές του πραγματικού κόσμου, τα αντικείμενα πρέπει όχι μόνο να ανιχνεύονται, αλλά και να εντοπίζονται, έως ότου ανιχνευθούν και εντοπιστούν όλα τα αντικείμενα της εικόνας. Αυτή είναι η αποστολή της ανίχνευσης αντικειμένων στην εικόνα.

Επί του παρόντος, οι αλγόριθμοι ανίχνευσης αντικειμένων εικόνας με βάση τα

CNN χωρίζονται κυρίως σε δύο κατηγορίες: αλγόριθμοι δύο σταδίων και αλγόριθμοι ενός σταδίου. Οι αλγόριθμοι δύο σταδίων βασίζονται στην τεχνική πρότασης περιοχής (regional proposal). Πρώτον, δημιουργούνται προτάσεις περιοχής (πλαίσια οριοθέτησης, bounding boxes) για τον εντοπισμό όλων των αντικειμένων. Στη συνέχεια, κάθε πρόταση περιοχής ταξινομείται χρησιμοποιώντας έναν αλγόριθμο ταξινόμησης εικόνας με βάση το CNN. Οι αλγόριθμοι βασισμένοι στο R-CNN είναι αντιπροσωπευτικοί αλγόριθμοι δύο σταδίων. Οι αλγόριθμοι ενός σταδίου επεξεργάζονται απευθείας την εικόνα εισόδου και βγάζουν ταυτόχρονα τη θέση και την κατηγορία του αντικειμένου-αντιπροσωπεύονται κυρίως από τον αλγόριθμο YOLO και SSD.

Στην παρούσα ενότητα παρουσιάζονται αρχικά οι μετρικές αξιολόγησης των αλγορίθμων ανίχνευσης αντικειμένων εικόνας και στη συνέχεια παρουσιάζονται λεπτομερώς τρεις αλγόριθμοι ανίχνευσης αντικειμένων εικόνας, συμπεριλαμβανομένων της σειράς R-CNN, της σειράς YOLO και του αλγορίθμου SSD.

#### 3.3.1 Μετρικές που χρησιμοποιούνται στους αλγορίθμους ανίχνευσης αντικειμένων

##### Τομή επί της ένωσης (Intersection over Union, IoU)

Υποθέτοντας ότι υπάρχει μόνο ένα αντικείμενο στην εικόνα εισόδου, όταν εντοπίζουμε αυτό το αντικείμενο, το αποτέλεσμα εξόδου θα πρέπει να είναι το πλαίσιο οριοθέτησης (bounding box) αυτού του αντικειμένου. Η πραγματική θέση του σκύλου στο Σχήμα 3.5 είναι το τετράγωνο πλαίσιο B. Εάν ο

εντοπισμός είναι ανακριβής, η θέση μπορεί να είναι το τετράγωνο πλαίσιο A. Για τη μέτρηση της ακρίβειας του εντοπισμού εικόνας με ένα μόνο αντικείμενο, χρησιμοποιείται συνήθως ως μετρική αξιολόγησης η τομή επί της ένωσης (Intersection over Union, IoU). Η IoU υπολογίζεται ως η περιοχή τομής μεταξύ των κουτιών A και B διαιρούμενη με την περιοχή ένωσης μεταξύ των κουτιών A και B (σχέση 3.1):

$$IoU = \frac{A \cap B}{A \cup B} \quad (3.1)$$

Εάν ο εντοπισμός είναι ακριβής και τα bounding boxes A και B είναι πλήρως επικαλυμμένα,  $IoU = 1$ . Αν το αντικείμενο δεν μπορεί να εντοπιστεί καθόλου και τα bounding boxes A και B δεν επικαλύπτονται καθόλου,  $IoU = 0$ . Αν εντοπιστεί μόνο ένα μέρος, τα bounding boxes A και B επικαλύπτονται και  $IoU \in (0, 1)$ . Συνήθως, αν  $IoU \geq 0,5$ , ο εντοπισμός θεωρείται ακριβής. Αυτό το πρότυπο μπορεί επίσης να αναπροσαρμοστεί ανάλογα με συγκεκριμένα σενάρια τα οποία απαιτούν μεγαλύτερη ή μικρότερη ακρίβεια στις προβλέψεις.



Σχήμα 3.5 Αναπαράσταση του IoU

### Ανάκληση και ακρίβεια (precision and recall) - Mean Average Precision (mAP)

Ο μέσος της μέσης Ακρίβειας (mAP) είναι μια μετρική που χρησιμοποιείται για την αξιολόγηση μοντέλων ανίχνευσης αντικειμένων, τα οποία βασίζονται σε CNN. Ο μέσος όρος των τιμών μέσης ακρίβειας (AP) για κάθε κλάση υπολογίζεται σε τιμές ανάκλησης από 0 έως 1.

Ο τύπος mAP βασίζεται στις ακόλουθες επιμέρους μετρικές: Μήτρα σύγχυσης (Confusion Matrix), Τομή επί της Ένωσης (IoU), Ανάκληση (Recall), Ακρίβεια (Precision).

Κατά την ανίχνευση αντικειμένων, εάν υπάρχουν πολλά αντικείμενα στην εικόνα εισόδου, μπορεί να υπάρχουν πολλά πλαίσια οριοθέτησης. Σε αυτή την περίπτωση, ορισμένα πλαίσια μπορεί να εντοπίζουν με ακρίβεια ένα αντικείμενο, ενώ ορισμένα άλλα δεν το κάνουν, ή κάποια κουτιά μπορεί ακόμη και να εντοπίσουν το λάθος αντικείμενο. Εάν ένα κουτί εντοπίσει το αντικείμενο, π.χ.  $IoU \geq 0,5$ , και η ταξινόμηση είναι σωστή, ο εντοπισμός του αντικειμένου θεωρείται ακριβής. Η ακρίβεια της ανίχνευσης όλων των αντικειμένων στο σύνολο δοκιμών της εικόνας μετράτε συνήθως με τη το μέσο της μέσης ακρίβειας (mAP).

### Μήτρα σύγχυσης

Για να δημιουργήσουμε έναν πίνακα σύγχυσης, χρειαζόμαστε τέσσερα χαρακτηριστικά:

- Αληθινά θετικά ("True Positives", TP): Το μοντέλο προέβλεψε μια ετικέτα και ταιριάζει σωστά σύμφωνα με τη θεμελιώδη αλήθεια.



- Αληθινά αρνητικά ("True Negatives", TN): Το μοντέλο δεν προβλέπει την ετικέτα και δεν αποτελεί μέρος της θεμελιώδους αλήθειας.
- False Positives ("False Positives", FP): Το μοντέλο προέβλεψε μια ετικέτα, αλλά δεν αποτελεί μέρος της θεμελιώδους αλήθειας.
- Ψευδώς αρνητικά ("False Negatives", FN): Το μοντέλο δεν προβλέπει μια ετικέτα, αλλά αποτελεί μέρος της θεμελιώδους αλήθειας.

**Ανάκληση (Recall):** Μεταξύ των επιλεγμένων  $N$  δειγμάτων, η ανάκληση είναι η αναλογία των  $k$  αληθώς θετικών δειγμάτων προς το σύνολο των  $M$  θετικών δειγμάτων:

$$\text{Recall} = k/M = TP/(TP + FN) \quad (3.2)$$

**Ακρίβεια (Precision):** Μεταξύ των επιλεγμένων  $N$  δειγμάτων, η ακρίβεια είναι η αναλογία των  $k$  αληθών θετικών δειγμάτων:

$$\text{Precision} = k/N = TP/(TP + FP) \quad (3.3)$$

		Actual	
		Positive	Negative
Predicted	Positive	True Positive	False Positive
	Negative	False Negative	True Negative

Σχήμα 3.6 Ο πίνακας σύγχυσης (confusion matrix)

Για τον υπολογισμό της mAP θα πρέπει να δημιουργήσουμε και την καμπύλη ακρίβειας-ανάκλησης

Καμπύλη ακρίβειας-ανάκλησης: Για κάθε κλάση, υπολογίζετε τις τιμές ακρίβειας και ανάκλησης σε διαφορετικά όρια εμπιστοσύνης της IoU, συνήθως στις τιμές 0.5 και 0.95.

Μέση τιμή ακρίβειας (AP): Η καμπύλη ακρίβειας-ανάκλησης χρησιμοποιείται για τον υπολογισμό της μέσης ακρίβειας για κάθε κλάση. Η AP είναι η περιοχή κάτω από την καμπύλη ακρίβειας-ανάκλησης και ποσοτικοποιεί την ποιότητα της ανίχνευσης αντικειμένων για τη συγκεκριμένη κλάση. Η AP υπολογίζεται με τον υπολογισμό της ακρίβειας σε κάθε μοναδική τιμή ανάκλησης και στη συνέχεια με τον υπολογισμό του μέσου όρου τους.

Μέσος της μέσης τιμή ακρίβειας (mAP): Τέλος, οι τιμές AP σε όλες τις κλάσεις υπολογίζονται κατά μέσο όρο για να προκύψει η μέση ακρίβεια. Η mAP παρέχει ένα συνολικό μέτρο απόδοσης για την ανίχνευση αντικειμένων.

Ακολουθεί μια βήμα προς βήμα ανάλυση του υπολογισμού του mAP:

1. Υπολογίστε το IoU μεταξύ κάθε προβλεπόμενου οριοθετημένου πλαισίου και του αντίστοιχου οριοθετημένου πλαισίου της θεμελιώδους αλήθειας.
2. Ταξινομήστε τα προβλεπόμενα πλαίσια οριοθέτησης σε φθίνουσα σειρά με βάση τις βαθμολογίες εμπιστοσύνης τους.
3. Ξεκινήστε με μια κενή λίστα αληθώς θετικών (TP) και ψευδώς θετικών (FP) για κάθε κλάση.
4. Για κάθε προβλεπόμενο πλαίσιο οριοθέτησης:
  - a. Εάν το IoU με οποιοδήποτε πλαίσιο βασικής αλήθειας υπερβαίνει ένα προκαθορισμένο όριο (π.χ. 0,5), θεωρήστε το ως αληθώς θετικό (TP) και αφαιρέστε το αντίστοιχο πλαίσιο βασικής αλήθειας από την αξιολόγηση.
  - b. Εάν το IoU είναι κάτω από το κατώφλι, θεωρήστε το ψευδώς θετικό (FP).
5. Υπολογίστε την ακρίβεια και την ανάκληση σε διαφορετικά κατώφλια εμπιστοσύνης με βάση τα συσσωρευμένα αληθώς θετικά (TP) και ψευδώς θετικά (FP) για κάθε κλάση.
6. Κατασκευάστε την καμπύλη ακρίβειας-ανάκλησης για κάθε κλάση.
7. Υπολογίστε τη μέση ακρίβεια (AP) για κάθε κλάση υπολογίζοντας την περιοχή κάτω από την καμπύλη ακρίβειας-ανάκλησης.
8. Υπολογίζουμε τον μέσο όρο των τιμών AP σε όλες τις κλάσεις για να προκύψει η μέση ακρίβεια (mAP).

Αξίζει να σημειωθεί ότι η ακριβής υλοποίηση και οι επιλογές παραμέτρων μπορεί να διαφέρουν ανάλογα με το συγκεκριμένο framework ή τη βιβλιοθήκη ανίχνευσης αντικειμένων που χρησιμοποιείται. Επιπλέον, ορισμένες παραλλαγές του mAP ενδέχεται να χρησιμοποιούν διαφορετικές στρατηγικές υπολογισμού του μέσου όρου ή για την καμπύλη ακρίβειας-ανάκλησης.

### **F1 Score - ο αρμονικός μέσος όρος της ακρίβειας και της ανάκλησης**

Η βαθμολογία F1 είναι μια μετρική που χρησιμοποιείται για την αξιολόγηση της ποιότητας των μοντέλων ταξινόμησης στη μηχανική μάθηση. Υπολογίζεται ως ο αρμονικός μέσος όρος της ακρίβειας και της ανάκλησης [24]. Ο τύπος για τον υπολογισμό της βαθμολογίας F1 είναι ο εξής:

$$F1 \text{ Score} = \frac{2 * (Precision * Recall)}{Precision + Recall} \quad (3.4)$$

Η βαθμολογία F1 είναι καλύτερη από την ακρίβεια και την ανάκληση εάν χρησιμοποιούταν η κάθε μετρική μεμονωμένα, επειδή εξισορροπεί και τις δύο μετρικές και δίνει μια ενιαία βαθμολογία που αντικατοπτρίζει τη συνολική απόδοση του μοντέλου. Μια υψηλή βαθμολογία F1 σημαίνει ότι το μοντέλο έχει τόσο υψηλή ακρίβεια όσο και υψηλή ανάκληση, πράγμα που σημαίνει ότι κάνει λίγα ψευδώς θετικά και λίγα ψευδώς αρνητικά αποτελέσματα. Μια χαμηλή βαθμολογία F1 σημαίνει ότι είτε η ακρίβεια είτε η ανάκληση είτε και τα δύο είναι χαμηλά, πράγμα που σημαίνει ότι το μοντέλο κάνει πολλά σφάλματα.

### 3.3.2 Το μοντέλο YOLO

Το You Only Look Once (YOLO) είναι ένας πρωτοποριακός αλγόριθμος ανίχνευσης ενός σταδίου, ενοποιώντας την ταξινόμηση και τον εντοπισμό αντικειμένων σε ένα νευρωνικό δίκτυο για να επιτύχει ολοκληρωμένη ανίχνευση αντικειμένων. Η [25] αρχική υλοποίηση του μοντέλου έγινε το 2015 και από τότε έχουν γίνει πάνω από 20 τροποποιήσεις στον τρόπο λειτουργίας σε διάφορα στρώματα του μοντέλου και σε τεχνικές βελτιστοποίησης για την εξαγωγή καλύτερων αποτελεσμάτων. Στην συγκεκριμένη διπλωματική εργασία έχουν χρησιμοποιηθεί οι εκδόσεις YOLOv2 (YOLO 9000), YOLOv5 και YOLOv7.

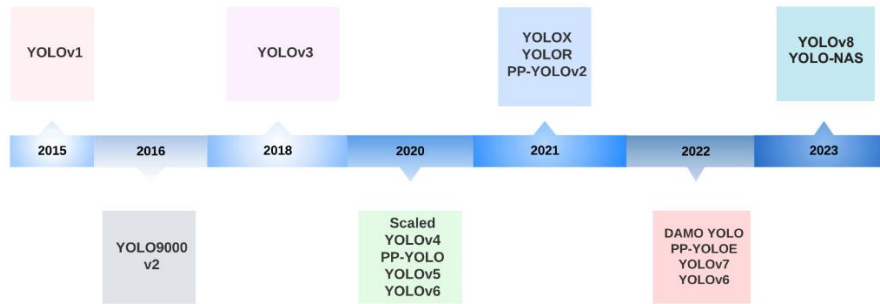
Όλες οι παραλλαγές του μοντέλου YOLO βασίζονται στην ενοποιημένη τεχνική ανίχνευσης, η οποία ομαδοποιεί τα διάφορα στοιχεία της ανίχνευσης αντικειμένων σε μια ενιαία ροή τροφοδοσίας του νευρωνικού δικτύου. Το μοντέλο διαιρεί μια εισερχόμενη εικόνα σε πολυάριθμα πλέγματα και υπολογίζει την πιθανότητα ένα αντικείμενο να βρίσκεται μέσα σε αυτό το πλέγμα. Για κάθε κελί πλέγματος (kernel) το μοντέλο προβλέπει έναν σταθερό αριθμό bounding boxes και τα αντίστοιχα σκορ εμπιστοσύνης, τα οποία υποδεικνύουν πόσο πιθανό είναι το κουτί να περιέχει ένα αντικείμενο και πόσο καλά ανταποκρίνεται στο αντικείμενο. Το μοντέλο προβλέπει επίσης τις πιθανότητες κατηγορίας για κάθε κελί πλέγματος, οι οποίες δείχνουν την πιθανότητα παρουσίας κάθε κατηγορίας αντικειμένου στο κελί.

Η τελική έξοδος του μοντέλου είναι ένας τένσορας που περιέχει όλες τις προβλέψεις του bounding box, τις βαθμολογίες εμπιστοσύνης και τις πιθανότητες κλάσης για κάθε κελί πλέγματος. Στη συνέχεια, το μοντέλο εφαρμόζει ένα κατώτατο όριο για να φιλτράρει τις προβλέψεις χαμηλής εμπιστοσύνης και εκτελεί μη μέγιστη καταστολή (NMS) για την εξάλειψη των επικαλυπτόμενων κουτιών. Τα εναπομείναντα κουτιά είναι οι τελικές προβλέψεις του μοντέλου.

Ως εκ τούτου, οι επιδόσεις της οικογένειας YOLO είναι πολλά υποσχόμενες. Η πρώτη έκδοση μπορεί να επιτύχει 45 fps (frames per second) σε ορισμένες GPU, γεγονός που μπορεί να ικανοποιήσει τις απαιτήσεις των εφαρμογών πραγματικού χρόνου. Οι νεότερες [25] εκδόσεις του YOLO είναι ικανές με συγκεκριμένες παραμετροποιήσεις να φτάσουν και τα 120 fps διατηρώντας την ακρίβεια του μοντέλου.

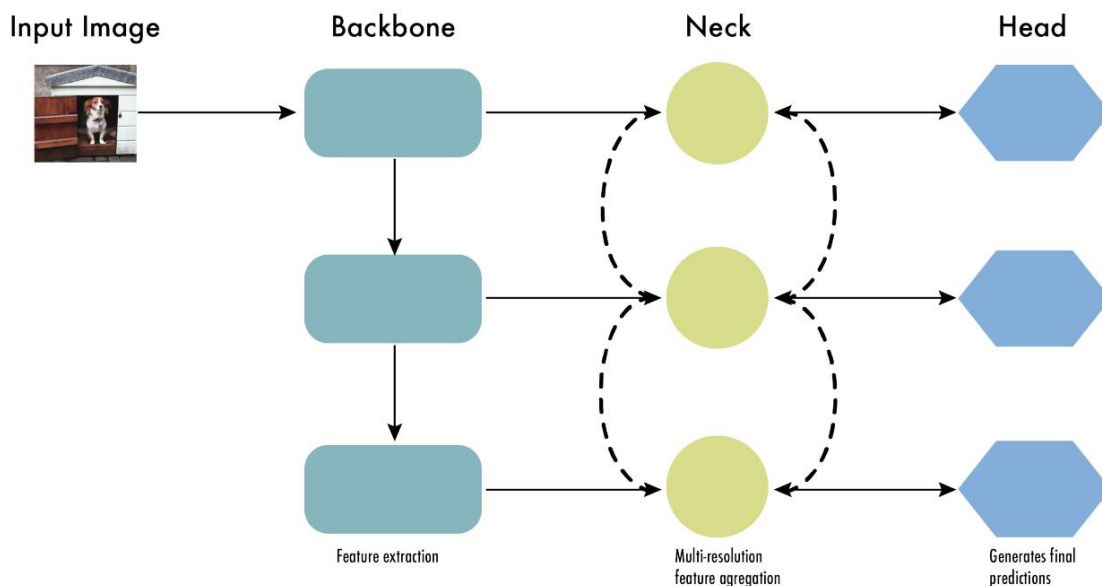
### 3.3.3 Αρχιτεκτονική των μοντέλων ανίχνευσης αντικειμένων – Αρχιτεκτονική του YOLOv2, YOLOv5 και YOLOv7

Η αρχιτεκτονική των ανιχνευτών αντικειμένων άρχισε να περιγράφεται σε τρία μέρη: τη ραχοκοκαλιά (backbone), τον λαιμό (neck) και το κεφάλι (head). Στο σχήμα 10 παρουσιάζεται ένα διάγραμμα υψηλού επιπέδου αρχιτεκτονικής [25].



Σχήμα 3.7 Η εξέλιξη και οι εκδόσεις της οικογένειας YOLO

**Backbone:** είναι υπεύθυνη για την εξαγωγή χρήσιμων χαρακτηριστικών από την εικόνα εισόδου. Πρόκειται συνήθως για ένα συνελκτικό νευρωνικό δίκτυο (CNN) που έχει εκπαιδευτεί σε μια εργασία ταξινόμησης εικόνων μεγάλης κλίμακας, όπως το ImageNet. Η ραχοκοκαλιά συλλαμβάνει ιεραρχικά χαρακτηριστικά σε διαφορετικές κλίμακες, με χαρακτηριστικά χαμηλότερου επιπέδου (π.χ. ακμές και υφές) που εξάγονται στα προηγούμενα στρώματα και χαρακτηριστικά υψηλότερου επιπέδου (π.χ. μέρη αντικειμένων και σημασιολογικές πληροφορίες) που εξάγονται στα βαθύτερα στρώματα.



Σχήμα 3.8 Η Αρχιτεκτονική των Object Detectors σύμφωνα με το [25]

**Neck:** είναι ένα ενδιάμεσο στοιχείο που συνδέει το backbone layer με το head. Συγκεντρώνει και βελτιώνει τα χαρακτηριστικά που εξάγει το backbone, εστιάζοντας συχνά στην ενίσχυση των χωρικών και σημασιολογικών πληροφοριών σε διαφορετικές κλίμακες. Το στρώμα του neck μπορεί να περιλαμβάνει πρόσθετα επίπεδα συνελκτικής ανάλυσης, δίκτυα πυραμίδας χαρακτηριστικών (FPN) [26] ή άλλους μηχανισμούς για να βελτιώσουν την αναπαράσταση των χαρακτηριστικών.

**Head:** Η κεφαλή είναι το τελικό συστατικό ενός ανιχνευτή αντικειμένων- είναι υπεύθυνη για την πραγματοποίηση προβλέψεων με βάση τα χαρακτηριστικά που παρέχονται από τη σπονδυλική στήλη και το λαιμό. Συνήθως αποτελείται από ένα ή περισσότερα υποδίκτυα ειδικών εργασιών που εκτελούν ταξινόμηση, τον εντοπισμό και, πιο πρόσφατα, την κατάτμηση στιγμιότυπων και την εκτίμηση της στάσης. Το κεφάλι επεξεργάζεται τα χαρακτηριστικά που παρέχει ο λαιμός, δημιουργώντας προβλέψεις για κάθε υποψήφιο αντικείμενο. Στο τέλος, ένα βήμα post processing, όπως η non maximum suppression (NMS), φιλτράρει τις επικαλυπτόμενες προβλέψεις και διατηρεί μόνο τις πιο σίγουρες ανιχνεύσεις.

Έχοντας υπόψιν το περίγραμμα της αρχιτεκτονικής backbone neck head μπορούμε να περιγράψουμε την αρχιτεκτονική των μοντέλων YOLO v2, v5 και v7.

- YOLOv2

Η αρχιτεκτονική κορμού που χρησιμοποιείται από το YOLOv2 ονομάζεται Darknet-19 και περιέχει 19 επίπεδα συνελκτικής ανάλυσης και πέντε επίπεδα max-pooling. Παρόμοια με την αρχιτεκτονική του YOLOv1, είναι εμπνευσμένη από το Network in Network χρησιμοποιώντας  $1 \times 1$  συνελίξεις μεταξύ των  $3 \times 3$  συνελίξεων για να μειωθεί ο αριθμός των παραμέτρων. Επιπλέον, όπως αναφέρθηκε παραπάνω, χρησιμοποιούν ομαλοποίηση παρτίδας για να κανονικοποιήσουν και να βοηθήσουν τη σύγκλιση. Τα στρώματα φαίνονται στον παρακάτω πίνακα

Πίνακας 1 Στρώματα του YOLO v2

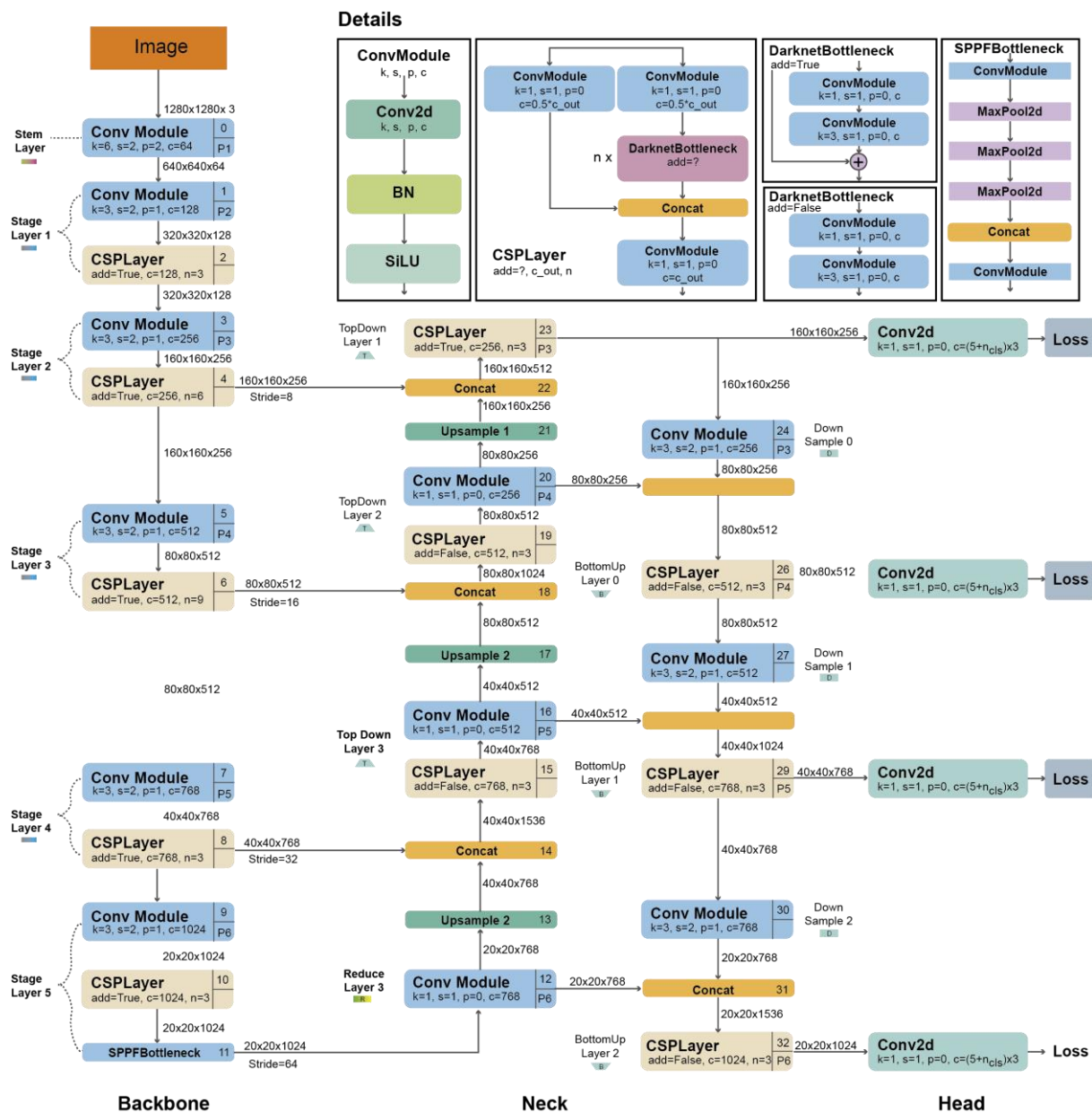
Num	Type	Filters	Size/Stride	Output
1	Conv/BN	32	$3 \times 3 / 1$	$416 \times 416 \times 32$
2	Max Pool		$2 \times 2 / 2$	$208 \times 208 \times 32$
3	Conv/BN	64	$3 \times 3 / 1$	$208 \times 208 \times 64$
4	Max Pool		$2 \times 2 / 2$	$104 \times 104 \times 64$
5	Conv/BN	128	$3 \times 3 / 1$	$104 \times 104 \times 128$
6	Conv/BN	64	$1 \times 1 / 1$	$104 \times 104 \times 64$
7	Conv/BN	128	$3 \times 3 / 1$	$104 \times 104 \times 128$
8	Max Pool		$2 \times 2 / 2$	$52 \times 52 \times 128$
9	Conv/BN	256	$3 \times 3 / 1$	$52 \times 52 \times 256$
10	Conv/BN	128	$1 \times 1 / 1$	$52 \times 52 \times 128$
11	Conv/BN	256	$3 \times 3 / 1$	$52 \times 52 \times 256$
12	Max Pool		$2 \times 2 / 2$	$52 \times 52 \times 256$
13	Conv/BN	512	$3 \times 3 / 1$	$26 \times 26 \times 512$
14	Conv/BN	256	$1 \times 1 / 1$	$26 \times 26 \times 256$
15	Conv/BN	512	$3 \times 3 / 1$	$26 \times 26 \times 512$
16	Conv/BN	256	$1 \times 1 / 1$	$26 \times 26 \times 256$
17	Conv/BN	512	$3 \times 3 / 1$	$26 \times 26 \times 512$
18	Max Pool		$2 \times 2 / 2$	$13 \times 13 \times 512$
19	Conv/BN	1024	$3 \times 3 / 1$	$13 \times 13 \times 1024$
20	Conv/BN	512	$1 \times 1 / 1$	$13 \times 13 \times 512$
21	Conv/BN	1024	$3 \times 3 / 1$	$13 \times 13 \times 1024$
22	Conv/BN	512	$1 \times 1 / 1$	$13 \times 13 \times 512$
23	Conv/BN	1024	$3 \times 3 / 1$	$13 \times 13 \times 1024$
24	Conv/BN	1024	$3 \times 3 / 1$	$13 \times 13 \times 1024$
25	Conv/BN	1024	$3 \times 3 / 1$	$13 \times 13 \times 1024$
26	Reorg layer 17			$13 \times 13 \times 2048$
27	Concat 25 and 26			$13 \times 13 \times 3072$
28	Conv/BN	1024	$3 \times 3 / 1$	$13 \times 13 \times 1024$
29	Conv	125	$1 \times 1 / 1$	$13 \times 13 \times 125$

- YOLOv5

Το backbone είναι ένα τροποποιημένο CSPDarknet53 [27] το οποίο ξεκινά με ένα στρώμα Stem, ένα στρώμα συνέλιξης με μεγάλο μέγεθος παραθύρου για τη μείωση της μνήμης και του υπολογιστικού κόστους και ακολουθείται από στρώματα συνέλιξης που εξάγουν τα σχετικά χαρακτηριστικά από την εικόνα εισόδου. Το στρώμα SPPF (spatial pyramid pooling fast) και τα επόμενα στρώματα συνέλιξης επεξεργάζονται τα χαρακτηριστικά σε διάφορες κλίμακες, ενώ τα στρώματα upsample αυξάνουν την ανάλυση των χαρτών χαρακτηριστικών.

Το στρώμα SPPF αποσκοπεί στην επιτάχυνση του υπολογισμού του δικτύου συγκεντρώνοντας χαρακτηριστικά διαφορετικών κλιμάκων σε έναν χάρτη χαρακτηριστικών σταθερού μεγέθους. Κάθε συνέλιξη ακολουθείται από κανονικοποίηση δέσμης (BN) και ενεργοποίηση SiLU [28]. Το neck χρησιμοποιεί SPPF και ένα τροποποιημένο CSP-PAN, ενώ το κεφάλι μοιάζει με το YOLOv3 [25].

Το YOLOv5 παρέχει πέντε κλιμακωτές εκδόσεις: YOLOv5n (nano), YOLOv5s (small), YOLOv5m (medium), YOLOv5l (large) και YOLOv5x (extra large), όπου το πλάτος και το βάθος των μονάδων συνέλιξης ποικίλλουν για να ταιριάζουν σε συγκεκριμένες εφαρμογές και απαιτήσεις υλικού. Για παράδειγμα, τα YOLOv5n και YOLOv5s είναι ελαφριά μοντέλα που απευθύνονται σε συσκευές με χαμηλούς πόρους όπως κινητά τηλέφωνα, ενώ το YOLOv5x είναι βελτιστοποιημένο για υψηλές επιδόσεις, αν και εις βάρος της ταχύτητας.



Σχήμα 3.9 Σχηματική αναπαράσταση των backbone, neck, head του YOLO v5

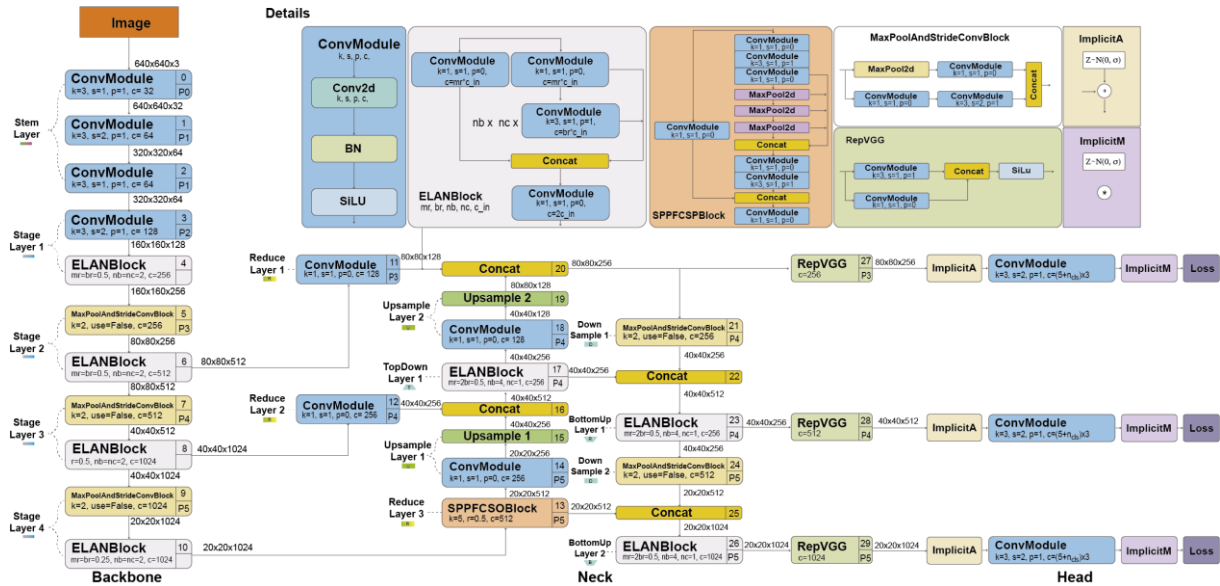
- YOLOv7

Η αρχιτεκτονική του v7 είναι παρόμοια με αυτή του v5 με 2 καινούργιες προσθήκες και διάφορες τεχνικές για την αύξηση της ταχύτητας εκπαίδευσης και τον χρόνο εξαγωγής μια πρόβλεψης (inference time):

- Διευρυμένο αποδοτικό δίκτυο συνάθροισης επιπέδων (E-ELAN). Το ELAN είναι μια στρατηγική που επιτρέπει ένα βαθύ μοντέλο να μαθαίνει και να συγκλίνει πιο αποτελεσματικά ελέγχοντας τη συντομότερη μεγαλύτερη gradient path. Η αρχιτεκτονική του YOLOv7 πρότεινε το E-ELAN που λειτουργεί για μοντέλα με απεριόριστα στοιβαγμένα υπολογιστικά μπλοκ. Το E-ELAN συνδυάζει τα χαρακτηριστικά διαφορετικών ομάδων με ανακάτεμα και συγχώνευση των κοινών χαρακτηριστικών για να ενισχύσει τη μάθηση του δικτύου χωρίς να καταστρέφει το αρχικό gradient path.

### Κεφάλαιο 3

- Κλιμάκωση μοντέλου για μοντέλα βασισμένα στη συνένωση. Η κλιμάκωση παράγει μοντέλα διαφορετικού μεγέθους προσαρμόζοντας ορισμένα χαρακτηριστικά του μοντέλου. Η αρχιτεκτονική του YOLOv7 είναι μια αρχιτεκτονική βασισμένη στη συνένωση, στην οποία τα τυπικά τεχνικές κλιμάκωσης, όπως η κλιμάκωση βάθους, προκαλούν μια αλλαγή αναλογίας μεταξύ του καναλιού εισόδου και του καναλιού εξόδου ενός μεταβατικού στρώματος, η οποία, με τη σειρά της, οδηγεί σε μείωση της χρήσης υλικού του μοντέλου.



Σχήμα 3.10 Η αρχιτεκτονική του YOLOv7 σύμφωνα με [29]

Σύμφωνα με τους δημιουργούς [30] του η έκδοση 7 της οικογένειας YOLO έχει την καλύτερη απόδοση σε όλες τις κλίμακες από τα 5 fps μέχρι και τα 160, καθώς εμφανίζει και την μεγαλύτερη μέση ακρίβεια στα benchmarks. Εκτός από την ακρίβεια η έκδοση 7 έχει και τους καλύτερους χρόνους εκτέλεσης μιας πρόβλεψης.



## Κεφάλαιο 4ο: Ανάπτυξη εφαρμογής για κινητά τηλέφωνα

### 4.1 Εισαγωγή στην ανάπτυξη εφαρμογών για κινητά τηλέφωνα

Η ανάπτυξη κινητών συσκευών είναι μια συμβολή της δημιουργικότητας, της τεχνικής ικανότητας και της τήρησης των βιομηχανικών προτύπων. Αρχικά θα αναφερθούμε στις γενικές αρχές που θα πρέπει να τηρεί μια εφαρμογή για κινητά τηλέφωνα, ανεξαρτήτως της πλατφόρμας της οποία χρησιμοποιεί Android, iOS ή cross-platform . Έπειτα θα εμβαθύνουμε συγκεκριμένα στις εφαρμογές για iOS, η πλατφόρμα η οποία επιλέχτηκε για την εκπόνηση αυτής της διπλωματικής και τέλος θα περιγράψουμε μερικές από τις αρχιτεκτονικές και τις βέλτιστες πρακτικές για την ανάπτυξη εφαρμογών.

Ο ακρογωνιαίος λίθος της αποτελεσματικής ανάπτυξης εφαρμογών για κινητά τηλέφωνα είναι, πρώτον, η καλή κατανόηση των γλωσσών προγραμματισμού. Για παράδειγμα, ενώ η Swift και η Objective-C είναι επιβεβλημένες για το iOS, η Java και η Kotlin έχουν προτεραιότητα στο Android. Ομοίως, η JavaScript ή η Dart μπορεί να χρησιμοποιηθεί για λύσεις cross-platform με τη χρήση React Native ή Flutter, αντίστοιχα. Η επιλογή της σωστής γλώσσας εξαρτάται συχνά από την πλατφόρμα και τις απαιτήσεις της εφαρμογής.

Η επιλογή των εργαλείων ανάπτυξης είναι ένας άλλος κρίσιμος παράγοντας. Οι προγραμματιστές iOS κλίνουν προς το Xcode ή το AppCode, ενώ το Android Studio είναι το προτιμώμενο για τους προγραμματιστές Android. Για την ανάπτυξη cross-platform, frameworks όπως το React Native και το Flutter διαθέτουν τα ειδικά IDE και εργαλεία τους. Η εξοικείωση με τα εργαλεία της επιλεγμένης πλατφόρμας, τις ενημερώσεις και τις επεκτάσεις τους είναι ζωτικής σημασίας για μια αποβλημάτιστη εμπειρία ανάπτυξης.

Η αρχιτεκτονική και τα πρότυπα σχεδίασης παίζουν σημαντικό ρόλο στη δημιουργία μιας επεκτάσιμης και συντηρήσιμης εφαρμογής. Αρχιτεκτονικές όπως το MVC (Model-View-Controller), το MVVM (Model-View-ViewModel) και πρότυπα όπως ο Singleton και ο Observer χρησιμοποιούνται συνήθως σε όλη την ανάπτυξη εφαρμογών για κινητά. Αυτά τα μοτίβα όχι μόνο διευκολύνουν τον καθαρό, δομημένο κώδικα αλλά και ενισχύουν τη συνεργασία μεταξύ των προγραμματιστών δημιουργώντας σταθερά μοτίβα για την ανάπτυξη της εφαρμογής. Τα πρότυπα κώδικα και οι συμβάσεις είναι θεμελιώδους σημασίας για κάθε διαδικασία ανάπτυξης. Η τήρηση αυτών των συμβάσεων - όπως το indentation, ο σχολιασμός του κώδικα και η περιγραφική ονομασία - βελτιώνει την αναγνωσιμότητα του κώδικα, μειώνει τα λάθη και προωθεί την αποτελεσματική συνεργασία μεταξύ των μελών της ομάδας σε συνεργατικά project.

Η βελτιστοποίηση των επιδόσεων είναι ένα άλλο κρίσιμο ζήτημα στην ανάπτυξη εφαρμογών για κινητά. Τεχνικές όπως το lazy loading, η ελαχιστοποίηση των frameworks, η μεταφόρτωση πολύπλοκων εργασιών σε νήματα παρασκήνιου και η αποδοτική χρήση της μπαταρίας συμβάλλουν σε μια πιο εύελκτη, φιλική προς τον χρήστη εφαρμογή. Η διαχείριση της μνήμης είναι ένας τομέας όπου η απόδοση της εφαρμογής μπορεί να επηρεαστεί σημαντικά. Ανεξάρτητα από την πλατφόρμα, η αποδοτική χρήση της μνήμης εξασφαλίζει την ομαλότερη λειτουργία των εφαρμογών. Για παράδειγμα, ενώ το iOS διαθέτει αυτόματη καταμέτρηση αναφορών (ARC), το Android βασίζεται στη συλλογή σκουπιδιών (Garbage Collection). Η κατανόηση αυτών των συστημάτων είναι ζωτικής σημασίας για την αποφυγή διαρροών μνήμης και την απόδοσης της εφαρμογής.

Η ασφάλεια είναι ένας τομέας που χρήζει σοβαρής προσοχής στην ανάπτυξη εφαρμογών για κινητά. Οι ασφαλείς πρακτικές κωδικοποίησης, η κρυπτογράφηση ευαίσθητων δεδομένων, η επικύρωση εισόδου,

η ασφαλής επικοινωνία μέσω δικτύου και η κατάλληλη διαχείριση δικαιωμάτων είναι μερικές μόνο από τις βέλτιστες πρακτικές που πρέπει να τηρούνται.

Η προσβασιμότητα και η διεθνοποίηση είναι ζωτικής σημασίας για την επέκταση της βάσης χρηστών μιας εφαρμογής. Με την ενσωμάτωση χαρακτηριστικών προσβασιμότητας, οι προγραμματιστές μπορούν να διασφαλίσουν ότι οι εφαρμογές τους ανταποκρίνονται σε χρήστες με διάφορες αναπηρίες. Επιπλέον, το localization - η προσαρμογή μιας εφαρμογής σε διάφορες γλώσσες και πολιτισμικές αποχρώσεις (π.χ. γραφή από δεξιά προς τα αριστερά στις αραβικές χώρες) - διευρύνει την παγκόσμια εμβέλεια της εφαρμογής.

Εν κατακλείδι, η τέχνη της ανάπτυξης εφαρμογών για κινητά περιλαμβάνει μια λεπτή ισορροπία μεταξύ της τεχνικής οξυδέρκειας, της κατανόησης των βέλτιστων πρακτικών και της κατανόησης προς τον τελικό χρήστη.

## 4.2 Εισαγωγή στην ανάπτυξη εφαρμογών για την πλατφόρμα του iOS

### 4.2.1 Το λειτουργικό σύστημα iOS και το iOS SDK

Το iOS είναι ένα λειτουργικό σύστημα για κινητά που δημιουργήθηκε και αναπτύχθηκε από την Apple Inc. αποκλειστικά για το hardware της. Είναι το λειτουργικό σύστημα που τροφοδοτεί πολλές από τις κινητές συσκευές της εταιρείας, όπως το iPhone, το iPad και το iPod Touch.

Η διεπαφή χρήστη (User Interface, UI) του iOS βασίζεται στον άμεσο χειρισμό, χρησιμοποιώντας χειρονομίες πολλαπλής αφής, όπως το σύρσιμο, το πάτημα, το τσίμπημα και το αντίστροφο τσίμπημα. Τα στοιχεία ελέγχου της διεπαφής περιλαμβάνουν ρυθμιστικά, διακόπτες και κουμπιά. Το iOS παρέχει διάφορα χαρακτηριστικά προσβασιμότητας για να βοηθήσει τους χρήστες με προβλήματα όρασης και ακοής. Το iOS χαρακτηρίζεται επίσης από μια ποικιλία αποκλειστικών εφαρμογών που αναπτύσσονται από την Apple, όπως το Safari, το Mail, οι Φωτογραφίες και τα Μηνύματα, μεταξύ πολλών άλλων. Περιλαμβάνει επίσης το App Store, όπου οι χρήστες μπορούν να ανακαλύψουν και να εγκαταστήσουν εφαρμογές τρίτων κατασκευαστών.

Για την ανάπτυξη εφαρμογών για iOS απαιτείται η χρήση του iOS SDK ή Software Development Kit, το οποίο είναι μια συλλογή εργαλείων, frameworks, γλωσσών και βιβλιοθηκών που παρέχονται από την Apple και χρησιμοποιούνται από τους προγραμματιστές για τη δημιουργία εφαρμογών για τις συσκευές iOS της Apple, οι οποίες περιλαμβάνουν το iPhone, το iPad και το iPod Touch.

Το iOS SDK περιλαμβάνει ένα ευρύ φάσμα στοιχείων [31]:

**Xcode:** είναι το ολοκληρωμένο περιβάλλον ανάπτυξης (IDE) που χρησιμοποιούν οι προγραμματιστές για τη δημιουργία εφαρμογών iOS. Παρέχει μια γραφική διεπαφή για τη σύνταξη κώδικα Swift (ή Objective-C), τη δοκιμή εφαρμογών και την αποσφαλμάτωση (Debugging). Περιλαμβάνει έναν σχεδιαστή διεπαφών τόσο για το UIKit όσο και για το SwiftUI, εργαλεία ανάλυσης επιδόσεων και πολλά άλλα.

**Swift** και Objective-C: Αυτές είναι οι κύριες γλώσσες που χρησιμοποιούνται για την ανάπτυξη iOS. Η Swift είναι μια πιο σύγχρονη γλώσσα, είναι type safe, πιο εύκολη στην ανάγνωση και διαθέτει σύγχρονα χαρακτηριστικά όπως τα optionals και τα closures (anonymous functions). Η Objective-C είναι μια παλαιότερη γλώσσα, η οποία αποτελεί υπερσύνολο της C, και ενώ χρησιμοποιείται λιγότερο συχνά για νέα έργα, πολλά υπάρχοντα έργα και βιβλιοθήκες είναι γραμμένα σε Objective-C. Αξίζει αν σημειωθεί ότι κώδικας Objective-C μπορεί να χρησιμοποιηθεί χωρίς περιορισμούς σε κώδικα Swift (backwards compatible)

**Βιβλιοθήκες και frameworks iOS:** περιέχουν προκατασκευασμένα κομμάτια κώδικα που εκτελούν κοινές εργασίες σε εφαρμογές iOS. Για παράδειγμα, το UIKit/SwiftUI για την κατασκευή διεπαφών χρήστη, το Core Data για τη διαχείριση persistent δεδομένων, το Core Animation για την απόδοση γραφικών και τις κινούμενες εικόνες, το Vision Framework για την χρήση AR/VR τεχνολογιών.

**Προσομοιωτής iOS:** είναι ένα εργαλείο που περιλαμβάνεται στο Xcode και επιτρέπει στους προγραμματιστές να προσομοιώσουν μια συσκευή iOS στον υπολογιστή τους. Αυτό τους επιτρέπει να δοκιμάζουν τις εφαρμογές τους χωρίς να χρειάζονται μια φυσική συσκευή.

**Interface Builder:** Πρόκειται για ένα γραφικό εργαλείο στο Xcode που επιτρέπει στους προγραμματιστές να δημιουργούν οπτικά το UI της εφαρμογής τους, αντί να γράφουν κώδικα γι' αυτήν.

**Instruments:** Πρόκειται για ένα εργαλείο μέτρησης και ανάλυσης επιδόσεων που συνοδεύει το Xcode. Παρέχει μια οπτική διεπαφή για την κατανόηση του τρόπου με τον οποίο χρησιμοποιούνται πόροι όπως η CPU, η μνήμη και το δίκτυο από την εφαρμογή.

**App Store Connect:** Πρόκειται για μια σουίτα διαδικτυακών εργαλείων για τη διαχείριση της εφαρμογής σας στο App Store. Μπορείτε να ανεβάζετε νέα builds και εκδόσεις, να διαχειρίζεστε τα μεταδεδομένα της εφαρμογής σας, να παρακολουθείτε τις πωλήσεις και τις λήψεις και πολλά άλλα.

#### 4.2.2 Τυπική δομή ενός project

Μόλις εγκατασταθεί το Xcode, το οποίο περιλαμβάνει το iOS SDK, το αρχικό βήμα στη ρύθμιση του έργου περιλαμβάνει τη διαμόρφωση των βασικών πληροφοριών για την εφαρμογή. Αυτό περιλαμβάνει τον καθορισμό του "Product Name", το οποίο αντιπροσωπεύει το όνομα της εφαρμογής, και τον προσδιορισμό της ομάδας "Team", η οποία θα αντιστοιχεί στην ομάδα Apple Developer στην οποία ανήκει ένας προγραμματιστής. Στην περίπτωση μεμονωμένων προγραμματιστών, αυτό το πεδίο περιέχει συνήθως τα δικά τους ονόματα. Απαιτείται επίσης ένα "Organization Identifier", το οποίο είναι συνήθως ένα string (συμβολοσειρά) τύπου reverse domain name (για παράδειγμα, com.example). Αυτό, σε συνδυασμό με το όνομα προϊόντος, σχηματίζει ένα μοναδικό "bundle identifier" για την εφαρμογή. Η επιλογή programming language είναι διαθέσιμη μεταξύ Swift ή Objective-C. Η Swift επιλέγεται συχνά λόγω της σύγχρονης, ασφαλούς και εκφραστικής φύσης της. Η επιλογή "User Interface" εξαρτάται από την έκδοση του Xcode, προσφέροντας μια επιλογή μεταξύ SwiftUI ή UIKit.

Με την πραγματοποίηση αυτών των ρυθμίσεων και την έναρξη του έργου, εμφανίζεται ένα προεπιλεγμένο σύνολο αρχείων και καταλόγων που σχηματίζουν τη δομή του project. Το "AppDelegate.swift" είναι μεταξύ αυτών των βασικών στοιχείων. Αποτελεί μέρος του κύκλου ζωής της εφαρμογής, επιφορτισμένο με την ευθύνη της λήψης συμβάντων του κύκλου ζωής. Ένα άλλο κρίσιμο αρχείο είναι το "SceneDelegate.swift", που εισήχθη στο iOS 13, το οποίο διαχειρίζεται το περιεχόμενο που εμφανίζεται στην οθόνη ελέγχοντας τις σκηνές των παραθύρων όταν αυτά είναι διαθέσιμα. Το "ViewController.swift" είναι η θέση όπου γράφεται ο κώδικας για τον έλεγχο του view. Αυτό περιλαμβάνει εργασίες όπως η φόρτωση των sub views, ο χειρισμός της αλληλεπίδρασης με τον χρήστη και η γενική συμπεριφορά του UI.

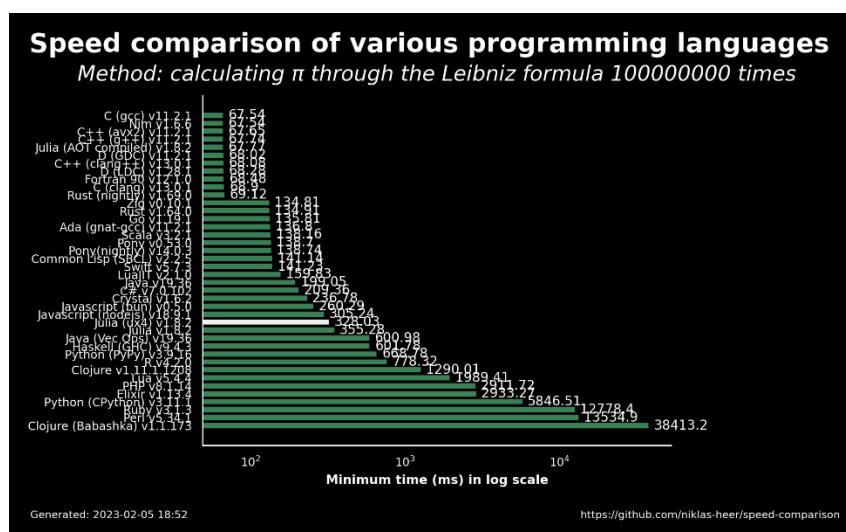
Το αρχείο "Main.storyboard" χρησιμοποιείται για το σχεδιασμό της διεπαφής χρήστη της εφαρμογής, ενεργοποιώντας μια λειτουργία drag-and-drop για στοιχεία UI πάνω σε σκηνές. Ο κατάλογος "Assets.xcassets" χρησιμεύει ως αποθετήριο για τα στατικά στοιχεία της εφαρμογής (εικόνες, χρώματα, γραμματοσειρές). Τέλος, το αρχείο "Info.plist" περιέχει μεταδεδομένα σχετικά με την εφαρμογή, συμπεριλαμβανομένων πληροφοριών όπως ο αριθμός έκδοσης και το όνομα εμφάνισης.

### 4.2.3 Η Γλώσσα προγραμματισμού Swift

Η Swift είναι μια στιβαρή και εύχρηστη γλώσσα προγραμματισμού που δημιουργήθηκε από την Apple για την κατασκευή εφαρμογών για iOS, Mac, Apple TV και Apple Watch. Έχει σχεδιαστεί για να παρέχει στους προγραμματιστές μεγαλύτερη ελευθερία από ποτέ. Η Swift είναι εύκολο στη χρήση και ανοιχτού κώδικα, οπότε όποιος έχει μια ιδέα μπορεί να δημιουργήσει να κάνει μια πρόταση προκειμένου να ενσωματωθεί στο project [32].

Ένα από τα σημαντικότερα πλεονεκτήματα της Swift είναι η καθαρή, συνοπτική και εκφραστική σύνταξή της, η οποία καθιστά τη γλώσσα πολύ ευανάγνωστη. Η δομή του κώδικα Swift είναι παρόμοια με τα φυσικά αγγλικά, γεγονός που μειώνει το γνωστικό φορτίο για τους αρχάριους προγραμματιστές και διευκολύνει την ανάγνωση και τη συγγραφή για τους έμπειρους προγραμματιστές. Η Swift ενσωματώνει επίσης πολλά σύγχρονα χαρακτηριστικά προγραμματισμού, όπως closures, generics, tuples, optionals και type inference. Αυτά τα χαρακτηριστικά οδηγούν σε λιγότερο κώδικα, λιγότερα σφάλματα και πιο εκφραστικές και ευέλικτες δομές κώδικα. Αυτή η αναγνωσιμότητα συμβάλλει στη συνολική φιλικότητα της Swift προς τον χρήστη και είναι ένας από τους κύριους λόγους που έχει γίνει τόσο δημοφιλής επιλογή για την ανάπτυξη εφαρμογών λογισμικού.

Επιπλέον, η Swift είναι γρήγορη. Οι επιδόσεις της συχνά ξεπερνούν τις επιδόσεις συγκρίσιμων γλωσσών σε τομείς όπως complex object sorting και string manipulation. Σύμφωνα με το [31] και [33], η Swift είναι έως και 2,6 φορές ταχύτερη από την Objective-C και 8,4 φορές ταχύτερη από την Python 2.7.



Σχήμα 4.1 Η κατάταξη της ταχύτητας των πιο διαδεδομένων γλωσσών προγραμματισμού σύμφωνα με το [33]

Εκτός από την αναγνωσιμότητά της, η Swift έχει σχεδιαστεί με μεγάλη έμφαση στην ασφάλεια. Περιλαμβάνει χαρακτηριστικά ασφαλείας, όπως τα optionals και το strict typing, που αποτρέπουν πολλά κοινά προγραμματιστικά λάθη. Το Automatic Reference Counting (ARC) χειρίζεται μεγάλο μέρος της διαχείρισης της μνήμης των εφαρμογών Swift, γεγονός που μειώνει σημαντικά κοινά προβλήματα όπως οι διαρροές μνήμης (memory leaks).

Η Swift έχει επίσης σημαντική υποστήριξη από την κοινότητα και τις εταιρείες. Ως γλώσσα ανοικτού κώδικα που αναπτύχθηκε από την Apple, η Swift ενημερώνεται, βελτιώνεται και συντηρείται συνεχώς από μία από τις κορυφαίες εταιρείες τεχνολογίας στον κόσμο. Η υποστήριξη αυτή επεκτείνεται σε μια πολύ ενεργή και χρήσιμη κοινότητα προγραμματιστών που συμβάλλουν συνεχώς στην ανάπτυξή της και προσφέρουν υποστήριξη σε νέους προγραμματιστές.

Μερικά παραδείγματα κώδικα για βασικές λειτουργίες όπως η αρχικοποίηση μιας δομής, η επέκταση μιας δομής με παραπάνω λειτουργίες, προσαρμογή σε πρωτόκολλα βάση του protocol oriented programming, ταξινόμηση και mapping όπως φαίνεται στα παρακάτω αποσπάσματα κώδικα.

```

struct Player {
    var name: String
    var highScore: Int = 0
    var history: [Int] = []

    init(_ name: String) {
        self.name = name
    }
}

var player = Player("Tomas")

```

Εικόνα 1 Κώδικας δημιουργίας δομής με μεταβλητές οι οποίες έχουν αρχικοποιηθεί απευθείας

```

extension Player {
    mutating func updateScore(_ newScore: Int) {
        history.append(newScore)
        if highScore < newScore {
            print("\(newScore)! A new high score for \(name)! 🎉")
            highScore = newScore
        }
    }
}

player.updateScore(50)

```

Εικόνα 2 Κώδικας που περιγράφει την λειτουργία των επεκτάσεων στην Swift, με μια συνάρτηση η οποία τροποποιεί την δομή (mutating).

```

let players = getPlayers()

// Sort players, with best high scores first
let ranked = players.sorted(by: { player1, player2 in
    player1.highScore > player2.highScore
})

// Create an array with only the players' names
let rankedNames = ranked.map { $0.name }
// ["Erin", "Rosana", "Tomas"]

```

Εικόνα 3 Εφαρμογή ταξινόμησης και μετατροπής με την χρήση των συναρτήσεων sort και map οι οποίες περιλαμβάνονται στο Foundation framework.

Τέλος, η εφαρμογή της Swift δεν περιορίζεται μόνο στην ανάπτυξη iOS ή macOS. Κερδίζει επίσης έδαφος στην ανάπτυξη server side applications με την χρήση του Vapor [34], ακόμη και σε τομείς όπως η μηχανική μάθηση με το έργο Swift for TensorFlow [35]. Αυτή η ποικιλομορφία στις περιπτώσεις χρήσης καθιστά τη Swift μια ευέλικτη γλώσσα, κατάλληλη για ένα ευρύ φάσμα εργασιών προγραμματισμού.

#### 4.2.4 Οδηγίες ανθρώπινης διεπαφής (human interface guidelines HIG)

Οι οδηγίες ανθρώπινης διεπαφής (human interface guidelines HIG) της Apple για το iOS είναι ένα σύνολο αρχών και οδηγιών σχεδιασμού για τη δημιουργία εφαρμογών και UI για συσκευές iOS, όπως το iPhone και το iPad. Οι οδηγίες ανθρώπινης διεπαφής καλύπτουν θέματα όπως ο σχεδιασμός διεπαφής, η πλοήγηση, η τυπογραφία και η προσβασιμότητα, με συγκεκριμένες προτάσεις για διαφορετικές συσκευές. Οι κατευθυντήριες γραμμές για την ανθρώπινη διεπαφή ενημερώνονται τακτικά και περιλαμβάνουν σχεδιαστικούς πόρους, όπως πρότυπα, εικονίδια και στοιχεία σχεδιασμού διεπαφής χρήστη, τα οποία οι προγραμματιστές μπορούν να χρησιμοποιήσουν στα γραφικά των εφαρμογών τους. Οι οδηγίες ανθρώπινης διεπαφής έχουν βαθιές ρίζες στις αρχές σχεδιασμού που η Apple βοήθησε να γίνουν πρωτοπόροι πριν από δεκαετίες και αποτελούν ένα ζωντανό έγγραφο που παρέχει τις πιο πρόσφατες οδηγίες σχεδιασμού για τις πλατφόρμες iOS. Οι βασικές αρχές που ακολουθεί το HIG [36] είναι οι παρακάτω:

1. Σαφήνεια: Η διεπαφή χρήστη πρέπει να καθιστά τις αλληλεπιδράσεις προφανείς και εύκολα κατανοητές. Το κείμενο θα πρέπει να είναι ευανάγνωστο σε κάθε μέγεθος, τα εικονίδια θα πρέπει να είναι ακριβή και σαφή, οι διακοσμήσεις θα πρέπει να είναι διακριτικές και η λειτουργικότητα θα πρέπει να είναι σαφής.
2. Παραπομπή: Το UI βοηθά τους χρήστες να κατανοήσουν και να αλληλεπιδράσουν με το περιεχόμενο, αλλά ποτέ δεν το ανταγωνίζεται. Η ρευστή κίνηση και μια καθαρή, όμορφη διεπαφή βοηθούν τους χρήστες να κατανοήσουν και να αλληλεπιδράσουν με το περιεχόμενο, αλλά το UI θα πρέπει να υποστηρίζει το περιεχόμενο και όχι να είναι το κύριο επίκεντρο.
3. Βάθος: Η στρατηγική χρήση των επιπέδων και της κίνησης μπορεί να ζωντανέψει μια εφαρμογή. Οι αλληλεπιδράσεις πρέπει να έχουν βάθος που μιμείται τον φυσικό κόσμο και να παρέχουν την αίσθηση του χώρου.
4. Συνέπεια: Μια συνεπής εφαρμογή εφαρμόζει οικεία πρότυπα χρησιμοποιώντας στοιχεία διεπαφής που παρέχονται από το σύστημα, γνωστά εικονίδια, τυποποιημένα στυλ κειμένου και ομοιόμορφη ορολογία. Η εφαρμογή θα πρέπει να ενσωματώνει αυτές τις παραδοχές, παρέχοντας παράλληλα μια ελκυστική εμπειρία.
5. Ανατροφοδότηση: Η ανατροφοδότηση αναγνωρίζει τις ενέργειες του χρήστη και δείχνει τα αποτελέσματα για να ενημερώνει τους ανθρώπους. Οι ενσωματωμένες εφαρμογές iOS παρέχουν κατανοητό feedback ως απάντηση σε κάθε ενέργεια του χρήστη.
6. Μεταφορές: Οι άνθρωποι μαθαίνουν πιο γρήγορα όταν τα εικονικά αντικείμενα και οι ενέργειες μιας εφαρμογής είναι μεταφορές για οικείες εμπειρίες. Η χρήση μεταφορών επιτρέπει στους χρήστες να μεταφέρουν τις γνώσεις και τις δεξιότητές τους από τον φυσικό κόσμο στον ψηφιακό κόσμο και μπορεί να είναι ιδιαίτερα χρήσιμη όταν σχεδιάζετε ένα νέο είδος προϊόντος ή λειτουργίας.
7. Έλεγχος από τον χρήστη: Οι χρήστες αισθάνονται πιο άνετα και έχουν τον έλεγχο όταν γνωρίζουν ότι μια ενέργεια είναι απλό να αντιστραφεί. Θα πρέπει να είναι σε θέση να ξεκινούν και να ελέγχουν τις ενέργειες, χωρίς η εφαρμογή να αναλαμβάνει ποτέ απροσδόκητα τον έλεγχο ή να επιβάλλει την πρόοδο.
8. Προσαρμοστικότητα και ευελιξία: Οι εφαρμογές θα πρέπει να στοχεύουν στην παροχή μιας απρόσκοπτης εμπειρίας σε όλες τις συσκευές της Apple, ενώ παράλληλα θα πρέπει να σέβονται τις μοναδικές πτυχές κάθε πλατφόρμας. Επιπλέον, οι εφαρμογές θα πρέπει να μπορούν να προσαρμόζονται σε διαφορετικά περιβάλλοντα, όπως διαφορετικές τοποθεσίες ή γλώσσες.

#### 4.2.5 Βασικά Frameworks που χρησιμοποιούνται στην ανάπτυξη εφαρμογών

Darwin: Το Darwin είναι το βασικό λειτουργικό σύστημα που τροφοδοτεί το iOS και άλλες πλατφόρμες της Apple. Βασίζεται στο UNIX και παρέχει υπηρεσίες χαμηλού επιπέδου, όπως ο πυρήνας (kernel),

σύστημα αρχείων, δικτύωση, ασφάλεια κλπ. Το Darwin είναι ως επί το πλείστον γραμμένο σε C και C++ και εκθέτει POSIX APIs για τους προγραμματιστές

**Core Foundation:** Το Core Foundation είναι ένα framework που παρέχει θεμελιώδεις τύπους δεδομένων και βοηθητικά προγράμματα για το iOS και άλλες πλατφόρμες της Apple. Είναι γραμμένο σε C και παρέχει μια γέφυρα μεταξύ των χαμηλού επιπέδου APIs του Darwin και των υψηλού επιπέδου APIs των Objective-C και Swift. Το Core Foundation περιλαμβάνει τύπους όπως strings, collections, dates, timers, streams κλπ.

**Foundation:** Το Foundation είναι ένα framework που παρέχει βασικούς τύπους δεδομένων και βοηθητικά προγράμματα για το iOS και άλλες πλατφόρμες της Apple. Είναι γραμμένο σε Objective-C και Swift και βασίζεται πάνω στο Core Foundation. Το Foundation περιλαμβάνει numbers, arrays, dictionaries, sets, URLs, data, files, threads, operations, notifications κλπ.

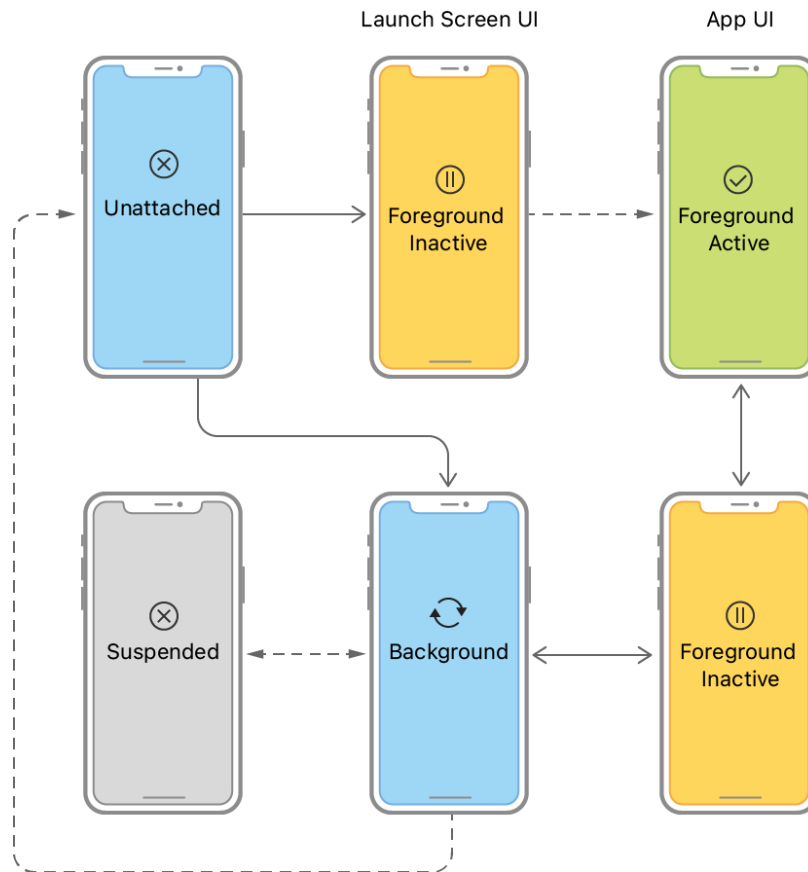
**UIKit:** Το UIKit είναι ένα framework που παρέχει τα βασικά αντικείμενα του UI που χρειάζεστε για τη δημιουργία εφαρμογών για iOS και tvOS. Είναι γραμμένο σε Objective-C και Swift και βασίζεται πάνω στο Foundation. Το UIKit περιλαμβάνει views, controls, windows, viewControllers, gestures, animations, fonts, images.

Συνοψίζοντας, το Darwin παρέχει τα θεμέλια σε επίπεδο λειτουργικού συστήματος στα οποία βασίζονται όλα αυτά τα frameworks. Το Core Foundation και το Foundation παρέχουν τους τύπους δεδομένων χαμηλού και υψηλού επιπέδου, τις δομές και τα βοηθητικά προγράμματα που είναι απαραίτητα για πιο σύνθετες λειτουργίες και το UIKit παρέχει τα στοιχεία που είναι απαραίτητα για τη δημιουργία της διεπαφής χρήστη και την απόκριση στις αλληλεπιδράσεις του χρήστη.

### **4.2.6 Η Εφαρμογή και το περιβάλλον**

Στην ανάπτυξη εφαρμογών iOS, ο χειρισμός μιας εφαρμογής και του περιβάλλοντός της επιτυγχάνεται μέσω μιας σειράς καταστάσεων στον κύκλο ζωής της εφαρμογής. Η κατανόηση αυτού του κύκλου ζωής από έναν προγραμματιστή είναι ζωτικής σημασίας για τη δημιουργία, το debugging και τη βελτιστοποίηση της εφαρμογής ώστε να παρέχει μια ομαλή εμπειρία χρήσης. Παρακάτω παρατίθεται μια σύντομη επεξήγηση των βασικών καταστάσεων όπως φαίνεται στο σχήμα 4.2 [31]:





Σχήμα 4.2 Ο κύκλος ζωής της εφαρμογής

- **Not running (Δεν εκτελείται):** Πρόκειται για την κατάσταση κατά την οποία η εφαρμογή δεν έχει ξεκινήσει ή εκτελούνταν αλλά τερματίστηκε από το σύστημα.
- **Inactive (Ανεργή):** Σε αυτή την κατάσταση, η εφαρμογή εκτελείται στο προσκήνιο, αλλά προς το παρόν δεν λαμβάνει συμβάντα. Αυτό μπορεί να συμβεί για ένα σύντομο χρονικό διάστημα όταν μια εφαρμογή μεταβαίνει στο ή από το παρασκήνιο ή όταν το σύστημα προτρέπει τον χρήστη να ανταποκριθεί σε κάποιο συμβάν, όπως μια εισερχόμενη τηλεφωνική κλήση ή ένα μήνυμα SMS.
- **Active (Ενεργό):** Σε αυτή την κατάσταση η εφαρμογή βρίσκεται στο προσκήνιο και λαμβάνει συμβάντα. Αυτή είναι η τυπική κατάσταση λειτουργίας για εφαρμογές που βρίσκονται σε πρώτο πλάνο.
- **Background (παρασκήνιο):** Εδώ, η εφαρμογή βρίσκεται στο παρασκήνιο και εκτελεί κώδικα. Οι περισσότερες εφαρμογές εισέρχονται για λίγο σε αυτή την κατάσταση πριν την αναστολή της λειτουργίας τους. Ωστόσο, μια εφαρμογή που ζητά επιπλέον χρόνο εκτέλεσης μπορεί να παραμείνει σε αυτή την κατάσταση για μεγαλύτερο χρονικό διάστημα.
- **Suspended (σε αναστολή):** Σε αυτή την κατάσταση, η εφαρμογή βρίσκεται στο παρασκήνιο αλλά δεν εκτελεί κώδικα. Το σύστημα μεταβαίνει αυτόματα τις εφαρμογές σε αυτή την κατάσταση. Ενώ βρίσκεται σε αναστολή, μια εφαρμογή παραμένει στη μνήμη αλλά δεν εκτελεί κώδικα.

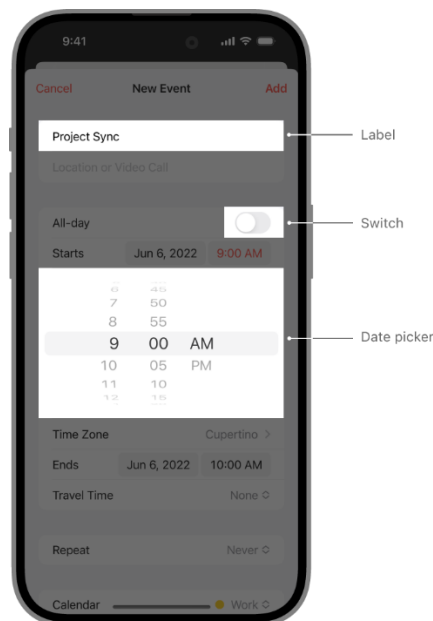
Όταν ένας χρήστης εκκινεί μια εφαρμογή, αυτή μεταβαίνει από την κατάσταση "Δεν εκτελείται" στην κατάσταση "Ενεργή". Εάν ο χρήστης πατήσει το home button, η εφαρμογή μεταβαίνει στην κατάσταση "Background" και λίγο αργότερα στην κατάσταση "Suspended". Όταν η εφαρμογή ξυπνήσει ή επανεκκινήσει, μεταβαίνει ξανά στην κατάσταση "Ενεργή".

Οι εφαρμογές iOS μπορούν επίσης να ανταποκρίνονται σε αλλαγές στο περιβάλλον τους -όπως αλλαγές στον προσανατολισμό της συσκευής ή προειδοποιήσεις για χαμηλή μνήμη- υλοποιώντας κατάλληλες μεθόδους στο αντικείμενο `app delegate`. Η κλάση `AppDelegate` επικοινωνεί με το αντικείμενο `UIApplication` για τη διαχείριση των μεταβάσεων μεταξύ των καταστάσεων του κύκλου ζωής.

- `application:didFinishLaunchingWithOptions`: Αυτή η μέθοδος καλείται όταν η εφαρμογή έχει ξεκινήσει και έχει φορτωθεί το κύριο storyboard ή το αρχείο  `nib`, αλλά δεν έχει γίνει ακόμη αποκατάσταση της κατάστασης.
- `applicationDidBecomeActive`: Αυτή η μέθοδος καλείται όταν η εφαρμογή μεταβαίνει από μια ανενεργή κατάσταση σε μια ενεργή κατάσταση.
- `applicationWillResignActive`: Αυτή η μέθοδος καλείται όταν η εφαρμογή πρόκειται να μετακινηθεί από μια ενεργή κατάσταση σε μια ανενεργή κατάσταση.
- `applicationDidEnterBackground`: Αυτή η μέθοδος καλείται όταν η εφαρμογή εισέρχεται στο παρασκήνιο.

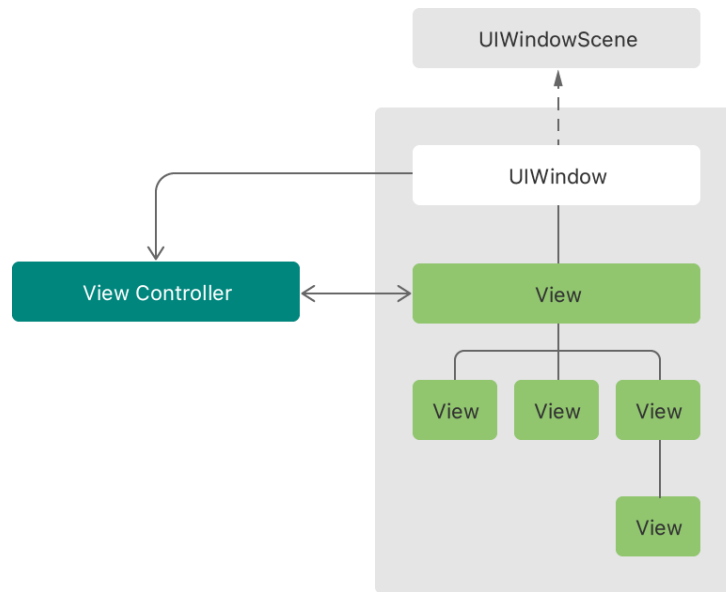
### 4.2.7 UI Components – UIKit

Οι προβολές (Views) και τα στοιχεία ελέγχου (Controls) είναι τα οπτικά δομικά στοιχεία του UI της εφαρμογής. χρησιμοποιούνται για να σχεδιάζουν και να οργανώνουν το περιεχόμενο της εφαρμογής στην οθόνη. Η `UIView` είναι η βασική κλάση για όλες τις προβολές (Views) και ορίζει την κοινή συμπεριφορά τους. Η `UIControl` ορίζει πρόσθετες συμπεριφορές που αφορούν ειδικά κουμπιά, διακόπτες και άλλες προβολές που έχουν σχεδιαστεί για αλληλεπιδράσεις με τον χρήστη.



Σχήμα 4.3 Τα Views σε μια τυπική οθόνη της εφαρμογής ημερολογίου. Κάθε components είναι υποκλάση του `UIView`. Τα Views που έχουν και αλληλεπίδραση με τον χρήστη είναι υποκλάση του `UIControl`.

Οι προβολές μπορούν να φιλοξενούν άλλες προβολές. Η ενσωμάτωση μιας προβολής μέσα σε μια άλλη δημιουργεί μια σχέση περιορισμού μεταξύ της προβολής που φιλοξενεί (γνωστή ως `superview`) και της ενσωματωμένης προβολής (γνωστή ως `subview`). Οι ιεραρχίες προβολών διευκολύνουν τη διαχείριση των προβολών.



Σχήμα 4.4 Βασική δομή που ακολουθούν όλες οι οθόνες σε μια εφαρμογή iOS

Κάθε σκηνή στο περιβάλλον εργασίας της εφαρμογής σας περιέχει ένα αντικείμενο παραθύρου UIWindow και ένα ή περισσότερα views. Το παράθυρο χρησιμεύει ως αόρατο κοντέινερ για το υπόλοιπο UI, λειτουργεί ως κοντέινερ ανώτατου επιπέδου για τα views και δρομολογεί συμβάντα σε αυτές. Τα views παρέχουν το πραγματικό περιεχόμενο που βλέπουν οι χρήστες στην οθόνη, σχεδιάζοντας κείμενο, εικόνες και άλλους τύπους προσαρμοσμένου περιεχομένου. Τα παράθυρα είναι αντικείμενα μεγάλης διάρκειας ζωής και τα καταστρεφете μόνο όταν κατεβάζετε ολόκληρο το UI της σκηνής σας. Αντίθετα, μπορεί να αλλάζετε συχνά τις προβολές σε αυτό το παράθυρο, ιδιαίτερα όταν θέλετε να εμφανίσετε νέο περιεχόμενο ή πληροφορίες.

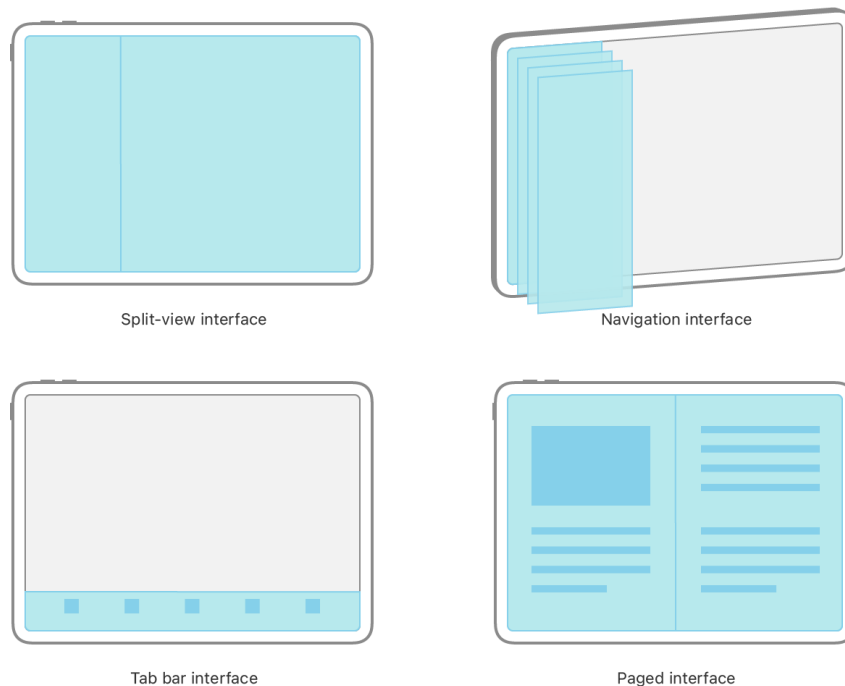
Για να διαχειρίζεστε αποτελεσματικά τα views σας και για να είναι εύκολη η προσθήκη και η αφαίρεσή τους από ένα παράθυρο, το UIKit παρέχει View Controllers. Ένας view controller διαχειρίζεται ένα ενιαίο σύνολο από views για την εφαρμογή σας και διατηρεί τις πληροφορίες σε αυτά τα views ενημερωμένες. Κάθε παράθυρο έχει έναν root view controller, τον οποίο χρησιμοποιείτε για να καθορίσετε το αρχικό σύνολο προβολών του παραθύρου. Όταν θέλετε να αλλάξετε αυτό το σύνολο προβολών, δίνετε εντολή στο UIKit να παρουσιάσει ή να απορρίψει πρόσθετους controllers. Το UIKit χειρίζεται τη μετάβαση από ένα σύνολο προβολών σε ένα άλλο και διαχειρίζεται ολόκληρη τη διεπαφή της εφαρμογής σας μέσω των αντικειμένων view controller. Ως αποτέλεσμα, οι ελεγκτές προβολής παίζουν σημαντικό ρόλο στην υλοποίηση του UI σας.

Εκτός από τα θεμελιώδη μέρη του UIKit, που αφορούν τον σχεδιασμό του UI και την διαχείριση της ζωής της εφαρμογής το framework περιέχει επίσης και λειτουργίες που εξυπηρετούν τα παρακάτω:

**Γραφικά και κινούμενα σχέδια:** Το UIKit περιλαμβάνει εργαλεία για τη σχεδίαση και την κίνηση προσαρμοσμένων στοιχείων διεπαφής χρήστη. Παρέχει υποστήριξη τόσο για 2D όσο και για 3D μετασχηματισμούς, καθώς και APIs για την παρουσίαση κινούμενων εικόνων αυτών των μετασχηματισμών και άλλων ιδιοτήτων, όπως η διαφάνεια και το χρώμα.

**Χειρισμός αφής:** Το UIKit παρέχει εξελιγμένες δυνατότητες χειρισμού αφής, υποστηρίζοντας συμβάντα πολλαπλής αφής, χειρονομίες όπως τσιμπήματα, σαρώσεις και περιστροφές, καθώς και εργαλεία για προσαρμοσμένους αναγνωριστές χειρονομιών.

**Πλοήγηση και διάταξη:** Το UIKit προσφέρει διάφορα στοιχεία για τη διαχείριση της πλοήγησης και της διάταξης της εφαρμογής σας. Μπορείτε να χρησιμοποιήσετε ελεγκτές πλοήγησης, μπάρες καρτελών, ελεγκτές διαχωρισμένης προβολής για να ορίσετε τη ροή της εφαρμογής σας (σχήμα 4.5). Οι κλάσεις αυτόματης διάταξης και μεγέθους επιτρέπουν στη διεπαφή σας να προσαρμόζεται σε διαφορετικά μεγέθη και προσανατολισμούς συσκευών.



Σχήμα 4.5 Προσανατολισμός και τύποι προβολών σε διαφορετικές διατάξεις

**Διαχείριση και εμφάνιση κειμένου:** Το UIKit παρέχει εργαλεία για την εμφάνιση κειμένου, επιτρέποντας εξελιγμένη τυπογραφία, και προσφέρει ισχυρά εργαλεία για τη δημιουργία, επεξεργασία και εμφάνιση κειμένου.

**Προσβασιμότητα:** Το UIKit έχει σχεδιαστεί με γνώμονα την προσβασιμότητα, παρέχοντας εργαλεία που επιτρέπουν στις εφαρμογές σας να χρησιμοποιούνται από άτομα με διάφορες ικανότητες και αναπηρίες.

**Πρόσβαση στη φωτογραφική μηχανή και τη βιβλιοθήκη φωτογραφιών:** Το UIKit προσφέρει API (Application Programming Interface) για πρόσβαση στην κάμερα και τη βιβλιοθήκη φωτογραφιών της συσκευής, επιτρέποντας στους χρήστες να τραβούν φωτογραφίες και να επιλέγουν υπάρχουσες εικόνες.

### 4.3 Αρχιτεκτονική εφαρμογής – βέλτιστες πρακτικές

Τα τελευταία χρόνια έχουν προταθεί διάφορες αρχιτεκτονικές για την δημιουργία συστημάτων πληροφορικής, τα οποία έχουν τουλάχιστον μια μοντελοποίηση ενός επιχειρησιακού κανόνα και μια αλληλεπίδραση με τον χρήστη και βρίσκονται είτε σε τεχνολογίες web, desktop ή mobile. Τα περισσότερα συστήματα χρησιμοποιούν σε κάποιο σημείο τους μια γλώσσα η οποία είναι αντικειμενοστραφής.

### 4.3.1 Γενικές αρχές του αντικειμενοστραφή προγραμματισμού ( object oriented programming OOP)

Σύμφωνα με την Microsoft [37] ο αντικειμενοστραφής προγραμματισμός (OOP) είναι ένα παράδειγμα προγραμματισμού που βασίζεται στην έννοια των αντικειμένων, τα οποία μπορούν να περιέχουν δεδομένα και κώδικα . Τα δεδομένα έχουν τη μορφή (συντάσσονται ως χαρακτηριστικά ή ιδιότητες), και ο κώδικας έχει τη μορφή εκτελέσιμων μπλοκ (συντάσσονται ως μέθοδοι). Ένα κοινό χαρακτηριστικό των αντικειμένων είναι ότι οι μέθοδοι συνδέονται με αυτά και μπορούν να έχουν πρόσβαση και να τροποποιούν τα πεδία δεδομένων του αντικειμένου.

Στην OOP, τα προγράμματα υπολογιστών σχεδιάζονται με την σύνταξη τους από αντικείμενα που αλληλεπιδρούν μεταξύ τους . Οι γλώσσες OOP είναι ποικίλες, αλλά οι πιο δημοφιλείς είναι βασισμένες σε κλάσεις, που σημαίνει ότι τα αντικείμενα είναι περιπτώσεις κλάσεων, οι οποίες καθορίζουν και τους τύπους τους. Οι κλάσεις είναι αφηρημένα προσχέδια (blueprints) που δημιουργούν συγκεκριμένα αντικείμενα . Οι κλάσεις συχνά αντιπροσωπεύουν ευρείες κατηγορίες, όπως το αυτοκίνητο ή ο σκύλος που έχουν κοινά χαρακτηριστικά . Αυτές οι κλάσεις καθορίζουν ποια χαρακτηριστικά θα έχει ένα instance αυτού του τύπου, όπως το χρώμα, αλλά όχι την τιμή αυτών των χαρακτηριστικών για ένα συγκεκριμένο αντικείμενο . Οι κλάσεις μπορούν επίσης να περιέχουν συναρτήσεις που ονομάζονται μέθοδοι και είναι διαθέσιμες μόνο σε αντικείμενα αυτού του τύπου .

Οι τέσσερις βασικές αρχές του αντικειμενοστραφούς προγραμματισμού σύμφωνα με το [37] είναι οι εξής:

- **Abstraction:** Μοντελοποίηση των σχετικών χαρακτηριστικών και αλληλεπιδράσεων των οντοτήτων ως κλάσεων για τον ορισμό μιας αφηρημένης αναπαράστασης ενός συστήματος.
- **Encapsulation:** Απόκρυψη της εσωτερικής κατάστασης και της λειτουργικότητας ενός αντικειμένου και δυνατότητα πρόσβασης μόνο μέσω ενός δημόσιου συνόλου λειτουργιών.
- **Inheritance:** Επιτρέπει σε μια κλάση να κληρονομεί χαρακτηριστικά και μεθόδους από μια άλλη κλάση, σχηματίζοντας μια ιεραρχική σχέση μεταξύ των κλάσεων.
- **Polymorphism:** Επιτρέπει τη χρήση ενός ονόματος ή ενός τελεστή με διαφορετικούς τύπους αντικειμένων με διαφορετικούς τρόπους ανάλογα με την κλάση τους.

### 4.3.2 Αρχιτεκτονικές και αρχές πολύπλοκων συστημάτων

Τα έργα και τα συστήματα πληροφορικής γίνονται ολοένα και πιο περίπλοκα και συχνά περιλαμβάνουν περισσότερα από 1 άτομα τα οποία καλούνται να συνεργαστούν για να δομήσουν μια εφαρμογή. Για την επιτυχή δημιουργία ενός τέτοιου έργου συχνά χρησιμοποιούνται προγράμματα που βοηθούν την συνεργασία μεταξύ των προγραμματιστών (Version Control, git), και γίνονται παραδοχές για την χρήση αρχών, αρχιτεκτονικών και ενιαίου στυλ κώδικα.

#### SOLID

Η πλέον διαδεδομένη αρχή του OOP είναι η αρχή SOLID [38], η οποία βασίζεται σε 5 παραδοχές, 1 για κάθε γράμμα της λέξης, για την δημιουργία κώδικα ο οποίος είναι πιο κατανοητός, ανεξάρτητος των τμημάτων του, και επεκτάσιμος.

Single Responsibility Principle (SPR), Η αρχή της ενιαίας ευθυνής: Κάθε κλάση έχει μια συγκεκριμένη ευθύνη, που σημαίνει ότι σε περίπτωση που θα πρέπει να αλλάξει θα πρέπει να αλλάξει για έναν και μόνο λόγο. Για αυτό τον λόγο θα πρέπει να δημιουργούνται κλάσεις οι οποίες είναι μικρές σε μέγεθος

και έχουν ένα συγκεκριμένο σκοπό. Μια κλάση θα πρέπει να είναι υπεύθυνη για μια συγκεκριμένη εργασία ή λειτουργία.

**Open Closed Principle (OCP):** Η OCP ορίζει ότι οι οντότητες λογισμικού (κλάσεις, modules, συναρτήσεις κ.λπ.) πρέπει να είναι ανοικτές για επέκταση αλλά κλειστές για τροποποίηση. Στο πλαίσιο της ανάπτυξης iOS, αυτό σημαίνει ότι θα πρέπει ο κώδικας να σχεδιάζεται με τρόπο που να επιτρέπει την εύκολη επέκταση ή προσθήκη νέων χαρακτηριστικών χωρίς να τροποποιείται ο υπάρχων κώδικας. Αυτό μπορεί να επιτευχθεί με τη χρήση πρωτοκόλλων και την τήρηση των αρχών του dependency injection το οποίο θα αναλυθεί παρακάτω.

**Liskov Substitution Principle (LSP):** Η LSP δηλώνει ότι τα αντικείμενα μιας υπερκλάσης (parent class) πρέπει να μπορούν να αντικατασταθούν με αντικείμενα των υποκλάσεων (child/sub class) της χωρίς να επηρεάζεται η ορθή λειτουργία του προγράμματος. Στην ανάπτυξη του iOS, η αρχή αυτή τονίζει ότι οι υποκλάσεις θα πρέπει να μπορούν να αντικαταστήσουν τη γονική τους κλάση χωρίς να εισάγουν απροσδόκητη συμπεριφορά ή να παραβιάζουν το συμβόλαιο που ορίζει η γονική κλάση. Προωθεί τη χρήση της σωστής κληρονομικότητας και την τήρηση της σχέσης "(child class) is-a (parent class)".

**Interface Segregation Principle (ISP):** Η ISP ορίζει ότι οι κλάσεις και οι δομές δεν πρέπει να αναγκάζονται να εξαρτώνται από διεπαφές που δεν χρησιμοποιούν. Με το διαχωρισμό των διεπαφών, αποτρέπετε την επιβάρυνση των πελατών με περιττές εξαρτήσεις.

**Dependency Inversion Principle (DIP):** Η DIP δηλώνει ότι οι μονάδες υψηλού επιπέδου δεν πρέπει να εξαρτώνται από μονάδες χαμηλού επιπέδου. Και οι δύο θα πρέπει να εξαρτώνται από αφαιρέσεις. Προωθεί τη χαλαρή σύζευξη και το dependency injection. Στην ανάπτυξη του iOS, η αρχή αυτή ενθαρρύνει τη χρήση πρωτοκόλλων ως αφαιρέσεις, επιτρέποντας στις ενότητες να εξαρτώνται από αφαιρέσεις αντί για συγκεκριμένες υλοποιήσεις. Αυτό προάγει την ευελιξία, τη δυνατότητα ελέγχου και την πολυμορφικότητα.

### **Dependency Injection (DI)**

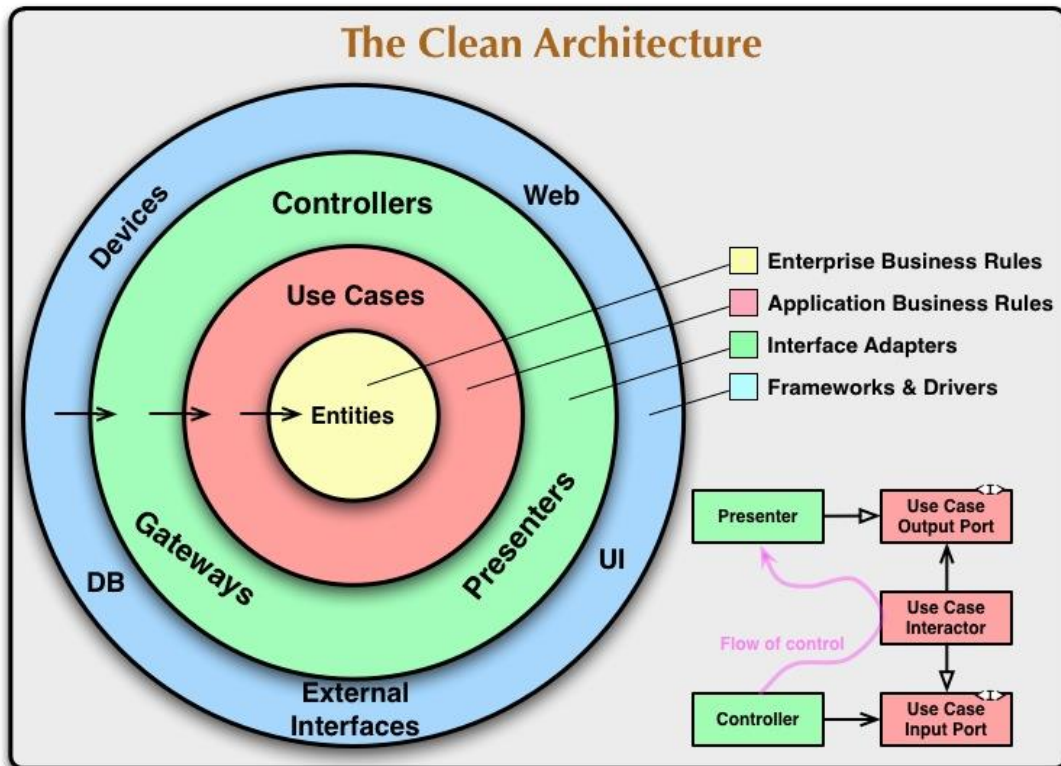
Η έγχυση εξαρτήσεων είναι μια τεχνική προγραμματισμού κατά την οποία ένα αντικείμενο ή μια συνάρτηση λαμβάνει άλλα αντικείμενα ή συναρτήσεις από τα οποία εξαρτάται. Αυτές οι εξαρτήσεις παρέχονται συνήθως από ένα εξωτερικό κομμάτι κώδικα, όπως ένα πλαίσιο ή ένας container. Το DI αποσκοπεί στο separation of concerns (βασισμένο στο SPR) της κατασκευής αντικειμένων και της χρήσης τους, οδηγώντας σε χαλαρά συνδεδεμένες μονάδες.

### **Design Patterns - Architecture**

Τα πρότυπα σχεδίασης είναι τυπικές λύσεις σε συχνά εμφανιζόμενα προβλήματα στη σχεδίαση λογισμικού. Είναι προκατασκευασμένα σχέδια που μπορείτε να προσαρμόσετε για να λύσετε ένα επαναλαμβανόμενο πρόβλημα σχεδίασης στον κώδικα [39]

Επιπλέον, όλα τα πρότυπα μπορούν να κατηγοριοποιηθούν ανάλογα με την πρόθεσή τους ή τον σκοπό τους. Δημιουργικά πρότυπα παρέχουν μηχανισμούς δημιουργίας αντικειμένων που αυξάνουν την ευελιξία και την επαναχρησιμοποίηση του υπάρχοντος κώδικα. Τα δομικά πρότυπα εξηγούν πώς να συναρμολογήσετε αντικείμενα και κλάσεις σε μεγαλύτερες δομές, διατηρώντας παράλληλα αυτές τις δομές ευέλικτες και αποδοτικές. Τα πρότυπα συμπεριφοράς φροντίζουν για την αποτελεσματική επικοινωνία και την ανάθεση αρμοδιοτήτων μεταξύ αντικειμένων.

## Clean Architecture



Εικόνα 4 Η Αρχιτεκτονική ενός συστήματος που ακολουθεί την καθαρή αρχιτεκτονική (Clean architecture)

Η καθαρή αρχιτεκτονική είναι ένα πρότυπο σχεδίασης λογισμικού που ακολουθεί τις έννοιες του καθαρού κώδικα και εφαρμόζει τις αρχές SOLID. Η καθαρή αρχιτεκτονική προτάθηκε από τον Robert C. Martin (γνωστό και ως Uncle Bob) στο βιβλίο του The Clean Coder [40]. Βασίζεται στην ιδέα του διαχωρισμού των ανησυχιών (separation of concerns) των διαφόρων τμημάτων και στρωμάτων ενός συστήματος και στην επιβολή ενός κανόνα εξάρτησης που **εξασφαλίζει ότι οι εξαρτήσεις μπορούν να δείχνουν μόνο προς τα μέσα** (Εικόνα 4), από τα εξωτερικά στρώματα στα εσωτερικά στρώματα.

- **Οντότητες (Entities):** Ο εσωτερικότερος κύκλος περιέχει τις οντότητες, οι οποίες είναι τα επιχειρησιακά αντικείμενα που περιέχουν τους πιο γενικούς και υψηλού επιπέδου κανόνες. Είναι ανεξάρτητες από οποιαδήποτε συγκεκριμένη εφαρμογή και μπορούν να χρησιμοποιηθούν από πολλές διαφορετικές εφαρμογές.
- **Περιπτώσεις χρήσης (Use cases):** Ο επόμενος κύκλος περιέχει τις περιπτώσεις χρήσης, οι οποίες είναι συγκεκριμένοι για κάθε εφαρμογή επιχειρησιακοί κανόνες που περιγράφουν τις αλληλεπιδράσεις μεταξύ των οντοτήτων και των χρηστών ή εξωτερικών συστημάτων. Είναι επίσης ανεξάρτητες από οποιοδήποτε UI, βάση δεδομένων ή εξωτερικό οργανισμό.
- **Προσαρμογείς διεπαφής:** Ο επόμενος κύκλος περιέχει τους προσαρμογείς διεπαφής, οι οποίοι είναι οι κλάσεις που μετατρέπουν τα δεδομένα από τη μορφή που είναι πιο βολική για τις περιπτώσεις χρήσης και τις οντότητες, στη μορφή που είναι πιο βολική για κάποιον εξωτερικό οργανισμό, όπως η βάση δεδομένων ή ο ιστός.
- **Πλαίσια (Frameworks) και οδηγοί (drivers):** Ο ακραίος κύκλος περιλαμβάνει τα Frameworks και τους οδηγούς, τα οποία είναι τα εργαλεία που παρέχουν γενική λειτουργικότητα, όπως Frameworks ιστού, βάσεων δεδομένων, UI κ.λπ. Αποτελούν πρόσθετα plugins της εφαρμογής και μπορούν εύκολα να αλλάξουν χωρίς να επηρεάσουν τους εσωτερικούς κύκλους.



### 4.3.3 Σχεδιαστικά πρότυπα MVC, MVVM, VIPER

Η ανάπτυξη εφαρμογών iOS απαιτεί την επιλογή μιας κατάλληλης αρχιτεκτονικής για την οργάνωση του κώδικα και τη διασφάλιση της ποιότητας, της δυνατότητας συντήρησης και της επεκτασιμότητάς του. Υπάρχουν πολλά αρχιτεκτονικά μοτίβα διαθέσιμα για τους προγραμματιστές iOS, αλλά τρία από τα πιο δημοφιλή είναι τα MVC, MVVM και VIPER. Σε αυτή την ενότητα, θα συγκρίνουμε αυτά τα τρία μοτίβα και θα παρουσιάσουμε μερικά παραδείγματα για την υλοποίησή τους στη Swift.

#### MVC: Model View Controller

Το MVC είναι ένα ευρέως χρησιμοποιούμενο μοτίβο σχεδίασης για την ανάπτυξη εφαρμογών, και είναι το πρότυπο το οποίο πρότεινε η Apple μέχρι το 2019 και την δημιουργία του SwiftUI το οποίο ακολουθεί το μοντέλο του MVVM το οποίο θα αναφερθούμε παρακάτω. Το MVC είναι ιδανικό για δημιουργία γρήγορων πρωτοτύπων και λειτουργιών καθώς δεν απαιτεί πολλές κλάσεις και boilerplate κώδικα (κώδικας ο οποίος γράφεται ξανά και ξανά). Είναι ιδανικός για δημιουργία εφαρμογών οι οποίες θα πρέπει να αναπτυχθούν γρήγορα. Δεν συνιστάτε για την χρήση από μεγάλες ομάδες προγραμματιστών καθώς η λειτουργίες του controller, σε μια εφαρμογή μεγάλου μεγέθους μπορούν να φτάσουν τις κλάσεις των controllers σε αρκετές χιλιάδες γραμμών κώδικα (όπως έχει μείνει εθιμοτυπικά Massive View Controller)

- **Model:** Αντιπροσωπεύει τα δεδομένα και την επιχειρησιακή λογική της εφαρμογής.
- **View:** Αντιπροσωπεύει τη διεπαφή χρήστη και είναι υπεύθυνη για την εμφάνιση των δεδομένων από το Μοντέλο.
- **Controller:** Λειτουργεί ως γέφυρα μεταξύ του Μοντέλου και της Προβολής, χειρίζεται την είσοδο του χρήστη και ενημερώνει αναλόγως την Προβολή.

```

// Model
struct Person {
    let name: String
    let age: Int
}

// View
class PersonView: UIView {
    @IBOutlet weak var nameLabel: UILabel!
    @IBOutlet weak var ageLabel: UILabel!

    func configure(with person: Person) {
        nameLabel.text = person.name
        ageLabel.text = "\(person.age)"
    }
}

// Controller
class PersonViewController: UIViewController {
    @IBOutlet weak var personView: PersonView!
    var person: Person!

    override func viewDidLoad() {
        super.viewDidLoad()
        personView.configure(with: person)
    }
}

```

Σχήμα 4.6 MVC Patern in swift



**MVVM: Model View View Model**

```

// Model
struct Person {
    let name: String
    let age: Int
}
// ViewModel
class PersonViewModel {
    private let person: Person
    var name: String {
        return person.name
    }
    var age: String {
        return "\(person.age)"
    }
    init(person: Person) {
        self.person = person
    }
}
// View
class PersonView: UIView {
    @IBOutlet weak var nameLabel: UILabel!
    @IBOutlet weak var ageLabel: UILabel!

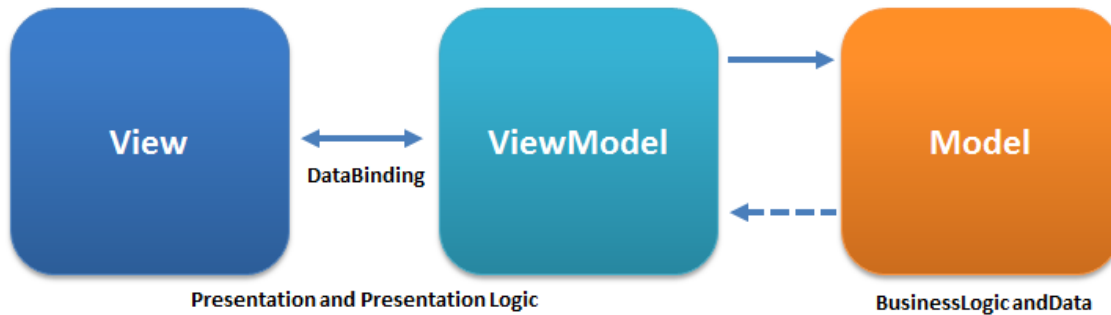
    func configure(with viewModel: PersonViewModel) {
        nameLabel.text = viewModel.name
        ageLabel.text = viewModel.age
    }
}
// Controller
class PersonViewController: UIViewController {
    @IBOutlet weak var personView: PersonView!
    var personViewModel: PersonViewModel!

    override func viewDidLoad() {
        super.viewDidLoad()
        personView.configure(with: personViewModel)
    }
}

```

Σχήμα 4.7 MVVM Patern In swift

Το MVVM είναι μια εξέλιξη του προτύπου MVC περιέχει και τα 3 επίπεδα του, και έχει ως πρωταρχικό στόχο τη βελτίωση του separation of concerns με την απόκρυψη του μοντέλου από το επίπεδο του View. Εισάγει ένα νέο συστατικό που ονομάζεται ViewModel:



Σχήμα 4.8 Η επικοινωνία μεταξύ των στρωμάτων στην αρχιτεκτονική MVVM

**ViewModel:** Λειτουργεί ως επίπεδο μετασχηματισμού δεδομένων και παρουσίασης μεταξύ του Model και της View. Είναι υπεύθυνο για την έκθεση των δεδομένων σε μορφή κατάλληλη για την εμφάνιση των δεδομένων σε κατάλληλη μορφή για προβολή από το View (Προβολή). Στο επίπεδο του View Model εντάσσεται όλο το business logic της εφαρμογής και τα Use Cases τα οποία θα πρέπει να καλύψουμε. Ένα άλλο πλεονέκτημα είναι ότι αποδεσμεύει το επίπεδο του View και το Model, καθώς επικοινωνούν μέσω του ViewModel. Αυτό καθιστά τον κώδικα πιο modular και προσαρμόσιμο σε διαφορετικά πλαίσια.

Ωστόσο, το MVVM επιφέρει κάποιες προκλήσεις. Μία από αυτές είναι ότι εισάγει περισσότερες κλάσεις και αρχεία προς διαχείριση, γεγονός που μπορεί να αυξήσει το γνωστικό φορτίο για τους προγραμματιστές. Μια άλλη πρόκληση είναι ότι απαιτεί έναν καλό μηχανισμό επικοινωνίας μεταξύ της προβολής και του ViewModel, όπως bindings or reactive frameworks.

#### **VIPER: View-Interactor-Presenter-Entity-Router**

Το VIPER είναι ένα προηγμένο αρχιτεκτονικό πρότυπο που βασίζεται στην αρχή Single Responsibility. Είναι ένα πρότυπο το οποίο ενδείκνυται για έργα πολύ μεγάλου μεγέθους και μεγάλο αριθμό μελών ομάδας. Το πρότυπο VIPER, ενώ παρέχει την μέγιστη δυνατή τμηματοποίηση του κώδικα, εισάγει και ένα επίπεδο πολυπλοκότητας που κάνει τους προγραμματιστές σε εφαρμογές μικρού και μεσαίου μεγέθους να το αποφεύγουν.

Διαχωρίζει την εφαρμογή σε πέντε κύρια στοιχεία:

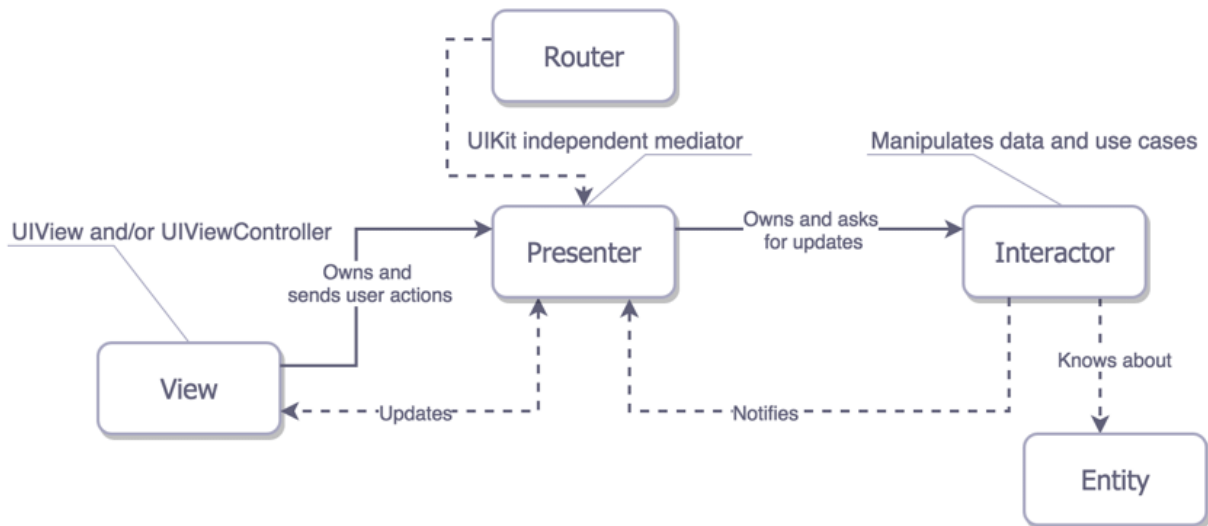
**View:** Εμφανίζει τα δεδομένα και χειρίζεται την είσοδο του χρήστη.

**Interactor:** Περιέχει την επιχειρησιακή λογική και διαχειρίζεται τα δεδομένα.

**Presenter:** Λειτουργεί ως γέφυρα μεταξύ του στοιχείου view της προβολής και του interactor, ενημερώνοντας το View με δεδομένα από τον Interactor.

**Entity:** Αντιπροσωπεύει το μοντέλο δεδομένων.

**Router:** Χειρίζεται την πλοήγηση μεταξύ διαφορετικών οθονών.



Σχήμα 4.9 Το διάγραμμα ροής πληροφορίας της αρχιτεκτονικής VIPER

Εν κατακλείδι, τα MVC, MVVM και VIPER είναι τρία δημοφιλή αρχιτεκτονικά μοτίβα για την ανάπτυξη εφαρμογών iOS. Έχουν διαφορετικά πλεονεκτήματα και μειονεκτήματα, ανάλογα με το περιβάλλον και τις απαιτήσεις του έργου. Το MVC είναι απλό και εύκολο στην υλοποίηση, αλλά μπορεί να οδηγήσει σε μεγάλους και πολύπλοκους Controllers. Το MVVM είναι μια βελτίωση σε σχέση με το MVC, καθώς μειώνει την πολυπλοκότητα του Controller και αποσυνδέει το επίπεδο του View και το Model, αλλά απαιτεί έναν καλό μηχανισμό επικοινωνίας μεταξύ του επιπέδου View και του ViewModel. Το VIPER είναι ένα προηγμένο πρότυπο που ακολουθεί την αρχή Single Responsibility Principle και την αρχή Dependency Inversion Principle, αλλά εισάγει πολύ κώδικα boilerplate και πολυπλοκότητα.

Δεν υπάρχει οριστική απάντηση στο ποιο πρότυπο είναι το καλύτερο για κάθε έργο. Οι προγραμματιστές θα πρέπει να εξετάσουν τα αντισταθμιστικά οφέλη και να επιλέξουν το καταλληλότερο για τις ανάγκες τους. Θα πρέπει επίσης να είναι ανοιχτοί στο να προσαρμόζουν και να εξελίσσουν την αρχιτεκτονική τους καθώς το έργο αναπτύσσεται και αλλάζει.

#### 4.3.4 Αρχιτεκτονικές που υιοθετεί η Swift

Η Swift αποτελεί μια Object Oriented γλώσσα προγραμματισμού ή οποία βασίζεται στις αρχές της κληρονομικότητας αλλά προωθεί τον προγραμματισμό με την χρήση πρωτοκόλλων (Protocol oriented programming) για την επέκταση των λειτουργιών των κλάσεων. Σύνθεση έναντι κληρονομικότητας (Composition over Inheritance): Αυτό σημαίνει ότι μπορείτε να επιτύχετε επαναχρησιμοποίηση του κώδικα και τμηματοποίηση του κώδικα σε μικρότερα τμήματα (modules) συνθέτοντας αντικείμενα με πρωτόκολλα αντί να βασίζεστε σε ιεραρχίες κλάσεων. Αυτή η προσέγγιση προωθεί πιο ευέλικτο και συντηρήσιμο κώδικα, επιτρέποντας στα αντικείμενα να συμμορφώνονται με πολλαπλά πρωτόκολλα, επιτρέποντάς τους να παρουσιάζουν διαφορετικές συμπεριφορές σε διάφορα πλαίσια.

Η Swift επιτρέπει σε έναν τύπο να συμμορφώνεται με πολλαπλά πρωτόκολλα ταυτόχρονα παρέχοντας αυξημένη ευελιξία και προωθώντας την οργάνωση του κώδικα, η Swift επιτρέπει στα αντικείμενα να υιοθετούν συμπεριφορές από πολλαπλά πρωτόκολλα ταυτόχρονα, ελαχιστοποιώντας τη σύζευξη μεταξύ διαφορετικών περιοχών της βάσης κωδικών σας. Με τη συμμόρφωση σε πολλαπλά πρωτόκολλα ταυτόχρονα, η Swift επιτρέπει ειδικευμένη λειτουργικότητα.

Η Swift ενθαρρύνει μια προσέγγιση που ονομάζεται *protocol oriented programming* και περιλαμβάνει τον σχεδιασμό κώδικα γύρω από τα πρωτόκολλα και τη σύνθεσή τους. Αυτή η προσέγγιση επικεντρώνεται στον ορισμό πρωτοκόλλων που περιγράφουν δυνατότητες και τη χρήση τους για τη σύνθεση τύπων, οδηγώντας τελικά σε ιδιαίτερα αποσυνδεδεμένο, επαναχρησιμοποιήσιμο και ελέγξιμο κώδικα (*testable code*).

Στην συγκεκριμένη διπλωματική εργασία έχουν εφαρμοστεί τεχνικές του κλασικού αντικειμενοστραφή προγραμματισμού και σχεδίασης όπως η αρχή SOLID, το Clean Architecture, το MVVM και το MVC.

### 4.3.5 Εφαρμογές Computer Vision στο iOS

Το Vision Framework της Apple είναι ένα σύνολο εργαλείων *computer vision* που οι προγραμματιστές μπορούν να χρησιμοποιήσουν για να δημιουργήσουν εφαρμογές με ισχυρές δυνατότητες ανάλυσης και αναγνώρισης εικόνων. Αξιοποιεί τη μηχανική μάθηση (ML) για να κάνει αυτές τις πολύπλοκες εργασίες πιο προσιτές και αποτελεσματικές. Το Vision Framework με το Object Detection API μπορεί να αναγνωρίζει και να εντοπίζει διάφορα αντικείμενα μέσα σε μια εικόνα με βάση *pre trained* μοντέλα ή προσαρμοσμένα από τον χρήστη. Στην ML, αυτό μπορεί να χρησιμοποιηθεί για την εκπαίδευση και την επαλήθευση μοντέλων που στοχεύουν στην ανίχνευση αντικειμένων. Εκτός από

Το Vision Framework παρουσιάζει επίσης μια εξαιρετική ικανότητα στην ανίχνευση προσώπων και χαρακτηριστικών προσώπου. Είναι σε θέση να ανιχνεύει πρόσωπα μαζί με μεμονωμένα χαρακτηριστικά του προσώπου μέσα σε εικόνες, και μπορεί ακόμη και να διακρίνει την κατεύθυνση προς την οποία ένα πρόσωπο είναι στραμμένο. Αυτά τα χαρακτηριστικά χρησιμοποιούνται συνήθως στη μηχανική μάθηση για την εκπαίδευση μοντέλων για εργασίες που περιλαμβάνουν την αναγνώριση προσώπου, την ανίχνευση συναισθημάτων ή την ανίχνευση υπνηλίας σε εφαρμογές ασφαλείας.

Στον τομέα της ανίχνευσης κειμένου, το Vision Framework υπερέχει στην αναγνώριση και τον εντοπισμό κειμένου μέσα σε εικόνες. Αυτό το αναγνωρισμένο κείμενο μπορεί στη συνέχεια να μεταβιβαστεί σε άλλα API για κατανάλωση. Αυτό το χαρακτηριστικό αποδεικνύεται ζωτικής σημασίας για έργα μηχανικής μάθησης που σχετίζονται με την οπτική αναγνώριση χαρακτήρων (OCR).

Από την άλλη πλευρά, το TensorFlow είναι ένα πιο ευέλικτο *framework* ανοικτού κώδικα που αναπτύχθηκε από την Google. Μπορεί να χρησιμοποιηθεί σε πολλές διαφορετικές πλατφόρμες, όχι μόνο στο iOS, και επιτρέπει μεγαλύτερη προσαρμογή και έλεγχο των μοντέλων μηχανικής μάθησης. Το TensorFlow απευθύνεται σε επιστήμονες δεδομένων και μηχανικούς μηχανικής μάθησης που απαιτούν ένα λεπτομερές επίπεδο ελέγχου και θέλουν να δημιουργούν, να εκπαιδεύουν και να αναπτύσσουν μοντέλα μηχανικής μάθησης από την αρχή. Το TensorFlow Lite είναι μια ελαφριά λύση από τη Google για κινητές συσκευές. Επιτρέπει στους προγραμματιστές να εκτελούν τα μοντέλα TensorFlow στη συσκευή, επιτρέποντας χαμηλή καθυστέρηση και *real time* αποτελέσματα.

Όσον αφορά τη σύγκριση, το TensorFlow (και το TensorFlow Lite) προσφέρει μεγαλύτερη ευελιξία και έλεγχο από το Vision Framework της Apple. Υποστηρίζει μεγαλύτερο εύρος προ-εκπαιδευμένων μοντέλων και σας επιτρέπει να σχεδιάσετε τα δικά σας μοντέλα, κάτι που δεν υποστηρίζεται άμεσα από το Vision, η λειτουργία αυτή παρέχεται μέσω του Xcode. Ωστόσο, αυτή η ευελιξία έρχεται με το κόστος μιας πιο απότομης καμπύλης εκμάθησης και της ανάγκης για περισσότερη εμπειρία στη μηχανική μάθηση.

## Κεφάλαιο 5ο: Ενσωμάτωση – Integration

### 5.1 Εισαγωγή

Στο κεφάλαιο αυτό θα αναλυθεί την διαδικασία για την ολοκληρωμένη ανάπτυξη του συστήματος, από την εκπαίδευση του μοντέλου YOLO, την λειτουργία του μοντέλου για την εξαγωγή αποτελεσμάτων σε στατικές εικόνες και βίντεο, την ανάπτυξη της εφαρμογής iOS και τις λειτουργίες της καθώς και το κρίσιμο κομμάτι την εξαγωγή του μοντέλου αναγνώρισης σε φορητή μορφή για χρήση σε πραγματικό χρόνο από την κάμερα του κινητού και την εξαγωγή συμπερασμάτων για τα μνημεία.

Η ανάπτυξη του μοντέλου έχει πραγματοποιηθεί παράλληλα με την ανάπτυξη της εφαρμογής. Ακολουθεί ένα διάγραμμα (σχήμα 5.1) με τις επιμέρους διεργασίες τις οποίες ακολουθήθηκαν για την ολοκλήρωση του συστήματος

**Ανάπτυξη μοντέλου μηχανικής μάθησης**



**Ανάπτυξη εφαρμογής**

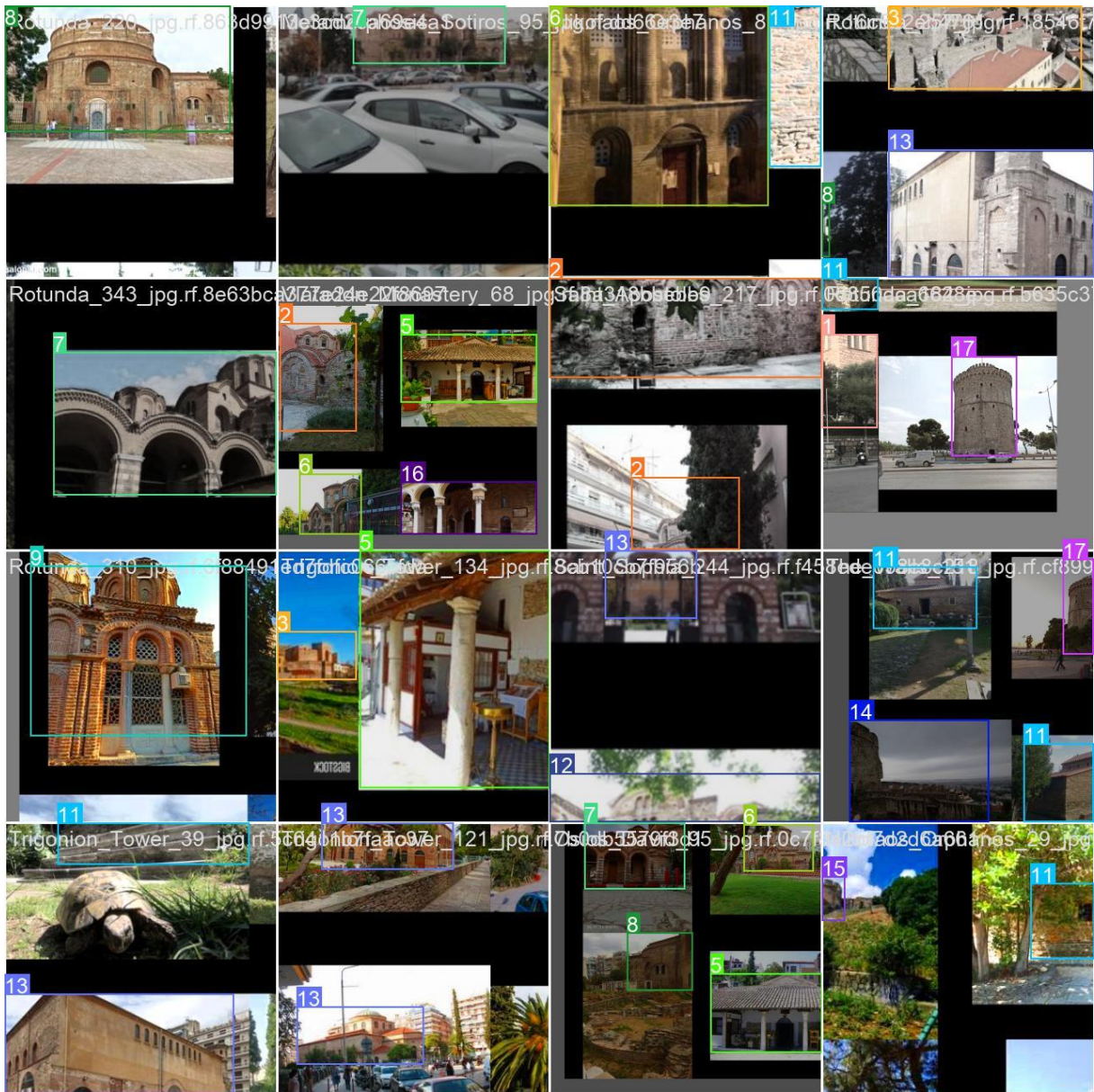
Σχήμα 5.1 Σχεδιασμός της εφαρμογής και περιγραφή εργασιών

## 5.2 Εκπαίδευση του μοντέλου

Στο κεφάλαιο αυτό θα αναλυθεί η διαδικασία για την επιλογή της έκδοσης του μοντέλου, τις τεχνικές για την προετοιμασία των δεδομένων, και την διαδικασία της εκπαίδευσης.

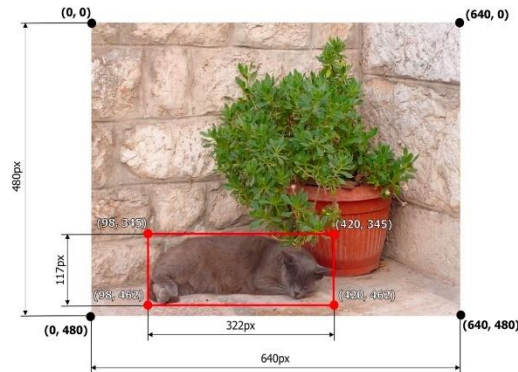
### 5.2.1 Προετοιμασία δεδομένων.

Τα δεδομένα τα οποία παρέχουμε ως είσοδο για την εκπαίδευση του μοντέλου αποτελούνται από 3843 φωτογραφίες των μνημείων της Θεσσαλονίκης κανονικοποιημένα ως προς το μέγεθος και τον αριθμό με 379 φωτογραφίες για κάθε μνημείο με την χρήση του COCO format όπου κάθε φωτογραφία, έχει το αντίστοιχο αρχείο όπου περιέχονται οι πληροφορίες για τα bounding boxes και την κλάση στην οποία ανήκει το αντικείμενο. Για να ορίσουμε την πληροφορία bounding box ενός αντικειμένου θα πρέπει να λάβουμε υπόψη 4 παραμέτρους, την αρχή στο μήκος και το πλάτος της φωτογραφίας (minX, minY), το ύψος και το πλάτος (height, width) του box σε σύγκριση με το μέγεθος της εικόνας.





Η μορφή των δεδομένων αντιστοιχεί 1 εικόνα με τις αντίστοιχες πληροφορίες για τα αντικείμενα τα οποία περιέχει. Π.χ. το αρχείο WhiteTower1.jpg , αντιστοιχεί με το WhiteTower1.txt, όπου το txt αρχείο περιέχει 5 πληροφορίες σε σειρά χωρισμένες με κόμμα, στην πρώτη θέση ο αριθμός της κλάσης η οποία περιέχεται, στην 2<sup>η</sup> και 3<sup>η</sup> θέση αντίστοιχα η κανονικοποιημένη συντεταγμένη του κέντρου του box όπου το (0,0) σημαίνει την πάνω αριστερά γωνία της φωτογραφίας και το (1,1) σημαίνει η κάτω δεξιά. Αντίστοιχα το ύψος και το πλάτος είναι και αυτά κανονικοποιημένα με βάση το μέγεθος εισόδου της φωτογραφίας, δηλαδή σε μια φωτογραφία 640 x 640 pixel, το 0.5h, 0.5w σημαίνει (640\*0.5) 320x320 μέγεθος box [41].

**pascal\_voc**

[x\_min, y\_min, x\_max, y\_max] → [98, 345, 420, 462]

**albumentations**

normalized [x\_min, y\_min, x\_max, y\_max] → [0.153125, 0.71875, 0.65625, 0.9625]

**coco**

[x\_min, y\_min, width, height] → [98, 345, 322, 117]

**yolo**

normalized [x\_center, y\_center, width, height] → [0.4046875, 0.8614583, 0.503125, 0.24375]

Σχήμα 5.2 Ένα παράδειγμα εικόνας με πλαίσιο οριοθέτησης από το σύνολο δεδομένων COCO [41] και η ερμηνεία της σε διάφορες μορφές αναπαράστασης.

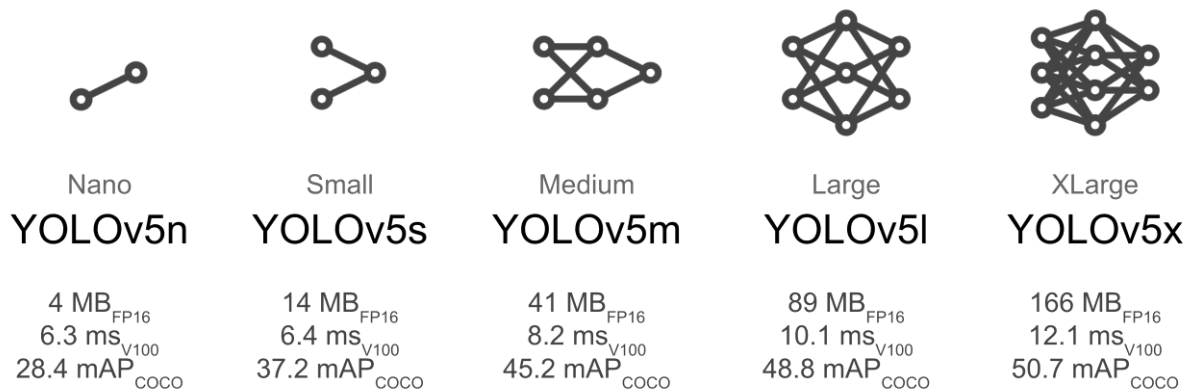
Στην συνέχεια τα δεδομένα χωρίστηκαν σε διαφορετικούς φακέλους για την εκπαίδευση την επαλήθευση και την δοκιμή για την εκπαίδευση του μοντέλου με διαχωρισμό 80/10/10 αντίστοιχα. Για τα δεδομένα εκπαίδευσης εφαρμόστηκαν τεχνικές augmentation, η μία ήταν η εφαρμογή αντικατοπτρισμού (mirror flip) και ενίσχυση της φωτεινότητας στις περιοχές των οριοθετημένων περιοχών.

## 5.2.2 Εκπαίδευση

Για την εκπαίδευση του μοντέλου χρησιμοποιήθηκαν 2 διαφορετικά περιβάλλοντα για 2 διαφορετικά μοντέλα και αρχιτεκτονικές μέχρι να καταλήξουμε στην τελική επιλογή όπου ικανοποιούνται όλες οι ανάγκες οι οποίες αναφέρθηκαν στην εισαγωγή του κεφαλαίου.

Το πρώτο μοντέλο, το YOLOv5 αντίστοιχα εκπαιδεύτηκε με την χρήση Python και του Pytorch framework χρησιμοποιώντας απαραίτητες βοηθητικές μεθόδους για την προετοιμασία των δεδομένων, την δημιουργία αρχείων με την παραμετροποίηση των μοντέλων, την επίβλεψη της εκπαίδευσης, την ανάλυση της εκπαίδευσης και την εξαγωγή σε μορφή ικανή για αξιοποίηση από την εφαρμογή μέσω των αποθετηρίων κώδικα (repositories) των αντίστοιχων διαθέσιμων επίσημων εκδόσεων των μοντέλων.

Για το YOLOv5 χρησιμοποιήθηκε η επίσημη έκδοση του μοντέλου που υποστηρίζεται από την Ultralytics [42] και μπορεί να βρεθεί εδώ [github.com/ultralytics/yolov5](https://github.com/ultralytics/yolov5). Περιέχει τις απαραίτητες βοηθητικές μεθόδους για την παραμετροποίηση, την προετοιμασία των δεδομένων, την εκπαίδευση και την αξιολόγηση του μοντέλου, καθώς και ένα set από ρυθμίσεις του μοντέλου σε διαφορετικά μεγέθη (τα οποία εισάγουν συμβιβασμούς μεταξύ ταχύτητας και ακρίβειας). Για τις δικές μας ανάγκες επιλέξαμε την Medium (ΠΑΡΑΡΤΗΜΑ Α) έκδοση καθώς οι μεγαλύτερες εκδόσεις του μοντέλου καθυστερούσαν στην εξαγωγή αποτελεσμάτων, αυξάνουν το μέγεθος της εφαρμογής οδηγώντας σε μια δυσάρεστη εμπειρία χρήσης.



Σχήμα 5.3 Τα διαφορετικά μεγέθη του YOLOv5 και η απόδοσή τους στο COCO dataset

Για την εκπαίδευση του μοντέλου θα πρέπει να προσδιοριστούν οι παράμετροι εκπαίδευσης. Αυτοί αποτελούν τον αριθμό των εποχών για τον οποίο θα γίνει η εκπαίδευση, τον τρόπο αρχικοποίησης των βαρών του μοντέλου (χρήση pre-trained weights ή όχι) και το μέγεθος της εικόνας εισόδου. Σύμφωνα με τους δημιουργούς του μοντέλου για την εξαγωγή των καλύτερων αποτελεσμάτων, συνιστάται η χρήση τουλάχιστον 200 εποχών και η χρήση των pre trained βαρών του μοντέλου στο COCO dataset. Το μέγεθος της εικόνας σε pixel επηρεάζει τα αποτελέσματα με περισσότερα από 1 αντικείμενα για αναγνώριση σε κάθε εικόνα ή μικρά αντικείμενα. Σαν παραμέτρους την βέλτιστη απόδοση απέδωσαν ή χρήση 200 εποχών εκπαίδευσης, μέγεθος εικόνας 320 και χρήση του προ εκπαιδευμένου μοντέλου.

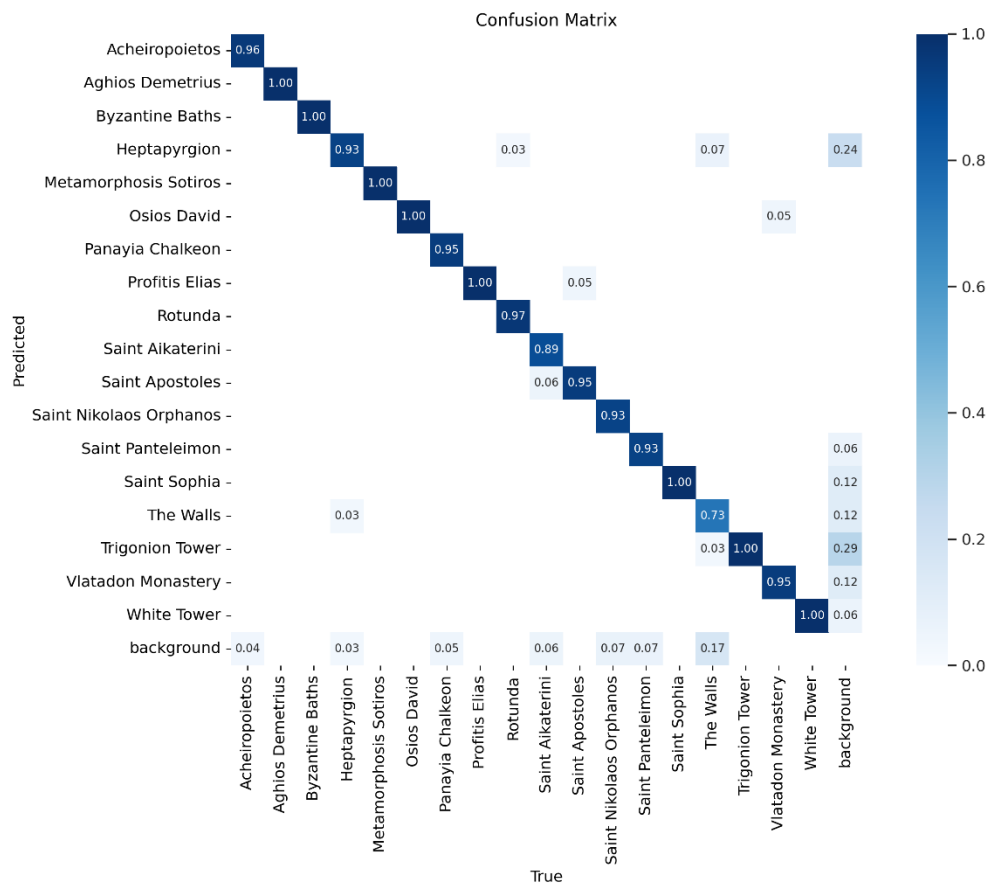
Εκτός από τις βασικές παραμέτρους υπάρχουν και οι υπερπαραμέτροι οι οποίες χρησιμοποιούνται για την ενίσχυση των δεδομένων κατά την διάρκεια της εκπαίδευσης πέρα από τις εργασίες που πραγματοποιήθηκαν για το augmentation πριν την εκκίνηση της διαδικασίας εκπαίδευσης, την αποφυγή του overfitting του μοντέλου και την βελτιστοποίηση της εκπαίδευσης οι οποίοι αναφέρονται στο ΠΑΡΑΡΤΗΜΑ Β.

Η εκπαίδευση του μοντέλου έγινε με την χρήση του Google Colab και την επιτάχυνση μέσω GPU. Για την ολοκλήρωση της εκπαίδευσης των 300 εποχών χρειάστηκαν 7.5 ώρες και αποτέλεσμα ήταν ένα μοντέλο με μέγεθος 48mb.

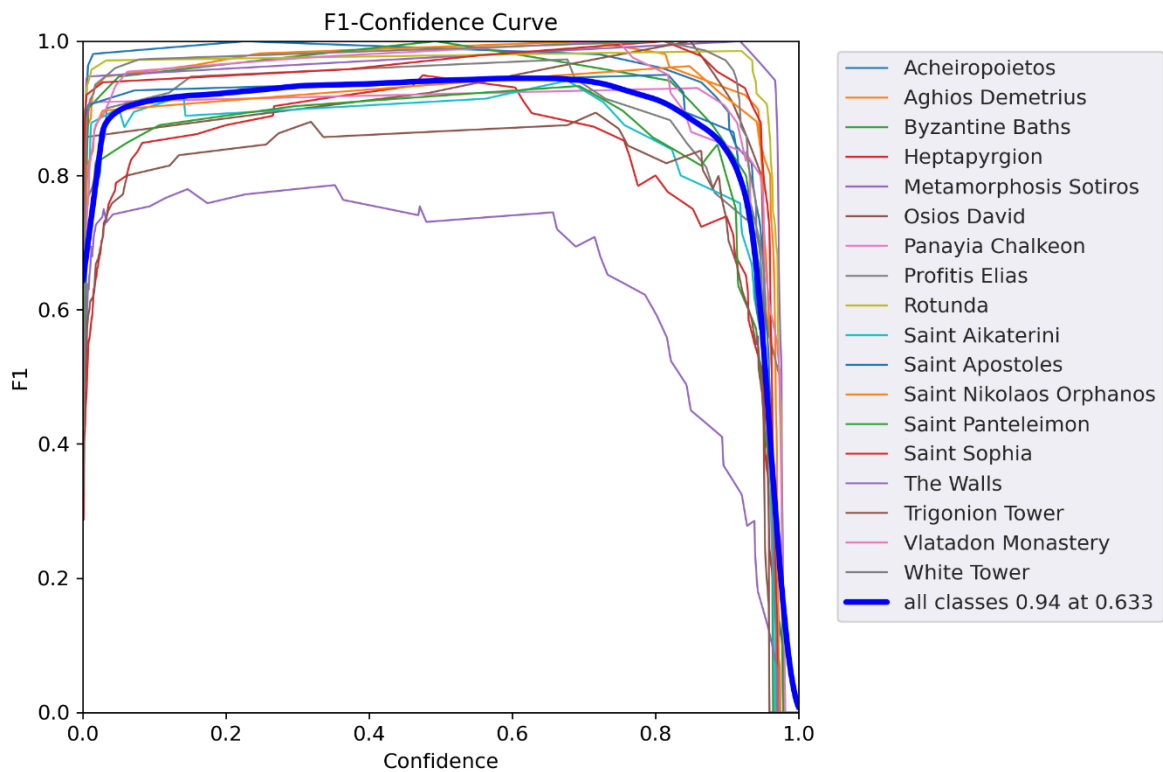
Τα αποτελέσματα τα οποία έδειξε το μοντέλο στα δεδομένα φαίνονται να είναι εξαιρετικά ικανοποιητικά στο εύρος του IoU 0.5...0.95, με τιμή mAP για κοντά στο 94% για την καλύτερη διαμόρφωση του μοντέλου και IoU > 0.95 με τιμή 0.76. Στην συνέχεια ακολουθούν τα διαγράμματα με τις σχετικές μετρικές της εκπαίδευσης: Confusion Matrix( Πίνακας 2) F1 Curve (Πίνακας 3) και τα αποτελέσματα κατά την χρονική εξέλιξη της εκπαίδευσης (Πίνακας 4). Τέλος παραδείγματα από την εφαρμογή του μοντέλου για την εξαγωγή αποτελεσμάτων σε άγνωστες εικόνες του test set.



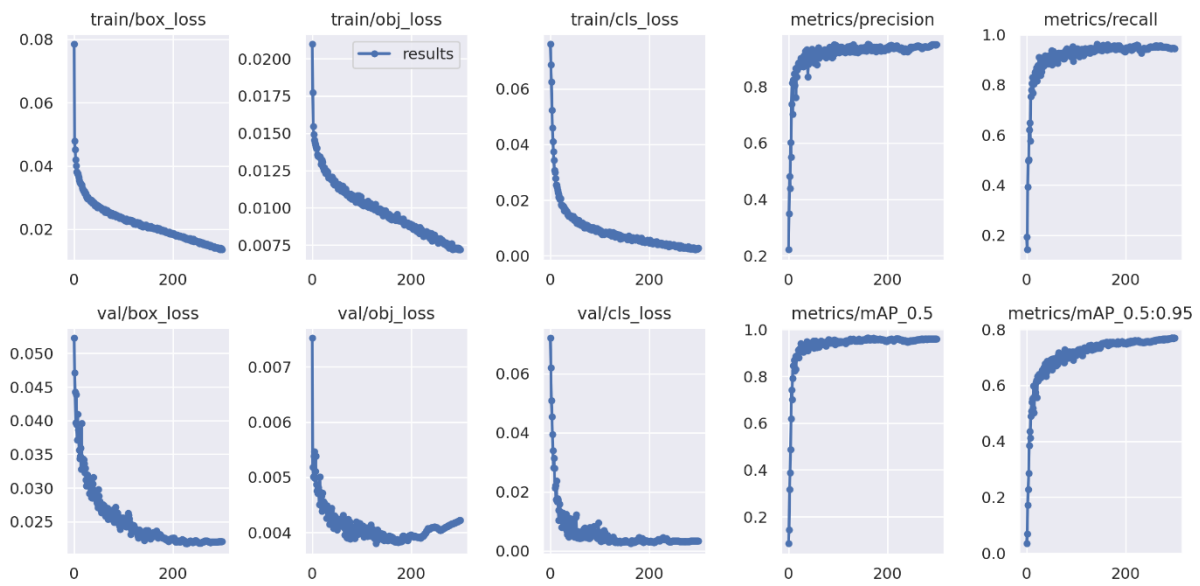
Πίνακας 2 Αποτελέσματα εκπαίδευσης - Confusion Matrix



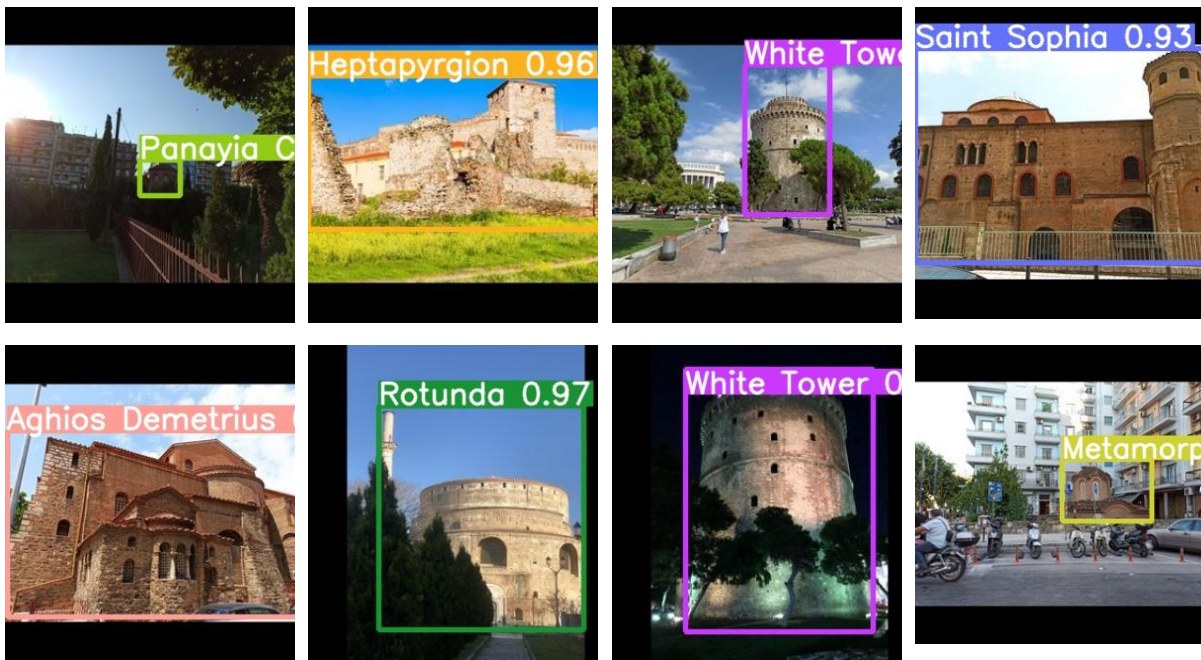
Πίνακας 3 Αποτελέσματα εκπαίδευσης - F1



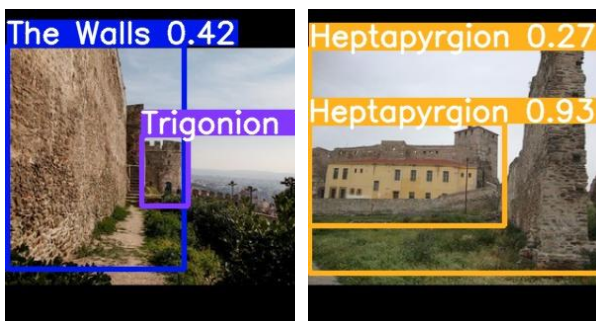
Πίνακας 4 Αποτελέσματα εκπαίδευσης - Χρονική Εξέλιξη Μετρικών



Εντοπίσεις των μνημείων από διαφορές οπτικές γωνίες, αποστάσεις και συνθήκες φωτισμού.



Πολλαπλές Εντοπίσεις στην ίδια εικόνα



Μερικός σωστή αναγνώριση



Σχήμα 5.4 Αποτελέσματα inference στο test set

Εκτός από την χρήση στατικών εικόνων το μοντέλο χρησιμοποιήθηκε και σε video με ικανοποιητικά αποτελέσματα ώστε να προσομοιάσουμε όσο το δυνατόν πιο καλά την χρήση της κάμερας και την εξαγωγή αποτελεσμάτων σε πραγματικό χρόνο.

### 5.2.3 Εκπαίδευση με το CreateML

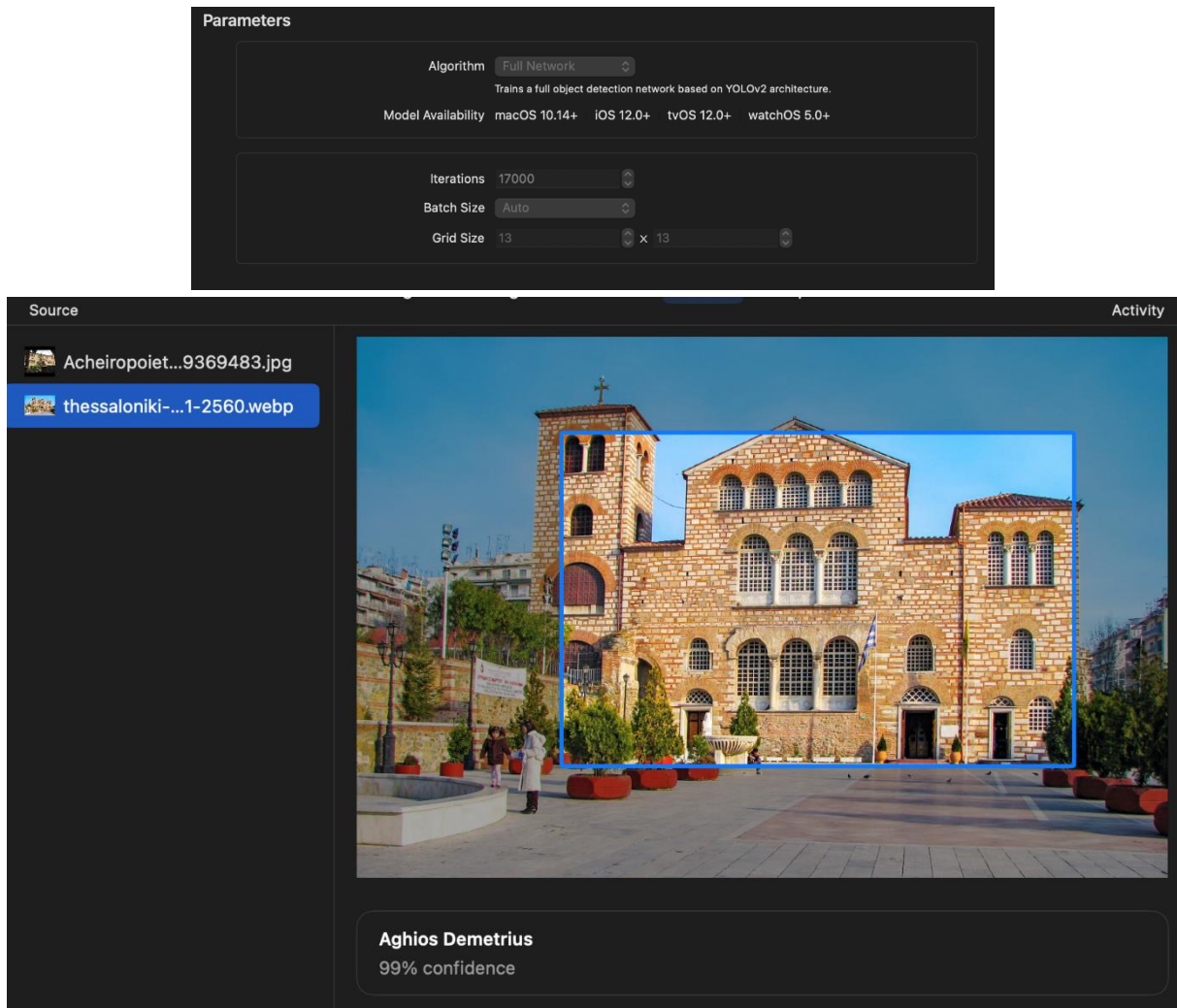
Για την δημιουργία ενός γρήγορου πρωτοτύπου για την χρήση για την δοκιμή από την εφαρμογή χρησιμοποιήθηκε με τα ίδια δεδομένα εκπαίδευσής και ένα μοντέλο με την χρήση του CreateML εργαλείου της Apple το οποίο είναι ενσωματωμένο στο Xcode IDE. Δίνει την δυνατότητα στον χρήστη με απλό τρόπο και μέσω UI να δημιουργήσει και να εκπαιδεύσει μοντέλα μηχανικής μάθησης (image classification, object detection, sound classification, sentiment analysis, tabular regression and classification) για άμεση εξαγωγή και χρήση από το Vision Framework στο περιβάλλον της Apple. Λόγω της ευκολίας στην εκπαίδευση και στην ταχύτητα της ενσωμάτωσης με την οποία συνεπάγεται η δημιουργία μοντέλων με την χρήση του εργαλείου αποτελεί μια πολύ καλή αρχή για proof of concepts και γρήγορη ανάπτυξη εφαρμογών.

Πίνακας 5 Σύγκριση των μοντέλων YOLO v2 και v5

YOLOv5	mAP50	mAP50-95	YOLOv2	mAP50	mAP50-95	Ποσοστιαία διαφορά	
<i>Acheiropoietos</i>	0.995	0.827	Acheiropoietos	0.81	0.31	0.186	0.625
<i>Aghios Demetrius</i>	0.995	0.841	Aghios Demetrius	0.81	0.41	0.186	0.512
<i>Byzantine Baths</i>	0.995	0.716	Byzantine Baths	0.2	0.07	0.799	0.902
<i>Heptapyrgion</i>	0.918	0.621	Heptapyrgion	0.6	0.2	0.346	0.678
<i>Metamorphosis Sotiros</i>	0.995	0.879	Metamorphosis Sotiros	0.86	0.52	0.136	0.408
<i>Osios David</i>	0.995	0.723	Osios David	1	0.52	-0.005	0.281
<i>Panayia Chalkeon</i>	0.938	0.73	Panayia Chalkeon	0.66	0.28	0.296	0.616
<i>Profitis Elias</i>	0.972	0.731	Profitis Elias	0.79	0.33	0.187	0.549
<i>Rotunda</i>	0.994	0.797	Rotunda	0.84	0.44	0.155	0.448
<i>Saint Aikaterini</i>	0.972	0.758	Saint Aikaterini	1	0.52	-0.029	0.314
<i>Saint Apostoles</i>	0.95	0.782	Saint Apostoles	0.75	0.43	0.211	0.450
<i>Saint Nikolaos Orphanos</i>	0.93	0.73	Saint Nikolaos Orphanos	0.68	0.37	0.269	0.493
<i>Saint Panteleimon</i>	0.937	0.601	Saint Panteleimon	0.81	0.23	0.136	0.617

<i>Saint Sophia</i>	0.995	0.857	Saint Sophia	0.91	0.39	0.085	0.545
<i>The Walls</i>	0.716	0.422	The Walls	0.49	0.2	0.316	0.526
<i>Trigonion Tower</i>	0.852	0.614	Trigonion Tower	0.65	0.27	0.237	0.560
<i>Vlatadon Monastery</i>	0.995	0.783	Vlatadon Monastery	0.81	0.41	0.186	0.476
<i>White Tower</i>	0.994	0.817	White Tower	0.94	0.61	0.054	0.253
<i>Average</i>						0.20838615	0.514193992

Το μοντέλο το οποίο δημιουργήθηκε με την χρήση του CreateML, είναι βασισμένο στο YOLOv2 το οποίο έχει αναλυθεί σε προηγούμενο κεφάλαιο. Τα αποτελέσματα του YOLOv2 συγκριτικά με το YOLOv5 είναι σχεδόν 20% καλύτερο στο mAP 0.5 και 50%. Παρ' όλα αυτά αποτέλεσε μια σημαντική βάση για να μπορέσει να δημιουργηθεί ένα πρωτότυπο για το proof of concept.



Εικόνα 5 Το περιβάλλον του Create ML

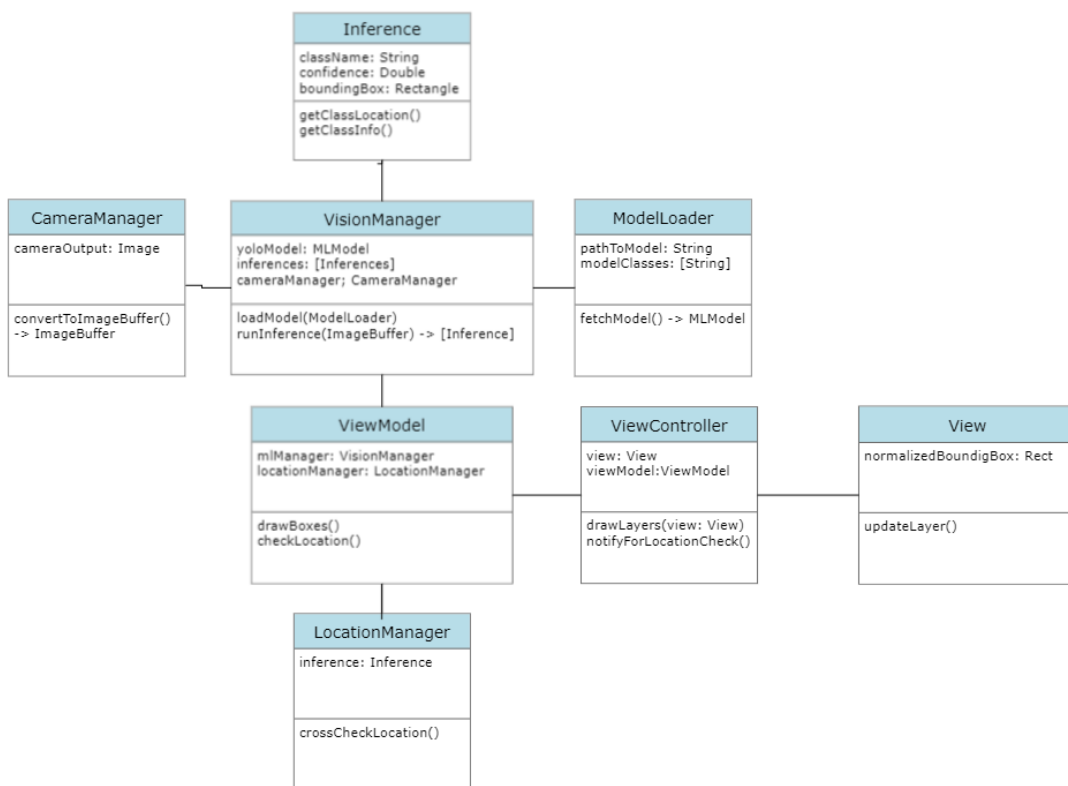
### 5.3 Εξαγωγή του μοντέλου για χρήση από την εφαρμογή

Για την εξαγωγή του μοντέλου μπορούν να χρησιμοποιηθούν διάφορες μέθοδοι για την βελτιστοποίηση σε χρήση συσκευών οι οποίες διαθέτουν περιορισμένες υπολογιστικές και χωρικές δυνατότητες.

Για τις χωρικές βελτιστοποιήσεις, στην συγκεκριμένη περίπτωση δεν εφαρμόστηκε κάποια από τις τεχνικές κβαντοποίησης του μοντέλου (Quantization, για μετατροπή από ακεραίους 32 bit σε ακεραίους 8 bit), καθώς το μέγεθος του μοντέλου ήταν 48 megabyte, το οποίο είναι απολύτως αποδεκτό για μια κινητή συσκευή. Αρχικά οι εξαγωγές έγιναν σε TensorFlow για την χρήση από την βιβλιοθήκη TensorFlowLite με ικανοποιητικά αποτελέσματα αλλά εμφάνιση καθυστερήσεων στο inference time που πολλές φορές μείωνε την απόδοση των λειτουργιών σε κάτω από 30 fps το οποίο καθιστούσε την λειτουργία της κάμερας και των προβλέψεων σε ζωντανό χρόνο μη δυνατό. Για αυτό για την βελτίωση των αποτελεσμάτων, έγινε εξαγωγή του μοντέλου σε CoreML, το native framework του iOS για την ενσωμάτωση μοντέλων σε εφαρμογές, το οποίο βελτίωσε δραματικά τον χρόνο εξαγωγής αποτελεσμάτων.

#### 5.4 Ενσωμάτωση στην εφαρμογή

Για την χρήση από την εφαρμογή απαιτείται η κατάλληλη αρχιτεκτονική για την δημιουργία μιας κλάσης η οποία θα φορτώνει το μοντέλο από τον δίσκο, θα λαμβάνει σαν είσοδο δεδομένα από την κάμερα του χρήστη, θα τα μετατρέπει τα δεδομένα στην μορφή η οποία απαιτείται από την είσοδο του μοντέλου, θα τρέχει ένα inference και ως αποτέλεσμα θα παράγει μια σειρά από 100 αποτελέσματα με τις προβλέψεις του για το αν υπάρχει κάποιο μνημείο ή όχι στην εικόνα η οποία συλλέχτηκε από την κάμερα του χρήστη. Τέλος σε περίπτωση που αναγνωριστεί ένα μνημείο ο location manager, κάνει cross check με την τοποθεσία του χρήστη ώστε να προσδιοριστεί αν ο χρήστης βρίσκεται όντως στο μνημείο και δεν έχει γίνει κάποιο λάθος στο εύρος εμπιστοσύνης του μοντέλου ή σκανάρει κάποια εικόνα στον υπολογιστή του.



Σχήμα 5.5 Διάγραμμα UML με την ροή πληροφορίας

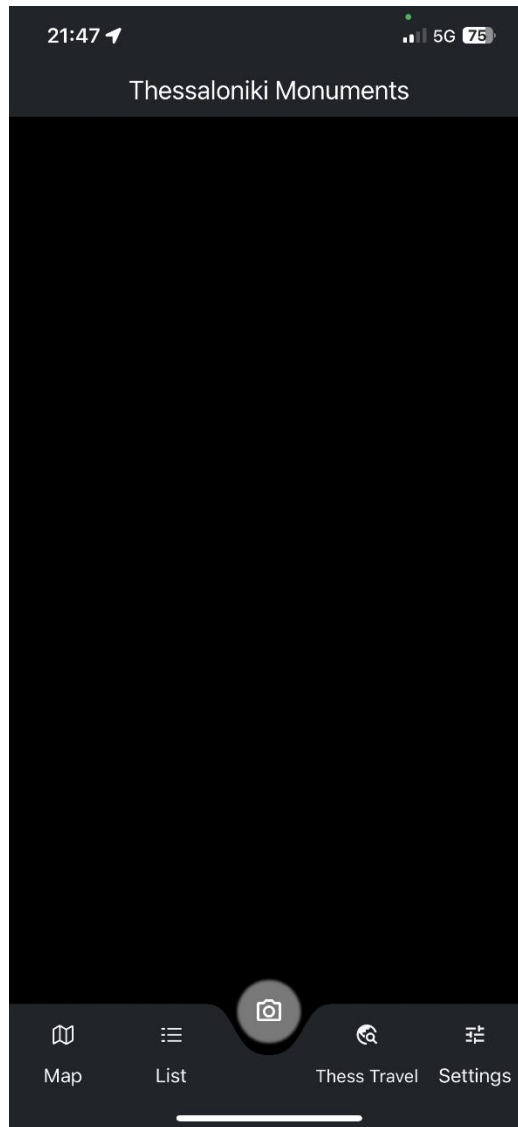
Η ροή πληροφορίας στην εφαρμογή βασίζεται στην αρχιτεκτονική του MVVM. Οι κλάσεις που είναι απαραίτητες για την λειτουργία της εφαρμογής περιγράφονται στο παραπάνω διάγραμμα UML.



## 5.5 Ανάπτυξη του UI της εφαρμογής για την αξιοποίηση του μοντέλου

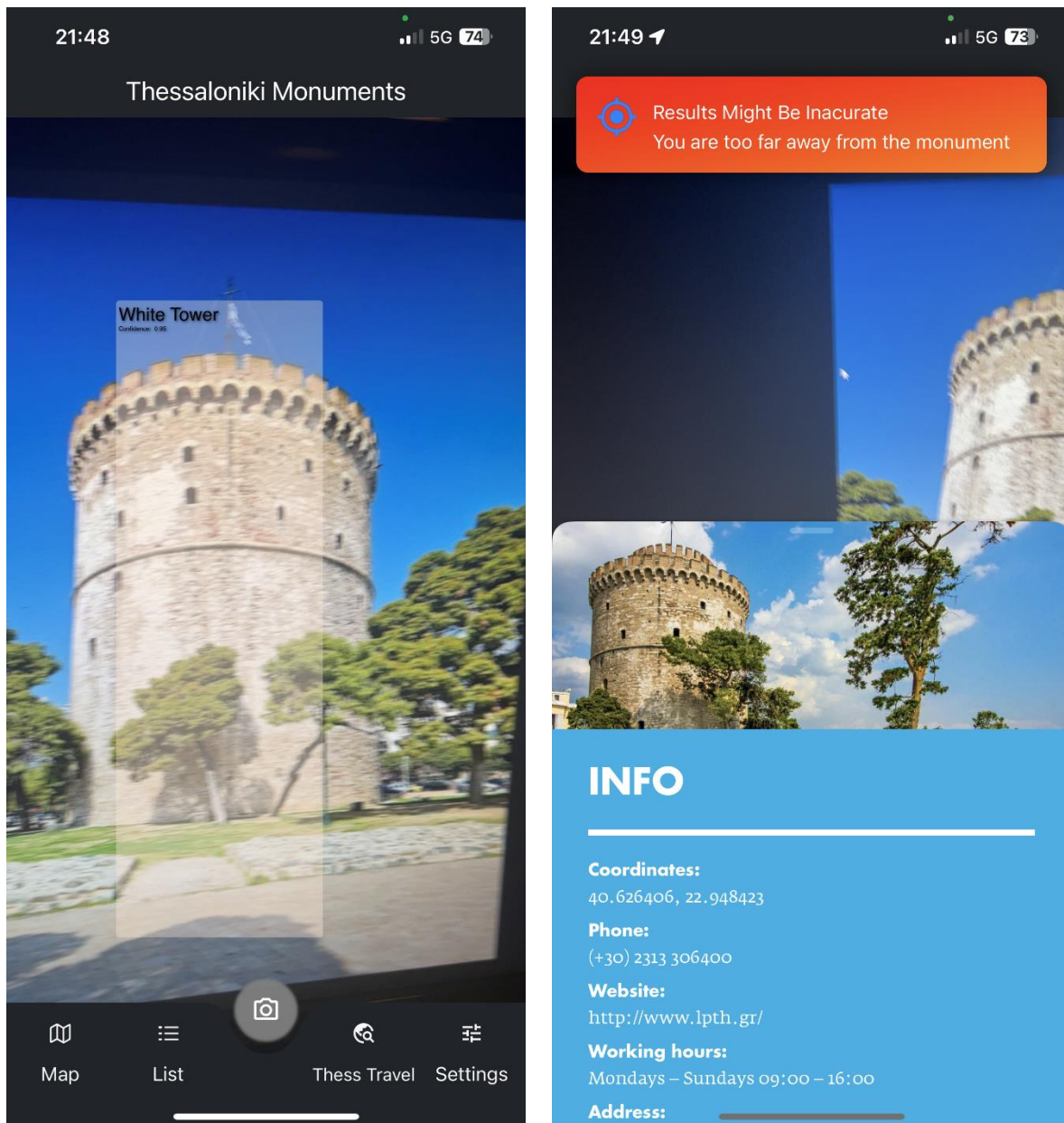
Η εφαρμογή iOS προσφέρει μια ολοκληρωμένη λύση για την τουριστική καθοδήγηση συνδυάζοντας την αναγνώριση αντικειμένων βάσει ML, τις υπηρεσίες που βασίζονται στην τοποθεσία και ένα UI που ακολουθεί το HIG.

Το bottom tab bar αποτελεί ένα οικείο σύστημα πλοήγησης για τους χρήστες του iOS. Αυτή η διάταξη επιτρέπει την εύκολη πρόσβαση και την απρόσκοπτη μετάβαση μεταξύ των τριών κύριων λειτουργιών της εφαρμογής: του Camera View Controller, του List Info View Controller και του Map View Controller.



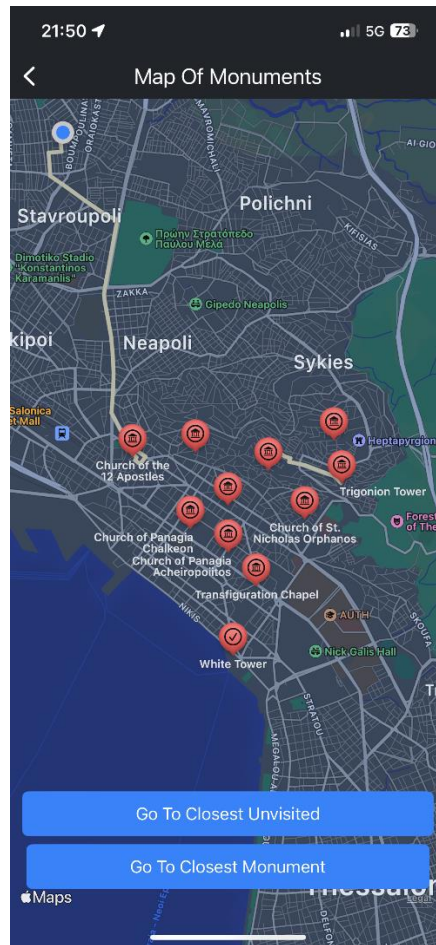
Εικόνα 6 Tab Bar Navigation

**Camera View Controller:** χρησιμεύει ως παράθυρο στον κόσμο σε πραγματικό χρόνο για τους χρήστες της εφαρμογής. Χρησιμοποιεί την τροφοδοσία της κάμερας της συσκευής ως είσοδο για τον ενσωματωμένο διαχειριστή ML, ο οποίος με τη σειρά του εκτελεί τις εργασίες αναγνώρισης αντικειμένων. Όταν ένας χρήστης στρέφει την κάμερα προς ένα πιθανό σημείο ενδιαφέροντος (POI), όπως ένα από τα διαθέσιμα μνημεία, ο αλγόριθμος ML αναγνωρίζει το αντικείμενο, παρέχοντας πληροφορίες ή ενδιαφέροντα στοιχεία για το εντοπισμένο POI. Αυτή η real time πληροφορία βελτιώνει την εμπειρία του χρήστη και παρέχει καθηλωτική αλληλεπίδραση με το περιβάλλον του.



Εικόνα 7 Camera View Controller, με τις λειτουργίες αναγνώρισης και εμφάνισης άμεσης πληροφορίας

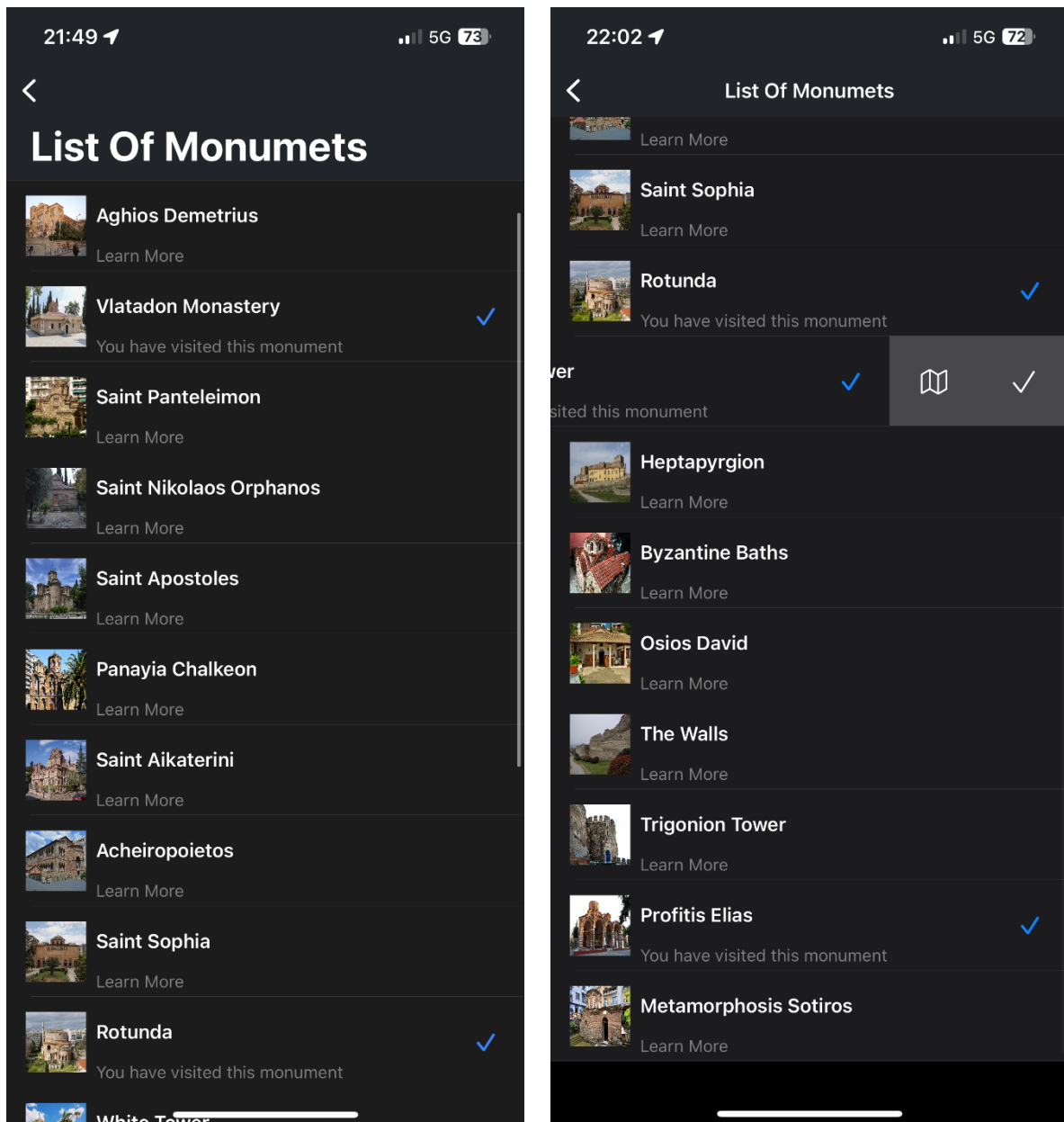
**Map View Controller:** αξιοποιεί τη δύναμη της τεχνολογίας GPS και τοποθετεί τα διαθέσιμα μνημεία σε έναν χάρτη. Ενσωματώνει επίσης ένα module καθοδήγησης, το οποίο, μετά την επιλογή ενός μνημείου από τον χρήστη, παρέχει οδηγίες πλοήγησης προς την επιθυμητή τοποθεσία. Παρουσιάζει μια αποτελεσματική λύση στις προκλήσεις προσανατολισμού που συχνά αντιμετωπίζουν οι τουρίστες σε άγνωστες τοποθεσίες. Η οπτική αναπαράσταση των μνημείων και των τοποθεσιών τους εντός της εφαρμογής βοηθά τους χρήστες στον σχεδιασμό διαδρομών και στην κατανόηση του γεωγραφικού πλαισίου κάθε POI.



Εικόνα 8 Map View Controller , με όλες τις πινέξες-μνημεία τα οποία μπορεί να επισκεφτεί ο χρήστης

**List Info View Controller:** παρουσιάζει έναν ολοκληρωμένο κατάλογο όλων των μνημείων που μπορεί να αναγνωρίσει η εφαρμογή. Αυτό χρησιμεύει όχι μόνο ως κατάλογος πιθανών προορισμών για τον χρήστη, αλλά και ως διεπαφή γρήγορης πρόσβασης σε σχετικές πληροφορίες για κάθε μνημείο. Μέσω της λίστας, ο χρήστης μπορεί να επιλέξει ένα μνημείο, προτρέποντας την εφαρμογή να εμφανίσει λεπτομερείς πληροφορίες, όπως ιστορικά στοιχεία, ώρες επίσκεψης και άλλες σχετικές λεπτομέρειες. Η λειτουργία αυτή προσφέρει στους χρήστες έναν εύκολο και οργανωμένο τρόπο για να προγραμματίζουν τις περιηγήσεις τους στα αξιοθέατα.

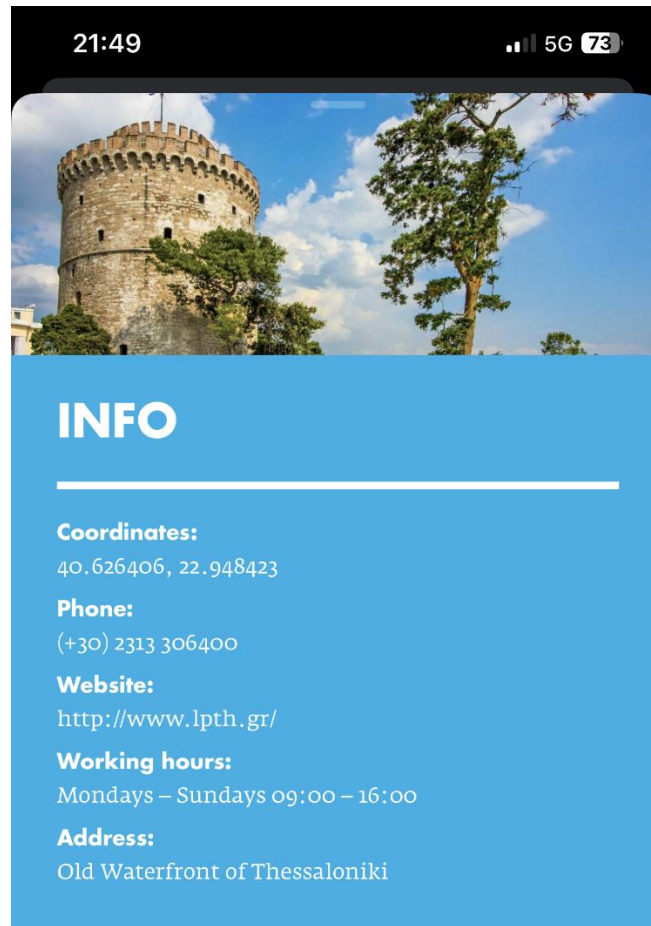




Εικόνα 9 ListViewController με τις λειτουργίες για άμεση πλοήγηση, επιβεβαίωση επίσκεψης, και παρουσίασης περισσότερων πληροφοριών με επιλογή ενός μνημείου

Προσθέτοντας ένα επιπλέον επίπεδο διαδραστικότητας και λεπτομέρειας στην εφαρμογή, ο **Info Details View Controller** είναι ένα στοιχείο, σχεδιασμένο για να βελτιώσει περαιτέρω την εμπειρία του χρήστη. Αυτή η λειτουργία είναι εγγενώς συνδεδεμένη με τον List Info View Controller και τον Camera View Controller. Ενεργοποιείται όταν ο χρήστης πατήσει σε ένα συγκεκριμένο μνημείο από τη λίστα ή αν το ανιχνεύσει με την κάμερα και σύρει το παράθυρο με τις πληροφορίες στο μέγιστο δυνατό σημείο. Προσφέρει πληροφορίες πέραν της γενικής επισκόπησης που είναι διαθέσιμη στο List Info View Controller, παρέχοντας ολοκληρωμένα δεδομένα, όπως λεπτομερές ιστορικό πλαίσιο, αρχιτεκτονική σημασία και μοναδικά χαρακτηριστικά του μνημείου. Οι λεπτομερείς παρέχονται μέσω τεχνικής Web Scrapping από το Website του οργανισμού Θεσσαλονίκης και φορτώνονται σαν WebViews δίνοντας

την δυνατότητα το περιεχόμενο της σελίδας να αλλάζει δυναμικά σε περίπτωση που π.χ. αλλάξει το ωράριο λειτουργίας ενός μνημείου.



**M**eeet Thessaloniki's impressive **history** by entering the city's most emblematic structure, the imposing White Tower and enjoy the lovely view from the roof top! The permanent exhibition of White Tower displays the city's history in a concise way, from its foundation period in 316/15 BC until today, via multimedia applications,

Εικόνα 10 Info Details View Controller

Εκτός από τους κύριους controllers, η εφαρμογή ενσωματώνει επίσης πρόσθετες βοηθητικές προβολές για μια πιο ολοκληρωμένη εμπειρία χρήσης. Μια τέτοια προβολή είναι ο **SettingsViewController**, ο οποίος παρέχει στους χρήστες επιλογές προσαρμογής για την εξατομίκευση της αλληλεπίδρασής τους με την εφαρμογή. Τέλος ο **WebViewController** που περιέχει πληροφορίες σχετικά με τον Οργανισμό Τουρισμού Θεσσαλονίκης. Αυτή η προβολή προσφέρει ένα ενσωματωμένο πρόγραμμα περιήγησης στο

διαδίκτυο μέσα στην εφαρμογή, κατευθύνοντας τους χρήστες στην επίσημη ιστοσελίδα του οργανισμού. Παρέχει εκτενείς πληροφορίες σχετικά με την αποστολή, τις πρωτοβουλίες, τις εκδηλώσεις και τις υπηρεσίες του Οργανισμού Τουρισμού Θεσσαλονίκης.

## Κεφάλαιο 6ο: Αξιολόγηση

### 6.1 Περιπτώσεις Χρήσης– Use Cases

Το σύστημα το οποίο δημιουργήθηκε είναι μια βάση για να αναπτυχθούν περαιτέρω εφαρμογές και λειτουργίες επάνω σε αυτό προσδίδοντας αξία στις εμπειρίες των επισκεπτών της Θεσσαλονίκης. Μερικές περιπτώσεις χρήσης τις οποίες μπορεί να καλύψει είναι οι παρακάτω:

1. Ιστορική εκπαίδευση και εξερεύνηση: Η εφαρμογή iOS μπορεί να χρησιμοποιηθεί ως εκπαιδευτικό εργαλείο τόσο για τους τουρίστες όσο και για τους ντόπιους που ενδιαφέρονται για την ιστορία της Θεσσαλονίκης. Χρησιμοποιώντας την κάμερα για τον εντοπισμό ιστορικών μνημείων και τοποθεσιών, οι χρήστες μπορούν να μάθουν για το πλούσιο παρελθόν της πόλης. Για παράδειγμα, ένας χρήστης που επισκέπτεται τον Λευκό Πύργο μπορεί να στρέψει την κάμερά του στο μνημείο για να μάθει για την ιστορική του σημασία, τον αρχιτεκτονικό του ρυθμό και τα κομβικά γεγονότα που έχει παρακολουθήσει.
2. Η εφαρμογή μπορεί να βοηθήσει τους χρήστες στο **σχεδιασμό εξατομικευμένων διαδρομών** περιήγησης στα αξιοθέατα. Ο ελεγκτής προβολής List Info παρέχει έναν ολοκληρωμένο κατάλογο μνημείων και σημείων ενδιαφέροντος στη Θεσσαλονίκη. Οι χρήστες μπορούν να επιλέξουν τα σημεία ενδιαφέροντος τους, να τα χαρακτηρίσουν ως προς επίσκεψη ή ότι τα έχουν ήδη επισκεφτεί και να σχεδιάσουν τη διαδρομή τους χρησιμοποιώντας τον Map View Controller.
3. Καθοδήγηση: Ένας χρήστης που δεν είναι εξοικειωμένος με την πόλη μπορεί να χρησιμοποιήσει τον Map View Controller για να πλοηγηθεί στα διάφορα σημεία ενδιαφέροντος. Για παράδειγμα, ένας χρήστης που μόλις έφτασε στο σιδηροδρομικό σταθμό της Θεσσαλονίκης μπορεί να επιλέξει ένα μνημείο από τη λίστα και η εφαρμογή θα του παρέχει οδηγίες προς τον επιλεγμένο προορισμό του.
4. Οι χρήστες που ενδιαφέρονται να ενημερωθούν για τις τουριστικές πρωτοβουλίες και εκδηλώσεις της Θεσσαλονίκης μπορούν να χρησιμοποιήσουν το ενσωματωμένο WebSite στο WebView για να αποκτήσουν πρόσβαση στην επίσημη ιστοσελίδα του Οργανισμού Τουρισμού Θεσσαλονίκης. Αυτό μπορεί να τους βοηθήσει να ενημερωθούν για τρέχουσες ή επερχόμενες εκδηλώσεις, και άλλες υπηρεσίες που σχετίζονται με τον τουρισμό, όπως και να βρουν στοιχεία επικοινωνίας και άλλες υπηρεσίες που σχετίζονται με τον τουρισμό.

### 6.2 Αποτελέσματα

Αποτέλεσμα της διπλωματικής εργασίας αυτής είναι ένα ολοκληρωμένο σύστημα το οποίο βασίζεται σε 3 βασικά υποσυστήματα, το σύστημα εκπαίδευσης και εξαγωγής του μοντέλου ML, την διαδικασία εξαγωγής του μοντέλου και την ενσωμάτωση στην εφαρμογή και τέλος στην σχεδίαση και την ανάπτυξη των λειτουργιών και του UI της εφαρμογής. Η συνεργασία και ο συνδυασμός αυτών των 3 συστημάτων έχει ως αποτέλεσμα την δημιουργία μιας χρηστικής και ελκυστικής εφαρμογής, ικανής να κερδίσει το ενδιαφέρον του χρήστη, να του δώσει πληροφορίες, έναν καινούργιο τρόπο να εξερευνήσει τα μνημεία UNESCO της Θεσσαλονίκης και να ενισχύσει αυτή την πολιτιστική κληρονομιά μέσω της τεχνολογίας. Τέλος στόχος παραμένει η παρουσίαση μιας εφαρμογής για κινητά τηλέφωνα στο ευρύ κοινό, με την μορφή ενός ολοκληρωμένου προϊόντος, με την δημοσίευση της εφαρμογής στο AppStore.

## Κεφάλαιο 7ο: Συμπεράσματα και προτάσεις βελτίωσης

### 7.1 Το μέλλον της τεχνολογίας στον τουρισμό

Συνοψίζοντας, το μέλλον της τεχνολογίας στον τουρισμό είναι γεμάτο προκλήσεις και ευκαιρίες. Οι τουρίστες, οι επιχειρήσεις και οι ταξιδιωτικοί προορισμοί, μέσω των εκπροσώπων τους, θα πρέπει να προσαρμόζονται συνεχώς στις νέες τάσεις και να αξιοποιούν την τεχνολογία για να βελτιώσουν την ποιότητα στον τουρισμό. Με την έλευση της Τέταρτης Βιομηχανικής Επανάστασης, η ενσωμάτωση της τεχνητής νοημοσύνης (AI) στις εφαρμογές κινητής τηλεφωνίας έχει γίνει όλο και πιο διαδεδομένη. Τα συστήματα τεχνολογίας γίνονται ολοένα πιο «έξυπνα» και πολύπλοκα, πολλές φορές αφαιρώντας από τον χρήστη την χαρά της εξερεύνησης ή της ανακάλυψης η οποία αποτελεί εγγενές κομμάτι της τουριστική εμπειρίας. Είναι σημαντικό να επισημάνουμε ότι η τεχνολογία θα πρέπει να ενισχύει την εμπειρία του χρήστη και να μην την υπερκαλύπτει. Οι τεχνολογίες για κινητά τηλέφωνα οπου διατηρούν μια ισορροπία στο κομμάτι της εμπύθισης και της εξάρτησης στην τεχνολογία αποτελούν ερεθίσματα για μια απολαυστική εμπειρία.

### 7.2 Πρόσθετη μελέτη, προτάσεις για το μέλλον

Οι νέες τεχνολογίες οι οποίες δίνουν την δυνατότητα στις τεχνολογίες VR και το metaverse φαίνονται ως πολλά υποσχόμενες για το πως θα διαμορφώσουν το μέλλον του τουρισμού και την επιρροή της τεχνολογίας επάνω του. Οι μέτοχοι στην τουριστική βιομηχανία θα πρέπει να το λάβουν σοβαρά υπόψιν και να προσπαθήσουν να χτίσουν πάνω σε αυτές τις αναπτυσσόμενες, μέρα με την μέρα, τεχνολογίες καθώς η βιομηχανία μετά την πανδημία του COVID 19 έχει εντείνει την ανάγκη επίλυσης των ανθρώπινων προβλημάτων μέσω της τεχνολογίας, και ο τουρισμός δεν αποτελεί εξαίρεση.

Επίσης εξελίξεις στους αλγορίθμους μηχανικής μάθησης για τον εντοπισμό δεδομένων συμβαίνουν συνεχώς, με νέες τεχνικές και μεθόδους που βελτιώνουν τα αποτελέσματα της εξαγωγής αποτελεσμάτων, μειώνοντας το περιθώριο λάθους. Η αρχιτεκτονική της εφαρμογής δίνει την δυνατότητα της αξιοποίησης τους με μικρές αλλαγές στον κώδικα της εφαρμογής χωρίς μεγάλο τεχνικό κόστος.

Η πρώτη πρόταση για μελλοντική χρήση μπορούσε να εξετάσει την εισαγωγή ενός συστήματος υποβοήθησης προσβασιμότητας με στόχο να μετατραπεί το σύστημα σε ένα εργαλείο για χρήστες με προβλήματα ακοής ή όρασης. Για παράδειγμα Με την λειτουργία voice over, επιτρέπει στους χρήστες αυτούς να πάρουν πληροφορίες και βοηθηθούν στην ξενάγηση τους και την μετακίνηση τους από μνημείο σε μνημείο.

Η Δεύτερη πρόταση είναι η εισαγωγή ενός κοινωνικού χαρακτήρα στην εφαρμογή, για παράδειγμα εάν κάποιος χρήστης βρήκε ενδιαφέρουσα την επίσκεψη του να μπορεί να μοιραστεί πληροφορίες με φίλους ή την οικογένεια του μέσω της εφαρμογής. Αυτό καθιστά την εφαρμογή ένα εργαλείο κοινωνικής δικτύωσης, που διαδίδει την ευαισθητοποίηση σχετικά με την πλούσια πολιτιστική κληρονομιά της Θεσσαλονίκης.

Τρίτη πρόταση είναι η εισαγωγή ενός χαρακτήρα gamification στην εφαρμογή δίνοντας, για παράδειγμα με την μορφή ενός παιχνιδιού συμπλήρωσης όλων των μνημείων ώστε ο χρήστης να κερδίσει ένα ψηφιακό βραβείο, και να δώσει ένα κίνητρο για την ολοκλήρωση της περιήγησης σε έναν X αριθμό μνημείων

## ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] H. Zhang, S. Gordon, D. Buhalis and X. Ding, "Experience Value Cocreation on Destination Online Platforms," *Journal of Travel Research*, vol. 57, no. 8, pp. 1093-1107, 11 2018.
- [2] Z. Xiang, Q. Du, Y. Ma and W. Fan, "A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism," *Tourism Management*, vol. 58, pp. 51-65, 2 2017.
- [3] D. A. Guttentag, "Virtual reality: Applications and implications for tourism," *Tourism Management*, vol. 31, no. 5, pp. 637-651, 10 2010.
- [4] C. D. Kounavis, A. E. Kasimati and E. D. Zamani, "Enhancing the Tourism Experience through Mobile Augmented Reality: Challenges and Prospects," *International Journal of Engineering Business Management*, vol. 4, p. 10, 1 2012.
- [5] U. G. U. Stankov, "Tourism 4.0 technologies and tourist experiences: a human-centered design perspective," *Inf Technol Tourism*, vol. 22, p. 477–488, 2020.
- [6] P. Taylor, "statista," 31 May 2023. [Online]. Available: <https://www.statista.com/statistics/272698/global-market-share-held-by-mobile-operating-systems-since-2009/>. [Accessed 17 June 2023].
- [7] Statista Research Department, "statista," Statista Research Department, 12 June 2023. [Online]. Available: <https://www.statista.com/topics/962/global-tourism/#topicOverview>. [Accessed 23 June 2023].
- [8] U. G. U. Stankov, "Tourism 4.0 technologies and tourist experiences: a human-centered design perspective," *Inf Technol Tourism* 22, p. 477–488, 2020.
- [9] "National Institute of Standards and Technology," Information Technology Laboratory / Information Access Division, 3 May 2021. [Online]. Available: <https://www.nist.gov/itl/iad/visualization-and-usability-group/human-factors-human-centered-design>. [Accessed 18 June 2023].
- [10] International Design Foundation, "Beyond AR vs. VR: What is the Difference between AR vs. MR vs. VR vs. XR?," May 2023. [Online]. Available: <https://www.interaction-design.org/literature/article/beyond-ar-vs-vr-what-is-the-difference-between-ar-vs-mr-vs-vr-vs-xr>. [Accessed 15 June 2023].
- [11] Z. Xiang and U. Gretzel, "Role of social media in online travel information search," 4 2010. [Online].
- [12] A. L. M. A. O. V. P. Garcia, "Intelligent Routing System for a Personalised Electronic Tourist Guide," *Information and Communication Technologies in Tourism 2009*. Springer, 2009.

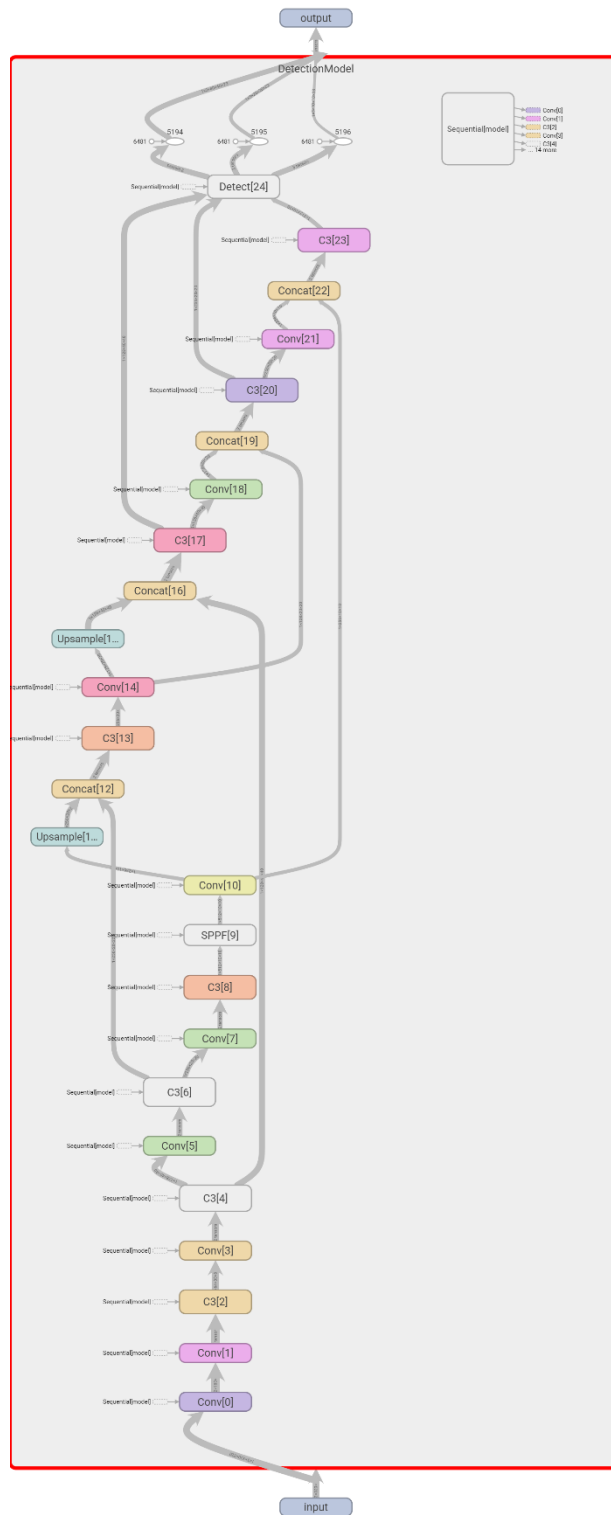
- [13] Microsoft, "Smart cities: The cities of the future," Microsoft, [Online]. Available: <https://www.microsoft.com/en-us/industry/government/resources/smart-cities>. [Accessed 16 June 2023].
- [14] P. K. S. B. Y. M. S. G. H. C. I.-H. R. Saurabh Singh, "Convergence of blockchain and artificial intelligence in IoT network for the sustainable smart city," *Sustainable Cities and Society*, vol. Volume 63, no. 102364, 2020.
- [15] G. U. D. B. Koo C, "Special issue on smart tourism cities," *Asia Pacific Journal of Tourism Research*, vol. 26, no. 4, pp. 349-351, 2021.
- [16] U. S. M. X. Z. Gretzel, "Smart tourism: foundations and developments," *Electron Markets*, pp. 179-188, 2015.
- [17] IBM, "What is machine learning?," IBM, [Online]. Available: <https://www.ibm.com/topics/machine-learning>. [Accessed 20 May 2023].
- [18] S. Brown, "Machine learning, explained," 21 April 2021. [Online]. Available: <https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained>. [Accessed 20 May 2023].
- [19] IBM, "AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What's the Difference?," 7 May 2020. [Online]. Available: <https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks><https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks>. [Accessed 20 June 2023].
- [20] O. Afolabi, "Neural Networks vs. Deep Learning: How Are They Different?," 9 March 2023. [Online]. Available: <https://www.makeuseof.com/neural-network-vs-deep-learning-are-they-different/>. [Accessed 15 June 2023].
- [21] S. B. Yoram Reich, "Evaluating machine learning models for engineering problems," *Artificial Intelligence in Engineering*, vol. Volume 13, no. Issue 3, 199.
- [22] K. Nighania, "Various ways to evaluate a machine learning model's performance," 30 Decemer 2018. [Online]. Available: <https://towardsdatascience.com/various-ways-to-evaluate-a-machine-learning-models-performance-230449055f15>. [Accessed 15 June 2023].
- [23] L. L. W. L. Q. G. Z. D. Z. X. Yunji Chen, *AI Computing Systems*, 2023.
- [24] Zach, "How to Calculate F1 Score in R (Including Example)," 8 September 2021. [Online]. Available: <https://www.statology.org/f1-score-in-r/>. [Accessed 18 June 2023].
- [25] D. C.-E. Juan Terven, "A Comprehensive Review of YOLO: From YOLOv1 and Beyond," 2023.
- [26] P. D. R. G. K. H. B. H. S. B. Tsung-Yi Lin, "Feature Pyramid Networks for Object Detection".
- [27] A. B. a. C.-Y. W. a. H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," 2020.
- [28] D. H. a. K. Gimpe, "Gaussian error linear units (gelus)," *arXiv preprint*, 2016.

- [29] "YOLOv7 by MMYOLO," 2023. [Online]. Available: <https://github.com/open-mmlab/mmyolo/tree/main/configs/yolov7>. [Accessed 22 June 2023].
- [30] A. B. H.-Y. M. L. Chien-Yao Wang, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object," *arXiv*, 2022.
- [31] Apple Inc., "Planning your iOS app," Apple Inc, 2023. [Online]. Available: <https://developer.apple.com/ios/>. [Accessed 10 May 2023].
- [32] Swift.org, "Swift.org," 2023. [Online]. Available: <https://www.swift.org/>. [Accessed 10 May 2023].
- [33] N. Heer, "Speed comparison of programming languages," github, 2023. [Online]. Available: <https://github.com/niklas-heer/speed-comparison>. [Accessed 11 May 2023].
- [34] vapor, "github.com/vapor," 2023. [Online]. Available: <https://github.com/vapor/vapor>.
- [35] tensorflow.org, "Swift for TensorFlow (In Archive Mode)," 2021. [Online]. Available: <https://www.tensorflow.org/swift/guide/overview>. [Accessed 11 May 2023].
- [36] Nona Digital, "iOS Human Interface Guidelines," 18 December 2017. [Online]. Available: <https://medium.com/nona-web/ios-human-interface-guideline-f012a5ba962e>. [Accessed 11 May 2023].
- [37] Microsoft, "Object-Oriented programming," [Online]. Available: <https://learn.microsoft.com/en-us/dotnet/csharp/fundamentals/tutorials/oop>. [Accessed 1 June 2023].
- [38] G. K. Arora, SOLID Principles, Morrisville, NC 27560: Syncfusion Inc., 2016.
- [39] "What's a design pattern?," [Online]. Available: <https://refactoring.guru/design-patterns/what-is-pattern>. [Accessed 1 June 2023].
- [40] R. C. M. (. Bob), "The Clean Architecture," 13 August 2012. [Online]. Available: <https://blog.cleancoder.com/uncle-bob/2012/08/13/the-clean-architecture.html>. [Accessed 23 June 2023].
- [41] Alumentations, "Alumentations documentation," [Online]. Available: <https://alumentations.ai/docs/>. [Accessed 23 June 2023].
- [42] S. W. Glenn Jocher, "Comprehensive Guide to Ultralytics YOLOv5," 29 March 2023. [Online]. Available: <https://docs.ultralytics.com/yolov5/>. [Accessed 23 June 2023].





# ΠΑΡΑΡΤΗΜΑ Α : Αρχιτεκτονική του μοντέλου



## ΠΑΡΑΡΤΗΜΑ Β : Υπερπαράμετροι του Yolov5

```
lr0: 0.01 # initial learning rate (SGD=1E-2, Adam=1E-3)
lrf: 0.01 # final OneCycleLR learning rate (lr0 * lrf)
momentum: 0.937 # SGD momentum/Adam beta1
weight_decay: 0.0005 # optimizer weight decay 5e-4
warmup_epochs: 3.0 # warmup epochs (fractions ok)
warmup_momentum: 0.8 # warmup initial momentum
warmup_bias_lr: 0.1 # warmup initial bias lr
box: 0.05 # box loss gain
cls: 0.5 # cls loss gain
cls_pw: 1.0 # cls BCELoss positive_weight
obj: 1.0 # obj loss gain (scale with pixels)
obj_pw: 1.0 # obj BCELoss positive_weight
iou_t: 0.20 # IoU training threshold
anchor_t: 4.0 # anchor-multiple threshold
# anchors: 3 # anchors per output layer (0 to ignore)
fl_gamma: 0.0 # focal loss gamma (efficientDet default gamma=1.5)
hsv_h: 0.015 # image HSV-Hue augmentation (fraction)
hsv_s: 0.7 # image HSV-Saturation augmentation (fraction)
hsv_v: 0.4 # image HSV-Value augmentation (fraction)
degrees: 0.0 # image rotation (+/- deg)
translate: 0.1 # image translation (+/- fraction)
scale: 0.5 # image scale (+/- gain)
shear: 0.0 # image shear (+/- deg)
perspective: 0.0 # image perspective (+/- fraction), range 0-0.001
flipud: 0.0 # image flip up-down (probability)
fliplr: 0.5 # image flip left-right (probability)
mosaic: 1.0 # image mosaic (probability)
mixup: 0.0 # image mixup (probability)
copy_paste: 0.0 # segment copy-paste (probability), κτλ.
```