



ΔΙΕΘΝΕΣ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΤΗΣ ΕΛΛΑΔΟΣ

ΔΙΕΘΝΕΣ ΠΑΝΕΠΙΣΤΗΜΙΟ ΤΗΣ ΕΛΛΑΔΟΣ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ενισχυτική και Βαθιά Μάθηση στον Έλεγχο Μη Επανδρωμένου Αεροχήματος: Τάσεις και Προκλήσεις



Φοιτητής:
Μεμλικάϊ Γεώργιος
Αριθμός Μητρώου: 2019206

Επιβλέπων:
Τσιντώτας Κωνσταντίνος
Επίκουρος Καθηγητής

Φεβρουάριος 2026

Ενισχυτική και Βαθιά Μάθηση στον Έλεγχο Μη Επανδρωμένου Αεροχήματος: Τάσεις και Προκλήσεις

Κωδικός Διατριβής: **25250**

Όνοματεπώνυμο φοιτητή: **Μεμλικαϊ Γεώργιος**

Όνοματεπώνυμο επιβλέποντα καθηγητή: **Τσιντώτας Κωνσταντίνος**

Ημερομηνία ανάληψης: 04-04-2025

Ημερομηνία ολοκλήρωσης: 04-09-2025

Με το παρόν έγγραφο επιβεβαιώνουμε τη συγγραφή του παρόντος κειμένου, καθώς και την αναγνώριση της όποιας βοήθειας λάβαμε στη σύνθεσή του. Επιπλέον, επισημάναμε τις διάφορες πηγές από τις οποίες αντλήσαμε δεδομένα, ιδέες, οπτικό ή γραπτό υλικό, σε παράφραση ή ακριβή παράθεση. Επιπλέον, βεβαιώνουμε την αποκλειστική σύνταξη της παρούσας εργασίας από εμένα και μόνο, με σκοπό την εκπόνησή της ως διπλωματικής εργασίας, στο *Τμήμα Μηχανικών Πληροφορικής και Ηλεκτρονικών Υπολογιστών* του Διεθνούς Πανεπιστημίου της Ελλάδος.

Η παρούσα εργασία αποτελεί πνευματική ιδιοκτησία του *Μεμλικαϊ Γεώργιου*, του φοιτητή που τη σύνταξε. Σύμφωνα με την πολιτική ανοικτής πρόσβασης, ο συγγραφέας/συντάκτης προσφέρει στο Διεθνές Πανεπιστήμιο της Ελλάδος την άδεια χρήσης του δικαιώματος αναπαραγωγής, δανεισμού, δημόσιας παρουσίασης και ψηφιακής διανομής της εργασίας παγκοσμίως, σε ηλεκτρονική μορφή και σε μέσα κάθε είδους, για διδακτικούς ή ερευνητικούς σκοπούς, εθελοντικά. Η ανοικτή πρόσβαση στο πλήρες κείμενο, σε καμία περίπτωση δεν παρέχει το δικαίωμα καταπάτησης της πνευματικής ιδιοκτησίας του συγγραφέα/συντάκτη, ούτε επιτρέπει την αναπαραγωγή, αναδημοσίευση, αναπαραγωγή, πώληση, εμπορική χρήση, διανομή, δημοσίευση, λήψη, μεταφόρτωση, μετάφραση, τροποποίηση κάθε είδους, εν μέρει ή περιληπτικά της εργασίας, χωρίς τη ρητή γραπτή συγκατάθεση των συγγραφέων.

Η έγκριση της παρούσας διατριβής από το *Τμήμα Μηχανικών Πληροφορικής και Ηλεκτρονικών Υπολογιστών* του Διεθνούς Πανεπιστημίου της Ελλάδος, δεν συνεπάγεται απαραίτητα την υιοθέτηση των απόψεων του συγγραφέα, εκ μέρους του Τμήματος.

Αφιέρωση

Αφιερώνω αυτό το έργο στην οικογένειά μου, η οποία μου παρείχε ακλόνητη υποστήριξη με αγάπη, υπομονή και κατανόηση κατά τη διάρκεια του ταξιδιού μου. Χωρίς τη συνεχή ενθάρρυνση και την αδιάκοπη παρουσία τους, η ολοκλήρωση αυτής της προσπάθειας δεν θα ήταν δυνατή. Ακόμη, θα ήθελα να ευχαριστήσω τον επίκουρο καθηγητή Τσιντώτα Κωνσταντίνο για την επίβλεψη της διατριβής μου και τις γνώσεις που μοιράστηκε.

Πρόλογος

Οι πρόσφατες τεχνολογικές καινοτομίες οδήγησαν στην ευρεία χρήση μη επανδρωμένων εναέριων οχημάτων (Unmanned Aerial Vehicles - UAVs), τα οποία χρησιμοποιούνται σήμερα για διάφορους σκοπούς, από την επιτήρηση και τη χαρτογράφηση έως την παράδοση προϊόντων και τις επιχειρήσεις διάσωσης. Παράλληλα, η ταχύτατη εξέλιξη στον τομέα της τεχνητής νοημοσύνης, και ειδικότερα της μηχανικής μάθησης, έχει ανοίξει νέους ορίζοντες στον τρόπο με τον οποίο τα μη επανδρωμένα εναέρια οχήματα μπορούν να λειτουργούν αυτόνομα και ευφυώς.

Σε αυτό το πλαίσιο, ο συνδυασμός της ενισχυτικής μάθησης (Reinforcement Learning - RL) και της βαθιάς μάθησης (Deep Learning - DL) έχει αναδειχθεί ως μια βιώσιμη στρατηγική για τον ευφυή και δυναμικό έλεγχο των αεροχημάτων σε πολύπλοκα και μεταβαλλόμενα περιβάλλοντα. Οι προσεγγίσεις βαθιάς ενισχυτικής μάθησης επιτρέπουν να λαμβάνουν αποφάσεις μέσω της εμπειρίας και να μαθαίνουν αποτελεσματικές τακτικές για την πλοήγηση, την αποφυγή εμποδίων και την εκτέλεση αποστολών, ακόμη και μπροστά σε συνθήκες αβεβαιότητας ή περιορισμών.

Αυτή η διατριβή εξετάζει τις τεχνικές ενισχυτικής και βαθιάς μάθησης που αφορούν τον έλεγχο μη επανδρωμένων εναέριων οχημάτων, με στόχο να αναδείξει τις θεμελιώδεις αρχές, τους χρησιμοποιούμενους αλγορίθμους και τις πρακτικές εφαρμογές. Ταυτόχρονα, δίνει έμφαση στις κύριες δυσκολίες που πρέπει να επιλύσει η έρευνα, όπως η ασφάλεια, η γενίκευση, η αποτελεσματική εκπαίδευση και η προσαρμογή σε πραγματικές συνθήκες.

Η εν λόγω εργασία επιδιώκει να παρουσιάσει ένα πλήρες θεωρητικό υπόβαθρο, να καταγράψει τις σύγχρονες ερευνητικές τάσεις και να συμβάλει στην κατανόηση του ρόλου που μπορεί να παίξει η συνδυασμένη μηχανική μάθηση στον έξυπνο έλεγχο των μη επανδρωμένων αεροχημάτων. Μέσα από βιβλιογραφική ανάλυση, επιχειρείται η αποτύπωση της τρέχουσας κατάστασης στον τομέα καθώς και των μελλοντικών του προοπτικών.

Abstract

This thesis focuses on the use of reinforcement and deep learning techniques to operate unmanned aerial vehicles (UAVs). The usage of UAVs for task execution has grown in popularity across diverse domains in recent years. Ensuring robust autonomy in dynamic environments requires advanced control techniques, among which reinforcement learning (RL) and deep reinforcement learning (DRL) have emerged as promising paradigms. This work provides a systematic review of RL - and DRL - methods applied to UAV control. Particular attention is given to simulation environments that serve as benchmarks, as well as to the gap between simulation and real-world deployment. Despite their advantages, RL-based methods face notable challenges, including high computational demands, dependence on large-scale training data, and limited generalization under environmental variability. This study contributes to UAV control by comparing RL/DRL methodologies, elucidating obstacles and opportunities, and highlighting trends that are likely to shape future research.

Περίληψη

Η παρούσα διπλωματική εργασία επικεντρώνεται στην εφαρμογή τεχνικών βαθιάς και ενισχυτικής μάθησης για τον έλεγχο μη επανδρωμένων αεροχημάτων (Unmanned Aerial Vehicles - UAVs). Η αξιοποίηση μη επανδρωμένων εναέριων οχημάτων για την εκτέλεση εργασιών κερδίζει ολο ένα και μεγαλύτερη απήχηση τα τελευταία χρόνια. Οι μέθοδοι που βασίζονται στη βαθιά ενισχυτική μάθηση (Deep Reinforcement Learning - DRL) που εκπαιδεύονται από χρήσιμα μοτίβα μέσα από μαζικά δεδομένα, παρουσιάζουν αξιοσημείωτα πλεονεκτήματα. Παράλληλα, αναλύονται προκλήσεις όπως οι υψηλές υπολογιστικές απαιτήσεις, η περιορισμένη γενίκευση σε δυναμικά περιβάλλοντα και η ανάγκη για εκτεταμένα δεδομένα εκπαίδευσης, που εμποδίζουν την αποτελεσματικότητά τους. Τέλος, η διατριβή συμβάλλει στον τομέα του ελέγχου των μη επανδρωμένων αεροχημάτων διερευνώντας τις τεχνικές RL (Reinforcement Learning), παρέχοντας πληροφορίες για τις προκλήσεις και τις ευκαιρίες αυτής της προσέγγισης για το συγκεκριμένο πρόβλημα, με στόχο την αποτύπωση της προόδου και των μελλοντικών κατευθύνσεων.

Περιεχόμενα

1	Εισαγωγή	1
1.1	Μη επανδρωμένα εναέρια οχήματα και ο ρόλος του έξυπνου ελέγχου.	1
1.2	Ιστορική αναδρομή των μη επανδρωμένων εναέριων οχημάτων	3
1.3	Παρουσίαση της ενισχυτικής και βαθιάς μάθησης.	5
1.4	Μεθοδολογία ανασκόπησης βιβλιογραφίας	7
1.4.1	Σχεδιασμός της ανασκόπησης	7
1.4.2	Διεξαγωγή της ανασκόπησης	8
1.5	Σκοπός της διατριβής	10
1.6	Δομή της εργασίας	10
2	Θεωρητικό υπόβαθρο	11
2.1	Επισκόπηση της μηχανικής μάθησης	11
2.2	Βασικές αρχές της ενισχυτικής μάθησης	12
2.2.1	Μαρκοβιανή διαδικασία απόφασης	12
2.3	Εισαγωγή στην βαθιά μάθηση	14
2.3.1	Τεχνητά νευρωνικά δίκτυα	15
2.3.2	Νευρωνικά δίκτυα στην εκμάθηση πολιτικών ελέγχου	15
2.4	Βασικοί αλγόριθμοι και μοντέλα ενισχυτικής μάθησης	16
2.4.1	Q-Learning	16
2.4.2	Deep Q-Network	17
2.4.3	Αλγόριθμος βαθιάς ντετερμινιστικής κλίσης πολιτικής	18
2.4.4	Αλγόριθμος βελτιστοποίησης εγγύς πολιτικής	19
2.5	Αρχιτεκτονικά μοντέλα	22
3	Εφαρμογές βαθιάς και ενισχυτικής μάθησης	28
3.1	Αυτόνομη πλοήγηση και αποφυγή εμποδίων	28
3.1.1	Έλεγχος μη επανδρωμένων αεροχημάτων	30
3.1.2	Αποφυγή εμποδίων	31
3.2	Συστήματα προσγείωσης και απογείωσης	31
3.3	Συνεργασία σμηγών	33
3.3.1	Ενισχυτική μάθηση πολλαπλών πρακτόρων	36
3.4	Βελτιστοποίηση ενεργειακής κατανάλωσης	37
3.4.1	Βελτιστοποίηση μεταφοράς ωφέλιμου φορτίου	38
3.5	Περιβάλλοντα προσομοίωσης και εκπαίδευσης	40

4	Σύγκριση και αξιολόγηση των μεθόδων	42
4.1	Ανασκόπηση υφιστάμενων προσεγγίσεων στον έλεγχο μη επανδρωμένων αεροχημάτων	42
4.2	Πλεονεκτήματα και μειονεκτήματα της ενισχυτικής και βαθιάς μάθησης	43
4.3	Σύγκριση με παραδοσιακές μεθόδους ελέγχου	46
4.4	Συνολική αξιολόγηση και σύγκριση μεθόδων	49
4.4.1	Ανάλυση καταλληλότητας μεθόδων ανά σενάριο χρήσης	49
4.4.2	Εφαρμογές και περιορισμοί σε πραγματικά περιβάλλοντα	52
4.5	Ποσοτική αξιολόγηση από βιβλιογραφία	52
5	Μελέτες περίπτωσης και προκλήσεις	54
5.1	Περιπτώσεις χρήσης	55
5.1.1	Μη επανδρωμένα αεροχήματα σταθερού πτερυγίου	55
5.1.2	Συνεργατική λειτουργία αεροχημάτων	56
5.1.3	Λειτουργία μη επανδρωμένων αεροχημάτων σε αστικά περιβάλλοντα	57
5.2	Προκλήσεις και περιορισμοί	59
5.2.1	Αντιμετώπιση αβεβαιότητας και πραγματικών συνθηκών	59
5.2.2	Ενεργειακή αποδοτικότητα και περιορισμοί υλικού	60
5.2.3	Δεδομένα εκπαίδευσης και προσομοίωση	62
6	Σύνοψη και μελλοντικές κατευθύνσεις	64
6.1	Ανακεφαλαίωση βασικών ευρημάτων	64
6.2	Προτάσεις για μελλοντική έρευνα	65
6.2.1	Βελτίωση αλγοριθμικών μεθόδων	66
6.2.2	Εκπαίδευση σε πραγματικό χρόνο και αποδοτικότητα δεδομένων	67
6.2.3	Προσαρμογή σε αβεβαιότητα και δυναμικά περιβάλλοντα	68
6.2.4	Ενεργειακή αποδοτικότητα και υπολογιστικοί περιορισμοί	69
6.2.5	Συνεργασία και σμήνη	69
6.3	Συμπέρασμα	70

Κατάλογος σχημάτων

1.1	Διαφορετικές κατηγορίες μη επανδρωμένων αεροσκαφών βάσει επιχειρησιακών δυνατοτήτων και μεγέθους: α) Micro και Nano UAV για αποστολές μικρής εμβέλειας, β) Tactical UAV (Τακτικά) για επιτήρηση, γ) MALE (Medium Altitude Long Endurance) για αποστολές μακράς διάρκειας και δ) Υβριδικά VTOL που συνδυάζουν κάθετη απογείωση με πτήση σταθερής πτέρυγας [3].	2
1.2	Το Sperry Aerial Torpedo (1917), το οποίο αναγνωρίζεται ως μία από τις πρώτες πειραματικές υλοποιήσεις μη επανδρωμένου αεροχήματος σταθερού πτερυγίου, σχεδιασμένο αρχικά ως αυτόνομη εναέρια τορπίλη [3].	4
1.3	Τα επίπεδα αυτονομίας των μη επανδρωμένων αεροσκαφών. Η κλίμακα ξεκινά από το Επίπεδο 0 (100 % χειροκίνητος έλεγχος) και κλιμακώνεται έως την πλήρη αυτονομία, όπου το σύστημα διαχειρίζεται εξ ολοκλήρου την πλοήγηση, τη λήψη αποφάσεων και την αποφυγή εμποδίων χωρίς ανθρώπινη παρέμβαση [11].	5
1.4	Διάγραμμα διαδικασίας λήψης αποφάσεων της ενισχυτικής μάθησης [6].	6
1.5	Διάγραμμα ροής της βιβλιογραφικής ανασκόπησης.	8
2.1	Γραφική αναπαράσταση μιας Μαρκοβιανής Διαδικασίας Απόφασης (MDP). Απεικονίζονται οι καταστάσεις (μπλέ κόμβοι), οι πιθανές ενέργειες (πράσινοι κόμβοι) και οι ανταμοιβές υποδεικνύονται από τα κυματιστά κόκκινα βέλη. Τα μαύρα βέλη και οι αριθμοί δίπλα τους περιγράφουν τις πιθανότητες μετάβασης [18]. . . .	13
2.2	Αρχιτεκτονική ενός Πολυεπίπεδου Perceptron (MLP) με ένα κρυφό στρώμα. Το δίκτυο δέχεται n χαρακτηριστικά εισόδου, τα επεξεργάζεται μέσω των βαρέων συνδέσεων στο κρυφό επίπεδο για την εκμάθηση μη γραμμικών σχέσεων και παράγει j τιμές εξόδου [12].	14
2.3	Διάγραμμα ροής αλγορίθμου βελτιστοποίησης εγγύς πολιτικής (Proximal Policy Optimization - PPO). Απεικονίζεται η αρχιτεκτονική Actor-Critic, η διαδικασία συλλογής εμπειριών (replay buffer) και ο μηχανισμός αποκοπής (clipping) στη συνάρτηση αντικειμενικού σκοπού για τη σταθεροποίηση της εκπαίδευσης [20], [21].	19
2.4	Ροή ελέγχου με την χρήση βαθιάς ενισχυτικής μάθησης: η αντίληψη του περιβάλλοντος (Perception) τροφοδοτεί την υψηλού επιπέδου πολιτική (Policy), η οποία παράγει εντολές για τον έλεγχο χαμηλού επιπέδου (PID/LQR). Το διάγραμμα συνδέει θεωρητικούς αλγορίθμους με την πρακτική εφαρμογή τους. [21]	22
2.5	Κύκλος εκπαίδευσης ενός πράκτορα βαθιάς ενισχυτικής μάθησης. Οι εμπειρίες αποθηκεύονται στο replay buffer και χρησιμοποιούνται για την εκτίμηση πλεονεκτημάτων (GAE), και την ενημέρωση των δικτύων πολιτικής/αξίας, όπως σε αλγόριθμους PPO με GAE[21].	22

2.6	Βασική αρχιτεκτονική ενός Συνελκτικού Νευρωνικού Δικτύου (CNN). Η διαδικασία ξεκινά με την εικόνα εισόδου, ακολουθείται από επίπεδα συνέλιξης (convolution) για την αυτόματη εξαγωγή χωρικών χαρακτηριστικών, συνεχίζεται με τα επίπεδα μέγιστης δειγματοληψίας (max pooling) για τη μείωση της διάστασης και καταλήγει σε πλήρως συνδεδεμένα επίπεδα για την τελική απόφαση [12].	23
2.7	Βασική αρχιτεκτονική αναδρομικού νευρωνικού δικτύου (RNN). Οι είσοδοι (μωβ κόμβοι) τροφοδοτούν τους κρυφούς νευρώνες (γκρι κόμβοι), οι οποίοι διαθέτουν αναδρομικές συνδέσεις (βρόχους). Αυτή η δομή επιτρέπει στο δίκτυο να διατηρεί «μνήμη» προηγούμενων καταστάσεων, καθιστώντας το κατάλληλο για την επεξεργασία χρονικών ακολουθιών [12].	24
2.8	Η διαδικασία ξεδιπλώματος (unfolding) ενός Αναδρομικού Νευρωνικού Δικτύου στον χρόνο. Αριστερά φαίνεται η συνοπτική αναπαράσταση με τον βρόχο ανατροφοδότησης, ενώ δεξιά η ανεπτυγμένη μορφή όπου η κρυφή κατάσταση h_t εξαρτάται από την τρέχουσα είσοδο x_t και την προηγούμενη κατάσταση h_{t-1} , επιτρέποντας τη μάθηση ακολουθιών [12].	25
3.1	Αρχιτεκτονική βαθιάς ενισχυτικής μάθησης τύπου end-to-end για τον έλεγχο μη επανδρωμένου αεροχήματος. Το σύστημα επεξεργάζεται ακατέργαστα δεδομένα αισθητήρων υψηλής διάστασης (π.χ. εικόνες) μέσω ενός Βαθέος Νευρωνικού Δικτύου (DNN) για την άμεση παραγωγή εντολών ελέγχου, παρακάμπτοντας τη χειροκίνητη εξαγωγή χαρακτηριστικών [4].	29
3.2	Διαφορετικοί τύποι μη επανδρωμένων αεροχημάτων: (1) Πολυρότορο, (2) Σταθερού πτερυγίου, (3) Υβριδικό VTOL και (4) Μονορότορο, όπως ταξινομούνται βάσει της μηχανικής τους διαμόρφωσης [3].	30
3.3	Ιεραρχική αρχιτεκτονική διαχείρισης σμήνους, απεικονίζοντας τη ροή πληροφορίας από τον σχεδιασμό αποστολής και τροχιάς έως τον έλεγχο σχηματισμού και το δίκτυο επικοινωνίας [25], [26].	34
3.4	Σχηματική ταξινόμηση των στρατηγικών αντιμετώπισης των ενεργειακών περιορισμών στα μη επανδρωμένα αεροχήματα, συμπεριλαμβανομένων των τεχνικών σχεδιασμού/αντικατάστασης, της εκμετάλλευσης ηλιακής ενέργειας και της ασύρματης φόρτισης [27].	38
5.1	Ταξινόμηση της δομής των εφαρμογών UAV στη διαχείριση αστικού περιβάλλοντος. Οι εφαρμογές κατηγοριοποιούνται σε τρεις κύριους τομείς: (α) Έλεγχος Υποδομών (π.χ. τρισδιάστατη χαρτογράφηση), (β) Αντιμετώπιση Καταστροφών (π.χ. πλημμύρες, σεισμοί) και (γ) Περιβαλλοντικός Έλεγχος (π.χ. ατμοσφαιρική ρύπανση), αναδεικνύοντας τον ρόλο τους στις Έξυπνες Πόλεις [3].	58

Κατάλογος πινάκων

1.1	Κριτήρια ένταξης και αποκλεισμού στη συστηματική ανασκόπηση.	9
2.1	Συγκριτική απεικόνιση των κυριότερων αλγορίθμων ενισχυτικής μάθησης (Q-Learning, DQN, DDPG, PPO) ως προς την καταλληλότητά τους για εφαρμογές UAV, τη σταθερότητα εκπαίδευσης και τις υπολογιστικές τους απαιτήσεις.	21
2.2	Συγκριτική ανάλυση των βασικών αρχιτεκτονικών Βαθιάς Μάθησης (CNN, RNN, Transformers) που εφαρμόζονται σε UAVs, με εστίαση στα λειτουργικά τους χαρακτηριστικά και τους υπολογιστικούς περιορισμούς.	26
3.1	Συγκριτική ταξινόμηση αλγορίθμων μηχανικής και ενισχυτικής μάθησης ανά πεδίο εφαρμογής σε μη επανδρωμένα αεροχήματα, με έμφαση στη λειτουργική ετοιμότητα σε πραγματικά σενάρια.	39
3.2	Συγκριτική επισκόπηση περιβαλλόντων προσομοίωσης και συνόλων δεδομένων (datasets) που χρησιμοποιούνται στην εκπαίδευση UAV, εστιάζοντας στις παραμέτρους που διασφαλίζουν την πειραματική αναπαραγωγικότητα.	41
4.1	Συνοπτική συγχώνευση πτυχών ενισχυτικής και βαθιάς μάθησης στον έλεγχο μη επανδρωμένων αεροσκαφών.	45
4.2	Σύγκριση βαθιάς ενισχυτικής μάθησης και παραδοσιακών μεθόδων ελέγχου.	48
4.3	Ενδεικτικά σενάρια ελέγχου και κατάλληλες προσεγγίσεις.	51
4.4	Μετρικές για την πλοήγηση μη επανδρωμένων αεροχημάτων.	53
4.5	Μετρικές των μη επανδρωμένων αεροχημάτων για την αποφυγή εμποδίων.	53
4.6	Μετρικές των μη επανδρωμένων αεροχημάτων για προσγείωση.	53
4.7	Μετρικές σμηνών (Swarms) μη επανδρωμένων αεροχημάτων.	53
5.1	Σύνοψη των κυριότερων προκλήσεων κατά την ανάπτυξη συστημάτων ελέγχου UAV, οι προτεινόμενες μέθοδοι αντιμετώπισης και παραδείγματα εφαρμογής [7].	62

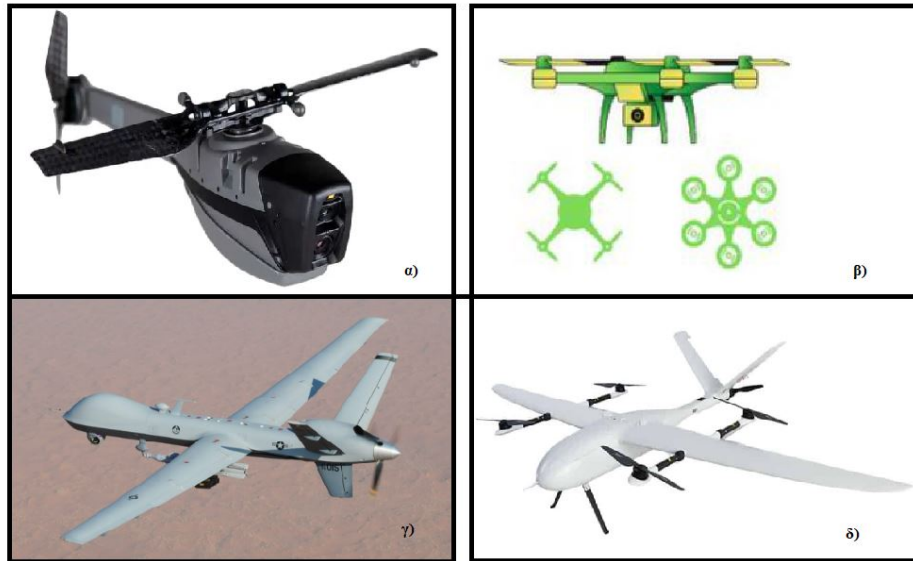
Κεφάλαιο 1

Εισαγωγή

1.1 Μη επανδρωμένα εναέρια οχήματα και ο ρόλος του έξυπνου ελέγχου.

Τα μη επανδρωμένα εναέρια οχήματα (Unmanned Aerial Vehicles – UAVs) αποτελούν σήμερα έναν από τους πλέον δυναμικά αναπτυσσόμενους κλάδους της αεροναυτικής και των αυτόνομων συστημάτων [1] [2]. Οι ραγδαίες εξελίξεις στην υπολογιστική όραση, την ασύρματη επικοινωνία, τη μικροηλεκτρονική και, κυρίως, τις τεχνολογίες τεχνητής νοημοσύνης (Artificial Inteligence - AI) σχετίζονται άμεσα με την ανάπτυξή τους [1]. Τα UAVs, είναι αεροσκάφη που πετούν χωρίς την παρουσία ανθρώπινου πιλότου και ελέγχονται εξ αποστάσεως από έναν χειριστή, αυτόνομα ή από προ-προγραμματισμένους αλγορίθμους, και χωρίζονται σε διάφορες ποικιλίες με βάση τον σκοπό και την αποστολή τους [2]. Τα τελευταία χρόνια έχουν αποκτήσει σημαντική αξία και βαρύτητα, λόγω της ευελιξίας τους και των προοπτικών τους για χρήση σε ένα ευρύ φάσμα εφαρμογών. Χάρη στα προαναφερθέντα πλεονεκτήματα, τα UAV χρησιμοποιούνται ευρέως σε διάφορα σενάρια εφαρμογών, όπως η υποστήριξη ασύρματης επικοινωνίας, η έρευνα και διάσωση, ο στρατός, η γεωργία, η επιτήρηση και άλλα. Μερικά απο αυτά είναι :

- **Micro και Nano UAVs:** Πολύ μικρά UAVs που χρησιμοποιούνται σε περιορισμένα περιβάλλοντα [3].
- **Tactical UAVs:** Μεσαίου μεγέθους UAVs που χρησιμοποιούνται κυρίως για στρατιωτικές επιχειρήσεις και επιτήρηση [3].
- **MALE (Medium Altitude, Long Endurance) και HALE (High Altitude, Long Endurance) UAVs:** Μεγαλύτερα UAVs που επιχειρούν σε υψηλά υψόμετρα και για μεγάλες χρονικές περιόδους [3].
- **Hybrid και Vertical Take-Off and Landing (VTOL) UAVs:** UAVs με δυνατότητα κάθετης απογείωσης και προσγείωσης, κατάλληλα για επιχειρήσεις σε περιορισμένους χώρους [3].



Σχήμα 1.1: Διαφορετικές κατηγορίες μη επανδρωμένων αεροσκαφών βάσει επιχειρησιακών δυνατοτήτων και μεγέθους: α) Micro και Nano UAV για αποστολές μικρής εμβέλειας, β) Tactical UAV (Τακτικά) για επιτήρηση, γ) MALE (Medium Altitude Long Endurance) για αποστολές μακράς διάρκειας και δ) Υβριδικά VTOL που συνδυάζουν κάθετη απογείωση με πτήση σταθερής πτέρυγας [3].

Στο Σχήμα 1.1 διακρίνονται οι τέσσερις βασικές κατηγορίες. Τα micro και nano UAVs (σχ. α) είναι μικρά, ελαφριά και ευκίνητα αεροχήματα που προορίζονται κυρίως για αποστολές αναγνώρισης σε εσωτερικούς χώρους ή περιορισμένους χώρους. Οι μεσαίας κλίμακας στρατιωτικές και πολιτικές εφαρμογές χρησιμοποιούν tactical UAVs (σχ. β), τα οποία προσφέρουν ισορροπία μεταξύ αυτονομίας, βάρους και ωφέλιμου φορτίου. Λόγω της εξαιρετικής τους ικανότητας να λειτουργούν σε μεγάλα υψόμετρα για παρατεταμένες χρονικές περιόδους, τα UAV MALE (Medium Altitude Long Endurance) και HALE (High Altitude Long Endurance) (σχ. γ) είναι ιδανικά για στρατηγική αναγνώριση και επιτήρηση. Τέλος, τα υβριδικά UAV VTOL (Vertical Take-Off and Landing) (σχ. δ) συνδυάζουν τα πλεονεκτήματα των πολυρότορων με τα αεροσκάφη σταθερού πτερυγίου, επιτρέποντας κάθετη απογείωση/προσγείωση και αποτελεσματική πτήση μεγάλων αποστάσεων.

Ο έλεγχος και η λήψη αποφάσεων σε πραγματικό χρόνο αποτελούν τις κύριες λειτουργικές προκλήσεις των μη επανδρωμένων αεροχημάτων. Εξαρτώνται από υπολογιστικά συστήματα που πρέπει να επιτυγχάνουν σταθερότητα, πλοήγηση, αποφυγή εμποδίων και βελτιστοποίηση αποστολών με ελάχιστη ή καθόλου ανθρώπινη παρέμβαση, σε αντίθεση με τα επανδρωμένα αεροσκάφη, όπου ο πιλότος λειτουργεί ως ο κύριος ελεγκτής [4]. Τα συμβατικά συστήματα ελέγχου έχουν χαμηλή απόδοση σε δυναμικά και απρόβλεπτα περιβάλλοντα, αλλά έχουν προσφέρει σημαντικές λύσεις σε σταθερές και γνωστές συνθήκες. Παράγοντες όπως οι ατμοσφαιρικές συνθήκες, οι μη γραμμικότητες της αεροδυναμικής, οι διαταραχές αισθητήρων και οι μεταβολές στο φορτίο απαιτούν μεθόδους ελέγχου που ξεπερνούν τις κλασικές προσεγγίσεις [5][6]. Σε αυτό το σημείο εισέρχεται ο ρόλος του έξυπνου ελέγχου (intelligent control) [7]. Αυτή η φράση περιγράφει προσεγγίσεις που χρησιμοποιούν τεχνικές βαθιάς μάθησης (Deep Learning - DL), ενισχυτικής μάθησης (Reinforcement Learning - RL) και μηχανικής μάθησης (Machine Learning - ML), ώστε να παρέχουν στο UAV

τη δυνατότητα να λαμβάνει αποφάσεις από μόνο του και να προσαρμόζεται [1]. Προκειμένου να παρέχεται ταυτόχρονα σταθερότητα, ασφάλεια και βέλτιστη απόδοση, ο έξυπνος έλεγχος χρησιμοποιείται συχνά σε υβριδικές αρχιτεκτονικές [7].

Η χρήση τεχνικών βαθιάς μάθησης επιτρέπει στο αεροσκάφος να επεξεργάζεται μεγάλους όγκους δεδομένων από πολλαπλούς αισθητήρες, συμπεριλαμβανομένων καμερών, αισθητήρων μέτρησης απόστασης μέσω λέιζερ (Light Detection and Ranging - LiDAR) και μονάδων αδρανειακής μέτρησης (Inertial Measurement Unit - IMU), προσφέροντάς του ανθρώπινες ικανότητες αντίληψης του περιβάλλοντος [1]. Αντίθετα, η ενισχυτική μάθηση προσφέρει ένα πλαίσιο δοκιμών και σφαλμάτων (trial-and-error) που επιτρέπει στο μη επανδρωμένο αεροσκάφος να βελτιώνει συνεχώς τη συμπεριφορά του, βελτιστοποιώντας μια συνάρτηση ανταμοιβής (reward function) που σχετίζεται με την ολοκλήρωση της αποστολής [8]. Όταν συνδυάζονται αυτές οι μέθοδοι, μπορούν να δημιουργηθούν αυτόνομα συστήματα πλοήγησης που μπορούν να προσαρμοστούν σε μεταβαλλόμενες συνθήκες χωρίς ανθρώπινη βοήθεια. Οι καταστάσεις υψηλής πολυπλοκότητας αναδεικνύουν τη σημασία του έξυπνου ελέγχου [5]. Ένα αεροσκάφος που λειτουργεί σε αστικό περιβάλλον, για παράδειγμα [2], πρέπει να αποφεύγει τα εμπόδια, να αντέχει σε ισχυρούς ανέμους και να ακολουθεί τις οδηγίες ασφαλείας. Τα συστήματα που βασίζονται στη μάθηση μπορούν να «γενικεύσουν» τις νέες συνθήκες χρησιμοποιώντας προσομοιώσεις και προϋπάρχουσα εμπειρία, αλλά οι παραδοσιακοί ελεγκτές δυσκολεύονται να χειριστούν τέτοιες δυναμικές [9].

Σε κάθε περίπτωση, ο έξυπνος έλεγχος αναδεικνύεται σε καθοριστικό παράγοντα της τεχνικής ανάπτυξης. Ένα νέο επιχειρησιακό πρότυπο αναδύεται ως αποτέλεσμα του συνδυασμού σύγχρονων τεχνικών τεχνητής νοημοσύνης και συμβατικών μεθόδων επιθεώρησης, μεταβαίνοντας από εργαλεία που λειτουργούν εξ αποστάσεως σε πλήρως αυτόνομα εναέρια συστήματα. Αυτή η μεταβολή δεν περιλαμβάνει μόνο τη βελτίωση των επιχειρησιακών δεξιοτήτων, αλλά και την ασφάλεια, την ενεργειακή απόδοση, την αξιοπιστία και τη συνεργατική νοημοσύνη. Επομένως, ο έξυπνος έλεγχος δεν έχει απλώς υποστηρικτική λειτουργία, αλλά αποτελεί το θεμέλιο για τις πιθανές και μελλοντικές χρήσεις των μη επανδρωμένων αεροσκαφών [9].

1.2 Ιστορική αναδρομή των μη επανδρωμένων εναέριων οχημάτων

Η ανάπτυξη των μη επανδρωμένων εναέριων οχημάτων έχει τις ρίζες της στα πρώτα χρόνια της αεροπορίας. Όταν έκαναν το ντεμπούτο τους στις αρχές της δεκαετίας του 1900, χρησιμοποιούνταν κυρίως για στρατιωτικά καθήκοντα όπως η ανίχνευση και οι αποστολές παρακολούθησης [3]. Κατά τη διάρκεια του πρώτου Παγκοσμίου Πολέμου, το πρώτο αυτόνομο αεροσκάφος [10] στις Ηνωμένες Πολιτείες έθεσε τα θεμέλια για τη συνεχή ανάπτυξη των αεροχημάτων.



Σχήμα 1.2: Το Sperry Aerial Torpedo (1917), το οποίο αναγνωρίζεται ως μία από τις πρώτες πειραματικές υλοποιήσεις μη επανδρωμένου αεροχήματος σταθερού πτερυγίου, σχεδιασμένο αρχικά ως αυτόνομη εναέρια торπίλη [3].

Μετά τον δεύτερο Παγκόσμιο Πόλεμο, παρουσιάστηκαν τα πρώτα μη επανδρωμένα αεροσκάφη και αναπτύχθηκαν τα συστήματα GPS και αισθητήρων, σηματοδοτώντας σημαντική πρόοδο στην τεχνολογία των μη επανδρωμένων αεροσκαφών. Στο πλαίσιο του Ψυχρού Πολέμου, χρησιμοποιήθηκαν εκτενώς για στρατιωτική κατασκοπεία και συλλογή πληροφοριών.

Στις αρχές του 21ου αιώνα, η ανάπτυξη τους γνώρισε εκρηκτική άνοδο καθώς άρχισαν να αποκτούν ευρύτερη εμπορική και πολιτική χρήση. Φτάνοντας στο σήμερα, με την εξέλιξη την τεχνητής νοημοσύνης, εξακολουθούν να αναπτύσσονται καθιστώντας δυνατή τη δημιουργία αυτόνομων συστημάτων που είναι ιδιαίτερα αποτελεσματικά και ευέλικτα σε δύσκολες καταστάσεις [2].

Το Σχήμα 1.3 απεικονίζει τα πέντε επίπεδα αυτονομίας των μη επανδρωμένων αεροσκαφών, τα οποία κατηγοριοποιούνται ανάλογα με τον βαθμό μηχανικής και ανθρώπινης αλληλεπίδρασης, καθώς και με την ικανότητα τους να αναγνωρίζουν και να αποφεύγουν εμπόδια [3]. Στο Επίπεδο 0, το αεροσκάφος δεν έχει αυτονομία και λειτουργεί εξ ολοκλήρου χειροκίνητα. Στο Επίπεδο 1, εξακολουθεί να ελέγχεται άμεσα από τον χρήστη, αλλά διαθέτει τουλάχιστον μία υποβοηθούμενη λειτουργία. Σε ορισμένες περιπτώσεις, το Επίπεδο 2 εισάγει μερική ανάληψη της πλοήγησης, αν και ο οδηγός εξακολουθεί να είναι υπεύθυνος για την ασφαλή λειτουργία του οχήματος. Στο επίπεδο 3, έχει την ικανότητα να αποφεύγει εμπόδια και να εκτελεί εργασίες ανεξάρτητα εντός προκαθορισμένων παραμέτρων, ενώ ένα άτομο παρέχει βοήθεια. Η πλήρης αυτονομία του και οι εφεδρικοί μηχανισμοί που εγγυώνται τη συνέχιση της λειτουργίας του ακόμη και σε περίπτωση μερικής βλάβης είναι χαρακτηριστικά του επιπέδου 4. Τα αεροσκάφη που χρησιμοποιούν μεθόδους τεχνητής νοημοσύνης και μηχανικής μάθησης είναι τα πιο αυτόνομα. Μπορούν να προσαρμόζονται στο περιβάλλον τους και να προγραμματίζονται χωρίς ανθρώπινη βοήθεια.

Τα 5 επίπεδα αυτονομίας των εναέριων οχημάτων

Επίπεδο αυτονομίας	Επίπεδο 0	Επίπεδο 1	Επίπεδο 2	Επίπεδο 3	Επίπεδο 4	Πλήρης αυτονομία
Ανθρώπινη παρέμβαση						
Μηχανική παρέμβαση						
Βαθμός αυτονομίας	Καμία αυτονομία	Χοιμνή αυτονομία	Μέσαλα αυτονομία	Υψηλή αυτονομία	Πλήρης αυτονομία	Πλήρης αυτονομία
Περιγραφή	Ο χειριστής του είναι 100% χειροκίνητος	Ο χειριστής παραμένει υπό τον έλεγχο. Το drone έχει τουλάχιστον μια λειτουργία υπό τον έλεγχο του.	Ο χειριστής παραμένει υπεύθυνος για την ασφαλή λειτουργία. Το drone μπορεί να αναλάβει την πορεία ,το υψόμετρο υπό συνθήκες	Ο χειριστής ενεργεί ως σύστημα εφεδρείας. Το drone μπορεί να εκτελέσει ενέργειες υπό ορισμένες συνθήκες.	Ο χειριστής είναι εκτός λούπας. Το drone διαθέτει εφεδρικά συστήματα, έτσι ώστε εάν ένα από αυτά αποτύχει, η πλαιφόρμα να εξακολουθεί να είναι λειτουργική.	Τα drones θα είναι σε θέση να χρησιμοποιούν εργαλεία τεχνητής νοημοσύνης για τον προγραμματισμό των πτήσεων τους ως αυτόνομα συστήματα μάθησης.
Αποφυγή εμποδίων	ΚΑΜΙΑ	ΑΝΤΙΛΑΜΒΑΝΟΝΤΑΙ ΚΑΙ ΠΡΟΕΙΔΟΠΟΙΟΥΝ	ΑΝΤΙΛΑΜΒΑΝΟΝΤΑΙ ΚΑΙ ΑΠΟΦΕΥΓΟΥΝ	ΑΝΤΙΛΑΜΒΑΝΟΝΤΑΙ ΚΑΙ ΠΛΗΘΙΝΟΥΝΤΑΙ	ΑΝΤΙΛΑΜΒΑΝΟΝΤΑΙ ΚΑΙ ΠΛΗΘΙΝΟΥΝΤΑΙ	ΑΝΤΙΛΑΜΒΑΝΟΝΤΑΙ ΚΑΙ ΠΛΗΘΙΝΟΥΝΤΑΙ

Σχήμα 1.3: Τα επίπεδα αυτονομίας των μη επανδρωμένων αεροσκαφών. Η κλίμακα ξεκινά από το Επίπεδο 0 (100 % χειροκίνητος έλεγχος) και κλιμακώνεται έως την πλήρη αυτονομία, όπου το σύστημα διαχειρίζεται εξ ολοκλήρου την πλοήγηση, τη λήψη αποφάσεων και την αποφυγή εμποδίων χωρίς ανθρώπινη παρέμβαση [11].

1.3 Παρουσίαση της ενισχυτικής και βαθιάς μάθησης.

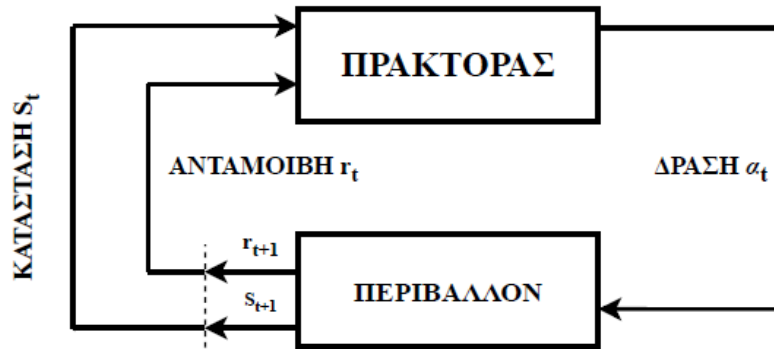
Ο συνδυασμός των πλεονεκτημάτων της ευφυούς λήψης αποφάσεων με τα νευρωνικά δίκτυα που βασίζονται σε δεδομένα, έχει πυροδοτήσει μια εξελικτική αλλαγή που ξεπερνά τα συμβατικά όρια. Τα τελευταία χρόνια, με την συγχώνευση δύο ισχυρών τομέων της τεχνητής νοημοσύνης: των βαθιών νευρωνικών δικτύων και της ενισχυτικής μάθησης, έχει παρατηρηθεί μια εξελικτική αλλαγή που ξεπερνά τα συμβατικά όρια σε ένα ευρύ φάσμα τομέων.

Η ενισχυτική μάθηση, η οποία επικεντρώνεται στη μάθηση μέσω της αλληλεπίδρασης με το περιβάλλον, είναι μια από τις πιο δυναμικές και δύσκολες τεχνικές της μηχανικής μάθησης. Σε αντίθεση με τις κλασικές μεθόδους επιβλεπόμενης μάθησης, οι οποίες χρησιμοποιούν προκαθορισμένα δεδομένα εισόδου και εξόδου για την εκπαίδευση του μοντέλου, επιτρέπει στους πράκτορες να μαθαίνουν κάνοντας λάθη με σκοπό τη μεγιστοποίηση της μακροπρόθεσμης ανταμοιβής[1].

Τα βασικά στοιχεία είναι :

- **Πράκτορας (Agent):** Ο εκπαιδευόμενος και κατά συνέπεια ο φορέας λήψης αποφάσεων.
- **Περιβάλλον (Environment):** Το σύστημα με το οποίο αλληλεπιδρά ο πράκτορας.
- **Συνάρτηση Ανταμοιβής (Reward Function):** Ορίζει τον στόχο του πράκτορα παρέχοντας την τιμή που καθορίζει την επιτυχία μιας δράσης.
- **Συνάρτηση Αξίας (Value Function):** Εκτιμά τη μακροπρόθεσμη ανταμοιβή από μια κατάσταση ή ενέργεια.

- **Πολιτική (Policy):** Ο κανόνας που καθορίζει ποιες δράσεις θα επιλεγούν σε κάθε κατάσταση.
- **Μοντέλο (Model):** Περιγράφει τη δυναμική του περιβάλλοντος, προβλέποντας την επόμενη κατάσταση και ανταμοιβή.
- **Κατάσταση (State):** Η αναπαράσταση της παρούσας κατάστασης του περιβάλλοντος.



Σχήμα 1.4: Διάγραμμα διαδικασίας λήψης αποφάσεων της ενισχυτικής μάθησης [6].

Ο βασικός μηχανισμός της ενισχυτικής μάθησης απεικονίζεται στο διάγραμμα 1.4, όπου ένας πράκτορας και το περιβάλλον του εμπλέκονται σε έναν επαναληπτικό βρόχο. Ο πράκτορας παρακολουθεί τις τρέχουσες συνθήκες S_t του περιβάλλοντος και επιλέγει την ενέργεια A_t που θεωρεί πιο κατάλληλη με βάση την πολιτική $\pi(a | s)$. Το περιβάλλον μεταβάλλεται από αυτή τη δράση και στη συνέχεια επιστρέφει στην αρχική του κατάσταση S_{t+1} και στην αντίστοιχη ανταμοιβή R_{t+1} . Λειτουργώντας ως μηχανισμός ανατροφοδότησης, η ανταμοιβή επιτρέπει στον πράκτορα να αξιολογήσει πόσο καλά λειτουργούν οι ενέργειές του και να προσαρμόσει τη στρατηγική του για να μεγιστοποιήσει την αθροιστική του απόδοση. Αυτός ο κύκλος συνεχίζεται μέχρις ότου ο πράκτορας ανακαλύψει μια πολιτική που οδηγεί στην καλύτερη δυνατή συμπεριφορά για το συγκεκριμένο περιβάλλον.

Αντιθέτως, η βαθιά μάθηση επικεντρώνεται στην εκμάθηση περίπλοκων χαρακτηριστικών από τεράστια σύνολα δεδομένων με τη χρήση πολυεπίπεδων νευρωνικών δικτύων [12]. Τα βαθιά νευρωνικά δίκτυα αναπαράγουν τον τρόπο λειτουργίας του ανθρώπινου εγκεφάλου, κάτι που τα καθιστά ιδιαίτερα αποτελεσματικά στην αναγνώριση πολύπλοκων προτύπων και σχέσεων στα δεδομένα, όπως εικόνες, ήχοι και χρονικές ακολουθίες. Ο συνδυασμός των δύο προσεγγίσεων, γνωστός ως βαθιά ενισχυτική μάθηση (Deep Reinforcement Learning - DRL), αξιοποιεί τη δύναμη των βαθιών δικτύων για τη διαχείριση προβλημάτων με υψηλή διάσταση.

Η εφαρμογή της βαθιάς ενισχυτικής μάθησης έχει οδηγήσει σε σημαντικές καινοτομίες, παρέχοντας την δυνατότητα εξελιγμένων τύπων αυτόνομης λειτουργίας, που είναι αδύνατο να εκτελεστούν με τη χρήση παραδοσιακών μεθόδων. Στην παρούσα μελέτη, διεξάγεται διερεύνηση της χρήσης των παραπάνω μεθόδων μάθησης για τον έλεγχο τους. Με την χρήση της βαθιάς ενισχυτικής μάθησης, έχουν καταστεί σημαντικές εξελίξεις καθιστώντας δυνατές προηγμένες μορφές αυτόνομης συμπεριφοράς που δύσκολα υλοποιούνται με παραδοσιακούς αλγορίθμους. Ορισμένες από τις βασικές εφαρμογές περιλαμβάνουν :

- **Αυτόνομη πλοήγηση:** χωρίς εξωτερική βοήθεια ή προ-προγραμματισμένες διαδρομές, η βαθιά ενισχυτική μάθηση εκπαιδεύει τα μη επανδρωμένα αεροσκάφη να πλοηγούνται σε νέες περιοχές, να αποφεύγουν τα εμπόδια και να φτάνουν στον προορισμό τους.
- **Αποφυγή εμποδίων σε πραγματικό χρόνο:** Ανιχνεύοντας και αποφεύγοντας κινούμενα ή στάσιμα αντικείμενα, μπορούν να τροποποιούν δυναμικά την πορεία τους χρησιμοποιώντας οπτικά δεδομένα (όπως αυτά που λαμβάνονται από κάμερες ή LiDAR).
- **Συνεργατική πτήση και νοημοσύνη σμηνών:** Με την χρήση βαθιάς ενισχυτικής μάθησης, πολλά αεροχήματα μπορούν να μάθουν να συνεργάζονται για την εκτέλεση εργασιών όπως η μεταφορά φορτίων, η αναγνώριση στόχων και η επιτήρηση περιοχών.

1.4 Μεθοδολογία ανασκόπησης βιβλιογραφίας

Η παρούσα μελέτη αποτελεί μια εκτενή βιβλιογραφική ανασκόπηση (Comprehensive Literature Review), επειδή βασίζεται σε μια μεθοδική και αναπαραγώγιμη προσέγγιση [13]. Επιδιώκει να εντοπίσει, να αξιολογήσει, να συγκεντρώσει και να εξετάσει κάθε μελέτη σχετική με έναν συγκεκριμένο τομέα της έρευνας. Αξιοποιώντας βασικές αρχές του προτύπου PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [14] σε συνδυασμό με την προσέγγιση των Tranfield [15], η μέθοδος αυτή ορίζεται ως μια διαφανής και επιστημονική διαδικασία που επιδιώκει τη μείωση του αποπροσανατολισμού με τη διεξαγωγή ενδεδειγμένων βιβλιογραφικών αναζητήσεων [16] [17].

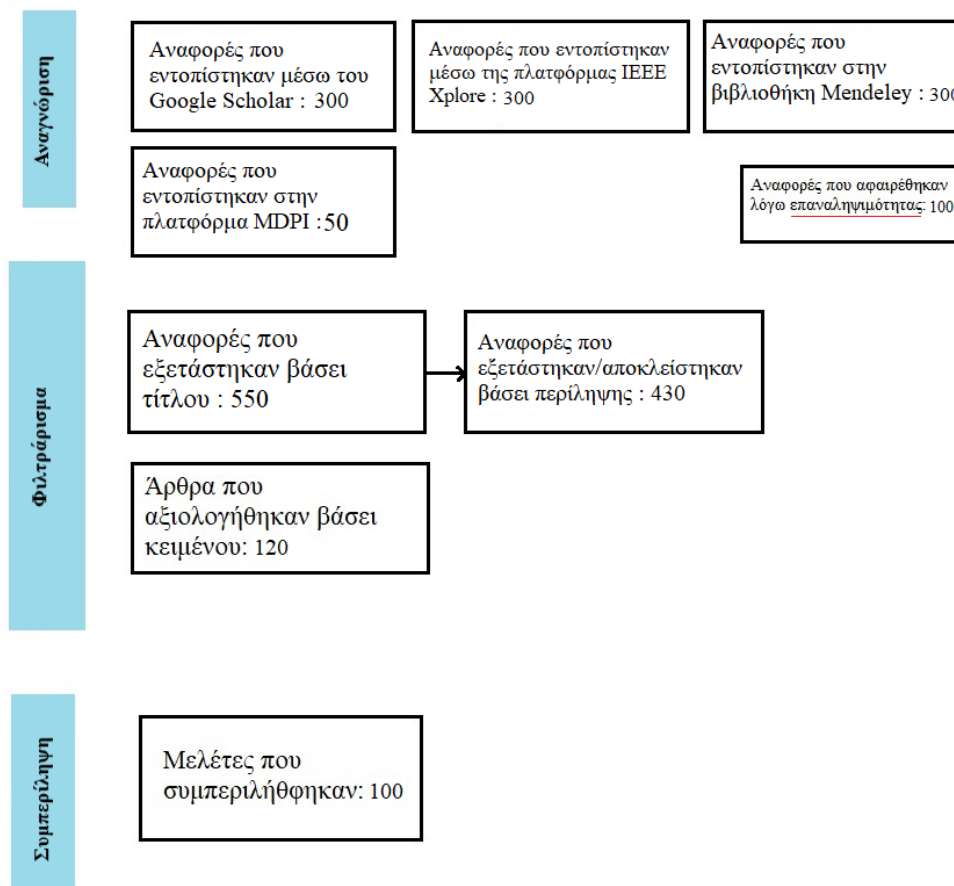
1.4.1 Σχεδιασμός της ανασκόπησης

Ο καθορισμός του στόχου της μελέτης και των βασικών ερωτημάτων είναι το πρώτο βήμα. Η ανάλυση των μεθόδων βαθιάς και ενισχυτικής μάθησης για τον έλεγχο μη επανδρωμένων αεροχημάτων είναι το κύριο θέμα της παρούσας εργασίας, εστιάζοντας κυρίως :

- Ποιές αλγοριθμικές προσεγγίσεις εφαρμόζονται στα μη επανδρωμένα αεροχήματα.
- Ποια είναι τα πλεονεκτήματα και μειονεκτήματα αυτών έναντι των παραδοσιακών τεχνικών ελέγχου.
- Ποιες είναι οι προκλήσεις και μελλοντικές τάσεις στον τομέα αυτό.

1.4.2 Διεξαγωγή της ανασκόπησης

Η διαδικασία συλλογής και επιλογής της σχετικής βιβλιογραφίας έγινε με αυστηρό και αναπαραγώγιμο τρόπο, ακολουθώντας τόσο τη φάση “Conducting the Review” της μεθοδολογίας Tranfield, όσο και τις οδηγίες του πλαισίου PRISMA για συστηματικές ανασκοπήσεις. Το πρώτο βήμα ήταν η διεξοδική αναζήτηση επιστημονικών δημοσιεύσεων σε αξιόπιστες πηγές, συμπεριλαμβανομένων των Google Scholar, IEEE Xplore, MDPI, Elsevier, SpringerLink και Mendeley Library. Προκειμένου να συμπεριληφθεί ένα ευρύ φάσμα σχετικών μελετών που συνδέουν τις τεχνολογίες βαθιάς και ενισχυτικής μάθησης με τα μη επανδρωμένα εναέρια οχήματα, οι όροι αναζήτησης περιλάμβαναν φράσεις όπως «UAV control», «Reinforcement Learning for UAVs», «Deep Learning in UAVs», «Swarm UAVs» και «Autonomous UAV navigation». Κατά την επιλογή των άρθρων, χρησιμοποιήθηκαν ορισμένα κριτήρια ένταξης, περιορισμένα σε αγγλόφωνες δημοσιεύσεις που εκδόθηκαν μεταξύ 2018 και 2025, και έχουν ρητό τεχνικό ή ακαδημαϊκό περιεχόμενο σχετικά με τη χρήση των σύγχρονων αλγορίθμων.



Σχήμα 1.5: Διάγραμμα ροής της βιβλιογραφικής ανασκόπησης.

Πίνακας 1.1: Κριτήρια ένταξης και αποκλεισμού στη συστηματική ανασκόπηση.

Κατηγορία	Κριτήρια ένταξης	Κριτήρια αποκλεισμού
Θεματικό πεδίο	Μελέτες που αφορούν βαθιά ενισχυτική μάθηση (πλοήγηση, έλεγχος, αποφυγή εμποδίων, σμήνη).	Μελέτες που αφορούν μη επανδρωμένα αεροσκάφη χωρίς βαθιά ενισχυτική μάθηση.
Τύπος δημοσίευσης	Άρθρα σε επιστημονικά περιοδικά με κριτική επιτροπή, συνέδρια, συστηματικές ανασκοπήσεις.	Blogs, white papers, τεχνικές αναφορές χωρίς επιτροπή.
Χρονολογία	2018–2025.	Μελέτες πριν το 2018.
Γλώσσα	Δημοσιεύσεις στην Αγγλική γλώσσα.	Μη αγγλικές πηγές.
Δεδομένα & αξιολόγηση	Βιβλιογραφικά άρθρα, Μελέτες που παρουσιάζουν προσομοίωση ή πειραματικά αποτελέσματα με UAVs.	Θεωρητικά πλαίσια χωρίς αξιολόγηση ή μη σχετιζόμενα με αεροχήματα.

Απειλές εγκυρότητας

Παρά την προσπάθεια να δοθεί μια συστηματική και αντικειμενική επισκόπηση της υπάρχουσας έρευνας σχετικά με τη χρήση της βαθιάς ενισχυτικής μάθησης στον έλεγχο των μη επανδρωμένων αεροχημάτων, πρέπει να αναγνωριστούν ορισμένα ζητήματα εγκυρότητας. Πρώτον, ο κίνδυνος για την εσωτερική εγκυρότητα πηγάζει από το γεγονός ότι οι διαφορετικές παράμετροι προσομοίωσης και τα αεροδυναμικά μοντέλα που χρησιμοποιούνται σε διαφορετικές πλατφόρμες προσομοίωσης (όπως Gazebo, AirSim και PyBullet) μπορούν να περιορίσουν την άμεση συγκρισιμότητα των αποτελεσμάτων. Δεύτερον, όσον αφορά την εξωτερική εγκυρότητα, τα περισσότερα αποτελέσματα δεν είναι πάντα πλήρως γενικεύσιμα σε πρακτικά περιβάλλοντα, καθώς βασίζονται σε σενάρια προσομοίωσης και όχι σε πραγματικές πτήσεις. Επιπλέον, υπάρχει κίνδυνος για την εγκυρότητα της δομής, καθώς διαφορετικές έρευνες χρησιμοποιούν διαφορετικές μετρήσεις (όπως RMSE, ποσοστό επιτυχίας και κατανάλωση ενέργειας), γεγονός που καθιστά δύσκολη την πλήρη ομογενοποίηση και σύγκρισή τους. Τέλος, ο κίνδυνος μεροληψίας στην επιλογή των μελετών και η πιθανή αδυναμία εύρεσης όλων των σχετικών εργασιών, δεδομένου ότι η αναζήτηση περιορίστηκε σε συγκεκριμένες βάσεις δεδομένων και χρονικό πλαίσιο (2018+), αποτελούν απειλή για την εγκυρότητα των συμπερασμάτων. Είναι σημαντικό να αναγνωριστούν αυτοί οι περιορισμοί, προκειμένου να ερμηνευθούν τα αποτελέσματα με προσοχή και να χρησιμοποιηθούν ως βάση για μελλοντικές μελέτες που θα χρησιμοποιούν καλύτερες τεχνικές.

1.5 Σκοπός της διατριβής

Στόχος της παρούσας διατριβής είναι να αναδείξει τις τρέχουσες ερευνητικές τάσεις και προκλήσεις που προκύπτουν στον τομέα του ελέγχου μη επανδρωμένων αεροσκαφών, εξετάζοντας και αναλύοντας τη χρήση προσεγγίσεων βαθιάς μάθησης και ενισχυτικής μάθησης. Πιο συγκεκριμένα, η εργασία επιδιώκει να αποσαφηνίσει τις θεμελιώδεις αρχές των τεχνικών βαθιάς ενισχυτικής μάθησης, τους τομείς χρήσης τους στα UAV και τους πρακτικούς περιορισμούς που αντιμετωπίζουν μέσω μιας εκτενούς επισκόπησης της βιβλιογραφίας και της θεωρητικής τεκμηρίωσης. Η κύρια συμβολή αυτής της ανασκόπησης είναι η οργανωμένη σύνθεση των υφιστάμενων γνώσεων, η κατηγοριοποίηση των στρατηγικών ελέγχου τόσο με βάση τον αλγοριθμικό τους τύπο όσο και την αρχιτεκτονική τους προσέγγιση, ενώ παράλληλα αξιολογείται η ετοιμότητά τους για εφαρμογή στον πραγματικό κόσμο.

1.6 Δομή της εργασίας

Η παρούσα εργασία αποτελείται από 6 Κεφάλαια :

- **Κεφάλαιο 1. Εισαγωγή:** Παρουσιάζεται το θέμα της έρευνας, μαζί με τον στόχο της μελέτης και το ερευνητικό ερώτημα, τη μεθοδολογία της συστηματικής ανασκόπησης και το ερευνητικό πλαίσιο.
- **Κεφάλαιο 2. Θεωρητικό υπόβαθρο:** Εξετάζονται οι βασικές ιδέες και οι αλγόριθμοι της βαθιάς και της ενισχυτικής μάθησης, και πώς σχετίζονται με τον έλεγχο των μη επανδρωμένων αεροσκαφών.
- **Κεφάλαιο 3. Εφαρμογές βαθιάς & ενισχυτικής μάθησης:** Συζητούνται οι κύριοι τομείς εφαρμογής, μαζί με επεξηγηματικές μεθόδους, έρευνα και εφαρμογές.
- **Κεφάλαιο 4. Σύγκριση και αξιολόγηση των μεθόδων:** Πραγματοποιείται συγκριτική ανάλυση των σύγχρονων τεχνικών προσεγγίσεων, με έμφαση στα πλεονεκτήματα, τα μειονεκτήματα και την απόδοσή τους.
- **Κεφάλαιο 5. Μελέτες περίπτωσης και προκλήσεις:** Εξετάζονται οι κύριες προκλήσεις και περιορισμοί, μαζί με την έρευνα περιπτώσεων χρήσης για περαιτέρω έρευνα.
- **Κεφάλαιο 6. Σύνοψη και μελλοντικές κατευθύνσεις:** Παρουσιάζονται και συζητούνται τα συνολικά ευρήματα και προτείνονται τρόποι περαιτέρω διερεύνησης.

Κεφάλαιο 2

Θεωρητικό υπόβαθρο

2.1 Επισκόπηση της μηχανικής μάθησης

Η μηχανική μάθηση χρησιμοποιείται για να προσπαθήσει να λύσει μια σειρά από ζητήματα τεχνητής νοημοσύνης που συνδέονται με την ανθρώπινη νοημοσύνη. Το βασικό στοιχείο των αλγορίθμων μηχανικής μάθησης είναι η ευελιξία τους, η οποία τους επιτρέπει να προσαρμόζονται χωρίς ρητό προγραμματισμό. Σε αντίθεση με τις παραδοσιακές μεθόδους, χρησιμοποιεί στατιστικά μοντέλα, στρατηγικές βελτιστοποίησης και μεθοδολογίες της θεωρίας πιθανοτήτων.

Με απώτερο στόχο την εξαγωγή μοτίβων, γνώσεων και πληροφοριών από τα δεδομένα, η πλειονότητα των πιο προηγμένων αλγορίθμων μηχανικής μάθησης λειτουργεί με παρόμοιο τρόπο: λαμβάνουν πολλά δεδομένα, εκπαιδεύονται σε αυτά και στη συνέχεια τα προσαρμόζουν και τα βελτιώνουν.

Μια κοινή μέθοδος αναγνώρισης για τη συμβατική μηχανική μάθηση είναι η διαδικασία εκπαίδευσης ενός αλγορίθμου για τη βελτίωση της ακρίβειας των προβλέψεών του. Ένας αλγόριθμος μηχανικής μάθησης θεωρείται επιτυχημένος εάν διαθέτει την ιδιότητα της γενίκευσης, που σημαίνει ότι τα συμπεράσματα που έχει λάβει από τα δεδομένα εκπαίδευσης μπορούν να εφαρμοστούν με επιτυχία σε νέα, μη αναγνωρισμένα δεδομένα. Υπάρχουν διάφορες μέθοδοι μάθησης, αλλά οι πιο βασικές περιλαμβάνουν την επιβλεπόμενη μάθηση (Supervised Learning), την μη-επιβλεπόμενη μάθηση (Unsupervised Learning), την ημι-επιβλεπόμενη μάθηση (Semi-supervised Learning), και την ενισχυτική μάθηση.

- **Επιβλεπόμενη μάθηση:** Σε αυτόν τον τύπο, παρέχονται στους αλγορίθμους επισημασμένα δεδομένα εκπαίδευσης και καθορίζονται οι μεταβλητές που θέλουν ο υπολογιστής να αναζητήσει. Ο στόχος είναι να μάθουν τη συσχέτιση μεταξύ αυτών, ώστε να μπορούν να προβλέψουν την έξοδο για νέες, άγνωστες εισόδους.
- **Μη επιβλεπόμενη μάθηση:** Στην περίπτωση αυτή, οι αλγόριθμοι αναζητούν συστάδες δεδομένων ή μοτίβα χωρίς ετικέτες. Χρησιμοποιείται συχνά για την ανάλυση τεράστιων ποσοτήτων, μη δομημένων δεδομένων, χωρίς να υπάρχει επίβλεψη ως προς το αποτέλεσμα που θα παραχθεί.
- **Ημι-επιβλεπόμενη μάθηση:** Αυτός ο τύπος μηχανικής μάθησης περιλαμβάνει το συνδυασμό των δύο προαναφερθέντων προσεγγίσεων. Το μοντέλο είναι ελεύθερο να διερευνήσει τα δεδομένα ανεξάρτητα και να αναπτύξει τη δική του κατανόηση του συνόλου δεδομένων.

- **Ενισχυτική Μάθηση:** Η ενισχυτική μάθηση διαφέρει από τις προηγούμενες κατηγορίες. Ένα πιο εξειδικευμένο είδος μάθησης στο οποίο οι πράκτορες μαθαίνουν μέσω αλληλεπίδρασης με το περιβάλλον, βασιζόμενοι σε ανταμοιβές (rewards) και ποινές (penalties). Με την πάροδο του χρόνου, οι πράκτορες προσπαθούν να μεγιστοποιήσουν τη συνολική τους απόδοση. Το σύστημα δεν καθοδηγείται από κάποιον επιβλέποντα, αλλά πρέπει μόνο του να ανακαλύψει ποιες είναι οι πιο επικερδείς ενέργειες.

2.2 Βασικές αρχές της ενισχυτικής μάθησης

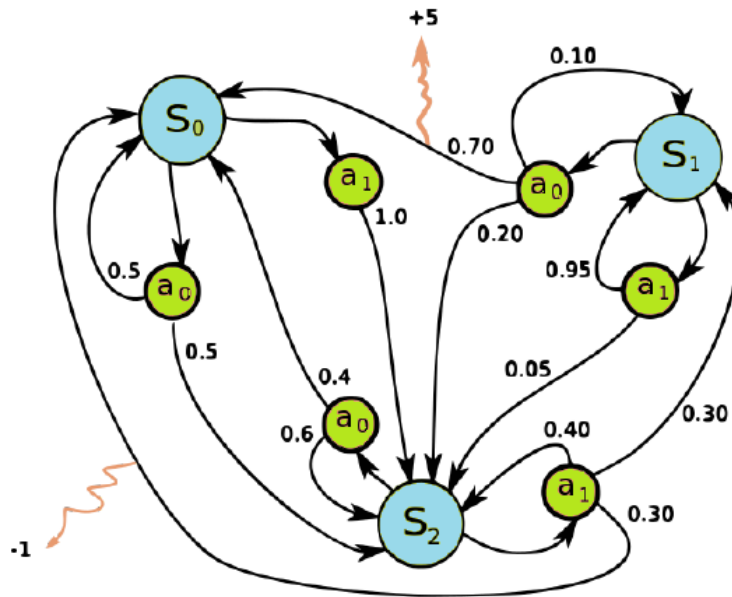
Η ενισχυτική μάθηση διακρίνεται από άλλες μεθόδους μηχανικής μάθησης λόγω της ικανότητάς της να χειρίζεται την αβεβαιότητα. Επιπλέον, δίνει μεγάλη έμφαση στη γενίκευση ή στην ικανότητα του πράκτορα να εφαρμόζει τις πληροφορίες που μαθαίνει σε νέες, μη δοκιμασμένες συνθήκες. Λόγω του ότι τα μη επανδρωμένα αεροχήματα πρέπει συχνά να λειτουργούν σε αχαρτογράφητες ή άγνωστες καταστάσεις, αυτό το χαρακτηριστικό είναι απαραίτητο. Η εξερεύνηση νέων στρατηγικών είναι επίσης κεντρικό ζήτημα, καθώς επηρεάζουν άμεσα την αποδοτικότητα και την αξιοπιστία των συστημάτων αυτών. Στην παρούσα εργασία, θα εξεταστούν και θα αναλυθούν μερικοί από τους πιο δημοφιλείς αλγόριθμους που χρησιμοποιούνται ευρέως για την ανάπτυξη αυτόνομων αεροχημάτων.

2.2.1 Μαρκοβιανή διαδικασία απόφασης

Ένα δημοφιλές μαθηματικό παράδειγμα για την τυποποίηση ζητημάτων επιλογής σε στοχαστικά πλαίσια στην ενισχυτική μάθηση είναι η μαρκοβιανή διαδικασία επιλογής (Markov Decision Process - MDP). Μια τέτοια διαδικασία εξηγεί πώς ένας πράκτορας μπορεί να κάνει διαδοχικές επιλογές ανάλογα με τις συνθήκες του περιβάλλοντος, τις ενέργειες που μπορεί να κάνει και τις ανταμοιβές που λαμβάνει. Ο πράκτορας που εκπαιδεύεται πρέπει να είναι σε θέση να "αισθάνεται" την κατάσταση του περιβάλλοντός του και να αναλαμβάνει δράσεις που μεταβάλλουν τις συνθήκες αυτές. Και τέλος, πρέπει να έχει έναν στόχο που να συνδέεται με την κατάσταση του περιβάλλοντος. Οι μαρκοβιανές διαδικασίες απόφασης είναι έτσι ορισμένες ώστε να εκφράζουν ακριβώς αυτά τα βασικά χαρακτηριστικά και αποτελούν ένα σαφές πλαίσιο του προβλήματος της μάθησης μέσω αλληλεπίδρασης. Πιο συγκεκριμένα :

- **\mathcal{S} (Καταστάσεις, States):** Το σύνολο όλων των πιθανών καταστάσεων στις οποίες μπορεί να βρεθεί το σύστημα.
- **\mathcal{A} (Δράσεις, Actions):** $\mathcal{A} = \{a_1, a_2, \dots, a_N\}$ ορίζεται ως το σύνολο των ενεργειών που μπορεί να πραγματοποιήσει ο πράκτορας.
- **\mathcal{P} (Μεταβατική πιθανότητα, Transition Probability):** $P(s_{t+1} | s_t, a_t)$ είναι η πιθανότητα να μεταβούμε από την κατάσταση s_t στην κατάσταση s_{t+1} υπό τη συνθήκη πως ο πράκτορας έχει πραγματοποιήσει την ενέργεια a_t .
- **\mathcal{R} (Ανταμοιβή, Reward):** $R(s_t, a_t, s_{t+1}) = r_t$ ορίζεται ως η επιβράβευση που λαμβάνει ο πράκτορας όταν αυτός πραγματοποιεί την ενέργεια a_t και το περιβάλλον μεταβαίνει από την κατάσταση s_t στην κατάσταση s_{t+1} .

- **γ (Συντελεστής Απόσβεσης, Discount Factor)** : Ένας αριθμός μεταξύ 0 και 1 που καθορίζει τη σημασία των μελλοντικών ανταμοιβών σε σχέση με τις άμεσες ανταμοιβές. Τα μελλοντικά οφέλη όσο πιο κοντά στο 1 είναι, τόσο πιο σημαντικά.



Σχήμα 2.1: Γραφική αναπαράσταση μιας Μαρκοβιανής Διαδικασίας Απόφασης (MDP). Απεικονίζονται οι καταστάσεις (μπλέ κόμβοι), οι πιθανές ενέργειες (πράσινοι κόμβοι) και οι οι ανταμοιβές υποδεικνύονται από τα κυματιστά κόκκινα βέλη. Τα μαύρα βέλη και οι αριθμοί δίπλα τους περιγράφουν τις πιθανότητες μετάβασης [18].

Η μαρκοβιακή διαδικασία απόφασης αποτελεί ένα μαθηματικό πλαίσιο για τη μοντελοποίηση προβλημάτων λήψης αποφάσεων υπό συνθήκες αβεβαιότητας, στα οποία ένας πράκτορας αλληλεπιδρά με ένα δυναμικό περιβάλλον. Η θεμελιώδης στοχαστικότητα της και η αλληλεπίδραση των καταστάσεων, των ενεργειών και των ανταμοιβών απεικονίζονται στο διάγραμμα 2.1, το οποίο αποτελεί ένα βασικό πλαίσιο για τη λήψη αποφάσεων σε συνθήκες αβεβαιότητας και χρησιμοποιείται συχνά σε εφαρμογές ενισχυτικής μάθησης. Ένα σύνολο καταστάσεων S , ένα σύνολο ενεργειών A , μια στοχαστική συνάρτηση μετάβασης $P(s' | s, a)$ και μια συνάρτηση ανταμοιβής $R(s, a, s')$ αποτελούν τον πυρήνα. Με βάση την τρέχουσα κατάσταση s , ο πράκτορας επιλέγει μια ενέργεια a σε κάθε χρονικό βήμα. Ο στόχος είναι να μεγιστοποιηθεί η συσσωρευμένη ανταμοιβή με την πάροδο του χρόνου, συνήθως μέσω της βελτιστοποίησης μιας πολιτικής $\pi(s)$.

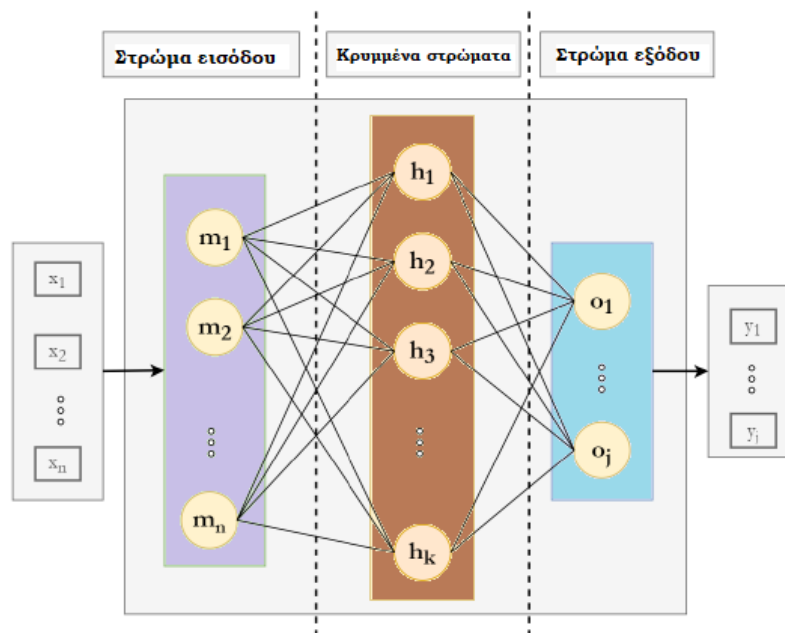
Η χρήση των MDPs είναι ζωτικής σημασίας, διότι επιτρέπει τη δημιουργία αλγορίθμων που είναι ικανοί να παίρνουν μόνοι τους αποφάσεις σε πραγματικό χρόνο, λαμβάνοντας υπόψη μεταβλητές όπως ο άνεμος, τα εμπόδια και η χρήση ενέργειας. Μπορεί να χρησιμοποιηθεί για τη μεγιστοποίηση της αποτελεσματικότητας της αποστολής, την ελαχιστοποίηση του κινδύνου σύγκρουσης και τη βελτιστοποίηση της τροχιάς ενός αεροχήματος. Ο στόχος είναι να προσδιοριστεί η βέλτιστη στρατηγική (Policy) που μεγιστοποιεί τη συνολική αναμενόμενη ανταμοιβή, γνωστή ως συνάρτηση αξίας (Value Function) :

$$V^\pi(s) = R(s, a) + \gamma \sum_{s'} P(s' | s, a) V^\pi(s')$$

Ωστόσο, έχουν μειονεκτήματα παρά την αποτελεσματικότητά τους, όπως η απαίτηση για πλήρη κατανόηση του μοντέλου μετάβασης και των ανταμοιβών, κάτι που δεν είναι πάντα εφικτό σε ένα πραγματικό περιβάλλον. Οι μεγάλοι χώροι καταστάσεων μπορούν επίσης να οδηγήσουν σε υπολογιστικές απαιτήσεις που αυξάνονται εκθετικά, καθιστώντας τη συμβατική μέθοδο δύσκολη.

2.3 Εισαγωγή στην βαθιά μάθηση

Μια προηγμένη τεχνική μηχανικής μάθησης είναι η βαθιά μάθηση (Deep Learning). Η βασική διαφορά έγκειται στην αυτόματη εξαγωγή χαρακτηριστικών. Η βαθιά μάθηση χρησιμοποιεί πολλαπλά επίπεδα επεξεργασίας (layers) που μαθαίνουν αυτόματα ποια χαρακτηριστικά είναι σημαντικά, χωρίς ανθρώπινη παρέμβαση, ειδικά σε πολύπλοκα δεδομένα. Δεδομένου ότι ένα νευρωνικό δίκτυο διαθέτει πολλαπλά επίπεδα μεταξύ των επιπέδων εισόδου και εξόδου, ο όρος «βαθύς» χρησιμοποιείται στο πλαίσιο της βαθιάς μάθησης. Το κυριότερο είναι ότι μια διαδικασία βαθιάς μάθησης μπορεί να επιλέξει ανεξάρτητα ποια χαρακτηριστικά ανήκουν σε ποιο επίπεδο. Πριν από αυτό, οι μέθοδοι μηχανικής μάθησης απαιτούσαν συχνά τη χειροκίνητη δημιουργία χαρακτηριστικών για να αλλάξουν τα δεδομένα σε μια αναπαράσταση που θα μπορούσε να χρησιμοποιήσει ένα σύστημα ταξινόμησης. Αντιθέτως, με τις τεχνικές βαθιάς μάθησης, το μοντέλο εξάγει αυτόματα πολύτιμες αναπαραστάσεις χαρακτηριστικών από τα δεδομένα. Τα νευρωνικά δίκτυα προσφέρουν μια σειρά μαθηματικών συναρτήσεων που μπορούν να χρησιμοποιηθούν για ευρείς σκοπούς για να εκφράσουν μη γραμμικούς μετασχηματισμούς μεταξύ διαφορετικών συνόλων δεδομένων.



Σχήμα 2.2: Αρχιτεκτονική ενός Πολυεπίπεδου Perceptron (MLP) με ένα κρυφό στρώμα. Το δίκτυο δέχεται n χαρακτηριστικά εισόδου, τα επεξεργάζεται μέσω των βαρέων συνδέσεων στο κρυφό επίπεδο για την εκμάθηση μη γραμμικών σχέσεων και παράγει j τιμές εξόδου [12].

Στο Σχήμα 2.2 φαίνεται ένα πολυεπίπεδο perceptron (Multi Layer Perceptron - MLP), μια βασική αρχιτεκτονική τεχνητού νευρωνικού δικτύου που έχει δημιουργηθεί ειδικά για τη διαχείριση σύνθετων εισροών και τη λήψη αποφάσεων. Περιλαμβάνει ένα στρώμα εισόδου με n νευρώνες, έναν αριθμό χαρακτηριστικών εισόδου όπως οι μετρήσεις από αισθητήρες, ένα και μοναδικό κρυφό στρώμα, και ένα στρώμα εξόδου με j νευρώνες που αποδίδουν τα τελικά αποτελέσματα του δικτύου. Το MLP κατηγοριοποιείται ως μοντέλο επιβλεπόμενης μάθησης. Σε κάθε νευρωνική σύνδεση αποδίδεται ένα βάρος, το οποίο αλλάζει κατά τη διάρκεια της εκπαίδευσης για τη μεγιστοποίηση της απόδοσης με τη χρήση τεχνικών όπως η οπισθοδρόμηση. Μέσω της εφαρμογής συναρτήσεων ενεργοποίησης (όπως η Rectified Linear Unit - ReLU), το κρυφό στρώμα προσθέτει μη γραμμικότητα στο σύστημα, επιτρέποντας στο δίκτυο να αναπαριστά πιο περίπλοκες σχέσεις εισόδου-εξόδου. Το δίκτυο perceptron πολλαπλών στρωμάτων παρουσιάζεται ως ρητή μορφή της συσχέτισης. Αυτή η κατηγορία δικτύων έχει αποτελέσει τη βάση για την πλειονότητα των εφαρμογών νευρωνικών δικτύων στην πράξη μέχρι σήμερα.

2.3.1 Τεχνητά νευρωνικά δίκτυα

Η βαθιά εκμάθηση βασίζεται σε μεγάλο βαθμό σε τεχνητά νευρωνικά δίκτυα (Artificial Neural Network - ANNs). Για να επιλύσουν πολύπλοκα προβλήματα που είναι δύσκολα να μοντελοποιηθούν με συμβατικές τεχνικές, αυτά τα μοντέλα επιδιώκουν να μάθουν μη γραμμικές συσχετίσεις μεταξύ εισόδων και αποτελεσμάτων. Η βασική δομή ενός τεχνητού νευρωνικού δικτύου αποτελείται από τρία διαφορετικά είδη επιπέδων: το στρώμα εισόδου, ένα ή περισσότερα κρυφά επίπεδα και το στρώμα εξόδου. Κάθε στρώση αποτελείται από τεχνητούς νευρώνες που εκτελούν έναν υπολογισμό βάσει μιας γραμμικής μετασχηματιστικής λειτουργίας και μιας μη γραμμικής συνάρτησης ενεργοποίησης (όπως η ReLU ή η sigmoid). Οι αλγόριθμοι βελτιστοποίησης, κυρίως της οπισθοδρόμησης, σε συνδυασμό με μεθόδους όπως η gradient descent, χρησιμοποιούνται στη διαδικασία εκπαίδευσης.

Η ικανότητα των τεχνητών νευρωνικών δικτύων να γενικεύουν από παραδείγματα, ακόμη και σε πολύ περίπλοκες καταστάσεις, είναι το κύριο πλεονέκτημά τους. Τα ANN είναι ελκυστικά για τη μοντελοποίηση δυναμικών συστημάτων σε εφαρμογές ελέγχου λόγω της ικανότητάς τους να προσεγγίζουν αυθαίρετες συναρτήσεις. Επιπλέον, παρέχουν μια δομική βάση για πιο περίπλοκες αρχιτεκτονικές όπως τα αναδρομικά ή τα συνελκτικά δίκτυα. Επομένως, η κατανόηση των πιο εξελιγμένων μεθόδων που ακολουθούν απαιτεί ως προϋπόθεση την σταθερή κατανόηση των θεμελιωδών ικανοτήτων και της λειτουργίας τους.

2.3.2 Νευρωνικά δίκτυα στην εκμάθηση πολιτικών ελέγχου

Η χρήση νευρωνικών δικτύων στην εκμάθηση πολιτικών ελέγχου μη επανδρωμένων αεροσκαφών έχει αναδειχθεί ως μία από τις πλέον ισχυρές προσεγγίσεις, ιδιαίτερα σε εφαρμογές ενισχυτικής μάθησης. Στην πράξη, ένα αεροσκάφος λειτουργεί σε ένα δυναμικό και συχνά απρόβλεπτο περιβάλλον, λαμβάνοντας αποφάσεις με βάση συνεχώς μεταβαλλόμενα δεδομένα. Τα νευρωνικά δίκτυα, και συγκεκριμένα τα βαθιά νευρωνικά δίκτυα (Deep Neural Networks - DNN), χρησιμοποιούνται για την προσέγγιση πολύπλοκων συναρτήσεων όπως η πολιτική ή η συνάρτηση αξίας ενός πράκτορα. Η ικανότητα των νευρωνικών δικτύων να γενικεύουν σε άγνωστες συνθήκες, καθώς και η ικανότητά τους να λειτουργούν με σύνθετες εισόδους όπως εικόνες ή δεδομένα αισθητήρων, τα

καθιστούν κρίσιμα εργαλεία για την ευφυή και αυτόνομη πλοήγηση των μη επανδρωμένων αεροσκαφών.

2.4 Βασικοί αλγόριθμοι και μοντέλα ενισχυτικής μάθησης

Η ενισχυτική μάθηση περιλαμβάνει ένα σύνολο αλγορίθμων και μοντέλων που επιτρέπει σε πράκτορες -όπως τα μη επανδρωμένα αεροσκάφη- να μαθαίνουν πώς να αλληλεπιδρούν με το περιβάλλον τους και να κάνουν τις καλύτερες επιλογές. Σε αντίθεση με άλλες τεχνικές μηχανικής μάθησης, επικεντρώνεται στη μεγιστοποίηση της συνολικής ανταμοιβής μέσα από μια σειρά ενεργειών

On-policy και off-policy

Ένα κριτήριο διαχωρισμού των μεθόδων ενισχυτικής μάθησης είναι ο τρόπος με τον οποίο ενημερώνουν την πολιτική τους.

- **On-policy** : Οι on-policy μέθοδοι αξιολογούν ή βελτιώνουν την πολιτική συμπεριφοράς τους. Ο αλγόριθμος SARSA, για παράδειγμα, προσαρμόζει τη συνάρτηση αξίας-ενέργειας στην υπάρχουσα πολιτική. Δηλαδή, αξιολογεί την πολιτική χρησιμοποιώντας δείγματα της ίδιας πολιτικής και στη συνέχεια κάνει άπληστες (greedy) βελτιώσεις της πολιτικής με βάση τις τιμές ενέργειας. Ουσιαστικά είναι ένας αλγόριθμος ενισχυτικής μάθησης που χρησιμοποιείται για την εκπαίδευση ενός πράκτορα σε ένα άγνωστο περιβάλλον. Το όνομά του προέρχεται από την πεντάδα s, a, r', s', a' που χρησιμοποιούνται για την ανανέωση της συνάρτησης Q (Q Function), η οποία είναι ως εξής :

$$Q(s, a) \leftarrow Q(s, a) + \alpha_t(s, a) [r' + \gamma Q(s', a') - Q(s, a)]$$

- **Off-policy** : Με την χρήση μεθόδων off-policy ο πράκτορας ακολουθεί μια δυνητικά άσχετη πολιτική συμπεριφοράς ενώ μαθαίνει μια βέλτιστη συνάρτηση αξίας. Χωρίς απαραίτητα να προσαρμόζεται στην πολιτική που παράγεται από τα δείγματα από τα οποία μαθαίνει, ο αλγόριθμος Q-Learning για παράδειγμα, προσπαθεί να βρει τις τιμές αξίας-ενέργειας για την βέλτιστη πολιτική άμεσα.

2.4.1 Q-Learning

Ένας από τους πιο δημοφιλείς και απλούς αλγορίθμους ενισχυτικής μάθησης είναι ο αλγόριθμος Q-Learning. Δεδομένου ότι μαθαίνει τη βέλτιστη συνάρτηση αξίας (Q-function) ανεξάρτητα από την πολιτική που ακολουθεί ο πράκτορας κατά την εκπαίδευση, εντάσσεται στην κατηγορία των off-policy προσεγγίσεων. Για να μπορέσει να εκτελεστεί ο αλγόριθμος, είναι απαραίτητος ο ορισμός μιας συνάρτησης επιβράβευσης η οποία σχετίζεται με το πρόβλημα προς επίλυση και το περιβάλλον στο οποίο ενεργεί ο πράκτορας. Έχει ορισμένους περιορισμούς στην εφαρμογή των μη επανδρωμένων αεροσκαφών · παραμένει χρήσιμος σε πιο απλά ή διακριτά περιβάλλοντα πτήσης. Πιο συγκεκριμένα, πρόκειται για την αδυναμία χειρισμού δράσεων. Τα αεροσκάφη συνήθως χρειάζονται λεπτομερή έλεγχο των στροφών και της ώθησης, κάτι που δεν διαχειρίζεται καλά.

Σε ένα απλό παράδειγμα, ο Q-Learning μπορεί να χρησιμοποιηθεί για την ελαχιστοποίηση της κατανάλωσης ενέργειας και την αποφυγή εμποδίων βελτιστοποιώντας τις διαδρομές ενός μη επανδρωμένου αεροχήματος που περιπολεί σε μια προκαθορισμένη περιοχή. Έτσι, μαθαίνει να συντονίζει τις κινήσεις του ώστε να καλύπτει την περιοχή όσο το δυνατόν περισσότερο, χρησιμοποιώντας παράλληλα τη μικρότερη δυνατή ποσότητα ενέργειας.

2.4.2 Deep Q-Network

Ένα από τα σημαντικότερα επιτεύγματα στην ανάπτυξη της ενισχυτικής μάθησης ήταν η δημιουργία του Deep Q-Network (DQN), ενός αλγορίθμου που ενσωματώνει τα νευρωνικά δίκτυα στις παραδοσιακές μεθόδους Q-learning. Ο DQN εκπαιδεύει με επιτυχία πράκτορες σε πολύπλοκα περιβάλλοντα με εισόδους υψηλής διάστασης, χρησιμοποιώντας τεχνικές όπως η εκμάθηση μέσω εμπειρίας (experience learning) και τα target networks για τη διατήρηση της σταθερότητας της μάθησης.

- **Εκμάθηση μέσω εμπειρίας:** Η βασική ιδέα είναι η αποθήκευση παλαιών εμπειριών s, a, r, s' και δειγματοληψία που κάνει πιο ομαλές τις αλλαγές στην κατανομή των δεδομένων.
- **Target Networks:** Ξεχωριστό Q-δίκτυο που ενημερώνεται περιοδικά, πράγμα που μειώνει την συσχέτιση μεταξύ των Q-values. Η συνάρτηση απώλειας (loss function) που χρησιμοποιείται για την ενημέρωση του αλγορίθμου είναι η εξής :

$$L(\theta) = \mathbb{E}_{(s,a,r,s')} \left[\left(r + \gamma \max_{a'} Q(s', a'; \theta^-) - Q(s, a; \theta) \right)^2 \right]$$

Ο DQN παρά την απλότητά του, έχει περιορισμούς όταν πρόκειται για συνεχείς χώρους ενεργειών. Ωστόσο, θεωρείται ως σημείο εκκίνησης για πιο περίπλοκες τεχνικές βαθιάς ενισχυτικής μάθησης, οι οποίες εξετάζονται παρακάτω.

2.4.3 Αλγόριθμος βαθιάς ντετερμινιστικής κλίσης πολιτικής

Όταν πρόκειται για τον έλεγχο μη επανδρωμένων αεροχημάτων, ο αλγόριθμος βαθιάς ντετερμινιστικής κλίσης πολιτικής (Deep Deterministic Policy Gradient - DDPG) είναι πολύ πιο αποτελεσματικός από τον Q-Learning, καθώς μπορεί να διαχειρίζεται συνεχείς περιοχές δράσης. Αντί να χρησιμοποιηθούν πιθανότητες για την αξιολόγηση των ενεργειών, η τεχνική αυτή χρησιμοποιεί τις εξόδους του δικτύου απευθείας ως ενέργειες που υλοποιεί ο πράκτορας. Επιπλέον, χρησιμοποιείται ένα δίκτυο-στόχος για να εξασφαλιστεί η σύγκλιση με τη χρήση της χρονικής καθυστέρησης για την αξιολόγηση του μοντέλου.

Όπως και σε άλλους αλγόριθμους, γίνεται χρήση ενός buffer για την ενημέρωση των παραμέτρων των δικτύων. Με αυτόν τον τρόπο, δίνεται μεγαλύτερη σημασία στις ενέργειες που επέφεραν υψηλή αλλαγή ανταμοιβής, βελτιώνοντας την απόδοση και την σταθερότητα. Λόγω των δυσκολιών της κατάστασης ελέγχου, απαιτείται συνοχή στους χώρους δράσης και καταστάσεων, καθιστώντας τον αλγόριθμο βαθιάς ντετερμινιστικής κλίσης πολιτικής έναν κατάλληλο υποψήφιο για ένα actor-critic πεδίο. Συνδυάζοντας τις ικανότητες των μεθόδων actor-critic και των βαθιών νευρωνικών δικτύων, οδηγεί στην εκμάθηση βέλτιστων πολιτικών σε συνεχείς χώρους δράσης. Μια ντετερμινιστική πολιτική μπορεί να οριστεί ως μια συνάρτηση :

$$a = \pi(s; \theta)$$

Η διαδικασία λειτουργίας μπορεί να παρουσιαστεί ως εξής :

1. **Αναπαράσταση πολιτικής και δράσης** : Μετατροπή της τρέχουσας κατάστασης του περιβάλλοντος σε μια κατάλληλη αριθμητική αναπαράσταση, όπως ακατέργαστες αισθητηριακές εισόδους ή προεπεξεργασμένα χαρακτηριστικά. Αντίστοιχα, αναπαράσταση των ενεργειών ως συνεχείς τιμές.
2. **Αρχιτεκτονική actor - critic** : Σχεδιασμός δύο βαθιών νευρωνικών δικτύων: Το actor δίκτυο, το οποίο μέσω μιας συνάρτησης πολιτικής χαρτογραφεί καταστάσεις σε ντετερμινιστικές δράσεις. Το critic δίκτυο, που μαθαίνει μια συνάρτηση τιμής δράσης για την εκτίμηση της αναμενόμενης μελλοντικής ανταμοιβής.
3. **Replay buffer** : Αποθηκεύει τις τρέχουσες πλειάδες (κατάσταση, δράση, ανταμοιβή, επόμενη κατάσταση) για ασύγχρονη εκπαίδευση και μείωση των χρονικών συσχετίσεων.
4. **Εξερεύνηση και εκμετάλλευση** : Η εξισορρόπηση της εξερεύνησης και της εκμετάλλευσης, γίνεται με την χρήση ντετερμινιστικής πολιτικής με προσθήκη θορύβου (π.χ. Ornstein-Uhlenbeck), για να διατηρεί την εξερεύνηση στις συνεχείς δράσεις.
5. **Δίκτυα στόχοι** : Διατήρηση ξεχωριστών δικτύων-στόχων για το actor και το critic, τα οποία ενημερώνονται με αργό ρυθμό.
6. **Εκπαίδευση** : Κατά τη διάρκεια της φάσης εκπαίδευσης, γίνεται δειγματοληψία εμπειριών από τον replay buffer,
7. **Επανάληψη των παραπάνω βημάτων**, αλληλεπίδραση με το περιβάλλον, συλλογή εμπειριών, ενημέρωση των δικτύων, μέχρις ότου η πολιτική να συγκλίνει σε μια βέλτιστη στρατηγική.

2.4.4 Αλγόριθμος βελτιστοποίησης εγγύς πολιτικής

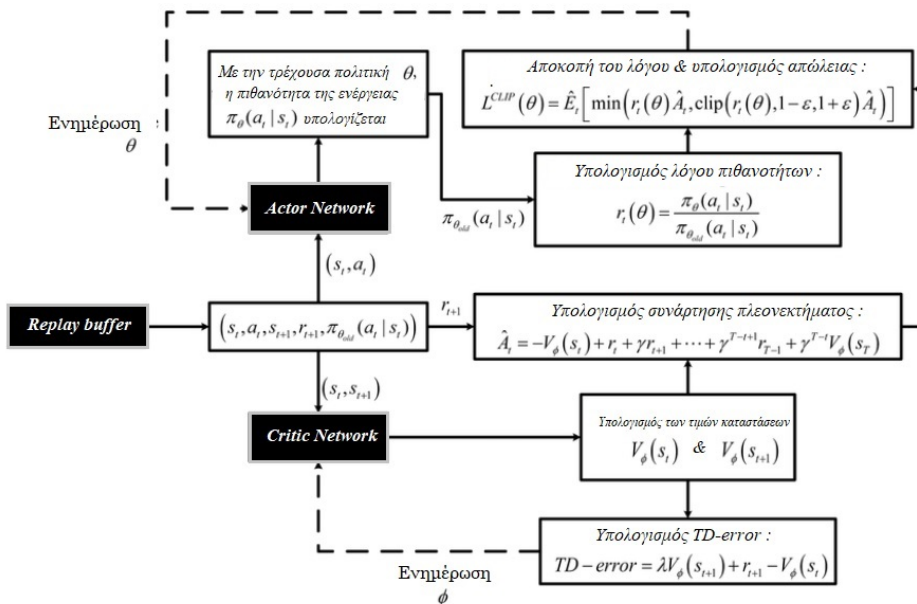
Ο αλγόριθμος βελτιστοποίησης εγγύς πολιτικής (Proximal Policy Optimization - PPO) είναι ένας on-policy αλγόριθμος που βελτιώνει την απόδοση με απλούστερη υλοποίηση και μεγαλύτερη σταθερότητα. Παρόμοια με άλλες μεθόδους βαθμίδων πολιτικής, ο PPO είναι μια τεχνική βαθμίδων πολιτικής που βελτιστοποιεί άμεσα την συνάρτηση μιας πολιτικής. Για να αποφευχθούν οι μεγάλες και ασταθείς ενημερώσεις κατά την εκπαίδευση, χρησιμοποιεί μια τροποποιημένη συνάρτηση, σε αντίθεση με τις συμβατικές προσεγγίσεις. Αυτό επιτυγχάνεται με την εφαρμογή ενός μηχανισμού «αποκοπής» (clipping) που περιορίζει τις αλλαγές πολιτικής και εγγυάται ότι αυτές παραμένουν εντός ενός λογικού και ασφαλούς εύρους, οδηγώντας σε πιο σταθερή μάθηση και καλύτερη συνολική απόδοση. Η συνάρτηση ορίζεται ως :

$$L_{CLIP}(\theta) = \hat{E} \left[\min \left(r_t(\theta) \hat{A}_t, \text{clip} \left(r_t(\theta), 1 - \epsilon, 1 + \epsilon \right) \hat{A}_t \right) \right]$$

όπου

$$r_t(\theta) = \frac{\pi_{\theta}(a_t | s_t)}{\pi_{\theta_{\text{old}}}(a_t | s_t)}$$

υποδηλώνει το λόγο της πιθανότητας της νέας πολιτικής προς την πιθανότητα της παλιάς πολιτικής, ϵ είναι μία υπερπαράμετρος που καθορίζει το εύρος αποκοπής (clipping threshold), και \hat{A}_t είναι ένας εκτιμητής του πλεονεκτήματος (advantage) τη χρονική στιγμή t .



Σχήμα 2.3: Διάγραμμα ροής αλγορίθμου βελτιστοποίησης εγγύς πολιτικής (Proximal Policy Optimization - PPO). Απεικονίζεται η αρχιτεκτονική Actor-Critic, η διαδικασία συλλογής εμπειριών (replay buffer) και ο μηχανισμός αποκοπής (clipping) στη συνάρτηση αντικειμενικού σκοπού για τη σταθεροποίηση της εκπαίδευσης [20], [21].

Στο Σχήμα 2.3 απεικονίζεται η λειτουργία του αλγορίθμου βελτιστοποίησης εγγύς πολιτικής, ο οποίος κατατάσσεται στις τεχνικές actor-critic της ενισχυτικής μάθησης. Ο πρωταρχικός στόχος είναι η βελτιστοποίηση της πολιτικής ενός πράκτορα, προκειμένου να καταστεί δυνατή η σταθερή και αποτελεσματική λήψη αποφάσεων σε εξαιρετικά περίπλοκες καταστάσεις. Ο PPO συνδυάζει ένα critic δίκτυο, το οποίο υπολογίζει την αξία των καταστάσεων, με ένα actor δίκτυο, το οποίο κάνει προτάσεις για ενέργειες. Το διάγραμμα δείχνει πώς τα δεδομένα εμπειριών αποθηκεύονται σε buffer, πώς αξιολογούνται μέσω των συναρτήσεων πλεονεκτήματος και TD-error, και πώς ενημερώνεται η πολιτική με τη χρήση της clipped συνάρτησης απώλειας.

Ενώ το critic δίκτυο υπολογίζει την αξία των σεναρίων με βάση τις εμπειρίες $(s_t, a_t, r_{t+1}, s_{t+1})$ του πράκτορα, το actor δίκτυο είναι υπεύθυνο για τη δειγματοληψία ενεργειών σύμφωνα με την τρέχουσα πολιτική $\pi_\theta(a|s)$. Αυτές αντλούνται για εκπαίδευση και ενημέρωση παραμέτρων από το replay buffer, όπου και αποθηκεύονται. Μια βασική καινοτομία του αλγορίθμου είναι η χρήση της συνάρτησης αποκοπής, η οποία εισάγει έναν περιορισμό στον λόγο πιθανοτήτων

$$r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)}.$$

Η διαδικασία υπολογισμού της συνάρτησης πλεονεκτήματος (advantage function) βασίζεται στον υπολογισμό

$$\hat{A}_t = -V_\phi(s_t) + r_t + \gamma r_{t+1} + \dots + \gamma^T r_{t+T-1} + \gamma^T V_\phi(s_T),$$

ο οποίος μετρά το πόσο ευνοϊκή είναι μια ενέργεια σε σχέση με την εκτίμηση του critic. Παράλληλα, ο critic ενημερώνεται μέσω του σφάλματος χρονικής διαφοράς

$$TD\text{-error} = r_t + \gamma V_\phi(s_{t+1}) - V_\phi(s_t),$$

το οποίο αποτελεί μέτρο για την ακρίβεια των εκτιμήσεών του. Συνολικά, το διάγραμμα αποτυπώνει τη στενή αλληλεπίδραση actor-critic, τη χρήση του replay buffer και τη σταθεροποιητική επίδραση της αποκοπής.

Γενικευμένη εκτίμηση πλεονεκτήματος

Για τον αποτελεσματικό υπολογισμό των εκτιμήσεων πλεονεκτήματος \hat{A}_t , ο αλγόριθμος βελτιστοποίησης πολιτικής χρησιμοποιεί συχνά τη γενικευμένη εκτίμηση πλεονεκτήματος (Generalized Advantage Estimation - GAE), η οποία είναι μια τεχνική που προσφέρει έναν συμβιβασμό μεταξύ υψηλής ακρίβειας και χαμηλής διακύμανσης στους υπολογισμούς του πλεονεκτήματος λαμβάνοντας υπόψη τις μακροπρόθεσμες ανταμοιβές. Υπολογίζεται ως :

$$\hat{A}_t = \delta_t + (\gamma\lambda)\delta_{t+1} + \dots + (\gamma\lambda)^{T-t+1}\delta_{T-1}$$

όπου

$$\delta_t = r_t + \gamma V_\nu(s_{t+1}) - V_\nu(s_t),$$

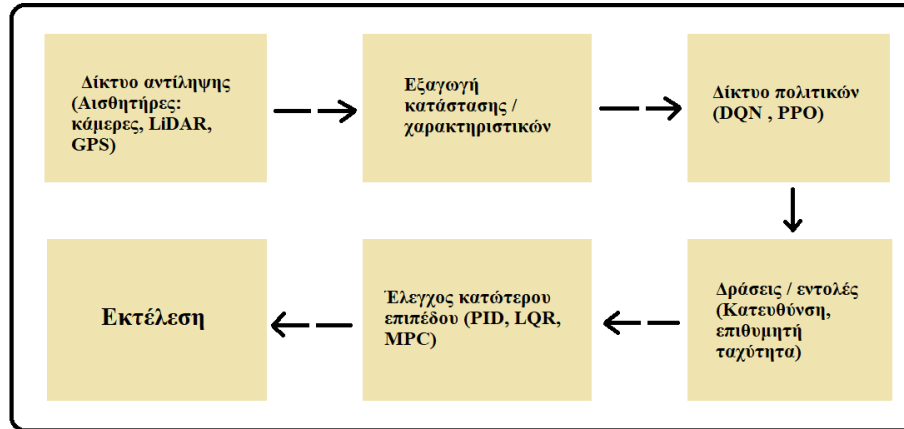
δ_t είναι το σφάλμα χρονικής διαφοράς (temporal difference error), γ είναι ο discount factor, λ είναι η παράμετρος του GAE ($0 \leq \lambda \leq 1$), και T είναι ο ορίζοντας (horizon) της τροχιάς (trajectory), $V_\nu(s_t)$ είναι η εκτιμώμενη αξία της κατάστασης s_t , και $V_\nu(s_{t+1})$ είναι η εκτιμώμενη αξία της επόμενης κατάστασης s_{t+1} .

Σύγκριση των αλγορίθμων

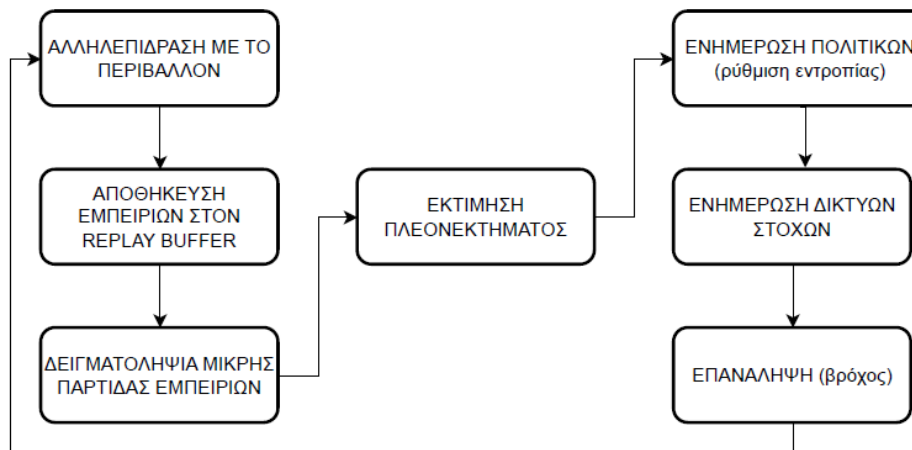
Αλγόριθμος	Εφαρμογές	Σταθερότητα	Υπολογιστικός κόστος
Q-Learning	Απλή πλοήγηση	Μέτρια	Χαμηλό
DQN	Απλή πλοήγηση	Υψηλότερη απο Q-Learning	Μέτριο
DDPG	Ακριβής έλεγχος	Μέτρια/Χαμηλή	Μέτριο
PPO	Πολύπλοκα περιβάλλοντα	Υψηλή	Υψηλή
SAC	Ρομποτικός έλεγχος	Υψηλή	Πολύ υψηλό

Πίνακας 2.1: Συγκριτική απεικόνιση των κυριότερων αλγορίθμων ενισχυτικής μάθησης (Q-Learning, DQN, DDPG, PPO) ως προς την καταλληλότητά τους για εφαρμογές UAV, τη σταθερότητα εκπαίδευσης και τις υπολογιστικές τους απαιτήσεις.

Στον πίνακα 2.1 συγκρίνονται το εύρος, η σταθερότητα και το υπολογιστικό κόστος τεσσάρων αντιπροσωπευτικών αλγορίθμων ενισχυτικής μάθησης. Λόγω της ευκολίας χρήσης και του χαμηλού υπολογιστικού κόστους, ο βασικός αλγόριθμος Q-Learning χρησιμοποιείται γενικά σε απλά σενάρια πλοήγησης με διακριτούς και περιορισμένους χώρους κατάστασης και δράσης. Ωστόσο, επειδή εξαρτάται από την ποιότητα του πίνακα Q και την πολιτική εξερεύνησης, η σταθερότητά του θεωρείται μέτρια. Τα Deep Q-Networks είναι κατάλληλα για πιο περίπλοκες καταστάσεις, επειδή επεκτείνουν το Q-learning εκτιμώντας τη συνάρτηση Q χρησιμοποιώντας βαθιά νευρωνικά δίκτυα. Λόγω μεθόδων όπως εκμάθησης μέσω εμπειρίας και τα δίκτυα στόχους, είναι πιο σταθερό από το Q-Learning. Ταυτόχρονα, το υπολογιστικό κόστος είναι μέτριο λόγω της εκπαίδευσης των δικτύων. Ο αλγόριθμος DDPG είναι προσανατολισμένος στον ακριβή έλεγχο σε συνεχή πεδία ενεργειών, καθιστώντας τον κατάλληλο για εργασίες όπως η ρύθμιση ταχύτητας, υψομέτρου και προσανατολισμού. Αν και είναι ικανός για τέτοιες εφαρμογές, η σταθερότητά του είναι μέτρια έως χαμηλή λόγω της ευαισθησίας του στις αρχικές συνθήκες και τις υπερπαραμέτρους, ενώ το υπολογιστικό του κόστος είναι μέτριο. Ο PPO αποτελεί μια ιδιαίτερα ισχυρή μέθοδο για πολύπλοκα και δυναμικά περιβάλλοντα, συμπεριλαμβανομένων συνεργατικών αποστολών μη επανδρωμένων αεροσκαφών. Μέσω του μηχανισμού περιορισμένων ενημερώσεων πολιτικής που αποτρέπει μεγάλες αποκλίσεις, του προσδίδει μεγάλη σταθερότητα. Ωστόσο, λόγω των απαιτήσεων σε δειγματοληψία, το υπολογιστικό του κόστος είναι υψηλό. Τέλος, ο αλγόριθμος SAC (Soft Actor-Critic) είναι αξιοσημείωτος για την ικανότητά του να ενισχύει την εντροπία πολιτικής, η οποία προάγει τη βέλτιστη εξερεύνηση, και την υψηλή αποτελεσματικότητα δειγματοληψίας. Ωστόσο, η αρχιτεκτονική του έχει το υψηλότερο υπολογιστικό κόστος σε σχέση με άλλους αλγόριθμους.



Σχήμα 2.4: Ροή ελέγχου με την χρήση βαθιάς ενισχυτικής μάθησης: η αντίληψη του περιβάλλοντος (Perception) τροφοδοτεί την υψηλού επιπέδου πολιτική (Policy), η οποία παράγει εντολές για τον έλεγχο χαμηλού επιπέδου (PID/LQR). Το διάγραμμα συνδέει θεωρητικούς αλγορίθμους με την πρακτική εφαρμογή τους. [21]



Σχήμα 2.5: Κύκλος εκπαίδευσης ενός πράκτορα βαθιάς ενισχυτικής μάθησης. Οι εμπειρίες αποθηκεύονται στο replay buffer και χρησιμοποιούνται για την εκτίμηση πλεονεκτήματος (GAE), και την ενημέρωση των δικτύων πολιτικής/αξίας, όπως σε αλγορίθμους PPO με GAE[21].

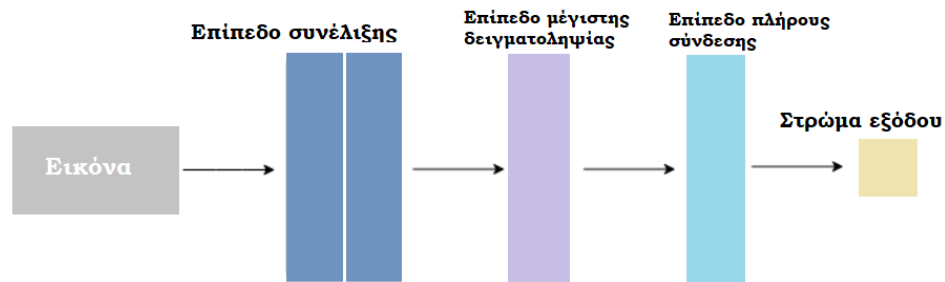
2.5 Αρχιτεκτονικά μοντέλα

Συνελκτικά νευρωνικά δίκτυα

Τα συνελκτικά νευρωνικά δίκτυα (Convolutional Neural Networks - CNN) είναι ιεραρχικά μοντέλα που βασίζονται στον τρόπο λειτουργίας του ανθρώπινου εγκεφάλου, με την αρχιτεκτονική του να είναι εμπνευσμένη από την οπτική αντίληψη. Μπορούν να εκπαιδευτούν για την εκτέλεση πολλών διαφορετικών εργασιών, όπως η ανίχνευση, η αναγνώριση και η τμηματοποίηση. Είναι μοντέλα νευρωνικών δικτύων που αποτελούνται από πολυάριθμα επίπεδα, με υψηλότερα επίπεδα

να στοιβάζονται πάνω σε χαμηλότερα για να αποτυπώνουν πιο αφηρημένες αναπαραστάσεις των δεδομένων εισόδου. Σε αντίθεση με τις συμβατικές τεχνικές εξαγωγής χαρακτηριστικών, ένα συνελκτικό νευρωνικό δίκτυο εξαλείφει την ανάγκη για χειροκίνητη εξαγωγή χαρακτηριστικών. Η δομή του αποτελείται συνήθως από ένα βήμα εξαγωγής χαρακτηριστικών που ακολουθείται από έναν ταξινομητή.

Ενώ έχει γίνει εκτεταμένη έρευνα για τη μείωση της πολυπλοκότητας των καλά μελετημένων μοντέλων συνελκτικών δικτύων με τη μορφή συμπίεσης παραμέτρων και κβαντισμού, δεν έχει υπάρξει μεγάλη προσπάθεια για την ανάπτυξη εξειδικευμένων σχεδίων για μη επανδρωμένα αεροχήματα, τα οποία απαιτούν ριζικές αλλαγές λόγω του εξαιρετικά περιορισμένου δυναμικού των πόρων για υπολογιστική ισχύ και του εύρους ζώνης.



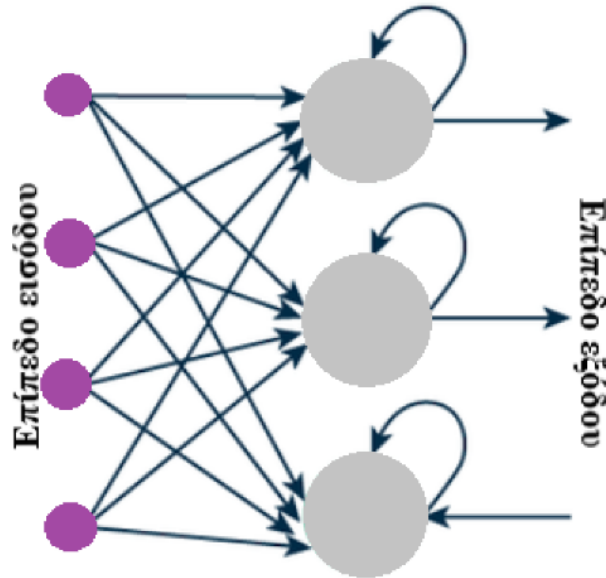
Σχήμα 2.6: Βασική αρχιτεκτονική ενός Συνελκτικού Νευρωνικού Δικτύου (CNN). Η διαδικασία ξεκινά με την εικόνα εισόδου, ακολουθείται από επίπεδα συνέλιξης (convolution) για την αυτόματη εξαγωγή χωρικών χαρακτηριστικών, συνεχίζεται με τα επίπεδα μέγιστης δειγματοληψίας (max pooling) για τη μείωση της διάστασης και καταλήγει σε πλήρως συνδεδεμένα επίπεδα για την τελική απόφαση [12].

Προκειμένου να αναλυθούν τα οπτικά ερεθίσματα από ένα μη επανδρωμένο αεροσκάφος, το Σχήμα 2.6 δείχνει τη βασική ροή δεδομένων μέσω των επιπέδων ενός συνελκτικού νευρωνικού δικτύου. Για παράδειγμα, μια εικόνα που τραβήχτηκε από την κάμερα του αεροχήματος χρησιμεύει αρχικά ως είσοδος στο δίκτυο. Προκειμένου να αναγνωριστούν σημαντικά χαρακτηριστικά όπως άκρα ή διάφορες μορφές, αυτή η εικόνα μεταφέρεται στο επίπεδο συνελκτικής επεξεργασίας, όπου εφαρμόζονται ορισμένα φίλτρα. Στη συνέχεια, το μέγιστο επίπεδο δειγματοληψίας αυξάνει τη γενίκευση, μειώνει την προσπάθεια επεξεργασίας και ελαχιστοποιεί τις πληροφορίες διατηρώντας τα πιο αντιπροσωπευτικά στοιχεία. Μετά από αυτό, τα δεδομένα μεταφέρονται μέσω του πλήρως συνδεδεμένου επιπέδου, όπου όλοι οι νευρώνες είναι πλήρως συνδεδεμένοι και πραγματοποιείται ανάλυση και λήψη αποφάσεων υψηλότερου επιπέδου. Τέλος, το επίπεδο εξόδου προσφέρει την τελευταία εκτίμηση που απαιτείται για την πλοήγηση ή τον έλεγχο του UAV, όπως η αναγνώριση περιοχής ή στόχου.

Αναδρομικά νευρωνικά δίκτυα

Τα αναδρομικά νευρωνικά δίκτυα (Recurrent Neural Networks - RNN) λειτουργούν λίγο διαφορετικά από τα κανονικά νευρωνικά δίκτυα. Σε ένα νευρωνικό δίκτυο, η πληροφορία ρέει προς μία κατεύθυνση από την είσοδο στην έξοδο. Ωστόσο, στα αναδρομικά νευρωνικά δίκτυα οι πληροφορίες επιστρέφουν στο σύστημα μετά από κάθε βήμα. Είναι σχεδιασμένα για να επεξεργάζονται

σειριακά δεδομένα και χρονικές ακολουθίες, επιτρέποντας στο δίκτυο να «θυμάται» προηγούμενες πληροφορίες τροφοδοτώντας την έξοδο μιας φάσης στην επόμενη. Αυτό βοηθά το δίκτυο να κατανοήσει το πλαίσιο των όσων έχουν ήδη συμβεί και να κάνει προβλέψεις με βάση αυτά.



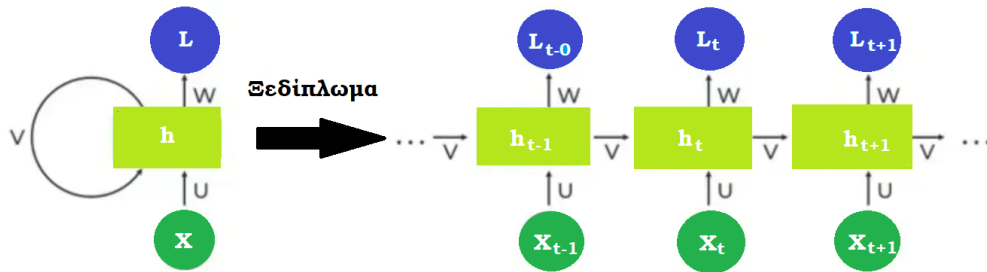
Σχήμα 2.7: Βασική αρχιτεκτονική αναδρομικού νευρωνικού δικτύου (RNN). Οι είσοδοι (μωβ κόμβοι) τροφοδοτούν τους κρυφούς νευρώνες (γκρι κόμβοι), οι οποίοι διαθέτουν αναδρομικές συνδέσεις (βρόχους). Αυτή η δομή επιτρέπει στο δίκτυο να διατηρεί «μνήμη» προηγούμενων καταστάσεων, καθιστώντας το κατάλληλο για την επεξεργασία χρονικών ακολουθιών [12].

Το Σχήμα 2.7 απεικονίζει την αρχιτεκτονική ενός επαναλαμβανόμενου νευρωνικού δικτύου. Οι επαναλαμβανόμενες συνδέσεις στα RNN, σε αντίθεση με τα τυπικά νευρωνικά δίκτυα προώθησης (feedforward), επιτρέπουν τη μεταφορά πληροφοριών από το ένα χρονικό βήμα στο επόμενο. Στα αριστερά βρίσκονται οι είσοδοι (μωβ κόμβοι), οι οποίες τροφοδοτούνται προς τα κρυφά επίπεδα (γκρι κόμβοι). Κάθε κρυφός κόμβος λαμβάνει πληροφορίες τόσο από τη δική του έξοδο στο προηγούμενο χρονικό βήμα (βρόχος) όσο και από το τρέχον βήμα. Λόγω της ικανότητάς του να «θυμάται» προηγούμενες καταστάσεις, το δίκτυο είναι ιδανικό για την ερμηνεία ακολουθιακών δεδομένων. Το RNN έχει καλύτερη απόδοση σε εφαρμογές όπως η πρόβλεψη και ο έλεγχος συστημάτων σε πραγματικό χρόνο, λόγω της επαναλαμβανόμενης δομής του, η οποία του επιτρέπει να αναγνωρίζει χρονικά μοτίβα και συσχετίσεις.

Τα αναδρομικά νευρωνικά δίκτυα μοιράζονται ομοιότητες στις δομές εισόδου και εξόδου με άλλες αρχιτεκτονικές βαθιάς μάθησης, αλλά διαφέρουν δραματικά στον τρόπο με τον οποίο οι πληροφορίες ρέουν από την είσοδο στην έξοδο. Τα βασικά στοιχεία λειτουργίας ενός τέτοιου δικτύου είναι:

1. Αναδρομική μονάδα, που κατέχει μια κρυφή κατάσταση που διατηρεί πληροφορίες σχετικά με τις προηγούμενες εισόδους σε μια ακολουθία.
2. Λειτουργία ξεδιπλώματος, όπου κατά τη διάρκεια του ξεδιπλώματος κάθε βήμα της ακολουθίας αναπαρίσταται ως ξεχωριστό επίπεδο σε μια σειρά που απεικονίζει τον τρόπο με τον

οποίο οι πληροφορίες ρέουν σε κάθε χρονικό βήμα.



Σχήμα 2.8: Η διαδικασία ξεδιπλώματος (unfolding) ενός Αναδρομικού Νευρωνικού Δικτύου στον χρόνο. Αριστερά φαίνεται η συνοπτική αναπαράσταση με τον βρόχο ανατροφοδότησης, ενώ δεξιά η ανεπτυγμένη μορφή όπου η κρυφή κατάσταση h_t εξαρτάται από την τρέχουσα είσοδο x_t και την προηγούμενη κατάσταση h_{t-1} , επιτρέποντας τη μάθηση ακολουθιών [12].

$$\begin{aligned} h(t) &= \tanh(a(t)) \\ L(t) &= c + Vh(t) \\ \hat{y}(t) &= \text{softmax}(L(t)) \end{aligned}$$

Οι τιμές των κρυφών μονάδων ορίζονται από την ακόλουθη εξίσωση :

$$h(t) = f(h(t-1), x(t); \theta)$$

Η μεταβλητή h στην εξίσωση, η οποία αντιπροσωπεύει την κατάσταση, αντιπροσωπεύει τα κρυφά στρώματα του δικτύου. Ένα επαναλαμβανόμενο δίκτυο συχνά μαθαίνει να χρησιμοποιεί την $h(t)$, ένα είδος σύννοψης με απώλειες των τμημάτων της προηγούμενης ακολουθίας εισόδων μέχρι το t που σχετίζονται με την εργασία, όταν εκπαιδεύεται να εκτελεί μια εργασία που απαιτεί την πρόβλεψη του μέλλοντος με βάση το παρελθόν. Επειδή μετατρέπει μια ακολουθία αόριστου μήκους σε έναν διάνυσμα σταθερού μήκους $h(t)$, αυτή η σύννοψη είναι συχνά με απώλειες. Ορισμένα τμήματα της προηγούμενης ακολουθίας μπορεί να διατηρούνται με μεγαλύτερη ακρίβεια σε αυτή την περίληψη από άλλα, ανάλογα με το κριτήριο εκπαίδευσης. Μπορεί να υπάρχουν περισσότερα από ένα κρυφά στρώματα στο μεσαίο στρώμα h , το καθένα με τις δικές του συναρτήσεις ενεργοποίησης και βάρη.

Transformers

Οι Transformers, γνωστοί για τον μηχανισμό προσοχής (attention mechanism) και τα πλεονεκτήματά τους στη μοντελοποίηση ακολουθιών, εφαρμόστηκαν αρχικά στον τομέα της επεξεργασίας φυσικής γλώσσας (Natural Language Processing - NLP). Οι αρχιτεκτονικές των Transformers, με την υψηλής ποιότητας σφαιρική εκμάθηση και τον αποδοτικό παράλληλο υπολογισμό τους, έχουν

αναδειχθεί ως μια ισχυρή εναλλακτική λύση στις κλασικές προσεγγίσεις ακολουθιών, όπως τα αναδρομικά νευρωνικά δίκτυα. Από την άποψη της ρομποτικής και του ελέγχου, υπερτερούν των επαναλαμβανόμενων μοντέλων στην αποτύπωση λειτουργιών μεγάλης εμβέλειας και στην ταχεία προσαρμογή στις δυναμικές αλλαγές του συστήματος.

Ένα Transformer αποτελείται από δύο μέρη: τον κωδικοποιητή (encoder) και τον αποκωδικοποιητή (decoder). Ο κωδικοποιητής επεξεργάζεται την ακολουθία εισόδου μετατρέποντάς την σε διάνυσμα, ενώ ο αποκωδικοποιητής μετατρέπει το διάνυσμα πίσω σε ακολουθία. Κάθε στρώμα κωδικοποιητή και αποκωδικοποιητή περιέχει στρώματα self-attention και feed-forward. Στον αποκωδικοποιητή, προστίθεται ένα στρώμα προσοχής κωδικοποιητή-αποκωδικοποιητή για να εστιάζει στα σχετικά bits της εισόδου.

Αρχιτεκτονική	Κύρια χαρακτηριστικά	Πλεονεκτήματα	Μειονεκτήματα
CNN (Συνελικτικά Δίκτυα)	Κατάλληλα για επεξεργασία εικόνας, video και χωρικών δεδομένων (spatial data).	Αποδοτική και αυτόματη εξαγωγή χαρακτηριστικών από εικόνες [22].	Περιορισμένη κατανόηση χρονικών συσχετίσεων και εξάρτηση από μεγάλα datasets.
RNN (Αναδρομικά Δίκτυα)	Διαχείριση ακολουθιακών δεδομένων και χρονοσειρών (time-series).	Ανάκληση παρελθοντικής πληροφορίας, κατάλληλα για πρόβλεψη τροχιάς [23].	Δύσκολη εκπαίδευση (vanishing gradient), αδυναμία διατήρησης πολύ μακρινών εξαρτήσεων.
Transformers (Μηχανισμοί Προσοχής)	Μάθηση σχέσεων από απόσταση μέσω μηχανισμών προσοχής (Self-Attention).	Καλύτερη επεξεργασία μακροχρόνιων σχέσεων, παραλληλοποίηση εκπαίδευσης [24].	Υψηλό υπολογιστικό κόστος και απαιτήσεις μνήμης, δύσκολη εφαρμογή σε edge συσκευές.

Πίνακας 2.2: Συγκριτική ανάλυση των βασικών αρχιτεκτονικών Βαθιάς Μάθησης (CNN, RNN, Transformers) που εφαρμόζονται σε UAVs, με εστίαση στα λειτουργικά τους χαρακτηριστικά και τους υπολογιστικούς περιορισμούς.

Στον πίνακα 2.2 συγκρίνονται τρεις βασικές αρχιτεκτονικές βαθιάς μάθησης, τα συνελικτικά νευρωνικά δίκτυα, τα επαναλαμβανόμενα νευρωνικά δίκτυα και οι αρχιτεκτονικές Transformer, που χρησιμοποιούνται στον έλεγχο των μη επανδρωμένων αεροσκαφών και στην επεξεργασία δεδομένων αισθητήρων. Τα συνελικτικά νευρωνικά δίκτυα είναι εξαιρετικά στην επεξεργασία οπτικών δεδομένων και δεδομένων εικόνας, καθώς χρησιμοποιούν τη χωρική συσχέτιση των εικονοστοιχείων για την αποτελεσματική εξαγωγή χαρακτηριστικών. Ωστόσο, δεν είναι ιδιαίτερα αποτελεσματικά στην κατανόηση των χρονικών συσχετίσεων. Αν και τα επαναλαμβανόμενα νευρωνικά δίκτυα αντιμετωπίζουν δυσκολίες κατά τη διάρκεια της εκπαίδευσης, έχουν σχεδιαστεί για την αξιολόγηση ακολουθιακών δεδομένων και μπορούν να διατηρήσουν προηγούμενη γνώση μέσω εσωτερικών καταστάσεων. Τα Transformers μπορούν να μάθουν αποτελεσματικά μακροπρόθεσμες εξαρτήσεις μέσω μηχανισμών αυτο-προσοχής, παρέχοντας μεγάλη ακρίβεια σε περίπλοκες συνθήκες, με τίμημα το αυξημένο υπολογιστικό κόστος. Λαμβάνοντας υπόψη τον τύπο των δεδομένων,

την ανάγκη για χρονική κατανόηση και τους διαθέσιμους υπολογιστικούς πόρους, η ανάλυση αυτή επιτρέπει την επιλογή της καλύτερης αρχιτεκτονικής με βάση τις ανάγκες κάθε αεροχήματος.

Κεφάλαιο 3

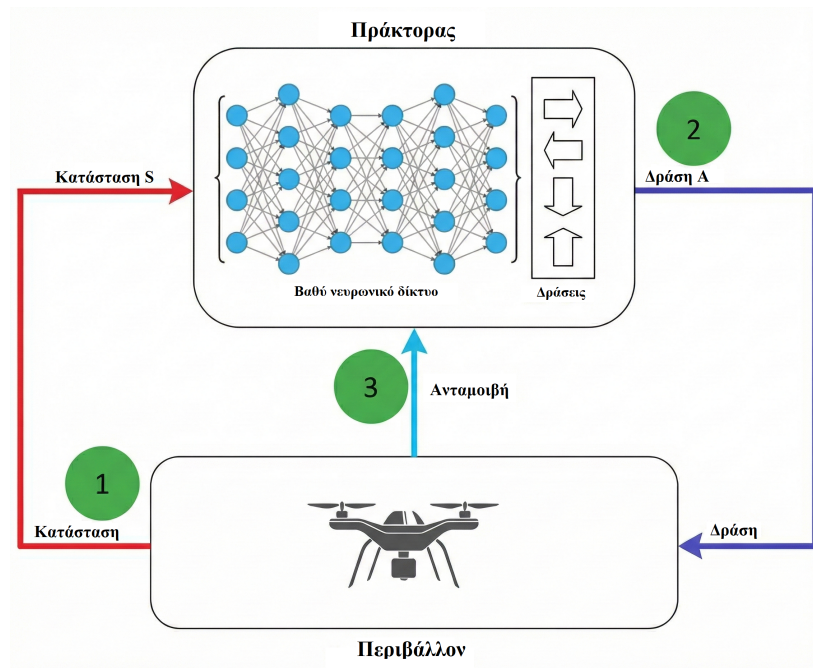
Εφαρμογές βαθιάς και ενισχυτικής μάθησης

Η πρόοδος της βαθιάς και ενισχυτικής μάθησης έχει δημιουργήσει νέες δυνατότητες για τον ευφυή έλεγχο μη επανδρωμένων αεροσκαφών. Τα μη επανδρωμένα αεροσκάφη με την αλληλεπίδραση με το περιβάλλον τους, μπορούν να προσαρμόσουν τη συμπεριφορά τους με τη βοήθεια αυτών των στρατηγικών. Αυτό καθιστά δυνατή τη δημιουργία αυτόνομων συστημάτων, εκτελώντας αποστολές με μεγαλύτερη ακρίβεια και προσαρμοστικότητα. Από την πλοήγηση και την αποφυγή εμποδίων έως και τη συνεργασία σε ομάδες και τη διαχείριση ενέργειας, οι εφαρμογές βαθιάς ενισχυτικής μάθησης αναδεικνύουν την καθοριστική συμβολή τους.

3.1 Αυτόνομη πλοήγηση και αποφυγή εμποδίων

Μια από τις πιο σημαντικές και συχνά μελετημένες χρήσεις στα μη επανδρωμένα αεροσκάφη είναι η αυτόνομη πλοήγηση και η αποφυγή εμποδίων. Χρησιμοποιώντας δεδομένα από τις κάμερες που διαθέτουν, ορισμένες έρευνες έχουν συνδυάσει τα DQNs (Deep Q Networks) με τα συνελκτικά νευρωνικά δίκτυα για να προβλέψουν την κατεύθυνση του αεροχήματος καθώς και την πιθανότητα σύγκρουσης. Η εκπαίδευση πλοήγησης διεξάγεται συνήθως σε εικονικά τρισδιάστατα περιβάλλοντα, επειδή έχουν περιορισμένους υπολογιστικούς πόρους και παροχή ενέργειας, και η αντικατάσταση εξαρτημάτων μπορεί να είναι δαπανηρή. Δεδομένου ότι οι ενδείξεις επηρεάζονται άμεσα από τον τρέχοντα εντοπισμό θέσης, και ο μελλοντικός εντοπισμός θέσης επηρεάζεται από τις τρέχουσες ενδείξεις, ο συνδυασμός των πληροφοριών από τις εικόνες και τις ενδείξεις άλλων αισθητήρων κατά τη διάρκεια της πτήσης, είναι συχνά ένα πολύπλοκο πρόβλημα για την λήψη αποφάσεων.

Για την επίλυση των ποικίλων προβλημάτων της αυτόνομης πλοήγησης χρησιμοποιούνται διάφοροι αλγόριθμοι και μεθοδολογίες βαθιάς ενισχυτικής μάθησης. Ο πράκτορας δρά ως "εμπειρογνώμονας", επιλέγοντας την πορεία δράσης που μεγιστοποιεί την ανταμοιβή.



Σχήμα 3.1: Αρχιτεκτονική βαθιάς ενισχυτικής μάθησης τύπου end-to-end για τον έλεγχο μη επανδρωμένου αεροχήματος. Το σύστημα επεξεργάζεται ακατέργαστα δεδομένα αισθητήρων υψηλής διάστασης (π.χ. εικόνες) μέσω ενός Βαθούς Νευρωνικού Δικτύου (DNN) για την άμεση παραγωγή εντολών ελέγχου, παρακάμπτοντας τη χειροκίνητη εξαγωγή χαρακτηριστικών [4].

Η θεμελιώδης αρχιτεκτονική εκπαίδευσης ενός μη επανδρωμένου αεροχήματος χρησιμοποιώντας έναν πράκτορα βαθιάς ενισχυτικής μάθησης απεικονίζεται στην παραπάνω εικόνα 3.1. Η διαδικασία περιλαμβάνει την αλληλεπίδραση μεταξύ του αεροχήματος (περιβάλλον) και του πράκτορα (Deep Agent), ο οποίος καθοδηγείται από ένα βαθύ νευρωνικό δίκτυο. Αρχικά, ο πράκτορας λαμβάνει την τρέχουσα κατάσταση του περιβάλλοντος (κατάσταση S), η οποία μπορεί να περιλαμβάνει τηλεμετρία, πληροφορίες πτήσης, εικόνες από κάμερες και δεδομένα από αισθητήρες (βήμα 1). Αφού αναλύσει αυτά τα δεδομένα, το βαθύ νευρωνικό δίκτυο του πράκτορα επιλέγει τη σωστή ενέργεια (A) από ένα εύρος πιθανών ενεργειών (όπως κατεύθυνση, ταχύτητα και υψόμετρο) (βήμα 2). Το αεροσκάφος αλλάζει κατάσταση εκτελώντας μια ενέργεια στον πραγματικό ή εικονικό κόσμο. Το αποτέλεσμα της ενέργειας αξιολογείται χρησιμοποιώντας ένα σύστημα ανταμοιβής και επιστρέφεται στον πράκτορα (βήμα 3). Ο πράκτορας μαθαίνει σταδιακά την καλύτερη πορεία δράσης για την εκτέλεση της εργασίας χρησιμοποιώντας αυτή την ανατροφοδότηση για να υποστηρίξει ή να απορρίψει τις αποφάσεις του.

Αυτή η μέθοδος επιτρέπει την εκπαίδευση του αεροχήματος για περίπλοκες εργασίες όπως ο σχηματισμός σμήνους, η αποφυγή εμποδίων, η πλοήγηση σε δύσκολα περιβάλλοντα και η παρακολούθηση χωρίς την ανάγκη χειροκίνητου προγραμματισμού. Η βαθιά ενισχυτική μάθηση είναι κατάλληλη για χρήση σε αυτόνομα αεροχήματα επειδή συνδυάζει τις δυνατότητες της βαθιάς μάθησης για την εξαγωγή χαρακτηριστικών και της ενισχυτικής μάθησης για τη λήψη αποφάσεων.

3.1.1 Έλεγχος μη επανδρωμένων αεροχημάτων

Ανάλογα με τον τύπο του UAV, γίνονται διάφορες τροποποιήσεις στη μηχανική του προκειμένου να ρυθμιστούν οι κινήσεις του στο περιβάλλον. Μπορούν να ταξινομηθούν με βάση τη μηχανική τους σε :

- Πολυρότορα (Multi-rotor)
- Σταθερής φτερωτής (Fixed-wing)
- Μονορότορα (Single rotor)
- Υβριδικά κάθετης απογείωσης και προσγείωσης (Hybrid Vertical Take-Off and Landing).



Σχήμα 3.2: Διαφορετικοί τύποι μη επανδρωμένων αεροχημάτων: (1) Πολυρότορο, (2) Σταθερού πτερυγίου, (3) Υβριδικό VTOL και (4) Μονορότορο, όπως ταξινομούνται βάσει της μηχανικής τους διαμόρφωσης [3].

Ο έλεγχος των μη επανδρωμένων αεροχημάτων είναι μια πρόκληση, καθώς λειτουργούν σε απρόβλεπτες και συχνά μεταβαλλόμενες συνθήκες, όπως ο άνεμος και οι αιφνίδιες μεταβολές φορτίου, επηρεάζοντας τη συμπεριφορά τους. Κατά τη διάρκεια της εκπαίδευσης ο πράκτορας διδάσκεται μια πολιτική που επιλέγει ενέργειες για τη διατήρηση της στάσης του αεροχήματος (κλίση, κύλιση και περιστροφή) γύρω από τις επιθυμητές τιμές, ακόμη και όταν το σύστημα «σπρώχνεται» εκτός ισορροπίας. Η αυτόνομη εκμάθηση πολιτικών ελέγχου καθίσταται εφικτή με την ενισχυτική μάθηση, όπου ένας αλγόριθμος «πειραματίζεται» με ενέργειες σε ένα πραγματικό ή προσομοιωμένο περιβάλλον και μαθαίνει την καλύτερη τακτική βάσει των ανταμοιβών. Στη συνέχεια, μέσω των κατάλληλων αλγορίθμων (π.χ. DDPG ή PPO), αποφασίζει ποια δύναμη ή ροπή να εφαρμόσει στους έλικες ώστε να σταθεροποιηθεί.

Ως λειτουργική «επίδραση» στον έλεγχο, η βαθιά μάθηση βοηθά κυρίως στην αναπαράσταση του περιβάλλοντος και στην πρόβλεψη συνθηκών. Η πρόβλεψη της μελλοντικής κατάστασης με βάση

αισθητηριακά δεδομένα είναι ιδιαίτερα σημαντική σε περιπτώσεις όπου η παραδοσιακή μοντελοποίηση είναι ανεπαρκής. Η κλιμάκωση σε πιο περίπλοκες εφαρμογές, συμπεριλαμβανομένης της συνεργασίας σμήνους ή της αποστολής σε συνθήκες χωρίς GPS, καθίσταται δυνατή μέσω της εκπαίδευσης, επεκτείνοντας σημαντικά τις δυνατότητές τους. Είναι σημαντικό να σημειωθεί ότι συχνά απαιτούνται δύο βρόχοι ελέγχου για τη λειτουργία ενός αυτόνομου αεροχήματος. Ενώ ο εξωτερικός είναι για τα σχέδια τροχιάς και τις απαιτήσεις επικοινωνίας, ο εσωτερικός είναι εξειδικευμένος στην ενεργοποίηση των ελιγμών διαχειριζόμενος τη θέση, την ταχύτητα και ούτω καθεξής. Η διαχείριση σμηνών μπορεί επίσης να χρησιμοποιήσει αυτόν τον εξωτερικό βρόχο.

3.1.2 Αποφυγή εμποδίων

Για να μπορεί ένα μη επανδρωμένο αεροσκάφος να λειτουργεί με ασφάλεια και αυτονομία σε δύσκολες συνθήκες, βασικό προαπαιτούμενο είναι να αποφεύγει εμπόδια σε πραγματικό χρόνο. Σε σύγκριση με τον βασικό προγραμματισμένο έλεγχο, ο οποίος χρησιμοποιεί προκαθορισμένες διαδρομές, τα σύγχρονα UAVs πρέπει να είναι σε θέση να χειρίζονται δυσμενείς συνθήκες. Σε αυτή την περίπτωση, η αποφυγή των εμποδίων μετατρέπεται σε ένα συνεχές ζήτημα αξιολόγησης του περιβάλλοντος της πτήσης.

Η διαδικασία προσέγγισης για την αντίληψη του χώρου που εκτελεί μια αποστολή ένα αερόχημα, και για την τροποποίηση της στρατηγικής της πτήσης, αξιοποιούνται πληροφορίες από διάφορους αισθητήρες. Η προσέγγιση αυτή μπορεί να ενισχυθεί με τη χρήση μεθόδων εκμάθησης που επιτρέπουν στο αεροσκάφος να αναπτύξει συμπεριφορά αποφυγής μέσω της εμπειρίας ή με την ενσωμάτωση μαθηματικών μοντέλων για την πρόβλεψη των κινήσεων.

Σε αυτό το σημείο, τα μοντέλα βαθιάς μάθησης τα οποία μαθαίνουν απευθείας από τις αισθητηριακές εισροές, μειώνουν την ανάγκη για χειροκίνητη μοντελοποίηση σε καταστάσεις όπου οι κρίσεις πρέπει να γίνουν γρήγορα και με ακρίβεια. Ωστόσο, στην πραγματικότητα, η δυσκολία δεν έγκειται μόνο στην αποφυγή συγκρούσεων αλλά και στη διατήρηση του τελικού στόχου. Η πλοήγηση και η σταθεροποίηση της πτήσης συνδέονται στενά με την αποφυγή εμποδίων, η οποία δεν είναι μια μεμονωμένη λειτουργία. Η πρόοδος της τεχνολογίας για την αποφυγή εμποδίων είναι καθοριστική.

3.2 Συστήματα προσγείωσης και απογείωσης

Για τα μη επανδρωμένα εναέρια οχήματα, η απογείωση και η προσγείωση είναι κρίσιμα στάδια πτήσης, επειδή απαιτούν ευαίσθητους ελιγμούς σε μικρή περιοχή και υπό μεταβαλλόμενες συνθήκες. Δεδομένου ότι τα αεροσκάφη συνεχίζουν να ενσωματώνονται σε διάφορους τομείς, η κατανόηση και η προώθηση των τεχνικών απογείωσης και προσγείωσης τόσο σε ακίνητες όσο και των κινούμενες πλατφόρμες, είναι απαραίτητη για την πλήρη αξιοποίηση των δυνατοτήτων τους σε ένα ευρύ φάσμα βιομηχανιών. Με την χρήση τεχνικών ανίχνευσης υψηλής ακρίβειας για ακριβή έλεγχο και προγραμματισμό διαδρομής, η προσγείωση πρέπει να ολοκληρωθεί σε περιορισμένο χρόνο και χώρο. Μέχρι τώρα, η ολοκλήρωση αυτού του έργου εκτελούνταν με τη χρήση χειροκίνητης ανάλυσης χαρακτηριστικών. Οι προσεγγίσεις βαθιάς και ενισχυτικής μάθησης συμβάλλουν σημαντικά στην αυτοματοποίηση αυτών των διαδικασιών, καθώς απαιτείται επακριβής αντίληψη για την λήψη αποφάσεων. Για την επιτυχή εκτέλεση αυτών των διαδικασιών, τα μη επανδρωμένα αεροσκάφη ενδέχεται να χρειάζονται διάφορα συστήματα και αλγοριθμικές τεχνικές, ανάλογα με τη φύση και την αποστολή τους. Για παράδειγμα, τα σταθερού πτερυγίου χρησιμοποιούν συμπλη-

ρωματικούς κινητήρες ή μηχανισμούς εκτόξευσης για μικρές αποστάσεις απογείωσης, και δίχτυα ή αλεξίπτωτα για προσγείωση. Από την άλλη πλευρά, η ικανότητα κάθετης απογείωσης και προσγείωσης των πολυρότορων αεροχημάτων, απαιτεί τη χρήση διαφορετικές τεχνικές βελτιστοποίησης της δυναμικής πτήσης κατά τη μετάβαση από το έδαφος στην αώρηση και αντίστροφα.

Οι προσεγγίσεις βαθιάς και ενισχυτικής μάθησης συμβάλλουν σημαντικά στην αυτοματοποίηση αυτών των διαδικασιών, καθώς απαιτείται επακριβής αντίληψη για την λήψη αποφάσεων. Με την εφαρμογή της ενισχυτικής μάθησης σε αυτό το πλαίσιο, έχουν αναπτυχθεί συστήματα με χρήση αλγορίθμων βελτιστοποίησης πολιτικής όπου η πολιτική ενισχύεται από παρατηρήσεις όπως ταχύτητα, υψόμετρο και προσανατολισμός του UAV, μεγιστοποιώντας την προβλεπόμενη συνολική ανταμοιβή. Επιπλέον, η ενσωμάτωση της βαθιάς μάθησης έχει εξεταστεί ενεργά. Για τον ακριβή προσδιορισμό της κατάλληλης θέσης απογείωσης ή προσγείωσης, χρησιμοποιούνται μοντέλα βαθιάς μάθησης, όπως τα συνελκτικά νευρωνικά δίκτυα, για την ερμηνεία των εισερχόμενων δεδομένων. Αυτό συμπεριλαμβάνεται σε μια ακόμη τεχνική πρόκληση, όσον αναφορά την διαχείριση εξωτερικών διαταραχών. Επιπλέον, οι αισθητήρες LiDAR και στερεοσκοπικής όρασης συνεργάζονται για να προσφέρουν τρισδιάστατη χαρτογράφηση του πεδίου προσγείωσης, η οποία είναι απαραίτητη και για την αποφυγή συγκρούσεων με εμπόδια.

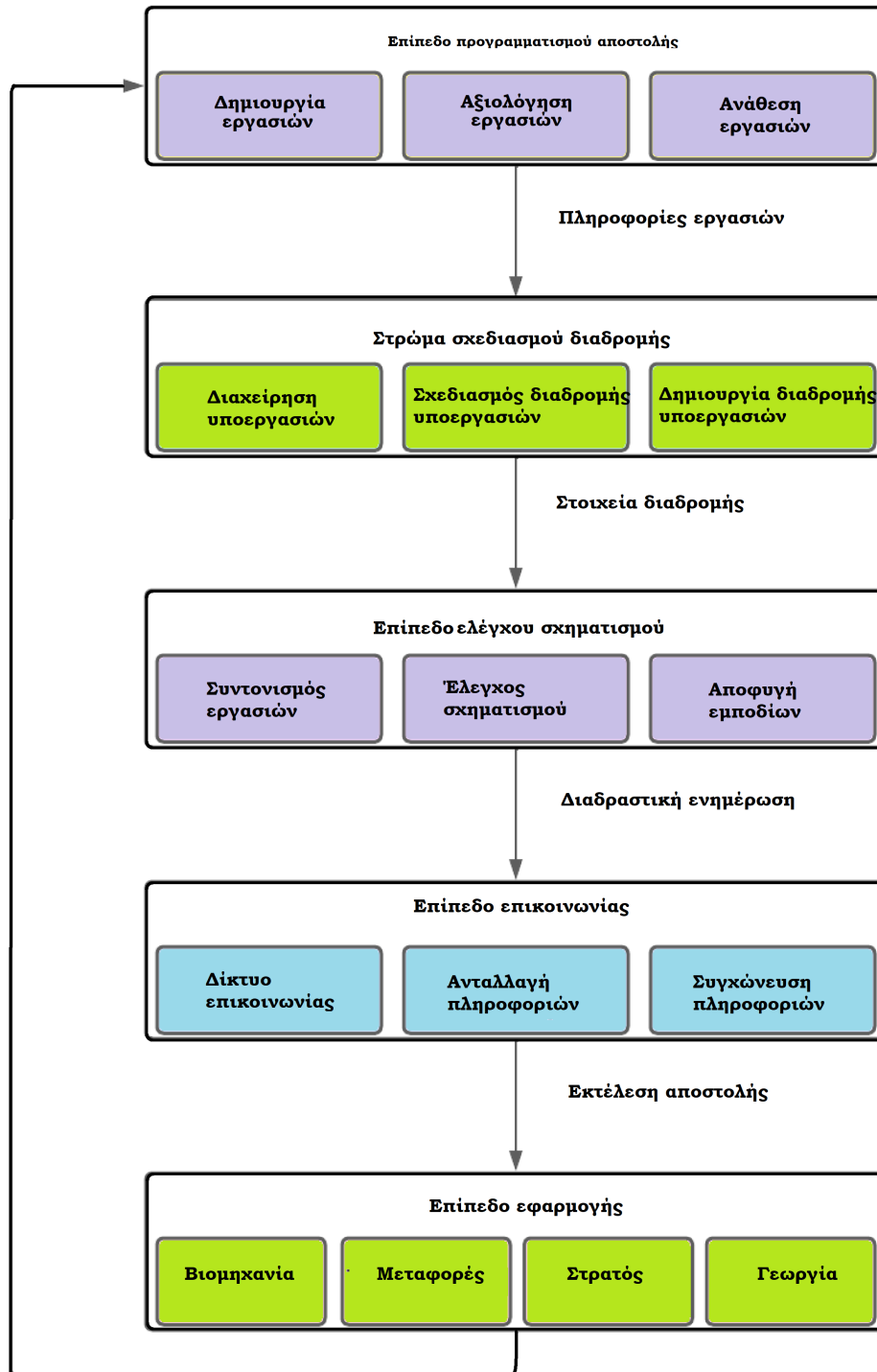
Σε αντίθεση με την βαθιά μάθηση, η βαθιά ενισχυτική μάθηση είναι μια πρόσθετη τεχνική που παρουσιάζει μεγάλες δυνατότητες. Λαμβάνοντας υπόψη την βαθιά ενισχυτική μάθηση για καθήκοντα αυτόνομης προσγείωσης μέσω οπτικών πληροφοριών, ο αλγόριθμος βαθιάς ντετερμινιστικής πολιτικής θεωρείται κατάλληλος για τον υπολογισμό των ταχυτήτων κίνησης του αεροσκάφους για την προσέγγιση του στόχου προσγείωσης. Σε επίπεδο υποδομών, η έρευνα επικεντρώνεται στη δημιουργία «έξυπνων πλατφορμών προσγείωσης», οι οποίες περιλαμβάνουν συστήματα επικοινωνίας για ακριβή καθοδήγηση, πινακίδες αναγνώρισης και συστήματα επαγωγικής φόρτισης. Τα μη επανδρωμένα αεροσκάφη μπορούν να μοιράζονται εμπειρίες απογείωσης και προσγείωσης όταν χρησιμοποιούνται τεχνικές Ομοσπονδιακής Μάθησης (Federated Learning), γεγονός που επιταχύνει τη μάθηση μέσω της συνεργατικής βελτιστοποίησης. Ο στόχος για το μέλλον είναι να ενσωματωθούν οι προαναφερθείσες τεχνολογίες σε πλήρως αυτόνομα συστήματα που θα μπορούν να πετούν και να προσγειώνονται σε άγνωστα εδάφη χωρίς ανθρώπινη επίβλεψη. Οι εφαρμογές αυτές επεκτείνουν σημαντικά τη χρησιμότητά τους σε πτήσεις άμεσης επέμβασης, μεταφοράς εμπορευμάτων ή επιτήρησης, επιτρέποντας την αυτόνομη απογείωση και προσγείωση ακόμη και σε απρόβλεπτες περιοχές.

3.3 Συνεργασία σμηνών

Λόγω της χρησιμότητας και της προσαρμοστικότητάς τους σε πραγματικές συνθήκες, τα σμήνη μη επανδρωμένων εναέριων οχημάτων έχουν αποτελέσει αντικείμενο πολυάριθμων ερευνητικών έργων τα τελευταία χρόνια. Ένα σμήνος είναι μια ομάδα από εναέρια οχήματα που συνεργάζονται για την επίτευξη ενός συγκεκριμένου στόχου. Ανάλογα με την εκάστοτε υπηρεσία, υπάρχουν αμέτρητες διαθέσιμες τακτικές σμηνών.

Στο Σχήμα 3.3 ακολουθεί μια πιο λεπτομερής επεξήγηση των στρωμάτων μιας αρχιτεκτονικής σμήνους.

- Το επίπεδο σχεδιασμού αποστολής είναι ένα επίπεδο υψηλού επιπέδου που δημιουργεί δεδομένα αποφάσεων για το επίπεδο σχεδιασμού διαδρομής και είναι υπεύθυνο για την αξιολόγηση και την κατανομή καθηκόντων σε μεμονωμένα ή και σε ομάδες.
- Το στρώμα σχεδιασμού διαδρομής ανήκει στο μεσαίο επίπεδο που επιβλέπει τις εργασίες και, χρησιμοποιώντας τα δεδομένα απόφασης, δημιουργεί κατάλληλες διαδρομές σχεδιασμού εργασιών.
- Το επίπεδο ελέγχου σχηματισμών εφαρμόζει αυτόνομη αποφυγή εμποδίων και έλεγχο σχηματισμού και συντονίζει τις εργασίες μεταξύ πολλαπλών αεροχημάτων με βάση τις πληροφορίες διαδρομής.
- Το επίπεδο επικοινωνίας υλοποιεί την απαιτούμενη ανταλλαγή πληροφοριών μεταξύ των προσώπων με τη διεξαγωγή δικτυακής επικοινωνίας με βάση τα διαδραστικά δεδομένα που παράγονται από το επίπεδο ελέγχου.
- Τέλος, το επίπεδο εφαρμογής ολοκληρώνει το βρόχο και καθιστά δυνατή την προσαρμογή στην μεταβαλλόμενη κατάσταση, τροφοδοτώντας το στρώμα σχεδιασμού διαδρομής με τα αντίστοιχα δεδομένα περιβάλλοντος βάσει των διαφόρων καταστάσεων εφαρμογής.



Σχήμα 3.3: Ιεραρχική αρχιτεκτονική διαχείρισης σμήνους, απεικονίζοντας τη ροή πληροφορίας από τον σχεδιασμό αποστολής και τροχιάς έως τον έλεγχο σχηματισμού και το δίκτυο επικοινωνίας [25], [26].

Για την αποτελεσματική διαχείριση ενός αυτόνομου σμήνους UAV απαιτείται η αρμονική συνεργασία πολλαπλών επιπέδων απόφασης και ελέγχου. Αυτή η διαδικασία οργανώνεται σε πέντε διαφορετικά λειτουργικά επίπεδα από την πολυεπίπεδη αρχιτεκτονική διαχείρισης που φαίνεται στο σχήμα. Αυτά τα επίπεδα συνεργάζονται για να εξασφαλίσουν την ασφαλή και αποτελεσματική εκτέλεση της αποστολής σε πραγματικές συνθήκες. Το υψηλότερο επίπεδο, που αναφέρεται ως επίπεδο προγραμματισμού αποστολής, είναι υπεύθυνο για τη στρατηγική διαχείριση της αποστολής. Σε αυτό το στάδιο καθορίζονται οι στόχοι, και οι εργασίες αξιολογούνται βάσει μιας σειράς κριτηρίων, συμπεριλαμβανομένου του κινδύνου, της προτεραιότητας και της διαθεσιμότητας των αεροχημάτων του σμήνους.

Το επόμενο βήμα είναι να ανατεθούν καθήκοντα στα επιμέρους μη επανδρωμένα αεροσκάφη με βάση τα χαρακτηριστικά τους. Το επιχειρησιακό πλαίσιο της αποστολής καθορίζεται σε αυτό το επίπεδο. Οι εργασίες στη συνέχεια χωρίζονται σε ξεχωριστά και διαχειρίσιμα υποέργα από το επίπεδο σχεδιασμού διαδρομής. Σε αυτό το επίπεδο πραγματοποιείται η διαχείριση των υποεργασιών με στόχο τη βελτιστοποίηση της επιχειρησιακής απόδοσης ενώ τα δρομολόγια των εναέριων οχημάτων σχεδιάζονται και αναπτύσσονται για να εγγυηθούν την ασφαλή διέλευσή τους από τον εναέριο χώρο, αποφεύγοντας κινδύνους και απρόβλεπτα εμπόδια.

Καθ' όλη τη διάρκεια της αποστολής, η συνεργασία τους συντονίζεται από το επίπεδο ελέγχου σχηματισμού. Περιλαμβάνονται ο συντονισμός εργασιών μεταξύ των μελών του σμήνους, ο έλεγχος σχηματισμού για τη διατήρηση προκαθορισμένων γεωμετρικών δομών που ενισχύουν την ασφάλεια πτήσης, η σταθερότητα και η ενεργειακή απόδοση, καθώς και η αποφυγή εμποδίων σε πραγματικό χρόνο με τη χρήση εξελιγμένων αλγορίθμων απόκρισης σε εξωτερικά ερεθίσματα.

Η υποδομή που απαιτείται για την ανταλλαγή πληροφοριών μεταξύ του σμήνους και των σταθμών ελέγχου εδάφους παρέχεται στη συνέχεια από το επίπεδο επικοινωνίας. Εκτός από την παροχή αξιόπιστης και ασφαλούς μεταφοράς δεδομένων, το δίκτυο επικοινωνιών διευκολύνει την ανταλλαγή και συγχώνευση πληροφοριών για την επιτυχή και αποτελεσματική εκτέλεση της αποστολής. Η επιτυχία και η συνοχή της αποστολής εξαρτώνται σε μεγάλο βαθμό από την αξιοπιστία αυτού του επιπέδου.

Τέλος, το επίπεδο εφαρμογής αναφέρεται στις πιθανές εφαρμογές της τεχνολογίας σμηνών. Ο στρατός χρησιμοποιεί τα μη επανδρωμένα αεροσκάφη για αποστολές υποστήριξης ή αναγνώρισης, η βιομηχανία τα χρησιμοποιεί για τη συντήρηση υποδομών και επιθεωρήσεις, οι μεταφορικές τα χρησιμοποιούν για την ταχεία διανομή ελαφρών εμπορευμάτων, και η γεωργία τα χρησιμοποιεί για την παρακολούθηση καλλιέργειών και την εφαρμογή γεωργικών τεχνικών. Η προσαρμοστική λειτουργία του σμήνους καθίσταται δυνατή από αυτή την πολυεπίπεδη μεθοδολογία, η οποία ενσωματώνει περιβαλλοντικές πληροφορίες στη λήψη αποφάσεων σε κάθε επίπεδο.

Συντονισμός και επικοινωνία σε περιβάλλον αβεβαιότητας

Ο συντονισμός και η επικοινωνία μεταξύ των μονάδων της συστάδας ενόψει της αβεβαιότητας και των περιορισμένων πόρων αποτελούν πρόκληση. Τα κατανεμημένα βαθιά νευρωνικά δίκτυα (Distributed Deep Neural Networks - DDNN) και οι αλγόριθμοι ενισχυτικής μάθησης πολλαπλών πρακτόρων (Multi-Agent RL) παρέχουν αποτελεσματικές λύσεις σε αυτή την περίπτωση. Για να λαμβάνουν αποφάσεις, η εκπαίδευση τους γίνεται με την περιορισμένη ανταλλαγή πληροφοριών με κοντινά αεροχήματα.

Υπάρχουν διάφοροι τρόποι για να οργανωθεί ένα σμήνος από μη επανδρωμένα αεροσκάφη. Αρχικά είναι τα αποκεντρωμένα σμήνη, όπου τα εναέρια οχήματα μπορούν να λαμβάνουν αποφά-

σεις μόνα τους, και τα συγκεντρωτικά σμήνη, όπου ένας μόνο ελεγκτής σμήνους έχει τον πλήρη έλεγχο των αεροχημάτων. Στις συγκεντρωτικές δομές ελέγχου μια μοναδική πλατφόρμα επιβλέπει άμεσα ολόκληρο το σμήνος, εκτελώντας όλους τους υπολογισμούς που απαιτούνται για την εκτέλεση των καθηκόντων. Εξαιτίας αυτού, η συγκεντρωτική αρχιτεκτονική είναι ουσιαστικά ο πιο απλός τύπος συστήματος ελέγχου. Αυτή η αρχιτεκτονική χρησιμοποιεί έναν ελεγκτή σμήνους για τη λήψη αποφάσεων αποστολής υψηλού επιπέδου. Ακόμη και σε κεντροποιημένες αρχιτεκτονικές, τα αεροσκάφη πρέπει συνήθως να λαμβάνουν αποφάσεις αυτόνομα κατά τη λειτουργία τους, και οι αποφάσεις αυτές έχουν συνήθως να κάνουν με καθήκοντα ελέγχου χαμηλότερου επιπέδου. Τα συγκεντρωτικά συστήματα έχουν το μειονέκτημα ότι εξαρτώνται σε μεγάλο βαθμό από τη διαθεσιμότητα ασφαλών καναλιών επικοινωνίας, γεγονός που οδηγεί σε μεγάλο όγκο κίνησης του δικτύου επικοινωνιών περιορίζοντας την επεκτασιμότητα.

3.3.1 Ενισχυτική μάθηση πολλαπλών πρακτόρων

Η ενισχυτική μάθηση πολλαπλών πρακτόρων (Multi-Agent Reinforcement Learning - MARL) έχει αναδειχθεί βασικό στοιχείο για τη διευκόλυνση των συνεργατικών επιχειρήσεων μη επανδρωμένων αεροσκαφών, ειδικά όταν πρόκειται για αποτελεσματικότητα του σμήνους, κατανομή εργασιών και αποφυγή συγκρούσεων. Σε αυτές τις καταστάσεις, κάθε αεροσκάφος λειτουργεί ως πράκτορας που έχει εκπαιδευτεί είτε μόνος του είτε σύμφωνα με μια κεντρικά σχεδιασμένη στρατηγική. Τεχνικές όπως οι παραλλαγές που βασίζονται σε αλγορίθμους βελτιστοποίησης πολιτικής με κεντρική - αποκεντρωμένη εκτέλεση (Centralized Training – Decentralized Execution - CTDE), και DDPG πολλαπλών πρακτόρων, επιτρέπουν τη χρήση πρακτικών τακτικών που λαμβάνουν υπόψη τόσο τα ατομικά όσο και τα συλλογικά συμφέροντα. Αυτές οι μέθοδοι αποφεύγουν το πρόβλημα της μη-στατικής δυναμικής, το οποίο προκύπτει όταν οι πολιτικές όλων των παραγόντων αλλάζουν ταυτόχρονα, και επιτρέπουν στα εναέρια οχήματα να συντονίζουν τις δραστηριότητές τους με ελάχιστη ανταλλαγή πληροφοριών. Σε αντίθεση με τα κεντροποιημένα συστήματα ελέγχου, ένα άλλο τεχνολογικό πλεονέκτημα είναι ότι η πολυπλοκότητα αυξάνεται εκθετικά με το μέγεθος του στόλου. Έτσι, οι τεχνικές ενισχυτικής μάθησης πολλαπλών πρακτόρων επιτρέπουν τη σταδιακή εκπαίδευση και γενίκευση πολιτικών που μπορούν να εφαρμοστούν σε μη επανδρωμένα αεροσκάφη με διαφορετικούς ρόλους και δυνατότητες. Μια άλλη σημαντική προσθήκη είναι οι προσεγγίσεις των νευρωνικών δικτύων γραφημάτων (Graph Neural Networks - GNN), οι οποίες επιτρέπουν την αναπαράσταση της τοπολογίας ενός σμήνους ως γράφημα, βελτιώνοντας την ικανότητα των πρακτόρων για αποτελεσματική ανταλλαγή πληροφοριών. Παρά τα δυναμικά εμπόδια ή την απώλεια σύνδεσης, αυτή η σύμπραξη έχει εφαρμογές σε αποστολές παρακολούθησης, επιτήρησης και αναγνώρισης περιοχής.

Η αυξημένη αυτονομία του σμήνους είναι δυνατή με τις αποκεντρωμένες δομές ελέγχου. εάν οι υπολογισμοί εκτελούνται με τη χρήση ενσωματωμένων ελεγκτών πτήσης. Τα αποκεντρωμένα συστήματα έχουν γενικά το πλεονέκτημα ότι είναι κλιμακούμενα και εξαρτώνται λιγότερο από τη σταθερότητα των συνδέσεων. Διακρίνονται δύο τύποι αποκεντρωμένων αρχιτεκτονικών: οι πλήρως καταναμημένες και οι ιεραρχικές. Τα ιεραρχικά συστήματα μπορούν θεωρούνται ως μια επέκταση των συγκεντρωτικών συστημάτων στα οποία δίνεται μια σχετική αυτονομία λήψης αποφάσεων. Στα ιεραρχικά συστήματα, η δομή της ακολουθίας των εντολών έχει τη μορφή δέντρου, στο οποίο ένας ή περισσότεροι ηγέτες είναι υπεύθυνοι για μια ομάδα, επίσης γνωστοί ως ακόλουθοι, οι οποίοι έχουν αυτονομία αλλά υπόκεινται στις οδηγίες του ηγέτη. Στα πλήρως καταναμημένα σχέδια, δεν υπάρχει αρχηγός σμήνους. Αντ'αυτού, όλοι οι πράκτορες είναι αυτόνομοι με ίδιο επίπεδο ευφυΐας.

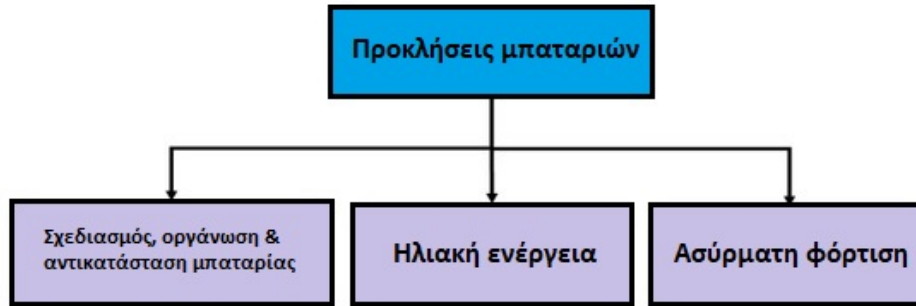
Εν κατακλείδι, μέθοδοι όπως αποκεντρωμένοι αλγόριθμοι βελτίωσης πολιτικών, νευρωνικά δίκτυα γραφημάτων και μηχανισμοί προσοχής, έχουν χρησιμοποιηθεί αποτελεσματικά για τον καταμερισμό της εργασίας, τον δυναμικό σχηματισμό σημνών και αποφυγή συγκρούσεων μεταξύ μονάδων.

3.4 Βελτιστοποίηση ενεργειακής κατανάλωσης

Η κύρια χρήση των μη επανδρωμένων εναέριων οχημάτων είναι η βραχυπρόθεσμη παράδοση και μετάδοση σημάτων για δημόσιες και εμπορικές εφαρμογές. Προκειμένου να αντιμετωπιστεί η ζήτηση για ασφαλείς και έξυπνες εναέριες μεταφορές, τα συνεργαζόμενα αεροσκάφη αναμένεται να επενδύσουν περισσότερη ενέργεια στην επικοινωνία και τη συνεργασία σμήνους λόγω της αυξανόμενης χρήσης τους και της ανάπτυξης έξυπνων πόλεων. Για την αποδοτική χρήση τους η ενεργειακή απόδοση είναι απαραίτητη, ιδίως για αποστολές μεγάλης διάρκειας. Ως εκ τούτου, είναι απαραίτητη η ρύθμιση της κατανάλωσης ενέργειας κατά τη λειτουργία τους.

Η ανθεκτικότητα τους καθορίζεται σε μεγάλο βαθμό από τις μπαταρίες, οι οποίες παρέχουν επίσης μια πηγή ενέργειας και χρήσιμη ενέργεια. Επίσης, η διάρκεια των πτήσεων μπορεί επίσης να περιορίζεται από εμπόδια, την αντίσταση του ανέμου και το ωφέλιμο φορτίο. Άλλες έρευνες έδειξαν ότι με τη χρήση φωτοβολταϊκού συστήματος διαχείρισης ηλεκτρικής ενέργειας, το οποίο προσφέρει πρόσθετη ηλεκτρική ενέργεια από την μπαταρία, αύξησε τη διάρκεια λειτουργίας. Κατόπιν αυτού, αποφασίστηκε να στραφεί η προσοχή των ερευνητών στην ανάπτυξη λογισμικού. Εάν οι εξωτερικοί παράγοντες παραμένουν σταθεροί, η ενισχυτική μάθηση μπορεί να είναι ένας από τους καλύτερους τρόπους για την αύξηση της ενεργειακής απόδοσης. Το αεροσκάφος μπορεί να εκπαιδευτεί μέσω δοκιμής και σφάλματος για ποιες ενέργειες, όπως η επιλογή ταχύτητας, κατεύθυνσης, ύψους ή τρόπου επιτάχυνσης, οδηγούν σε μείωση της συνολικής κατανάλωσης ενέργειας, ενώ παράλληλα ολοκληρώνεται η αποστολή, μοντελοποιώντας το θέμα ως μια μαρκοβιανή διαδικασία απόφασης. Σε αυτή την κατάσταση έχουν εφαρμοστεί αποτελεσματικά αλγόριθμοι όπως οι DDPG, PPO και SAC, οι οποίοι εκμεταλλεύονται το συνεχές χώρο ενεργειών και ενσωματώνουν περιορισμούς αποδοτικότητας ή ασφάλειας.

Η εκπαίδευση τέτοιων μοντέλων είναι ασφαλής και επαναλαμβανόμενη όταν γίνεται προσομοίωση με πραγματικά σενάρια πτήσης, είτε με φυσικά μοντέλα είτε με προσομοιωτές. Η μεταφορά της πολιτικής από το εικονικό περιβάλλον στο πραγματικό (sim-to-real transfer), αν και παραμένει μια ερευνητική πρόκληση, σημειώνει σημαντική πρόοδο.



Σχήμα 3.4: Σχηματική ταξινόμηση των στρατηγικών αντιμετώπισης των ενεργειακών περιορισμών στα μη επανδρωμένα αεροχήματα, συμπεριλαμβανομένων των τεχνικών σχεδιασμού/αντικατάστασης, της εκμετάλλευσης ηλιακής ενέργειας και της ασύρματης φόρτισης [27].

Στην εικόνα 3.4 γίνεται επίδειξη κάποιων προτεινόμενων λύσεων. Η χωρητικότητα της μπαταρίας ενός μη επανδρωμένου αεροχήματος θεωρείται κρίσιμος παράγοντας που επηρεάζει τον χρόνο που απαιτείται για τις αποστολές. Όταν αυξάνεται η χωρητικότητα της μπαταρίας, αυξάνεται και το βάρος της. Επομένως, το αεροσκάφος καταναλώνει περισσότερη ενέργεια στις αποστολές.

1. Ο σχεδιασμός, ο προγραμματισμός και η αντικατάσταση της μπαταρίας αποτελούν μέρος της διαχείρισης της μπαταρίας, προκειμένου να ολοκληρωθούν οι απαραίτητες αποστολές. Η κατανομή και ο προγραμματισμός της μπαταρίας μπορούν να εφαρμοστούν για να ελαχιστοποιηθεί η μείωση της διάρκειας ζωής της μπαταρίας. Η εναλλαγή συστημάτων μπαταριών είναι μια εξαιρετική λύση για τα αυτόνομα αεροχήματα. Αυτά τα συστήματα περιλαμβάνουν έναν μικροελεγκτή, πλατφόρμα προσγείωσης, αποθήκευση μπαταρίας, και φορτιστή μπαταρίας. Για να αποφευχθεί η απώλεια δεδομένων, πρώτα πραγματοποιείται η διαδικασία ανταλλαγής και στη συνέχεια η αντικατάσταση της μπαταρίας. Μετά από αυτό, το αεροσκάφος αποσυνδέεται από την πηγή ενέργειας για να συνεχίσει την αποστολή του.
2. Τα αεροχήματα που λειτουργούν με ηλιακή ενέργεια μπορούν να πετούν σε μεγάλα υψόμετρα για παρατεταμένες χρονικές περιόδους, με την ηλιακή ενέργεια να χρησιμεύει ως ο κύριος μηχανισμός ώθησης.
3. Η ασύρματη φόρτιση έχει παρουσιαστεί από μερικούς ερευνητές στη βιβλιογραφία. Μια προτεινόμενη λύση είναι η παροχή ενός αυτόματου συστήματος σταθμών φόρτισης κατανομής κατά μήκος της διαδρομής που ακολουθεί το αεροσκάφος. Τέσσερα εξαρτήματα συνθέτουν αυτό το σύστημα σταθμών: μια μπαταρία, ένας μετατροπέας ισχύος, ένα ασύρματο pad φόρτισης και ένα ηλιακό πάνελ.

3.4.1 Βελτιστοποίηση μεταφοράς ωφέλιμου φορτίου

Με την πρόοδο της τεχνολογίας των αεροχημάτων, η σημασία της χωρητικότητας του ωφέλιμου φορτίου αυξάνεται. Η ανάγκη για μη επανδρωμένα αεροσκάφη που μπορούν να μεταφέρουν μεγαλύτερο ωφέλιμο φορτίο έχει αυξηθεί λόγω της ανάπτυξης προηγμένων αισθητήρων, καμερών και άλλων συστημάτων ωφέλιμου φορτίου. Κατά συνέπεια, η ακριβής αξιολόγηση και σύγκριση της χωρητικότητας φορτίου διαφόρων μοντέλων είναι εξαιρετικά σημαντική.

Με τη χρήση ενισχυτικής μάθησης και ειδικότερα αλγορίθμων όπως οι PPO και DDPG, μπορούν να δημιουργηθούν πολιτικές πτήσης που προσαρμόζονται στην τοπολογία του ωφέλιμου φορτίου και τη στόχευση χαμηλής κατανάλωσης. Σε προχωρημένες περιπτώσεις, παρατηρείται συνδυασμός με βαθιάς μάθησης με συνελκτικά νευρωνικά δίκτυα για οπτική είσοδο για παράδειγμα, για να μπορούν τα αεροσκάφη να αντιλαμβάνονται την κατανομή του φορτίου σε πραγματικό χρόνο και να τροποποιούν τη συμπεριφορά τους.

Ο συγκριτικός πίνακας 3.1 που ακολουθεί παρέχει μια επισκόπηση των κύριων εφαρμογών της βαθιάς μάθησης και της ενισχυτικής μάθησης. Καλύπτουν ένα ευρύ φάσμα, από την απλή απογείωση και σταθεροποίηση έως τη συνεργασία πολλών αεροσκαφών. Συνοψίζει βασικές περιοχές εφαρμογής, τους τύπους μάθησης και τους σχετικούς αλγορίθμους, καθώς και ενδείξεις για την εφαρμογή τους σε πραγματικά συστήματα.

Εφαρμογή UAV	Τύπος μάθησης	Αλγόριθμοι	Περιγραφή / Σκοπός	Εφαρμογή σε πραγματικά σενάρια
Γενικός έλεγχος αεροχημάτων	RL & DL	DDPG, PPO, CNN	Σταθερή και ακριβής πτήση σε δυναμικά περιβάλλοντα [21], [28].	Ναι
Συστήματα απογείωσης, προσγείωσης	RL & DL	PPO, SAC	Ομαλή και ασφαλής απογείωση ή προσγείωση σε περιορισμένους χώρους [29], [30].	Ναι (σε logistics)
Συνεργασία σμηνών	Ενισχυτική μάθηση πολλαπλών πρακτόρων	MADDPG, GNN	Συντονισμός UAVs, συνεργασία και κάλυψη μεγάλης περιοχής [31].	Περιορισμένη (κυρίως σε έρευνα)
Βελτιστοποίηση ενεργειακής κατανάλωσης	Ενισχυτική μάθηση	PPO, DQN, SAC	Μείωση κατανάλωσης ενέργειας και βελτιστοποίηση διαδρομών πτήσης [32], [33].	Ναι

Πίνακας 3.1: Συγκριτική ταξινόμηση αλγορίθμων μηχανικής και ενισχυτικής μάθησης ανά πεδίο εφαρμογής σε μη επανδρωμένα αεροχήματα, με έμφαση στη λειτουργική ετοιμότητα σε πραγματικά σενάρια.

Όπως φαίνεται, οι τεχνικές ενισχυτικής και βαθιάς μάθησης εφαρμόζονται σε ποικίλες λειτουργίες των μη επανδρωμένων εναέριων οχημάτων, ανάλογα με τις απαιτήσεις κάθε περίπτωσης. Ο συνδυασμός αυτών των μεθόδων επιτρέπει ευέλικτα και αποδοτικά συστήματα, ειδικά σε σύνθετα περιβάλλοντα όπου η παραδοσιακή ρύθμιση ελέγχου αποτυγχάνει.

3.5 Περιβάλλοντα προσομοίωσης και εκπαίδευσης

Η διεξαγωγή πραγματικών πειραμάτων στα συστήματα που βασίζονται σε μη επανδρωμένα εναέρια οχήματα και ιδιαίτερα στα συστήματα σμηνών μπορεί να είναι εξαιρετικά δαπανηρή και περίπλοκη. Προσομοιωτές, εξομοιωτές και κατάλληλα πλαίσια χρησιμοποιούνται για να καλύψουν αυτό το κενό, προσφέροντας ταχεία και οικονομικά αποδοτική αξιολόγηση της απόδοσης και της λειτουργικής δυνατότητας τους. Η ανάπτυξη αποτελεσματικών πολιτικών πτήσης μέσω της ενισχυτικής μάθησης απαιτεί ρεαλιστικά και κλιμακούμενα περιβάλλοντα προσομοίωσης. Πλατφόρμες όπως οι AirSim (της Microsoft), Gazebo και DroneGym παρέχουν τη δυνατότητα ρεαλιστικής φυσικής προσομοίωσης, επιτρέποντας παράλληλη εκπαίδευση αλγορίθμων ενισχυτικής μάθησης. Όταν συνδυάζονται με τους αλγόριθμους DQN, PPO, A3C και SAC, έχουν δείξει εξαιρετική απόδοση σε εργασίες όπως η αποφυγή εμποδίων, η πλοήγηση και η συνεργασία πολλαπλών αεροχημάτων. Επιπλέον, για την επίλυση της λεγόμενης μετάβασης από την προσομοίωση στην πραγματικότητα, αντιμετωπίζεται όλο και πιο αποτελεσματικά μέσω fine-tuning, κάτι που επιτρέπει την ομαλή μετάβαση των πολιτικών που έχουν εκπαιδευτεί σε προσομοίωση στον πραγματικό κόσμο.

Ωστόσο, η εφαρμογή της στην ερευνητική διαδικασία πρέπει να διακατέχεται από ακριβώς καθορισμένους παραμέτρους, ώστε να διασφαλίζεται η συγκρισιμότητα και η αξιοπιστία των αποτελεσμάτων. Σε αυτό το σημείο, ένα βασικό στοιχείο είναι η αναπαραγωγιμότητα. Οι ασυμβατότητες που συχνά προκύπτουν από τροποποιήσεις στη φυσική ή στο rendering (διαδικασία παραγωγής οπτικής αναπαράστασης) μπορούν να αποφευχθούν με τη ρητή αναφορά στην έκδοση του προσομοιωτή. Η ακριβής αναπαραγωγή της αρχικής κατάστασης (π.χ. κατανομή εμποδίων, ανάπτυξη αεροχήματος) είναι εγγυημένη όταν χρησιμοποιούνται συγκεκριμένα seed values για τυχαίους αριθμούς. Η σύγκριση με άλλες έρευνες είναι δυνατή χάρη στην αναλυτική εξήγηση του αριθμού των επεισοδίων, της διάρκειας κάθε δοκιμής και των μεμονομένων παραγόντων (όπως η ταχύτητα του ανέμου). Δεδομένου ότι ακόμη και μικρές προσαρμογές των παραμέτρων μπορεί να οδηγήσουν σε δραστικά διαφορετικές πολιτικές ελέγχου, ο σχεδιασμός της συνάρτησης ανταμοιβής πρέπει να δημοσιοποιείται στο κοινό. Ταυτόχρονα, τα μοντέλα αντίληψης εκπαιδεύονται ή επικυρώνονται χρησιμοποιώντας πραγματικά σύνολα δεδομένων πτήσεων, τα οποία αποτελούν ουσιαστικό συστατικό της μεταφοράς από το προσομοιωμένο στο πραγματικό περιβάλλον. Σε «ιδανικά» ειδικά περιβάλλοντα, η ενσωμάτωση προσομοιωμένων και πραγματικών δεδομένων βελτιώνει τη γενίκευση και μειώνει το πρόβλημα της υπερεκπαίδευσης.

Εργαλείο / Dataset	Βασικά χαρακτηριστικά	Παράμετροι αναπαραγωγιμότητας	Ενδεικτική χρήση
AirSim (v1.6)	Φυσική υψηλής πιστότητας, υποστήριξη αισθητήρων, ενσωμάτωση του Unreal Engine	Seed values, weather profiles, αριθμός επεισοδίων, reward shaping scripts	Εκπαίδευση βαθιάς ενισχυτικής μάθησης για αποφυγή εμποδίων σε πολυκόπτερα [34]
Gazebo (v11)	Συμβατότητα με ROS, modular plugins, 3D περιβάλλοντα	Σταθερές εκδόσεις παγκόσμιων αρχείων, fixed seeds, ρύθμιση παραμέτρων φυσικής	Έλεγχος σχηματισμών (formation control) σε σμήνη UAVs [35]
UAV123 Dataset	123 ακολουθίες αρχείων, πραγματικές πτήσεις αεροχημάτων	Annotation consistency, σταθερή κατανομή εκπαίδευσης	Ανίχνευση και παρακολούθηση στόχων [12]
VisDrone Dataset	>10.000 εικόνες, αστικά περιβάλλοντα	Σταθερά σημεία αναφοράς, αναπαραγώγιμες μετρικές	Ανίχνευση αντικειμένων σε πόλεις [12]
Blackbird Dataset	Πτήσεις υψηλής ταχύτητας σε εσωτερικούς χώρους με καταγραφή κινήσεων	Επαναλήψιμες τροχιές, δεδομένα βαθμονόμησης	Έρευνα για sim-to-real transfer σε DRL ελεγκτές [36]

Πίνακας 3.2: Συγκριτική επισκόπηση περιβαλλόντων προσομοίωσης και συνόλων δεδομένων (datasets) που χρησιμοποιούνται στην εκπαίδευση UAV, εστιάζοντας στις παραμέτρους που διασφαλίζουν την πειραματική αναπαραγωγιμότητα.

Κεφάλαιο 4

Σύγκριση και αξιολόγηση των μεθόδων

Οι ραγδαίες εξελίξεις στους κλάδους της ενισχυτικής μάθησης και της βαθιάς μάθησης έχουν αλλάξει σημαντικά τον τρόπο ελέγχου των μη επανδρωμένων αεροσκαφών, σε αντίθεση με τις παραδοσιακές προσεγγίσεις ελέγχου που απαιτούν αυστηρή μοντελοποίηση. Αυτό το κεφάλαιο στοχεύει να τονίσει τα πλεονεκτήματα, τα μειονεκτήματα και τη γενική εφαρμοστικότητα των στρατηγικών που περιγράφηκαν προηγουμένως για εφαρμογή στον έλεγχο εναέριων οχημάτων, συγκρίνοντάς τις και αξιολογώντας τις μεθοδικά. Οι αλγόριθμοι βαθιάς ενισχυτικής μάθησης συγκαταλέγονται στις τεχνικές που μελετώνται, μαζί με υβριδικές στρατηγικές που χρησιμοποιούν την βαθιά μάθηση για την αναπαράσταση του χώρου καταστάσεων ή την αντίληψη. Τα οφέλη και τα μειονεκτήματα κάθε στρατηγικής εξετάζονται επίσης συγκρίνοντάς τις με πιο συμβατικές μεθόδους ελέγχου, συμπεριλαμβανομένων των ελεγκτών PID. Μέσα από αυτή τη συγκριτική μελέτη, οι στόχοι είναι η κατανόηση του πότε και πού οι τεχνολογίες μηχανικής μάθησης αποδίδουν καλύτερα στον έλεγχο των μη επανδρωμένων αεροσκαφών, καθώς και ο εντοπισμός καταστάσεων όπου οι συμβατικές προσεγγίσεις είναι ακόμα πιο αποτελεσματικές ή αξιόπιστες.

4.1 Ανασκόπηση υφιστάμενων προσεγγίσεων στον έλεγχο μη επανδρωμένων αεροχημάτων

Ένα τυπικό σύστημα βαθιάς ενισχυτικής μάθησης αποτελείται από έναν αριθμό διακριτών υποσυστημάτων που αλληλεπιδρούν στενά μεταξύ τους, επιτρέποντας σε έναν πράκτορα να μπορεί να μάθει από την εμπειρία χρησιμοποιώντας βαθιά νευρωνικά δίκτυα για να προσεγγίσει σύνθετες συναρτήσεις. Σε αντίθεση με τις συμβατικές τεχνικές εκμάθησης, τα συστήματα βαθιάς ενισχυτικής μάθησης χρησιμοποιούν την ικανότητα από τα βαθιά νευρωνικά δίκτυα να διαχειρίζονται συνεχείς ή πολύπλοκες καταστάσεις και εισόδους υψηλών διαστάσεων. Τα προκαθορισμένα δυναμικά μοντέλα αποτέλεσαν τη βάση για την παραδοσιακή προσέγγιση στον έλεγχο, ενώ ο συνδυασμός των προσεγγίσεων της βαθιάς και της ενισχυτικής μάθησης έχει συγκεντρώσει μεγάλη προσοχή από τους ερευνητές. Οι τεχνικές που βασίζονται στη βαθιά ενισχυτική μάθηση υιοθετήθηκαν σταδιακά ως αποτέλεσμα της απαίτησης για ευέλικτους, προσαρμοστικούς και αυτόνομους πράκτορες.

Με έμφαση στη γενίκευση και τη μάθηση μέσω επαφής με το περιβάλλον, οι υπάρχουσες τεχνικές κυμαίνονται ευρέως, από απλούς αλγόριθμους πλοήγησης μέχρι πολυεπίπεδο έλεγχο σε περίπλοκες καταστάσεις. Μεταξύ των τεχνικών βαθιάς ενισχυτικής μάθησης, αλγόριθμοι όπως αυτός της βελτιστοποίησης πολιτικής και τα Deep Q-Networks έχουν χρησιμοποιηθεί σε προσομοιωμένα

και ελεγχόμενα περιβάλλοντα για την ανάπτυξη πολιτικών αποφυγής εμποδίων και σταθεροποίησης πτήσης, με αξιολογικά αποτελέσματα. Παρόμοια με αυτό, αλγόριθμοι συνεχούς δράσης όπως οι DDPG και SAC έχουν χρησιμοποιηθεί για λεπτομερή έλεγχο κίνησης σε δυναμικές καταστάσεις, παρέχοντας καλύτερη απόδοση σε εφαρμογές που απαιτούν ακριβείς ελιγμούς. Επιπλέον, η ενισχυτική μάθηση πολλαπλών στόχων ενσωματώνεται σε πολιτικές που λαμβάνουν υπόψη περιορισμούς όπως η κατανάλωση ενέργειας και η ασφάλεια πτήσεων, εκτός από την επίτευξη στόχων.

Σε πιο σύνθετες εφαρμογές, παρατηρούνται επίσης υβριδικά συστήματα που συνδυάζουν την βαθιά μάθηση με συμβατικά συστήματα ελέγχου, στα οποία τα νευρωνικά δίκτυα εκτελούν προγνωστικό έλεγχο ή αντισταθμιστικές λειτουργίες για τη βελτίωση της σταθερότητας και της ευελιξίας. Προκειμένου να ξεπεραστούν ζητήματα όπως η αποκεντρωμένη εκπαίδευση, η επίλυση συγκρούσεων και η κοινή εκμάθηση πολιτικών, η ερευνητική κοινότητα έχει καταφύγει επίσης σε σενάρια πολλαπλών πρακτόρων για τον συντονισμό σμηνών αεροχημάτων. Οι συνθήκες για μια πιο συστηματική αξιολόγηση και σύγκριση των προσεγγίσεων, δημιουργούνται πλέον από την κίνηση προς ασφαλέστερες και πιο κατανοητές πολιτικές.

4.2 Πλεονεκτήματα και μειονεκτήματα της ενισχυτικής και βαθιάς μάθησης

Οι τεχνικές βαθιάς μάθησης και ενισχυτικής μάθησης έχουν μοναδικά πλεονεκτήματα και μειονεκτήματα που θα πρέπει να λαμβάνονται υπόψη κατά την ανάπτυξη συστημάτων μη επανδρωμένων αεροσκαφών, όπως ακριβώς κάθε άλλη στρατηγική τεχνητής νοημοσύνης. Αυτές οι στρατηγικές ξεπερνούν τους περιορισμούς των συμβατικών προσεγγίσεων, ιδιαίτερα όταν αντιμετωπίζουν αβέβαιες καταστάσεις ή μεταβαλλόμενες ανάγκες. Ωστόσο, έχουν μειονεκτήματα, όπως σημαντικό κόστος υπολογιστικής ισχύος, την ανάγκη για εντατική εκπαίδευση και προβλήματα ερμηνευσιμότητας. Μια σύγκριση των κύριων πλεονεκτημάτων και μειονεκτημάτων τους αναλύεται παρακάτω.

Πλεονεκτήματα

- **Αυτονομία και προσαρμοστικότητα:** Επειδή οι πράκτορες της ενισχυτικής μάθησης έχουν σχεδιαστεί για να εκπαιδεύονται αλληλεπιδρώντας με το περιβάλλον τους, τα αεροσκάφη μπορούν να προσαρμοστούν σε μεταβαλλόμενες συνθήκες στις οποίες οι συμβατικές τεχνικές ελέγχου είναι αναποτελεσματικές. Για παράδειγμα, οι πράκτορες μπορούν να προσαρμόσουν άμεσα την προσέγγισή τους σε καταστάσεις με μεταβαλλόμενα εμπόδια ή άστατο καιρό.
- **Βελτιστοποίηση απόδοσης και χειρισμός πολύπλοκων καταστάσεων:** Οι πολιτικές ενισχυτικής μάθησης βελτιστοποιούν τη συνολική απόδοση μακροπρόθεσμα, σε αντίθεση με τα παραδοσιακά συστήματα ελέγχου που στοχεύουν σε άμεσες βελτιστοποιήσεις. Η βαθιά μάθηση, όταν χρησιμοποιείται σε συνδυασμό με την ενισχυτική μάθηση, χρησιμοποιεί αρχιτεκτονικές όπως τα CNN, RNN και Transformers για να εξάγει σημαντικά χαρακτηριστικά από αισθητηριακά δεδομένα. Αυτό εξαλείφει την ανάγκη για χειροκίνητη μοντελοποίηση του περιβάλλοντος και επιτρέπει την αντιμετώπιση των προβλημάτων πλοήγησης.

Μειονεκτήματα

- **Δυσκολία στην γενίκευση:** Όταν η κατανομή των γεγονότων μετά την ανάπτυξη είναι διαφορετική από εκείνη κατά τη διάρκεια της εκπαίδευσης, τα συστήματα βαθιάς ενισχυτικής μάθησης συχνά παρουσιάζουν μειωμένη γενίκευση. Πολιτικές που έχουν εκπαιδευτεί σε ένα περιβάλλον μπορεί να μην λειτουργούν καθόλου σε ένα άλλο που έχει τροποποιηθεί ελάχιστα. Σε πραγματικές συνθήκες πτήσης, αυτό θα μπορούσε να είναι επικίνδυνο και να οδηγήσει σε καταστροφική αποτυχία.
- **Ευαισθησία στην παραμετροποίηση:** Οι υπερπαραμέτροι που προσαρμόζονται εμπειρικά είναι ζωτικής σημασίας για την αποτελεσματικότητα των τεχνικών βαθιάς και ενισχυτικής μάθησης. Σε κρίσιμες καταστάσεις, ο στοχαστικός χαρακτήρας των πολιτικών μπορεί να οδηγήσει σε απρόβλεπτα αποτελέσματα. Σε εφαρμογές όπου η ασφάλεια και η προβλεψιμότητα είναι κρίσιμες, τα δίκτυα που χρησιμοποιούνται για την προσέγγιση των πολιτικών ελέγχου αναφέρονται συχνά ως "μαύρα κουτιά" λόγω της δυσκολίας στην ερμηνεύσή τους, γεγονός που καθιστά δύσκολη τη διάγνωση σφαλμάτων, τη βελτιστοποίηση της απόδοσης και τη δημιουργία πιστοποιημένων μοντέλων πτήσης.

Μια περιεκτική αλλά διεξοδική σύνοψη των κύριων χαρακτηριστικών που αναδείχθηκαν από την συνδυασμένη εφαρμογή προσεγγίσεων ενισχυτικής μάθησης και βαθιάς μάθησης στον έλεγχο μη επανδρωμένων αεροσκαφών παρέχεται στον πίνακα 4.1. Οι επαναλαμβανόμενες αρχιτεκτονικές βελτιώνουν την ανθεκτικότητα, ενώ η ευελιξία και η αυτονομία των συστημάτων ενισχυτικής μάθησης επιτρέπει την διαχείριση δυναμικών και αβέβαιων περιβάλλοντων. Παρά τις ανάγκες σε δεδομένα και τα ζητήματα γενίκευσης, η εκπαίδευση από άκρο σε άκρο καθίσταται δυνατή από την άμεση επεξεργασία αισθητηριακών δεδομένων που επιτυγχάνεται μέσω της εφαρμογής αλγορίθμων βαθιάς μάθησης. Επιπλέον, οι σύγχρονες αρχιτεκτονικές μάθησης όπως οι Transformers, τα GNN και τα LSTM (Long Short Term Memory), βελτιώνουν τη χρησιμότητά τους σε καταστάσεις που περιλαμβάνουν πολλαπλούς πράκτορες ή σμήνη. Οι πολιτικές συνήθως εκπαιδεύονται σε προσομοιωτές χρησιμοποιώντας τεχνικές μεταφοράς μάθησης που επιτρέπουν την εφαρμογή τους σε πραγματικές καταστάσεις. Αν και η εκπαίδευση απαιτεί υψηλό υπολογιστικό κόστος, μπορεί να πραγματοποιηθεί αποτελεσματικά σε μη επανδρωμένα αεροσκάφη με περιορισμένους πόρους.

Πτυχή	Συνδυαστική περιγραφή
Προσαρμοστικότητα και αυτονομία	Τα συστήματα ενισχυτικής μάθησης μπορούν να μαθαίνουν πολιτικές χωρίς εποπτεία και να προσαρμόζονται σε αβέβαια, δυναμικά ή στοχαστικά περιβάλλοντα. Η χρήση επαναλαμβανόμενων αρχιτεκτονικών ενισχύει την ανθεκτικότητα σε απώλειες αισθητήρων ή περιβαλλοντικό θόρυβο. [37]
Επεξεργασία δεδομένων και βαθιά μάθηση	Η βαθιά μάθηση επιτρέπει την απευθείας ανάλυση δεδομένων από αισθητήρες, καθιστώντας εφικτή την end-to-end εκπαίδευση. Παρόλα αυτά, απαιτούνται τεράστιοι όγκοι δεδομένων και υπάρχει πρόβλημα γενίκευσης. [38],[39]
Μαθησιακές αρχιτεκτονικές επέκταση	Προηγμένες αρχιτεκτονικές βαθιάς μάθησης (Transformers, GNNs, LSTMs) βελτιώνουν την απόδοση σε συνεργατικά σενάρια ή σε σμήνη αεροχημάτων. Οι πολιτικές ενισχυτικής μάθησης μπορούν να υποστηρίξουν multi-objective στόχους, π.χ. σταθερότητα, ενεργειακή απόδοση, αποφυγή σύγκρουσης. [40]
Εκπαίδευση και μεταφορά στην πραγματικότητα	Η εκπαίδευση γίνεται κυρίως μέσω προσομοιώσεων (AirSim, Gazebo), με δυνατότητα μεταφοράς σε πραγματικό περιβάλλον (μεταφορά μάθησης). Αυτό μειώνει το κόστος δοκιμών και αυξάνει την ασφάλεια. [36]
Υπολογιστικοί πόροι και εκτέλεση	Παρότι η εκπαίδευση απαιτεί ισχυρό υλικό, τα εκπαιδευμένα μοντέλα μπορούν να μεταφερθούν και να εκτελεστούν σε edge UAVs μετά από βελτιστοποίηση. [41]
Ασφάλεια, ερμηνευσιμότητα και εγγυήσεις	Τα συστήματα βαθιάς και ενισχυτικής μάθησης υποφέρουν από χαμηλή ερμηνευσιμότητα (μαύρο κουτί), δυσκολία πιστοποίησης και απουσία θεωρητικών εγγυήσεων για σταθερότητα και σύγκλιση.[42]
Επανάχρηση και μεταφορά μάθησης	Η χρήση μετα-μάθησης και μεταφορά μάθησης επιταχύνει τη μάθηση σε νέα σενάρια, αλλά η απόδοση εξαρτάται από τη συνάφεια μεταξύ πηγής και στόχου.[43]

Πίνακας 4.1: Συνοπτική συγχώνευση πτυχών ενισχυτικής και βαθιάς μάθησης στον έλεγχο μη επανδρωμένων αεροσκαφών.

4.3 Σύγκριση με παραδοσιακές μεθόδους ελέγχου

Για πολλά χρόνια, το θεμέλιο για την πλοήγηση και τη σταθεροποίηση των μη επανδρωμένων αεροσκαφών έχει τεθεί από συμβατικές τεχνικές ελέγχου, όπως ο έλεγχος αναλογικού-ολοκληρωτικού-παραγωγικού, στρατηγικές βέλτιστου ελέγχου όπως LQR (Linear Quadratic Regulator) και MPC (Model Predictive Control), καθώς και προσαρμοστικούς ελεγκτές. Η ανάλυση σταθερότητας, η μοντελοποίηση δυναμικών συστημάτων και μια διεξοδική διαδικασία παραμετρικής ρύθμισης αποτελούν τη βάση αυτών των τεχνικών. Οι συμβατικοί ελεγκτές λειτουργούν καλύτερα σε περιβάλλοντα που είναι προ-καθορισμένα, στατικά και έχουν σε μεγάλο βαθμό σταθερές συνθήκες και μοντέλα συστημάτων. Ωστόσο, οι συμβατικές μέθοδοι έχουν σημαντικά μειονεκτήματα καθώς οι περιπτώσεις χρήσης των μη επανδρωμένων αεροχημάτων μετακινούνται προς δυναμικά, απρόβλεπτα και μη γραμμικά περιβάλλοντα. Τα συστήματα ελέγχου που βασίζονται σε ενισχυτική και βαθιά μάθηση μπορούν να μάθουν βέλτιστες στρατηγικές απευθείας από δεδομένα ή εμπειρία, επιτρέποντας συνεχή προσαρμογή.

Κριτήρια σύγκρισης

Απόδοση σε δυναμικά περιβάλλοντα

Στις εφαρμογές μη επανδρωμένων αεροχημάτων, ο έλεγχος της απροβλεπτότητας και των αιφνίδιων περιβαλλοντικών αλλαγών είναι απαραίτητος. Οι συμβατικές τεχνικές ελέγχου, όπως οι PID, LQR και MPC, βασίζονται σε προκαθορισμένες συνθήκες λειτουργίας ή στη σταθερότητα του μοντέλου. Από την άλλη πλευρά, οι προσεγγίσεις βαθιάς ενισχυτικής μάθησης χρησιμοποιούν εμπειρική μάθηση για να ενσωματώσουν τη δυναμική προσαρμογή, μέσω της οποίας η πολιτική τους αλλάζει μέσω εμπειριών και ανταμοιβών από αλληλεπίδραση με το περιβάλλον. Ακόμα και σε πραγματικό χρόνο, αυτό επιτρέπει την αποτελεσματική διαχείριση μεταβλητών όπως βλάβες αισθητήρων, ριπές ανέμου ή παρεμβολές. Ένα πλεονέκτημα διαφοροποίησης σε πραγματικές εφαρμογές παρέχεται από την ικανότητα των πρακτόρων που βασίζονται στη βαθιά ενισχυτική μάθηση να διατηρούν σταθερή πτήση ή να ολοκληρώνουν αποστολές ακόμα και σε καταστάσεις που διαφέρουν από το αρχικό σενάριο εκπαίδευσης.

Μοντελοποίηση συστημάτων

Ένα σημαντικό πρόβλημα για τις κλασικές προσεγγίσεις είναι η απαίτηση για λεπτομερή μοντελοποίηση. Για παράδειγμα, ένα ακριβές μαθηματικό μοντέλο που εξηγεί τη δυναμική συμπεριφορά του αεροχήματος, τις εξισώσεις κίνησης και χαρακτηριστικά όπως η αεροδυναμική αντίσταση, η ώθηση, το βάρος και το περιβάλλον λειτουργίας είναι απαραίτητο για την κατασκευή ενός ελεγκτή MPC. Η εξαγωγή και η επικύρωση αυτών των μοντέλων είναι χρονοβόρα, επιρρεπείς σε λάθη και συχνά μη γενικεύσιμη. Τα συστήματα βαθιάς ενισχυτικής μάθησης, ειδικά τα model-free, χρησιμοποιούν μόνο προσομοιώσεις ή πραγματικά περιβάλλοντα για να συλλέξουν εμπειρία. Χωρίς να γνωρίζει τις εξισώσεις κίνησης του φυσικού συστήματος ή άλλες λεπτομέρειες, ο πράκτορας μαθαίνει την καλύτερη στρατηγική ελέγχου. Αυτό μειώνει σημαντικά την απαίτηση για πρόιμη μηχανική μοντελοποίηση, αυξάνοντας την προσαρμοστικότητα των λύσεων σε διάφορα είδη αεροσκαφών.

Προσαρμοστικότητα σε μη γραμμικά δυναμικά συστήματα

Η αεροδυναμική και τα χαρακτηριστικά πτήσης ενός μη επανδρωμένου αεροσκάφους είναι από τη φύση τους πολύπλοκα και μη γραμμικά. Για να είναι υπολογιστικά βιώσιμοι, οι παραδοσιακοί ελεγκτές συχνά χρειάζονται απλοποιήσεις, όπως η γραμμικοποίηση γύρω από σημεία ισορροπίας, και αποδίδουν καλύτερα με γραμμικά ή ημιγραμμικά μοντέλα. Όταν τα εναέρια οχήματα λειτουργούν μακριά από ιδανικές συνθήκες λειτουργίας, αυτό μειώνει την ακρίβειά τους. Επειδή οι τεχνικές βαθιάς ενισχυτικής μάθησης βασίζονται σε παραμετρικά μοντέλα (όπως τα βαθιά νευρωνικά δίκτυα) που μπορούν να απεικονίσουν περίπλοκες και εξαιρετικά μη γραμμικές σχέσεις μεταξύ των παραμέτρων ελέγχου, της κατάστασης του συστήματος και της απόκρισης του συστήματος, αποδίδουν εξαιρετικά καλά. Επιπλέον, ο πράκτορας μαθαίνει να λειτουργεί διεξοδικά χωρίς να εξαρτάται από αυστηρές παραδοχές σχετικά με τη δυναμική του αεροχήματος, ρυθμίζοντας κατάλληλα τη συνάρτηση ανταμοιβής και την πολιτική. Ακόμη και σε καταστάσεις που είναι δύσκολο να απεικονιστούν εννοιολογικά, αυτό παρέχει ακρίβεια και ευελιξία.

Υπολογιστικό κόστος και υλοποίηση

Οι συμβατικοί ελεγκτές, όπως οι μικροελεγκτές χαμηλής κατανάλωσης, είναι κατασκευασμένοι για να λειτουργούν σε περιβάλλοντα με περιορισμένους πόρους. Το κόστος υπολογιστικής ισχύος είναι γνωστό, σταθερό και συνήθως ελάχιστο. Αυτό καθιστά εύκολη την ενσωμάτωσή του σε αεροσκάφους με περιορισμένες διαστάσεις, βάρος και κατανάλωση ενέργειας. Ωστόσο, τα συστήματα βαθιάς ενισχυτικής μάθησης χρειάζονται μεγάλη υπολογιστική ισχύ κατά την εκπαίδευσή τους, κάτι που συχνά γίνεται εκτός του αεροχήμα. Το κόστος κατά τη φάση λειτουργίας εξαρτάται από την πολυπλοκότητα του μοντέλου · παρόλα αυτά, υπάρχουν μέθοδοι για τη μείωσή του. Χάρη στις συνεχείς εξελίξεις στις ενσωματωμένες πλατφόρμες τεχνητής νοημοσύνης, η απόδοση μπορεί να διατηρηθεί κατά την σταδιακή εφαρμογή των ελεγκτών βαθιάς ενισχυτικής μάθησης σε πραγματικά αεροχήματα.

Ασφάλεια και αξιοπιστία

Σε αποστολές υψηλού κινδύνου ή κατά τη λειτουργία σε πυκνοκατοικημένες περιοχές και απρόβλεπτες συνθήκες, η ασφάλεια και η αξιοπιστία είναι ζωτικής σημασίας παράγοντες για τον έλεγχο των μη επανδρωμένων αεροχημάτων. Οι παραδοσιακές μέθοδοι ελέγχου πλεονεκτούν καθώς είναι πλήρως ερμηνεύσιμες, καθώς παρέχουν εγγυήσεις θεωρητικής σταθερότητας που έχουν αποδειχθεί και προσφέρουν προβλέψιμη συμπεριφορά ακόμα και σε περίπτωση αστοχίας αισθητήρων ή διακοπής επικοινωνίας. Αν και οι τεχνικές βαθιάς ενισχυτικής μάθησης μπορούν να επιτύχουν αξιοσημείωτα αποτελέσματα, δεν είναι εγγενώς ασφαλείς και η διαδικασία λήψης αποφάσεων είναι αδιαφανής (ένα "μαύρο κουτί"). Σε δεδομένα εκτός κατανομής, δηλαδή σενάρια που δεν καλύπτονται στην εκπαίδευση, μπορεί επίσης να παρουσιάσουν αστάθεια. Η επίσημη πιστοποίηση και η αξιοπιστία των συστημάτων DRL εξακολουθούν να αποτελούν αντικείμενο έρευνας, παρά τις εξελίξεις.

Κριτήριο σύγκρισης	DRL μέθοδοι	Παραδοσιακές μέθοδοι	Ενδεικτικές μελέτες (έτος)
Απόδοση σε δυναμικά περιβάλλοντα	Προσαρμόζονται σε μη στατικά ή απρόβλεπτα σενάρια μέσω συνεχούς μάθησης και εξερεύνησης.	Περιορισμένη προσαρμοστικότητα χωρίς ανασχεδιασμό ή αλλαγή παραμέτρων.	2024 [28]
Μοντελοποίηση συστήματος	Δεν απαιτούν ρητό μοντέλο – λειτουργούν με model-free προσέγγιση.	Απαιτείται λεπτομερές, συχνά μη ρεαλιστικό, μαθηματικό μοντέλο.	2025 [38]
Προσαρμογή σε μη γραμμικά συστήματα	Εξαιρετική ικανότητα διαχείρισης πολύπλοκων, μη γραμμικών συστημάτων.	Χρειάζονται ειδικά σχεδιασμένες ή γραμμικοποιημένες προσεγγίσεις.	2025 [44]
Υπολογιστικό κόστος και υλοποίηση	Υψηλό κόστος σε εκπαίδευση, ανάγκη για GPU/TPU.	Χαμηλότερες απαιτήσεις ειδικά σε ενσωματωμένα συστήματα.	2025 [41]
Ασφάλεια και Αξιοπιστία	Δύσκολη ερμηνεία αποφάσεων, μη εγγυημένη συμμόρφωση με περιορισμούς.	Υψηλή προβλεψιμότητα, ευκολότερη επικύρωση και θεωρητικές εγγυήσεις.	2025 [42]
Δειγματοληπτική αποδοτικότητα	Απαιτούν πολλές δοκιμές/επεισόδια, συνήθως σε προσομοιωτές.	Λιγότερη ανάγκη για δεδομένα, λόγω ενσωματωμένων θεωρητικών μοντέλων.	2025 [35]
Ενσωμάτωση περιορισμών	Γίνεται μέσω ποινών στη συνάρτηση ανταμοιβής, με περιορισμένη θεωρητική εγγύηση.	Άμεση ενσωμάτωση μέσω βελτιστοποίησης με περιορισμούς και σταθερότητα.	2025 [40]
Κλιμακωσιμότητα	Υψηλή δυνατότητα επέκτασης σε πολυπρακτορικά και υψηλής διάστασης σενάρια.	Περιορισμένη κλιμάκωση, αυξημένη πολυπλοκότητα σε μεγάλα συστήματα.	2025 [33]

Πίνακας 4.2: Σύγκριση βαθιάς ενισχυτικής μάθησης και παραδοσιακών μεθόδων ελέγχου.

Με έμφαση στα τεχνολογικά χαρακτηριστικά που έχουν άμεση σχέση με την επιχειρησιακή εφαρμογή στα UAV, ο πίνακας προσφέρει μια μεθοδική χαρτογράφηση των κύριων διακρίσεων μεταξύ των δύο στρατηγικών. Ο πίνακας παρουσιάζει τα σχετικά ισχυρά και ασθενή σημεία κάθε προσέγγισης, με κάθε γραμμή να υποδηλώνει έναν σημαντικό παράγοντα. Ωστόσο, είναι προφανές ότι τα κύρια εμπόδια στη χρήση του DRL σε πρακτικές εφαρμογές UAV είναι οι υψηλές υπολογιστικές απαιτήσεις του, η χαμηλή αποδοτικότητα δειγματοληψίας και τα ζητήματα ασφάλειας της εκπαίδευσης. Η πιθανή συμπληρωματικότητα των δύο προσεγγίσεων στα υβριδικά συστήματα ελέγχου τονίζεται από το γεγονός ότι οι παραδοσιακοί ελεγκτές, όπως οι PID και MPC, έχουν μια καλά

εδραιωμένη θεωρητική βάση και αξιοπιστία, καθιστώντας τους την προτιμώμενη επιλογή σε καταστάσεις που απαιτούν υψηλή ασφάλεια, προβλεψιμότητα και εμπεριστατωμένη κατανόηση του συστήματος.

4.4 Συνολική αξιολόγηση και σύγκριση μεθόδων

Ο έλεγχος των μη επανδρωμένων αεροσκαφών έχει γίνει ένα από τα πιο δύσκολα ζητήματα στη ρομποτική και στα συστήματα ελέγχου λόγω των εξελίξεων σε αυτόν τον τομέα. Οι μέθοδοι που χρησιμοποιούνται σήμερα διαφέρουν σημαντικά όσον αφορά τον σχεδιασμό τους, καθώς και τις υποθέσεις και τις προδιαγραφές τους. Οι συμβατικές τεχνικές όπως οι PID, LQR και MPC αποδίδουν καλύτερα σε καταστάσεις με προβλέψιμες συνθήκες πτήσης και ακριβή κατανόηση του δυναμικού μοντέλου. Παρόλα αυτά, όταν συμβαίνουν απροσδόκητες διαταραχές ή όταν χρειάζεται γενίκευση εκτός του προγραμματισμένου εύρους, αντιμετωπίζουν δυσκολίες. Η αδυναμία προσαρμογής στις μεταβαλλόμενες παραμέτρους, η οποία επιβάλλει επανασχεδιασμό ή επαναβαθμολόγηση, είναι άλλο ένα μειονέκτημα.

Η ενισχυτική μάθηση προσφέρει μια εναλλακτική μέθοδο δεδομένου ότι βασίζεται στην εμπειρική βελτιστοποίηση στρατηγικών μέσω δοκιμής και σφάλματος. Αυτό καθιστά δυνατή την εκμάθηση αποτελεσματικών στρατηγικών ακόμα και σε καταστάσεις με αβεβαιότητα ή με κρυφές μεταβλητές. Ωστόσο, χωρίς σημαντικές υποδομές, τα περισσότερα συστήματα RL είναι μη λειτουργικά λόγω των υψηλών απαιτήσεων σε πόρους, είτε σε προσομοίωση είτε σε πραγματικό χρόνο. Η βαθιά μάθηση, παρότι δεν χρησιμοποιείται συχνά αυτόνομα για τον πυρήνα του ελέγχου, είναι μια σημαντική μονάδα εξαγωγής χαρακτηριστικών υψηλού επιπέδου, κυρίως από οπτικά δεδομένα ή σήματα αισθητήρων. Η βαθιά μάθηση ενσωματώνεται συχνά σε δομές ελέγχου ή πρόβλεψης σε άτυπα ή περιβάλλοντα με χαμηλή παρατηρησιμότητα, με την ικανότητά της να χαρτογραφεί μη γραμμικές εισόδους, να προσφέρει μεγάλη ευελιξία. Συνδυάζοντας τα δυνατά σημεία της βαθιάς μάθησης για την αναπαράσταση και τη γενίκευση με τη στρατηγική βελτιστοποίησης της ενισχυτικής μάθησης, η βαθιά ενισχυτική μάθηση είναι το αποτέλεσμα της συνένωσης των δύο προηγούμενων τεχνολογιών. Οι πράκτορες βαθιάς ενισχυτικής μάθησης μπορούν να επιτύχουν υψηλή απόδοση σε περιβάλλοντα που παραδοσιακές μέθοδοι αποτυγχάνουν.

Ο βαθμός γενίκευσης της λύσης είναι ένας κρίσιμος παράγοντας που διακρίνει τις τεχνικές: Οι πολιτικές βαθιάς ενισχυτικής μάθησης μπορούν να προσαρμοστούν (ή ακόμα και να επαναπροσαρμοστούν) σε νέες συνθήκες χωρίς να απαιτείται επανασχεδιασμός, ενώ οι παραδοσιακοί ελεγκτές είναι βελτιστοποιημένοι για ένα περιορισμένο σύνολο συνθηκών. Επομένως θεωρούνται ιδιαίτερα ελκυστικές για αποστολές με ασταθείς συνθήκες πτήσης. Τέλος, ένα άλλο κρίσιμο σημείο απόκλισης είναι η μετάβαση από τη προσομοίωση στην πραγματικότητα. Εφόσον πληρούνται οι υποθέσεις του μοντέλου, οι παραδοσιακές μέθοδοι, οι οποίες από τη φύση τους είναι αναλυτικές, μπορούν να εφαρμοστούν άμεσα. Από την άλλη, οι πράκτορες χρειάζονται περαιτέρω μεθόδους για να αποτρέψουν τα λεγόμενα κενά πραγματικότητας, δηλαδή τις διαφορές μεταξύ των ρυθμίσεων εκπαίδευσης και του πραγματικού κόσμου.

4.4.1 Ανάλυση καταλληλότητας μεθόδων ανά σενάριο χρήσης

Στην ουσία, το λειτουργικό σενάριο και οι τεχνολογικές ανάγκες ενός μη επανδρωμένου αεροσκάφους καθορίζουν ποια μέθοδος ελέγχου είναι η καλύτερη. Διαφορετικά είδη μη επανδρωμένων

αεροσκαφών, καθώς και διαφορετικά περιβάλλοντα λειτουργίας, εγείρουν ασύμβατες απαιτήσεις από πλευράς ασφάλειας, ενέργειας, και απόδοσης ελέγχου. Καμία μεμονωμένη μέθοδος – είτε βασισμένη σε μάθηση είτε συμβατική – δεν είναι πάντα καλύτερη. Η καταλληλότητα ποικίλλει ανάλογα με την αποστολή και τις συνθήκες του συστήματος.

- **UAV σταθερού πτερυγίου & αποστολές μακράς εμβέλειας**

Η ακρίβεια στην παρακολούθηση πορείας και η διατήρηση σταθερής δυναμικής πτήσης, είναι απαραίτητες για αποστολές μεγάλης διάρκειας, συμπεριλαμβανομένης της παρακολούθησης συνόρων ή εκτεταμένων γεωργικών εφαρμογών με χρήση αεροσκαφών σταθερού πτερυγίου. Εδώ, συνιστώνται συμβατικές τεχνικές όπως PID ή LQR επειδή παρέχουν καλή σταθερότητα και έχουν χαμηλό υπολογιστικό κόστος, κάτι που είναι ζωτικής σημασίας για τα UAV με περιορισμένους ενσωματωμένους επεξεργαστές.

- **Δυναμικές αποστολές με αβέβια στοιχεία**

Επειδή στερούνται ευελιξία, οι παραδοσιακές τεχνικές είναι συχνά ανεπαρκείς σε καταστάσεις με υψηλό βαθμό αβεβαιότητας (όπως περιβαλλοντικοί παράγοντες όπως οι ριπές ανέμου ή κινούμενα εμπόδια). Ακόμη και σε μεταβαλλόμενες συνθήκες λειτουργίας, οι προσεγγίσεις ενισχυτικής μάθησης (Adaptive RL, Meta-RL) επιτρέπουν τη σταδιακή βελτιστοποίηση των πολιτικών πτήσης καθ' όλη τη διάρκεια μιας αποστολής.

- **Συνεργατικά συστήματα μη επανδρωμένων αεροχημάτων (Multi-UAV)**

Οι τεχνικές της ενισχυτικής μάθησης πολλαπλών πρακτόρων (Multi-agent reinforcement learning - MARL) επιτρέπουν την αλληλεπίδραση σε πραγματικό χρόνο και την εκμάθηση συνεργατικών μεθόδων όταν πολλά αεροχήματα πρέπει να συνεργάζονται, όπως σε επιχειρήσεις έρευνας και διάσωσης ή δυναμικής κάλυψης περιοχής. Αυτές οι τεχνικές ακόμη και με την υψηλότερη υπολογιστική τους πολυπλοκότητα, είναι αρκετά αποτελεσματικές επειδή επιτρέπουν την προσαρμοστική κατανομή ρόλων και την αποφυγή συγκρούσεων μέσω της εμπειρίας,

- **Αστικά περιβάλλοντα – Πλοήγηση με εμπόδια**

Η εφαρμογή της βαθιάς ενισχυτικής μάθησης με αλγορίθμους όπως οι DDPG, PPO ή SAC, προσφέρει αξιοσημείωτα πλεονεκτήματα στην αποφυγή συγκρούσεων, στον ασφαλή σχεδιασμό διαδρομών και στην εκμάθηση κίνησης σε κακώς μοντελοποιημένα περιβάλλοντα. Η ενσωμάτωση αισθητήρων (LiDAR, κάμερες) με νευρωνικά δίκτυα επιτρέπει λήψη αποφάσεων υψηλού επιπέδου ακόμη και χωρίς πλήρες μοντέλο του περιβάλλοντος.

- **Εσωτερικοί χώροι**

Η πλήρης επίγνωση του περιβάλλοντος και η ικανότητα πλοήγησης σε περιορισμένους χώρους είναι απαραίτητες σε εφαρμογές εντός κτιρίων, αποθήκες ή εγκαταστάσεις με κακή ακρίβεια τοποθεσίας. Σε αυτές τις περιπτώσεις, το συμβατικό μοντέλο ελέγχου μπορεί να αντικατασταθεί από την εκμάθηση μίμησης (Imitation learning) ή την περιορισμένη ενισχυτική μάθηση (Constrained RL), η οποία εγγυάται ασφαλή και εύκολη πλοήγηση ακόμα και χωρίς πλήρη χαρτογράφηση.

Σενάριο χρήσης	Τεχνικές απαιτήσεις	Κατάλληλη προσέγγιση	Σχόλια / Ιδιαιτερότητες
Fixed-Wing UAV – Μακράς διάρκειας πτήση	Σταθερότητα, απόδοση ενέργειας, χαμηλό υπολογιστικό φορτίο	Παραδοσιακοί ελεγκτές (PID, LQR)	Αποτελεσματικοί σε καλά μοντελοποιημένα συστήματα· κατάλληλοι για απλές, προκαθορισμένες πορείες. [45]
Swarm UAVs – Συνεργατικά συστήματα	Συντονισμός, αποφυγή σύγκρουσης, δυναμική ανάθεση στόχων	Multi-Agent RL (MARL)	Υψηλή υπολογιστική απαίτηση· κατάλληλο για μάθηση σε πραγματικό χρόνο και συνεργασία. [46]
Αστικά περιβάλλοντα – Πλοήγηση με εμπόδια	Αντίληψη περιβάλλοντος, real-time αποφάσεις, αποφυγή εμποδίων	DRL (π.χ. PPO, SAC, DDPG)	Εύκαμπτες πολιτικές πτήσης· δεν απαιτούν πλήρες δυναμικό μοντέλο. [23] [47]
Εσωτερικά περιβάλλοντα	Πλοήγηση με αισθητήρες, περιορισμένος χώρος, ασφάλεια	Imitation Learning	Αξιοποιεί αισθητηριακή είσοδο και ιστορικά δεδομένα· κατάλληλο για σταθερές, επαναλαμβανόμενες διαδρομές. [48]
Δυναμικές / αβέβαιες συνθήκες (π.χ. άνεμος, κινούμενα εμπόδια)	Ανθεκτικότητα σε μεταβολές, σταδιακή μάθηση	Meta-RL / προσαρμοστική RL	Υποστηρίζει διαδικασίες συνεχούς μάθησης· κατάλληλο για απρόβλεπτα σενάρια. [37]
Υβριδικά συστήματα	Συνδυασμός ακριβείας και ευελιξίας	DRL + κλασικοί ελεγκτές (π.χ. DRL policy με PID εκτελεστή)	Εξισορροπεί σταθερότητα με προσαρμοστικότητα· εφαρμόσιμο σε ελιγμούς υψηλής ακριβείας. [29]

Πίνακας 4.3: Ενδεικτικά σενάρια ελέγχου και κατάλληλες προσεγγίσεις.

Μια συγκριτική τεχνική επισκόπηση των καλύτερων στρατηγικών ελέγχου για κάθε σενάριο χρήσης δίνεται στον παραπάνω Πίνακα 4.3, λαμβάνοντας υπόψη σημαντικούς παράγοντες όπως η δυναμική του περιβάλλοντος, οι υπολογιστικοί περιορισμοί και οι ανάγκες προσαρμοστικότητας. Λόγω της απλότητάς τους και των ελάχιστων απαιτήσεων υπολογιστικής ισχύος, οι κλασικοί ελεγκτές συνεχίζουν να είναι αποτελεσματικοί σε στατικές και καλά μοντελοποιημένες συνθήκες. Από την άλλη πλευρά, οι τεχνικές βαθιάς ενισχυτικής μάθησης αποδίδουν καλά σε εξαιρετικά πολύπλοκα και αβέβαια περιβάλλοντα, συμπεριλαμβανομένων των σμηνών UAV, επειδή μπορούν να μαθαίνουν από τα λάθη και να προσαρμόζονται στις μεταβαλλόμενες συνθήκες. Ορισμένες εφαρμογές, όπως ο υβριδικός έλεγχος ή η πλοήγηση σε εσωτερικούς χώρους, απαιτούν εξειδικευμένες ή συνδυαστικές προσεγγίσεις.

4.4.2 Εφαρμογές και περιορισμοί σε πραγματικά περιβάλλοντα

Ένα κρίσιμο βήμα στην αξιολόγηση της χρησιμότητας των στρατηγικών ελέγχου που χρησιμοποιούν τη βαθιά και την ενισχυτική μάθηση, είναι η μεταφορά από προσομοιωμένες σε πραγματικές εφαρμογές. Η πλήρης ενσωμάτωση αυτών των τεχνικών σε πραγματικά εναέρια οχήματα είναι γεμάτη λειτουργικές και τεχνικές δυσκολίες, παρά τις αξιοσημείωτες προόδους στις προσομοιώσεις. Διατίθενται ρεαλιστικές πλατφόρμες εικονικής πραγματικότητας για την ασφαλή εκπαίδευση αλγορίθμων, οι οποίοι στη συνέχεια μεταφέρονται σε πραγματικά μη επανδρωμένα αεροχήματα. Στην πράξη, οι αλγόριθμοι ενισχυτικής και βαθιάς ενισχυτικής μάθησης έχουν χρησιμοποιηθεί κυρίως σε ελεγχόμενα περιβάλλοντα ή για συγκεκριμένους σκοπούς. Σε άλλες περιπτώσεις, έχουν χρησιμοποιηθεί οι μέθοδοι DDPG ή SAC για εκπαίδευση σε συνεχείς χώρους δράσης σε πραγματικές εφαρμογές σε τετρακόπτερα.

Παρά αυτά τα επιτεύγματα, μια σειρά τεχνικών ζητημάτων εμποδίζουν την ευρεία χρήση των αλγορίθμων σε εμπορικά αεροχήματα. Για να συγκλίνουν σε βέλτιστες πολιτικές, οι περισσότερες τεχνικές ενισχυτικής μάθησης χρειάζονται εκατομμύρια αλληλεπιδράσεις, κάτι που είναι σχεδόν ανέφικτο σε πραγματικές πτήσεις. Είναι δύσκολο να ενσωματωθούν τα μοντέλα βαθιάς ενισχυτικής μάθησης, ιδίως όταν περιλαμβάνουν συνελκτικές ή αναδρομικές αρχιτεκτονικές, λόγω των υψηλών απαιτήσεων υπολογιστικής ισχύος τους. Επίσης, η σύγκριση και η αξιολόγηση είναι δύσκολες επειδή δεν υπάρχει επί του παρόντος ένα ευρέως αναγνωρισμένο σύνολο κριτηρίων για την βαθιά ενισχυτική μάθηση σε πραγματικά δεδομένα, και οι κατασκευαστές δεν έχουν ενσωματώσει ακόμα υποστήριξη για βιβλιοθήκες βαθιάς ενισχυτικής μάθησης. Η προσέγγιση "trial and error" της ενισχυτικής μάθησης αυξάνει την πιθανότητα αποτυχίας ή σύγκρουσης κατά τη φάση εκμάθησης. Συνεπώς, συχνά απαιτούνται περιορισμένη ενίσχυτική μάθηση ή υβριδικές τεχνικές.

4.5 Ποσοτική αξιολόγηση από βιβλιογραφία

Προηγουμένως έγινε αναφορά στο γεγονός ότι δεν μπορεί να επιτευχθεί συγκρισιμότητα και αντικειμενική αποτίμηση των παραπάνω τεχνικών, επειδή υπάρχουν συγκεκριμένα εμπόδια. Παρά τις δυσκολίες, η ερευνητική κοινότητα έχει επικεντρωθεί στην εξέλιξη των μεθοδολογιών ποσοτικής αξιολόγησης. Τα άρθρα που αναλύουν σενάρια πλοήγησης, αποφυγή εμποδίων, αυτόνομη προσγείωση ή συνεργατικών πτήσεων περιλαμβάνουν συχνά μετρήσεις απόδοσης, όπως το ποσοστό επιτυχίας της αποστολής, το μέσο τετραγωνικό σφάλμα (RMSE) στη διατήρηση της πορείας ή της προσγείωσης, ο χρόνος σύγκλισης πολιτικής και η κατανάλωση ενέργειας ανά χιλιόμετρο. Η σύ-

γκριση διαφορετικών προσεγγίσεων σε δημοφιλή συστήματα προσομοίωσης όπως το AirSim και το Gazebo είναι δυνατή χάρη στη μεθοδική καταγραφή τέτοιων μετρικών, η οποία συμβάλλει στη δημιουργία μιας πιο αξιόπιστης βάσης αναφοράς. Αυτή η ενότητα συγκεντρώνει αξιοσημείωτα παραδείγματα από τη βιβλιογραφία, σε μια προσπάθεια να απεικονίσει το εύρος των επιδόσεων που επιτυγχάνονται με τη χρήση διαφορετικών προσεγγίσεων και συνθηκών.

Πίνακας 4.4: Μετρικές για την πλοήγηση μη επανδρωμένων αεροχημάτων.

Μελέτη	Μέθοδος & Περιβάλλον	Μετρική	Τιμή
2024, [28]	PPO (Προσομοίωση Unity)	Ποσοστό επιτυχίας	92.5%
2024, [28]	Σύγκριση DDPG vs SAC	Σύγκλιση	~200 επεισόδια
2025, [38]	End-to-End DRL (Sim)	RMSE	0.38 m
2025, [38]	End-to-End DRL (Sim)	Ποσοστό επιτυχίας	94.1%

Πίνακας 4.5: Μετρικές των μη επανδρωμένων αεροχημάτων για την αποφυγή εμποδίων.

Μελέτη	Μέθοδος & Περιβάλλον	Μετρική	Τιμή
2025, [35]	PPO (Δυναμικά εμπόδια)	Ποσοστό επιτυχίας	89.3%
2023, [49]	DRL (Άγνωστο περιβάλλον)	Επεισόδια εκπαίδευσης	~2000
2025, [50]	DRL (Complex Sim)	Ποσοστό επιτυχίας	> 90%
2023, [34]	Improved PPO	Εκπαίδευση σύγκλισης	~5000 επεισόδια

Πίνακας 4.6: Μετρικές των μη επανδρωμένων αεροχημάτων για προσγείωση.

Μελέτη	Μέθοδος & Περιβάλλον	Μετρική	Τιμή
2024, [51]	DQN (Κινούμενη πλατφόρμα)	Επιτυχία προσγείωσης	100%
2024, [51]	Vision-based RL	Σφάλμα θέσης (std)	0.12 m

Πίνακας 4.7: Μετρικές σμηνών (Swarms) μη επανδρωμένων αεροχημάτων.

Μελέτη	Μέθοδος & Περιβάλλον	Μετρική	Τιμή
2025, [52]	MADDPG (Συνεργατικό)	Ποσοστό ολοκλήρωσης	91.0%
2025, [52]	Multi-Agent RL	Μέσο σφάλμα σχηματισμού	0.27 m

Κεφάλαιο 5

Μελέτες περίπτωσης και προκλήσεις

Μια κρίσιμη φάση στη διαδικασία έρευνας και ανάπτυξης, είναι η μετάβαση από τη θεωρητική δημιουργία και αξιολόγηση αλγορίθμων βαθιάς και ενισχυτικής μάθησης στην πραγματική εφαρμογή τους σε επιχειρησιακά περιβάλλοντα. Η χρήση αυτών των τεχνολογιών σε πραγματικές συνθήκες δημιουργεί μια σειρά από λειτουργικές και τεχνολογικές δυσκολίες που απαιτούν μεθοδική έρευνα και βελτίωση, ακόμη και με την αξιοσημείωτη πρόοδο που έχει σημειωθεί σε περιβάλλοντα προσομοίωσης. Ο στόχος αυτής της ενότητας είναι να συσχετίσει τα αποτελέσματα που έχουν επιτευχθεί σε ελεγχόμενα περιβάλλοντα με τις απαιτήσεις του πεδίου. Η εμπειρία έχει δείξει ότι το λεγόμενο χάσμα πραγματικότητας (reality gap), δηλαδή η απόκλιση μεταξύ του ιδανικού, ελεγχόμενου μοντέλου και της αβεβαιότητας, της πολυπλοκότητας και του θορύβου που χαρακτηρίζουν το πραγματικό περιβάλλον, μπορεί να προκαλέσει δραστική επιδείνωση της απόδοσης ενός αλγορίθμου όταν εφαρμόζεται σε ένα πραγματικό μη επανδρωμένο αεροσκάφος.

Η βελτίωση της αυτονομίας, της ευελιξίας και της προσαρμοστικότητας σε δυναμικές καταστάσεις είναι δυνατή με τη χρήση των προσεγγίσεων βαθιάς αλλά και βαθιάς ενισχυτικής μάθησης. Ωστόσο, υπάρχουν πολλές προκλήσεις που πρέπει να επιλυθούν κατά τη μετάβαση από το σχεδιασμό στο περιβάλλον λειτουργίας. Μερικές είναι:

- **Διαχείριση αβεβαιότητας** λόγω καιρικών συνθηκών ή παρεμβολών
- **Ενεργειακή κατανάλωση** που επηρεάζει άμεσα τον χρόνο πτήσης και την αυτονομία
- **Περιορισμοί υπολογιστικής ισχύος** καθώς δεν υποστηρίζουν όλα τα συστήματα τους ίδιους υπολογιστικούς πόρους
- **Νομικοί περιορισμοί** που αφορούν την επιχειρησιακή χρήση των αεροχημάτων.

Σε αυτό το πλαίσιο, οι μελέτες των περιπτώσεων χρήσης είναι σημαντικές για δύο λόγους. Αφενός, καθιστούν δυνατή την αξιολόγηση της αποτελεσματικότητας συγκεκριμένων εφαρμογών σε πρακτικές συνθήκες Αφετέρου, λειτουργούν ως σημείο αναφοράς για τον προσδιορισμό των κύριων μειονεκτημάτων και την πρόταση βελτιώσεων για το μέλλον. Ενώ η εκπαίδευση σε προσομοιωτές μπορεί να μειώσει το κόστος και τον κίνδυνο των δοκιμών, μελέτες που χρησιμοποιούν την βαθιά ενισχυτική μάθηση έχουν δείξει ότι απαιτούνται ορισμένες στρατηγικές μεταφοράς μάθησης (transfer learning) και προσαρμογής για να γεφυρωθεί το χάσμα μεταξύ προσομοίωσης και πραγματικότητας.

Προκειμένου να καταδειχθεί το εύρος των τεχνικών αυτών, το παρόν κεφάλαιο είναι δομημένο κατά τέτοιο τρόπο ώστε να παρουσιάζει πρώτα μερικά παραδείγματα περιπτώσεων που καλύπτουν διάφορους τύπους μη επανδρωμένων αεροχημάτων και αποστολών. Στη συνέχεια, ακολουθεί η ανάλυση των προκλήσεων που εμποδίζουν την πλήρη αξιοποίηση των τεχνολογιών αυτών σε πραγματικές συνθήκες. Τέλος, συζητούνται οι κύριες κατευθύνσεις της έρευνας για το μέλλον, προκειμένου να ξεπεραστούν αυτές οι προκλήσεις και να ενσωματωθεί πλήρως η αυτονομία των UAV σε ένα ευρύ φάσμα εφαρμογών. Επομένως, εστιάζοντας τόσο στις τεχνικές (αλγοριθμικές, υπολογιστικές, ενεργειακές) όσο και στις πρακτικές (νομικές, αξιοπιστία, κόστος) πτυχές, αυτό το μέρος λειτουργεί ως σύνδεσμος μεταξύ της τεχνολογικής βιωσιμότητας και της λειτουργικής βιωσιμότητας. Η μελλοντική ανάπτυξη καθορίζεται από τη σχέση μεταξύ αυτών των χαρακτηριστικών.

5.1 Περιπτώσεις χρήσης

5.1.1 Μη επανδρωμένα αεροχήματα σταθερού πτερυγίου

Για αποστολές μεγάλης εμβέλειας και διάρκειας, όπως η χαρτογράφηση ή η επιτήρηση μεγάλων εκτάσεων, τα αεροχήματα σταθερού πτερυγίου αποτελούν την ιδανική πλατφόρμα. Σε σύγκριση με τα περιστροφικών πτερυγίων, τα αεροσκάφη σταθερού πτερυγίου προσφέρουν σημαντικό πλεονέκτημα όσον αφορά την αεροδυναμική απόδοση, επιτρέποντας ταχύτερες πτήσεις και μεγαλύτερη εμβέλεια, γεγονός που τα καθιστά κατάλληλα για αποστολές που απαιτούν κάλυψη μεγάλων αποστάσεων. Οι τεχνικές προδιαγραφές ενός συμβατικού σταθερού πτερυγίου είναι ένα ελαφρύ αεροδυναμικό αεροσκάφος με κινητήρα εσωτερικής καύσης ή έναν ηλεκτροκινητήρα, συστήματα αισθητήρων όπως GPS, IMU, κάμερες υψηλής ανάλυσης και LIDAR για ακριβή πλοήγηση και συλλογή δεδομένων.

Με τη βελτιστοποίηση συγκεκριμένων στόχων, όπως η μεγιστοποίηση της κάλυψης της περιοχής ή η μείωση της κατανάλωσης ενέργειας, αναπτύσσονται αλγόριθμοι που μπορούν να μάθουν πολιτικές πτήσης ως μέρος της εφαρμογής τεχνικών βαθιάς ενισχυτικής μάθησης στη λειτουργία τέτοιων αεροσκαφών. Συνήθως, το πρόβλημα ελέγχου εκφράζεται ως μαρκοβιανή διαδικασία απόφασης, όπου η κατάσταση (state) περιλαμβάνει πληροφορίες όπως η θέση, ταχύτητα και προσανατολισμός, οι ενέργειες (actions) αφορούν τις εντολές στις επιφάνειες ελέγχου και στον κινητήρα, και η συνάρτηση ανταμοιβής (reward) έχει ως στόχο να προωθεί ασφαλείς, αποτελεσματικές και ομαλές πτήσεις. Προκειμένου να διευκολυνθεί η γρήγορη αξιολόγηση και βελτιστοποίηση, η εκπαίδευση πραγματοποιείται πρώτα σε προσομοιωτές με ακριβή δυναμικά μοντέλα. Στη συνέχεια, χρησιμοποιούνται τεχνικές προσαρμογής για τη διαχείριση των παραλλαγών του μοντέλου όταν η πολιτική μεταφέρεται στο πραγματικό αεροσκάφος (sim to real transfer).

Σύμφωνα με τα αποτελέσματα που παρουσιάζονται σε μελέτες [45] [53], οι ελεγκτές που βασίζονται σε βαθιά ενισχυτική μάθηση μπορούν να αποδώσουν αισθητά καλύτερα από τις πιο συμβατικές τεχνικές ελέγχου, όπως PID ή LQR, ιδιαίτερα σε καταστάσεις που περιλαμβάνουν μεταβαλλόμενες συνθήκες περιβάλλοντος. Αυτές οι τεχνικές είναι υπερέχουσες και παρουσιάζουν μεγαλύτερη αντοχή σε απρόβλεπτες διαταραχές, όπως αποδεικνύεται από μετρήσεις όπως ο χρόνος πτήσης, η ακρίβεια πλοήγησης και η κατανάλωση ενέργειας. Επιπλέον, οι αλγόριθμοι μπορούν να προσαρμόζονται δυναμικά μαθαίνοντας από δεδομένα σε πραγματικό χρόνο, γεγονός που εξαλείφει την ανάγκη για εκτεταμένη μοντελοποίηση και παραμετροποίηση.

5.1.2 Συνεργατική λειτουργία αεροχημάτων

Η ιδέα πίσω από την συνεργατική και ταυτόχρονη λειτουργία πολλών μη επανδρωμένων αεροσκαφών που εκτελούν αποστολές σύμφωνα με αρχές συντονισμού, έχουν ως πρότυπο φυσικά φαινόμενα όπως οι κινήσεις σμήνους εντόμων ή πουλιών. Οι εξελίξεις στην βαθιά ενισχυτική μάθηση έχουν καταστήσει δυνατή τη δημιουργία πιο προσαρμόσιμων και αυτοδιαχειριζόμενων συστημάτων σμήνους που μπορούν να ανταλλάσσουν πληροφορίες, να προσαρμόζονται σε μεταβαλλόμενες συνθήκες και να λαμβάνουν συλλογικές αποφάσεις με ελάχιστη ανθρώπινη παρέμβαση. Η ενισχυτική μάθηση πολλαπλών πρακτόρων, στην οποία κάθε αερόχημα λειτουργεί ως πράκτορας που αλληλεπιδρά με το περιβάλλον και άλλους πράκτορες, χρησιμοποιείται συχνά στην τεχνική προσέγγιση για την επίλυση του προβλήματος διαχείρισης σμήνους. Η εκπαίδευση μπορεί να ακολουθεί μοντέλο κεντροποιημένης εκπαίδευσης με αποκεντρωμένη εκτέλεση (CTDE), επιτρέποντας τη συλλογική μάθηση σε στάδιο προσομοίωσης και την αποκεντρωμένη λήψη αποφάσεων κατά την πραγματική αποστολή. Για τη μοντελοποίηση και τη βελτιστοποίηση των αλληλεπιδράσεων μεταξύ των αεροχημάτων, χρησιμοποιούνται συχνά τα νευρωνικά δίκτυα γραφημάτων επειδή κωδικοποιούν τις τοπολογικές πληροφορίες του σμήνους και διευκολύνουν τον πιο αποτελεσματικό συντονισμό σε πραγματικό χρόνο.

Ένας από τους σημαντικότερους παράγοντες επιτυχίας είναι η επικοινωνία των αεροσκαφών. Τα ad-hoc δίκτυα mesh και τα πρωτόκολλα χαμηλής καθυστέρησης είναι ζωτικής σημασίας σε περιβάλλοντα με περιορισμένο εύρος ζώνης ή υψηλή πιθανότητα παρεμβολών. Επίσης, η ανεκτικότητα σε σφάλματα μπορεί να εξασφαλιστεί με την εκπαίδευση αλγορίθμων μηχανικής μάθησης για τη διαχείριση αποσυνδέσεων ή απώλειας κόμβων. Τεχνικές όπως η συνεργατική χαρτογράφηση (Simultaneous Localization and Mapping - SLAM) επιτρέπουν σε όλο το σμήνος να δημιουργεί και να ενημερώνει έναν κοινό χάρτη της περιοχής, βελτιώνοντας την ακρίβεια πλοήγησης. Η καλύτερη κάλυψη της περιοχής και η χαμηλότερη κατανάλωση ενέργειας επιτυγχάνονται χάρη στη συνεργατική πλοήγηση με τη χρήση βελτιστοποιημένων σχηματισμών αεροσκαφών. Επιπλέον, η αποκεντρωμένη λήψη αποφάσεων εξαλείφει την ανάγκη για συνεχή σύνδεση με έναν κεντρικό σταθμό ελέγχου και εγγυάται ταχύτερη αντίδραση σε απειλές ή αλλαγές αποστολής.

Η τεχνολογία σμηνών έχει πρακτικές εφαρμογές σε διάφορους κλάδους. Στον τομέα της αναζήτησης και διάσωσης, τα σμήνη μπορούν να σαρώσουν γρήγορα εκτεταμένες περιοχές και να ανταλλάξουν δεδομένα σε πραγματικό χρόνο για τον εντοπισμό αγνοουμένων. Τα αεροχήματα που είναι εξοπλισμένα με πολυφασματικούς αισθητήρες μπορούν να εντοπίζουν αλλαγές στη βλάστηση, να χαρτογραφούν καμμένες περιοχές και να παρακολουθούν την εξάπλωση ρύπων. Στον στρατιωτικό τομέα, χρησιμοποιούνται για αναγνώριση, παραπλάνηση εχθρικών συστημάτων και ακόμη και για συντονισμένες επιθέσεις με κέντρα εμπλοκής. Στον βιομηχανικό τομέα, μπορούν να αυτοματοποιήσουν τις επιθεωρήσεις σημαντικών υποδομών (π.χ. αγωγοί, δίκτυα υψηλής τάσης) μέσω κατανεμημένης κάλυψης. Η αξιοποίηση της νοημοσύνης σμήνους με βαθιά ενισχυτική μάθηση παρέχει μια νέα κατηγορία μη επανδρωμένων αεροσκαφών που ξεπερνά τα μεμονωμένα αεροσκάφη σε όρους ευελιξίας και αυτονομίας, κάτι που θα ήταν αδύνατο να επιτευχθεί με τις παραδοσιακές μεθόδους ελέγχου.

5.1.3 Λειτουργία μη επανδρωμένων αεροχημάτων σε αστικά περιβάλλοντα

Ένα από τα πιο δύσκολα και περίπλοκα προβλήματα στον τομέα του αυτόνομου ελέγχου είναι η λειτουργία των αεροσκαφών σε αστικές περιοχές. Τα υψηλά επίπεδα χωρικής πολυπλοκότητας, η υψηλή πυκνότητα εμποδίων (κτίρια, γέφυρες και πυλώνες), ο περιορισμένος ανοιχτός χώρος και τα δυναμικά κινούμενα αντικείμενα, όπως αυτοκίνητα και πεζοί, είναι χαρακτηριστικά των πόλεων. Από τεχνική άποψη, το πρόβλημα ξεκινά με τον τρόπο που αντιλαμβάνεται κανείς το περιβάλλον. Η τρισδιάστατη χαρτογράφηση του χώρου καθίσταται δυνατή χάρη στην ενσωμάτωση αισθητήρων όπως LiDAR, στερεοσκοπικές κάμερες και ραντάρ μικρής εμβέλειας. Ακόμη και σε περιβάλλοντα με χαμηλό φωτισμό ή ομίχλη, τα εναέρια μέσα μπορούν να αναγνωρίσουν και να κατηγοριοποιήσουν με ακρίβεια αντικείμενα, να εντοπίσουν σημεία ενδιαφέροντος και να εντοπίσουν πιθανούς κινδύνους με τη χρήση συνελκτικών νευρωνικών δικτύων. Σε συνδυασμό με αλγορίθμους σημασιολογικού διαχωρισμού (Semantic Segmentation) μπορούν να δημιουργηθούν πλούσιοι χάρτες που διαφοροποιούν δρόμους, κτίρια, και ζώνες περιορισμένης πτήσης.

Αλγόριθμοι βαθιάς ντετερμινιστικής κλίσης πολιτικής, βελτιστοποίησης πολιτικής ή Soft Actor-Critic (SAC) βελτιώνουν την πλοήγηση και την αποφυγή συγκρούσεων σε αστικές περιοχές. Αυτοί οι αλγόριθμοι εκπαιδεύονται χρησιμοποιώντας προσομοιώσεις υψηλής πιστότητας (όπως AirSim και Gazebo) με πραγματικά μοντέλα πόλεων. Τα αεροσκάφη μαθαίνουν πώς να πετούν με ασφάλεια κατά τη διάρκεια αυτής της εκπαίδευσης, λαμβάνοντας υπόψη περιορισμούς όπως το μέγιστο επιτρεπόμενο υψόμετρο, τις ασφαλείς αποστάσεις από κατασκευές και την ελάχιστη κατανάλωση ενέργειας. Είναι επίσης απαραίτητη η συνεργασία με την υποδομή επικοινωνιών. Μπορούν να επεξεργάζονται δεδομένα τοπικά, ενώ παραμένουν συνδεδεμένα με κεντρικούς σταθμούς ή το cloud για να λαμβάνουν νέες εντολές, να ενημερώνουν χάρτες και να στέλνουν ζωτικές πληροφορίες χάρη στον συνδυασμό των τεχνολογιών 5G και Edge AI. Η χαμηλή καθυστέρηση του 5G καθιστά εφικτή την ανταλλαγή δεδομένων μεταξύ πολλών μη επανδρωμένων αεροχημάτων που λειτουργούν στην ίδια αστική περιοχή βελτιώνοντας τον συντονισμό και την αποφυγή συγκρούσεων.

Υπάρχουν πολλές χρήσεις σε αστικές περιοχές. Η εφαρμογή της βαθιάς ενισχυτικής μάθησης στις εμπορικές παραδόσεις επιτρέπει τη δημιουργία διαδρομών που είναι βελτιστοποιημένες για τη μείωση του χρόνου παράδοσης και της κατανάλωσης ενέργειας. Τα αεροσκάφη μπορούν να χρησιμοποιούν μοντέλα βαθιάς μάθησης για την επεξεργασία εικόνων στην αξιολόγηση υποδομών, προκειμένου να εντοπίζουν ζημιές σε ενεργειακά δίκτυα, κτίρια και γέφυρες. Με την αποστολή φωτογραφιών και αναλυτικών δεδομένων σε πραγματικό χρόνο σε κέντρα ελέγχου, μπορούν να βοηθήσουν στη διαχείριση της κυκλοφορίας, την ανίχνευση πυρκαγιών και την παρακολούθηση του πλήθους στον τομέα της δημόσιας ασφάλειας. Η ασφαλής και αποτελεσματική ενσωμάτωση τους στον αστικό ιστό καθίσταται δυνατή χάρη στις δυνατότητες αντίληψης, σχεδιασμού και εκτέλεσης σε πραγματικό χρόνο, οι οποίες υπόκεινται σε αυστηρούς νομικούς και χωρικούς περιορισμούς.



Σχήμα 5.1: Ταξινόμηση της δομής των εφαρμογών UAV στη διαχείριση αστικού περιβάλλοντος. Οι εφαρμογές κατηγοριοποιούνται σε τρεις κύριους τομείς: (α) Έλεγχος Υποδομών (π.χ. τρισδιάστατη χαρτογράφηση), (β) Αντιμετώπιση Καταστροφών (π.χ. πλημμύρες, σεισμοί) και (γ) Περιβαλλοντικός Έλεγχος (π.χ. ατμοσφαιρική ρύπανση), αναδεικνύοντας τον ρόλο τους στις Έξυπνες Πόλεις [3].

Το Σχήμα 5.1 αναδεικνύει βασικούς τομείς εφαρμογής για τη χρήση τεχνολογιών αιχμής σε αστικά περιβάλλοντα, όπου η χρήση των μη επανδρωμένων εναέριων οχημάτων γίνεται όλο και πιο σημαντική. Συγκεκριμένα, μπορούν να βοηθήσουν στην τρισδιάστατη χαρτογράφηση πόλεων σε πραγματικό χρόνο, στην παρακολούθηση δρόμων, γεφυρών και ψηλών κτιρίων, καθώς και στην αξιολόγηση ζωτικού ηλεκτρικού εξοπλισμού. Σε σύγκριση με τις συμβατικές μεθόδους, μπορούν επίσης να παρέχουν ταχύτερες, ακριβέστερες και ασφαλέστερες επιθεωρήσεις των αστικών υποδομών. Χρησιμοποιούνται για την ανάλυση της άμεσης κατάστασης σε περίπτωση αστικών καταστροφών, όπως σεισμοί και πλημμύρες. Παρέχουν επίσης εκτεταμένη παρακολούθηση του πλήθους για ασφάλεια και βελτιωμένη διαχείριση των πόρων κατά τη διάρκεια καταστάσεων έκτακτης ανάγκης. Ταυτόχρονα μπορούν να βοηθήσουν στην παρακολούθηση της ατμοσφαιρικής ρύπανσης και της ποιότητας του νερού στο πλαίσιο της αστικής οικολογικής παρακολούθησης, παρέχοντας ζωτικές πληροφορίες για τη βελτίωση του βιοτικού επιπέδου.

Η χρήση τους σε αυτούς τους τομείς με διάφορους τρόπους: σε αστικές περιοχές, χρησιμοποιούνται για συνεχή παρακολούθηση και ακριβή χαρτογράφηση· σε αποστολές σμήνους, χρησιμοποιούνται για συλλογική και καταναμημένη επιτήρηση μεγάλων περιοχών· και στα αεροσκάφη σταθερού πτερυγίου, παρέχουν μεγαλύτερη αυτονομία πτήσης και κάλυψη μεγάλων περιοχών, γεγονός που

τα καθιστά κατάλληλα για περιβαλλοντική παρακολούθηση και επιθεωρήσεις μεγάλης κλίμακας.

5.2 Προκλήσεις και περιορισμοί

Η μετάβαση από το ελεγχόμενο περιβάλλον της προσομοίωσης στην πολυπλοκότητα του πραγματικού κόσμου εξακολουθεί να είναι αρκετά δυσχερής, παρά τις ραγδαίες εξελίξεις στην εφαρμογή της ενισχυτικής και της βαθιάς μάθησης. Η πλειονότητα των ερευνών που παρουσιάζουν τις δυνατότητες των τεχνικών διεξάγεται σε προσομοιωμένα περιβάλλοντα με γνωστή ή ελεγχόμενη δυναμική των αεροχημάτων, μεταβλητές περιβάλλοντος και παράγοντες αβεβαιότητας. Στην πράξη ωστόσο, η εφαρμογή γίνεται πιο περίπλοκη λόγω υλικών περιορισμών (π.χ. χρόνοι απόκρισης, υπολογιστικοί πόροι) και αβεβαιότητας (π.χ. ριπές ανέμου). Επιπλέον, για την εκπαίδευση εξελιγμένων αλγορίθμων απαιτείται μεγάλος όγκος δεδομένων, τα οποία συχνά δεν είναι διαθέσιμα ή είναι ανεπαρκούς ποιότητας. Η απόσταση μεταξύ προσομοίωσης και πραγματικότητας εξακολουθεί να αποτελεί σημαντικό εμπόδιο, παρόλο που οι τεχνικές προσομοίωσης καθιστούν δυνατή τη δημιουργία εκτεταμένων συνόλων δεδομένων. Η αναντιστοιχία αυτή μπορεί να οδηγήσει σε ανεπαρκή ή μη ασφαλή συμπεριφορά όταν ο αλγόριθμος εφαρμόζεται σε πραγματικά σενάρια.

5.2.1 Αντιμετώπιση αβεβαιότητας και πραγματικών συνθηκών

Ένα από τα μεγαλύτερα εμπόδια στην εφαρμογή αλγορίθμων της ενισχυτικής και της βαθιάς μάθησης σε πραγματικά περιστατικά είναι η αβεβαιότητα. Η λειτουργία στον πραγματικό κόσμο χαρακτηρίζεται από μη γραμμικές, δυναμικές και συχνά απρόβλεπτες καταστάσεις, ακόμη και αν η πλειονότητα των προσεγγίσεων έχουν δοκιμαστεί σε ιδανικές προσομοιώσεις. Υπάρχουν πολλές διαφορετικές πηγές αβεβαιότητας, οι οποίες μπορούν να χωριστούν σε λειτουργικές, συστημικές και περιβαλλοντικές. Η πτήση επηρεάζεται άμεσα από περιβαλλοντικούς παράγοντες όπως ο άνεμος, οι αλλαγές θερμοκρασίας και οι βροχοπτώσεις. Για παράδειγμα, οι διακυμάνσεις της θερμοκρασίας μπορούν να επηρεάσουν αρνητικά την απόδοση των αισθητήρων ή να μειώσουν τη διάρκεια ζωής της μπαταρίας, ενώ οι ισχυρές ριπές ανέμου προκαλούν μη γραμμικές αλλαγές στην αεροδυναμική συμπεριφορά. Για την εκπαίδευση των μοντέλων βαθιάς και ενισχυτικής μάθησης στη διαχείριση αυτών των παραμέτρων χρησιμοποιείται είτε η τεχνική τυχαιοποίηση πεδίου (domain randomization), στην οποία η εκπαίδευση πραγματοποιείται σε προσομοιώσεις με διαφορετικά περιβαλλοντικά χαρακτηριστικά, είτε η ισχυρή ενισχυτική μάθηση (robust reinforcement learning), στην οποία ο αλγόριθμος ενσωματώνει μηχανισμούς ανθεκτικότητας έναντι απρόβλεπτων διαταραχών.

Μια άλλη σημαντική αιτία απρόβλεπτων καταστάσεων είναι οι απρόβλεπτες δυσλειτουργίες του εξοπλισμού. Μπορεί να προκύψουν βλάβες ή αστάθειες στους αισθητήρες, στα συστήματα ώθησης (κινητήρες, έλικες) ή στις γραμμές επικοινωνίας. Όταν, για παράδειγμα, ο αισθητήρας GPS υποστεί απότομη απώλεια σήματος, ένα σύστημα ενισχυτικής μάθησης που έχει εκπαιδευτεί υπό ιδανικές συνθήκες μπορεί να αποτύχει εντελώς. Προκειμένου να καταστεί δυνατό το αεροσκάφος να συνεχίσει την αποστολή του σε περίπτωση βλάβης, η έρευνα σε αυτόν τον τομέα έχει επικεντρωθεί σε συστήματα ελέγχου ανεκτικά σε σφάλματα (fault-tolerant control), που χρησιμοποιούν βαθιά μάθηση, όπου τα νευρωνικά δίκτυα περιέχουν αλγόριθμους αναδιαμόρφωσης. Επιπλέον, η πρόβλεψη της αβεβαιότητας παρουσιάζει μια μοναδική δυσκολία. Η πραγματική δυναμική ενός αεροσκάφους δεν είναι ποτέ πλήρως κατανοητή, ακόμη και με τις προόδους στη φυσική προσομοίωση και την

αεροδυναμική. Τα χαρακτηριστικά πτήσης μπορούν να αλλάξουν δραστικά από μικρές αλλαγές στο φορτίο, τη φθορά των εξαρτημάτων ή την κατασκευή. Σε αυτόν τον τομέα έχουν προταθεί τεχνικές βασισμένες σε μοντέλα ενισχυτικής μάθησης, με στόχο τη συνεχή προσαρμογή του αλγορίθμου στις νέες συνθήκες μέσω της εκμάθησης ενός εσωτερικού μοντέλου του περιβάλλοντος σε πραγματικό χρόνο.

Η αισθητηριακή αβεβαιότητα είναι μια επιπλέον διάσταση. Οι μετρήσεις των αισθητήρων που βρίσκονται επί του οχήματος υπόκεινται πάντα σε θόρυβο, καθυστερήσεις και παρεκκλίσεις. Τα συστήματα βαθιάς μάθησης που βασίζονται σε αισθητηριακές εισόδους για την πλοήγηση (όπως κάμερες RGB-D και οπτική ροή) πρέπει να διαθέτουν τρόπους φιλτραρίσματος και εξαγωγής αξιόπιστων χαρακτηριστικών. Έχει προταθεί ότι μέθοδοι όπως τα μπαγεςιανά νευρωνικά δίκτυα (bayesian neural networks) [54] ή αρχιτεκτονικές με ενσωματωμένη εκτίμηση αβεβαιότητας μπορούν να προσφέρουν τόσο ποσοτικά μέτρα εμπιστοσύνης όσο και προβλέψεις. Οι δυσλειτουργίες στην επικοινωνία αποτελούν σημαντικό παράγοντα, ιδίως σε περιπτώσεις συνεργασίας αεροσκαφών. Η αποτελεσματικότητα των αλγορίθμων συνεργασίας μπορεί να μειωθεί λόγω απώλειας πακέτων, καθυστερήσεων και περιορισμών στο εύρος ζώνης. Επειδή περιορίζουν την εξάρτηση από τη συνεχή και αξιόπιστη επικοινωνία, οι τεχνικές ενισχυτικής μάθησης που περιλαμβάνουν αποκεντρωμένες πολιτικές (decentralized policies) και τοπικές διαδικασίες λήψης αποφάσεων (local decision-making) έχουν δείξει μεγαλύτερη ανθεκτικότητα από αυτή την άποψη.

Συνολικά, η αντιμετώπιση της αβεβαιότητας στα πραγματικά μη επανδρωμένα αεροχήματα απαιτεί ένα συνδυασμό μεθόδων για εκπαίδευση σε μεταβαλλόμενες συνθήκες, δημιουργία αξιόπιστων και ανθεκτικών σε σφάλματα μοντέλων, και ενσωμάτωση αρχιτεκτονικών που υπολογίζουν και αξιοποιούν την αβεβαιότητα. Αυτές οι κατευθυντήριες γραμμές είναι απαραίτητες για τη μετάβαση από την εργαστηριακή έρευνα σε ασφαλείς και αξιόπιστες εφαρμογές στον πραγματικό κόσμο.

5.2.2 Ενεργειακή αποδοτικότητα και περιορισμοί υλικού

Η ενεργειακή αποδοτικότητα αποτελεί κρίσιμο παράγοντα στον σχεδιασμό και την λειτουργία των UAVs, καθώς καθορίζει άμεσα τη διάρκεια πτήσης, το ωφέλιμο φορτίο και την αξιοπιστία των αποστολών. Η εφαρμογή αλγορίθμων ενισχυτικής και βαθιάς μάθησης, οι οποίοι χαρακτηρίζονται από υψηλές απαιτήσεις σε πόρους, επηρεάζεται από τους περιορισμούς του υλικού εξοπλισμού, είτε αυτοί σχετίζονται με την μπαταρία είτε με την υπολογιστική ισχύ των ενσωματωμένων συστημάτων. Η χωρητικότητα της μπαταρίας των μη επανδρωμένων αεροχημάτων, ιδιαίτερα των μικρών πολυρότορων, είναι περιορισμένη και συνήθως εξαντλείται εντός 20 έως 40 λεπτών από την απογείωση. Η αυτονομία περιορίζεται περαιτέρω από τη μεγάλη αύξηση της κατανάλωσης ενέργειας που προκαλείται από την εκτέλεση περίπλοκων αλγορίθμων σε ενσωματωμένους επεξεργαστές. Τα μεγάλα βαθιά νευρωνικά δίκτυα ή η συνεχής επεξεργασία εικόνων υψηλής ανάλυσης έχουν αποδειχθεί ότι μειώνουν τον χρόνο πτήσης κατά 15-25 % . Ως αποτέλεσμα, δημιουργούνται τεχνικές αρχιτεκτονικής με υψηλή ενεργειακή απόδοση, όπως τα ελαφρού υπολογιστικού φορτίου συνελκτικά δίκτυα (όπως το MobileNet και το SqueezeNet) [55], τα οποία καταναλώνουν λιγότερη ενέργεια διατηρώντας παράλληλα εξαιρετική απόδοση σε εργασίες όπως η ανίχνευση εμποδίων ή η πλοήγηση.

Λόγω των περιορισμών βάρους και ισχύος, σπάνια διαθέτουν ισχυρούς ενσωματωμένους επεξεργαστές σε επίπεδο υπολογιστικής ισχύος. Σε αντίθεση με τις GPU πλήρους κλίμακας, η πλειονότητα των συστημάτων χρησιμοποιεί μικροσκοπικούς ενσωματωμένους υπολογιστές, όπως το Raspberry Pi ή το NVIDIA Jetson Nano/Xavier, οι οποίοι διαθέτουν λιγότερη μνήμη και ισχύ. Ως αποτέλε-

σμα, η πολυπλοκότητα των μοντέλων που μπορούν να εκπαιδευτούν εκτός σύνδεσης και εκείνων που μπορούν να λειτουργούν σε πραγματικό χρόνο κατά την πτήση διαφέρει. Επομένως, ιδιαίτερη προσοχή δίνεται στις στρατηγικές συμπίεσης και βελτιστοποίησης των νευρωνικών δικτύων:

- Η κβάντιση είναι η διαδικασία περιορισμού του ενεργειακού φορτίου μέσω της μείωσης της ακρίβειας των αριθμητικών πράξεων (για παράδειγμα, από 32-bit σε 8-bit).
- Αραίωση μοντέλου (Model pruning), είναι η διαδικασία εξάλειψης περιττών παραμέτρων, διατηρώντας παράλληλα τις πιο σημαντικές.
- Απόσταξη γνώσης (Knowledge distillation), είναι η μεταφορά γνώσης από μεγάλα μοντέλα (teacher) σε μικρότερα (student) που τρέχουν αποδοτικά σε ενσωματωμένο υλικό.

Μια από τις κύριες προκλήσεις είναι η βελτιστοποίηση των αλγορίθμων για λειτουργία σε πραγματικό χρόνο. Για να συγκλίνουν οι αλγόριθμοι ενισχυτικής μάθησης, συχνά απαιτείται μεγάλος αριθμός δειγμάτων και επαναλήψεων. Αυτό δεν είναι δυνατό εν πτήση, γι' αυτό και αναπτύσσονται μέθοδοι όπως η offline-ενισχυτική μάθηση ή η μεταφορά από προσομοίωση σε πραγματικότητα. Αυτές οι μέθοδοι καταναλώνουν λιγότερη ενέργεια και χρόνο, επειδή η εκπαίδευση γίνεται σε προσομοίωση και μόνο η τελική ρύθμιση γίνεται στο πραγματικό αεροσκάφος. Ιδιαίτερη σημασία αποκτά η συνεργασία υλικού-αλγορίθμων. Η δυναμική βελτιστοποίηση της λειτουργίας μπορεί να επιτευχθεί με το σχεδιασμό αλγορίθμων που λαμβάνουν υπόψη τη θερμοκρασία του επεξεργαστή ή την τρέχουσα κατάσταση της μπαταρίας. Για παράδειγμα, ένα μη επανδρωμένο αεροσκάφος μπορεί να εξισορροπήσει την απόδοση και την κατανάλωση τροποποιώντας το frame rate της επεξεργασίας βίντεο ή την ακρίβεια των παρεμβολών, με βάση το επίπεδο ενέργειας. Η πιο γενική κατηγορία υπολογισμών με γνώμονα την ενέργεια περιλαμβάνει αυτές τις προσαρμοστικές μεθόδους.

Τέλος, αναπτύσσεται ένας νέος τομέας του edge computing στα συνεργατικά σμήνη, στον οποίο πολλά εναέρια οχήματα μοιράζονται το φορτίο επεξεργασίας. Ως αποτέλεσμα, τα αεροσκάφη με μικρότερη υπολογιστική ισχύ μπορούν να χειριστούν απλούστερες εργασίες, ενώ οι πιο ισχυρές μονάδες χειρίζονται τα πιο σύνθετα δεδομένα, με αποτέλεσμα την αύξηση της συνολικής αποδοτικότητας. Συμπερασματικά, η επιτυχία των μελλοντικών αυτόνομων εναέριων συστημάτων θα εξαρτηθεί σε μεγάλο βαθμό από τον συνδυασμό ενεργειακά αποδοτικών υλικών και έξυπνων τεχνικών κατανομής πόρων.

5.2.3 Δεδομένα εκπαίδευσης και προσομοίωση

Για να είναι αποτελεσματικοί οι αλγόριθμοι ενισχυτικής μάθησης και βαθιάς μάθησης στα μη επανδρωμένα αεροσκάφη, είναι απαραίτητη τόσο η ποσότητα όσο και η ποιότητα των δεδομένων εκπαίδευσης. Στην πραγματικότητα, η συλλογή πραγματικών δεδομένων πτήσης είναι δαπανηρή, χρονοβόρα και συχνά επικίνδυνη, ιδίως σε επικίνδυνες ή πολύπλοκες τοποθεσίες. Για τον λόγο αυτό, οι προσομοιώσεις αποτελούν αναπόσπαστο μέρος της διαδικασίας εκπαίδευσης και αξιολόγησης, επειδή παρέχουν ένα ασφαλές και ρυθμιζόμενο περιβάλλον για την εκπαίδευση πολιτικών ελέγχου. Τα φυσικά στοιχεία (άνεμος, θερμοκρασία, υψόμετρο) και τα δυναμικά εμπόδια σε αστικές ή φυσικές συνθήκες μπορούν να αναπαρασταθούν πλήρως σε σύγχρονα συστήματα προσομοίωσης όπως το AirSim ή το Gazebo, όπως αναφέρθηκαν σε προηγούμενη ενότητα. Οι προσεγγίσεις βαθιάς ενισχυτικής μάθησης μπορούν να χρησιμοποιηθούν για την εκπαίδευση των αεροχημάτων σε αυτά, επιτρέποντάς τους να τροποποιήσουν τις πολιτικές τους για να βελτιστοποιήσουν την αποδοτικότητα, την ασφάλεια και την ακρίβεια της αποστολής χωρίς να θέτουν πραγματικά σε κίνδυνο.

Η χρήση προσομοιώσεων υποδηλώνει «μεταφορά μάθησης», στην οποία τα μοντέλα που έχουν διδαχθεί σε ψηφιακό περιβάλλον μετατρέπονται σταδιακά σε πραγματικά αεροσκάφη με ελάχιστες δοκιμές στο πεδίο. Η ικανότητα του αλγορίθμου να διαχειρίζεται την αβεβαιότητα σε πραγματικές πτήσεις, βελτιώνεται με την ενσωμάτωση τυχαίων διακυμάνσεων και απρόβλεπτων γεγονότων σε όλη τη διάρκεια της προσομοίωσης.

Πίνακας 5.1: Σύνοψη των κυριότερων προκλήσεων κατά την ανάπτυξη συστημάτων ελέγχου UAV, οι προτεινόμενες μέθοδοι αντιμετώπισης και παραδείγματα εφαρμογής [7].

Υποενότητα	Κύριες προκλήσεις	Τεχνικές / Μέθοδοι	Παράδειγμα / Σχόλιο
5.2.1 Αντιμετώπιση αβεβαιότητας	Περιβαλλοντικές μεταβολές, αστοχίες υλικού	Τυχαία κατανομή πεδίου, ισχυρός έλεγχος, σχεδιασμός ανθεκτικός σε σφάλματα	Πολιτικές που προσαρμόζονται σε διαφορετικές ταχύτητες ανέμου
5.2.2 Ενεργειακή αποδοτικότητα	Περιορισμένη μπαταρία και onboard υπολογισμός	Δίκτυα χαμηλής πολυπλοκότητας, κβάντιση, προγραμματισμός	DDPG για ενεργειακά αποδοτικούς ελιγμούς σε αεροχήματα σταθερού πτερυγίου
5.2.3 Δεδομένα εκπαίδευσης	Έλλειψη πραγματικών δεδομένων, ρίσκο δοκιμών στο πεδίο	Προσομοίωση (AirSim, Gazebo), sim-to-real transfer, εμπλουτισμός δεδομένων	Εκπαίδευση σμηνών σε προσομοίωση με εμπόδια

Ο Πίνακας 5.1 παρέχει μια επισκόπηση των κύριων εμποδίων στη χρήση μεθόδων ενισχυτικής μάθησης και βαθιάς μάθησης για τον έλεγχο μη επανδρωμένων αεροσκαφών. Δεδομένου ότι υπόκεινται σε δυναμικές περιβαλλοντικές αλλαγές (άνεμος, θερμοκρασία, βροχόπτωση) και πιθανές βλάβες υλικού (αισθητήρες, κινητήρες), είναι επιτακτική ανάγκη να αντιμετωπιστεί η αβεβαιότητα. Για να επιτευχθεί αυτό, στρατηγικές όπως οι προσεγγίσεις ισχυρού ελέγχου και domain randomization έχουν αποδειχθεί αποτελεσματικές, επιτρέποντας στους αλγόριθμους βαθιάς ενισχυτικής μάθησης να αποδίδουν καλύτερα σε δυναμικά περιβάλλοντα και να γενικεύονται πιο αποτελεσματικά. Παρόμοια με αυτό, καθώς τα αεροχήματα εξαρτώνται από περιορισμένης χωρητικότητας μπαταρίες και ελαφριά υπολογιστικά συστήματα, η ενεργειακή αποδοτικότητα και οι περιορισμοί της τεχνολογίας επί του αεροχήματος αποτελούν σημαντικά ζητήματα. Σε αυτό το πλαίσιο, έχουν προταθεί μέθοδοι όπως κβαντισμός και χαμηλής πολυπλοκότητας αρχιτεκτονικές δικτύων οι οποίες μειώνουν την κατανάλωση ενέργειας και τους υπολογιστικούς πόρους που απαιτούνται χωρίς να θυσιάζουν σε μεγάλο βαθμό την ακρίβεια. Επιπλέον, μέθοδοι όπως η κατανεμημένη επεξεργασία (edge computing) και τα πρωτόκολλα ελέγχου ενεργειακής απόδοσης έχουν εφαρμοστεί σε διάφορες κατηγορίες UAV για να αυξήσουν την αυτονομία τους.

Τέλος, η ευρεία χρήση προσομοιωτών υψηλής πιστότητας, όπως το AirSim και το Gazebo, που επιτρέπουν τη δημιουργία και την αξιολόγηση πολιτικών βαθιάς ενισχυτικής μάθησης σε ελεγχόμενα περιβάλλοντα πριν από την εφαρμογή τους στο πεδίο, είναι αποτέλεσμα της έλλειψης επαρκών δεδομένων εκπαίδευσης και του υψηλού κινδύνου των δοκιμών στον πραγματικό κόσμο. Η αντιμετώπιση του ζητήματος της διαφοράς μεταξύ προσομοίωσης και πραγματικότητας κατέστη δυνατή χάρη στην προσέγγιση της μεταφοράς από το εικονικό στο πραγματικό (sim-to-real transfer) και της τεχνικής του εμπλουτισμού των δεδομένων (data augmentation). Στην πράξη, η εφαρμογή αυτών των τεχνικών έχει καταστήσει δυνατή την επιτυχή δημιουργία σμηνών σε αστικά περιβάλλοντα και περιοχές με πολλά εμπόδια, διακρίνοντας μια αξιοσημείωτη αύξηση της προσαρμοστικότητας και της ανθεκτικότητας. Η πληρότητα του πίνακα διευκολύνει τη μετάβαση σε μια πιο εις βάθος εξέταση κάθε θέματος, τονίζοντας τη στενή σχέση μεταξύ των δεδομένων εκπαίδευσης, των ενεργειακών περιορισμών και της αβεβαιότητας ως των κύριων παραγόντων που επηρεάζουν τον τομέα της έρευνας.

Κεφάλαιο 6

Σύνοψη και μελλοντικές κατευθύνσεις

6.1 Ανακεφαλαίωση βασικών ευρημάτων

Η μελέτη αυτή υπογράμμισε τη σημασία των μη επανδρωμένων αεροσκαφών στο σημερινό τεχνολογικό και ερευνητικό περιβάλλον. Η αυξανόμενη εφαρμογή τους σε τομείς όπως οι μεταφορές, η γεωργία ακριβείας, οι επιχειρήσεις έρευνας και διάσωσης και η επιτήρηση υπογραμμίζει την ανάγκη για συστήματα ελέγχου που ξεπερνούν τους περιορισμούς των συμβατικών τεχνικών. Η πολυπλοκότητα και η απρόβλεπτη φύση των εφαρμογών στον πραγματικό κόσμο απαιτούν τη χρήση πιο «ευφύων» διαδικασιών, αν και οι παραδοσιακές προσεγγίσεις βελτιστοποίησης έχουν αποδειχθεί αξιόπιστες σε απλά ή καλά μοντελοποιημένα περιβάλλοντα. Σε αυτό το πλαίσιο, οι τεχνικές βαθιάς μάθησης και ενισχυτικής μάθησης έχουν καταστεί απαραίτητα εργαλεία για τη δημιουργία ανθεκτικών, προσαρμοστικών και αυτόνομων πολιτικών ελέγχου. Τα βασικά συμπεράσματα του άρθρου υποδεικνύουν όλα το γεγονός ότι τα μη επανδρωμένα αεροχήματα επωφελούνται σε μεγάλο βαθμό από τα πλεονεκτήματα αυτών των τεχνικών.

Η ικανότητα των μη επανδρωμένων αεροσκαφών να αναλύουν περίπλοκα δεδομένα από κάμερες, LiDAR, ραντάρ ή περιβαλλοντικούς αισθητήρες βελτιώθηκε σημαντικά με τη χρήση της βαθιάς μάθησης. Τα βαθιά συνελκτικά ή επαναλαμβανόμενα νευρωνικά δίκτυα επιτρέπουν στα αεροσκάφη να «κατανοούν» τον χώρο και να ανακτούν αμέσως ζωτικά δεδομένα. Αυτή η πρόοδος είναι ιδιαίτερα εμφανής σε εφαρμογές αναγνώρισης αντικειμένων και χαρτογράφησης σε μη δομημένα περιβάλλοντα. Ταυτόχρονα, όμως, η χρήση της βαθιάς μάθησης παρουσιάζει δυσκολίες όσον αφορά την ερμηνευσιμότητα των μοντέλων, την υπολογιστική πολυπλοκότητα και την ανάγκη για τεράστια σύνολα δεδομένων. Η ενισχυτική μάθηση έχει αποτελέσει καθοριστικό παράγοντα στην εξέλιξη των τεχνικών ελέγχου των μη επανδρωμένων αεροσκαφών. Μπορούν να δημιουργήσουν σχέδια πτήσης που μεγιστοποιούν μεταβλητές όπως η χρήση ενέργειας, η σταθερότητα των καιρικών συνθηκών και η αποφυγή εμποδίων μέσω αλληλεπιδράσεων με το περιβάλλον τους. Οι παραδοσιακοί ελεγκτές PID ή LQR έχουν περιορισμούς, αλλά η ικανότητα «μάθησης μέσω δοκιμών και σφαλμάτων» έχει καταστήσει δυνατή τη δημιουργία ευέλικτων πολιτικών. Ωστόσο, η ανάγκη για μεγάλο αριθμό επεισοδίων εκπαίδευσης και η δυσκολία γενίκευσης αποτελούν βασικές προκλήσεις. Η ενσωμάτωση της ενισχυτικής μάθησης με τη βαθιά μάθηση σηματοδότησε μια κρίσιμη στιγμή. Τα βαθιά δίκτυα λειτουργούν ως συναρτήσεις προσέγγισης, επιτρέποντας στα αεροχήματα να διαχειρίζονται καταστάσεις υψηλής διάστασης, όπως δυναμικά σενάρια με πολλά εμπόδια ή μεταβλητές καιρικές συνθήκες. Η βαθιά ενισχυτική μάθηση προσφέρει τεράστια οφέλη, αλλά έχει και μειονεκτήματα, όπως η σταθερότητα της εκπαίδευσης και η πιθανότητα υπερπροσαρμογής σε

προσομοιωμένα περιβάλλοντα που δεν αντικατοπτρίζουν πλήρως την πραγματικότητα.

Ένα από τα πιο αξιοσημείωτα αποτελέσματα της μελέτης είναι η δυναμική φύση των υβριδικών συστημάτων, τα οποία συνδυάζουν κλασικούς ελεγκτές με νεότερες μεθοδολογίες βαθιάς και ενισχυτικής μάθησης. Αυτές οι προσεγγίσεις διατηρούν τη σταθερότητα και την ερμηνευσιμότητα των κλασικών ελεγκτών, ενώ ταυτόχρονα αξιοποιούν την ικανότητα των νευρωνικών δικτύων να μαθαίνουν και να προσαρμόζονται σε καταστάσεις που είναι δύσκολο να μοντελοποιηθούν. Ένα τέτοιο υβριδικό πλαίσιο μπορεί να μειώσει τους κινδύνους που συνδέονται με τις ασταθείς της ενισχυτικής μάθησης, επιτυγχάνοντας ταυτόχρονα έναν καλύτερο συνδυασμό ασφάλειας και ευελιξίας. Ταυτόχρονα, τονίζεται ότι η μετα-μάθηση αποτελεί μια πολύ υποσχόμενη προσέγγιση. Η μετα-μάθηση επιτρέπει στο αεροσκάφος να χρησιμοποιεί εμπειρίες από μια σειρά σεναρίων για να προσαρμοστεί σε νέες συνθήκες πιο γρήγορα από ό,τι με τις τυπικές μεθόδους εκπαίδευσης, οι οποίες απαιτούν χιλιάδες επαναλήψεις σε ένα συγκεκριμένο περιβάλλον. Σε καταστάσεις όπως οι επιχειρήσεις διάσωσης σε καταστροφές, όπου η αβεβαιότητα είναι συχνή, αυτό είναι απαραίτητο. Επιπλέον, η δυνατότητα των συστημάτων πολλαπλών πρακτόρων και της συνεργατικής μάθησης απέκτησε όλο και μεγαλύτερη σημασία. Όπως ήδη αναφέρθηκε, τα σμήνη αντιμετωπίζουν πολύπλοκα ζητήματα συντονισμού, επικοινωνίας και κατανομής πόρων. Η εκπαίδευση καταναμημένων μοντέλων καθίσταται δυνατή με τη χρήση αποκεντρωμένων μεθοδολογιών, οι οποίες εξαλείφουν την ανάγκη συλλογής όλων των δεδομένων σε ένα μόνο σημείο. Εκτός από τη βελτίωση της ασφάλειας και της αποτελεσματικότητας των δεδομένων, αυτό καθιστά δυνατή τη δημιουργία κανόνων συνεργατικής πτήσης που μπορούν να μεγιστοποιήσουν το συνολικό αποτέλεσμα της αποστολής.

Η ανάπτυξη αεροσκαφών που μπορούν να λειτουργούν με ασφάλεια και αξιοπιστία χωρίς άμεση ανθρώπινη επίβλεψη είναι το κύριο ζητούμενο για το μέλλον, όπως αποδεικνύεται από τη μελέτη των πλήρως αυτόνομων λειτουργιών σε μη ελεγχόμενες συνθήκες. Για να διευκολυνθεί η ευρεία υιοθέτηση και χρήση τους σε κρίσιμες αποστολές, αυτή η προσέγγιση απαιτεί όχι μόνο εξελιγμένους αλγόριθμους, αλλά και την ενσωμάτωση στοιχείων ασφάλειας, ερμηνευσιμότητας και πιστοποίησης. Εν κατακλείδι, τα κύρια συμπεράσματα αυτής της μελέτης καταδεικνύουν ότι, ούτε οι παραδοσιακές τεχνικές, ούτε οι αποκλειστικά επικεντρωμένες στα δεδομένα στρατηγικές, μπορούν να αποτελέσουν το μοναδικό μέσο για την επίτευξη «έξυπνου ελέγχου» των αεροχημάτων. Αντίθετα, απαιτείται ένας συνδυασμός μεθόδων για την εξισορρόπηση της ερμηνευσιμότητας, της αποτελεσματικότητας και της ακρίβειας.

6.2 Προτάσεις για μελλοντική έρευνα

Η τεχνολογική πρόοδος στον τομέα του ελέγχου μη επανδρωμένων αεροχημάτων και η ενσωμάτωση των μεθόδων βαθιάς ενισχυτικής μάθησης, έχουν αποδείξει ότι μπορούν να προσφέρουν ευελιξία, προσαρμοστικότητα και βελτιστοποίηση σε μια σειρά από περιπτώσεις χρήσης και καταστάσεις. Αν και ακόμα υπάρχουν ανοιχτές νέες κατευθύνσεις, οι προτάσεις αυτές επιδιώκουν να αντιμετωπίσουν τις αδυναμίες των υφιστάμενων στρατηγικών και να προσφέρουν λύσεις σε κρίσιμες προκλήσεις που εντοπίστηκαν προηγουμένως. Η ταχεία πρόοδος της ενισχυτικής και βαθιάς μάθησης στον έλεγχο αεροχημάτων, έχει ήδη αποδείξει τη δυναμική της, ωστόσο παραμένουν ανοιχτά ερευνητικά ζητήματα που περιορίζουν την πλήρη αξιοποίησή της σε πραγματικά περιβάλλοντα. Εκτός από τη βελτίωση της απόδοσης τους στην πλοήγηση, την αποφυγή εμποδίων ή τις συνεργατικές αποστολές, οι μελλοντικές ερευνητικές τάσεις θα επικεντρωθούν στη δημιουργία συστημάτων που θα είναι πιο ανθεκτικά, ενεργειακά αποδοτικά, και ικανά να λειτουργούν με ελάχιστη ή και

μηδενική ανθρώπινη παρέμβαση. Οι κατευθύνσεις αυτές αναπτύσσονται στρατηγικά με στόχο να μετατρέψουν τα μη επανδρωμένα αεροχήματα από απλά εργαλεία αποστολών σε αυτόνομα, αξιόπιστα, συνεργατικά συστήματα με ποικίλες χρήσεις. Επομένως, οι μελλοντικές μελέτες αναμένεται να γεφυρώσουν το χάσμα μεταξύ των θεωρητικών εξελίξεων και των πραγματικών εφαρμογών.

6.2.1 Βελτίωση αλγοριθμικών μεθόδων

Οι εξελίξεις στην βαθιά ενισχυτική μάθηση εξαρτώνται σε μεγάλο βαθμό από τη δημιουργία πιο ισχυρών και αξιόπιστων υπολογιστικών τεχνικών. Ακόμη και με την επιτυχία μεθόδων όπως PPO, DDPG και SAC, εξακολουθούν να υπάρχουν προβλήματα με την αργή σύγκλιση σε περίπλοκα περιβάλλοντα, την ευαισθησία των υπερπαραμέτρων και την αστάθεια της εκπαίδευσης. Ο κίνδυνος δυσμενών καταστάσεων κατά την πτήση μπορεί να μειωθεί με την ενσωμάτωση προσεγγίσεων κατανεμημένης ενισχυτικής μάθησης που βελτιώνουν την ικανότητα γενίκευσης την ανθεκτικότητα έναντι αβεβαιοτήτων. Επιπλέον, προσφέρει πιο σταθερή συμπεριφορά όταν συνδυάζεται με συμβατικές τεχνικές ελέγχου, ιδιαίτερα κατά τη διάρκεια κρίσιμων περιόδων πτήσης. Για πολλά χρόνια, οι συμβατικές τεχνικές ελέγχου και μη γραμμικές στρατηγικές έχουν χρησιμεύσει ως βάση για την ανάπτυξη ελεγκτών.

Μια πρώτη κατεύθυνση αφορά τη δημιουργία υβριδικών δομών ελέγχου, που συνδυάζουν συμβατικές μεθόδους με μαθησιακές προσεγγίσεις, με τα μοντέλα βαθιάς ενισχυτικής μάθησης να αποκτούν πολιτικές υψηλού επιπέδου, ενώ οι παραδοσιακοί ελεγκτές εκτελούν ακριβή έλεγχο χαμηλού επιπέδου. Προκειμένου να επιτευχθεί ισορροπία μεταξύ προβλεψιμότητας και προσαρμοστικότητας, αναμένεται να διεξαχθούν περισσότερες έρευνες σε τέτοια σχήματα για την αξιόπιστη λειτουργία σε συνθήκες αβεβαιότητας. Η ενοποίηση μπορεί να πάρει διάφορες μορφές :

- **Ιεραρχικές αρχιτεκτονικές ελέγχου:** Η βαθιά ενισχυτική μάθηση διαχειρίζεται αποφάσεις υψηλού επιπέδου (πλοήγηση, αποφυγή εμποδίων, επιλογή διαδρομής), ενώ οι παραδοσιακοί ελεγκτές χειρίζονται καθήκοντα χαμηλού επιπέδου (σταθεροποίηση θέσης, έλεγχος στάσης κ.λπ.). Αυτό διατηρεί την αξιοπιστία του παραδοσιακού ελέγχου, ενώ αξιοποιεί την προσαρμοστικότητα του DRL για στρατηγικές αποφάσεις.
- **Ρύθμιση κέρδους με βάση τη μάθηση:** Μπορεί να είναι δύσκολο να ρυθμιστούν οι παράμετροι ελέγχου (κέρδη) σε συστήματα PID/LQR για διαφορετικά περιβάλλοντα. Η απόδοση μπορεί να βελτιωθεί χωρίς να αντικατασταθεί πλήρως ο ελεγκτής, χρησιμοποιώντας την DRL για να προσαρμόζονται δυναμικά τα κέρδη σε πραγματικό χρόνο.
- **Υπολειμματική ενισχυτική μάθηση:** Σε αυτήν την προσέγγιση μια βασική πολιτική εφαρμόζεται από έναν κλασικό ελεγκτή και η βαθιά ενισχυτική μάθηση εκπαιδεύεται είτε για να «διορθώσει» το σφάλμα είτε για να προσθέσει έναν υπολειμματικό (residual) όρο που βελτιώνει τη συμπεριφορά σε δύσκολες καταστάσεις. Με αυτόν τον τρόπο, η ικανότητα μάθησης συνδυάζεται με τη σταθερότητα του κλασικού ελέγχου. Με την προσέγγιση αυτή επιτρέπεται ταχύτερη σύγκλιση, μεγαλύτερη σταθερότητα και ασφαλή μάθηση, αφού ο κλασικός ελεγκτής παρέχει ένα βασικό, αξιόπιστο πλαίσιο λειτουργίας.
- **Μοντέλο προγνωστικού ελέγχου με ενσωμάτωση βαθιάς μάθησης:** Προκειμένου να προβλέψει μελλοντικές συνθήκες και να μεγιστοποιήσει τον έλεγχο, το μοντέλο προγνωστικού ελέγχου χρειάζεται ακριβή μοντέλα. Τα μοντέλα βαθιάς μάθησης μπορούν να παρέχουν πιο ακριβείς προβλέψεις, μαθαίνοντας την ακριβή δυναμική του αεροχήματος.

Αν και οι υβριδικές προσεγγίσεις μειώνουν την ανάγκη για τεράστιους όγκους δεδομένων, σε σύγκριση με τις «καθαρές» τεχνικές μάθησης, εξακολουθούν να απαιτούν πολλές προσομοιώσεις υψηλής πιστότητας για τη βελτιστοποίηση και την εκπαίδευση των αλγορίθμων. Η αναντιστοιχία μεταξύ του περιβάλλοντος προσομοίωσης και των πραγματικών συνθηκών μπορεί να οδηγήσει σε υποβαθμισμένη απόδοση. Επίσης, πρέπει να αντιμετωπίσουν και τον περιορισμό των υπολογιστικών πόρων. Για την ενσωμάτωση μοντέλων βαθιάς μάθησης σε εναέρια οχήματα με περιορισμένη μνήμη, ισχύ και υπολογιστικές δυνατότητες απαιτούνται αποδοτικές αρχιτεκτονικές. Εφόσον οι τρέχουσες προκλήσεις ξεπεραστούν μέσω της προόδου σε αλγοριθμικές, υπολογιστικές και θεωρητικές κατευθύνσεις, η υβριδική λογική δύναται να εξελιχθεί σε βασικό παράδειγμα ελέγχου για τα μη επανδρωμένα αεροσκάφη του μέλλοντος.

6.2.2 Εκπαίδευση σε πραγματικό χρόνο και αποδοτικότητα δεδομένων

Οι μέθοδοι βαθιάς ενισχυτικής μάθησης απαιτεί συχνά εκατομμύρια επεισόδια, κάτι που είναι πρακτικά ανέφικτο σε πραγματικά αεροχήματα λόγω του κόστους, των χρονικών περιορισμών και της φθοράς των υλικών. Μια από τις πιο πρωτοποριακές προσεγγίσεις για την αύξηση της αυτονομίας είναι η εφαρμογή της μετα-μάθησης (Meta learning). Επιδιώκει να δημιουργήσει αλγόριθμους που μπορούν να προσαρμοστούν γρήγορα σε νέες συνθήκες ή εργασίες με λίγα δεδομένα, σε αντίθεση με τα παραδοσιακά μοντέλα ενισχυτικής μάθησης που απαιτούν μακρές φάσεις εκπαίδευσης σε ένα συγκεκριμένο περιβάλλον. Σε τεχνικό επίπεδο, μία από τις πιο διαδεδομένες προσεγγίσεις είναι η μετα-μάθηση ανεξάρτητη από το μοντέλο (Model-Agnostic meta learning - MAML), η οποία βελτιστοποιεί τις παραμέτρους ενός νευρωνικού δικτύου ώστε αυτό να προσαρμόζεται σε νέες εργασίες. Επιπλέον, μέθοδοι όπως η μετα-ενισχυτική μάθηση (meta-RL) επιτρέπουν στον πράκτορα να αναγνωρίζει άμεσα τις αλλαγές στο περιβάλλον και να κάνει τροποποιήσεις ανάλογα τη πολιτική του. Η έρευνα μπορεί να επικεντρωθεί σε πλαίσια προσαρμογής μερικών λήψεων (few-shot adaptation frameworks) για δύσκολες καταστάσεις, και με τη δημιουργία μηχανισμών προσαρμογής σε πραγματικό χρόνο, η συμπεριφορά του αεροχήματος μπορεί να βελτιώνεται συνεχώς κατά την εκτέλεση αποστολών, εξαλείφοντας την ανάγκη για εντατική προ-εκπαίδευση. Η εκμάθηση με λίγες δοκιμές (FewShot Learning - FSL), μια εξελιγμένη μέθοδος για την εκπαίδευση μοντέλων βαθιάς εκμάθησης με περιορισμένο αριθμό δειγμάτων εκπαίδευσης, έχει πρόσφατα επιδείξει κορυφαία απόδοση σε βιομηχανικές εφαρμογές και κατηγοριοποίηση εικόνων [43]. Συγκεκριμένα, ο όρος «few» μπορεί να μειωθεί σε ένα ή ακόμα και σε μηδέν, υποδηλώνοντας ότι τα μοντέλα μπορούν να προβλέψουν τις ετικέτες των στόχων, με μόνο ένα ή ακόμα και χωρίς δείγματα εκπαίδευσης ανά κατηγορία χωρίς υπερπροσαρμογή.

Εξίσου σημαντική είναι η σύνδεση μεταξύ προσομοίωσης και πραγματικότητας (μεταφορά από την προσομοίωση στην πραγματικότητα). Τεχνικές όπως η προσαρμογή τομέα και η τυχαιοποίηση τομέα επιχειρούν να γεφυρώσουν το χάσμα, για παράδειγμα, αλλάζοντας τυχαία τη φυσική, τον αέρα ή τον φωτισμό της προσομοίωσης, έτσι ώστε το μοντέλο να είναι πιο εφαρμόσιμο σε σενάρια του πραγματικού κόσμου. Ως αποτέλεσμα, το σύστημα έχει την ικανότητα να γενικεύει και να ανταποκρίνεται σε καταστάσεις που δεν καλύφθηκαν πλήρως κατά τη διάρκεια της εκπαίδευσης. Επιπλέον, η on-board εκπαίδευση σε πραγματικό χρόνο (online RL) είναι ένας τομέας με τεράστιες προοπτικές. Υπό ασφαλείς συνθήκες, η ικανότητα των μη επανδρωμένων αεροσκαφών να μαθαίνουν κατά τη διάρκεια της πτήσης μπορεί τελικά να οδηγήσει σε πλήρως αυτόνομα συστήματα που βελτιώνονται συνεχώς με βάση τις εμπειρίες τους. Η προοδευτική ανάπτυξη της αυτονομίας υπό επίβλεψη, αποτελεί μια ρεαλιστική στρατηγική. Τα πλήρως αυτόνομα συστήματα προσφέρουν

πολλά πλεονεκτήματα, όπως τη δυνατότητα εκτέλεσης πολύπλοκων εργασιών χωρίς συνεχή ανθρώπινη επίβλεψη, ταχύτερους χρόνους απόκρισης σε απρόβλεπτες καταστάσεις και ενεργειακή απόδοση μέσω της συνεχούς αξιολόγησης και τροποποίησης της πτήσης.

6.2.3 Προσαρμογή σε αβεβαιότητα και δυναμικά περιβάλλοντα

Η λειτουργία των μη επανδρωμένων αεροσκαφών σε πραγματικές συνθήκες συνοδεύεται αναπόφευκτα από αβεβαιότητες που προκύπτουν τόσο από τη φύση των ίδιων των αισθητήρων όσο και από τις μεταβαλλόμενες περιβαλλοντικές συνθήκες. Οι μετρήσεις από IMU, GPS και κάμερες είναι συχνά θορυβώδεις, με αποτέλεσμα η εκτίμηση της κατάστασης να αποκλίνει από την πραγματικότητα. Επιπλέον, οι ατμοσφαιρικές συνθήκες, όπως οι ισχυροί άνεμοι, οι βροχοπτώσεις ή τα θερμικά ρεύματα, δημιουργούν δυναμικά σενάρια που είναι δύσκολο να προβλεφθούν εκ των προτέρων. Η βαθιά ενισχυτική μάθηση, αν και είναι ικανή να μαθαίνει πολύπλοκες πολιτικές πλοήγησης και ελέγχου, συχνά υπερ-εξειδικεύεται στα χαρακτηριστικά του περιβάλλοντος εκπαίδευσης. Αυτό σημαίνει ότι ένα μη επανδρωμένο αεροσκάφος που έχει εκπαιδευτεί σε ένα συγκεκριμένο προσομοιωμένο σενάριο μπορεί να αποτύχει εντελώς όταν αντιμετωπίζει νέες, άγνωστες καταστάσεις.

Προκειμένου να διασφαλιστεί ότι η πολιτική παραμένει σταθερή ακόμη και σε περίπτωση απρόβλεπτων περιστάσεων, η έρευνα έχει στραφεί προς ισχυρές τεχνικές ενισχυτικής μάθησης που στοχεύουν να συμπεριλάβουν μοντέλα αβεβαιότητας (π.χ. μέσω Bayesian προσεγγίσεων). Η προσαρμογή σε δυναμικά περιβάλλοντα εξακολουθεί να αποτελεί πρόκληση παρά τις σημαντικές εξελίξεις, καθώς απαιτεί την εξισορρόπηση μεταξύ εξειδίκευσης και γενίκευσης, η οποία έχει υψηλό κόστος όσον αφορά τα δεδομένα και τους πόρους επεξεργασίας. Όπως και σε άλλες περιπτώσεις, τεχνικές όπως η προσαρμογή τομέα και η τυχαιοποίηση τομέα έχουν προταθεί για την αντιμετώπιση αυτών των προβλημάτων. Για να καταστεί η πολιτική που μαθαίνει ο πράκτορας πιο εφαρμόσιμη σε πραγματικές καταστάσεις, η προσομοίωση «εμπλουτίζεται» στο πρώτο σενάριο με ένα ευρύ φάσμα χαρακτηριστικών, όπως τυχαίες αλλαγές φωτισμού, θέσεις εμποδίων και ακόμη και φυσικά χαρακτηριστικά. Στο δεύτερο σενάριο, οι στρατηγικές τελειοποιούνται χρησιμοποιώντας μικρά δείγματα πραγματικών δεδομένων, σε μια προσπάθεια μεταφοράς γνώσης από τον προσομοιωμένο στον πραγματικό χώρο.

6.2.4 Ενεργειακή αποδοτικότητα και υπολογιστικοί περιορισμοί

Η έλλειψη ενέργειας και υπολογιστικής ισχύος αποτελεί ένα από τα κύρια εμπόδια στην ανάπτυξη πλήρως αυτόνομων αεροχημάτων. Η πλειονότητα τους τροφοδοτείται από μπαταρίες περιορισμένης χωρητικότητας, οι οποίες επηρεάζονται από το ωφέλιμο φορτίο, το περιβάλλον και τη διαχείριση ισχύος του υποσυστήματος. Οι κινητήρες με μπαταρία προσφέρουν μεγάλο χρόνο πτήσης και μεγαλύτερη εμβέλεια, καθώς είναι πιο ευέλικτοι. Αν και ανθεκτικοί, οι κινητήρες εσωτερικής καύσης είναι σημαντικά βαρύτεροι και χρησιμοποιούνται μόνο σε μοντέλα σταθερών πτερύγων, γεγονός που περιορίζει τη χρήση τους. Ωστόσο, και οι δύο έχουν πρόβλημα στην παροχή ρεύματος αιχμής όταν χρειάζεται, και η παροχή αυτή εξαντλεί γρήγορα την πηγή. Οι πολιτικές βαθιάς ενισχυτικής μάθησης είναι εξαιρετικά δαπανηρές από υπολογιστική άποψη για την εκπαίδευση και την ανάπτυξη, γεγονός που τις καθιστά δύσκολο να εκτελεστούν σε CPU χαμηλής ισχύος ή σε ελαφριά ενσωματωμένα υπολογιστικά συστήματα όπως το NVIDIA Jetson. Ως αποτέλεσμα, πολλά προηγμένα συστήματα συνεχίζουν να βασίζονται σε ισχυρά clusters ή επίγειους σταθμούς για την εξαγωγή συμπερασμάτων και την εκπαίδευση, γεγονός που περιορίζει την αυτονομία τους.

Τα μοντέλα μπορούν να λειτουργούν με λιγότερη μνήμη και επεξεργαστική ισχύ, δημιουργώντας ενεργειακά αποδοτικούς αλγόριθμους με μεθόδους κβαντοποίησης και απόσταξη γνώσης (knowledge distillation). Παράλληλα, η ενσωμάτωση ενεργειακών μετρικών μέσα στην συνάρτηση-στόχος της ενισχυτικής μάθησης, επιτρέπει την εκμάθηση πολιτικών που εξισορροπούν την αποδοτικότητα καυσίμων/ενέργειας με την επιτυχία της αποστολής. Σε αυτή την περίπτωση, τα υβριδικά συστήματα προσφέρουν γενικά ένα σημαντικό πλεονέκτημα σε σύγκριση με όλα τα άλλα συστήματα. Για να αξιοποιήσουν τα μοναδικά πλεονεκτήματα που προσφέρει κάθε πηγή ενέργειας, μπορούν να χρησιμοποιούν πολλές πηγές ενέργειας. Μικρά προβλήματα όπως παρατεταμένες περιόδους φόρτισης, σύντομη διάρκεια πτήσης, ανεπαρκής παροχή ισχύος κ.λπ. επιλύονται επίσης συχνά με αυτές τις λύσεις. Οι υπερπυκνωτές είναι η πιο συχνά χρησιμοποιούμενη πηγή ενέργειας στα υβριδικά συστήματα, καθώς τα πλεονεκτήματά τους συνήθως υπερτερούν των μειονεκτημάτων των εναλλακτικών πηγών ενέργειας (ηλιακή ενέργεια, κινητήρες εσωτερικής καύσης).

6.2.5 Συνεργασία και σμήνη

Η συνεργασία μεταξύ πολλών αεροχημάτων διευρύνει τις δυνατότητες σε εφαρμογές όπως η αναζήτηση και διάσωση, η εκτεταμένη επιτήρηση και οι στρατιωτικές δραστηριότητες. Για τα σμήνη μη επανδρωμένων αεροσκαφών είναι απαραίτητες τεχνικές που ενσωματώνουν αυτονομία, επικοινωνία και κατανομημένο έλεγχο. Σε αυτόν τον τομέα, η βαθιά ενισχυτική μάθηση έχει δείξει σημαντικές προοπτικές, ιδίως με την έλευση της ενισχυτικής μάθησης πολλαπλών πρακτόρων. Αυτές οι μέθοδοι χρησιμοποιούν είτε ανεξάρτητες είτε κοινές ανταμοιβές για να διδάξουν σε κάθε αεροσκάφος μια πολιτική συντονισμού ή μια ατομική πολιτική. Το πρόβλημα εδώ είναι η διαχείριση της πολυπλοκότητας. Καθώς ο αριθμός των αεροσκαφών αυξάνεται, οι αβεβαιότητες στις αλληλεπιδράσεις και την επικοινωνία καθιστούν την κατάσταση σημαντικά πιο δύσκολη.

Για αποστολές που απαιτούν κατανομή εργασίας, όπως η χαρτογράφηση μιας μεγάλης έκτασης, είναι ιδιαίτερα σημαντικό να εφαρμόζονται πολιτικές συνεργασίας που μεγιστοποιούν την αποδοτικότητα. Έχουν αναπτυχθεί προηγμένες μεθοδολογίες, όπως η κεντρικοποιημένη εκπαίδευση με αποκεντρωμένη εκτέλεση, για την εκπαίδευση των μη επανδρωμένων αεροχημάτων χρησιμοποιώντας ένα κοινό μοντέλο, ενώ παράλληλα επιτρέπεται η αυτόνομη λήψη αποφάσεων κατά τη διάρκεια των επιχειρήσεων. Επιπλέον, διερευνώνται μηχανισμοί συνεργατικής αντίληψης, που επι-

τρέπουν στα αεροχήματα να ανταλλάσσουν δεδομένα αισθητήρων για μια πιο ολοκληρωμένη κατανόηση του περιβάλλοντος. Ωστόσο, η αποτελεσματικότητα αυτών των τεχνολογιών σε πρακτικά σενάρια είναι περιορισμένη; καθυστερήσεις στην επικοινωνία, απώλεια πακέτων και περιορισμοί στο εύρος ζώνης ενδέχεται να θέσουν σε κίνδυνο ολόκληρη την αποστολή.

6.3 Συμπέρασμα

Η εργασία αυτή τόνισε τη κρίσιμη σημασία της ενισχυτικής και της βαθιάς μάθησης στη διαχείριση και τη λειτουργία των μη επανδρωμένων αεροσκαφών. Η μελέτη έδειξε ότι οι σύγχρονες τεχνικές μηχανικής μάθησης παρέχουν ευελιξία, προσαρμοστικότητα και ικανότητα γενίκευσης σε σύνθετα και αβέβαια περιβάλλοντα, σε αντίθεση με τις κλασικές μεθόδους ελέγχου που βασίζονται σε αυστηρά μοντελοποιημένη δυναμική και απαιτούν σημαντική παραμετροποίηση. Η μετάβαση από τα μη επανδρωμένα αεροχήματα που λειτουργούν υπό στενή ανθρώπινη εποπτεία σε αυτά που μπορούν να λαμβάνουν αποφάσεις από μόνα τους σε πραγματικό χρόνο, εξαρτάται σε μεγάλο βαθμό από αυτή την ικανότητα. Τα αποτελέσματα καταλήγουν σε τρία βασικά συμπεράσματα. Πρώτον, η χρήση της ενισχυτικής μάθησης επιτρέπει την αυτόνομη βελτιστοποίηση των τεχνικών ελέγχου με βάση την εμπειρία και την αλληλεπίδραση με το περιβάλλον, παρέχοντας απαντήσεις σε ζητήματα για τα οποία δεν υπάρχουν ακριβή μαθηματικά μοντέλα. Δεύτερον, οι τεχνικές βαθιάς μάθησης βελτιώνουν τις αντιληπτικές ικανότητες των αεροχημάτων μέσω της επεξεργασίας και της ερμηνείας δεδομένων υψηλής διάστασης, συμπεριλαμβανομένων εικόνων και μετρήσεων αισθητήρων. Τρίτον, ο συνδυασμός βαθιάς και ενισχυτικής μάθησης έχει ως αποτέλεσμα συστήματα που είναι σε θέση να μαθαίνουν και να δημιουργούν «συμπεριφορές» που βελτιώνονται συνεχώς, αντί να εκτελούν απλώς προγραμματισμένες εντολές.

Η διερεύνηση των ζητημάτων έδειξε ταυτόχρονα ότι η πλήρης ενσωμάτωση αυτών των τεχνολογιών αντιμετωπίζει πολύπλευρα εμπόδια. Επειδή οι αστοχίες μπορεί να οφείλονται σε διαφορές μεταξύ των προσομοιωμένων και των πραγματικών δεδομένων, το πρόβλημα της μεταφοράς από την προσομοίωση στην πραγματικότητα παραμένει κεντρικό. Επιπλέον, για τα αεροχήματα με αυστηρούς περιορισμούς ενέργειας και χρόνου, οι υπολογιστικές απαιτήσεις αποτελούν σημαντικό εμπόδιο. Τέλος, καθώς η ανεξέλεγκτη λήψη αποφάσεων από «μαύρα κουτιά» εγείρει σοβαρές ανησυχίες σχετικά με την αξιοπιστία και την υιοθέτησή τους σε κρίσιμες εφαρμογές, η ασφάλεια και η εξηγησιμότητα των μοντέλων γίνονται όλο και πιο σημαντικά ζητήματα. Η υβριδική προσέγγιση, η οποία συνδυάζει κλασικούς ελεγκτές με αλγόριθμους μηχανικής μάθησης, φαίνεται να είναι η πιο πρακτική μέθοδος για ασφαλή και σταθερή ενσωμάτωση σε συστήματα πραγματικού χρόνου. Η ευελιξία και η γνωστική ικανότητα της μηχανικής μάθησης συνδυάζονται με τη σταθερότητα και την αξιοπιστία των παραδοσιακών μεθόδων για τη δημιουργία συστημάτων που είναι εξαιρετικά ευφυή και συμπεριφέρονται με ελεγχόμενο τρόπο. Ταυτόχρονα, νέες δυνατότητες καθίστανται εφικτές χάρη στην ανάπτυξη τεχνικών συνεργατικής μάθησης και μετα-μάθησης, οι οποίες επιτρέπουν τη κατανομημένη νοημοσύνη μεταξύ πολλών UAV και τη γρήγορη προσαρμοστικότητα σε νέες καταστάσεις.

Ίσως η πιο φιλόδοξη πιθανότητα, αλλά και η πιο ενδιαφέρουσα, είναι η δυνατότητα πλήρως αυτόνομων μη επανδρωμένων αεροσκαφών που λειτουργούν σε μη ελεγχόμενες καταστάσεις. Για την πρακτική εφαρμογή της απαιτούνται διάφορες εξελίξεις, από τη δημιουργία νέων αλγορίθμων που εγγυώνται ανθεκτικότητα σε αβεβαιότητες έως τον σχεδιασμό ενεργειακά αποδοτικών αρχιτεκτονικών που επιτρέπουν την εκτέλεση σύνθετων υπολογισμών απευθείας στο ίδιο το αεροσκάφος.

Συμπερασματικά, η παρούσα εργασία δείχνει ότι η βαθιά μάθηση και η ενισχυτική μάθηση λειτουργούν ως καταλύτης για την πρόοδο των παρόντων τεχνικών ελέγχου των μη επανδρωμένων εναέριων οχημάτων, αντί να τις αντικαθιστούν πλήρως. Η σύγκλιση αυτών των τεχνολογιών δημιουργεί ένα νέο πρότυπο ευφυούς ελέγχου, στο οποίο συνυπάρχουν η αποδοτικότητα, η ασφάλεια και η αυτονομία. Στο μέλλον, μπορεί να δημιουργηθεί ένα εξελιγμένο οικοσύστημα μη επανδρωμένων αεροσκαφών που θα είναι τόσο τεχνικά προηγμένο όσο και αξιόπιστο, εξυπηρετώντας εφαρμογές από την πολιτική αεροπορία και τη βιομηχανία έως την πολιτική προστασία και την έρευνα και διάσωση.

Βιβλιογραφία

- [1] A. Carrio, C. Sampedro, A. Rodriguez-Ramos και P. Campoy, “A Review of Deep Learning Methods and Applications for Unmanned Aerial Vehicles”, *Journal of Sensors*, τόμ. 2017, αρθμ. 1, σ. 3 296 874, 2017. DOI: <https://doi.org/10.1155/2017/3296874> eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1155/2017/3296874>. διεύθν.: <https://onlinelibrary.wiley.com/doi/abs/10.1155/2017/3296874>
- [2] S. A. H. Mohsan, M. A. Khan, F. Noor, I. Ullah και M. H. Alsharif, “Towards the Unmanned Aerial Vehicles (UAVs): A Comprehensive Review”, *Drones*, τόμ. 6, αρθμ. 6, 2022, ISSN: 2504-446X. DOI: 10.3390/drones6060147 διεύθν.: <https://www.mdpi.com/2504-446X/6/6/147>
- [3] N. Elmeseiry, N. Alshaer και T. Ismail, “A Detailed Survey and Future Directions of Unmanned Aerial Vehicles (UAVs) with Potential Applications”, *Aerospace*, τόμ. 8, αρθμ. 12, 2021, ISSN: 2226-4310. DOI: 10.3390/aerospace8120363 διεύθν.: <https://www.mdpi.com/2226-4310/8/12/363>
- [4] F. AlMahamid και K. Grolinger, “Autonomous Unmanned Aerial Vehicle navigation using Reinforcement Learning: A systematic review”, *Engineering Applications of Artificial Intelligence*, τόμ. 115, σ. 105 321, 2022, ISSN: 0952-1976. DOI: <https://doi.org/10.1016/j.engappai.2022.105321> διεύθν.: <https://www.sciencedirect.com/science/article/pii/S095219762200358X>
- [5] S. A. H. Mohsan, N. Q. H. Othman, Y. Li, M. H. Alsharif και M. A. Khan, “Unmanned aerial vehicles (UAVs): practical aspects, applications, open challenges, security issues, and future trends”, *Intelligent Service Robotics*, τόμ. 16, αρθμ. 1, σσ. 109–137, 2023. DOI: 10.1007/s11370-022-00452-4
- [6] Z Fan, “An exploration of reinforcement learning and deep reinforcement learning”, *Applied and Computational Engineering*, τόμ. 73, σσ. 154–159, 2024.
- [7] A. T. Azar κ.ά., “Drone Deep Reinforcement Learning: A Review”, *Electronics*, τόμ. 10, αρθμ. 9, 2021, ISSN: 2079-9292. DOI: 10.3390/electronics10090999 διεύθν.: <https://www.mdpi.com/2079-9292/10/9/999>
- [8] S. Kulkarni και D. D. Patil, “Reinforcement Learning for Autonomous Systems”, στο *2025 4th International Conference on Sentiment Analysis and Deep Learning (ICSADL)*, 2025, σσ. 816–820. DOI: 10.1109/ICSADL65848.2025.10933414
- [9] K. Telli κ.ά., “A Comprehensive Review of Recent Research Trends on Unmanned Aerial Vehicles (UAVs)”, *Systems*, τόμ. 11, αρθμ. 8, 2023, ISSN: 2079-8954. DOI: 10.3390/systems11080400 διεύθν.: <https://www.mdpi.com/2079-8954/11/8/400>

- [10] NASA, *History of Unmanned Aerial Systems*, <https://www.slideserve.com/hoshi/history-of-unmanned-aerial-systems>.
- [11] M. Radovic, *Tech Talk: The 5 Levels of Drone Autonomy*, <https://droneii.com/drone-autonomy>, Accessed: 2025-09-12, 2019.
- [12] N. Al-Iqubaydhi κ.ά., “Deep learning for unmanned aerial vehicles detection: A review”, *Computer Science Review*, τόμ. 51, σ. 100 614, 2024, ISSN: 1574-0137. DOI: <https://doi.org/10.1016/j.cosrev.2023.100614> διεύθυν.: <https://www.sciencedirect.com/science/article/pii/S1574013723000813>
- [13] D. Cabrera, L. Cabrera και E. Cabrera, “The Steps to Doing a Systems Literature Review (SLR)”, *Journal of Systems Thinking Preprints*, 2023. DOI: 10.54120/jost.pr000019.v1
- [14] M. J. Page κ.ά., “The PRISMA 2020 statement: an updated guideline for reporting systematic reviews”, *BMJ*, τόμ. 372, 2021. DOI: 10.1136/bmj.n71 eprint: <https://www.bmj.com/content/372/bmj.n71.full.pdf>. διεύθυν.: <https://www.bmj.com/content/372/bmj.n71>
- [15] D. Tranfield, D. Denyer και P. Smart, “Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review”, *British Journal of Management*, τόμ. 14, αρθμ. 3, 2003. DOI: 10.1111/1467-8551.00375
- [16] M. Abelha, S. Fernandes, D. Mesquita, F. Seabra και A. T. Ferreira-Oliveira, “Graduate Employability and Competence Development in Higher Education—A Systematic Literature Review Using PRISMA”, *Sustainability*, τόμ. 12, αρθμ. 15, 2020, ISSN: 2071-1050. DOI: 10.3390/su12155900 διεύθυν.: <https://www.mdpi.com/2071-1050/12/15/5900>
- [17] L. Gualtieri, E. Rauch και R. Vidoni, “Emerging research fields in safety and ergonomics in industrial collaborative robotics: A systematic literature review”, *Robotics and Computer-Integrated Manufacturing*, τόμ. 67, σ. 101 998, 2021, ISSN: 0736-5845. DOI: <https://doi.org/10.1016/j.rcim.2020.101998> διεύθυν.: <https://www.sciencedirect.com/science/article/pii/S073658452030209X>
- [18] R. S. Sutton και A. G. Barto, *Reinforcement Learning: An Introduction*, Second. MIT press, 2018.
- [19] Y. Zhang, “Path Planning and Control of UAV using Machine Learning and Deep Reinforcement Learning Techniques”, Unpublished, Διδακτορική διατρ., Concordia University, 2022. διεύθυν.: <https://spectrum.library.concordia.ca/id/eprint/991887/>
- [20] W. Shi, Y. Feng, H. Huang, Z. Liu, J. Huang και G. Cheng, “Efficient hierarchical policy network with fuzzy rules”, *International Journal of Machine Learning and Cybernetics*, τόμ. 13, αρθμ. 9, σσ. 447–459, 2022. DOI: 10.1007/s13042-021-01417-2
- [21] W. Xue, H. Wu, H. Ye και S. Shao, “An Improved Proximal Policy Optimization Method for Low-Level Control of a Quadrotor”, *Actuators*, τόμ. 11, αρθμ. 4, 2022, ISSN: 2076-0825. DOI: 10.3390/act11040105 διεύθυν.: <https://www.mdpi.com/2076-0825/11/4/105>
- [22] A. Rouhi κ.ά., “An Overview of Deep Learning in UAV Perception”, στο *2024 IEEE International Conference on Consumer Electronics (ICCE)*, IEEE, 2024, σσ. 1–6.

- [23] T. GUO, N. JIANG, B. LI, X. ZHU, Y. WANG και W. DU, “UAV navigation in high dynamic environments: A deep reinforcement learning approach”, *Chinese Journal of Aeronautics*, τόμ. 34, αρθμ. 2, σσ. 479–489, 2021, ISSN: 1000-9361. DOI: <https://doi.org/10.1016/j.cja.2020.05.011> διεύθν.: <https://www.sciencedirect.com/science/article/pii/S1000936120302247>
- [24] C. Sun κ.ά., “Advancing UAV Communications: A Comprehensive Survey of Cutting-Edge Machine Learning Techniques”, *IEEE Open Journal of Vehicular Technology*, τόμ. 5, σσ. 825–854, 2024. DOI: 10.1109/OJVT.2024.3401024
- [25] R. Arranz, D. Carramiñana, G. d. Miguel, J. A. Besada και A. M. Bernardos, “Application of Deep Reinforcement Learning to UAV Swarming for Ground Surveillance”, *Sensors*, τόμ. 23, αρθμ. 21, 2023, ISSN: 1424-8220. DOI: 10.3390/s23218766 διεύθν.: <https://www.mdpi.com/1424-8220/23/21/8766>
- [26] S. Javaid κ.ά., “Communication and Control in Collaborative UAVs: Recent Advances and Future Trends”, *IEEE Transactions on Intelligent Transportation Systems*, τόμ. 24, αρθμ. 6, σσ. 5719–5739, 2023. DOI: 10.1109/TITS.2023.3248841
- [27] A. Townsend, I. N. Jiya, C. Martinson, D. Bessarabov και R. Gouws, “A comprehensive review of energy sources for unmanned aerial vehicles, their shortfalls and opportunities for improvements”, *Heliyon*, τόμ. 6, αρθμ. 11, e05285, 2020, ISSN: 2405-8440. DOI: <https://doi.org/10.1016/j.heliyon.2020.e05285> διεύθν.: <https://www.sciencedirect.com/science/article/pii/S2405844020321289>
- [28] Y. Sheng, H. Liu, J. Li και Q. Han, “UAV Autonomous Navigation Based on Deep Reinforcement Learning in Highly Dynamic and High-Density Environments”, *Drones*, τόμ. 8, αρθμ. 9, 2024, ISSN: 2504-446X. DOI: 10.3390/drones8090516 διεύθν.: <https://www.mdpi.com/2504-446X/8/9/516>
- [29] L. Wu, C. Wang, P. Zhang και C. Wei, “Deep Reinforcement Learning with Corrective Feedback for Autonomous UAV Landing on a Mobile Platform”, *Drones*, τόμ. 6, αρθμ. 9, 2022, ISSN: 2504-446X. DOI: 10.3390/drones6090238 διεύθν.: <https://www.mdpi.com/2504-446X/6/9/238>
- [30] J. Amendola, L. R. Cenkeramaddi και A. Jha, “Drone Landing and Reinforcement Learning: State-of-Art, Challenges and Opportunities”, *IEEE Open Journal of Intelligent Transportation Systems*, τόμ. 5, σσ. 520–539, 2024. DOI: 10.1109/OJITS.2024.3444487
- [31] C. C. Ekechi, T. Elfouly, A. Alouani και T. Khattab, “A Survey on UAV Control with Multi-Agent Reinforcement Learning”, *Drones*, τόμ. 9, αρθμ. 7, 2025, ISSN: 2504-446X. DOI: 10.3390/drones9070484 διεύθν.: <https://www.mdpi.com/2504-446X/9/7/484>
- [32] Z. Wang και Y. Xing, “Energy consumption optimisation for unmanned aerial vehicle based on reinforcement learning framework”, *International Journal of Powertrains*, τόμ. 13, αρθμ. 1, σσ. 75–94, 2024. DOI: 10.1504/IJPT.2024.138001
- [33] K. Xue, L. Zhai, Y. Li, Z. Lu και W. Zhou, “Task offloading and multi-cache placement based on DRL in UAV-assisted MEC networks”, *Vehicular Communications*, τόμ. 53, σ. 100 900, 2025, ISSN: 2214-2096. DOI: <https://doi.org/10.1016/j.vehcom.2025.100900> διεύθν.: <https://www.sciencedirect.com/science/article/pii/S2214209625000270>

- [34] A. P. Kalidas, C. J. Joshua, A. Q. Md, S. Basheer, S. Mohan και S. Sakri, “Deep Reinforcement Learning for Vision-Based Navigation of UAVs in Avoiding Stationary and Mobile Obstacles”, *Drones*, τόμ. 7, αρθμ. 4, 2023, ISSN: 2504-446X. DOI: 10.3390/drones7040245 διεύθν.: <https://www.mdpi.com/2504-446X/7/4/245>
- [35] B. Lei, W. Hu, Z. Ren και S. Ji, “DRL-Based UAV Autonomous Navigation and Obstacle Avoidance with LiDAR and Depth Camera Fusion”, *Aerospace*, τόμ. 12, αρθμ. 9, 2025, ISSN: 2226-4310. DOI: 10.3390/aerospace12090848 διεύθν.: <https://www.mdpi.com/2226-4310/12/9/848>
- [36] H. Hu, K. Zhang, A. H. Tan, M. Ruan, C. Agia και G. Nejat, “A Sim-to-Real Pipeline for Deep Reinforcement Learning for Autonomous Robot Navigation in Cluttered Rough Terrain”, *IEEE Robotics and Automation Letters*, τόμ. 6, αρθμ. 4, σσ. 6569–6576, 2021. DOI: 10.1109/LRA.2021.3093551
- [37] K. Zhao, Y. Li και Z. Liu, “A Meta-reinforcement Learning Framework for Adaptive Quadrotor UAV Attitude Control”, στο *Parallel and Distributed Computing, Applications and Technologies*, Y. Li, Y. Zhang και J. Xu, επιμελητές, Singapore: Springer Nature Singapore, 2025, σσ. 26–40, ISBN: 978-981-96-4207-6.
- [38] H. Wu, W. Wang, T. Wang και S. Suzuki, “Model-Free UAV Navigation in Unknown Complex Environments Using Vision-Based Reinforcement Learning”, *Drones*, τόμ. 9, αρθμ. 8, 2025, ISSN: 2504-446X. DOI: 10.3390/drones9080566 διεύθν.: <https://www.mdpi.com/2504-446X/9/8/566>
- [39] S. Chen, W. Zhou, A.-S. Yang, H. Chen, B. Li και C.-Y. Wen, “An End-to-End UAV Simulation Platform for Visual SLAM and Navigation”, *Aerospace*, τόμ. 9, αρθμ. 2, 2022, ISSN: 2226-4310. DOI: 10.3390/aerospace9020048 διεύθν.: <https://www.mdpi.com/2226-4310/9/2/48>
- [40] X. Hong, Z. Wang, Y. Wang, C. Xue και Y. Gao, “Multi-UAV Dynamic Target Search Based on Multi-Potential-Field Fusion Reward Shaping MAPPO”, *Drones*, τόμ. 9, αρθμ. 11, 2025, ISSN: 2504-446X. DOI: 10.3390/drones9110770 διεύθν.: <https://www.mdpi.com/2504-446X/9/11/770>
- [41] D. Rossi, G. Borghi και R. Vezzani, *TakuNet: an Energy-Efficient CNN for Real-Time Inference on Embedded UAV systems in Emergency Response Scenarios*, 2025. arXiv: 2501.05880 [cs.CV]. διεύθν.: <https://arxiv.org/abs/2501.05880>
- [42] Q. Zhang, M. D. Furqan, T. Nutzhat, F. Machida και E. Andrade, “Dependability of UAV-Based Networks and Computing Systems: A Survey”, *arXiv preprint arXiv:2506.16786*, 2025, Survey article on dependability of UAV computing and networking systems. arXiv: 2506.16786 [cs.PF]. διεύθν.: <https://arxiv.org/abs/2506.16786>
- [43] P. Zhang, Y. Bai, D. Wang, B. Bai και Y. Li, “Few-Shot Classification of Aerial Scene Images via Meta-Learning”, *Remote Sensing*, τόμ. 13, αρθμ. 1, 2021, ISSN: 2072-4292. DOI: 10.3390/rs13010108 διεύθν.: <https://www.mdpi.com/2072-4292/13/1/108>

- [44] G. Sharma και S. Jain, “Deep Reinforcement Learning-Based Framework for Path Planning of AUAVs”, *Procedia Computer Science*, τόμ. 258, σσ. 1112–1122, 2025, International Conference on Machine Learning and Data Engineering, ISSN: 1877-0509. DOI: <https://doi.org/10.1016/j.procs.2025.04.346> διεύθν.: <https://www.sciencedirect.com/science/article/pii/S1877050925014486>
- [45] M. Theile, A. R. Z. Rodriguez, M. Caccamo και A. L. Sangiovanni-Vincentelli, *Continuous World Coverage Path Planning for Fixed-Wing UAVs using Deep Reinforcement Learning*, 2025. arXiv: 2505.08382 [cs.R0]. διεύθν.: <https://arxiv.org/abs/2505.08382>
- [46] T. T. Nguyen, N. D. Nguyen και S. Nahavandi, “Deep Reinforcement Learning for Multiagent Systems: A Review of Challenges, Solutions, and Applications”, *IEEE Transactions on Cybernetics*, τόμ. 50, αριθμ. 9, σσ. 3826–3839, 2020. DOI: 10.1109/TCYB.2020.2977374
- [47] K. Chikhaoui, H. Ghazzai και Y. Massoud, “PPO-based Reinforcement Learning for UAV Navigation in Urban Environments”, στο *2022 IEEE 65th International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2022, σσ. 1–4. DOI: 10.1109/MWSCAS54063.2022.9859287
- [48] Y. Guo, L. Tang, X. Zhou, J. Zhang και L. Chen, “Vision-Based Indoor UAV Navigation with Deep Reinforcement Learning”, στο *Proceedings of 4th 2024 International Conference on Autonomous Unmanned Systems (4th ICAUS 2024)*, L. Liu, Y. Niu, W. Fu και Y. Qu, επιμελητές, Singapore: Springer Nature Singapore, 2025, σσ. 101–112, ISBN: 978-981-96-3572-6.
- [49] G. A. Jiménez, A. de la Escalera Hueso και M. J. Gómez-Silva, “Reinforcement Learning Algorithms for Autonomous Mission Accomplishment by Unmanned Aerial Vehicles: A Comparative View with DQN, SARSA and A2C”, *Sensors*, τόμ. 23, αριθμ. 21, 2023, ISSN: 1424-8220. DOI: 10.3390/s23219013 διεύθν.: <https://www.mdpi.com/1424-8220/23/21/9013>
- [50] H. Kim, J. Choi, H. Do και G. T. Lee, “A Fully Controllable UAV Using Curriculum Learning and Goal-Conditioned Reinforcement Learning: From Straight Forward to Round Trip Missions”, *Drones*, τόμ. 9, αριθμ. 1, 2025, ISSN: 2504-446X. DOI: 10.3390/drones9010026 διεύθν.: <https://www.mdpi.com/2504-446X/9/1/26>
- [51] Y. Chen, R. Chen, Y. Huang, Z. Xiong και J. Li, “DRL-Based Improved UAV Swarm Control for Simultaneous Coverage and Tracking with Prior Experience Utilization”, *Drones*, τόμ. 8, αριθμ. 12, 2024, ISSN: 2504-446X. DOI: 10.3390/drones8120784 διεύθν.: <https://www.mdpi.com/2504-446X/8/12/784>
- [52] G. Yao, L. Guo, H. Liao και F. Wu, “Fusing Adaptive Game Theory and Deep Reinforcement Learning for Multi-UAV Swarm Navigation”, *Drones*, τόμ. 9, αριθμ. 9, 2025, ISSN: 2504-446X. DOI: 10.3390/drones9090652 διεύθν.: <https://www.mdpi.com/2504-446X/9/9/652>
- [53] J. Li και Y. Liu, “Deep Reinforcement Learning based Adaptive Real-Time Path Planning for UAV”, στο *2021 8th International Conference on Dependable Systems and Their Applications (DSA)*, 2021, σσ. 522–530. DOI: 10.1109/DSA52907.2021.00077

- [54] S. Gong, M. Wang, B. Gu, W. Zhang, D. T. Hoang και D. Niyato, “Bayesian Optimization Enhanced Deep Reinforcement Learning for Trajectory Planning and Network Formation in Multi-UAV Networks”, *IEEE Transactions on Vehicular Technology*, τόμ. 72, αρθμ. 8, σσ. 10 933–10 948, 2023. DOI: 10.1109/TVT.2023.3262778
- [55] A. Kumar, S. Wang, A. M. Shaikh, H. Bilal, B. Lu και S. Song, “Building on prior lightweight CNN model combined with LSTM-AM framework to guide fault detection in fixed-wing UAVs”, *International Journal of Machine Learning and Cybernetics*, τόμ. 15, αρθμ. 9, σσ. 4175–4191, Ιούν. 2024. DOI: 10.1007/s13042-024-02141-3
- [56] B. Guembe, A. Azeta, S. Misra, V. C. Osamor, L. Fernandez-Sanz και V. Pospelova, “The Emerging Threat of Ai-driven Cyber Attacks: A Review”, *Applied Artificial Intelligence*, τόμ. 36, 1 2022, ISSN: 10876545. DOI: 10.1080/08839514.2022.2037254
- [57] Node.js Foundation, *About | Node.js*. επίσκεψη 1 Μάι. 2024. διεύθν.: <https://nodejs.org/en/about>
- [58] GeeksforGeeks, *GeeksforGeeks: A Computer Science Portal for Geeks*, Accessed: 2025-07-09, 2025. διεύθν.: <https://www.geeksforgeeks.org/>
- [59] □. □. Τσιακμάκης, “Βελτίωση της αδρανειακής πλοήγησης UAV με αποδοτική χρήση βαθιάς ενισχυτικής μάθησης”, 2022.
- [60] S.-E. Shen και Y.-C. Huang, “Application of Reinforcement Learning in Controlling Quadrotor UAV Flight Actions”, *Drones*, τόμ. 8, αρθμ. 11, 2024, ISSN: 2504-446X. DOI: 10.3390/drones8110660 διεύθν.: <https://www.mdpi.com/2504-446X/8/11/660>
- [61] Z. Zuo, C. Liu, Q.-L. Han και J. Song, “Unmanned Aerial Vehicles: Control Methods and Future Challenges”, *IEEE/CAA Journal of Automatica Sinica*, τόμ. 9, αρθμ. 4, σσ. 601–614, 2022. DOI: 10.1109/JAS.2022.105410
- [62] S. I. Abdelmaksoud, M. Mailah και A. M. Abdallah, “Control Strategies and Novel Techniques for Autonomous Rotorcraft Unmanned Aerial Vehicles: A Review”, *IEEE Access*, τόμ. 8, σσ. 195 142–195 169, 2020. DOI: 10.1109/ACCESS.2020.3031326
- [63] J. Wang, Z. Yu, D. Zhou, J. Shi και R. Deng, *Vision-Based Deep Reinforcement Learning of UAV Autonomous Navigation Using Privileged Information*, 2024. arXiv: 2412.06313 [cs.R0]. διεύθν.: <https://arxiv.org/abs/2412.06313>
- [64] Z. Xue και T. Gonsalves, “Vision Based Drone Obstacle Avoidance by Deep Reinforcement Learning”, *AI*, τόμ. 2, αρθμ. 3, σσ. 366–380, 2021, ISSN: 2673-2688. DOI: 10.3390/ai2030023 διεύθν.: <https://www.mdpi.com/2673-2688/2/3/23>
- [65] D. Mourtzis, J. Angelopoulos και N. Panopoulos, “Uavs for industrial applications: Identifying challenges and opportunities from the implementation point of view”, *Procedia Manufacturing*, τόμ. 55, σσ. 183–190, 2021.
- [66] Y. Wang, J. Li, X. Yang και Q. Peng, “UAV–Ground Vehicle Collaborative Delivery in Emergency Response: A Review of Key Technologies and Future Trends”, *Applied Sciences*, τόμ. 15, αρθμ. 17, σ. 9803, 2025.

- [67] Y. Gao, M. Liu, X. Yuan, Y. Hu, P. Sun και A. Schmeink, “Federated deep reinforcement learning based trajectory design for UAV-assisted networks with mobile ground devices”, *Scientific Reports*, τόμ. 14, αρθμ. 1, σ. 22 753, 2024. DOI: 10.1038/s41598-024-72654-y διεύθν.: <https://doi.org/10.1038/s41598-024-72654-y>
- [68] W. Koch, R. Mancuso, R. West και A. Bestavros, “Reinforcement Learning for UAV Attitude Control”, *ACM Trans. Cyber-Phys. Syst.*, τόμ. 3, αρθμ. 2, Φεβ. 2019, ISSN: 2378-962X. DOI: 10.1145/3301273 διεύθν.: <https://doi.org/10.1145/3301273>
- [69] Y. Yang, X. Xiong και Y. Yan, “UAV Formation Trajectory Planning Algorithms: A Review”, *Drones*, τόμ. 7, αρθμ. 1, 2023, ISSN: 2504-446X. DOI: 10.3390/drones7010062 διεύθν.: <https://www.mdpi.com/2504-446X/7/1/62>
- [70] S. Feng, X. Li, L. Ren και S. Xu, “Reinforcement learning with parameterized action space and sparse reward for UAV navigation”, *Intelligence & Robotics*, τόμ. 3, αρθμ. 2, 2023, ISSN: 2770-3541. DOI: 10.20517/ir.2023.10 διεύθν.: <https://www.oaepublish.com/articles/ir.2023.10>
- [71] P. Nguyen, H. Nguyen και Q. Ha, “Obstacle avoidance for UAVs using deep reinforcement learning”, *Engineering Letters*, τόμ. 27, αρθμ. 3, σσ. 867–876, 2019. διεύθν.: https://www.engineeringletters.com/issues_v32/issue_3/EL_32_3_20.pdf
- [72] D. Tsiakmakis, N. Passalis και A. Tefas, “Data efficient Deep Reinforcement Learning for robust inertial-based UAV localization”, *Robotics and Autonomous Systems*, τόμ. 194, σ. 105 139, 2025, ISSN: 0921-8890. DOI: <https://doi.org/10.1016/j.robot.2025.105139> διεύθν.: <https://www.sciencedirect.com/science/article/pii/S0921889025002362>
- [73] A. Gupta, K. Sharma και P. Kumar, “SwarmHive: Autonomous UAV Swarms for Coordinated Landing on a Moving Platform”, *arXiv preprint*, 2022. arXiv: 2206.08856 [cs.R0]. διεύθν.: <https://arxiv.org/abs/2206.08856>
- [74] A. Gupta, E. Dorzhieva, A. Baza, M. Alper, A. Fedoseev και D. Tsetserukou, *SwarmHawk: Self-Sustaining Multi-Agent System for Landing on a Moving Platform through an Agent Supervision*, 2022. arXiv: 2206.08874 [cs.R0]. διεύθν.: <https://arxiv.org/abs/2206.08874>
- [75] F. Prochazka, D. Latal και J. Hromkovic, “Nonlinear MPC-based landing strategy for UAVs on moving platforms”, *Ocean Engineering*, τόμ. 310, σ. 117 579, 2025. DOI: 10.1016/j.oceaneng.2024.117579 διεύθν.: <https://www.sciencedirect.com/science/article/abs/pii/S0029801824025022>
- [76] Z. Zhao κ.ά., *Towards Robust Multi-UAV Collaboration: MARL with Noise-Resilient Communication and Attention Mechanisms*, 2025. arXiv: 2503.02913 [cs.MA]. διεύθν.: <https://arxiv.org/abs/2503.02913>
- [77] Y. Gao, L. Ren, T. Shi, T. Xu και J. Ding, “Autonomous Obstacle Avoidance Algorithm for Unmanned Aerial Vehicles Based on Deep Reinforcement Learning”, *Engineering Letters*, τόμ. 32, αρθμ. 3, σσ. 650–660, 2024, Accessed: 2025-09-17. DOI: 10.38028/vol32iss3.2024.20 διεύθν.: https://www.engineeringletters.com/issues_v32/issue_3/EL_32_3_20.pdf

- [78] D. Arce, J. Solano και C. Beltrán, “A Comparison Study between Traditional and Deep-Reinforcement-Learning-Based Algorithms for Indoor Autonomous Navigation in Dynamic Scenarios”, *Sensors*, τόμ. 23, αρθμ. 24, 2023, ISSN: 1424-8220. DOI: 10.3390/s23249672
διεύθυν.: <https://www.mdpi.com/1424-8220/23/24/9672>
- [79] J. Zhou, H. Zhang, M. Hua, F. Wang και J. Yi, “P-DRL: A Framework for Multi-UAVs Dynamic Formation Control under Operational Uncertainty and Unknown Environment”, *Drones*, τόμ. 8, σ. 475, Σεπτ. 2024. DOI: 10.3390/drones8090475
- [80] H. Xu κ.ά., “A Survey on UAV Applications in Smart City Management: Challenges, Advances, and Opportunities”, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, τόμ. 16, σσ. 8982–9010, 2023. DOI: 10.1109/JSTARS.2023.3317500
- [81] J. Tsai, P.-C. Lu και M.-H. Tsai, “Accuracy Improvement of Straight Take-off, Flying Forward and Landing of a Drone with Reinforcement Learning”, στο *2019 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW)*, 2019, σσ. 1–2. DOI: 10.1109/ICCE-TW46550.2019.8991683
- [82] M. G. Mateus, R. B. Grando και P. L. J. Drews, “Active Perception Applied To Unmanned Aerial Vehicles Through Deep Reinforcement Learning”, στο *2022 Latin American Robotics Symposium (LARS), 2022 Brazilian Symposium on Robotics (SBR), and 2022 Workshop on Robotics in Education (WRE)*, 2022, σσ. 1–6. DOI: 10.1109/LARS/SBR/WRE56824.2022.9995830
- [83] P. Sha και Q. Wang, “Autonomous Navigation of UAVs in Resource Limited Environment Using Deep Reinforcement Learning”, στο *2022 37th Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, 2022, σσ. 36–41. DOI: 10.1109/YAC57282.2022.10023581
- [84] C. Park, H. Lee, W. J. Yun, S. Jung και J. Kim, “Coordinated Multi-Agent Reinforcement Learning for Unmanned Aerial Vehicle Swarms in Autonomous Mobile Access Applications”, *arXiv preprint arXiv:2304.08493v1*, 2022. arXiv: 2304.08493v1 [cs.MA].
- [85] X. Mao, G. Wu, M. Fan, Z. Cao και W. Pedrycz, “DL-DRL: A Double-Level Deep Reinforcement Learning Approach for Large-Scale Task Scheduling of Multi-UAV”, *IEEE Transactions on Automation Science and Engineering*, τόμ. 22, σσ. 1028–1044, 2025. DOI: 10.1109/TASE.2024.3358894
- [86] L. A. Fagundes-Junior, K. B. de Carvalho, R. S. Ferreira και A. S. Brandão, “Machine Learning for Unmanned Aerial Vehicles Navigation: An Overview”, *SN Computer Science*, τόμ. 5, αρθμ. 2, σ. 256, 2024. DOI: 10.1007/s42979-023-02592-5
- [87] L. Canese κ.ά., “Multi-Agent Reinforcement Learning: A Review of Challenges and Applications”, *Applied Sciences*, τόμ. 11, αρθμ. 11, 2021, ISSN: 2076-3417. DOI: 10.3390/app11114948
διεύθυν.: <https://www.mdpi.com/2076-3417/11/11/4948>
- [88] W. Gu, K. P. Valavanis, M. J. Rutherford και A. Rizzo, “UAV Model-based Flight Control with Artificial Neural Networks: A Survey”, *Journal of Intelligent & Robotic Systems*, τόμ. 100, αρθμ. 3-4, σσ. 1469–1491, 2020. DOI: 10.1007/s10846-020-01227-8

- [89] E. Aldao, L. M. González-deSantos, H. Michinel και H. González-Jorge, “UAV Obstacle Avoidance Algorithm to Navigate in Dynamic Building Environments”, *Drones*, τόμ. 6, αρθμ. 1, 2022, ISSN: 2504-446X. DOI: 10.3390/drones6010016 διεύθν.: <https://www.mdpi.com/2504-446X/6/1/16>
- [90] F. WANG, X. ZHU, Z. ZHOU και Y. TANG, “Deep-reinforcement-learning-based UAV autonomous navigation and collision avoidance in unknown environments”, *Chinese Journal of Aeronautics*, τόμ. 37, αρθμ. 3, σσ. 237–257, 2024, ISSN: 1000-9361. DOI: <https://doi.org/10.1016/j.cja.2023.09.033> διεύθν.: <https://www.sciencedirect.com/science/article/pii/S1000936123003448>
- [91] S. A. H. Mohsan, N. Q. H. Othman, Y. Li, M. H. Alsharif και M. A. Khan, “Unmanned aerial vehicles (UAVs): practical aspects, applications, open challenges, security issues, and future trends”, *Intelligent Service Robotics*, τόμ. 16, αρθμ. 1, σσ. 109–137, 2023. DOI: 10.1007/s11370-022-00452-4
- [92] S. Sai, A. Garg, K. Jhavar, V. Chamola και B. Sikdar, “A Comprehensive Survey on Artificial Intelligence for Unmanned Aerial Vehicles”, *IEEE Open Journal of Vehicular Technology*, τόμ. 4, σσ. 713–738, 2023. DOI: 10.1109/OJVT.2023.3316181
- [93] J. Kim, S. Kim, C. Ju και H. I. Son, “Unmanned Aerial Vehicles in Agriculture: A Review of Perspective of Platform, Control, and Applications”, *IEEE Access*, τόμ. 7, σσ. 105 100–105 115, 2019. DOI: 10.1109/ACCESS.2019.2932119
- [94] H. Shakhathreh κ.ά., “Unmanned Aerial Vehicles (UAVs): A Survey on Civil Applications and Key Research Challenges”, *IEEE Access*, τόμ. 7, σσ. 48 572–48 634, 2019. DOI: 10.1109/ACCESS.2019.2909530
- [95] D. Wang, Q. Pan, Y. Shi, J. Hu και C. Zhao, “Efficient nonlinear model predictive control for quadrotor trajectory tracking: Algorithms and experiment”, *IEEE Transactions on Cybernetics*, τόμ. 51, αρθμ. 10, σσ. 5057–5068, 2021.
- [96] S. Y. Choi και D. Cha, “Unmanned aerial vehicles using machine learning for autonomous flight; state-of-the-art”, *Advanced Robotics*, τόμ. 33, αρθμ. 6, σσ. 265–277, 2019.
- [97] Y. Song, S. Naji, E. Kaufmann, A. Loquercio και D. Scaramuzza, “Flightmare: A Flexible Quadrotor Simulator”, στο *Proceedings of the 2020 Conference on Robot Learning*, J. Kober, F. Ramos και C. Tomlin, επιμελητές, σειρά Proceedings of Machine Learning Research, τόμ. 155, PMLR, 2021, σσ. 1147–1157. διεύθν.: <https://proceedings.mlr.press/v155/song21a.html>
- [98] N. Abbas κ.ά., “Survey of Advanced Nonlinear Control Strategies for UAVs: Integration of Sensors and Hybrid Techniques”, *Sensors*, τόμ. 24, αρθμ. 11, 2024, ISSN: 1424-8220. DOI: 10.3390/s24113286 διεύθν.: <https://www.mdpi.com/1424-8220/24/11/3286>
- [99] H. Samma και S. El-Ferik, “Autonomous UAV Visual Navigation Using an Improved Deep Reinforcement Learning”, *IEEE Access*, τόμ. 12, σσ. 79 967–79 977, 2024. DOI: 10.1109/ACCESS.2024.3409780
- [100] C. Chronis, G. Anagnostopoulos, E. Politi, G. Dimitrakopoulos και I. Varlamis, “Dynamic Navigation in Unconstrained Environments Using Reinforcement Learning Algorithms”, *IEEE Access*, τόμ. 11, σσ. 117 984–118 001, 2023. DOI: 10.1109/ACCESS.2023.3326435

- [101] Z. Fang και A. V. Savkin, “Strategies for Optimized UAV Surveillance in Various Tasks and Scenarios: A Review”, *Drones*, τόμ. 8, αρθμ. 5, 2024, ISSN: 2504-446X. DOI: 10.3390/drones8050193 διεύθν.: <https://www.mdpi.com/2504-446X/8/5/193>