



Μη Επανδρωμένο Αεροσκάφος (Μ.Ε.Α.) με Δυνατότητα Αναγνώρισης Στόχων για Χρήση από τις Ένοπλες Δυνάμεις



Μάριος Γρ. Μεσσήνης - Φίλιππας



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ
ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



Η Τριμελής Εξεταστική Επιτροπή

-Η-

Επιβλέπουσα
Καθηγήτρια

Ειρήνη Καρανάσιου

Καθηγήτρια

1^ο Μέλος

Ελευθέριος Δοϊτσίδης

ELEFTHERIOS
DOITSIDIS
01.07.2025 09:37

Αναπληρωτής
Καθηγητής

2^ο Μέλος

Μιχαήλ Γαργαλάκος

Αναπληρωτής
Καθηγητής



Περίληψη

Η παρούσα πτυχιακή εργασία ασχολείται με τη σχεδίαση ενός Μη Επανδρωμένου Αεροσκάφους (Μ.Ε.Α.) για την υποστήριξη επίγειων στρατιωτικών δυνάμεων. Ο στόχος είναι η έγκαιρη αναγνώριση θερμικών αποτυπωμάτων τεθωρακισμένων αρμάτων μάχης μέσω αλγορίθμου μηχανικής μάθησης και η ειδοποίηση των στρατιωτικών μονάδων. Η εργασία χωρίζεται σε δύο βασικά μέρη. Το πρώτο αφορά τις ασύρματες επικοινωνίες και περιλαμβάνει ανάλυση των δικτύων επικοινωνίας Μ.Ε.Α. και των VHF ασυρμάτων σταθμών, με έμφαση στη χρήση αναπηδούμενων συχνοτήτων. Το δεύτερο μέρος επικεντρώνεται στην αναγνώριση θερμικών αποτυπωμάτων, αναλύοντας τις αρχές της υπολογιστικής όρασης, την εκπαίδευση του αλγορίθμου YOLOv10 και την αξιολόγηση των αποτελεσμάτων. Η εργασία ολοκληρώνεται με τη σύνοψη των συμπερασμάτων και προτάσεις για μελλοντική έρευνα.

Λέξεις-κλειδιά: Μη Επανδρωμένο Αεροσκάφος, Τεχνητή Νοημοσύνη, Υπολογιστική Όραση, Θερμικά Αποτυπώματα, YOLOv10, Ασύρματες Επικοινωνίες, VHF, Στρατιωτικές Εφαρμογές.



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



Abstract

This master thesis focuses on designing an Unmanned Aerial Vehicle (UAV) to support ground military forces. The objective is the early detection of thermal signatures of armored battle tanks using a machine learning algorithm and the timely notification of military units. The study is divided into two main parts. The first part addresses wireless communications, analyzing UAV communication networks and VHF radio stations, with an emphasis on frequency hopping techniques. The second part focuses on thermal signature recognition, covering computer vision principles, the training of the YOLOv10 algorithm, and the evaluation of results. The thesis concludes with a summary of findings and suggestions for future research.

Keywords: Unmanned Aerial Vehicle, Artificial Intelligence, Computer Vision, Thermal Signatures, YOLOv10, Wireless Communications, VHF, Military Applications.



Ευχαριστίες

Κάθε τέλος είναι και μια νέα αρχή, κι όπως είπε ο Νίτσε, «Αυτός που έχει ένα γιατί για να ζήσει, μπορεί να αντέξει σχεδόν οποιοδήποτε πώς».

Η πτυχιακή αυτή εργασία δεν είναι μόνο αποτέλεσμα ακαδημαϊκής προσπάθειας, αλλά και μιας διαδρομής γεμάτης ερωτήματα, αμφιβολίες, στήριξη και έμπνευση.

Θα ήθελα να εκφράσω τις ειλικρινείς μου ευχαριστίες στην κα. Ειρήνη Καρανάσιου και στον κ. Ελευθέριο Δοϊτσίδη για την καθοδήγηση, τη γνώση και την υπομονή που έδειξαν καθ' όλη τη διάρκεια της εργασίας μου. Η συμβολή τους δεν περιορίστηκε μόνο σε επιστημονικό επίπεδο, αλλά αποτέλεσε και πηγή έμπνευσης για να προσεγγίσω τη γνώση με αφοσίωση και κριτική σκέψη.

Ευχαριστώ επίσης από καρδιάς την οικογένειά μου και τους φίλους μου, για την αμέριστη στήριξη, την υπομονή και την ενθάρρυνση που μου πρόσφεραν σε κάθε βήμα αυτής της διαδρομής.



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



Περιεχόμενα

Περίληψη	2
Abstract	3
Ευχαριστίες	4
Κατάλογος Εικόνων	7
Δήλωση Συγγραφέα Διπλωματικής Εργασίας	8
Σκοπός και Διάρθρωση Διπλωματικής Εργασίας	9
Εισαγωγή.....	11
1. Ασύρματες Ζεύξεις ΜΕΑ.....	12
1.1 Συστήματα Επικοινωνιών ΜΕΑ	12
2. Ασύρματες VHF Επικοινωνίες	15
2.1 Στρατιωτικές VHF Επικοινωνίες	15
2.2 VHF Διαυλοποίηση και Πρόσβαση στο Σ/Α	22
Μέθοδοι Αμφιδρόμησης	24
Πολλαπλή Πρόσβαση στο Μέσο Επικοινωνίας.....	25
3. Υπολογιστική Όραση.....	33
3.1 Ιστορική Αναδρομή.....	33
3.2 Θεωρητικό Υπόβαθρο για Νευρωνικά Δίκτυα – ΝΔ (Neural Networks - NN).....	35
3.2.1 Νευρωνικά Δίκτυα – Perceptron	35
3.2.2 Πολυεπίπεδα Δίκτυα Perceptron (Multiple Layer Perceptron – MLP) ..	38
3.2.3 Εκπαίδευση Νευρωνικών Δικτύων	39
3.2.3 Συνελκτικό Επίπεδο (Convolutional Layer)	42
3.2.4 Επίπεδα pooling (Τμήμα Υποδειγματοληψίας)	45
3.2.5 Οπισθοδιάδοση Σφάλματος (Backpropagation)	46
3.3 Αρχιτεκτονική των CNN δικτύων	47
3.4 Ανίχνευση Αντικειμένων (Object Detection).....	49
3.4.1 Μέθοδοι Ανίχνευσης	49
3.4.2 Μετρήσεις Αξιολόγησης	53
4. Μεθοδολογία.....	57
4.1 Θερμικό Αποτύπωμα Αρμάτων και Σύγχρονες Προκλήσεις	57
4.2 Σύνολο Δεδομένων	59
4.3 Υλοποίηση με Google Kolab	64
4.4 Αλγόριθμος YOLOv10.....	65
5. Υλοποίηση Εκπαίδευσης Αλγορίθμου	66
6. Αποτελέσματα	69



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ
ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



7. Συμπεράσματα – Μελλοντική Εργασία	74
Βιβλιογραφία	76



Κατάλογος Εικόνων

Εικόνα 1. Είδη Επικοινωνιών UAV Πηγή [1].....	13
Εικόνα 2. Διάγραμμα Uplink, Downlink, Real Link Πηγή [3].....	14
Εικόνα 3. Εύρος Line-of-Sight Σ/A με VHF συχνότητες Πηγή[14].....	16
Εικόνα 4. Διάθλαση σε VHF σήμα Πηγή [14].....	17
Εικόνα 5. Ανάκλαση Σήματος Πηγή [14]	18
Εικόνα 6. Διάδοση κύματος σε εσωτερικό χώρο Πηγή [14].....	19
Εικόνα 7. Διασπορά φάσματος με FHSS Πηγή [37]	21
Εικόνα 8. Διάγραμμα Διαυλοποίησης Πηγή [46]	22
Εικόνα 9. Διάγραμμα FDM Πηγή [38]	23
Εικόνα 10. Διάγραμμα TDM Πηγή[39].....	24
Εικόνα 11. Σχηματική Αναπαράσταση FDM Πηγή [37]	27
Εικόνα 12. . Σχηματική Αναπαράσταση TDMA Πηγή [37]	28
Εικόνα 13. Διάγραμμα TDMA Πηγή [37]	28
Εικόνα 14. Σχηματική Αναπαράσταση NTDMA Πηγή [37]	29
Εικόνα 15. Αναπαράσταση πειράματος γάτας Πηγή [40].....	33
Εικόνα 16. Διαδικασία επεξεργασίας εικόνας Πηγή [40]	34
Εικόνα 17. Απεικόνιση Νευρώνα Perceptron Πηγή [41].....	36
Εικόνα 18. Τρόπος Λειτουργίας Perceptron Πηγή [41].....	40
Εικόνα 19. Διδιάστατη συνέλιξη Πηγή [42]	42
Εικόνα 20. Απεικόνιση συνέλιξης εικόνας, με φίλτρο πίνακα 3x3 Πηγή [42]	43
Εικόνα 21. Απεικόνιση συνέλιξης RGB εικόνας με τρισδιάστατο φίλτρο Πηγή [42].	43
Εικόνα 22. Διαδικασία ανίχνευσης ακμών Πηγή [42]	44
Εικόνα 23. Παράδειγμα εφαρμογής padding με p=1 Πηγή [42]	44
Εικόνα 24. Απεικόνιση feature map σε εικόνα	45
Εικόνα 25. Παράδειγμα Εφαρμογής Max Pooling Πηγή [42].....	45
Εικόνα 26. Σύγκριση Shallow με Deep CNN Πηγή [43]	47
Εικόνα 27. Απεικόνιση Ανίχνευσης Αντικειμένων σε Εικόνα	49
Εικόνα 28. Διαδικασία Ανίχνευσης Αντικειμένων με το μοντέλο YOLO Πηγή [18].	51
Εικόνα 29. Αρχιτεκτονική Darknet ως backbone σε YOLOv3. Πηγή:[20].....	51
Εικόνα 30. Αρχιτεκτονική R-CNN Πηγή [21].....	53
Εικόνα 31. Αρχιτεκτονική Faster R-CNN με RPN Πηγή [22]	53
Εικόνα 32. Παράδειγμα μέτρησης IoU Πηγή [44]	55
Εικόνα 33. Πίνακας Σύγχυσης Πηγή [45]	55
Εικόνα 34. Θερμικό αποτύπωμα αρμάτων μάχης M1 Abrams (Αριστερά) και T-72 (Δεξιά).....	57
Εικόνα 35. Απόκρυψη Θερμικού Αποτυπώματος.....	58
Εικόνα 36. Διαδικασία Δημιουργίας Project για Επισημείωση Δεδομένων	61
Εικόνα 37. Εισαγωγή Δεδομένων	61
Εικόνα 38. Roflow Auto-Label.....	62
Εικόνα 39. Εφαρμογή Blur σε εικόνα του συνόλου δεδομένων.....	63
Εικόνα 40. Παράδειγμα σημειωματαρίου Google Colab	64
Εικόνα 41. Συνεπείς διπλές αναθέσεις για εκπαίδευση χωρίς NMS Πηγή [35].....	65



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ
ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



Δήλωση Συγγραφέα Διπλωματικής Εργασίας

Ο/η κάτωθι υπογεγραμμένος/η Μάριος Μεσσήνης - Φίλιππας του Γρηγορίου, με αριθμό μητρώου φοιτητής του ΠΜΣ Ευφυή Συστήματα του Τμήματος Μηχανικών Παραγωγής και Διοίκησης του Πολυτεχνείου Κρήτης και της Στρατιωτικής Σχολής Ευελπίδων, δηλώνω ότι:

«Είμαι συγγραφέας αυτής της πτυχιακής/διπλωματικής εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της, είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, οι όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε ακριβώς είτε παραφρασμένες, αναφέρονται στο σύνολό τους, με πλήρη αναφορά στους συγγραφείς, τον εκδοτικό οίκο ή το περιοδικό, συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο. Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από μένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου, όσο και του Ιδρύματος.

Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του πτυχίου μου».

Ο Δηλών

Μάριος Γρ. Μεσσήνης - Φίλιππας



Σκοπός και Διάρθρωση Διπλωματικής Εργασίας

Σκοπός της παρούσης διπλωματικής εργασίας αποτελεί η σχεδίαση ενός Μη Επανδρωμένου Αεροσκάφους (Μ.Ε.Α) για τη χρήση του από επίγειες στρατιωτικές δυνάμεις. Πιο συγκεκριμένα, κάνοντας χρήση αλγορίθμου μηχανικής μάθησης, αναγνωρίζει το θερμικό αποτύπωμα τεθωρακισμένων αρμάτων μάχης του αντιπάλου ειδοποιώντας εγκαίρως τις στρατιωτικές μονάδες που ακολουθούν. Παράλληλα, θα εξεταστεί σε θεωρητικό επίπεδο, η ιδανικότερη μέθοδος επικοινωνίας μεταξύ Μ.Ε.Α και επίγειων δυνάμεων, χρησιμοποιώντας τις VHF αναπηδούμενες συχνότητες του σταθμού ασυρμάτου TRC - 9200.

Προς την επίτευξη αυτού του σκοπού, κρίθηκε σκόπιμο, η εργασία να χωριστεί σε δύο διακριτά τμήματα. Το μεν πρώτο, αφορά τις Ασύρματες Επικοινωνιακές Ζεύξεις, με μία βιβλιογραφική ανασκόπηση των μεθόδων που χρησιμοποιούνται για την επικοινωνία των Μ.Ε.Α με έναν επίγειο σταθμό, αλλά και εκείνων που εφαρμόζονται από τα χερσαία τμήματα των Ενόπλων Δυνάμεων. Έτσι, στο 1^ο Κεφάλαιο αναφέρονται τα δίκτυα που χρησιμοποιούνται για την επικοινωνία των Μ.Ε.Α, με έμφαση στις συνδέσεις ελέγχου (CNPC) που διασφαλίζουν την ασφάλεια πτήσης και στις ζεύξεις δεδομένων (Data Links) για την επιχειρησιακή ανταλλαγή πληροφοριών. Η επικοινωνία αυτή, πραγματοποιείται μέσω εξειδικευμένων συχνοτήτων που ενισχύουν την αξιοπιστία και την ανθεκτικότητα της μετάδοσης των δεδομένων. Κατόπιν, στο 2^ο Κεφάλαιο αναλύονται οι ασύρματες επικοινωνίες μέσω VHF συχνοτήτων, με έμφαση στις στρατιωτικές. Παρουσιάζονται τα ειδικά χαρακτηριστικά των Σταθμών Ασυρμάτου (Σ/Α) που χρησιμοποιούν VHF συχνότητες, όπως για παράδειγμα ο περιορισμός της οπτικής επαφής- Line-of-Sight (LoS) μεταξύ δύο Σ/Α και η μέθοδος αναπήδησης συχνότητας – Frequency Hoping, για την προστασία από υποκλοπές. Ακόμη, αναφέρονται οι μέθοδοι πολλαπλής πρόσβασης στο μέσο, οι οποίες δίνουν τη δυνατότητα της μετάδοσης δεδομένων από διάφορους χρήστες, παρέχοντας άμεση και πλήρη ανταπόκριση στον κεντρικό σταθμό. Ολοκληρώνοντας, στο τέλος του Κεφαλαίου, πραγματοποιείται σύγκριση μεταξύ των μεθόδων πολλαπλής πρόσβασης στο μέσο, με βάση τα κριτήρια που τίθενται για τη χρήση τους από τις Ε.Δ.

Το δεύτερο τμήμα της παρούσης εργασίας, αφορά την αναγνώριση των θερμικών αποτυπωμάτων αρμάτων μάχης, από αλγόριθμο μηχανικής μάθησης και την παρουσίαση των αποτελεσμάτων. Απαραίτητο βήμα, για την πληρέστερη κατανόηση του τομέα της Υπολογιστικής Όρασης, είναι μία αναδρομή της εξέλιξής της στην πορεία των ετών και η ανάλυση της θεωρίας πάνω στην οποία βασίστηκε. Το 3^ο Κεφάλαιο λοιπόν, αποσκοπεί στην ανάπτυξη, από τον αναγνώστη, ενός θεωρητικού υποβάθρου, απαραίτητου για να εντρυφήσει στο συναρπαστικό μεν, απαιτητικό δε, τομέα της υπολογιστικής όρασης. Αναλύονται ορισμοί που αφορούν τα Τεχνητά Νευρωνικά Δίκτυα και την εκπαίδευση αυτών, όπως η μεθοδολογία οπισθοδιάδοσης σφάλματος (Backpropagation), καθώς και ορισμοί για την Ανίχνευση Αντικειμένων (Object Detection). Έπειτα, στο 4^ο Κεφάλαιο, παρουσιάζεται η μεθοδολογία που ακολουθήθηκε για την εκπαίδευση του αλγορίθμου μηχανικής όρασης. Πραγματοποιείται μία συνοπτική αναφορά των αιτίων που καθιστούν αναγκαία την έγκαιρη αναγνώριση του θερμικού αποτυπώματος των αρμάτων μάχης, ενώ ακολούθως παρουσιάζονται οι τεχνικές επεξεργασίας και επισημείωσης των δεδομένων



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



εκπαίδευσης, το περιβάλλον υλοποίησης Google Kolab και ο αλγόριθμος YOLOv10. Στο 5^ο Κεφάλαιο, παρατίθενται τμήματα του κώδικα που χρησιμοποιήθηκε για την εκπαίδευση του αλγορίθμου, ενώ στο 6^ο Κεφάλαιο, παρουσιάζονται και αξιολογούνται τα αποτελέσματα της υλοποίησης. Η εργασία ολοκληρώνεται με το 7^ο Κεφάλαιο, όπου εξάγονται τα συμπεράσματα από την παρούσα έρευνα και αναγνωρίζονται τα σημεία, στα οποία μπορεί να υπάρξει συμπληρωματική μελλοντική μελέτη.



Εισαγωγή

Τα τελευταία έτη και ιδιαίτερα, έπειτα από τις τελευταίες ένοπλες συρράξεις, αναδείχθηκε στην επικαιρότητα η χρήση των UAV ως οπλικό σύστημα. Στη σύρραξη μεταξύ Ουκρανίας και Ρωσίας, τα UAV χρησιμοποιούνται κατά κόρον, ως ένα φθινό μεν, αποτελεσματικό δε, οπλικό σύστημα, το οποίο επιφέρει καίρια πλήγματα στον αντίπαλο. Ενδεικτικό της σχέσης μεταξύ κόστους και δυνατοτήτων, αποτελεί το γεγονός ότι, ένα μόλις UAV μερικών εκατοντάδων – αν όχι δεκάδων – ευρώ, μπορεί να καταστρέψει ιδιαίτερα ακριβά οπτικά συστήματα, όπως τεθωρακισμένα άρματα μάχης ή αντιαεροπορικές συστοιχίες. Παράλληλα, καίριο ρόλο έχουν στην παροχή πληροφοριών για τον εχθρό. Τόσο σε τακτικό, όσο και σε επιχειρησιακό επίπεδο, ένα UAV μπορεί να παράσχει πληροφορίες για τη θέση και δραστηριότητα εχθρικών τμημάτων, από το κλιμάκιο της Διμοιρίας έως και του Τάγματος. Δίνει τη δυνατότητα, λοιπόν έγκαιρης πληροφόρησης, άρα και πληρέστερης σχεδίασης της επιχείρησης που θα εκτελεστεί, μειώνοντας δραστικά τις ενδεχόμενες απώλειες.

Πέρα από τη χρήση των UAV, ο αιώνας μας χαρακτηρίζεται από την ραγδαία ανάπτυξη της τεχνητής νοημοσύνης. Στον τομέα αυτό, ανήκει και η υπολογιστική όραση, η οποία μεταξύ άλλων, δίνει τη δυνατότητα της αναγνώρισης αντικειμένων. Ο αλγόριθμος εκπαιδεύεται σε ένα σύνολο δεδομένων και έπειτα δύναται να αναγνωρίσει τα αντικείμενα αυτά, είτε βρίσκονται στο σύνολο, είτε όχι. Η τεχνολογία αυτή, βρίσκει εφαρμογή στις Ένοπλες Δυνάμεις, μέσω των περιφερόμενων πυρομαχικών (loitering munitions). Παράλληλα, δύναται να χρησιμοποιηθεί για την έγκαιρη αναγνώριση στόχων υψηλής αξίας στο πεδίο της μάχης, λόγω της ευκινησίας που παρέχει το UAV. Παράδειγμα τέτοιων στόχων, είναι το άρμα μάχης, το οποίο λόγω της ισχύος πυρός του, αποτελεί σημαντικό οπλικό σύστημα. Προκειμένου να αναγνωρισθεί το άρμα μάχης από έτερα άρματα, σε μία αρματομαχία, χρησιμοποιείται το υπέρυθρο φάσμα, ώστε να είναι ορατό το θερμικό αποτύπωμα του άρματος. Το θερμικό αποτύπωμα, αποτελεί και την αδυναμία κάθε άρματος μάχης, καθώς είναι πολύ δύσκολο να ελαττωθεί, οπότε και αυτή εκμεταλλεύεται ένα UAV.



1. Ασύρματες Ζεύξεις ΜΕΑ

1.1 Συστήματα Επικοινωνιών ΜΕΑ

Τα τελευταία χρόνια, τα μη Επανδρωμένα Αεροσκάφη (ΜΕΑ) - UAVs, χρησιμοποιούνται με ολοένα και αυξανόμενη συχνότητα, τόσο σε πολιτικές όσο και σε στρατιωτικές εφαρμογές. Η απομακρυσμένη επικοινωνία μεταξύ αυτών με το κέντρο ελέγχου, αποτελεί κομβικής σημασίας έργο για την ολοκλήρωση της αποστολής. Προς τούτο, υπάρχουν 4 κύριοι τύποι επικοινωνιακών δικτύων για τις επικοινωνίες των UAV [1] :

- (1) δίκτυο από UAV-σε-UAV (U2U), για συνδέσεις δεδομένων και ελέγχου,
- (2) δίκτυο UAV με Σταθμό Βάσης στο έδαφος (U2GS), για σύνδεση ελέγχου και εντολών,
- (3) Ασύρματοι κόμβοι UAV-σε έδαφος, για συλλογή και μετάδοση δεδομένων που συλλέχθηκαν από το UAV
- (4) συστήματα UAV σε δορυφόρο.

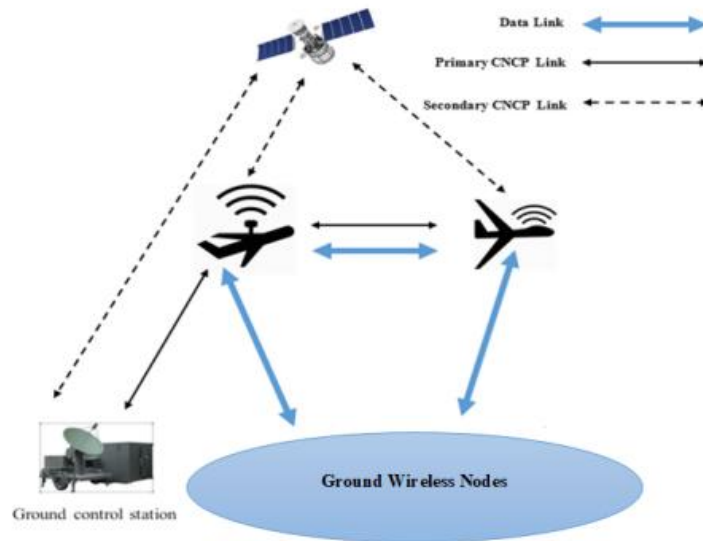
Οι επικοινωνίες των UAVs, αποτελούνται από δύο βασικούς τύπους τηλεπικοινωνιακών συνδέσεων, τη σύνδεση ελέγχου (control link), που αναφέρεται στη βιβλιογραφία ως Control and Non-Payload Communications (CNPC) Link και τη σύνδεση δεδομένων (data link).

Το CNPC, χρησιμοποιείται για την ασφάλεια πτήσης του UAV και την αποφυγή συγκρούσεων μεταξύ διαφόρων UAVs, που επιχειρούν στην ίδια περιοχή. Αναφέρεται δηλαδή, στις πληροφορίες που αφορούν την πλοήγηση του αεροσκάφους, αλλά δεν σχετίζονται με τη μεταφορά του κύριου φορτίου (payload) που το αεροσκάφος μπορεί να μεταφέρει. Έτσι, το CNPC δίκτυο κατηγοριοποιείται στις εξής:

1. Command & Control (C2): Αφορά δεδομένα σχετικά με την τηλεμετρία του UAV. Μεταδίδονται εντολές και δεδομένα από το χειριστή προς το UAV και αντίστροφα, καθώς και στοιχεία πτήσης, όπως η θέση, η ταχύτητα, η πορεία κ.ά
2. Air Traffic Control (ATC): Αφορά δεδομένα για την ασφάλεια πτήσης του UAV. Σε πληθώρα εφαρμογών τα UAVs, επιχειρούν σε ένα περιβάλλον αυξημένης αεροπορικής κίνησης και προς τούτο είναι κρίσιμη η διακίνηση δεδομένων ATC, για την διασφάλιση της ασφάλειας των πτήσεων.
3. Sense & Avoid: Αφορά δεδομένα κρίσιμα για την ασφαλή πλοήγηση του UAV, μακριά από έτερα UAV, φυσικά ή τεχνητά εμπόδια.

Παράλληλα, γίνεται χρήση δύο δικτύων CNPC [2], το πρωτεύον CNPC δίκτυο και το δευτερεύον. Το πρωτεύον δίκτυο, αποτελεί τον προτιμώμενο σύνδεσμο ελέγχου, είναι ο κύριος τρόπος επικοινωνίας και ελέγχου του UAV, καθώς καθιερώνεται άμεσα κατά την απογείωση και την προσγείωση, όπου η άμεση και αξιόπιστη επικοινωνία είναι κρίσιμη. Επίσης, δύναται να χρησιμοποιηθεί μέσω δορυφόρου, ώστε να ενισχυθεί η αξιοπιστία και η ανθεκτικότητα της επικοινωνίας σε περίπτωση που ο κύριος σύνδεσμος διακοπεί. Το δευτερεύον δίκτυο, αντίθετα, μπορεί να δημιουργηθεί μέσω

δορυφόρου, γεγονός που επιτρέπει την επικοινωνία με το UAV, ακόμη και σε μεγάλες αποστάσεις από τον χειριστή. Το συγκεκριμένο δίκτυο, αποτελεί έναν εφεδρικό σύνδεσμο, μία εναλλακτική λύση στον πρωτεύοντα και δύναται να ενεργοποιηθεί όταν το UAV βρίσκεται σε λειτουργία.

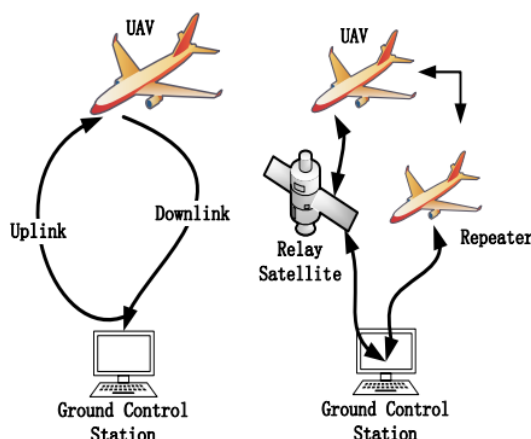


Εικόνα 1. Είδη Επικοινωνιών UAV
Πηγή [1]

Πέρα από το CNCP δίκτυο, υπάρχει και το Data Link. Το συγκεκριμένο δίκτυο χρησιμοποιείται για αποστολή και λήψη δεδομένων, όπως τη μετάδοση κατερχόμενης ζεύξης από το UAV στον επίγειο σταθμό ή δορυφόρο και τη μετάδοση ανερχόμενης ζεύξης από τον επίγειο σταθμό ή τον δορυφόρο στο UAV. Η χωρητικότητα της σύνδεσης αυτής, εξαρτάται από τις επιχειρησιακές απαιτήσεις, καθώς διακινούνται δεδομένα (timely mannered data), όπως φωτογραφίες, βίντεο, δεδομένα αισθητήρων κ.ά. Προκύπτει συνεπώς, η ακόλουθη κατηγοριοποίηση του Data Link, η οποία αποδίδεται και σχηματικά στην εικόνα 2:

1. Uplink: Η ανερχόμενη ζεύξη, χρησιμοποιείται κυρίως για την αποστολή και λήψη εντολών μεμακρυσμένου ελέγχου από τον σταθμό εδάφους στο UAV. Για την επίτευξη αυτού, υιοθετεί την τεχνολογία άμεσης επέκτασης για να εξασφαλίσει την ασφαλή μετάδοση των πληροφοριών. Όταν ο σταθμός ελέγχου στέλνει δεδομένα μέσω της ανερχόμενης σύνδεσης, το UAV τα λαμβάνει και τα εκτελεί, ελέγχοντάς το. Ο σταθμός δημιουργεί μια επικοινωνιακή σύνδεση και διατηρεί τον συγχρονισμό για να πραγματοποιήσει τον έλεγχο της στάσης και των παραμέτρων του UAV από τον σταθμό ελέγχου.
2. Downlink: Η καθοδική ζεύξη χρησιμοποιείται για τη μεταφορά των δεδομένων πτήσης και εικόνων από το UAV προς τον σταθμό εδάφους. Δεδομένου ότι, καλείται να διαχειριστεί μεγάλο όγκο δεδομένων, διαθέτει μεγαλύτερο εύρος ζώνης και υψηλότερο ρυθμό μετάδοσης σε σύγκριση με την ανερχόμενη σύνδεση.
3. Realy Link: Τα περισσότερα UAV βασίζονται σε προσαρμοσμένους συνδέσμους δεδομένων που λειτουργούν εντός οπτικής γραμμής. Όταν τα

UAV υπερβαίνουν την εμβέλεια της ασύρματης επικοινωνίας, είναι αναγκαία η χρήση αναμεταδότη ή σύνδεσης δεδομένων διαμέσου δορυφορικής επικοινωνίας για σύνδεση, πέρα από τον ορίζοντα [4] , δηλαδή μια σύνδεση αναμετάδοσης.



*Εικόνα 2. Διάγραμμα Uplink, Downlink, Relay Link
Πηγή [3]*

Προκειμένου να υλοποιηθούν οι ζεύξεις που περιγράφηκαν ανωτέρω, κρίνεται απαραίτητη η εκμετάλλευση ορισμένων ζωνών συχνοτήτων. Ο διαχωρισμός και η διάκριση των ζωνών, καθώς και οι χρήσεις της κάθε ζώνης καθορίζονται από Διεθνείς Οργανισμούς, όπως η Διεθνής Ένωση Τηλεπικοινωνιών του ΟΗΕ (International Telecommunication Union , ITU of the United Nations Organization), οπότε και προκύπτει ο ακόλουθος πίνακας.

ΟΝΟΜΑΣΙΑ ΖΩΝΗΣ	ΑΡΤ/ΕΟ	ITU ΖΩΝΗ	ΣΥΧΝΟΤΗΤΑ	ΜΗΚΟΣ ΚΥΜΑΤΟΣ	ΤΥΠΙΚΕΣ ΧΡΗΣΕΙΣ
Extremely Low	ELF	1	3-30 Hz	100,000km – 10,000 km	Υποβρυχιακές Επικοινωνίες
Super Low	SLF	2	30-300 Hz	10,000-1000km	Υποβρυχιακές Επικοινωνίες
Ultra Low	ULF	3	300-3000Hz	1000-100km	Επικ/νίες Ορυχείων
Very Low	VLF	4	30-30 kHz	100-10 km	Παλμογράφοι
Low	LF	5	30-300 kHz	10 – 1 km	ΑΜ Μεταδόσεις
Medium	MF	6	300–3000 kHz	1 km – 100 m	ΑΜ Μεταδόσεις
High	HF	7	3-30 MHz	100 – 10m	Ραδιοερασιτέχνες
Very High	VHF	8	30 – 300 MHz	10m – 1 m	TV Μεταδόσεις, Ένοπλες Δυνάμεις
Ultra High	UHF	9	300 – 3000 MHz	1m – 100 mm	TV, Κινητά, Επικ/νίες Αέρος-Αέρος
Super High	SHF	10	3 – 3- GHz	100 – 10mm	Ραντάρ, LAN
Extremely High	EHF	11	30 – 300 GHz	10 mm – 1mm	Αστρονομία



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



Οι συχνότητες που χρησιμοποιούνται για τις επικοινωνίες CNPC και Data Link, για τα UAV, καθορίζονται από διεθνείς και εθνικές κανονιστικές αρχές. Σε αυτό το πλαίσιο, για τη διασφάλιση της αξιόπιστης λειτουργίας των UAV, η διεθνής ένωση τηλεπικοινωνιών (International Telecommunication Union – ITU), καθόρισε το απαιτούμενο Bandwidth – BW , για επίγεια (terrestrial) και δορυφορική (satellite) επικοινωνία, σε 34 MHz και 56 MHz, αντίστοιχα. Καλύπτονται έτσι αμφότερες οι Line-of-Sight (LOS) και Beyond-Line-of-Sight (BLOS) επικοινωνίες για τα CNPC δίκτυα. Παράλληλα, στο World Radiocommunication Conference 2012, αποφασίστηκε η χρήση του εύρους συχνοτήτων C-Band (5030 – 5091 MHz), για την υποστήριξη των CNPC δικτύων. Το συγκεκριμένο εύρος συχνοτήτων, προσφέρει επαρκή χωρητικότητα για τη μεταφορά δεδομένων, ενώ ταυτόχρονα είναι ανθεκτικό σε παρεμβολές, προσφέροντας αξιόπιστες επικοινωνίες. Όσον αφορά το Data Link, καθώς η χωρητικότητα εξαρτάται από τα δεδομένα που διακινούνται, υπάρχει πληθώρα συχνοτήτων. Από 400 MHz έως 900 Mhz, για δεδομένα τηλεμετρίας, σε 3 – 8 GHz για μετάδοση βίντεο.

2. Ασύρματες VHF Επικοινωνίες

2.1 Στρατιωτικές VHF Επικοινωνίες

Στις στρατιωτικές ασύρματες επικοινωνίες, η ζώνη συχνοτήτων που εξακολουθεί να χρησιμοποιείται είναι η VHF. Παραμένει σημαντική στις επικοινωνίες μη Επανδρωμένων Αεροσκαφών, ιδιαίτερα ειδικά όταν τα πιο σύγχρονα δίκτυα είναι ευάλωτα σε παρεμβολές ή καταστροφή ή σε περιβάλλοντα με φυσικά εμπόδια.. Η ζώνη VHF για χρόνια, αποτελούσε μια επιλογή για Σταθμούς Ασυρμάτου (Σ/Α) τύπου manpack, οι οποίοι χρησιμοποιούνται από στρατεύματα στο έδαφος για να επικοινωνούν σε μια τοπική περιοχή, ακτίνας πέντε περίπου μιλίων [14]. Οι κεραίες και τα στοιχεία επιλεκτικής ρύθμισης των ραδιοφώνων VHF είναι πολύ μικρότερα από τα αντίστοιχα των HF. Παράλληλα, η αλματώδης εξέλιξη στη βιομηχανία ημιαγωγών, έχει βελτιώσει την απόδοση των ραδιοφώνων VHF, καθώς οι σύγχρονες μπαταρίες είναι μικρότερες, ελαφριές και μεγαλύτερης διάρκειας, συγκριτικά με αυτές του παρελθόντος. Εν αντιθέσει με το HF, οι μεταδόσεις VHF δεν έχουν τη δυνατότητα ανάκλασης μέσω της ιονόσφαιρας και περιορίζονται σε επικοινωνίες οπτικής επαφής (LOS). Αυτό, μειώνει την πολυπλοκότητα των ραδιοεκπομπών σε ένα μεγάλο πεδίο μάχης και την πιθανότητα υποκλοπής από μη φιλικούς παράγοντες.

Επιπλέον, το μεγαλύτερο εύρος ζώνης των ραδιοφώνων VHF ενισχύει την αποτελεσματικότητα των μεθόδων κωδικοποίησης και κρυπτογράφησης, επιτρέποντας υψηλότερη απόδοση δεδομένων συγκριτικά με τα ραδιόφωνα HF. Η μεγαλύτερη χωρητικότητα ζώνης και η περιορισμένη εμβέλειά τους τα καθιστούν ιδανικά για επικοινωνίες μεταξύ ομάδων.

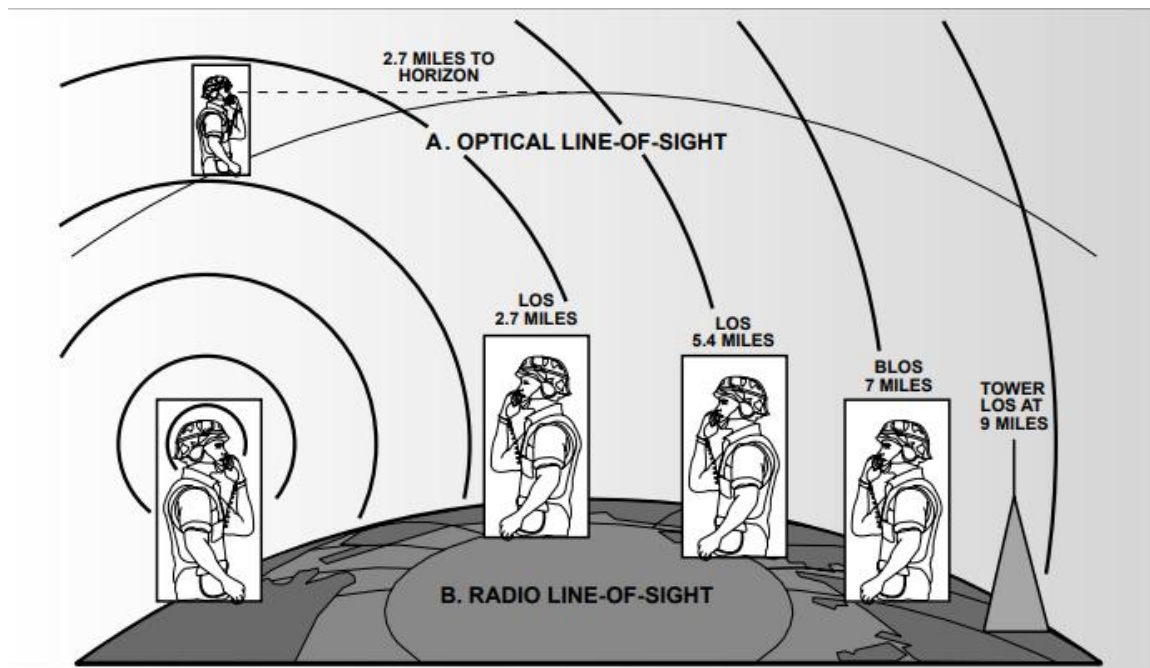
Όπως, αναφέρθηκε και προηγουμένως το εύρος της ζώνης VHF κυμαίνεται από 30 – 300 MHz. Η διαυλοποίηση (chanelling) που πραγματοποιείται είναι ανά 25 kHz, ωστόσο τα τελευταία έτη, προκειμένου να επιλυθεί το πρόβλημα της συμφόρησης με τις συχνότητες, χρησιμοποιείται η 8.33 kHz διαυλοποίηση. Το γεγονός αυτό, επιτρέπει στους VHF Σταθμούς Ασυρμάτου (Σ/Α), να διαθέτουν 3 φορές περισσότερους

σταθμούς εκπομπής στο διαθέσιμο φάσμα. Παρόλη την εξέλιξη σχετικά με το πλήθος των σταθμών στο bandwidth της VHF συχνότητας, αυτή υπόκειται σε σημαντικούς περιορισμούς όσον αφορά τις συνθήκες διάδοσης στο διάλυο. Στη ζώνη VHF, απαραίτητη προϋπόθεση για τη διάδοση ενός σήματος, αποτελεί η οπτική επαφή (Line-of-Sight LOS) μεταξύ πομπού και δέκτη.

Η σχέση, με την οποία καθορίζεται η εμβέλεια LOS για οποιουδήποτε είδους μεταδόσεις, είναι

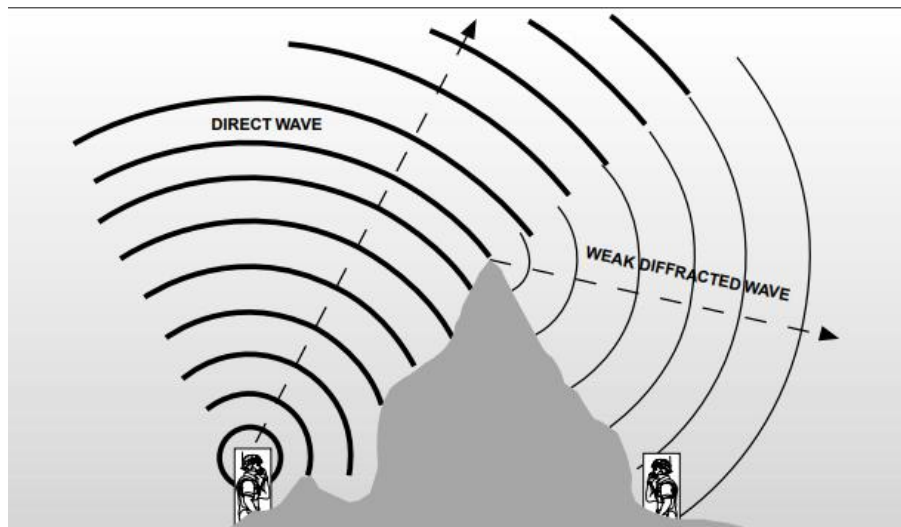
$$D_{LOS} = \sqrt{2 * H_1 * R + H_1^2} + \sqrt{2 * H_2 * R + H_2^2}$$

Όπου D_{LOS} , η εμβέλεια LOS, μετρούμενη σε ναυτικά μίλια, H_1 και H_2 το ύψος της 1^{ης} και 2^{ης} κεραίας αντίστοιχα και R η ακτίνα της Γης. Η ορατή γραμμή του ορίζοντα, όταν παρατηρείται από ύψος περίπου πέντε ποδιών πάνω από επίπεδη επιφάνεια, εκτείνεται σε απόσταση έως και 2,7 μιλίων. Αυτή, περίπου αποτελεί τη μέγιστη εμβέλεια ραδιοκυμάτων οπτικής επαφής (LOS) που μεταφέρεται από ένα όρθιο άτομο προς ένα άλλο Σ/Α τοποθετημένο στο έδαφος. Εάν και ο δεύτερος Σ/Α είναι τοποθετημένος στο ύψος ενός μέσου ατόμου, η εμβέλεια διπλασιάζεται σε 5,4 μίλια. Εάν ο δεύτερος άνθρωπος βρίσκεται στα 7 μίλια, αυτό θεωρείται πέραν της οπτικής επαφής (BLOS), οπότε δεν υπάρχει επαφή μέσω VHF. Καθίσταται, λοιπόν σαφές ότι το ύψος των κεραιών εκπομπής και λήψης έχει καθοριστικό ρόλο στην απόσταση της επικοινωνίας. Παραδείγματος χάριν, αν η κεραία λήψης ήταν τοποθετημένη σε πύργο ύψους 26 ποδιών, η συνολική απόσταση οπτικής επαφής (LOS) θα αυξανόταν στα 9 μίλια. Φυσικά, αν οι χειριστές ραδιοεπικοινωνίας βρίσκονταν στις κορυφές βουνών, η εμβέλεια της οπτικής επαφής θα μπορούσε να φτάσει από 50 έως 100 μίλια.



Εικόνα 3. Εύρος Line-of-Sight Σ/Α με VHF συχνότητες
Πηγή[14]

Όπως γίνεται εύκολα αντιληπτό, στο πεδίο σπανίζουν οι περιοχές όπου δεν υπάρχουν εμπόδια τα οποία θα εμποδίσουν την εκπομπή του Σ/Α. Στην πραγματικότητα, θα βρίσκονται κτίρια, εδαφικές εξάρσεις τα οποία θα επηρεάζουν την ποιότητα του ληφθέντος σήματος. Μία τέτοια περίπτωση αποτελεί και η εκπομπή VHF ραδιοκυμάτων πίσω από ράχες βουνών. Οι χαράδρες και οι λόφοι μπορούν να δημιουργήσουν σκιές, κατά τη μετάδοση VHF ραδιοκυμάτων. Ωστόσο, όταν υπάρχουν πολύ απότομες ράχες ή άλλοι τύποι απότομων εμποδίων, υπάρχει μία σημαντική εξαίρεση, στον προηγούμενο περιορισμό. Αυτό προκαλείται από ένα φαινόμενο που είναι γνωστό ως **Περίθλαση**. Όταν ένα VHF κύμα φτάσει σε μια απότομη άκρη, ένα μέρος του κύματος λυγίζει γύρω από την άκρη και συνεχίζει να διαδίδεται σαν να είχε τοποθετηθεί ένας ραδιοπομπός χαμηλής ισχύος στην κορυφή της ράχης. Ιδιαίτερα σημαντικό είναι, το σχήμα της ράχης να είναι σχετικά απότομο, καθώς ένας στρογγυλεμένος λόφος ή η καμπυλότητα της γης δεν είναι αρκετά για τη δημιουργία αυτού του φαινομένου. Η περίθλαση, αποτελεί ένα φαινόμενο ιδιαίτερα σημαντικό σε ένα πεδίο μάχης, όπου ένας στρατιώτης καλείται να βρει καταφύγιο πίσω από μια εδαφική έξαρση και παράλληλα να επικοινωνεί μέσω Σ/Α.



Εικόνα 4. Διάθλαση σε VHF σήμα
Πηγή [14]

Τα VHF και UHF ραδιοκύματα μπορούν να ανακλαστούν σε συνεκτικές επιφάνειες όπως βράχια ή αγώγιμο έδαφος, με ανάλογο τρόπο που μια δέσμη φωτός ανακλάται σε έναν τοίχο ή μια οροφή. Αρκετές φορές, υπάρχουν πολλές πιθανές διαδρομές μεταξύ μιας κεραίας εκπομπής και μιας κεραίας λήψης. Στο σχήμα, που ακολουθεί, υπάρχει μια άμεση διαδρομή LOS μεταξύ δύο Σ/Α, αλλά υπάρχει επίσης τουλάχιστον μια ανακλώμενη διαδρομή, που προέρχεται από το έδαφος μιας κοιλάδας μεταξύ τους.

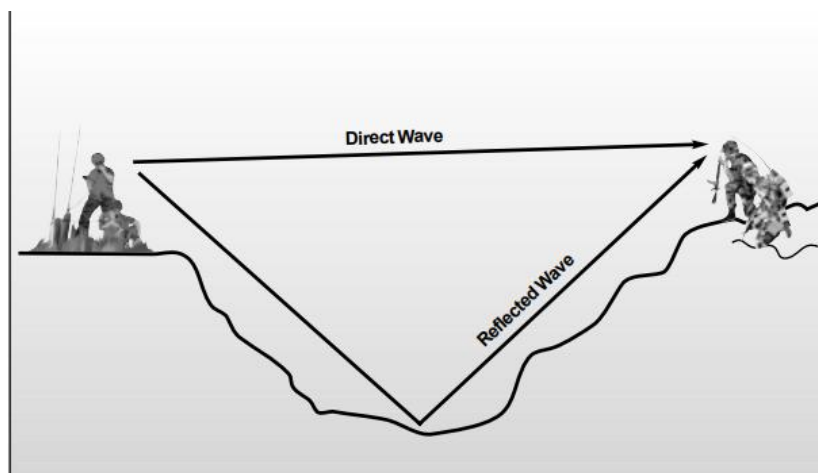
Όπως γίνεται αντιληπτό, αυτές οι δύο διαδρομές έχουν διαφορετικό μήκος, με την άμεση διαδρομή να είναι η πιο σύντομη από τις δύο. Δεδομένου ότι, τα ραδιοκύματα ταξιδεύουν με σταθερή ταχύτητα, το σήμα της άμεσης διαδρομής φτάνει στον δέκτη πριν από το ανακλώμενο κύμα. Αυτό, έχει ως αποτέλεσμα ότι οι ίδιες πληροφορίες φτάνουν στον δέκτη σε δύο διαφορετικές χρονικές στιγμές.

Το αποτέλεσμα αυτού του φαινομένου, μπορεί να παρομοιαστεί με τις ηχώ που ακούει κάποιος σε ένα δωμάτιο με κακή ακουστική. Αν οι ηχώ είναι αρκετά κοντά μεταξύ τους,

είναι δύσκολο να καταλάβει κανείς τι λέγεται. Στην ορολογία της ραδιοτηλεφωνίας, αυτό ονομάζεται **παραμόρφωση πολλαπλών διαδρομών** (multipath distortion). Παρόλο που αυτό το φαινόμενο στις φωνητικές επικοινωνίες, μπορεί να προκαλέσει όχληση, έχει καταστροφικά αποτελέσματα για την ψηφιακή επικοινωνία υψηλού ρυθμού δεδομένων.

Το “**Picket fencing**” (**Διαλείψεις Πολλαπλών Διαδρομών**) αποτελεί μια μορφή πολλαπλής διαδρομής που είναι συνηθισμένη σε Σ/Α τοποθετημένους σε οχήματα. Όσο υψηλότερη είναι η συχνότητα, τόσο πιο έντονα εκδηλώνεται το φαινόμενο. Συνήθως, προκαλείται από παρεμβολές ή ανακλάσεις σημάτων προερχόμενα από ανθρωπογενείς κατασκευές, όπως κτίρια και σπίτια. Αυτά τα αντικείμενα, οδηγούν στη δημιουργία περιοχών με ενισχυμένα και αποδυναμωμένα πεδία (ή ενισχυμένα και εξασθενημένα σήματα), με αποτέλεσμα όταν ένα όχημα κινείται μέσα από αυτά τα πεδία, να λαμβάνει εναλλάξ ισχυρότερα και ασθενέστερα σήματα. Συνήθως, υπάρχει ένας «συρτός» ήχος στον δέκτη, καθώς τα σήματα γίνονται γρήγορα πιο αδύναμα, μετά πιο ισχυρά, μετά πιο αδύναμα ξανά. Οι κορυφές και τα κενά του σήματος είναι συνάρτηση του μήκους κύματος. Ένα σήμα 50 MHz, το οποίο λαμβάνει ένα όχημα που ταξιδεύει με 60 μίλια/ώρα μπορεί να «τρεμοπαίζει» πολύ γρήγορα καθώς το όχημα κινείται εντός του κέντρου μιας πόλης. Μερικές φορές αυτό το ίδιο φαινόμενο προκαλείται από σήματα δύο σταθερών Σ/Α που ανακλώνται από ένα κινούμενο αεροσκάφος πάνω από αυτά

Ο παραπάνω αντίστοιχος αγγλικός όρος, προέρχεται από το φαινόμενο που περιεγράφηκε και δημιουργεί εναλλαγή ισχυρού και αδύναμου σήματος, σαν να περνά μέσα από ένα “φράχτη από δοκάρια” (picket fence). Προκαλείται από πολλαπλές διαδρομές διάδοσης του σήματος και την επίδραση του φαινομένου Doppler, λόγω της κίνησης. Τα αποτελέσματα του φαινομένου, είναι η πρόκληση διακοπών στην επικοινωνία, κυρίως σε κινητά τηλέφωνα, ασύρματα δίκτυα και ραδιοεπικοινωνίες. Σε ψηφιακές μεταδόσεις, δύναται να οδηγήσει σε απώλεια πακέτων ή καθυστερήσεις λόγω επαναμετάδοσης δεδομένων ενώ, σε αναλογικές μεταδόσεις (π.χ. ραδιόφωνο FM), μπορεί να δημιουργήσει παροδικά προβλήματα στην ποιότητα του ήχου. Η χρήση πολλαπλών κεραιών σε διαφορετικές θέσεις, η ανίχνευση και διόρθωση λαθών στα ψηφιακά σήματα και η ισοστάθμιση ου σήματος σε πραγματικό χρόνο για μείωση των επιπτώσεων του fading αποτελούν λύσεις μείωσής του.

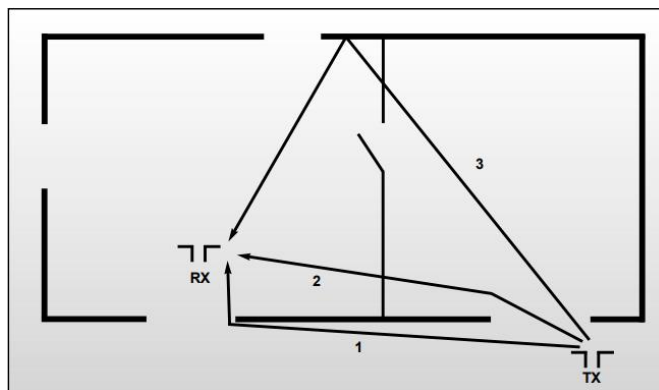


Εικόνα 5. Ανάκλαση Σήματος
Πηγή [14]

Σε τακτικές καταστάσεις, οι φορητοί Σ/Α συχνά λειτουργούν υπό κάλυψη μέσα σε κτίρια. Τα VHF κύματα δυσκολεύονται να διαπεράσουν εξωτερικούς τοίχους από οπλισμένο σκυρόδεμα, αλλά περνούν σχετικά εύκολα μέσα από παράθυρα και ελαφρά εσωτερικά χωρίσματα τοίχων. Η επόμενη εικόνα απεικονίζει έναν δέκτη που βρίσκεται σε ένα δωμάτιο ενός κτιρίου με έναν πομπό που βρίσκεται εκτός αυτού. Σε αυτή την περίπτωση, υπάρχουν τρεις διαφορετικές διαδρομές από τον πομπό προς τον δέκτη, από τις οποίες καμία δεν είναι άμεση.

Η Διαδρομή 1 διέρχεται διαμέσου του παραθύρου που βρίσκεται πιο κοντά στον δέκτη και εκτρέπεται γύρω από την αιχμηρή γωνία του παραθύρου, προτού καταλήξει στον δέκτη. Ομοίως, η Διαδρομή 2 σχεδόν ακολουθεί μία άμεση διαδρομή προς τον δέκτη, καθώς εκτρέπεται ελαφρώς από το πλαίσιο του παραθύρου που βρίσκεται πιο κοντά στον πομπό και στη συνέχεια περνά μέσα από έναν εσωτερικό τοίχο προτού φτάσει στον δέκτη. Η Διαδρομή 3, τέλος, διέρχεται μέσα από ένα παράθυρο και έναν εσωτερικό τοίχο και εν συνεχεία, προσκρούει σε έναν εξωτερικό τοίχο του κτιρίου και ανακλάται πίσω στον δέκτη.

Κάθε μία από αυτές τις διαδρομές διαθέτει διαφορετικό μήκος και, ως εκ τούτου, δύναται να προκαλέσει παραμόρφωση πολλαπλών διαδρομών. Συχνά, με τη μετακίνηση του δέκτη λίγα μέτρα προς κάποια κατεύθυνση μπορεί να αποφευχθεί μία ή περισσότερες από τις διαθέσιμες διαδρομές και να βελτιώσει σημαντικά την λήψη του σήματος.



Εικόνα 6. Διάδοση κύματος σε εσωτερικό χώρο
Πηγή [14]

Τα προτεινόμενα όρια στην εμβέλεια LOS μερικές φορές υπερβαίνονται στην πράξη. Ένας από τους κύριους λόγους για αυτό είναι ένα φαινόμενο που καλείται, **ducting** ή αγωγοποίηση /«μορφή κυματοδότησης». Τα VHF κύματα που ταξιδεύουν μέσω της ατμόσφαιρας κινούνται λίγο πιο αργά από ό,τι στο κενό, καθώς η πυκνότητα του αέρα τα επιβραδύνει. Όσο μεγαλύτερη είναι η πυκνότητα του αέρα, τόσο πιο κινούνται και τα κύματα. Υπό κανονικές συνθήκες, η πυκνότητα του αέρα είναι μεγαλύτερη στην επιφάνεια της γης και μειώνεται σταδιακά, όσο αυξάνεται και το υψόμετρο. Υπό συνθήκες καλοκαιρίας με ξηρή ατμόσφαιρα και ήπιο καιρό, οι όποιες μικρές διακυμάνσεις στην πυκνότητα του αέρα έχουν αμελητέα αποτελέσματα στην πορεία των ραδιοκυμάτων που κινούνται εντός αυτού.

Συχνά, όμως, παρατηρούνται έντονες αλλαγές στην πυκνότητα του αέρα λόγω μετεωρολογικών φαινομένων ή βαριών νεφώσεων. Σε αυτές τις περιπτώσεις, τα VHF



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



κύματα μπορούν να υποστούν κάμψη ή να «κυματοδηγηθούν» μεταξύ στρωμάτων αέρα με διαφορετική πυκνότητα. Μερικές φορές, αυτό το φαινόμενο προκαλεί καμπή των ραδιοκυμάτων προς το έδαφος, με αποτέλεσμα να τείνουν να ακολουθούν την καμπυλότητα της γης και ως εκ τούτου, η εμβέλεια LOS να είναι μεγαλύτερη από την οπτική εμβέλεια. Αυτός ο τύπος διάδοσης κυμάτων είναι αδύνατον να προβλεφθεί και όταν υπάρχουν συνθήκες αγωγοποίησης, αυτές συνήθως διαρκούν για αρκετές ώρες.

Από τα ανωτέρω, συνάγεται η σημασία της οπτικής επαφής LOS, μεταξύ δύο πομποδεκτών, ώστε να επιτευχθεί η μεταξύ τους VHF ζεύξη. Καταλυτικό ρόλο στην αύξηση της εμβέλειας, μπορούν να παίζουν τα μη Επανδρωμένα Αεροσκάφη (μηΕΑ ή UAV), πετυχαίνοντας ζεύξεις Beyond Line-of-Sight (BLOS). Το UAV, αποτελεί τον «μεσάζοντα», τον σύνδεσμο, ο οποίος μεταφέρει την πληροφορία μεταξύ των πομποδεκτών, λειτουργώντας ως αναμεταδότης. Παρακάμπτονται με αυτό τον τρόπο εμπόδια, τα οποία εξασθενούν το σήμα, ώστε να έχουμε ακόμα μεγαλύτερα δίκτυα επικοινωνιών.

Τα πλεονεκτήματα της χρήσης των συχνοτήτων VHF στις στρατιωτικές επικοινωνίες είναι πολλαπλά και καλύπτουν ένα ευρύ φάσμα επιχειρησιακών αναγκών, λόγω της φύσης της ζώνης και των ιδιοτήτων διάδοσης του σήματος. Παρέχουν αξιοπιστία σε επικοινωνίες μικρής εμβέλειας εφόσον υπάρχει οπτική επαφή (Line of Sight - LOS) καθιστώντας αυτή, ιδανική για στρατιωτικές μονάδες που επιχειρούν σε κοντινές αποστάσεις. Παράλληλα, τα VHF έχουν καλύτερη ποιότητα ήχου σε σύγκριση με άλλες ζώνες, όπως τα HF, η οποία οφείλεται στο ευρύτερο εύρος ζώνης που είναι διαθέσιμο στα VHF, επιτρέποντας πιο αποτελεσματικά συστήματα κωδικοποίησης και κρυπτογράφησης, κάτι που είναι σημαντικό σε στρατιωτικές επικοινωνίες όπου η σαφήνεια της πληροφορίας είναι κρίσιμη. Παρουσιάζουν ακόμη ευκολότερη διάδοση σήματος σε σύνθετα περιβάλλοντα, καθώς μπορούν να περάσουν μέσα από φυσικά εμπόδια, όπως δέντρα και κτίρια, αν και με μειωμένη ισχύ. Η ικανότητα για περίθλαση (diffraction) και ανάκλαση επιτρέπει κάποιες φορές τη μετάδοση σε περιοχές που δεν είναι απευθείας ορατές, βοηθώντας στην κάλυψη στρατιωτικών ομάδων σε σύνθετα εδάφη. Τέλος, με τεχνολογίες διαύλου, όπως η διαυλοποίηση στα 8.33 kHz επιτρέπει την καλύτερη εκμετάλλευση του διαθέσιμου φάσματος και τον ταυτόχρονο συντονισμό περισσότερων σταθμών επικοινωνίας σε περιορισμένο χώρο, παρέχοντας διαθεσιμότητα και ευελιξία.

Τα VHF παραμένουν σημαντικά στις επικοινωνίες μη επανδρωμένων, ιδιαίτερα για αποστολές μικρής εμβέλειας ή σε περιβάλλοντα με φυσικά εμπόδια. Ωστόσο, σε σύγχρονες εφαρμογές Beyond Line of Sight (BLOS), η θέση τους είναι περιορισμένη και συχνά υποστηρικτική, καθώς οι ζώνες υψηλότερων συχνοτήτων (π.χ. C-Band, Ku-Band) και τα δίκτυα 5G κυριαρχούν λόγω της ταχύτητας, του εύρους ζώνης και της ευελιξίας τους.

Αναπήδηση Συχνότητας (Frequency Hopping)

Στις στρατιωτικές επικοινωνίες, πρωταγωνιστικό ρόλο έχει η δυνατότητα των χρηστών του δικτύου, να μπορούν να μεταφέρουν σήματα, χωρίς να είναι ευάλωτοι σε υποκλοπές ή παρεμβολές σήματος. Με τη διασπορά της εκπεμπόμενης ισχύος σε μια μεγαλύτερη φασματική ζώνη, μειώνεται η πυκνότητα του φάσματος, με αποτέλεσμα να γίνεται αντιληπτή ως θόρυβος και να κρύβεται μέσα στον υπάρχοντα θόρυβο του



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

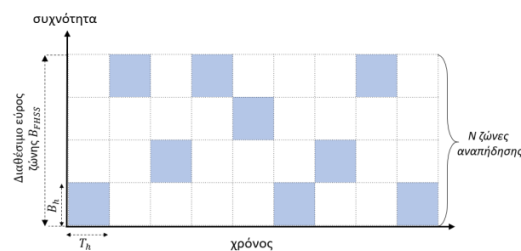
ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



περιβάλλοντος επικοινωνίας. Προς τούτο, δεν χρησιμοποιείται μία σταθερή συχνότητα, αλλά πολλές συχνότητες με μία διαδικασία που καλείται, αναπήδηση συχνότητας (frequency hopping).

Η τεχνική διασποράς φάσματος με αναπήδηση συχνότητας (Frequency Hopping Spread Spectrum – FHSS)[37], προσδίδει στο δίκτυο, μεγάλη ανοχή σε κάθε μορφή παρεμβολών. Αυτό, επιτυγχάνεται με την περιοδική αλλαγή της φέρουσας συχνότητας του μεταδιδόμενου σήματος, ώστε να αποφεύγονται οι παρεμβολές. Δύναται επίσης να αντιμετωπιστούν αποτελεσματικότερα με την κωδικοποίηση του καναλιού, κάνοντας έτσι το δίκτυο ανθεκτικό σε βάθος χρόνου, απέναντι σε κάθε κακόβουλη ενέργεια.

Σε ένα σύστημα FHSS (**F**requency-**H**opping **S**pread **S**pectrum), η αλλαγή των φερουσών συχνοτήτων πραγματοποιείται με βάση ένα μοτίβο αναπήδησης συχνότητας (frequency-hopping pattern), ενώ το σύνολο των διαθέσιμων φερουσών συχνοτήτων που μπορούν να χρησιμοποιηθούν ονομάζεται, σύνολο αναπήδησης (hopset). Το σύνολο αναπήδησης, αποτελείται από διάφορες φέρουσες συχνότητες, οι οποίες μπορούν να χρησιμοποιηθούν για τη δημιουργία του μοτίβου αναπήδησης. Ο ρυθμός με τον οποίο καθορίζεται η αλλαγή των φερουσών συχνοτήτων καθορίζεται από τον ρυθμό αναπήδησης (hop rate), η οποία πραγματοποιείται μέσα σε μία συγκεκριμένη ζώνη συχνοτήτων και διαχωρίζεται σε πολλά κανάλια. Τα κανάλια αυτά ονομάζονται ζώνη αναπήδησης (hopping band) και αποτελούν την περιοχή γύρω από κάθε φέρουσα συχνότητα και έχουν ένα συγκεκριμένο εύρος ζώνης, αρκετά μεγάλο ώστε να μεταφερθεί το σήμα.



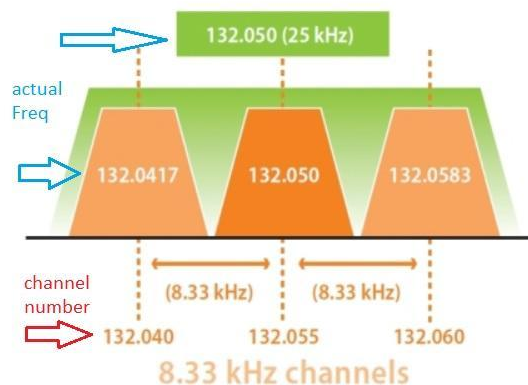
Εικόνα 7. Διασπορά φάσματος με FHSS
Πηγή [37]

Κύριο συστατικό ενός συστήματος FHSS είναι η δημιουργία του μοτίβου αναπήδησης συχνότητας. Μέσω μίας γεννήτριας μοτίβου, αυτό δημιουργείται με ψευδοτυχαίο και μη προβλέψιμο τρόπο για να μην μπορεί να υποκλαπεί. Η ασφάλεια αυτή που παρέχεται, γεννά την ανάγκη ενός ακριβούς συγχρονισμού μεταξύ των Σ/Α, ώστε να έχουν όλοι οι χρήστες το ίδιο μοτίβο αναπήδησης συχνότητας.

2.2 VHF Διαυλοποίηση και Πρόσβαση στο Σ/Α

Η διαδικασία της διαυλοποίησης (channeling) αποτελεί τη διαδικασία κατανομής του διαθέσιμου εύρους ζώνης στο VHF εύρος σε ξεχωριστά κανάλια. Διευκολύνεται έτσι, η ταυτόχρονη χρήση πολλαπλών συχνοτήτων από διαφορετικούς χρήστες ή συσκευές επικοινωνίας για την αποφυγή παρεμβολών και την αποδοτική χρήση του διαθέσιμου φάσματος συχνοτήτων. Μέσα στο υπάρχον εύρος ζώνης, δημιουργούνται κανάλια με διαφορετικό πλάτος το καθένα, ανάλογα την εφαρμογή και τη χρήση της συχνότητας. Κανάλια που χρησιμοποιούνται για αναλογική μετάδοση φωνής έχουν πλάτος 25 kHz ή 12.5 kHz, ενώ πιο σύγχρονες τεχνικές διαυλοποίησης, όπως η ψηφιακή επικοινωνία, επιτρέπουν τη χρήση μικρότερων καναλιών. Υπάρχουν δύο κύριες μέθοδοι διαυλοποίησης η FDMA (Frequency Division Multiple Access) Πολυπλεξία Πολλαπλής Πρόσβασης με Διάρθρωση Συχνότητας και η TDMA (Time Division Multiple Access) Πολυπλεξία Πολλαπλής Πρόσβασης με Διάρθρωση Χρόνου. Βασική διαφορά μεταξύ των δύο τεχνικών αποτελεί η υλοποίηση των δύο τεχνικών. Στη μεν πρώτη, κάθε χρήστης λαμβάνει μια ξεχωριστή συχνότητα για τη μετάδοση του σήματος, ενώ δε στη δεύτερη, η συνολική διαθέσιμη συχνότητα διαιρείται σε χρονικές σχισμές (time slots), όπου κάθε χρήστης μεταδίδει σε διαφορετικό χρονικό διάστημα, μοιράζοντας την ίδια συχνότητα με άλλους χρήστες.

Τα προηγούμενα έτη, το συνηθέστερο πλάτος καναλιού αποτελούσε τα 25kHz, η οποία εισήχθη τη δεκαετία του 1970. Επιτρέπει τη χρήση συνολικά 760 συχνοτήτων (19 MHz x 40 κανάλια ανά MHz). Για την κάλυψη των αναγκών των χρηστών, αυτές οι συχνότητες ίσως χρειαστεί να επαναχρησιμοποιηθούν. Για τον λόγο αυτό, εισήχθη η διαυλοποίηση σε 12,5 kHz, η οποία χωρίζει το κανάλι πλάτους 25 kHz σε δύο μικρότερα των 12,5 kHz. Αυξάνεται έτσι, η διαθεσιμότητα των διαύλων στο ίδιο εύρος συχνοτήτων, βελτιώνοντας παράλληλα την αποδοτικότητα του φάσματος, επιτρέποντας την ταυτόχρονη χρήση περισσότερων συχνοτήτων από διαφορετικούς χρήστες και εφαρμογές. Το νεότερο πλάτος διαυλοποίησης που χρησιμοποιείται στις VHF συχνοότητες είναι το 8.33 kHz. Με το πρότυπο απόστασης 8,33 kHz, κάθε υποζώνη των 25 kHz διαιρείται σε τρεις, αυξάνοντας σχεδόν τριπλάσια τον αριθμό των διαθέσιμων συχνοτήτων.



Εικόνα 8. Διάγραμμα Διαυλοποίησης
Πηγή [45]



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

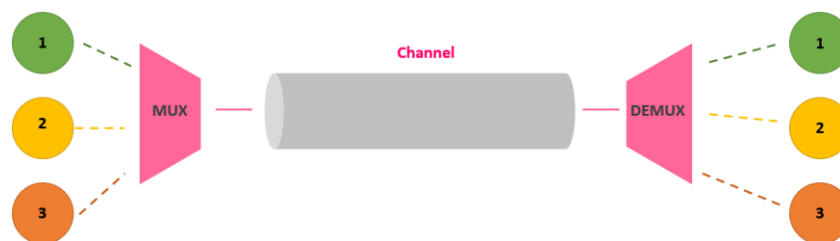
ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



Η πολυπλεξία (multiplexing) αποτελεί μία μέθοδο του φυσικού επιπέδου, διαμέσου της οποίας είναι δυνατή μετάδοση και λήψη πολλών διαφορετικών σημάτων δεδομένων μεταξύ πομπού και δέκτη, ταυτόχρονα. Η πολυπλεξία δύναται να εφαρμοστεί όταν ο ολικός αριθμός των σημάτων που θα διακινηθούν, είναι μικρότερος από το εύρος ζώνης. Σημαντικό γνώρισμα της πολυπλεξίας αποτελεί η διαφορετική φύση που μπορεί να έχουν τα σήματα που αποστέλλονται μεταξύ πομπού και δέκτη, καθώς αυτά ίσως περιέχουν διαφορετικά είδη πληροφοριών. Στον πομπό τα σήματα πολυπλέκονται με τρόπο, τέτοιοι ώστε να παραμένουν ασυσχέτιστα κατά την αποστολή τους, ενώ στο δέκτη το ληφθέν σήμα αποπολυπλέκεται. Οι δύο αυτές διαδικασίες καλούνται πολυπλεξία (multiplexing) και αποπολυπλεξία (demultiplexing) και πραγματοποιούνται μέσω συσκευών, του πολυπλέκτη (multiplexer) και αποπολυπλέκτη (demultiplexer). Παρακάτω θα παρουσιαστούν δύο βασικές μέθοδοι πολυπλεξίας, την πολυπλεξία διαίρεσης συχνότητας (Frequency division multiplexing – FDM) και η πολυπλεξία διαίρεσης χρόνου (Time division multiplexing – TDM)[36].

Πολυπλεξία Διαίρεσης Συχνότητας (FDM)

Η πολυπλεξία διαίρεσης συχνότητας (FDM), αποτελεί τεχνική η οποία πολυπλέκει συχνοτικά σήματα που μεταδίδονται από πομπό, σε σήμα με υψηλό εύρος ζώνης. Απαιτείται το εύρος ζώνης (bandwidth) του μέσου μετάδοσης να είναι μεγαλύτερο από το αντίστοιχο, του συνόλου των σημάτων που θα μεταδοθούν, ώστε να διαιρεθεί το διαθέσιμο το εύρος ζώνης σε μη επικαλυπτόμενες υποζώνες. Κυρίως, χρησιμοποιείται σε αναλογικά συστήματα, αλλά δύναται να εφαρμοστεί και σε ψηφιακά. Τα σήματα βασικής ζώνης, που επρόκειτο να μεταδοθούν, διαμορφώνονται με τέτοιο τρόπο ώστε να μεταφερθούν αυτά στις υποζώνες συχνοτήτων που εκχωρήθηκαν και να δημιουργηθούν ζωνοπερατά σήματα. Για να αποφευχθούν οι παρεμβολές μεταξύ των σημάτων, αφήνουμε κενό ανάμεσα σε γειτονικά σήματα, δηλαδή μία ζώνη φύλαξης (guard band), η οποία αποτελεί μια μικρή περιοχή συχνοτήτων που δεν χρησιμοποιείται για μετάδοση. Η επιτυχία της FDM οφείλεται σε αυτή, καθώς υπάρχει ορθογωνιότητα στο πεδίο των συχνοτήτων, ώστε να μην σχετίζονται τα σήματα συχνοτικά. Τέλος, με την άφιξη των σημάτων στον δέκτη, αυτά διαχωρίζονται από ένα ζωνοπερατό φίλτρο και εν συνεχεία διαμορφώνονται εκ νέου στο αρχικό σήμα.

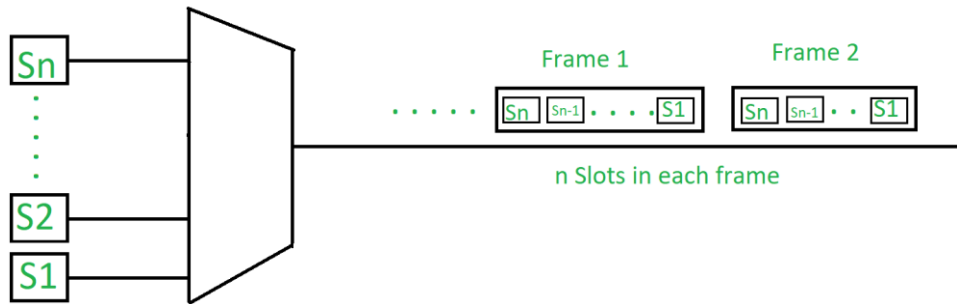


Εικόνα 9. Διάγραμμα FDM
Πηγή [38]

Πολυπλεξία Διαίρεσης Χρόνου (TDM)

Η συγκεκριμένη αποτελεί, μία ψηφιακή μέθοδο πολυπλεξίας που χρησιμοποιείται, για τον συνδυασμό πολλών σημάτων με χαμηλό ρυθμό μετάδοσης, σε σήμα με υψηλό ρυθμό μετάδοσης, το οποίο και μεταδίδεται. Το κανάλι μετάδοσης, διαχωρίζεται

χρονικώς σε πληθώρα μεταδιδόμενων σημάτων από ένα πομπό, χωρίζεται δηλαδή σε χρονοθυρίδες, στις οποίες κάθε σταθμός μεταδίδει μέρος των σημάτων του σε αυτές. Η TDM, διακρίνεται σε δύο μεθόδους τη synchronous TDM και asynchronous TDM. Στην μεν πρώτη, κάθε σήμα που μεταδίδεται αντιστοιχίζεται σε μία και μόνο χρονοθυρίδα που μπορεί να το μεταδώσει, με το μειονέκτημα ότι, εάν δεν υπάρχει σήμα προς μετάδοση, μένει ανεκμετάλλευτη. Στη δε δεύτερη, κάθε πομπός δίνει στον πολυπλέκτη δεδομένα, όταν αυτός έχει, τα οποία και μεταδίδονται στην πρώτη κενή χρονοθυρίδα. Έτσι, αυξάνεται η αποδοτικότητα του καναλιού, μεγαλώνοντας ωστόσο και την πολυπλοκότητα της πολυπλεξίας-αποπολυπλεξίας.



Εικόνα 10. Διάγραμμα TDM
Πηγή[39]

FDM	TDM
Υποστηρίζει αναλογική μετάδοση και υπό προϋποθέσεις και ψηφιακή μετάδοση.	Υποστηρίζει ψηφιακή μετάδοση.
Πολλάπλά σήματα συνδυάζονται παράλληλα.	Πολλάπλά σήματα στέλνονται διαδοχικά μέσω των χρονοθυρίδων.
Δεν απαιτείται συγχρονισμός.	Απαιτείται συγχρονισμός.
Δεν υπάρχουν καθυστερήσεις μετάδοσης.	Υφίστανται καθυστερήσεις μετάδοσης εξαιτίας της χρονοδρομολόγησης των δεδομένων.
Το κόστος υλοποίησης είναι μεγαλύτερο.	Το κόστος υλοποίησης είναι μικρότερο.
Χαμηλή αποδοτικότητα λόγω παρεμβολών.	Μεγαλύτερη αποδοτικότητα λόγω λιγότερων παρεμβολών.

Μέθοδοι Αμφιδρόμησης

Στις ασύρματες επικοινωνίες υπάρχουν τρεις βασικοί τρόποι επικοινωνίας μεταξύ δύο Σ/Α, η μονόπλευρη (simplex), ημι-αμφίδρομη (half-duplex) και αμφίδρομη (duplex) επικοινωνία. Στην simplex, γίνεται μετάδοση μόνο προς μία κατεύθυνση, το οποίο σημαίνει ότι όλο το εύρος ζώνης χρησιμοποιείται για τη μετάδοση δεδομένων, από το ένα εκ των δύο άκρων, από πομπό προς δέκτη. Αντίθετα, στην ημι-αμφίδρομη επικοινωνία, γίνεται μετάδοση και από τα δύο άκρα, αλλά όχι ταυτόχρονα, οπότε το εύρος ζώνης χρησιμοποιείται από τον πομποδέκτη που εκπέμπει κάθε φορά. Τέλος, στην αμφίδρομη επικοινωνία, μεταδίδεται πληροφορία ταυτόχρονα και από τα δύο άκρα, με τη διαδικασία που ονομάζεται αμφιδρόμηση. Προκειμένου να επιτευχθεί, υπάρχουν τρεις μέθοδοι, η αμφιδρόμηση διαίρεσης συχνότητας (FDD), η αμφιδρόμηση διαίρεσης χρόνου (TDD) και η ημι-αμφιδρόμηση διαίρεσης συχνότητας (Half FDD).

Αμφιδρόμηση διαίρεσης συχνότητας (FDD): Σε αυτή την τεχνική, οι επικοινωνίες μεταξύ πομπού και δέκτη γίνονται ταυτόχρονα, χρησιμοποιώντας δύο διαφορετικές συχνότητες—μία για την αποστολή (uplink) και μία για τη λήψη (downlink). Η ταυτόχρονη και αμφίδρομη, αυτή, επικοινωνία είναι ιδιαίτερα χρήσιμη σε εφαρμογές όπου απαιτείται συνεχής ροή δεδομένων και στις δύο κατευθύνσεις, όπως στις φωνητικές κλήσεις ή στις εφαρμογές συνεχούς ροής.



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



Αμφιδρόμηση Διαίρεσης Χρόνου (TDD): Η TDD χρησιμοποιεί την ίδια συχνότητα για αποστολή και λήψη, εναλλάσσοντας μεταξύ τους ανά συγκεκριμένα χρονικά διαστήματα. Το γεγονός αυτό, σημαίνει ότι η επικοινωνία είναι διαδοχική και όχι ταυτόχρονη. Η TDD προσφέρει ευελιξία στην κατανομή πόρων, επιτρέποντας τη δυναμική ρύθμιση του χρόνου που διατίθεται για uplink και downlink, ανάλογα με τις ανάγκες.

Ημι-αμφιδρόμηση Διαίρεσης Συχνότητας (Half FDD): Η συγκεκριμένη μέθοδος, συνδυάζει στοιχεία των προηγούμενων μεθόδων, χρησιμοποιώντας δύο διαφορετικές συχνότητες για uplink και downlink, αλλά χωρίς να γίνεται ταυτόχρονη επικοινωνία. Η αποστολή και η λήψη εναλλάσσονται, προσφέροντας ένα συμβιβασμό σε περιπτώσεις όπου η πλήρης αμφιδρόμηση δεν είναι απαραίτητη ή είναι υπερβολικά απαιτητική σε πόρους.

Πολλαπλή Πρόσβαση στο Μέσο Επικοινωνίας

Οι μέθοδοι της πολυπλεξίας και της αμφιδρόμησης που περιγράφηκαν ανωτέρω, καθορίζουν με σαφήνεια τα προβλήματα της ταυτόχρονης μετάδοσης σημάτων-δεδομένων και την επικοινωνία μεταξύ πομποδεκτών, αντίστοιχα. Ωστόσο, αφορούν μόνο δύο πομποδέκτες, σε πραγματικές συνθήκες, υπάρχει η ανάγκη για επικοινωνία πολλαπλών συσκευών και η μεταξύ τους μετάδοση δεδομένων. Παράδειγμα αποτελεί και η παρούσα εργασία, όπου ένα μέρος της εξετάζει την επικοινωνία ενός UAV με έναν αριθμό επίγειων χρηστών -αρμάτων μάχης. Για το σενάριο αυτό, απαραίτητη είναι η ταυτόχρονη χρήση του δικτύου από πολλαπλές συσκευές, μέσω μεθόδων πολλαπλής πρόσβασης στο μέσο. Υπάρχουν τρεις μέθοδοι, **η μέθοδος επιμερισμού του μέσου μετάδοσης, η μέθοδος τυχαίας πρόσβασης στο μέσο και η μέθοδος ελεγχόμενης πρόσβασης στο μέσο.**

Μέθοδος Επιμερισμού του Μέσου Μετάδοσης: Η μέθοδος αυτή, περιλαμβάνει την κατανομή των διαθέσιμων πόρων του μέσου μετάδοσης όπως συχνότητες, χρόνο ή κωδικούς σε διάφορους χρήστες, με τις εξής τεχνικές:

- **FDMA (Frequency Division Multiple Access):** Κάνει χρήση διαφορετικών συχνοτήτων για διαφορετικούς χρήστες.
- **TDMA (Time Division Multiple Access):** Κάνει χρήση διαφορετικών χρονικών διαστημάτων για διάφορους χρήστες.
- **CDMA (Code Division Multiple Access):** Χρησιμοποιεί διαφορετικούς κωδικούς για το διαχωρισμό της μετάδοσης από διαφορετικών χρηστών.
- **OFDMA (Orthogonal Frequency Division Multiple Access):** Χρησιμοποιούνται υπο-κανάλια ώστε πολλοί χρήστες να στέλνουν δεδομένα ταυτόχρονα.

Μέθοδος Τυχαίας Πρόσβασης στο Μέσο: Περιλαμβάνει προσεγγίσεις όπου οι χρήστες μπορούν να μεταδώσουν δεδομένα κάθε στιγμή, χωρίς να υπάρχει κάποια προγραμματισμένη διάθεση του μέσου. Ορισμένες κοινές μέθοδοι τυχαίας πρόσβασης περιλαμβάνουν:



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



- **ALOHA:** Οι χρήστες στέλνουν δεδομένα όποτε επιθυμούν και αν υπάρξει σύγκρουση, περιμένουν τυχαία χρονικά διαστήματα πριν προσπαθήσουν εκ νέου.
- **CSMA (Carrier Sense Multiple Access):** Οι χρήστες "ακούν" το μέσο πριν στείλουν δεδομένα και αν είναι ελεύθερο, τα στέλνουν. Αν υπάρχει σύγκρουση, περιμένουν και δοκιμάζουν εκ νέου.
 - **CSMA/CD (Collision Detection):** Χρησιμοποιείται σε ενσύρματα δίκτυα (π.χ. Ethernet) για ανίχνευση συγκρούσεων.
 - **CSMA/CA (Collision Avoidance):** Χρησιμοποιείται σε ασύρματα δίκτυα (π.χ. Wi-Fi) για να αποφυγή συγκρούσεων πριν την αποστολή των σημάτων.

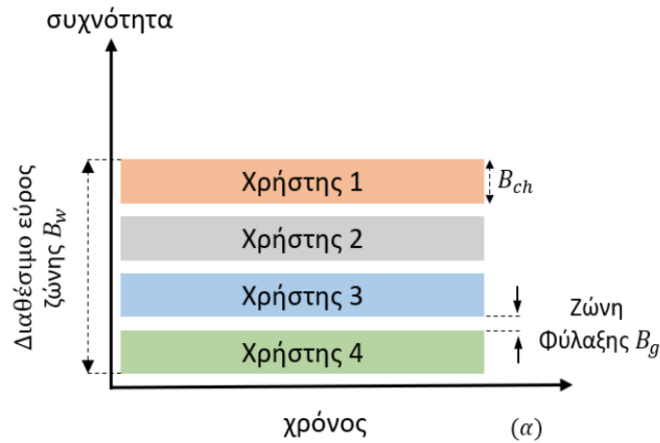
Μέθοδος Ελεγχόμενης Πρόσβασης στο Μέσο: Περιλαμβάνεται ο έλεγχος πρόσβασης στο μέσο από έναν κεντρικό διαχειριστή ή διαμέσου κανόνων που καθορίζουν το χρόνο και τρόπο, που οι χρήστες μπορούν να στείλουν δεδομένα. Οι τεχνικές που χρησιμοποιούνται είναι:

- **Polling:** Ο κεντρικός ελεγκτής προσκαλεί περιοδικά κάθε χρήστη αν έχει δεδομένα να μεταδώσει με μια προκαθορισμένη διαδοχική σειρά.
- **Token Passing:** Ένα "token" κυκλοφορεί μεταξύ των χρηστών. Μόνο ο χρήστης που κατέχει το token μπορεί να στείλει δεδομένα.
- **Time Division Access:** Προγραμματισμένα χρονικά διαστήματα για κάθε χρήστη, συνδυάζοντας στοιχεία του TDMA με ελεγχόμενη πρόσβαση.

Από τα παραπάνω, μπορούμε να καταλάβουμε ότι ιδανική μέθοδος για την περίπτωση χρήσης που εξετάζουμε, είναι η μέθοδος επιμερισμού του μέσου επικοινωνίας. Η επιλογή αυτή, είναι δικαιολογημένη καθώς το δίκτυο αποτελείται από χρήστες με αυξημένες απαιτήσεις όσον αφορά τη μεταφορά δεδομένων, οπότε και απαιτείται να αποφεύγονται συγκρούσεις κατά τη μετάδοση τους. Θα αναλυθούν λοιπόν παρακάτω οι δύο τεχνικές της μεθόδου αυτής, η FDMA και η TDMA.

Πολλαπλή Πρόσβαση Διαίρεσης Συχνότητας (Frequency Division Multiple Access)

Η πολλαπλή πρόσβαση διαίρεσης συχνότητας, αποτελεί την παλαιότερη και απλούστερη μέθοδο πολλαπλής πρόσβασης. Το διαθέσιμο εύρος ζώνης διαχωρίζεται σε πολλαπλές και μη επικαλυπτόμενες ζώνες συχνοτήτων, με την κάθε ζώνη να ανατίθεται σε έναν επιμέρους χρήστη, προκειμένου να μεταδίδει δεδομένα. Η ανάθεση αυτή, πραγματοποιείται στην έναρξη της μετάδοσης των δεδομένων και παραμένει ίδια έως το τέλος αυτής. Υπάρχει ένας κεντρικός ελεγκτής, ο οποίος αναθέτει τη ζώνη συχνοτήτων σε κάθε χρήστη, αναλόγως των αναγκών τους. Προκειμένου όμως, να μην υπάρξουν παρεμβολές, μεταξύ γειτονικών ζωνών υπάρχουν ζώνες φύλαξης.



Εικόνα 11. Σχηματική Αναπαράσταση FDMA
Πηγή [37]

Επίσης, στο FDMA απαιτείται αμφιδρόμηση για επικοινωνία ανάμεσα σε πομπό και δέκτη, οπότε και το FDMA συνδυάζεται με αμφιδρόμηση διαίρεσης συχνότητας (FDD). Η αρχιτεκτονική αυτή του FDMA, το καθιστά ιδανικό για επικοινωνίες φωνής, ωστόσο είναι προβληματικό όσον αφορά τη χρήση ισχύος και εύρους ζώνης. Στο FDMA, όταν ένα κανάλι δεν χρησιμοποιείται από χρήστη, δεν μπορεί να εκμεταλλευτεί από κάποιον άλλο, οπότε αδρανεί. Παρακάτω, παρουσιάζεται συγκεντρωτικός πίνακας με τα πλεονεκτήματα και μειονεκτήματα του FDMA.

Πίνακας 1. Αξιολόγηση FDMA

Πλεονεκτήματα	Μειονεκτήματα
Απλό στην υλοποίηση από άποψη υλικού	Υψηλότερη σπατάλη εύρους ζώνης σε περίπτωση αχρησίας
Χαμηλότερη πολυπλοκότητα συστήματος σε σύγκριση με το TDMA	Δεν είναι κατάλληλο για ψηφιακές μεταδόσεις
Αύξηση χωρητικότητας με αποδοτική κωδικοποίηση και χαμηλό ρυθμό δεδομένων	Οι ζώνες φύλαξης μειώνουν το διαθέσιμο εύρος ζώνης
Μειώνει τη διασυμβολική παρεμβολή	Σταθερός και μικρός μέγιστος ρυθμός δεδομένων ανά κανάλι
Δεν απαιτεί ισοστάθμιση λόγω μειωμένης διασυμβολικής παρεμβολής	Έλλειψη ευελιξίας σε υψηλούς ρυθμούς δεδομένων
Κατάλληλο όταν υπάρχουν λίγοι χρήστες και η κίνηση είναι σταθερή	Σπατάλη εύρους ζώνης όταν ένα κανάλι παραμένει αδρανές
Κατάλληλο για αναλογική μετάδοση δεδομένων	Δεν επιτρέπει εύκολη χρήση μεταβαλλόμενων ρυθμών δεδομένων

Πολλαπλή Πρόσβαση Διαίρεσης Χρόνου (Time Division Multiple Access)

Αντίθετα με το FDMA, η TDMA διαχωρίζει χρονικά τη χρήση του μέσου μετάδοσης από τους χρήστες. Το διαθέσιμο εύρος ζώνης κάθε καναλιού, χρησιμοποιείται εξ 'ολοκλήρου, από το σύνολο των χρηστών, με διαδοχική σειρά. Μοιράζεται το κανάλι δηλαδή, σε αλληλοεπικαλυπτόμενες χρονοθυρίδες (slots) και χρησιμοποιείται από κάθε χρήστη ιεραρχικά, ο οποίος κάνει χρήση μόνο της χρονοθυρίδας που του έχει διατεθεί. Απαραίτητη προϋπόθεση, αποτελεί η ακριβής γνώση της αρχής και τέλους της χρονοθυρίδας, για αυτό και απαιτείται ο συγχρονισμός μεταξύ των χρηστών. Αυτό διασφαλίζεται με τα διαστήματα φύλαξης, καθώς

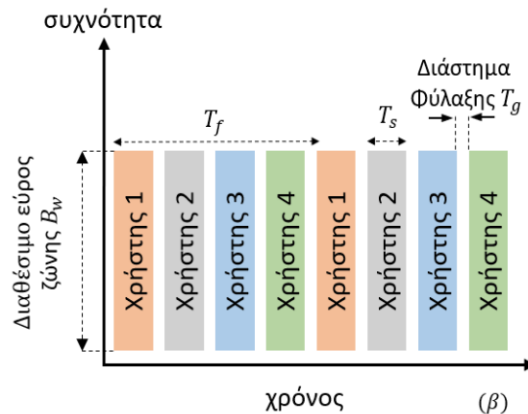


ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



εξασφαλίζουν ότι δεν θα επικαλύπτονται χρονικά οι μεταδόσεις και δεν θα υπάρχουν παρεμβολές, μεταξύ των χωρικά διασκορπισμένων χρηστών.



Εικόνα 12. . Σχηματική Αναπαράσταση TDMA
Πηγή [37]

Δύναται να χρησιμοποιηθεί και η αμφίδρομη διαίρεση συχνότητας (FDD) ή χρόνου (TDD), σε συνδυασμό με το TDMA, ώστε να διαχωριστεί το κανάλι ανερχόμενης ζεύξης με το αντίστοιχο της κατερχόμενης ζεύξης σε μια μετάδοση δεδομένων. Οπότε δημιουργούνται τα TDMA/FDD και TDMA/TDD αντίστοιχα.

Τα δεδομένα κάθε χρήστη αντιστοιχίζονται σε μία χρονοθυρίδα (slot), οι οποίες αποτελούν ένα πλαίσιο (frame). Όλες οι θυρίδες απαρτίζουν το μήνυμα της πληροφορίας. Στην αρχή του πλαισίου, βρίσκεται το προοίμιο, το οποίο αποτελείται από τις διευθύνσεις των πομποδεκτών και πληροφορίες σχετικά με τον συγχρονισμό. Κάθε χρήστης, μπορεί να χρησιμοποιήσει παραπάνω από μία χρονοθυρίδες για τη μετάδοση των δεδομένων, γεγονός που εξασφαλίζει ποικιλία διαθέσιμων ρυθμών μετάδοσης. Τέλος, για κάθε slot, υπάρχουν τα bits φύλαξης, που επιτρέπουν τον συγχρονισμό των δεκτών ανάμεσα σε διαφορετικές θυρίδες και πλαίσια.

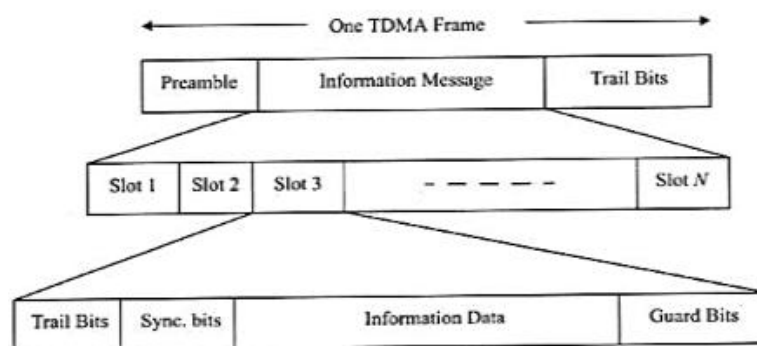


Fig: TDMA frame structure

Εικόνα 13. Διάγραμμα TDMA
Πηγή [37]

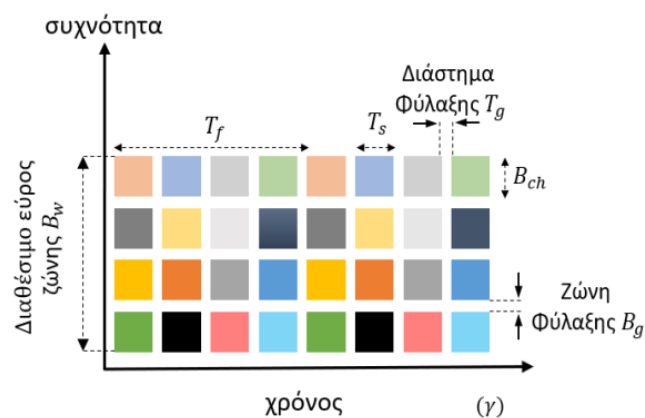
Παρακάτω ακολουθεί συγκεντρωτικός πίνακας με τα πλεονεκτήματα και μειονεκτήματα του TDMA

Πίνακας 2. Αξιολόγηση TDMA

Πλεονεκτήματα	Μειονεκτήματα
Επιτρέπει ευέλικτο ρυθμό μετάδοσης δεδομένων, τόσο πολλαπλάσιο όσο και υποπολλαπλάσιο του βασικού ρυθμού	Απαιτεί υψηλή ισχύ για τη μετάδοση, ειδικά στην ανερχόμενη ζεύξη
Προσφέρει παρακολούθηση ισχύος σήματος και ρυθμού δυαδικών σφαλμάτων πλαίσιο προς πλαίσιο	Απαιτεί επεξεργασία σήματος για συγχρονισμένη διαμόρφωση και αποδιαμόρφωση
Χρησιμοποιεί το εύρος ζώνης αποτελεσματικά, χωρίς συχνοτικές ζώνες φύλαξης μεταξύ των καναλιών	Απαιτεί συγχρονισμό, και σε περίπτωση απώλειας, τα κανάλια μπορεί να συγκρουστούν
Εκπέμπει με χρονικά διαστήματα φύλαξης για την αντιμετώπιση αστάθειας συγχρονισμού και καθυστερήσεων	Ο χρόνος διάδοσης ποικίλλει ανάλογα με την απόσταση, επηρεάζοντας τη λειτουργία του TDMA
Κατάλληλο για ψηφιακές μεταδόσεις σε κυψελοειδή δίκτυα	Υποστηρίζει μόνο ψηφιακή μετάδοση

Συνδυασμένη Πολλαπλή Πρόσβαση Διαίρεσης Συχνότητας και Χρόνου (Narrow Band TDMA – NTDMA)

Πέραν των δύο προαναφερθέντων μεθόδων πολλαπλής πρόσβασης σε μέσο, έχει δημιουργηθεί και υβριδική μέθοδος, η οποία συνδυάζει τα πλεονεκτήματα αυτών. Η μέθοδος αυτή, καλείται FDMA/TDMA, αλλά στη βιβλιογραφία αναφέρεται και ως TDMA στενής ζώνης (narrow band TDMA – NTDMA). Στο NTDMA, το εύρος ζώνης που διατίθεται διαχωρίζεται σε μικρότερες ζώνες συχνότητας, οι οποίες διαιρούνται σε αριθμό χρονοθυρίδων. Αξιοποιεί δηλαδή, την χρονική κατανομή του TDMA, χωρίζοντας το φάσμα σε χρονικές θυρίδες, ενώ πρώτα έχει δημιουργήσει μικρότερα εύρη ζώνης, στα πρότυπα του FDMA. Με αυτό τον τρόπο, καθίσταται εφικτός ο διαμοιρασμός του καναλιού από πληθώρα χρηστών, χωρίς την ύπαρξη παρεμβολών, με την ύπαρξη ωστόσο ακριβούς συγχρονισμού.



Εικόνα 14. Σχηματική Αναπαράσταση NTDMA
Πηγή [37]

Στο NTDMA, το ολικό εύρος ζώνης B_w διαιρείται σε N_{ch} υποζώνες, με εύρος ζώνης B_{ch} . Κάθε μία χωρίζεται από τη γείτονα με ζώνες φύλαξης B_g . Οπότε, το συνολικό εύρος ζώνης δίνεται από τη σχέση:



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



$$B_w = N_{ch}B_{ch} + (N_{ch} + 1)B_G$$

Η αποδοτικότητα, όσον αφορά το φάσμα, δίδεται

$$n_{NTDMA} = \frac{(T_f - T_p - T_t)N_I(B_w - (N_{ch} + 1)B_G)}{T_f N_S B_w}$$

Συνεπώς ο μέγιστος αριθμός που το NTDMA, δύναται να υποστηρίξει αποτελεί γινόμενο των υποζωνών με τις θυρίδες ανά κάθε υποζώνη, $N_{user}=N_{ch}N_{slot}$.

Σύγκριση Μεθόδων TDMA και FDMA σε Στρατιωτικές Επικοινωνίες

Παραπάνω, αναλύθηκαν οι τρεις βασικές μέθοδοι πολλαπλής πρόσβασης στο μέσο, TDMA, FDMA NTDMA. Η κύρια διαφορά μεταξύ TDMA και FDMA, είναι η διαφορετική προσέγγιση σχετικά με τη διαίρεση του διαθέσιμου φάσματος, σε χρονοθυρίδες και διαφορετικές υποζώνες, αντίστοιχα. Στο TDMA δηλαδή, κάθε χρήστης χρησιμοποιεί την ίδια συχνότητα, αλλά σε διαφορετική χρονική στιγμή, γεγονός που απαιτεί ακριβή συγχρονισμό μεταξύ των χρηστών, προσφέροντας υψηλή απόδοση φάσματος και χαμηλή κατανάλωση ενέργειας. Αντίθετα, στο FDMA σε κάθε χρήστη έχει εκχωρηθεί μία μοναδική συχνότητα ώστε, να επικοινωνεί ταυτόχρονα με άλλους χρήστες, το οποίο είναι μεν απλό στην υλοποίηση αλλά δεν αξιοποιεί το διαθέσιμο φάσμα, απαιτώντας μεγαλύτερη ζώνη συχνοτήτων για την υποστήριξη πολλαπλών χρηστών.

Στις στρατιωτικές επικοινωνίες, απαιτείται η άμεση ανταλλαγή δεδομένων από πλήθος χρηστών, οι οποίοι βρίσκονται στο ίδιο δίκτυο. Η δυνατότητα αυτή είναι κρίσιμη, καθώς οι χρήστες βρίσκονται σε ταχέως εξελισσόμενα περιβάλλοντα και επιχειρούν σε διάφορες αποστάσεις, οι οποίες συχνά δεν έχουν οπτική επαφή μεταξύ τους. Προκειμένου να επιλεγεί η κατάλληλη μέθοδος πολλαπλής πρόσβασης γίνεται ανάλυση πολλών παραγόντων, όπως ο αριθμός χρηστών, η ταχύτητα μετάδοσης και η ανθεκτικότητα σε παρεμβολές. Η συνηθέστερη μέθοδος που χρησιμοποιείται για την αντιμετώπιση παρεμβολών και υποκλοπών, είναι η διασπορά του φάσματος με την αναπήδηση συχνότητας. Όσον αφορά τις μεθόδους πολλαπλής πρόσβασης στο μέσο, το TDMA, προσφέρει υψηλή απόδοση φάσματος, λόγω της χρονικής οργάνωσης, που το καθιστά ιδανικό για περιβάλλοντα με υψηλή πυκνότητα χρηστών, όπως η επικοινωνία μεταξύ UAVs και αρμάτων. Απαιτείται, απόλυτη ακρίβεια συγχρονισμού για την αποφυγή παρεμβολών και την ομαλή ροή πληροφοριών. Αντίθετα, το FDMA αποτελεί απλή και σταθερή λύση σε εφαρμογές όπου δεν απαιτείται αυστηρός συγχρονισμός, απαιτεί ωστόσο, μεγαλύτερη ζώνη συχνοτήτων που το καθιστά λιγότερο αποδοτικό σε περιπτώσεις όπου το φάσμα πρέπει να είναι περιορισμένο. Τέλος, μία σύγχρονη προσέγγιση είναι το NTDMA, το οποίο συνδυάζει χαρακτηριστικά των δύο παραπάνω πρωτοκόλλων. Το NTDMA (Non-contiguous Time Division Multiple Access), επιτρέπει σε χρήστες να εκμεταλλεύονται διαφορετικά τμήματα του χρόνου και του φάσματος, δίχως να υπάρχει μία αυστηρή χρονική διαίρεση. Αποτελεί ουσιαστικά μια παραλλαγή του TDMA όπου οι χρονικές σχισμές (time slots) που αποδίδονται σε έναν χρήστη δεν είναι συνεχόμενες, αλλά κατανομημένες δυναμικά σε διαφορετικά χρονικά διαστήματα. Έτσι, παρέχεται μεγαλύτερη ευελιξία και ανθεκτικότητα σε συνθήκες ηλεκτρονικού πολέμου, γεγονός που την καθιστά κατάλληλη για στρατιωτικές εφαρμογές.



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



Εστιάζοντας στις στρατιωτικές εφαρμογές, το FDMA λειτουργεί τόσο σε αναλογικά όσο και σε ψηφιακά συστήματα, είναι ανθεκτικό σε παρεμβολές, αφού κάθε κανάλι είναι ανεξάρτητο και δεν απαιτεί αυστηρό συγχρονισμό, σε αντίθεση με το TDMA. Λόγω της βασικής αρχής λειτουργίας του έχει χαμηλότερη απόδοση φάσματος, επειδή όμως κάθε χρήστης δεσμεύει μια συχνότητα συνεχώς, είναι ευάλωτο σε παρεμβολές (jamming) και δεν διευκολύνει την κρυπτογράφηση. Συνήθως προτιμάται στις HF, VHF και UHF επικοινωνίες για στρατηγικές και τακτικές αποστολές, στις ραδιοεπικοινωνίες πεδίου μάχης για συντονισμό μονάδων και σε επικοινωνίες αεροσκαφών και πλοίων που απαιτούν μεγαλύτερη εμβέλεια.

Το TDMA υποστηρίζει κρυπτογράφηση και anti-jamming τεχνικές, κάνοντας αποδοτική χρήση του φάσματος, αφού οι χρήστες μοιράζονται την ίδια συχνότητα παρέχοντας, παράλληλα, υψηλή ασφάλεια, λόγω εύκολης ενσωμάτωσης κρυπτογράφησης και μεταβλητών σχισμών. Συνολικά, εξασφαλίζει καλύτερη αντοχή σε ηλεκτρονικό πόλεμο (EW). Λόγω της βασικής αρχής λειτουργίας του είναι ευαίσθητο σε χρονικές καθυστερήσεις, εξαιτίας της ανάγκης για συγχρονισμό, ενώ δεν υποστηρίζει πολύ υψηλές ταχύτητες δεδομένων λόγω του περιορισμένου χρόνου ανά χρήστη. Προς τούτο, προτιμάται σε στρατιωτικές δορυφορικές επικοινωνίες (MILSATCOM), συστήματα τακτικών επικοινωνιών μεταξύ στρατιωτικών μονάδων καθώς και σε αεροπορικές και ναυτικές επιχειρήσεις, όπου απαιτείται συντονισμός πολλαπλών χρηστών[48].

Το NTDMA βασίζεται στο TDMA, αλλά οι χρονικές σχισμές δεν είναι συνεχόμενες. Δίνει δυναμικά περισσότερες ή λιγότερες σχισμές σε κάθε χρήστη, ανάλογα με τις ανάγκες και χρησιμοποιείται σε ευέλικτα και προσαρμόσιμα δίκτυα.

Η δυναμική κατανομή πόρων εξασφαλίζει βελτιωμένη απόδοση φάσματος, μεγαλύτερη αντοχή σε ηλεκτρονικό πόλεμο (EW), καθώς οι σχισμές αλλάζουν συνεχώς καθιστώντας αυτό, κατάλληλο για εφαρμογές με υψηλές απαιτήσεις δεδομένων, όπως στρατιωτικά UAVs και δίκτυα αισθητήρων. Ο αλγόριθμος διαχείρισης όμως για τη δυναμική κατανομή των σχισμών είναι πολύπλοκος και απαιτεί ισχυρό συγχρονισμό, όπως το TDMA, ενώ χρειάζεται προηγμένα ραδιοσυστήματα για την υλοποίηση. Έτσι, είναι κατάλληλο για δορυφορικά στρατιωτικά δίκτυα (MILSATCOM) με υψηλό φορτίο δεδομένων, σε δίκτυα μάχης (Tactical Networks) για άμεση επικοινωνία και ανίχνευση απειλών και σε ασύρματα δίκτυα αισθητήρων για συλλογή και μετάδοση πληροφοριών πεδίου[49].



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



Πίνακας 3. Συγκριτική Αξιολόγηση FDMA, TDMA, NTDMA

Χαρακτηριστικό	TDMA	FDMA	NTDMA
Τύπος Διαμοιρασμού	Χρόνος	Συχνότητα	Δυναμικός Χρόνος
Ανθεκτικότητα σε Jamming	Υψηλή (με scrambling)	Χαμηλή	Πολύ Υψηλή
Απαιτήσεις Συγχρονισμού	Υψηλές	Χαμηλές	Πολύ Υψηλές
Αξιοποίηση Φάσματος	Υψηλή	Μέτρια	Πολύ Υψηλή
Ευελιξία & Κλιμάκωση	Μέτρια	Χαμηλή	Πολύ Υψηλή
Απαιτήσεις Υποδομής	Υψηλές	Χαμηλές	Πολύ Υψηλές
Εφαρμογή σε Δίκτυα Δεδομένων	Μέτρια	Χαμηλή	Πολύ Υψηλή
Ιδανικό για	Δορυφορικά & κινητά δίκτυα	Σταθερές επικοινωνίες	Δίκτυα υψηλής προσαρμοστικότητας

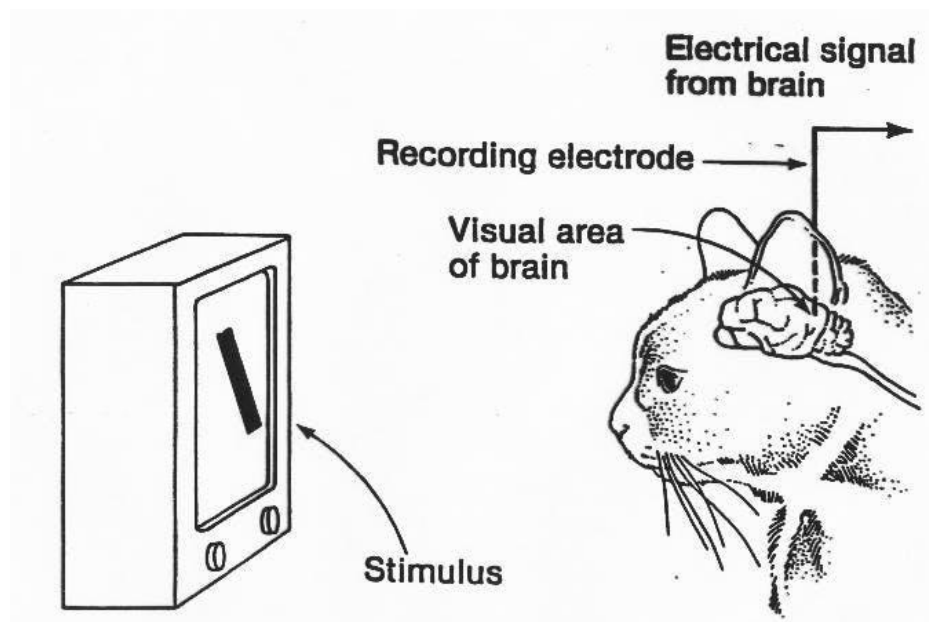
Συνοψίζοντας, καταλήγουμε στο συμπέρασμα ότι ιδανική για ένα περιβάλλον στρατιωτικών επιχειρήσεων ιδανικό είναι το πρωτόκολλο TDMA και το σύγχρονό του NTDMA. Το TDMA, παρέχει μια πιο δομημένη και αποδοτική χρήση του φάσματος, ενώ αποκλείει το ενδεχόμενο της ταυτόχρονης μετάδοσης δεδομένων, λόγω του διαχωρισμού του φάσματος σε χρονικές ζώνες για κάθε χρήστη. Έτσι, είναι δυνατή η μετάδοση δεδομένων μεταξύ χρηστών και UAV, ώστε να παραμένουν ενημερωμένοι με τα πιο πρόσφατα δεδομένα. Τέλος, το NTDMA να παρέχει έναν συνδυασμό ευελιξίας και απόδοσης, συνδυάζοντας το διαχωρισμό φάσματος και χρόνου.

3. Υπολογιστική Όραση

Η υπολογιστική όραση (computer vision), αποτελεί ένα πεδίο τεχνητής νοημοσύνης (Artificial Intelligence - AI) που αξιοποιώντας τις δυνατότητες της μηχανικής μάθησης και των Νευρωνικών Δικτύων, εκπαιδεύει υπολογιστές και συστήματα να αντλούν σημαντικές πληροφορίες από ψηφιακές εικόνες, βίντεο και άλλες οπτικές πηγές — και να προτείνουν λύσεις ή να αναλαμβάνουν δράση όταν ανιχνεύουν ελαττώματα ή προβλήματα. Εάν η τεχνητή νοημοσύνη δίνει τη δυνατότητα στους υπολογιστές να σκέφτονται, η υπολογιστική όραση τους επιτρέπει να βλέπουν, να παρατηρούν και να κατανοούν το ανθρώπινο περιβάλλον.

3.1 Ιστορική Αναδρομή

Για περίπου 60 χρόνια, επιστήμονες και μηχανικοί, επιχειρούν να αναπτύξουν τρόπους ώστε οι μηχανές να μπορούν να βλέπουν και κατανοούν οπτικά δεδομένα. Οι πρώτες απόπειρες ξεκίνησαν το 1959 από τους πρωτοπόρους νευροφυσιολόγους, David Hubel και Torsten Wiesel. Στην έρευνα τους «Receptive fields of single neurons in the cat's striate cortex» έδειξαν σε μια γάτα μια σειρά από εικόνες, προσπαθώντας να συσχετίσουν μια απόκριση στον εγκέφαλό της. Ανακάλυψαν, λοιπόν ότι ανταποκρινόταν πρώτα σε σκληρές γωνίες ή γραμμές και επιστημονικά, αυτό σήμαινε ότι η επεξεργασία εικόνας αρχίζει πρώτα, με απλά σχήματα, όπως απλές γωνίες.

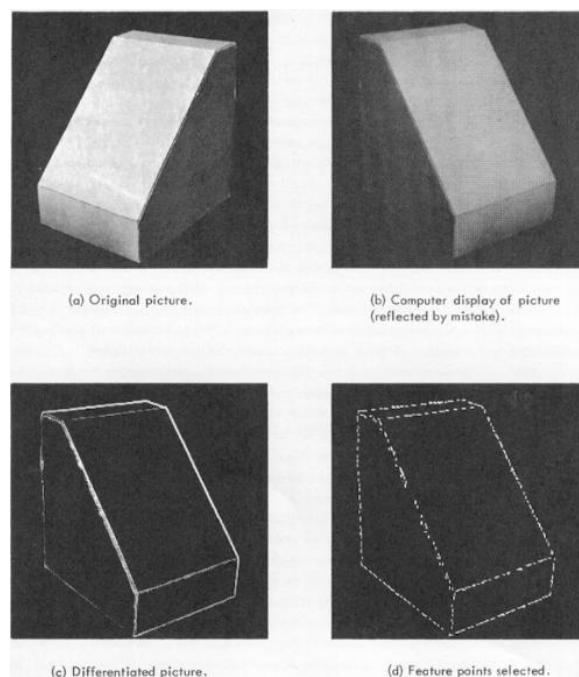


Εικόνα 15. Αναπαράσταση πειράματος γάτας
Πηγή [40]

Την ίδια περίπου εποχή, αναπτύχθηκε η πρώτη τεχνολογία σάρωσης εικόνων με τη χρήση υπολογιστή, η οποία επέτρεψε στους υπολογιστές να ψηφιοποιούν και να αποκτούν εικόνες. Ένα ακόμη, σημαντικό ορόσημο επιτεύχθηκε το 1963 όταν οι υπολογιστές κατάφεραν να μετατρέψουν τις δισδιάστατες εικόνες σε τρισδιάστατες μορφές. Το 1974, αποτελεί καθοριστικό έτος καθώς, εισήχθη η τεχνολογία οπτικής αναγνώρισης χαρακτήρων (Optical Character Recognition - OCR), μέσω της οποίας κατέστη εφικτή, η αναγνώριση κειμένου τυπωμένου σε οποιαδήποτε επιφάνεια ή

γραμματοσειρά. Παράλληλα, η έξυπνη αναγνώριση χαρακτήρων (Intelligent Character Recognition - ICR) δύναται να αποκρυπτογραφήσει χειρόγραφο κείμενο με τη χρήση Νευρωνικών Δικτύων. Έκτοτε OCR και ICR, έχουν χρησιμοποιούνται στην επεξεργασία εγγράφων και τιμολογίων, στην αναγνώριση πινακίδων οχήματος, στις πληρωμές μέσω κινητού τηλεφώνου, στη μετατροπή μηχανημάτων και σε άλλες κοινές εφαρμογές.

Το 1982, ο νευροεπιστήμονας David Marr, μέσω πειραμάτων που πραγματοποίησε, προχώρησε στη διαπίστωση ότι η όραση λειτουργεί ιεραρχικά, οπότε και εισήγαγε αλγόριθμους για μηχανές προκειμένου να ανιχνεύουν άκρες, γωνίες, καμπύλες και όμοια βασικά σχήματα. Την ίδια περίοδο, ο επιστήμονας υπολογιστών Kunihiro Fukushima ανέπτυξε ένα δίκτυο κυττάρων, τα οποία μπορούσαν να αναγνωρίσουν διάφορα μοτίβα. Το δίκτυο αυτό, ονομάζεται Neocognitron και περιελάμβανε συνελκτικά στρώματα σε ένα Νευρωνικό Δίκτυο (ΝΔ).



Εικόνα 16. Διαδικασία επεξεργασίας εικόνας
Πηγή [40]

Μέχρι το 2000, η αναγνώριση αντικειμένων αποτελούσε το επίκεντρο της μελέτης και έως το 2001, εμφανίστηκαν οι πρώτες εφαρμογές αναγνώρισης προσώπου σε πραγματικό χρόνο. Κατά συνέπεια της προόδου αυτής, στη δεκαετία του 2000, εμφανίστηκε η τυποποίηση του τρόπου επισήμανσης και σχολιασμού των συνόλων οπτικών δεδομένων. Το 2010, έγινε διαθέσιμο στο ευρύ κοινό, το σύνολο δεδομένων ImageNet, το οποίο περιείχε εκατομμύρια εικόνες με ετικέτα σε χιλιάδες κατηγορίες αντικειμένων. Παρείχε λοιπόν, τη βάση για Τεχνητά Νευρωνικά Δίκτυα – ΤΝΔ (Convolutional Neural Network – CNN) και μοντέλα βαθιάς μάθησης που χρησιμοποιούνται ακόμη και σήμερα. Το 2012, μια ομάδα από το Πανεπιστήμιο του Τορόντο συμμετείχε σε έναν διαγωνισμό αναγνώρισης εικόνας, με ένα CNN μοντέλο, εν ονόματι AlexNet. Το συγκεκριμένο μοντέλο, μείωσε σημαντικά το ποσοστό σφάλματος για την αναγνώριση εικόνας, με αποτέλεσμα, αυτή η ανακάλυψη να



αποτελέσει το εφελτήριο προόδου ώστε, τα ποσοστά σφαλμάτων έχουν μειωθεί σε λίγα μόνο τοις εκατό.[5]

3.2 Θεωρητικό Υπόβαθρο για Νευρωνικά Δίκτυα – ΝΔ (Neural Networks - NN)

Όπως αναφέρθηκε πρότερα, η υπολογιστική όραση επιτρέπει στον υπολογιστή να βλέπει και να κατανοεί το περιβάλλον γύρω του. Προκειμένου όμως, να αναπτυχθούν τα μοντέλα αυτά απαιτείται η ύπαρξη μεγάλου όγκου δεδομένων για την εκπαίδευση τους. Λύση στο πρόβλημα αυτό, δόθηκε το 2010 με την εφαρμογή των συνελκτικών νευρωνικών δικτύων (Convolutional Neural Networks – CNNs), τα οποία μέσω της βαθιάς μηχανικής μάθησης, αναλύουν το οπτικό υλικό διακρίνοντας πρώτα απλές γραμμές και έπειτα σημασιολογικά χαρακτηριστικά. Μιμούνται έτσι, τον τρόπο που ο άνθρωπος βλέπει ένα αντικείμενο από μακριά.

3.2.1 Νευρωνικά Δίκτυα – Perceptron

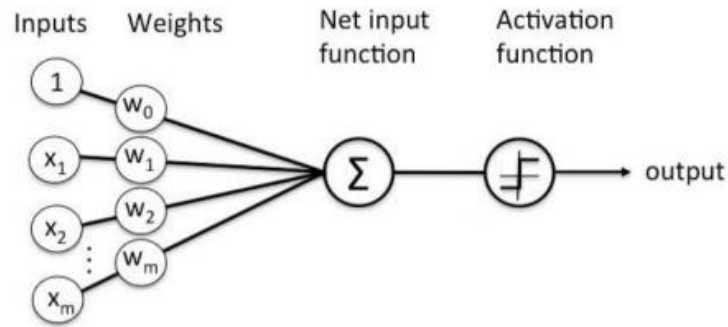
Τα νευρωνικά δίκτυα, ή Τεχνητά Νευρωνικά Δίκτυα (TNN), μιμούνται τη λειτουργία του ανθρώπινου εγκεφάλου, τον τρόπο που οι νευρώνες επικοινωνούν. Το νευρωνικό δίκτυο λαμβάνει γνώση από το περιβάλλον του, μέσω μίας διαδικασίας μάθησης και η αποκτηθείσα γνώση αποθηκεύεται στις ενώσεις μεταξύ των νευρώνων, οι οποίες ονομάζονται συναπτικά βάρη. Υπάρχουν τρεις κύριες κατηγορίες μηχανικής μάθησης του νευρωνικού δικτύου, ανάλογες του είδους των δεδομένων εισόδου:

Επιβλεπόμενη Μάθηση (Supervised Learning): Στην Επιβλεπόμενη μάθηση, χρησιμοποιούνται τα πρότυπα εισόδου x_1, x_2, x_n καθώς και οι αντίστοιχοι στόχοι, που καλούνται αντικειμενική αλήθεια –ground truth. Το μοντέλο εκπαιδεύεται δηλαδή, με δεδομένα που έχουν ετικέτες (labels), γνωρίζουμε την επιθυμητή έξοδο για κάθε είσοδο. Ο στόχος είναι να μάθει το μοντέλο να προβλέπει τις ετικέτες για νέα, άγνωστα δεδομένα.

Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning): Στη μη επιβλεπόμενη μάθηση χρησιμοποιούνται μόνο τα πρότυπα εισόδου, χωρίς να υπάρχει καμία πληροφορία για τους στόχους. Το μοντέλο εκπαιδεύεται με δεδομένα που δεν έχουν ετικέτες με στόχο να ανακαλύψει μοτίβα ή δομές στα δεδομένα, όπως ομάδες (clusters) ή μειώσεις διαστάσεων. Ο συγκεκριμένος τύπος μηχανικής μάθησης προσομοιάζει σε μεγάλο βαθμό την ανθρώπινη νοημοσύνη.

Ενισχυτική Μάθηση (Reinforcement Learning): Στην Ενισχυτική Μάθηση τα πρότυπα έχουν τη μορφή ακολουθίας ενώ οι στόχοι αποτελούν επιβράβευση και πρόστιμο, έτσι το μοντέλο μαθαίνει μέσω αλληλεπίδρασης με ένα περιβάλλον. Το μοντέλο ενεργεί και μαθαίνει από τα αποτελέσματα των ενεργειών του μέσω ανταμοιβών ή ποινών. Στόχος του είναι να μεγιστοποιήσει τη συνολική ανταμοιβή.

Ξεκινώντας την ανάλυση βασικών στοιχείων για τα ΤΝΔ, κρίνεται σκόπιμη η παρουσίαση του απλούστερου ΤΝΔ, του νευρώνα Perceptron. Ο νευρώνας Perceptron, είναι είδος τεχνητού νευρωνικού δικτύου και αποτελεί έναν γραμμικό νευρώνα για την επίλυση γραμμικώς διαχωρίσιμων προβλημάτων.



Εικόνα 17. Απεικόνιση Νευρώνα Perceptron
Πηγή [41]

Η αρχιτεκτονική του νευρώνα Perceptron, βασίζεται στο πρότυπο υπολογιστικού νευρώνα και προτάθηκε από τους MacCulloch και Pitts [9]. Αποτελείται από τα βάρη w – weights και την πόλωση b -bias. Έχοντας έναν νευρώνα $f(W,b)$ με $W \in R^M$ και $b \in R$, όταν του δώσουμε είσοδο $X=[x_1, x_2, \dots, x_m]$, $Q \in R^M$ τότε προκύπτει έξοδος

$$v = \sum_{i=0}^m w_i * x_i + b = W^T * X$$

Παρατηρούμε ότι η έξοδος του νευρώνα ισούται με το εσωτερικό γινόμενο της εισόδου και των βαρών. Το γεγονός αυτό, καθιστά των νευρώνα Perceptron ικανό να επιλύει γραμμικά προβλήματα διαχωρισμού κλάσεων, το οποίο και αποδείχθηκε μαθηματικά από τον Rosenblatt. Πιο συγκεκριμένα, εάν τα πρότυπα εισόδου αντλούνται από δύο κλάσεις, γραμμικώς διαχωριζόμενες μεταξύ τους, τότε ο αλγόριθμος συγκλίνει και ορίζει ένα υπερ-επίπεδο, που διαχωρίζει τις δύο κλάσεις.

Συναρτήσεις Ενεργοποίησης

Στο μοντέλο που αναλύθηκε ανωτέρω, ακόμη και εάν συνδυαστούν παραπάνω νευρώνες, το αποτέλεσμα θα παρέμενε γραμμικό, καθώς θα αποτελούσε μία σύνθεση γραμμικών συναρτήσεων. Ωστόσο, το μοντέλο θα κληθεί να επιλύσει διάφορα και πολύπλοκα προβλήματα, για αυτό το λόγο πρέπει να εισάγουμε την έννοια της μη γραμμικότητας. Χωρίς την έννοια αυτή το μοντέλο, θα αποτελούσε έναν απλό γραμμικό μετασχηματιστή δεδομένων. Οι συναρτήσεις ενεργοποίησης, δίνουν λύση στο πρόβλημα αυτό, καθώς λαμβάνουν ως είσοδο την έξοδο του νευρώνα και μετά την εφαρμογή μη-γραμμικών μετασχηματισμών παράγουν μία νέα έξοδο. Εάν θεωρήσουμε ως $\varphi(u)$ συνάρτηση ενεργοποίησης, η έξοδος του νευρώνα μετασχηματίζεται ως εξής [50]

$$y = \varphi(u) = \varphi(W^T * X)$$

Η διαδικασία αυτή αναπαρίσταται σχηματικά στην Εικόνα 17.

Ανάλογα τη φύση του προβλήματος και τη θέση των νευρώνων, ο αρχιτέκτονας του νευρωνικού δικτύου, καλείται να επιλέξει μεταξύ διαφόρων κατηγοριών συναρτήσεων ενεργοποίησης. Κυριότερες εξ αυτών είναι η σιγμοειδής (sigmoid), η υπερβολική εφαπτομένη (tanh), η ReLU, και η Softmax.



Σιγμοειδής Συνάρτηση

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

Η Σιγμοειδής Συνάρτηση ή Λογιστική Συνάρτηση, παράγει μία έξοδο στο εύρος μεταξύ 0 και 1 [50]. Χάριν σε αυτή την ιδιότητά της, χρησιμοποιείται για την μοντελοποίηση της πιθανότητας και για προβλήματα δυαδικής ταξινόμησης. Μεγάλο της πλεονέκτημα είναι η ευκολία στην ερμηνεία της ως πιθανότητα, ενώ δεν συνίσταται για την εκπαίδευση βαθιών δικτύων, καθώς πάσχει από το πρόβλημα του εξαφανιζόμενου βαθμωτού (vanishing gradient). Δηλαδή, καθώς απομακρύνεται από το 0, τείνει και η παράγωγος της στο 0, συνεπώς εξαφανίζεται η παράγωγος.

Συνάρτηση ReLU

$$f(x) = \max(0, x)$$

Η Ανορθωμένη Γραμμική Μονάδα (Rectified Linear Unit), αποτελεί τη πιο διαδεδομένη συνάρτηση ενεργοποίησης. Πρόκειται για μία μη γραμμική συνάρτηση, η οποία έχει έξοδο 0, εάν η είσοδος είναι αρνητική και ίση με την είσοδο, εάν είναι θετική [50]. Θα μπορούσαμε να πούμε ότι, λειτουργεί ως ένας διακόπτης, ο οποίος απενεργοποιεί τον νευρώνα όταν είναι αρνητική η είσοδος και τον ενεργοποιεί όταν είναι θετική. Η συνάρτηση ReLU, είναι διαφορίσιμη και εν μέρει επιλύει το πρόβλημα του εξαφανιζόμενου βαθμωτού. Παρά την απλότητα της όμως, κατά τη διάρκεια της εκπαίδευσης ελλοχεύει ο κίνδυνος να «πεθάνει» εάν οι νευρώνες δίνουν ως είσοδο σταθερά 0 (Dead ReLU Problem)[10]

Υπερβολική Εφαπτομένη (Tanh – Hyperbolic Tangent)

$$f(x) = \tanh(x) = \frac{e^x + e^{-x}}{e^x - e^{-x}}$$

Η συγκεκριμένη συνάρτηση ενεργοποίησης, χρησιμοποιείται σε κρυμμένα επίπεδα νευρώνων[50]. Παράγει έξοδο σε εύρος μεταξύ -1 και 1, οπότε χρησιμοποιείται σε περιπτώσεις όπου απαιτείται έξοδος με θετικές και αρνητικές τιμές. Μάλιστα, όσο η είσοδος μεγαλώνει, η υπερβολική εφαπτομένη τείνει στο +1, ενώ όταν η είσοδος μικραίνει τείνει στο -1. Κεντράρει τις εξόδους γύρω από το 0, γεγονός που βοηθάει στη σύγκλιση του μοντέλου. Ωστόσο, όπως και η σιγμοειδής, η παράγωγος γίνεται αμελητέα, οπότε εμφανίζεται το πρόβλημα το εξαφανιζόμενου βαθμωτού.

Συνάρτηση Softmax

Η συνάρτηση ενεργοποίησης Softmax, χρησιμοποιείται στους νευρώνες εξόδου, για την ταξινόμηση-κατηγοριοποίηση πολλών κλάσεων [50]. Λύνει ουσιαστικά το πρόβλημα της μοντελοποίησης της κατανομής πιθανότητας για περισσότερες από δύο κλάσεις, το οποίο δεν μπορεί να μοντελοποιηθεί με τη σιγμοειδή συνάρτηση. Μετασχηματίζει τις εξόδους του νευρωνικού δικτύου, σε κατανομή πιθανότητας, παράγοντας πιθανότητες για κάθε κατηγορία όπου το άθροισμά τους είναι 1.



$$f(x_i) = \frac{e^{x_i}}{\sum_j e^{x_j}}$$

Εάν έχουμε πρόβλημα ταξινόμησης με C κλάσεις, στην έξοδο θα έχουμε C νευρώνες με ενεργοποίηση softmax. Ως z_i θα ορίζεται η έξοδος του νευρώνα που είναι υπεύθυνος για την κλάση i .

3.2.2 Πολυεπίπεδα Δίκτυα Perceptron (Multiple Layer Perceptron – MLP)

Τα πολυεπίπεδα δίκτυα Perceptron, αποτελούν νευρωνικά δίκτυα που αποτελούνται από επίπεδα νευρώνων συνδεδεμένα μεταξύ τους. Στη βιβλιογραφία, καλούνται και ως πλήρως συνδεδεμένα επίπεδα (fully connected layers), κάθε νευρώνας είναι συνδεδεμένος με όλους τους νευρώνες κάθε προηγούμενου επιπέδου. Μπορούμε πλέον, να λύνουμε δυσκολότερα προβλήματα, μέσω των πιο πολύπλοκων τοπολογιών των πολυεπίπεδων δικτύων. Τα δίκτυα αυτά, αποτελούνται από ένα επίπεδο εισόδου, το οποίο αποτελείται από έναν νευρώνα ανά τιμή εισόδου, ένα ή περισσότερα κρυφά επίπεδα, τα οποία αυξάνουν την πολυπλοκότητα αναλύοντας περαιτέρω την είσοδο και το επίπεδο εξόδου, που είναι είτε κατανομή πιθανότητας, για προβλήματα ταξινόμησης είτε πρόβλεψη συνεχούς τιμής, για προβλήματα παλινδρόμησης. Στην περίπτωση που η έξοδος είναι κατανομή πιθανότητας, το επίπεδο εξόδου αποτελείται από αριθμό νευρώνων, ίσο με τις κλάσεις του προβλήματος και συνδυασμένων με συναρτήσεις ενεργοποίησης softmax, ενώ στη δεύτερη περίπτωση το επίπεδο εξόδου αποτελείται από έναν νευρώνα, χωρίς συνάρτηση ενεργοποίησης.

Τα πολυεπίπεδα νευρωνικά δίκτυα, με περισσότερα από ένα κρυφά επίπεδα, καλούνται βαθιά νευρωνικά δίκτυα (Deep Neural Networks) και η εκπαίδευση τους βαθιά μάθηση (Deep Learning). Όσο αυξάνεται το βάθος ενός νευρωνικού δικτύου, αυξάνεται ο αριθμός των συνάψεων μεταξύ των συναπτικών βαρών, των νευρώνων, οπότε μπορεί να επιλύσει πολύπλοκα προβλήματα, απαιτώντας ωστόσο, μεγάλη υπολογιστική ισχύ. Παράλληλα, τα δίκτυα αυτά είναι γνωστά ως δίκτυα διάδοσης προς τα εμπρός (feed-forward neural networks). Το σήμα εισόδου διαδίδεται, μέσω των συναπτικών βαρών προς τα εμπρός, έως το επίπεδο εξόδου, χωρίς να υπάρχει λειτουργία ανάδρασης.

Θεώρημα Universal Approximation

Σε κάθε πρόβλημα που αντιμετωπίζει το νευρωνικό δίκτυο, τα συναπτικά βάρη είναι παραμετροποιήσιμα με τον κατάλληλο αλγόριθμο μάθησης. Σύμφωνα με το Θεώρημα Universal Approximation, τα νευρωνικά δίκτυα μπορούν να προσεγγίσουν οποιαδήποτε συνάρτηση $f(x)$, το οποίο σημαίνει ότι σε κάθε συνάρτηση που μοντελοποιεί ένα πρόβλημα, αντιστοιχεί κατάλληλο δίκτυο που μπορεί να την προσεγγίσει. Η ικανότητα αυτή, είναι προϊόν της συνάρτησης ενεργοποίησης, λόγω της μη γραμμικότητας που προσφέρει στους υπολογισμούς[11]. Παράλληλα, το γεγονός αυτό, υποδηλώνει και την δυνατότητα των νευρωνικών δικτύων να εξάγουν χαρακτηριστικά από δεδομένα που δεν έχουν υποστεί επεξεργασία (raw data). Η πρακτική αυτή, καλείται μάθηση χαρακτηριστικών (feature learning) και επιτρέπει στο ΤΝΔ να ανακαλύψει χαρακτηριστικά και σχέσεις μεταξύ των δεδομένων και να κάνει προβλέψεις.



Συνάρτηση Απώλειας

Σημαντική παράμετρος κατά τη σχεδίαση ενός ΤΝΔ, αποτελεί η συνάρτηση απώλειας (loss function). Η συνάρτηση αυτή μετρά την απόκλιση της προβλεπόμενης, από το δίκτυο, τιμής με την πραγματική. Αποτελεί δηλαδή, μία συνάρτηση κόστους, η οποία μέσω μαθηματικών υπολογισμών αποτυπώνει το συνολικό σφάλμα του μοντέλου, κατά τη διάρκεια της εκπαίδευσής του και το χρησιμοποιεί για να βελτιστοποιήσει τις εκπαιδευσιμες παραμέτρους. Υπάρχουν διάφορα είδη συναρτήσεων απώλειας, από τα οποία επιλέγεται το κατάλληλο, ανάλογα τη φύση του προβλήματος.

Μέσο Τετραγωνικό Σφάλμα (Mean Squared Error – MSE)

Αποτελεί συνάρτηση που χρησιμοποιείται κυρίως σε προβλήματα παλινδρόμησης, για να εκτιμηθούν μία ή περισσότερες μεταβλητές που λαμβάνουν συνεχείς τιμές. Υπολογίζει τη μέση τετραγωνική διαφορά μεταξύ της τιμής που προβλέφθηκε και της πραγματικής, δίνοντας έμφαση στις μεγαλύτερες αποκλίσεις.

$$LMSE = \frac{1}{N} \sum_{i=1}^N (y_i - y^{\wedge}_i)^2$$

Κατηγορική Διασταυρωμένη Εντροπία (Categorical Cross Entropy – CCE)

Χρησιμοποιείται συνδυαστικά με τη συνάρτηση ενεργοποίησης Softmax, σε προβλήματα ταξινόμησης πολλών κλάσεων. Μετρείται η απόκλιση μεταξύ των κατανομών πιθανότητας στο επίπεδο της εξόδου.

$$LCCE = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C y_{ij} \log(y_{ij})$$

Για ένα σύνολο δεδομένων με N πρότυπα εκπαίδευσης, y_{ij} αντικειμενική αλήθεια για το i-οστό πρότυπο εκπαίδευσης, στην j-οστή κλάση και y^{\wedge}_{ij} πιθανότητα στην j-οστή κλάση στο i-οστό πρότυπο εκπαίδευσης, έχουμε τον παραπάνω τύπο.

3.2.3 Εκπαίδευση Νευρωνικών Δικτύων

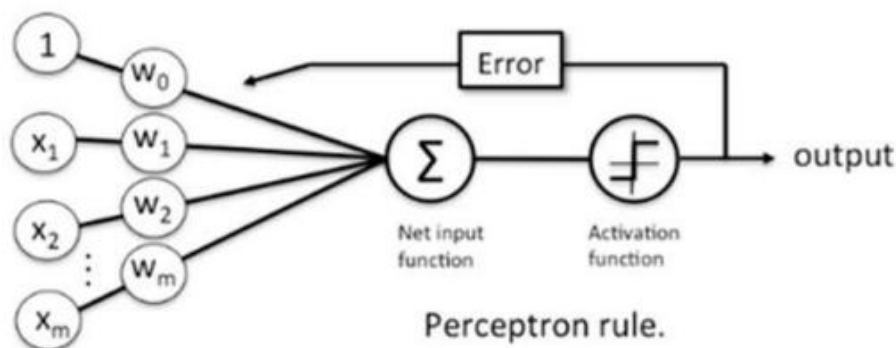
Προηγουμένως, αναλύθηκε η αρχιτεκτονική ενός ΤΝΔ και βασικές λειτουργίες που επιτελούνται στο εσωτερικό του. Στο παρόν υποκεφάλαιο θα γίνει αναφορά στη διαδικασία εκπαίδευσης ενός ΤΝΔ, ώστε αυτό να καταστεί ικανό να προβλέπει με ακρίβεια τις τιμές στην έξοδο του.

Απαραίτητη προϋπόθεση για την εκπαίδευση ενός ΤΝΔ, αποτελεί το σύνολο των δεδομένων, τα οποία θα χρησιμοποιηθούν για την εκπαίδευση του. Τα πρότυπα εισόδου, προωθούνται στο ΤΝΔ, ανά δέσμες (batches), μικρότερα υποσύνολα δηλαδή τα οποία έχουν πλήθος δυνάμεως του 2. Η πληροφορία αυτή, μεταδίδεται από το ένα επίπεδο νευρώνων στο άλλο, ενεργοποιώντας το, πραγματοποιώντας ένα **εμπρόσθιο πέρασμα (forward pass)**, μέχρις ότου να φτάσει στο επίπεδο εξόδου. Κατόπιν, πραγματοποιείται σύγκριση μεταξύ της προβλεπόμενης και πραγματικής τιμής, με τη συνάρτηση απώλειας. Η απώλεια που υπολογίστηκε συνολικά για τη δέσμη εισόδου, διαδίδεται προς τα πίσω (backward pass) με διαδικασία που καλείται οπισθοδιάδοση σφάλματος (backpropagation).

Μέσω της διαδικασίας αυτής, τα βάρη υφίστανται ενημέρωση και αλλαγή, η οποία καθορίζεται από το ρυθμό μάθησης (learning rate). Οι τιμές που λαμβάνει ο ρυθμός μάθησης είναι το δυνατόν μικρότερες 0,1 ή 0,01, διότι δεν είναι επιθυμητές οι μεγάλες και απότομες αλλαγές, αλλά οι μικρές έως ότου να επιτευχθεί το ελάχιστο της συνάρτησης απώλειας [12]. Η διαδικασία που περιγράφηκε επαναλαμβάνεται για όλα τα πρότυπα εισόδου, με την ενημέρωση να πραγματοποιείται για καθένα πρότυπο ξεχωριστά. Ο ολοκληρωμένος κύκλος ενημέρωσης βαρών για το σύνολο των δεδομένων εισόδου, καλείται εποχή (epoch). Ο αριθμός των εποχών, είναι ανάλογος της πολυπλοκότητας των δεδομένων και της αρχιτεκτονικής του δικτύου και κυμαίνεται από δεκάδες σε εκατοντάδες ή και χιλιάδες.

Αλγόριθμος Οπισθοδιάδοσης Σφάλματος

Ο αλγόριθμος οπισθοδιάδοσης σφάλματος, αποτελεί ένα σημαντικό εργαλείο στη φαρέτρα του σχεδιαστή ενός ΤΝΔ. Επιτυγχάνεται χάρις σε αυτόν, καταμερισμός του μετρήσιμου σφάλματος, κατά την έξοδο, σε κάθε ενδιάμεσο νευρώνα, ώστε να ενημερώσουμε τα βάρη του δικτύου για να ελαττωθεί το σφάλμα.



Εικόνα 18. Τρόπος Λειτουργίας Perceptron
Πηγή [41]

Χρησιμοποιούμε την παράγωγο της συνάρτησης απώλειας, ως προς τα υπόλοιπα βάρη του δικτύου [41]. Ωστόσο, οι νευρώνες ανήκουν σε διαφορετικά επίπεδα και για αυτό πρέπει να χρησιμοποιηθεί ο κανόνας της αλυσίδας του διαφορικού λογισμού, ώστε να υπολογισθεί η διανυσματική παράγωγος. Θα ακολουθήσει κώδικας σε ψευδογλώσσα, σχετικά με την οπισθοδιάδοση του σφάλματος, όπου (x, y) είναι πρότυπο εισόδου και \hat{y} η πρόβλεψη του νευρωνικού δικτύου.



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



Algorithm 1 [23] : Forward propagation through a typical deep neural network and the computation of the cost function. The loss $L(\hat{y}, y)$ depends on the output \hat{y} and on the target y .

Require: Network depth, l

Require: $W^i, i \in \{1, \dots, l\}$, the weight matrices of the model.

Require: $b^i, i \in \{1, \dots, l\}$, the bias parameters of the model

Require: x , the input to process and y , the target output

$h^{(0)} = x$

for $k = 1, \dots, l$ **do**

$a^{(k)} = b^{(k)} + W^{(k)} \cdot h^{(k-1)}$

$h^{(k)} = f(a^{(k)})$

end for

$\hat{y} = h^{(l)}$

$J = L(\hat{y}, y)$

Αλγόριθμος Βελτιστοποίησης (Optimizer)

Ο αλγόριθμος οπισθοδιάδοσης ρυθμίζει τον τρόπο και τη συχνότητα με την οποία, θα πραγματοποιείται η ενημέρωση των βαρών, αλλά και την αλλαγή του ρυθμού μάθησης. Μέσω των παραγώγων της συνάρτησης απώλειας, επιδιώκουμε να φτάσουμε στο ολικό ελάχιστο της. Προκειμένου να επιτευχθεί αυτό, πρέπει να επιλεγεί ο κατάλληλος ρυθμός μάθησης. Λόγω της μη κυρτότητας της συνάρτησης κόστους των νευρωνικών δικτύων, υπάρχουν τοπικά ελάχιστα και σαγματικά σημεία, τα οποία έχουν μηδενική είτε αμελητέα παράγωγο. Για να αποφεύγονται τα προβλήματα που δημιουργούνται από αυτά, οι παράμετροι αρχικοποιούνται τυχαία. Η κανονικοποίηση των βαρών αποτρέπει την αποστήθιση του συνόλου εκπαίδευσης, βελτιώνοντας τη γενίκευση σε νέα δεδομένα.

Υπάρχουν δύο διαδεδομένοι αλγόριθμοι βελτιστοποίησης η **Στοχαστική Κάθοδος Κλίσης (Stochastic Gradient Descent - SGD)** και η **Κάθοδος κλίσης (Gradient Descent – GD)**[13]. Ο αλγόριθμος GD, που χρησιμοποιείται για την ελαχιστοποίηση του σφάλματος, αποτελεί μία αποτελεσματική και απλή λύση που βρίσκει εφαρμογή σε ευρεία ποικιλία προβλημάτων. Ωστόσο, όπως αναφέρθηκε και ανωτέρω, η συνάρτηση κόστους σε ένα ΤΝΔ δεν είναι κυρτή (non-convex), με αποτέλεσμα ο αλγόριθμος αυτός να είναι ευαίσθητος σε τοπικά ελάχιστα. Παράλληλα, παρουσιάζει ευαισθησία σε μεγάλα σύνολα δεδομένων, καθώς απαιτεί τον υπολογισμό του σφάλματος για όλα τα μεγέθη σε κάθε επανάληψη. Αντίθετα, ο SGD αντί να χρησιμοποιεί το σύνολο των δεδομένων εκπαίδευσης, ενημερώνει τα βάρη χρησιμοποιώντας ένα μόνο δείγμα δεδομένων σε κάθε επανάληψη. Έτσι, υπάρχει πιο γρήγορη ενημέρωση των βαρών, με αποτέλεσμα να απαιτεί χαμηλό υπολογιστικό κόστος, αντιμετωπίζοντας βέλτιστα μεγάλα σύνολα δεδομένων και την εκμάθηση σε πραγματικό χρόνο. Επιπλέον, παρότι ο SGD δεν έχει τόσο ομαλή σύγκλιση, όπως ο GD, λόγω της τυχαίας επιλογής δειγμάτων, εντούτοις ο θόρυβος που περιέχει η συνάρτηση κόστους του, του επιτρέπει να αποφεύγει τα τοπικά ελάχιστα. Παράλληλα, οι στοχαστικές ενημερώσεις του SGD τον βοηθούν να αποφεύγει την προσαρμογή στο σύνολο δεδομένων (overfitting). Το πλεονέκτημα όμως του SGD, να ενημερώνει τα

βάρη μέσω ενός δείγματος, απαιτεί περισσότερες επαναλήψεις από τον GD, ώστε να μειώσει τον θόρυβο και να συγκλίνει

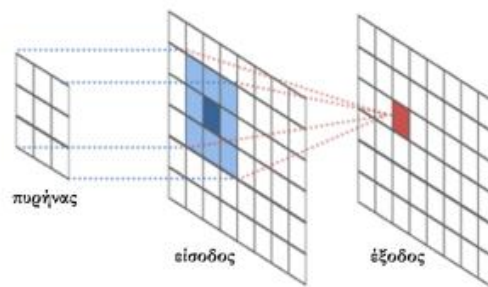
3.2.3 Συνελικτικό Επίπεδο (Convolutional Layer)

Προτού ξεκινήσουμε την ανάλυση, είναι σκόπιμο να παρουσιάσουμε τη σχέση συνέλιξης μεταξύ δύο συναρτήσεων $x(t)$ και $y(t)$, με τη σχέση συνέλιξης να ορίζεται ως

$$s(t) = (x * y)(t) = \int x(a)y(t - a)da$$

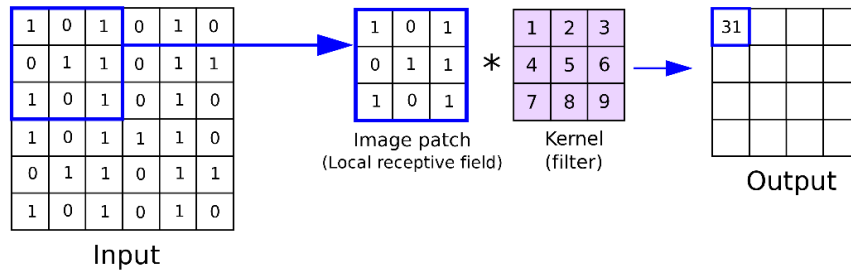
Μπορούμε να προχωρήσουμε σε μία γενίκευση της σχέσης συνέλιξης, στον τομέα της υπολογιστικής όρασης. Το x αντιπροσωπεύει την είσοδο, ενώ το y ονομάζεται πυρήνας ή φίλτρο και η έξοδος s , χάρτης χαρακτηριστικών (feature map). Τα προβλήματα που επιλύονται με τα συνελικτικά δίκτυα, έχουν ως είσοδο μία εικόνα, η οποία αποτελεί ένα διδιάστατο πίνακα, ενώ το φίλτρο έναν μικρότερο πίνακα. Με τη διαδικασία της συνέλιξης, υπολογίζεται το εσωτερικό γινόμενο των δύο πινάκων, μιας μικρής περιοχής (patch), δηλαδή, της εικόνας με το φίλτρο. Το εσωτερικό γινόμενο αυτό, η παραγόμενη έξοδος δηλαδή, είναι μία νέα εικόνα στην οποία κάθε στοιχείο αποτελεί ένα σταθμισμένο (weighted) άθροισμα.

Ένα συνελικτικό νευρωνικό δίκτυο, έχει παρόμοια λειτουργία με ένα απλό νευρωνικό. Αν θεωρήσουμε ως νευρώνα ένα εικονοστοιχείο, τότε καθένα από αυτά συνδέεται με τα αντίστοιχα εικονοστοιχεία του επόμενου επιπέδου, μέσω της συνέλιξης με έναν διδιάστατο πυρήνα. Ο πυρήνας ή φίλτρο, αποτελεί το βάρος της σύνδεσης, το οποίο και μοιράζονται όλα τα εικονοστοιχεία. Η διαδικασία περιγράφεται στην ακόλουθη εικόνα.



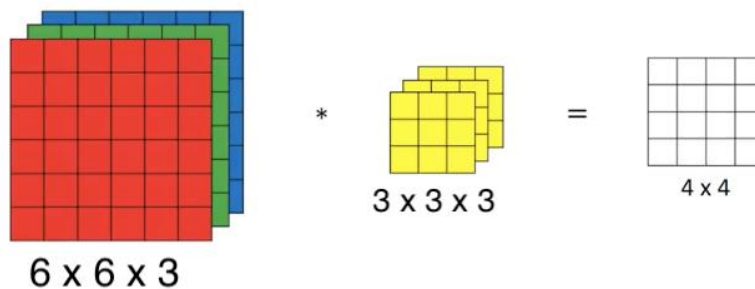
Εικόνα 19. Διδιάστατη συνέλιξη
Πηγή [42]

Το φίλτρο μετακινείται επάνω στην εικόνα εισόδου και υπολογίζει τα εσωτερικά γινόμενα [42]. Η μετακίνηση αυτή του φίλτρου, μπορεί να προσομοιαστεί με ένα παράθυρο. Το μέγεθος του παραθύρου, η διάσταση δηλαδή, του φίλτρου που συνελίσσεται με την εικόνα, καθορίζεται από το receptive field, το δεκτικό πεδίο. Ο αριθμός καναλιών των φίλτρων, είναι ανάλογος του αριθμού των εισόδων τους, έτσι σε μία εικόνα RGB με 3 κανάλια, το φίλτρο θα έχει 3 κανάλια, όπως φαίνεται και στην Εικόνα 20. Παράλληλα, παρατηρείται και η μείωση της διάστασης της εικόνας, που προκαλείται από τη διαδικασία της συνέλιξης.



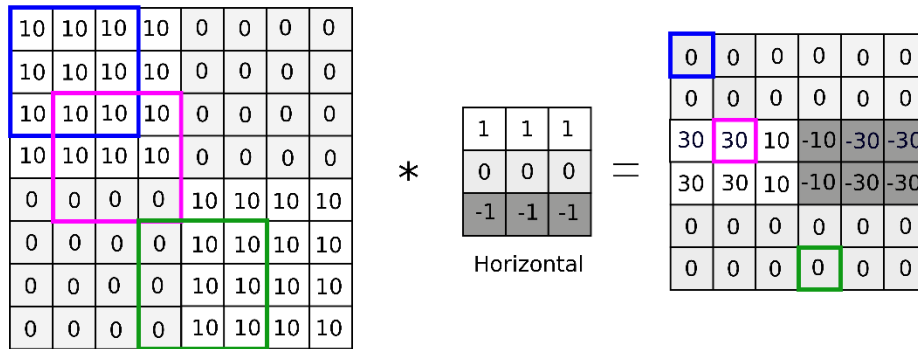
Εικόνα 20. Απεικόνιση συνέλιξης εικόνας, με φίλτρο πίνακα 3x3
Πηγή [42]

Τα σύγχρονα συνελκτικά δίκτυα, αποτελούνται από πολλά συνεχόμενα συνελκτικά επίπεδα, τα οποία και αυτά αποτελούνται από διαφορετικά φίλτρα. Η δυνατότητα των συνελκτικών δικτύων, να έχουν παραμετροποιήσιμα βάρη είναι αυτή που τους δίνει το πλεονέκτημα για τη χρήση τους στη μηχανική μάθηση, καθώς δύναται να τροποποιηθούν με τη χρήση αλγορίθμων. Επίσης, κατά την εκπαίδευσή τους, τα συνελκτικά δίκτυα, αποκτούν εξειδικεύσεις. Στα πρώτα επίπεδα, πραγματοποιείται επεξεργασία της πληροφορίας χαμηλού επιπέδου, όπως ακμές, ενώ τα φίλτρα των τελευταίων επιπέδων επεξεργάζονται πληροφορία υψηλού επιπέδου, δηλαδή το σημασιολογικό περιεχόμενο της εικόνας.



Εικόνα 21. Απεικόνιση συνέλιξης RGB εικόνας με τρισδιάστατο φίλτρο
Πηγή [42]

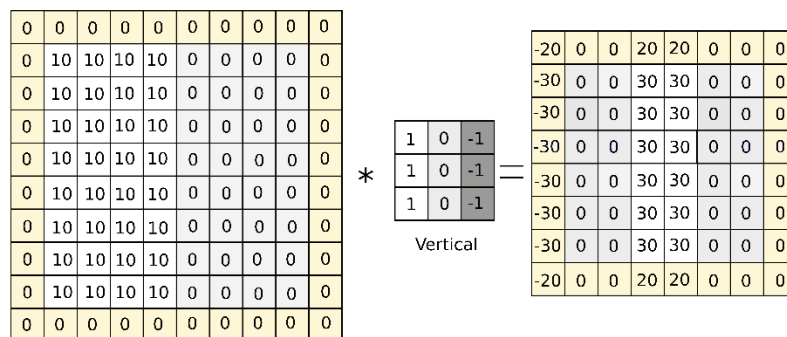
Στην Εικόνα 22, παρουσιάζεται η προαναφερθείσα διαδικασία των πρώτων επιπέδων με έναν οριζόντιο ανιχνευτή άκρων, σε μία ελαφρώς πιο περίπλοκη εικόνα. Η μη μηδενική ζώνη στη κέντρο αποκαλύπτει την οριζόντια ακμή στο μέσον της αρχικής εικόνας, αλλά ανακαλύπτει επίσης και την αλλαγή από φωτεινό σε σκοτεινό (αριστερή πλευρά) και από σκούρο σε φωτεινό (δεξιά πλευρά).



*Εικόνα 22. Διαδικασία ανίχνευσης ακμών
Πηγή [42]*

Μία ακόμα πολύ σημαντική παράμετρος, όσον αφορά τα συνελκτικά δίκτυα αποτελεί το γέμισμα ή padding [42]. Εάν η αρχική εικόνα έχει διαστάσεις $n \times n$ και το φίλτρο είναι μεγέθους $f \times f$, τότε το μέγεθος της παραγόμενης εικόνας είναι $(n-f+1) \times (n-f+1)$. Αυτό συνεπάγεται, τη συρρίκνωση του μεγέθους της εικόνας εισόδου έπειτα από κάθε συνελκτικό επίπεδο. Επιπλέον, λόγω του γεγονότος ότι, το άκρο της αρχικής εικόνας δεν χρησιμοποιείται τόσο συχνά κατά τη λειτουργία συνέλιξης, σε σχέση με τα κεντρικά εικονοστοιχεία (pixel), ορισμένες πολύτιμες πληροφορίες δύναται να χαθούν κατά τη διαδικασία αυτή.

Η λύση σε τούτο το πρόβλημα, είναι η συμπλήρωση του περιγράμματος της αρχικής εικόνας με p επιπλέον στρώμα(τα) μηδενικών προς πάσα κατεύθυνση. Με αυτόν τον τρόπο, οι διαστάσεις των εικόνων εισόδου και εξόδου γίνονται $(n+2p) \times (n+2p)$ και $(n+2p-f+1) \times (n+2p-f+1)$ αντίστοιχα. Όταν $p=0$, δηλαδή δεν υπάρχει padding, ονομάζεται "έγκυρη" συνέλιξη, ενώ όταν $p = (f-1)/2$, με τρόπο που οι διαστάσεις των εικόνων εισόδου και εξόδου να είναι τα ίδια, έχουμε "ίδια" συνέλιξη.



*Εικόνα 23. Παράδειγμα εφαρμογής padding με $p=1$
Πηγή [42]*

Όπως παρατηρείται και στην Εικόνα 23, έχουμε $p=1$, οπότε και το μέγεθος της εισόδου ταυτίζεται με εκείνο της εξόδου. Παράλληλα, παρατηρείται και ότι στο αποτέλεσμα της εξόδου, τα συνοριακά πίξελ λαμβάνονται υπόψιν κατά τους υπολογισμούς της συνέλιξης, όπως και τα πιο κεντρικά.



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



Μία ακόμη υπερ-παράμετρος συνέλιξης είναι το stride ή βήμα και συμβολίζεται με s . Το stride, καθορίζει το βήμα που κάνει το φίλτρο, καθώς αυτό μετακινείται επάνω στην εικόνα για να πραγματοποιήσει τον επόμενο υπολογισμό. Έτσι, όταν το βήμα ισούται με 1, το φίλτρο μετακινείται μία θέση, όταν ισούται με 2, δύο κ.ο.κ. Αυξάνοντας έτσι το βήμα, μειώνεται η διάσταση της εικόνας με αποτέλεσμα να μειώνεται και το πλήθος των υπολογισμών. Ωστόσο, η διαδικασία αυτή δύναται να δημιουργήσει προβλήματα, διότι θα αγνοούνται ορισμένα εικονοστοιχεία κατά τη μετακίνηση του φίλτρου.

Πολλαπλά φίλτρα μπορούν να εφαρμοστούν σε κάθε εικόνα. Δοσμένης μίας εικόνας διαστάσεων $n \times n \times n_c$ και n'_c φίλτρα διαστάσεων $f \times f \times n_c$ με βήμα s και γέμισμα p , οι διαστάσεις της παραγόμενης εικόνας υπολογίζονται ως ακολούθως :

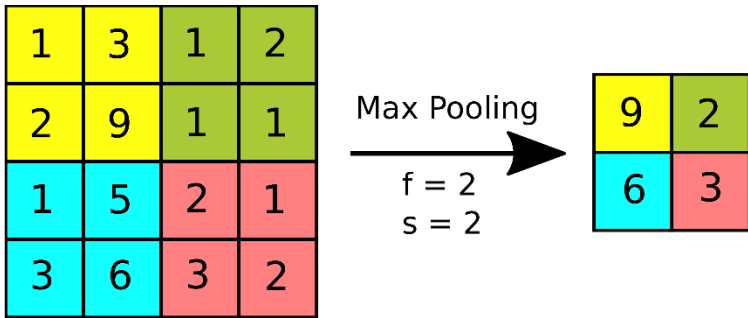
$$\left(\frac{n + 2p - f}{s} + 1\right) \times \left(\frac{n + 2p - f}{s} + 1\right) \times n'_c$$



Εικόνα 24. Απεικόνιση feature map σε εικόνα

3.2.4 Επίπεδα pooling (Τμήμα Υποδειγματοληψίας)

Τα επίπεδα συσσώρευσης (Pooling Layers), αποτελούν σημαντικό κομμάτι της αρχιτεκτονικής των CNN. Ο ρόλος τους έγκειται στην συγχώνευση της πληροφορίας, η οποία εμπεριέχεται στον χάρτη χαρακτηριστικών, με στόχο να μειώσουν τη διάσταση του. Δύνανται, ωστόσο, να παραλειφθεί, καθώς μπορεί να υλοποιηθεί κατά τη διάρκεια της συνέλιξης. Η χρήση τους, είναι παρόμοια με εκείνη της συνέλιξης, ένας μικρός πίνακας διαστάσεων $f \times f$ (παράθυρο) μετακινείται επάνω στην εικόνα με βήμα s και με γέμισμα $p=0$. Υπάρχουν διάφορα είδη τεχνικών υποδειγματοληψίας με πιο συνηθισμένα τη συσσώρευση μεγίστου (Max Pooling) και τη συσσώρευση μέσου όρου (Average Pooling). Στη συσσώρευση Max Pooling, η συσσώρευση εφαρμόζεται σε μία περιοχή του χάρτη και λαμβάνουμε μεγαλύτερο στοιχείο αυτής, ενώ η Average Pooling, το μέσο όρο των στοιχείων.



Εικόνα 25. Παράδειγμα Εφαρμογής Max Pooling
Πηγή [42]



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



Στην Εικόνα 25, παρουσιάζεται παράδειγμα εφαρμογής του Max Pooling για $f = 2$ και $s = 2$, σε πίνακα 4×4 ($n=4$). Η συσσώρευση, εν αντιθέσει με τη συνέλιξη, εφαρμόζεται σε κάθε κανάλι ξεχωριστά, με αποτέλεσμα να διατηρείται ο αριθμός των καναλιών. Οπότε, οι διαστάσεις της εξόδου, δεδομένης εισόδου $n_H \times n_W \times n_C$, είναι

$$\left(\frac{n_H - f}{s} + 1\right) \times \left(\frac{n_W - f}{s} + 1\right) \times n_C$$

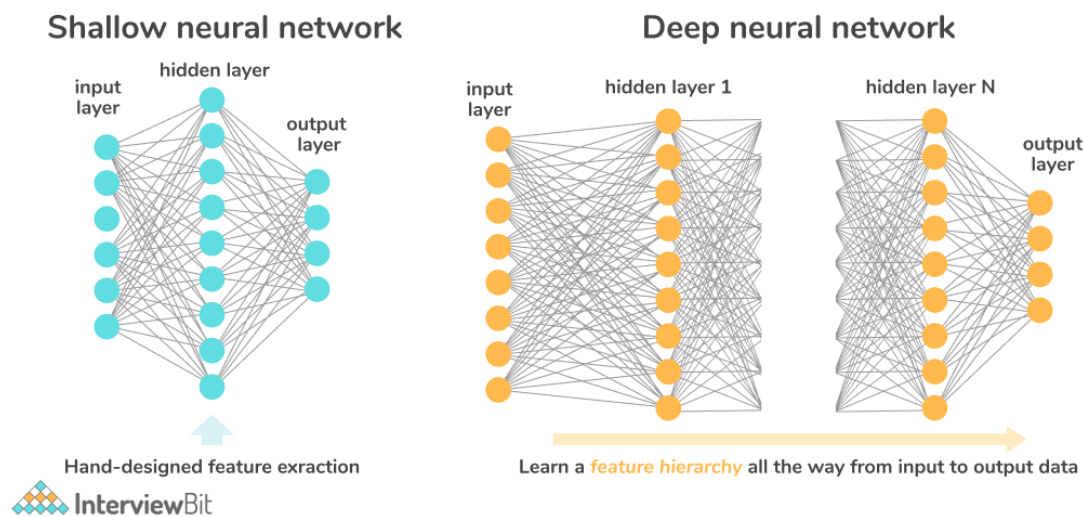
Πειραματικές μελέτες [6], έχουν αποδείξει ότι η τεχνική του max pooling, αποδίδει καλύτερα από εκείνη του average. Το αποτέλεσμα αυτό, είναι αναμενόμενο διότι, δεν είναι επιθυμητή η μείωση των μεγάλων τιμών, λόγω του μέσου όρου. Υπάρχουν περιπτώσεις δε, που σε περιοχές μεγέθους 4 εικονοστοιχείων και άνω, μπορεί να υπάρξει επικάλυψη των περιοχών δειγματοληψίας, οπότε και να βελτιώνεται η απόδοση του δικτύου[7].

3.2.5 Οπισθοδιάδοση Σφάλματος (Backpropagation)

Η τεχνική της οπισθοδιάδοσης σφάλματος, χρησιμοποιείται κατά κόρον στην εκπαίδευση των νευρωνικών δικτύων. Η τεχνική αυτή, βασίζεται στην προώθηση των σφαλμάτων που προέκυψαν κατά την κατηγοριοποίηση, από το τελευταίο επίπεδο στο πρώτο. Η αντίστοιχη τεχνική αυτή, χρησιμοποιείται στα CNN, όπου το σφάλμα μεταφέρεται από τα πλήρως συνδεδεμένα συνελικτικά επίπεδα στα αραιά έως το πρώτο επίπεδο. Ιδιαίτερη προσοχή στη διαδικασία αυτή, απαιτείται κατά τη μετάδοση του σφάλματος στο νευρώνα, ώστε να λάβει το σφάλμα μόνο από τους νευρώνες που είναι συνδεδεμένος. Παράλληλα, σημαντικό κομμάτι αποτελεί και το upscaling. Η συγκεκριμένη διαδικασία, αποσκοπεί στην επαναφορά στις αρχικές διαστάσεις, του αποτελέσματος της συνέλιξης όταν εκτελέστηκε η υποδειγματοληψία, ώστε να γίνει σωστή αντιστοίχιση του σφάλματος. Το αποτέλεσμα αυτό, επιτυγχάνεται με τις διαδικασίες του max και average pooling. Στο max-pooling, το σφάλμα κάθε εικονοστοιχείου, δίνεται σε εκείνο το εικονοστοιχείο με το μεγαλύτερο σφάλμα κατά την φάση της προώθησης, ενώ τα υπόλοιπα έχουν 0. Αντίθετα, στο average-pooling, το σφάλμα κάθε εικονοστοιχείου μοιράζεται σε γειτονικά αυτού. Συμπεραίνουμε λοιπόν, συγκρίνοντας τις δύο τεχνικές, ότι το max-pooling, αποτελεί μία πιο πολύπλοκη διαδικασία, καθώς απαιτείται η διατήρηση στη μνήμη της θέσης του εικονοστοιχείου με το μεγαλύτερο σφάλμα.

3.3 Αρχιτεκτονική των CNN δικτύων

Παραπάνω παρουσιάστηκαν τα βασικά συστατικά των συνελκτικών νευρωνικών δικτύων, καθώς και οι υπερ-παράμετροι τους. Τα στοιχεία αυτά συνδυάζονται προκειμένου να κατασκευαστεί το νευρωνικό δίκτυο, το οποίο μπορεί να διαθέτει ένα ή περισσότερα κρυφά επίπεδα. Έτσι, μπορούμε να κατασκευάσουμε Νευρωνικά Δίκτυα πολλαπλών επιπέδων, τα οποία μπορούν να αντιμετωπίσουν καλύτερα πιο σύνθετα προβλήματα. Έτσι, τα συνελκτικά δίκτυα διακρίνονται σε δύο μεγάλες κατηγορίες, τα Αβαθή Νευρωνικά Δίκτυα (Shallow Neural Networks) και τα Βαθιά Νευρωνικά Δίκτυα (Deep Neural Networks) [43]. Όσον αφορά τα Βαθιά NN, ο αριθμός των συνελίξεων που διαθέτουν μπορεί να είναι αυξημένος, με αποτέλεσμα τα διανύσματα εξόδου, να έχουν σε μεγάλο βαθμό ελαττωμένες διαστάσεις, μη χάνοντας όμως την πληροφορία που περιλαμβάνουν. Συνοψίζοντας, όσο περισσότερα επίπεδα έχει ένα ΣΝΔ τόσο πιο γενικά και αφηρημένα γίνονται αυτά τα χαρακτηριστικά. Στα πρώτα επίπεδα, τα χαρακτηριστικά είναι πιο συγκεκριμένα όπως άκρες και γωνίες. Σε βαθύτερα επίπεδα, τα χαρακτηριστικά είναι πιο σύνθετα και γενικά, π.χ., μπορεί να αντιπροσωπεύουν πιο περίπλοκα σχήματα ή μοτίβα που αντιστοιχούν σε ολόκληρα αντικείμενα.



Εικόνα 26. Σύγκριση Shallow με Deep CNN
Πηγή [43]

Η υπολογιστική όραση αποτελεί ένα σημαντικό εργαλείο με τεράστιες δυνατότητες. Η αύξηση της διαθέσιμης υπολογιστικής ισχύος, οδήγησε σε αλματώδη εξέλιξη της υπολογιστικής όρασης, με αποτέλεσμα να ξεπεράσει τις ανθρώπινες δυνατότητες στην επίλυση ορισμένων προβλημάτων. Ορισμένα από αυτά είναι η ταξινόμηση εικόνας, η σημασιολογική κατάτμηση, η ανίχνευση αντικειμένων, αλλά και πιο σύνθετα όπως η εκτίμηση πόζας, η εκτίμηση βάθους εικόνας και η συνεχής παρακολούθηση αντικειμένων.

Ανίχνευση Αντικειμένων (Object Detection)

Η ανίχνευση αντικειμένων, αποτελεί έναν συνδυασμό της ταξινόμησης με ταυτόχρονο εντοπισμό των αντικειμένων που αναγνωρίζονται. Το μοντέλο αυτό, έχει ως έξοδο όχι μόνο την παρουσία ενός αντικειμένου σε μια εικόνα, αλλά και την ακριβή τοποθεσία



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



του, περικλείοντας το εντός ενός «πλαισίου οριοθέτησης». Παράδειγμα χρήσης του συγκεκριμένου μοντέλου είναι η αναγνώριση της κυκλοφοριακής συμφόρησης σε δρόμους των ΗΠΑ[15]. Αναπτύσσοντας μοντέλα ανίχνευσης αυτοκινήτων, οι ερευνητές μπορούν να μοντελοποιήσουν την κίνηση στους δρόμους, από εικόνα που προέρχεται από UAV. Ομοίως, στον τομέα της βιομηχανία, η αναγνώριση αντικειμένων μπορεί να βοηθήσει στη βελτιστοποίηση της επιλογής ελαττωματικών προϊόντων.

Ταξινόμηση Εικόνας (Image Classification)

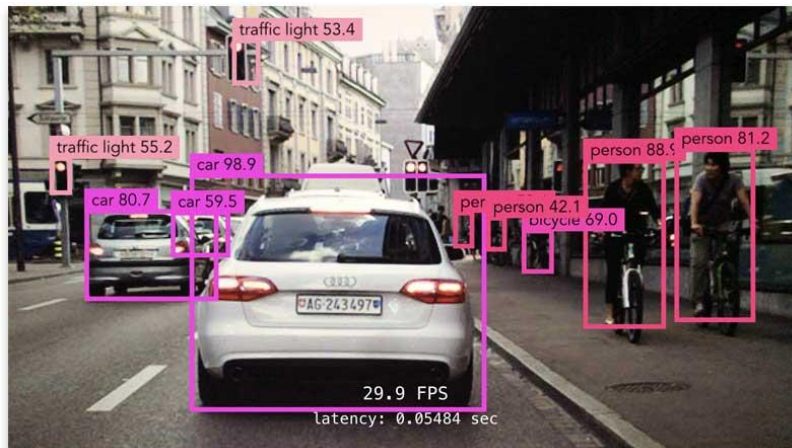
Στα συγκεκριμένα προβλήματα, το μοντέλο δύναται να ταξινομεί εικόνες, ανάλογα της κλάσης του αντικειμένου που απεικονίζεται. Τα σύνολα δεδομένων για τα προβλήματα αυτά, είναι αντικειμενοκεντρικά, όπως το σύνολο δεδομένων MNIST[16]. Το MNSIT, αποτελείται από ένα σύνολο χειρόγραφων ψηφίων, όπου η είσοδος είναι η ασπρόμαυρη εικόνα ενός ψηφίου, την οποία πρέπει να αναγνωρίσει το μοντέλο. Οι εφαρμογές ταξινόμησης, βρίσκουν εφαρμογή σε πληθώρα επιστημονικών πεδίων, με πιο χαρακτηριστικό την ιατρική, όπου μοντέλα υπολογιστικής όρασης χρησιμοποιούνται για να προβλέψουν την κατάσταση του ασθενούς μέσα από την μαγνητική τομογραφία MRI.

Σημασιολογική Κατάτμηση (Semantic Segmentation)

Στα προβλήματα σημασιολογικής κατάτμησης, το μοντέλο ταξινομεί καθένα από τα εικονοστοιχεία (πίξελ) της εικόνας εισόδου, στην κλάση που αντιστοιχεί. Το μοντέλο υπολογιστικής όρασης, κατανοεί το οπτικό ερέθισμα, αναγνωρίζοντας σε επίπεδο πίξελ το σχήμα, όρια και το ίδιο το αντικείμενο. Βρίσκει εφαρμογή σε προβλήματα όπως η αυτόνομη οδήγηση και η γεωργία, όπου απαιτείται κατάτμηση του σημασιολογικού χάρτη σε αντικείμενα απαραίτητα για την επίτευξη του στόχου. Στην αυτόνομη οδήγηση, ο σημασιολογικός χάρτης, χωρίζεται σε αντικείμενα όπως, δρόμος, οχήματα και άνθρωποι[17].

3.4 Ανίχνευση Αντικειμένων (Object Detection)

Η παρούσα εργασία, επιχειρεί να αντιμετωπίσει το πρόβλημα της ανίχνευσης του θερμικού αποτυπώματος τεθωρακισμένων αρμάτων μάχης. Η ανίχνευση του θερμικού αποτυπώματος, εντάσσεται στην ευρύτερη κατηγορία της ανίχνευσης αντικειμένων. Όπως προαναφέρθηκε, στόχος αποτελεί η εύρεση του αντικειμένου καθώς και η ακριβής τοποθεσία του. Στην πλειοψηφία των μοντέλων, η έξοδος συνοδεύεται και από το σκορ/βαθμολογία εμπιστοσύνης – confidence score, το οποίο αντιστοιχεί στη βεβαιότητα με την οποία πραγματοποιεί την πρόβλεψη το μοντέλο.



Εικόνα 27. Απεικόνιση Ανίχνευσης Αντικειμένων σε Εικόνα

Πλαίσια Οριοθέτησης– Bounding Boxes

Τα bounding boxes, αποτελούν τον τρόπο με τον οποίο ορίζεται η τοποθεσία των αντικειμένων, μέσα στη φωτογραφία εισόδου. Τα πλαίσια αυτά, μπορούν να είναι ορθογώνια (rectangle), αλλά και πολύγωνα (polygon). Όσον αφορά τα ορθογώνια, αυτά αναπαρίστανται με 4 διαφορετικά νούμερα (x_1 , x_2 , y_1 , y_2), τις συντεταγμένες των άκρων του κουτιού, είτε με (x , y , h , w) με τις συντεταγμένες του κεντρικού σημείου, μαζί με το ύψος και μήκος αυτού. Οι συντεταγμένες αυτές, υπόκεινται σε κανονικοποίηση τόσο στην αντικειμενική αλήθεια όσο και στην πρόβλεψη, με αποτέλεσμα να δίνονται αριθμοί από το 0 έως 1. Συνεπώς, προκειμένου να ευρεθούν τα ακριβή εικονοστοιχεία πρέπει να γίνει πολλαπλασιασμός με τις διαστάσεις της εικόνας. Επίσης, οι συντεταγμένες ή πιο σωστά μεταβλητές του bounding box, είναι συνεχείς μεταβλητές, οπότε και ο υπολογισμός τους αποτελεί ένα πρόβλημα παλινδρόμησης. Έτσι, τα μοντέλα ανίχνευσης αντικειμένων ταυτοχρόνως επιλύουν τρία διαφορετικά προβλήματα, παλινδρόμησης, ταξινόμησης και εκτίμησης πιθανότητας, τα οποία αντιστοιχούν στην εκτίμηση του bounding box, πρόβλεψη κλάσης αντικειμένου και εύρεση σκορ αυτοπεποίθησης.

3.4.1 Μέθοδοι Ανίχνευσης

Οι αλγόριθμοι ανίχνευσης αντικειμένων βασίζονται επάνω στην αρχιτεκτονική των CNN. Τα CNN είναι η κύρια αρχιτεκτονική που χρησιμοποιείται για την ανάλυση εικόνων και την εξαγωγή χαρακτηριστικών από αυτές. Οι CNN αλγόριθμοι μπορούν να εκπαιδευτούν για την αναγνώριση αντικειμένων σε διάφορα πλαίσια και να εφαρμοστούν για προβλήματα όπως η ταξινόμηση εικόνων, η ανίχνευση αντικειμένων



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



και η αναγνώριση προσώπων. Υπάρχουν δύο κύριες κατηγορίες αλγορίθμων ανίχνευσης, οι **ανιχνευτές δύο σταδίων (two-stage detectors)** και οι **ανιχνευτές ενός σταδίου (one-stage detectors)**.

Οι ανιχνευτές ενός σταδίου απαιτούν μία και μόνο πρόβλεψη, προκειμένου να προβούν σε εκτίμηση των bounding boxes και την κλάσης, ταυτοχρόνως. Δηλαδή, το δίκτυο επεξεργάζεται την εικόνα και παράγει απευθείας προβλέψεις για τις κατηγορίες και τις θέσεις των αντικειμένων χωρίς να χρειάζεται να εκτελέσει πρώτα ένα ξεχωριστό στάδιο εντοπισμού περιοχών ενδιαφέροντος (region proposal)

Χαρακτηριστικό παράδειγμα, ανιχνευτών ενός σταδίου, αποτελεί το μοντέλο YOLO (You Only Look Once). Όπως προδίδει και το όνομα του, το μοντέλο εκτελεί ένα και μοναδικό πέρασμα προς τα εμπρός, ώστε να παράξει αποτέλεσμα.

Το σύστημά προχωρά σε διαίρεση της εισερχόμενης εικόνας σε ένα πλέγμα $S \times S$. Εάν το κέντρο ενός αντικειμένου βρίσκεται εντός ενός κελιού του πλέγματος, τότε αυτό το κελί θα είναι υπεύθυνο για την ανίχνευση του αντικειμένου. Κάθε κελί πλέγματος μπορεί να προβλέψει B bounding boxes και confidence scores για αυτά τα boxes. Αυτές οι βαθμολογίες εμπιστοσύνης αντικατοπτρίζουν το πόσο σίγουρο είναι το μοντέλο ότι το bounding box περιέχει ένα αντικείμενο, καθώς επίσης και πόσο ακριβής θεωρεί ότι είναι το bounding box που προβλέπει. Συνήθως, ορίζουμε το confidence score ως **$\text{Pr}(\text{Object}) * \text{IOU}_{\text{truth_pred}}$** .

Εάν δεν υπάρχει αντικείμενο μέσα σε αυτό το κελί, τα confidence scores πρέπει να είναι μηδέν. Σε διαφορετική περίπτωση, είναι επιθυμητό η βαθμολογία εμπιστοσύνης να ισούται με την ένωση κατά τομή (Intersection over Union, IOU) προβλεπόμενης bounding box και πραγματικής bounding box (ground truth).

Καθένα από τα bounding boxes, αποτελείται από 5 προβλέψεις: x , y , w , h και μία βαθμολογία εμπιστοσύνης (confidence score). Οι συντεταγμένες (x, y) δηλώνουν το κέντρο του bounding box σε σχέση με τα όρια του κελιού στο πλέγμα. Επίσης, το πλάτος (w) και το ύψος (h) υπολογίζονται σε σχέση με ολόκληρη την εικόνα. Τέλος, η πρόβλεψη της βαθμολογίας εμπιστοσύνης, αντιπροσωπεύει το IOU μεταξύ προβλεπόμενου bounding box και πραγματικού.

Επιπλέον, κάθε κελί του πλέγματος προβλέπει C , υπό συνθήκη πιθανότητες κλάσεων, $\text{Pr}(\text{Class}|\text{Object})$. Αυτές οι πιθανότητες υποδηλώνουν την προϋπόθεση ότι στο κελί του πλέγματος βρίσκεται ένα αντικείμενο. Σε κάθε κελί του πλέγματος, προβλέπεται ένα μόνο σύνολο πιθανότητας κλάσεων, ανεξάρτητα από τον αριθμό των bounding boxes B , στην συγκεκριμένη περίπτωση $B=2[18]$.

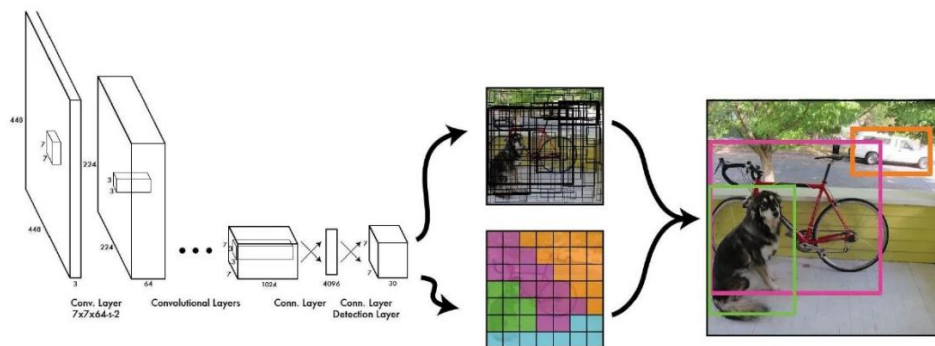
Κατά τη διάρκεια της δοκιμής, πολλαπλασιάζουμε τις υπό συνθήκη πιθανότητες κλάσεων με τις ατομικές προβλέψεις εμπιστοσύνης των bounding boxes, δηλαδή:

$$\text{Pr}(\text{Class}|\text{Object}) * \text{Pr}(\text{Object}) * \text{IOU}_{\text{truth-pred}} = \text{Pr}(\text{Class}) * \text{IOU}_{\text{truth-pred}}$$

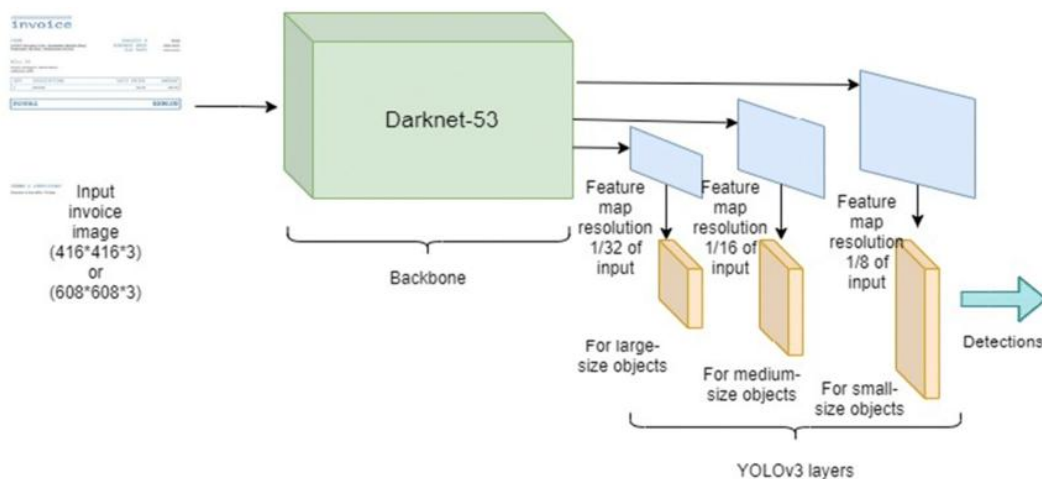
Αυτό μας δίνει βαθμολογίες εμπιστοσύνης ανά κλάση για κάθε box. Οι βαθμολογίες αυτές, κωδικοποιούν την πιθανότητα εμφάνισης της συγκεκριμένης κλάσης στο bounding box, αλλά και το πόσο καλά το bounding box, που προβλέφθηκε, ταιριάζει με το αντικείμενο. Μέσω του αλγορίθμου non max suppression, γίνεται μία διαλογή

των σημαντικότερων προβλέψεων, τις οποίες υλοποιεί μοντέλο, το οποίο στην πραγματικότητα αποτελεί ένα CNN, έχοντας εκατοντάδες εισόδους και μία τελική έξοδο, όπως φαίνεται και στην Εικόνα 28.

Η πλειονηφία των μεθόδων ανίχνευσης ενός σταδίου, βασίζεται στο YOLO, το οποίο είτε ενισχύουν με τη χρήση ισχυρού δικτύου, ως «ραχοκοκκαλιά» είτε με άλλες μεθόδους, ώστε να βελτιωθεί η απόδοσή του. Ένα τέτοιο παράδειγμα, αποτελεί η εργασία των Joseph Redmon και Ali Farhadi, σχετικά με την ενίσχυση του YOLOv3. Στη συγκεκριμένη εργασία, προτείνεται η αρχιτεκτονική Darknet-53, ως backbone (Εικόνα 22), η οποία διαθέτει 53 συνελκτικά επίπεδα για την εξαγωγή χαρακτηριστικών από τις εικόνες, ενώ ταυτόχρονα πραγματοποιούνται προβλέψεις σε διαφορετικές κλίμακες S.



Εικόνα 28. Διαδικασία Ανίχνευσης Αντικειμένων με το μοντέλο YOLO
Πηγή [18]



Εικόνα 29. Αρχιτεκτονική Darknet ως backbone σε YOLOv3.
Πηγή:[20]

Οι ανιχνευτές ενός σταδίου, είναι δημοφιλείς λόγω της ταχύτητας με την οποία πραγματοποιούν προβλέψεις, καθιστώντας τους ιδανικούς για εφαρμογές ανίχνευσης σε πραγματικό χρόνο. Ωστόσο, προκειμένου να αντιμετωπιστεί η έλλειψη του επιπέδου της εκτίμησης περιοχών, χρησιμοποιούνται ορισμένες μέθοδοι, ώστε να αυξηθεί η αξιοπιστία τους.



Μία συνήθη πρακτική αποτελεί η χρήση των **anchor boxes** [23]. Τα anchor boxes, γνωστά και ως anchor priors ή default boxes, είναι προκαθορισμένα bounding boxes με συγκεκριμένα μεγέθη, αναλογίες και θέσεις, τα οποία χρησιμοποιούνται ως πρότυπα αναφοράς κατά την ανίχνευση αντικειμένων. Αυτά τα anchor boxes τοποθετούνται σε διάφορες θέσεις σε μια εικόνα, συχνά με μοτίβο που μοιάζει με πλέγμα, ώστε να καλύπτουν αντικείμενα διαφόρων μεγεθών και σχημάτων. Κατά τη διάρκεια της εκπαίδευσης και της πρόβλεψης, τα anchor boxes χρησιμοποιούνται για την πρόβλεψη των τοποθεσιών και των σχημάτων των αντικειμένων σε σχέση με αυτά τα κουτιά αναφοράς. Σημαντική παράλληλα, είναι και η χρήση του αλγορίθμου **non max suppression**, για το φιλτράρισμα των εικόνων που καταλήγουν στον ανιχνευτή. Ο συγκεκριμένος αλγόριθμος, αρχικά απορρίπτει τα bounding boxes με confidence score χαμηλότερο από ένα κατώφλι, που καλείται confidence threshold. Έπειτα, προχωρά σε ταξινόμηση των bounding boxes που προβλέφθηκαν με βάση το confidence score τους, κατά φθίνουσα σειρά. Έτσι υπολογίζει τη μετρική IoU, μεταξύ του box με το μεγαλύτερο score και των υπολοίπων. Αν το αποτέλεσμα είναι μεγαλύτερο από το κατώφλι nms threshold, τότε διαγράφεται το box με τη μικρότερη αυτοπεποίθηση. Η διαδικασία αυτή, επαναλαμβάνεται για το σύνολο των προβλέψεων, έως ότου διατηρηθούν οι καλύτερες εκτιμήσεις για τα αντικείμενα στην κάθε εικόνα[24].

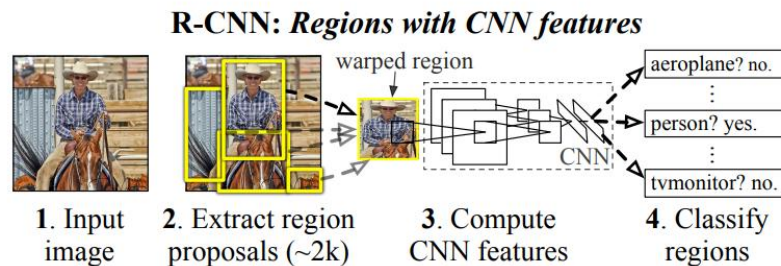
Ανιχνευτές Δύο Σταδίων (Two-Stage Detectors)

Οι ανιχνευτές δύο σταδίων, αποτελούν μία ακόμη κατηγορία αλγορίθμων ανίχνευσης αντικειμένων, οι οποίοι διαχωρίζουν τη διαδικασία ανίχνευσης στο στάδιο της πρότασης περιοχής (region proposal) και της ταξινόμησης βελτίωσης. Το πρώτο στάδιο, είναι υπεύθυνο για τη δημιουργία περιοχών ενδιαφέροντος (Regions of Interest – RoIs) για την κάθε εικόνα, οι οποίες είναι υποψήφιες να περιέχουν αντικείμενα. Τυπικό παράδειγμα αλγορίθμου που χρησιμοποιείται στο εν λόγω στάδιο, είναι ο Region Proposal Network (RPN), ο οποίος εξάγει περιοχές από την εικόνα που πιθανότατα περιέχουν αντικείμενα. Αυτές οι περιοχές έχουν διάφορα μεγέθη και αναλογίες (aspect ratios).

Στο δεύτερο στάδιο, της ταξινόμησης και βελτίωσης, οι περιοχές που προτάθηκαν από το πρώτο στάδιο περνούν από ένα νευρωνικό δίκτυο για να ταξινομηθούν (classification) και να βελτιωθούν (refinement). Με αυτό τον τρόπο, ο ανιχνευτής αποφασίζει αν οι περιοχές περιέχουν αντικείμενα, ποια αντικείμενα είναι αυτά (αντικείμενο ή φόντο), και στη συνέχεια βελτιώνει τα όρια (bounding boxes) για να είναι πιο ακριβή.

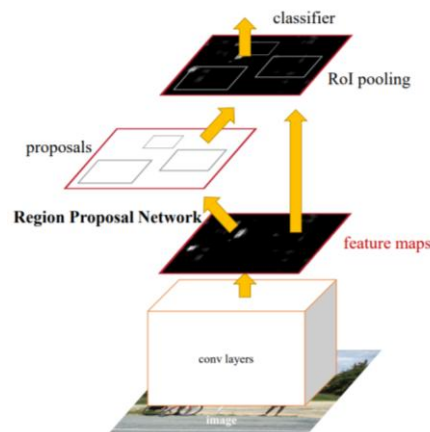
Έτσι, οι ανιχνευτές δύο σταδίων αρχικά προβλέπουν εκατοντάδες ή χιλιάδες πιθανές τοποθεσίες κουτιών (bounding boxes), και στη συνέχεια κάθε περιοχή δίνεται ως είσοδος σε ένα δίκτυο ταξινόμησης, για να καθοριστεί αν περιέχει αντικείμενο και με ποια πιθανότητα. Δημοφιλείς αλγόριθμοι βασιζόμενοι στην αρχιτεκτονική δύο σταδίων, είναι τα δίκτυα R-CNN και Faster R-CNN. Στο μεν πρώτο [21], χρησιμοποιείται ένας αλγόριθμος για την πρόταση πιθανών περιοχών, που αποτελούν τα bounding boxes. Αφού υπολογιστούν περίπου 2000 τέτοιες περιοχές, αυτές δίνονται μεμονωμένα σε ένα CNN για να ταξινομηθούν και να καθοριστεί αν περιέχουν κάποιο αντικείμενο ενδιαφέροντος. Η αρχιτεκτονική, όπως αντιλαμβάνεται κάποιος, δεν είναι

αποδοτική, καθώς απαιτεί την εκτίμηση μεγάλου αριθμού περιοχών και το CNN πρέπει να κάνει ισάριθμες προβλέψεις, καθιστώντας τη διαδικασία αργή και ακατάλληλη για εφαρμογές σε πραγματικό χρόνο.



Εικόνα 30. Αρχιτεκτονική R-CNN
Πηγή [21]

Στο δεύτερο δε [22], αντί για αλγόριθμο πρότασης περιοχής, χρησιμοποιείται το δίκτυο RPN, το οποίο βασίζεται σε ένα ισχυρό δίκτυο "ραχοκοκαλιάς" (backbone) και υπολογίζει τον χάρτη χαρακτηριστικών της εικόνας, που προωθείται στα επόμενα στάδια της μεθόδου. Αυτά τα χαρακτηριστικά χρησιμοποιούνται τόσο για την εκτίμηση των bounding boxes όσο και για την ταξινόμηση των κατηγοριών, επιταχύνοντας σημαντικά τον χρόνο πρόβλεψης, αφού το RPN συμμετέχει και στον υπολογισμό των χαρακτηριστικών της εικόνας εισόδου.



Εικόνα 31. Αρχιτεκτονική Faster R-CNN με RPN
Πηγή [22]

3.4.2 Μετρήσεις Αξιολόγησης

Για την αξιολόγηση της απόδοσης του μοντέλου που χρησιμοποιήθηκε για την ανίχνευση, δηλαδή το πόσο καλά εντοπίζει και ταξινομεί αντικείμενα σε μία εικόνα, χρησιμοποιούνται οι μετρικές ανίχνευσης. Οι μετρικές ανίχνευσης, ποσοτικοποιούν την ακρίβεια και αποδοτικότητα του μοντέλου και είναι οι ακόλουθες.

Precision (Ακρίβεια) [51]: Δηλώνει το ποσοστό των θετικών προβλέψεων που είναι σωστές, δηλαδή το ποσοστό των ορθώς ανιχνευμένων αντικειμένων σε σχέση με όλα τα αντικείμενα που ανιχνεύθηκαν.



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



$$Precision = \frac{TP}{TP + FP}$$

Mean Average Precision (mAP) [51]: Αποτελεί την πιο σημαντική μετρική αναφοράς, για την αξιολόγηση ενός μοντέλου. Υπολογίζεται ως ο μέσος όρος της ακρίβειας (average precision - AP) σε πολλαπλά thresholds (όρια) του IoU.

$$mAP = \frac{1}{C} \sum_{i=1}^C AP_i$$

Αν κατασκευάσουμε την καμπύλη ακρίβειας ανάκλησης, χρησιμοποιώντας τις αντίστοιχες μεταβλητές, μπορούμε να υπολογίσουμε το εμβαδόν μεταξύ της καμπύλης και του άξονα x. Το εμβαδόν αυτό, είναι η mAP και δίδεται από τον παραπάνω τύπο.

Για κάθε κατηγορία αντικειμένων, υπολογίζεται η ακρίβεια ανά threshold (π.χ., IoU > 0.5, IoU > 0.75), και η mAP είναι ο μέσος όρος αυτών των τιμών για όλες τις κατηγορίες. Μια κοινή μέθοδος είναι να υπολογίζεται η mAP για κατώφλι **IoU > 0.5**, η οποία τότε ονομάζεται $mAP_{0.5}$. Αντίθετα, πιο ακριβή μοντέλα χρησιμοποιούν την mAP για ένα εύρος IoU κατωφλίων από 0.5 έως 0.95 και έχουμε τη $mAP_{0.5 : 0.95}$. Υπολογίζεται δηλαδή, η mAP για διαφορετικά κατώφλια, από το 0.5 έως το 0.95 με βήμα 0.5. Κατόπιν, λαμβάνουμε τη μέση τιμή και υπολογίζουμε την τιμή που προκύπτει.

Recall (Ανάκληση) [51]: Ο εν λόγω δείκτης, επικεντρώνεται στη δυνατότητα του μοντέλου να βρίσκει μεταξύ των πραγματικά θετικών παραδειγμάτων, σωστές προβλέψεις. Είναι ο λόγος του αριθμού των θετικών προβλέψεων προς αυτόν των θετικών παραδειγμάτων, οπότε υψηλή τιμή του λόγου αντιστοιχεί σε αυξημένη ικανότητα του μοντέλου να εντοπίζει θετικές προβλέψεις.

$$Recall = \frac{TP}{TP + FN}$$

Average Recall [51]: Η μέση Ανάκληση AR, υπολογίζει το ποσοστό των σωστών ανιχνεύσεων (true positives) σε σχέση με το συνολικό αριθμό των πραγματικών αντικειμένων. Δύναται ωστόσο, να υπολογίζεται και ως ο μέσος για διάφορα thresholds, κατ' αντιστοιχία με τη mAP.

Intersection over Union (Τομή Δια την Ένωση) [44]: Ο εν λόγω δείκτης, μετρά το ποσοστό της επικάλυψης μεταξύ προβλεπόμενης και πραγματικής μάσκας κατάτμησης. Την επικαλυπτόμενη περιοχή, δηλαδή, μεταξύ του προτεινόμενου bounding box από το μοντέλο και του πραγματικού ορίου. Ένα μοντέλο θεωρείται ότι έχει κάνει σωστή πρόβλεψη όταν το IoU ξεπερνά κάποιο κατώφλι – threshold, συνήθως 0.5 ή 50%.

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union} = \frac{TP}{TP + FP + FN}$$

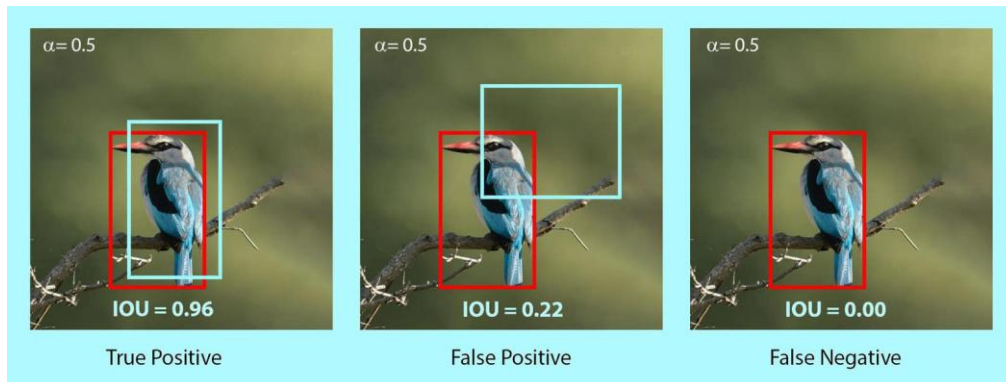
Όπου ,

TP : true positive, ο αριθμός των πραγματικών θετικών που εντοπίστηκαν ορθώς

FP : false positive, ο αριθμός των ψευδώς θετικών, που ανιχνεύθηκαν εσφαλμένως

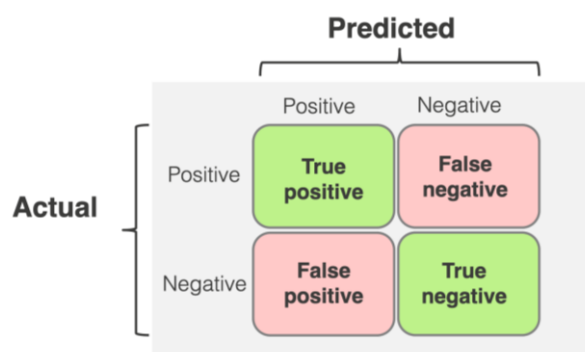
TN : true negative, ο αριθμός των πραγματικών αρνητικών που εντοπίστηκαν ορθώς

FN : false negative, ο αριθμός των ψευδώς αρνητικών, που δεν ανιχνεύονται ορθά



Εικόνα 32. Παράδειγμα μέτρησης IoU
Πηγή [44]

Για να κατανοήσουμε τα αποτελέσματα του δείκτη IoU, μπορούμε να κατασκευάσουμε τον πίνακα σύγχυσης – confusion matrix. Αν το μοντέλο ανιχνεύει ένα αντικείμενο με $\text{IoU} > 0.5$, τότε την προσθέτουμε στην κατηγορία True Positive (TP) στον πίνακα σύγχυσης. Αν το μοντέλο ανιχνεύει ένα αντικείμενο με $\text{IoU} < 0.5$ και το αντικείμενο υπάρχει πραγματικά, τότε προσθέτουμε στην κατηγορία False Negative (FN). Κατ' αντιστοιχία, αν το μοντέλο προτείνει ένα αντικείμενο ($\text{IoU} < 0.5$) και το αντικείμενο δεν υπάρχει στην πραγματικότητα, τότε προσθέτουμε στην κατηγορία False Positive (FP).



Εικόνα 33. Πίνακας Σύγχυσης
Πηγή [45]

Στόχος μας, αποτελεί να μεγιστοποιήσουμε τη διαγώνιο TP TN μειώνοντας παράλληλα τις εσφαλμένες προβλέψεις.

F1 Score [51]: Αποτελεί, συνολικό δείκτη αξιολόγησης, είναι ένας συνδυασμός των δεικτών της ανάκλησης και της ακρίβειας. Προκειμένου, να υπάρξει η εύρυθμη λειτουργία του μοντέλου, οι δύο αυτοί δείκτες πρέπει να βρίσκονται σε ισορροπία.



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall}$$

False Positives False Negatives : Η ελαχιστοποίηση των FN, FP αποτελεί βασικό παράγοντα βελτίωσης της συνολικής απόδοσης του μοντέλου.

Ημι-επιβλεπόμενη Μάθηση – Αυτοεκπαίδευση

Οι αλγόριθμοι και μεθοδολογίες που περιγράφηκαν ανωτέρω, εντάσσονται στην ευρύτερη κατηγορία της επιβλεπόμενης μάθησης. Απαραίτητη προϋπόθεση για την εκπαίδευση, δηλαδή, αποτελεί η ύπαρξη ενός επισημειωμένου συνόλου δεδομένων, είτε αυτό αφορά μία απλή ετικέτα κλάσης είτε bounding boxes. Εύκολα όμως, μπορούμε να υποθέσουμε, ότι δεν είναι πάντα δυνατό να υπάρχει ένα επαρκές σύνολο εκπαίδευσης.

4. Μεθοδολογία

Στο τρέχον αυτό κεφάλαιο, πραγματοποιείται μία ανάλυση της μεθοδολογίας που ακολουθήθηκε για την δημιουργία του συνόλου δεδομένων, καθώς και των εφαρμογών που χρησιμοποιήθηκαν για την εκπαίδευση του αλγορίθμου υπολογιστικής όρασης.

4.1 Θερμικό Αποτύπωμα Αρμάτων και Σύγχρονες Προκλήσεις

Στην παρούσα εργασία, επιχειρούμε να προσεγγίσουμε το πρόβλημα της αναγνώρισης τεθωρακισμένων αρμάτων μάχης, από εικόνα UAV, μέσω του θερμικού αποτυπώματός τους. Η φυσιογνωμία του πεδίου της μάχης, πολλές φορές -εάν όχι πάντοτε- καθιστά δύσκολη την αναγνώριση ενός εχθρικού στόχου, παρατηρώντας στο ορατό φάσμα. Προς τούτο, καθοριστικής σημασίας αποτελεί η χρήση οργάνων παρατήρησης στο υπέρυθρο φάσμα.



Όλα τα τεθωρακισμένα άρματα μάχης, παρά τις εγγενείς τους διαφορές, λόγω διαφορετικής χώρας κατασκευής ή περιβάλλοντος στο οποίο αναμένεται να επιχειρήσουν, παρουσιάζουν ορισμένες ομοιότητες. Στο υπέρυθρο φάσμα, κάθε άρμα μάχης παρουσιάζει θερμικά μοτίβα που συνδέονται με τη λειτουργία των μηχανικών του μερών, όπως η θερμότητα που εκπέμπεται από τον κινητήρα, τις εξαγωγές αερίων καύσης, τους εδαφικούς τροχούς της ερπύστριας λόγω της αυξημένης πίεσης. Βέβαια, σημαντικό ρόλο, έχουν και οι καιρικές συνθήκες, οι οποίες μπορεί να δυσχεράνουν την παρατήρηση. Με αυτό τον τρόπο, ένας εκπαιδευμένος πυροβολητής, μπορεί να αναγνωρίσει την ταυτότητα του άρματος που έχει απέναντί του.



Εικόνα 34. Θερμικό αποτύπωμα αρμάτων μάχης M1 Abrams (Αριστερά) και T-72 (Δεξιά)

Ωστόσο, η αδυναμία αυτή των αρμάτων μάχης, να εκπέμπουν ένα τόσο δυνατό θερμικό αποτύπωμα, δεν έχει μείνει χωρίς αντίμετρο. Πλέον, χρησιμοποιούνται προηγμένης τεχνολογίας δίχτυα παραλλαγής (anti-IR), τα οποία καλύπτουν το άρμα, μειώνοντας το θερμικό αποτύπωμα στο υπέρυθρο φάσμα, αλλά και αναμειγνύοντας το με τα χαρακτηριστικά του περιβάλλοντος που επιχειρεί. Παράλληλα, καινοτομία στον τομέα της κάλυψης του θερμικού αποτυπώματος, αποτελεί το σύστημα επικάλυψης της ελληνικής εταιρείας Intermat Defence[25], το οποίο είναι ένα σύστημα επίστρωσης προηγμένης τεχνολογίας, για την κάλυψη του θερμικού αποτυπώματος – thermal signature, που καλύπτει τις εκπομπές θερμότητας του κάθε μέσου, δίνοντας αποφασιστικό πλεονέκτημα στον χρήστη.



Εικόνα 35. Απόκρυψη Θερμικού Αποτυπώματος

Συνοψίζοντας, με βάση τα παραπάνω, μπορούμε να αντιληφθούμε τη σημασία που έχει το θερμικό αποτύπωμα για την ανίχνευση ενός στόχου στο πεδίο της μάχης. Ο έγκαιρος εντοπισμός αυτού, από χειριστές αντιαρματικών(A/T), πιλότους εχθρικών ελικοπτέρων (ΕΕΦ) ή άλλα άρματα μάχης (ΤΘ), δίνει τακτικό πλεονέκτημα εξασφαλίζοντας επιτυχημένη βολή. Στο σύγχρονο πεδίο της μάχης, τα τελευταία έτη, εντάχθηκε ένας ακόμη παράγοντας κινδύνου για τα άρματα μάχης -τα UAV.

Η ευρεία χρήση των UAV, εναντίον αρμάτων μάχης πραγματοποιήθηκε στον πόλεμο μεταξύ Ουκρανίας και Ρωσίας. Οι Ουκρανοί, προκειμένου να αντιμετωπίσουν τη ρωσική υπεροπλία, σε μέσα και προσωπικό, κατέφυγαν σε ανορθόδοξες, μέχρι τότε μεθόδους. Εισήγαγαν τη χρήση UAV First Person View (FPV), για τον εντοπισμό και καταστροφή αρμάτων μάχης. Το UAV αυτό φέρει εκρηκτικά και χειριζόμενο από ένα χειριστή αρκετά μακριά από το πεδίο της μάχης, μπορεί να καταστρέψει κρίσιμα σημεία ενός άρματος, είτε πέφτοντας επάνω στον κινητήρα του, είτε ρίχνοντας τα εκρηκτικά στο επάνω μέρος του πύργου του[26][27]. Και στις δύο περιπτώσεις, το κόστος του UAV, είναι μηδαμινό εμπρός σε αυτό του άρματος μάχης, επιφέροντας παράλληλα τεράστιο ψυχολογικό αντίκτυπο στο πλήρωμα αυτού, καθώς δύσκολα γίνεται αντιληπτό. Η τεχνολογία αυτή έχει επηρεάσει την τακτική του πολέμου, καθώς οι ρωσικές δυνάμεις αναγκάζονται να κρατούν τα άρματα τους πιο μακριά από την πρώτη γραμμή, για να αποφύγουν την καταστροφή από τα UAV[27].



4.2 Σύνολο Δεδομένων

Προκειμένου να πραγματοποιήσουμε την εκπαίδευση του αλγορίθμου ανίχνευσης αντικειμένων, απαιτείται η ύπαρξη ενός επισημειωμένου συνόλου δεδομένων. Το σύνολο αυτό αποτελείται από ξεχωριστά σύνολα εκπαίδευσης (training), επικύρωσης (validation) και ελέγχου (test). Λαμβάνοντας υπόψιν, την υπάρχουσα έρευνα στον τομέα της αναγνώρισης αρμάτων μάχης, παρατηρείται η έλλειψη συνόλων δεδομένων που να περιέχουν τα θερμικά αποτυπώματα αρμάτων μάχης. Η πλειοψηφία των υπαρκτών συνόλων αποτελούν ιδιοκτησία των Ε.Δ των ΗΠΑ, ενώ κάποιες έρευνες προτείνουν τη χρήση συνθετικών εικόνων προσομοίωσης των αρμάτων στο υπέρυθρο φάσμα. Αντίθετα, μεγάλη είναι η διαθεσιμότητα συνόλων δεδομένων που περιέχουν φωτογραφίες αρμάτων – είτε πραγματικών, είτε μοντέλων – στο ορατό φάσμα.

Στην παρούσα εργασία, επιχειρήσαμε τη δημιουργία του δικού μας συνόλου δεδομένων με βίντεο και φωτογραφίες αρμάτων στο υπέρυθρο φάσμα. Βασική επιδίωξη, αποτέλεσε η αναζήτηση υλικού το οποίο λήφθηκε από UAV ή μεγάλο ύψος, σε διάφορες καιρικές συνθήκες, ώστε να αποτελεί όσο το δυνατόν μία καλύτερη προσέγγιση της πραγματικότητας. Κύρια πηγή αποτέλεσαν, βίντεο μέσω πλατφορμών όπως το YouTube και X.

4.2.1 Τεχνικές Επεξεργασίας Δεδομένων

Για να καταφέρει το νευρωνικό δίκτυο, να αναγνωρίζει μοτίβα σε πληθώρα αντικειμένων, απαιτείται μεγάλος όγκος δεδομένων εκπαίδευσης. Η προεπεξεργασία εικόνας είναι η διαδικασία κατά την οποία, εκτελείται η μετατροπή των ακατέργαστων δεδομένων εικόνας σε μια χρησιμοποιήσιμη και ουσιαστική μορφή. Μας δίνει τη δυνατότητα, να αφαιρέσουμε τις ανεπιθύμητες παραμορφώσεις και να ενισχύσουμε συγκεκριμένες ιδιότητες που είναι καίριες για εφαρμογές υπολογιστικής όρασης. Η προεπεξεργασία, αποτελεί ένα κρίσιμο πρώτο στάδιο για την προετοιμασία των δεδομένων της εικόνας προτού τροφοδοτηθούν σε μοντέλα μηχανικής εκμάθησης.

Υπάρχουν διάφορες τεχνικές που χρησιμοποιούνται στην προεπεξεργασία εικόνας[8]:

- **Αλλαγή μεγέθους (Resizing):** Η αλλαγή μεγέθους εικόνων σε ομοιόμορφο μέγεθος είναι σημαντική για τη σωστή λειτουργία των αλγορίθμων μηχανικής μάθησης. Ένα παράδειγμα χρήσης είναι με τη μέθοδο `resize()` του OpenCV, με την οποία μπορούμε να αλλάξουμε το μέγεθος των εικόνων.
- **Μετατροπή σε κλίμακα του γκρι (Grayscale):** Η μετατροπή έγχρωμων εικόνων σε κλίμακα του γκρι δύναται να απλοποιήσει τα δεδομένα της εικόνας, με αποτέλεσμα να μειώσει τις υπολογιστικές ανάγκες για ορισμένους αλγόριθμους.
- **Μείωση θορύβου (Noise reduction):** Οι τεχνικές εξομάλυνσης (smoothing), θολώματος (blurring) και φιλτραρίσματος δύναται να εφαρμοστούν για να αφαιρέσουμε τον ανεπιθύμητο θόρυβο από διάφορες εικόνες του συνόλου. Οι μέθοδοι `GaussianBlur()` και `medianBlur()`, μπορούν να χρησιμοποιηθούν συνήθως για αυτό.
- **Κανονικοποίηση (Normalization):** Η κανονικοποίηση προσαρμόζει τις τιμές έντασης των εικονοστοιχείων σε ένα επιθυμητό εύρος, μεταξύ 0 και 1, βελτιώνοντας έτσι, την απόδοση των μοντέλων μηχανικής εκμάθησης.



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



- **Binarization**: Η Binarization μετατρέπει τις εικόνες σε κλίμακα του γκρι σε ασπρόμαυρες με όριο. Η μέθοδος threshold () χρησιμοποιείται για τη δυαδοποίηση εικόνων στο OpenCV.

- **Βελτίωση αντίθεσης (Contrast enhancement)**: Η αντίθεση των εικόνων μπορεί να ρυθμιστεί χρησιμοποιώντας την εξίσωση ιστογράμματος. Η μέθοδος equalizeHist () ενισχύει την αντίθεση των εικόνων.

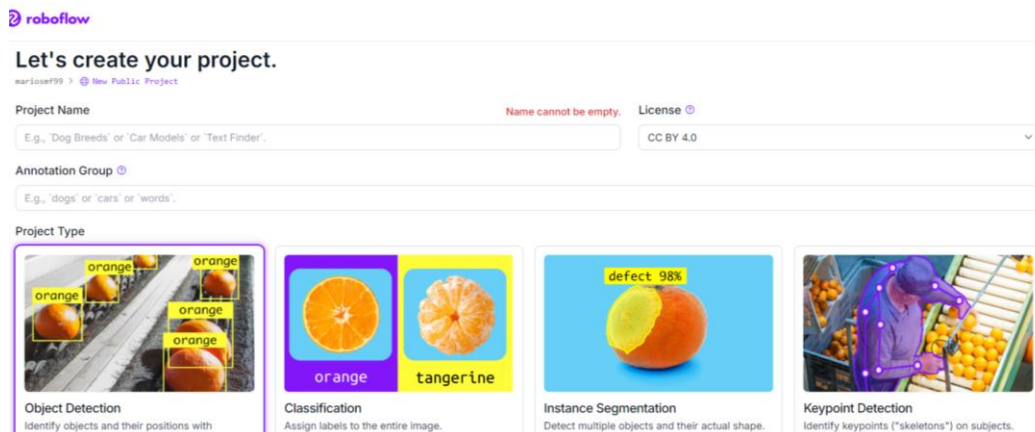
Με τον σωστό συνδυασμό αυτών των τεχνικών, έχουμε τη δυνατότητα να βελτιώσουμε αισθητά τα δεδομένα εικόνας και να δημιουργήσουμε καλύτερες εφαρμογές υπολογιστικής όρασης. Η προεπεξεργασία εικόνας, επιτρέπει να βελτιώσουμε τις ακατέργαστες εικόνες σε μια μορφή κατάλληλη για το εκάστοτε πρόβλημα.

4.2.2 Επισημείωση Δεδομένων

Η επισημείωση των δεδομένων, έχει πρωταγωνιστικό ρόλο στην επιβλεπόμενη εκπαίδευση αλγορίθμων υπολογιστικής όρασης. Η επισημείωση δεδομένων σχολιάζει τα ανεπεξέργαστα δεδομένα με σημαντικές ετικέτες, παρέχοντας πλαίσιο και κατηγοριοποίηση για τα μοντέλα μηχανικής μάθησης (ML) προς κατανόηση. Αυτές οι ετικέτες χρησιμεύουν ως βασικοί οδηγοί για τα μοντέλα ML, επιτρέποντάς τους να ερμηνεύουν αποτελεσματικά τα δεδομένα. Στην αναγνώριση εικόνων, ετικέτες όπως "γάτα" ή "σκύλος" ορίζουν κατηγορίες αντικειμένων, ενώ στην ανάλυση κειμένου, οι ετικέτες δηλώνουν συναισθήματα ή ονομαστικές οντότητες.

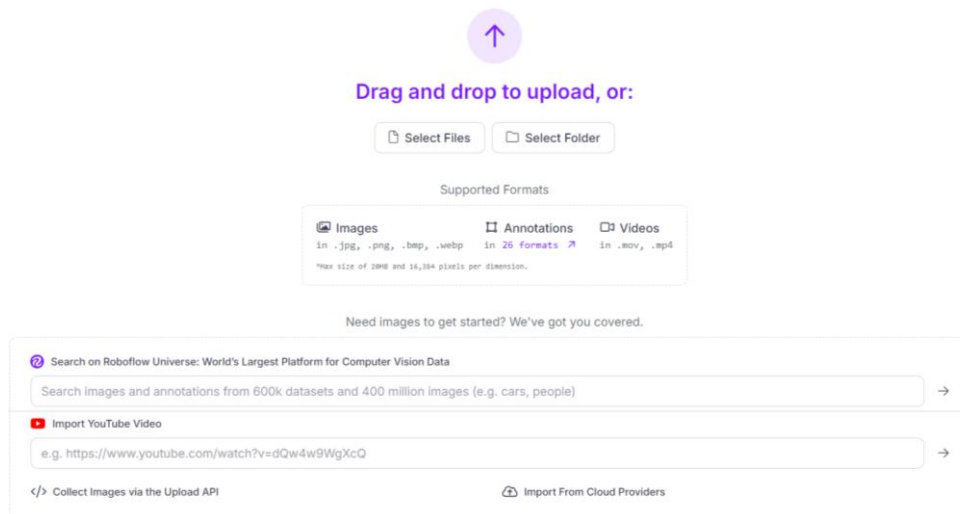
Η επισήμανση δεδομένων μετατρέπει τα ακατέργαστα δεδομένα σε μια κατανοητή μορφή για μοντέλα ML, διευκολύνοντας την αναγνώριση προτύπων και τις δυνατότητες πρόβλεψης. Στην παρούσα εργασία, χρησιμοποιήθηκε το εργαλείο Roboflow Annotator, της ομώνυμης εταιρείας Roboflow. Επιλέχθηκε η χειροκίνητη επισημείωση των εικόνων με ορθογώνιο bounding box αλλά και με πολύγωνο, όπου η ευκρίνεια του άρματος ήταν μειωμένη. Επιπλέον, ο επιλογέας κλάσης εντός του εργαλείου, επιτρέπει στους σχολιαστές να αντιστοιχίσουν την κατάλληλη ετικέτα σε κάθε bounding box. Η διαδικασία αυτή πραγματοποιήθηκε για 360 φωτογραφίες διασφαλίζοντας τη διαθεσιμότητα δεδομένων με ακριβή σήμανση για την μετέπειτα εκπαίδευση του μοντέλου ανίχνευσης αντικειμένων.

Αφού δημιουργήσαμε έναν λογαριασμό στο Roboflow, ξεκινήσαμε ένα νέο project με τίτλο mbt (main battle tank). Κατόπιν δίνεται η επιλογή στον χρήστη να επιλέξει αν θέλει, να εκπαιδεύσει ένα μοντέλο για **ανίχνευση αντικειμένων (object detection)**, **ταξινόμηση εικόνας (image classification)** ή **τμηματοποίηση (segmentation)**.



Εικόνα 36. Διαδικασία Δημιουργίας Project για Επισημείωση Δεδομένων

Το Roboflow, δίνει τη δυνατότητα να ανεβάσουμε το σύνολο των εικόνων που επιθυμούμε να επισημειώσουμε είτε από τον υπολογιστή, είτε απευθείας από άλλες πλατφόρμες όπως το Google Drive ή το Dropbox. Παράλληλα, δύναται να εισαχθούν απευθείας βίντεο από το YouTube, εισάγοντας απλώς τη διεύθυνση URL.



Εικόνα 37. Εισαγωγή Δεδομένων

Με την εισαγωγή βίντεο, μας δίνει τη δυνατότητα να επιλέξουμε τον αριθμό frame ανά δευτερόλεπτο, για να πραγματοποιηθεί η επεξεργασία του βίντεο. Με αυτό τον τρόπο, μπορούμε να εξάγουμε μεγάλο αριθμό εικόνων έτοιμων προς επισημείωση.

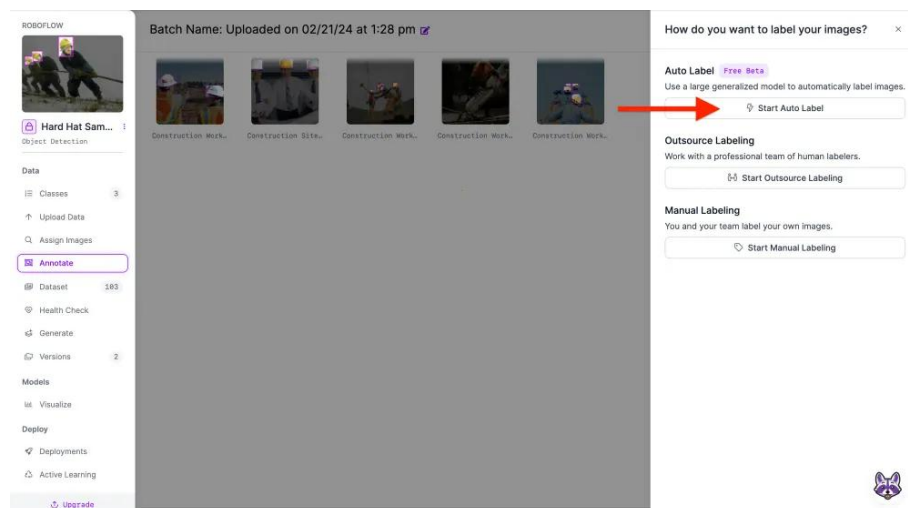
Η επισημείωση δεδομένων πραγματοποιείται με τα παρακάτω εργαλεία:

- **Bounding Boxes:** Για εκπαίδευση μοντέλου ανίχνευσης αντικειμένων, συνήθως χρησιμοποιούνται τα "Bounding Boxes" ώστε να περικλείσουμε τα αντικείμενα που μας ενδιαφέρουν.
- **Polygons:** Σε προβλήματα τμηματοποίησης (segmentation), μπορούμε να χρησιμοποιήσουμε ελεύθερες γραμμές για πιο ακριβή επισημείωση των αντικειμένων με πολύπλοκα σχήματα. Όμως, για να πετύχουμε μεγαλύτερη ακρίβεια και σε προβλήματα ανίχνευσης αντικειμένων, είναι δυνατόν να χρησιμοποιηθεί η συγκεκριμένη μέθοδος επισημείωσης. Η πολυγωνικές

επισημειώσεις, επιτυγχάνουν μεγαλύτερα mAP_{0.5} συγκριτικά με τα bounding boxes[28].

- **Tags:** Τέλος, για τα προβλήματα ταξινόμησης, μπορούμε απλώς να προσθέσουμε ετικέτες (tags) σε κάθε εικόνα.

Η χειροκίνητη επισημείωση δεδομένων αποτελεί μία χρονοβόρα διαδικασία, ειδικά όταν πρόκειται για μεγάλο όγκο δεδομένων ή αντικείμενα με πολύπλοκες μορφές. Η διαδικασία αυτή, ακόμη και όταν εκτελείται από έμπειρο προσωπικό μπορεί να φτάσει και έως μερικά λεπτά, για μία μόνο εικόνα. Σε αυτό το σημείο το Roboflow καινοτομεί και προσφέρει επίσης δυνατότητες **αυτόματης επισημείωσης** με προεκπαιδευμένα μοντέλα, ώστε να μειώσουμε το χρόνο που απαιτείται για την επισημείωση μεγάλου όγκου δεδομένων. Ωστόσο η αυτόματη επισημείωση, πραγματοποιείται με ήδη προεκπαιδευμένα μοντέλα, για αυτό είναι απαραίτητος ο έλεγχος και διόρθωση εάν απαιτείται.



Εικόνα 38. Robflow Auto-Label

Αφού ολοκληρωθεί η επισημείωση, μπορούμε να εξάγουμε τα δεδομένα σε διάφορες μορφές που υποστηρίζονται από μοντέλα όπως YOLO, TensorFlow, PyTorch, και άλλες δημοφιλείς βιβλιοθήκες μηχανικής μάθησης.

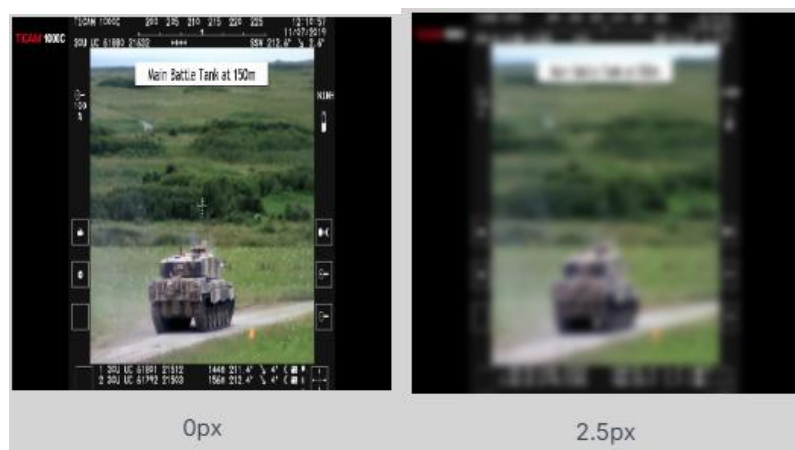
Παράλληλα, μέσω του Roboflow έχουμε τη δυνατότητα για data augmentation. **Επαύξηση δεδομένων (data augmentation)** είναι η διαδικασία κατά την οποία δημιουργούμε παραλλαγές των υπαρχουσών εικόνων σε ένα σύνολο δεδομένων, ώστε να δημιουργήσουμε νέες εικόνες. Για παράδειγμα, μπορούμε να αυξήσουμε δεδομένα αυξάνοντας ή μειώνοντας τη φωτεινότητα μιας εικόνας. Αυτό θα βοηθήσει το μοντέλο μας να μάθει να αναγνωρίζει καλύτερα αντικείμενα ενδιαφέροντος σε διαφορετικές συνθήκες φωτισμού. Μπορούμε επίσης να εφαρμόσουμε την επαύξηση δεδομένων αλλάζοντας τη γωνία της εικόνας, βοηθώντας το μοντέλο να ανιχνεύει αντικείμενα από διαφορετικές γωνίες.

Ας φανταστούμε ένα σενάριο όπου θέλουμε να εκπαιδεύσουμε ένα μοντέλο να ανιχνεύει οχήματα. Η εφαρμογή επαύξησης δεδομένων μέσω μεταβολής της φωτεινότητας θα μας ήταν χρήσιμη, επειδή το μοντέλο μας μπορεί να χρησιμοποιηθεί σε συνθήκες βροχής και ηλιοφάνειας, όπου η φωτεινότητα ποικίλλει. Προσθέτοντας

αυτή τη μέθοδο επαύξησης δεδομένων, θα έχουμε περισσότερα παραδείγματα από τα οποία το μοντέλο θα μπορεί να εκπαιδευτεί.

Υπάρχουν πολλοί διαφορετικοί τρόποι να αυξήσουμε δεδομένα σε ένα σύνολο εικόνων. Ακολουθούν μερικές από τις πιο κοινές μεθόδους:

- **Περιστροφή (Rotation):** Περιστρέφουμε τις εικόνες κατά διαφορετικές μοίρες (π.χ. 90° , 180°). Αυτό είναι χρήσιμο αν το μοντέλο αναμένεται να αναγνωρίζει αντικείμενα που βρίσκονται σε γωνία.
- **Τυχαία αποκοπή (Random crop):** Αποκόπτουμε την εικόνα σε τυχαία σημεία. Μπορεί να αποκόψουμε μέρος ενός αντικειμένου, επιτρέποντας στο μοντέλο να μάθει να αναγνωρίζει αντικείμενα που δεν είναι πλήρως ορατά.
- **Φωτεινότητα (Exposure):** Αυξάνουμε ή μειώνουμε τη φωτεινότητα μιας εικόνας, κάτι χρήσιμο για χρήση σε διαφορετικές συνθήκες φωτισμού.
- **Θάμπωμα (Blur):** Εφαρμόζουμε εφέ θολώματος σε μια εικόνα.
- **Αναστροφή (Flip):** Αναστρέφουμε μια εικόνα από πάνω προς τα κάτω ή από αριστερά προς τα δεξιά. Αυτό πρέπει να αποφεύγεται αν η εικόνα περιέχει κείμενο.
- **Κορεσμός (Saturation):** Αλλάζουμε την ένταση των χρωμάτων σε μια εικόνα. Αυτό είναι χρήσιμο αν οι συνθήκες φωτισμού μπορεί να ποικίλλουν.
- **Τυχαίος θόρυβος (Random noise):** Προσθέτουμε λευκά και μαύρα pixel σε μια εικόνα, κάνοντάς την πιο «θολή».
- **Μωσαϊκό (Mosaic augmentation):** Συνδυάζουμε διαφορετικές εικόνες μεταξύ τους. Αυτή η τεχνική είναι ιδιαίτερα χρήσιμη για έργα αεροφωτογραφιών.



Εικόνα 39. Εφαρμογή Blur σε εικόνα του συνόλου δεδομένων

Για την επαύξηση του συνόλου εκπαίδευσης του μοντέλου μας, χρησιμοποιήσαμε τις εξής μεθόδους

Flip: Horizontal (οριζόντιο)



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



Hue: Between -20° and $+20^\circ$

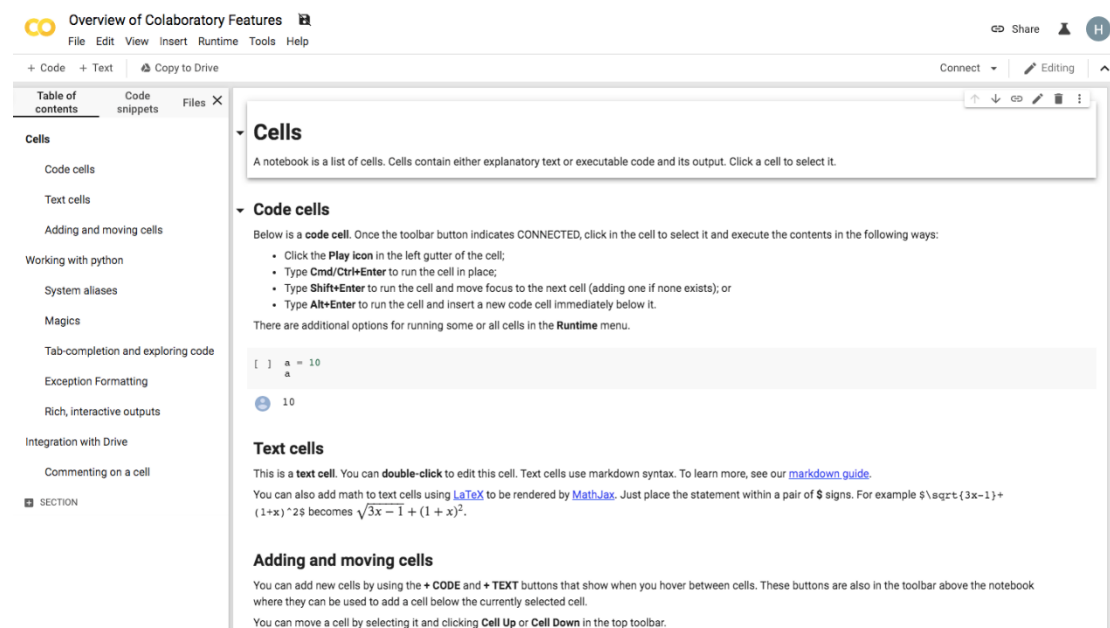
Exposure: Between -6% and $+6\%$

Blur: Up to 2.5px

Έτσι το επαυξημένο, πλέον, σύνολο δεδομένων μας ανέρχεται στις **625 εικόνες**.

4.3 Υλοποίηση με Google Colab

Η υλοποίηση του αλγορίθμου ανίχνευσης αντικειμένων υλοποιήθηκε με την εφαρμογή Google Colab. Το Google Colab, αποτελεί εφαρμογή cloud, της εταιρείας Google, διαθέσιμη από το 2017, η οποία χρησιμοποιώντας τα σημειωματάρια Jupyter, μπορεί να εκτελεστεί και συνταχθεί κώδικας Python. Με αυτόν τον τρόπο, καθένας χρήστης του διαδικτύου, μπορεί να έχει πρόσβαση σε περιβάλλον Python και να χρησιμοποιήσει βιβλιοθήκες μηχανικής μάθησης.



Εικόνα 40. Παράδειγμα σημειωματαρίου Google Colab

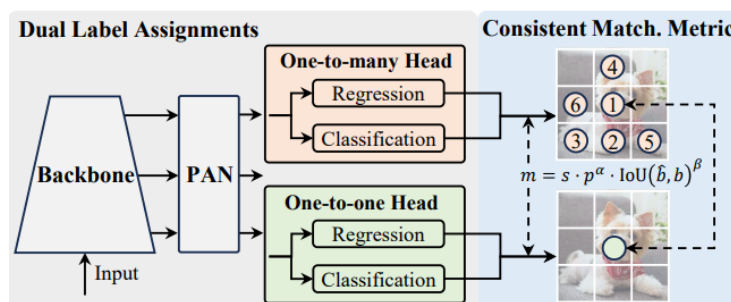
Το μεγάλο πλεονέκτημα της συγκεκριμένης εφαρμογής, είναι η εκτέλεση του κώδικα σε εικονική μηχανή. Τα παραδοσιακά σημειωματάρια Jupyter αποθηκεύονται τοπικά και ο κώδικας εκτελείται χρησιμοποιώντας το υλικό του τοπικού υπολογιστή κάθε χρήστη. Ακόμα κι αν διατίθεται ένας πολύ γρήγορος υπολογιστής, οι δυνατότητές του είναι περιορισμένες σε σύγκριση με την υπολογιστική ισχύ που παρέχεται από το Google Colab [30].

4.4 Αλγόριθμος YOLOv10

Ο αλγόριθμος YOLOv10, αποτελεί μία σημαντική πρόοδο στον τομέα της ανίχνευσης αντικειμένων σε πραγματικό χρόνο. Βασίζεται επάνω στην επιτυχία προηγούμενων εκδόσεων, αντιμετωπίζοντας παράλληλα μεγάλες και δυσεπίλυτες προκλήσεις, όπως η εξάρτηση από το Non-Maximum Suppression (NMS) για επεξεργασία μετά την ανίχνευση, που αύξανε την καθυστέρηση κατά την ανάπτυξη. Δια μέσου της εκπαίδευσης χωρίς NMS και της βελτίωσης του αρχιτεκτονικού σχεδιασμού, επιτυγχάνεται καλύτερη ακρίβεια και υψηλότερη αποδοτικότητα.

Όπως και αναφέρθηκε πρότερα, ο αλγόριθμος NMS, χρησιμοποιείται για το φιλτράρισμα των εικόνων που καταλήγουν στον ανιχνευτή ενός σταδίου. Κατά την εκπαίδευση, το YOLO χρησιμοποιούν συνήθως το TAL[31] για την κατανομή πολλαπλών θετικών δειγμάτων ανά αντικείμενο. Η υιοθέτηση της στρατηγικής αυτής παρέχει άφθονα εποπτικά σήματα, βελτιστοποιώντας την απόδοση. Ωστόσο, αυτό απαιτεί τη χρήση του NMS για την τελική επιλογή, μειώνοντας την αποδοτικότητα. Αν και προηγούμενες έρευνες, όπως οι [33][32], εξετάζουν τη στρατηγική "ένα προς ένα" ή one to one (o2o) για να περιορίσουν τις περιττές προβλέψεις, συχνά αυξάνουν το υπολογιστικό βάρος ή επιφέρουν υποδεέστερη απόδοση.

Σε αντίθεση με την ανάθεση «ένα προς πολλά» ή one to many – o2m, η αντιστοίχιση o2o αναθέτει μόνο μία πρόβλεψη σε κάθε αληθινό δείγμα, αποφεύγοντας την επεξεργασία NMS. Ωστόσο, αυτό οδηγεί σε αδύνατη εποπτεία, η οποία προκαλεί μη βέλτιστη ακρίβεια και ταχύτητα σύγκλισης. Ευτυχώς, αυτή η ανεπάρκεια μπορεί να αντισταθμιστεί από την ανάθεση o2m[34]. Για να το πετύχουμε αυτό, εισάγουμε διπλές αναθέσεις ετικετών για το YOLO προκειμένου να συνδυάσουμε τα καλύτερα από τις δύο στρατηγικές, ενσωματώνοντας μια άλλη κεφαλή "o2o" για τα YOLO. Διατηρεί την ίδια δομή και υιοθετεί τους ίδιους στόχους βελτιστοποίησης με τον αρχικό κλάδο "o2m", αλλά αξιοποιεί την αντιστοίχιση "o2o" για να αποκτήσει αναθέσεις ετικετών. Κατά τη διάρκεια της εκπαίδευσης, οι δύο κεφαλές βελτιστοποιούνται από κοινού με το μοντέλο, επιτρέποντας στο backbone και neck να έχουν εποπτεία λόγω της ανάθεσης "o2m". Κατά την εξαγωγή, απορρίπτουμε την κεφαλή "o2m" και χρησιμοποιούμε την κεφαλή "o2o" για να κάνουμε προβλέψεις. Αυτό επιτρέπει στα YOLO για την ανάπτυξη end-to-end χωρίς να επιβαρύνονται με επιπλέον κόστος εξαγωγής. Επιπλέον, στην αντιστοίχιση "o2o", υιοθετούμε την επιλογή του βέλτιστου ενός, η οποία επιτυγχάνει την ίδια απόδοση με την Hungarian αντιστοίχιση με λιγότερο πρόσθετο χρόνο εκπαίδευσης[35].



Εικόνα 41. Συνεπείς διπλές αναθέσεις για εκπαίδευση χωρίς NMS
Πηγή [35]



5. Υλοποίηση Εκπαίδευσης Αλγορίθμου

Στα προηγούμενα κεφάλαια, παρουσιάστηκε το θεωρητικό υπόβαθρο που είναι απαραίτητο για την κατανόηση των συνελκτικών νευρωνικών δικτύων, καθώς και της ανίχνευσης αντικειμένων. Παράλληλα, παρουσιάστηκε και το λογισμικό που θα χρησιμοποιηθεί για την υλοποίηση της αναγνώρισης των θερμικών αποτυπωμάτων αρμάτων. Στο παρόν κεφάλαιο, θα παρουσιαστεί η υλοποίηση και εκπαίδευση του αλγορίθμου στο σύνολο δεδομένων που δημιουργήσαμε.

✓ 1ο Βήμα

Προκειμένου να βελτιώσουμε τον χρόνο εκπαίδευσης του αλγορίθμου θα χρησιμοποιήσουμε την εντολή `nvidia-smi`, για να αποκτήσουμε πρόσβαση στην κάρτα γραφικών

```
!nvidia-smi
```

Η εντολή `nvidia-smi` (NVIDIA System Management Interface) χρησιμοποιείται για την παρακολούθηση και τη διαχείριση των καρτών γραφικών NVIDIA σε ένα υπολογιστικό σύστημα. Αποτελεί εργαλείο γραμμής εντολών για παροχή πληροφοριών σχετικά με την κατάσταση των GPU, τη χρήση τους και άλλες σημαντικές παραμέτρους.

```
[ ] import os  
HOME = os.getcwd()  
print(HOME)
```

Εισάγουμε το Python module “os”, ώστε να αποκτήσουμε πρόσβαση σε λειτουργίες για αλληλεπίδραση με το λειτουργικό σύστημα και με την εντολή `os.getcwd()` παίρνουμε τη διαδρομή του φακέλου που εκτελείται το script.

```
[ ] !pip install -q git+https://github.com/THU-MIG/yolov10.git
```

```
Installing build dependencies ... done  
Getting requirements to build wheel ... done  
Preparing metadata (pyproject.toml) ... done  
Building wheel for ultralytics (pyproject.toml) ... done
```

Εγκαθιστούμε το YOLOv10, τη νεότερη έκδοση του YOLO, απευθείας από το αποθετήριο του github.

```
[ ] !pip install -q supervision roboflow
```

```
151.5/151.5 kB 7.0 MB/s eta 0:00:00  
80.3/80.3 kB 7.5 MB/s eta 0:00:00  
66.8/66.8 kB 6.1 MB/s eta 0:00:00  
54.5/54.5 kB 4.8 MB/s eta 0:00:00
```

Εγκαθιστούμε, τα πακέτα `supervision` και `roboflow`. Με το `supervision` μπορούμε να επεξεργαστούμε και οπτικοποιήσουμε τις εξόδους των object detectors, όπως τα μοντέλα YOLO. Παράλληλα, μπορεί να βοηθήσει στον έλεγχο και διαχείριση των αποτελεσμάτων από το μοντέλο ανίχνευσης. Με το `roboflow`, εγκαθιστάται μια βιβλιοθήκη Python που μας επιτρέπει την αλληλεπίδραση με την πλατφόρμα **Roboflow**, η οποία προσφέρει εργαλεία για την εύκολη διαχείριση και την ανάπτυξη datasets για εφαρμογές υπολογιστικής όρασης, όπως ανίχνευση αντικειμένων,





ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



```
!sed -i '$d' {dataset.location}/data.yaml  
!sed -i '$d' {dataset.location}/data.yaml  
!sed -i '$d' {dataset.location}/data.yaml  
!sed -i '$d' {dataset.location}/data.yaml  
!sed -i '$d' {dataset.location}/data.yaml  
!echo -e "test: ../test/images\ntrain: ../train/images\nval: ../valid/images" >> {dataset.location}/data.yaml
```

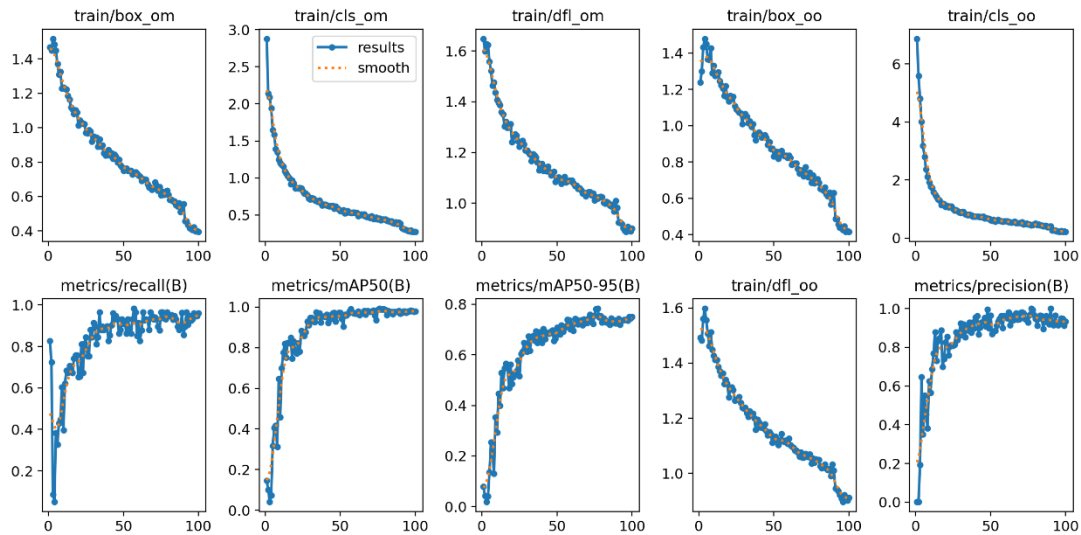
Επειδή χρησιμοποιούμε σύνολο δεδομένων εισηγμένο από το Roboflow, προσθέτουμε τις παραπάνω γραμμές κώδικα, ώστε να είναι δυνατές οι τροποποιήσεις του περιεχομένου του αρχείου data.yaml. Με τις εντολές !sed διαγράφονται οι τελευταίες τέσσερις γραμμές του αρχείου, ενώ με την εντολή !echo, προστίθενται στο αρχείο, νέες γραμμές με διαδρομές προς τα δεδομένα test, train, val.

```
[77] %cd {HOME}  
  
!yolo task=detect mode=train epochs=100 batch=32 plots=True\  
model={HOME}/weights/yolov10n.pt \  
data={dataset.location}/data.yaml
```

Εκπαιδεύουμε το YOLO με στόχο την ανίχνευση αντικειμένων για 100 εποχές και με batch size 16 εικόνων.

6. Αποτελέσματα

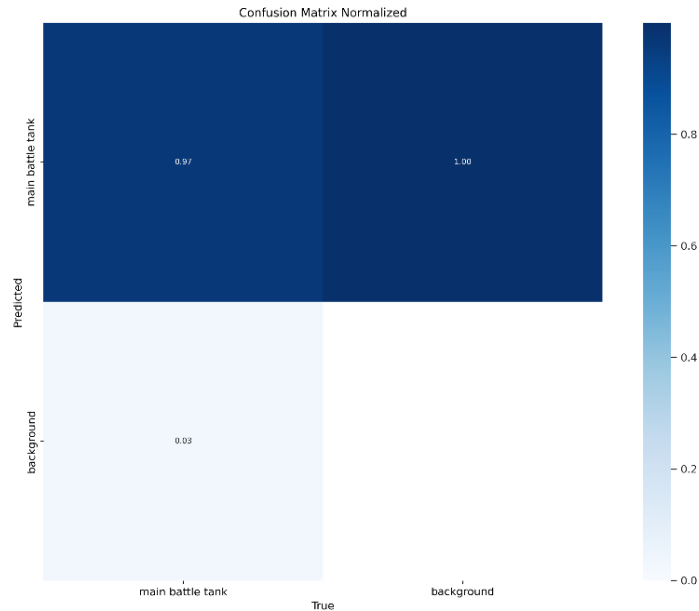
Με την ολοκλήρωση της εκπαίδευσης του μοντέλου, προκύπτουν τα ακόλουθα γραφήματα, τα οποία θα σχολιάσουμε, προκειμένου να εξάγουμε συμπεράσματα, τόσο για την ακρίβεια του μοντέλου, όσο και του συνόλου δεδομένων που χρησιμοποιήσαμε.



Διάγραμμα 1. Διαγράμματα Εκπαίδευσης

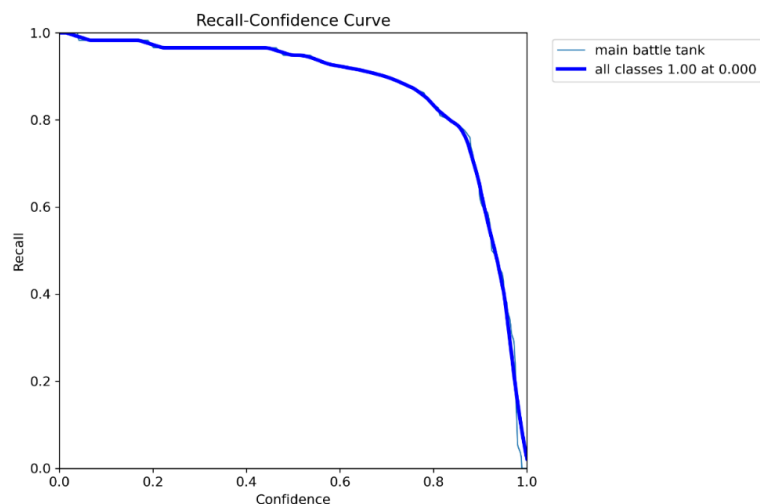
Αρχικά, τα 5 πρώτα διαγράμματα, τα οποία αφορούν την εκπαίδευση του μοντέλου (Training Metrics Evolution). Τα διαγράμματα αυτά, αντιπροσωπεύουν τη συμπεριφορά των αποτυχιών “loss”, των bounding boxes – box, classes – cls, Distribution Focal Loss – df_l, κατά τη διάρκεια της εκπαίδευσης. Ο δείκτης om – on-model, αφορά αντικείμενα που ανήκουν στην ίδια κλάση με αυτή που το μοντέλο έχει εκπαιδευτεί, δηλαδή τα άρματα μάχης, ενώ oo – out-of- αντικείμενα τα οποία δεν ανήκουν στις κλάσεις που εκπαιδεύεται το μοντέλο ή πρέπει να αγνοήσει. Παρατηρούμε λοιπόν, ότι στα διαγράμματα αυτά υπάρχει σταθερή μείωση του σφάλματος, με την πάροδο των εποχών εκπαίδευσης, το οποίο υποδηλώνει ότι το μοντέλο μαθαίνει από τα σφάλματά του, υποδεικνύοντας σταθερή εκπαίδευση.

Τα 4 διαγράμματα που ακολουθούν, αφορούν metrics recall, mAP₅₀, mAP₅₀₋₉₅, και precision. Παρατηρούμε μία σταθερή αύξηση των καμπυλών, η οποία καθώς το μοντέλο εκπαιδεύεται πλησιάζει στο 1, το οποίο μεταφράζεται σε υψηλή ακρίβεια του μοντέλου, καθώς και στην επιτυχία ανακλήσεων στις προβλέψεις που αυτό κάνει.



Διάγραμμα 2. Πίνακας Σύγκυσης (Confusion Matrix)

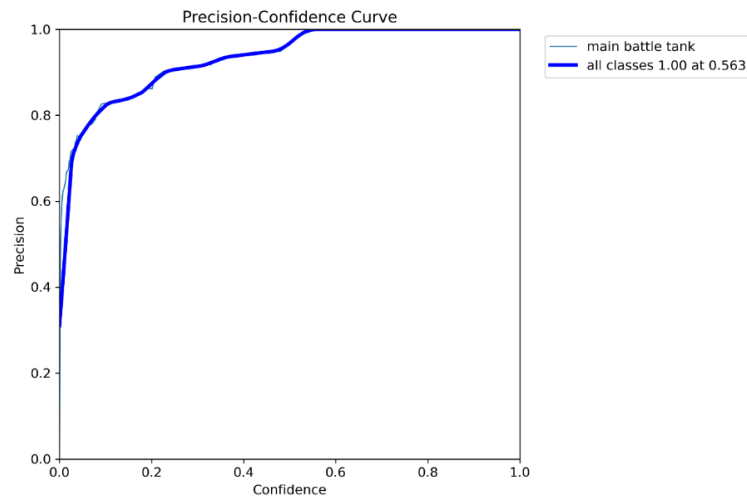
Ένα ιδιαίτερα χρήσιμο βοήθημα για την αξιολόγηση της αξιοπιστίας του μοντέλου, αποτελεί και ο πίνακας σύγκυσης – confusion matrix. Διαβάζοντας τον πίνακα μπορούμε να συνάγουμε το συμπέρασμα ότι το μοντέλο έχει πολύ καλή ακρίβεια στην αναγνώριση της κλάσης “main battle tank” και το background. Όσον αφορά τα κύρια άρματα μάχης (main battle tank - mbt), το μοντέλο έχει 97% σωστή πρόβλεψη, ενώ η εσφαλμένη αναγνώριση ενός mbt ως background, ανέρχεται στο 3 %.



Διάγραμμα 3. Διάγραμμα Ανάκλησης - Αυτοπεποίθησης (Recall - Confidence)

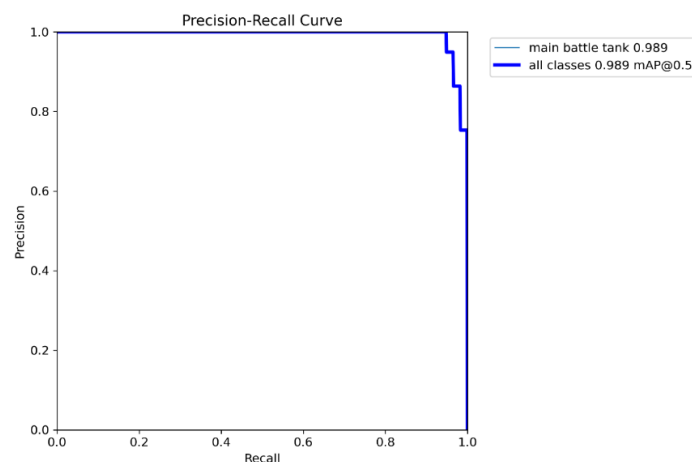
Το παραπάνω διάγραμμα **Recall-Confidence**, αναπαριστά τη σχέση μεταξύ της ανάκλησης και της αυτοπεποίθησης του μοντέλου για την κλάση main battle tank. Η καμπύλη αυτή, αναπαριστά το ποσοστό των πραγματικών θετικών δειγμάτων που ανιχνεύονται σωστά, με την αυτοπεποίθηση που έχει το μοντέλο για τις προβλέψεις του. Παρατηρούμε λοιπόν, ότι στην αρχή της καμπύλης, το recall είναι κοντά στο 1, για χαμηλές τιμές confidence, γεγονός που υποδηλώνει την ικανότητα του μοντέλου να ανιχνεύει σχεδόν όλα τα θετικά δείγματα ακόμα και με χαμηλή εμπιστοσύνη στις

προβλέψεις του. Καθώς η τιμή confidence αυξάνεται, το recall παραμένει υψηλό μέχρι περίπου 0.8 και πέραν αυτού υφίσταται απότομη πτώση, οπότε και το μοντέλο αρχίζει να χάνει κάποια θετικά δείγματα.



Διάγραμμα 4. Διάγραμμα Ακρίβειας - Αυτοπεποίθησης (Precision - Confidence)

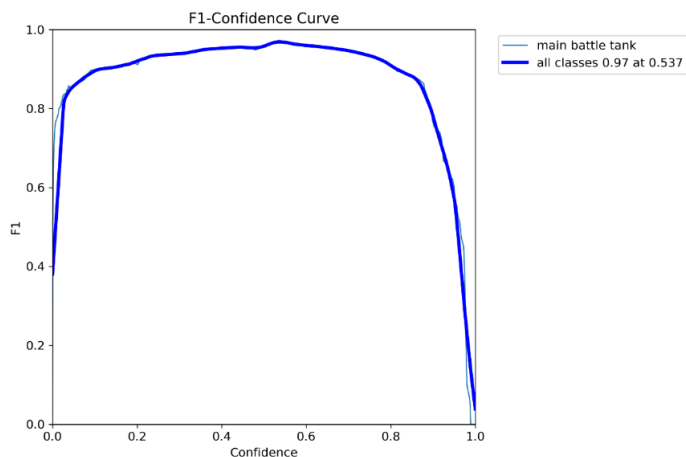
Ακολουθως, το γράφημα **Precision-Confidence**, απεικονίζει τη σχέση μεταξύ της ακρίβειας (precision) και confidence του μοντέλου για την κατηγορία "main battle tank". Αναπαριστά δηλαδή, το ποσοστό των θετικών προβλέψεων του μοντέλου που είναι πράγματι σωστές σε αναλογία με το confidence αυτού. Αρχικά, η ακρίβεια του μοντέλου είναι σχετικά χαμηλή, γύρω στο 0.4, που σημαίνει ότι κάνει πολλές λανθασμένες προβλέψεις. Καθόσον όμως, αυξάνεται το confidence (περίπου από 0.2 και πάνω), το precision ανεβαίνει σχεδόν έως το 1, δηλαδή το μοντέλο γίνεται πιο σίγουρο και σχεδόν όλες οι προβλέψεις του, είναι σωστές.



Διάγραμμα 5. Διάγραμμα Ακρίβειας - Ανάκλησης (Precision - Recall)

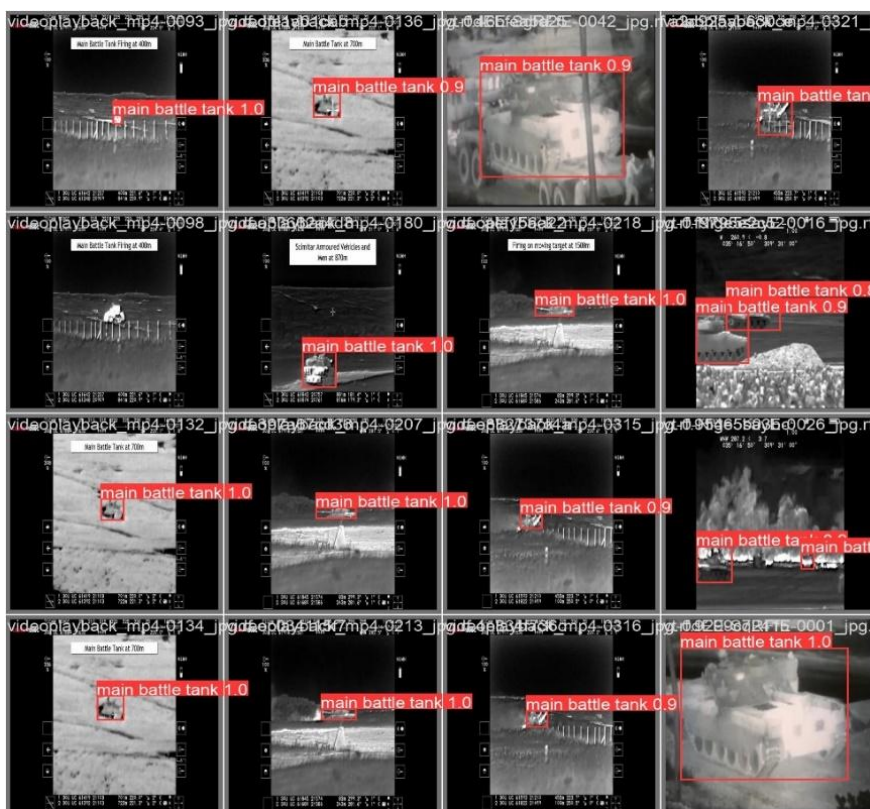
Το γράφημα δείχνει την **Precision-Recall Curve** για την κατηγορία "main battle tank", απεικονίζοντας την σχέσης μεταξύ της precision και recall του μοντέλου. Αντιπροσωπεύει δηλαδή, το ποσοστό των σωστών προβλέψεων από το σύνολο των θετικών προβλέψεων σε αναλογία με το ποσοστό των πραγματικών θετικών που αναγνωρίστηκαν σωστά. Προκύπτει λοιπόν η τιμή της $mAP_{50} = 0.989$, που υποδηλώνει

ότι το μοντέλο έχει πολύ καλή απόδοση, κάνοντας πολύ λίγα λάθη, συνδυάζοντας έτσι, υψηλή ακρίβεια και ανάκληση.



Διάγραμμα 6. Διάγραμμα F1 - Αυτοπεποίθησης (F1 - Confidence)

Το διάγραμμα F1 -Confidence Curve, αναπαριστά τη σχέση μεταξύ του συνολικού δείκτη αξιολόγησης F1 και της confidence του μοντέλου. Αναπαριστά δηλαδή, τη διακύμανση του F1 (συνδυασμός precision και recall), με την confidence του μοντέλου. Το μοντέλο όπως φαίνεται, επιτυγχάνει την καλύτερη απόδοση ($F1 = 0.97$) όταν η confidence βρίσκεται στο 0.537. Παρατηρείται επίσης ότι η καμπύλη έχει σχεδόν επίπεδο τμήμα στο μέσο, γεγονός που δείχνει ότι το μοντέλο διατηρεί μια σταθερή απόδοση σε ένα ευρύ φάσμα τιμών confidence, προτού αρχίσει να μειώνεται δραστικά κοντά στο 1.

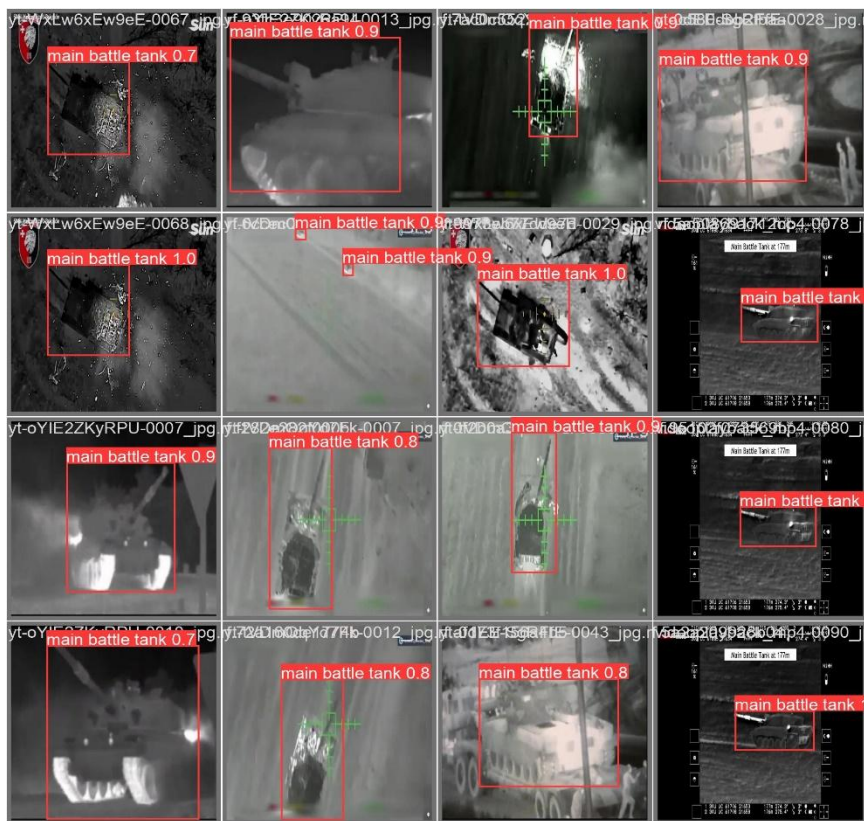


ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



Ενδεικτικά παρουσιάζονται οι προβλέψεις του μοντέλου με εικόνες που αντλήθηκαν από το validation set. Παράλληλα, παρουσιάζεται το confidence score της πρόβλεψης που πραγματοποιεί το μοντέλο, για κάθε διαφορετική εικόνα.





7. Συμπεράσματα – Μελλοντική Εργασία

Συμπερασματικά από τα ανωτέρω, μπορούμε να εξάγουμε το συμπέρασμα ότι το μοντέλο μας έχει ικανοποιητικές αποδόσεις, αναγνωρίζοντας με μεγάλη ακρίβεια τα άρματα – στόχους. Το μοντέλο αποδίδει εξαιρετικά, με **υψηλά Precision, Recall και F1-score**, υποδηλώνοντας ότι είναι ικανό να ανιχνεύει και να ταξινομεί αντικείμενα με μεγάλη ακρίβεια, παρόλο το μικρό μέγεθος του συνόλου δεδομένων. Αυτό που δεν εξετάζεται στην παρούσα εργασία και αποτελεί κομμάτι για μελλοντική μελέτη, είναι το over-fitting (υπερμοντελοποίηση) μοντέλου. Η καλύτερη αξιολόγηση του μοντέλου, δύναται να πραγματοποιηθεί με την τεχνική του cross-validation. Διαμέσου του cross-validation, χωρίζουμε το υπάρχον σύνολο δεδομένων σε μικρότερα και εκπαιδεύουμε το μοντέλο σε καθένα από αυτά, οπότε και εξάγουμε ένα μέσο όρο, η οποία όμως, ενέχει μεγάλο υπολογιστικό κόστος. Επίσης, η περαιτέρω βελτίωση του μοντέλου μπορεί να επιτευχθεί με την παραμετροποίηση του ή την αύξηση των κύκλων εκπαίδευσης, με την προϋπόθεση της διαθεσιμότητας υψηλής ποιότητας εξοπλισμού, προκειμένου να πραγματοποιηθούν πολλαπλά πειράματα.

Όσον αφορά δε το πρώτο τμήμα της εργασίας, καταδείχθηκε επαρκώς η αξία της χρήσης της μεθόδου TDMA για τις συχνότητες VHF. Το TDMA, παρέχει μια πιο δομημένη και αποδοτική χρήση του φάσματος, αποκλείοντας το ενδεχόμενο της ταυτόχρονης μετάδοσης δεδομένων από διαφορετικούς χρήστες. Έτσι, είναι δυνατή η μετάδοση δεδομένων μεταξύ χρηστών και UAV, ώστε να παραμένουν ενημερωμένοι με τα πιο πρόσφατα δεδομένα. Η δυνατότητα αυτή, αποτελεί καίριας σημασίας για το σύγχρονο περιβάλλον πολεμικών επιχειρήσεων, όπου κυριαρχούν η ευμεταβλητότητα και αβεβαιότητα. Παράλληλα, το σύγχρονό του, NTDMA παρέχει έναν συνδυασμό ευελιξίας και απόδοσης, συνδυάζοντας το διαχωρισμό φάσματος και χρόνου, παρέχοντας έτσι αυξημένες δυνατότητες για χρήση σε δορυφορικά στρατιωτικά δίκτυα (MILSATCOM) με υψηλό φορτίο δεδομένων, σε δίκτυα μάχης (Tactical Networks) για άμεση επικοινωνία και ανίχνευση απειλών και μετάδοση πληροφοριών πεδίου.

Παρότι η εργασία αποτελείται από δύο διακριτά τμήματα, δεν νοείται η ύπαρξη ενός αυτόνομου συστήματος ανίχνευσης ή και καταστροφής απειλών, χωρίς τη δυνατότητα επικοινωνίας με έναν επίγειο σταθμό. Η επικοινωνία για τη μεταβίβαση πληροφοριών σχετικά με την φύση ή την κατάσταση των απειλών, είναι σημαίνουσας σημασίας, διότι αποτελεί βοήθημα για την επιτυχή σχεδίαση και εκτέλεση οποιασδήποτε αποστολής. Με τη χρήση UAVs για αναγνώριση απειλών, εν προκειμένω αρμάτων μάχης, δίνεται η δυνατότητα έγκαιρης προειδοποίησης, τόσο των Σταθμών Διοίκησης αλλά και των προπορευόμενων τμημάτων. Μέσω των συχνοτήτων VHF και της μεθόδου TDMA, δίνεται η δυνατότητα του UAV να επικοινωνεί με τους διάφορους επίγειους χρήστες, μεταβιβάζοντας την πληροφορία για τη φύση και θέση των στόχων.

Με την ολοκλήρωση της παρούσας εργασίας, αναγνωρίζεται και η ανάγκη για μελλοντική μελέτη. Ενδιαφέρουσα περίπτωση, αποτελεί και η δυνατότητα χρήσης συνθετικών εικόνων, για την εκπαίδευση του μοντέλου. Η τεχνική αυτή, αποτελεί μια πολλά υποσχόμενη προσέγγιση, ειδικά σε τομείς όπου τα υπάρχοντα δεδομένα είναι περιορισμένα ή διαβαθμισμένα. Στις στρατιωτικές εφαρμογές, επί παραδείγματι, η πρόσβαση σε εικόνες ή βίντεο με τα πραγματικά θερμικά αποτυπώματα μέσων είναι συχνά περιορισμένη, λόγω απορρήτου ή του κόστους συλλογής δεδομένων. Έτσι, η



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



χρήση συνθετικών εικόνων παραγόμενων από λογισμικά προσομοίωσης, επιτρέπει την ανάπτυξη μεγάλων και πληρέστερων συνόλων δεδομένων, τα οποία μπορούν να χρησιμοποιηθούν για την εκπαίδευση αλγορίθμων υπολογιστικής όρασης. Η δυναμική αυτής της μεθόδου, αντικατοπτρίζεται και στη διεθνή βιβλιογραφία, η οποία περιλαμβάνει έρευνες με τη χρήση μοντέλων εκπαιδευμένων σε συνθετικά δεδομένα, τα οποία επιτυγχάνουν υψηλά ποσοστά ακρίβειας σε πραγματικά σενάρια. Καταφέρνουν έτσι, να αναγνωρίζουν θερμικά αποτυπώματα στόχων, όπως άρματα μάχης, αξιοποιώντας την προσαρμοστικότητα και την γενίκευση που παρέχει η εκπαίδευση σε συνθετικά δεδομένα, μειώνοντας το απαιτούμενο υπολογιστικό κόστος και χρόνο εκπαίδευσης.

Παρά τα ενθαρρυντικά αποτελέσματα, η χρήση συνθετικών δεδομένων για εφαρμογές υπολογιστικής όρασης, εγείρει προκλήσεις. Η απόδοση του μοντέλου είναι ανάλογη της αληθοφάνειας των δεδομένων και της ικανότητας του αλγορίθμου να γενικεύσει τη μάθηση σε πραγματικά δεδομένα. Συνεπώς, θεωρείται απαραίτητη η ανάπτυξη μεθόδων που αμβλύνουν το χάσμα μεταξύ συνθετικών και πραγματικών δεδομένων, όπως με την ενσωμάτωση πραγματικών δεδομένων στο στάδιο επαλήθευσης του μοντέλου. Συνολικά ωστόσο, η αξιοποίηση συνθετικών εικόνων στην εκπαίδευση μοντέλων υπολογιστικής όρασης ανοίγει νέους δρόμους για την επίλυση σύνθετων προβλημάτων σε περιβάλλοντα περιορισμένων δεδομένων, ενισχύοντας τις δυνατότητες και την αποτελεσματικότητα των συστημάτων αυτών.



Βιβλιογραφία

- [1] Sawalmeh, Ahmad & Othman, Noor. An Overview of Collision Avoidance Approaches and Network Architecture of Unmanned Aerial Vehicles (UAVs). *International Journal of Engineering & Technology*. 2018.
- [2] Y. Zeng, R. Zhang, and T. J. Lim, “Wireless communications with unmanned aerial vehicles: opportunities and challenges,” *IEEE Communications Magazine*, vol. 54, no. 5, pp. 36–42, (2016)
- [3] Wang, G., & Miao, H. (2021). UAV data link system: A survey. *Scientific Journal of Intelligent Systems Research*, 3(11), 93. ISSN: 2664-9640
- [4] Yeo S C, Kang B W, Bae K H, et al. Study on Data-link Antenna System for UAV[J]. *The Journal of the Korea institute of electronic communication sciences*, 2020, 15(1): 9-14
- [5] Demush, Rostyslav. "A Brief History of Computer Vision (and Convolutional Neural Networks)." *HackerNoon*, 26 Feb. 2019, <https://hackernoon.com/a-brief-history-of-computer-vision-and-convolutional-neural-networks-8fe8aacc79f3>.
- [6] Jarrett, K., K. Kavukcuoglu, M. Ranzato, and Y. LeCun: What is the best multistage architecture for object recognition In *Computer Vision*, 2009 IEEE 12th International Conference on, pages 2146-2153, Sept 2009.
- [7] Krizhevsky, Alex, Ilya Sutskever, and George E. Hinton: Imagenet classification with deep convolutional neural networks. In Bartlett, Peter L., Fernando C. N. Pereira, Christopher J. C. Burges, Leon Bottou, and Kilian Q. Weinberger (editors): *NIPS*, pages 1106-1114, 2012. <http://dblp.uni-trier.de/db/conf/nips/nips2012.html#KrizhevskySH12>.
- [8] Patel, Maahi. "The Complete Guide to Image Preprocessing Techniques in Python." *Medium*, 24 July 2023, <https://medium.com/@maahip1304/the-complete-guide-to-image-preprocessing-techniques-in-python-dca30804550c>. Accessed 19 Aug. 2024.
- [9] W. S. McCulloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” *The bulletin of mathematical biophysics*, vol. 5, pp. 115–133, 1943.
- [10] Lu, L., Shin, Y., Su, Y., & Karniadakis, G. E. (2019). *Dying ReLU and Initialization: Theory and Numerical Examples*. arXiv. <https://doi.org/10.48550/arXiv.1903.06733>
- [11] Online, “towardsdatascience.com/neural-networks-and-theuniversal-approximation-theorem,” <https://towardsdatascience.com/neural-networks-and-the-universal-approximation-theorem-8a389a33d30a>.
- [12] Online, “machinelearningmastery.com/neural-networks-crash-course,” <https://machinelearningmastery.com/neural-networks-crash-course/>.
- [13] Online, “Analyticsvidhya/ comprehensive guide on optimizers on deep learning,” <https://www.analyticsvidhya.com/blog/2021/10/a-comprehensive-guide-on-deep-learning-optimizers/>
- [14] Harris Corporation. *Radio Communications in the Digital Age, Volume Two: VHF/UHF Technology*. RF Communications Division, 2000. Library of Congress Catalog Card Number: 00-132465



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



- [15] Wang, Y., Zhang, J. & Zhou, J. Urban traffic tiny object detection via attention and multi-scale feature driven in UAV-vision. Sci Rep 14, 20614 (2024). <https://doi.org/10.1038/s41598-024-71074-2>
- [16] Y. LeCun, "The mnist database of handwritten digits," <http://yann.lecun.com/exdb/mnist/>, 1998.
- [17] Online, "Ibm/topics/semantic segmentation," <https://www.ibm.com/topics/semantic-segmentation>.
- [18] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2016.
- [19] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," arXiv, 2018.
- [20] Chazhoor, Anisha & Sarobin, Vergin. (2022). Intelligent automation of invoice parsing using computer vision techniques. Multimedia Tools and Applications. 81. 10.1007/s11042-022-12916-x.
- [21] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2014.
- [22] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," Advances in neural information processing systems, vol. 28, 2015.
- [23] Itakura, Kenta & Hosoi, Fumiki. (2020). Automatic Tree Detection from Three-Dimensional Images Reconstructed from 360° Spherical Camera Using YOLO v2. Remote Sensing. 12. 988. 10.3390/rs12060988.
- [24] Yanan Song, Quan-Ke Pan, Liang Gao, Biao Zhang, Improved non-maximum suppression for object detection using harmony search algorithm, Applied Soft Computing, Volume 81, 2019, 105478, ISSN 1568-4946, <https://doi.org/10.1016/j.asoc.2019.05.005>.
- [25] <https://intermatdefense.com/anti-thermal>
- [26] Kyiv Post, "Ukrainian Soldiers Destroy Russian Tank with FPV Drone," September 2024. [Link](#).
- [27] Reuters, "How drone combat in Ukraine is changing warfare," September 2024. [Link](#).
- [28] Ariuntuya, A. (July 19, 2023). *Improve Accuracy: Polygon Annotations for Object Detection*. Roboflow Blog. Retrieved from <https://blog.roboflow.com/polygon-vs-bounding-box-computer-vision-annotation/>
- [29] James Gallagher. (Aug 23, 2023). What is Data Augmentation? The Ultimate Guide.. Roboflow Blog: <https://blog.roboflow.com/data-augmentation/>
- [30] <https://www.androidpolice.com/google-colab-explainer/>
- [31] Chengjian Feng, Yujie Zhong, Yu Gao, Matthew R Scott, and Weilin Huang. Tood: Task-aligned one-stage object detection. In 2021 IEEE/CVF



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ
ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



- International Conference on Computer Vision (ICCV), pages 3490–3499. IEEE Computer Society, 2021.
- [32] Peize Sun, Yi Jiang, Enze Xie, Wenqi Shao, Zehuan Yuan, Changhu Wang, and Ping Luo. What makes for end-to-end object detection? In International Conference on Machine Learning, pages 9934–9944. PMLR, 2021.
- [33] Jianfeng Wang, Lin Song, Zeming Li, Hongbin Sun, Jian Sun, and Nanning Zheng. End-to-end object detection with fully convolutional network. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 15849–15858, 2021.
- [34] Yiqun Chen, Qiang Chen, Qinghao Hu, and Jian Cheng. Date: Dual assignment for end-to-end fully convolutional object detection. arXiv preprint arXiv:2211.13859, 2022.
- [35] Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., & Ding, G. *YOLOv10: Real-Time End-to-End Object Detection*. Tsinghua University. arXiv:2405.14458, 2024
- [36] Goldsmith, A. (2005). *Wireless Communications (Illustrated)*. Cambridge University Press.
- [37] Τσουκάτος, Κ., & Χατζούλης, Δ. (2023). Ασύρματες Επικοινωνίες [Προπτυχιακό εγχειρίδιο]. Κάλλιπος, Ανοικτές Ακαδημαϊκές Εκδόσεις. <http://dx.doi.org/10.57713/kallipos-202>
- [38] All About Electronics. (n.d.). *Frequency Division Multiplexing (FDM) Explained*. Retrieved from <https://www.allaboutelectronics.org/frequency-division-multiplexing-fdm-explained/>
- [39] GeeksforGeeks. (n.d.). *Frequency Division and Time Division Multiplexing*. Retrieved from <https://www.geeksforgeeks.org/frequency-division-and-time-division-multiplexing/>
- [40] Nguyen, Anh & Yosinski, Jason & Clune, Jeff. (2019). Understanding Neural Networks via Feature Visualization: A Survey. 10.1007/978-3-030-28954-6_4.
- [41] Online, “Simplilearn/tutorials/deep-learning-tutorial/perceptron,” [https:// www.simplilearn.com/tutorials/deep-learning-tutorial/perceptron](https://www.simplilearn.com/tutorials/deep-learning-tutorial/perceptron).
- [42] Online, “Anh reynolds/logs/cnn,” <https://anhreynolds.com/blogs/cnn.html>.
- [43] InterviewBit. "Top Deep Learning Interview Questions and Answers." Available at: <https://www.interviewbit.com/deep-learning-interview-questions/>.



ΠΜΣ: ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ

ΜΗ ΣΤΕΛΕΧΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ



- [44] LearnOpenCV. "Intersection over Union (IoU) in Object Detection and Segmentation." Available at: <https://learnopencv.com/intersection-over-union-iou-in-object-detection-and-segmentation/>
- [45] Evidently AI. "Confusion Matrix in Classification Metrics." Available at: <https://www.evidentlyai.com/classification-metrics/confusion-matrix>.
- [46] Eurocontrol. Available at: https://ext.eurocontrol.int/aixm_confluence/pages/viewpage.action?pageId=3087537
- [47] IDST. (2023). Satellite multiple access techniques: TDMA, FDMA, CDMA, and emerging schemes. IDST International Defence, Security & Technology. <https://idstch.com/space/satellite-multiple-access-techniques-tdma-fdma-cdma-and-emerging-schemes>
- [48] Murata. (2024). FDMA, TDMA, and CDMA multiple access: Effective utilization of frequency bandwidth. <https://article.murata.com/en-us/article/multiplexing-and-multiple-access-1>
- [49] Πανεπιστήμιο Αιγαίου. (2015). Ασύρματα Δίκτυα Επικοινωνιών. Πανεπιστήμιο Αιγαίου - Eclass. <https://eclass.aegean.gr/modules/document/file.php/ICSD122>
- [50] Ghosh, Anirudha & Sufian, A. & Sultana, Farhana & Chakrabarti, Amlan & De, Debashis. (2020). Fundamental Concepts of Convolutional Neural Network. 10.1007/978-3-030-32644-9_36.
- [51] Vedoveli, H. (2023) *Metrics matter: A deep dive into object detection evaluation*. Medium. <https://medium.com/@henriquevedoveli/metrics-matter-a-deep-dive-into-object-detection-evaluation-ef01385ec62>