**TẬP ĐOÀN CÔNG NGHIỆP - VIỄN THÔNG QUÂN ĐỘI**





**HỒ SƠ THUYẾT MINH**

**ĐỀ TÀI KHOA HỌC VÀ CÔNG NGHỆ CẤP CƠ SỞ**

**Tên đề tài:**

**Nghiên cứu các công nghệ lõi điều khiển và nền tảng chỉ huy thông minh các phương tiện không người lái trên bộ và trên không**

Mã số đề tài: 002-25-VAI-RDP-DS

**Đơn vị đăng ký chủ trì:** Trung tâm Dịch vụ dữ liệu và Trí tuệ nhân tạo Viettel

**Cá nhân đăng ký chủ nhiệm đề tài:** Nguyễn Hòa Bình.

Hà Nội, tháng 08/2025

| TẬP ĐOÀN CÔNG NGHIỆP - VIỄN THÔNG QUÂN ĐỘI  **TRUNG TÂM DỊCH VỤ DỮ LIỆU VÀ**  **TRÍ TUỆ NHÂN TẠO VIETTEL** | **CỘNG HOÀ XÃ HỘI CHỦ NGHĨA VIỆT NAM**  **Độc lập - Tự do – Hạnh phúc** |
| --- | --- |
|  | *Hà Nội, ngày tháng năm 2025* |

**THUYẾT MINH ĐỀ TÀI  
KHOA HỌC CÔNG NGHỆ CẤP CƠ SỞ**

**I. THÔNG TIN CHUNG VỀ ĐỀ TÀI**

**1. Tên đề tài:** Nghiên cứu các công nghệ lõi điều khiển và nền tảng chỉ huy thông minh các phương tiện không người lái trên bộ và trên không.

**2. Mã số**: 002-25-VAI-RDP-DS.

**3. Thời gian thực hiện:** 18 tháng.

(Từ tháng 09/2025 đến tháng 32/2027)

**4. Cấp quản lý:** Tập đoàn ☐ Cơ sở ☒

**5.** **Tổng kinh phí thực hiện:** 29.443.505.660 **đồng (bằng chữ: Hai mươi chín tỷ bốn trăm bốn mươi ba triệu năm trăm lẻ năm nghìn sáu trăm sáu mươi đồng, đã bao gồm VAT), trong đó:**

| **Nguồn** | **Kinh phí***(đồng)* |
| --- | --- |
| - Quỹ phát triển KHCN Viettel | 29.443.505.660 |
| - Từ nguồn tự có của cơ quan, đơn vị |  |
| - Từ nguồn kinh phí khác |  |

**6. Loại đề tài:**

☐ Chương trình (Ghi rõ tên chương trình, nếu có), Mã số:

☐ Dự án KH&CN

☒ Độc lập

☐ Khác

**7. Lĩnh vực:**

☒ Quân sự ☒ Dân sự

**8. Chủ nhiệm đề tài**

Họ và tên: Nguyễn Hòa Bình Số định danh cá nhân: 025086007202

Cấp bậc:

Ngày, tháng, năm sinh: 20/06/1986 Giới tính: Nam

Học hàm, học vị/Trình độ chuyên môn: Thạc sỹ

Chức danh nghề nghiệp: Quản lý.

Chức vụ: Phó Giám đốc Khối Công nghệ mới.

Điện thoại: 0869865453 E-mail: binhnh23@viettel.com.vn

Tên cơ quan, đơn vị đang công tác: Trung tâm Dịch vụ dữ liệu và Trí tuệ nhân tạo Viettel.

Địa chỉ cơ quan, đơn vị: Tòa nhà Trung tâm Đổi mới Sáng tạo Quốc gia, số 6 ngõ 7 Tôn Thất Thuyết, Cầu Giấy, Hà Nội.

**9. Thư ký khoa học của đề tài**

Họ và tên: Lê Hữu Thắng Số định danh cá nhân: 038095030936

Cấp bậc: Không

Ngày, tháng, năm sinh: 08/09/1995 Giới tính: Nam

Học hàm, học vị/Trình độ chuyên môn: Kỹ sư

Chức vụ: Nhân viên Khoa học công nghệ

Điện thoại: 0969904896 E-mail: thanglh11@viettel.com.vn

Tên cơ quan, đơn vị đang công tác: Trung tâm Dịch vụ dữ liệu và Trí tuệ nhân tạo Viettel.

Địa chỉ cơ quan, đơn vị: Tòa nhà Trung tâm Đổi mới Sáng tạo Quốc gia, số 6 ngõ 7 Tôn Thất Thuyết, Cầu Giấy, Hà Nội.

**10. Tổ chức chủ trì đề tài**

Tên đơn vị chủ trì đề tài: Trung tâm Dịch vụ dữ liệu và Trí tuệ nhân tạo Viettel.

Điện thoại: 02422826688

Website: viettelai.vn

Địa chỉ: Tòa nhà Trung tâm Đổi mới Sáng tạo Quốc gia, số 6 ngõ 7 Tôn Thất Thuyết, Cầu Giấy, Hà Nội.

Họ và tên thủ trưởng đơn vị: Nguyễn Mạnh Quý

Email: quynm@viettel.com.vn

Kho bạc Nhà nước/Ngân hàng: Ngân hàng TMCP Quân đội – PGD. Giảng Võ.

Tên cơ quan chủ quản đề tài: Tập đoàn Công nghiệp - Viễn thông Quân đội.

**11. Các tổ chức phối hợp chính thực hiện đề tài** *(nếu có)*

Không có.

**12. Thành viên thực hiện đề tài**

| **TT** | **Họ và tên,**  **học hàm, học vị** | **Chức danh thực hiện đề tài** | **Đơn vị công tác** |
| --- | --- | --- | --- |
| 1 | Nguyễn Hòa Bình, Thạc sỹ Khoa học. | Chủ nhiệm đề tài | Viettel AI |
| 2 | Lê Hữu Thắng | Thư ký khoa học | Viettel AI |
| 3 | Lê Xuân Tùng | Thành viên chính | Viettel AI |
| 4 | Lê Doãn Đạt | Thành viên chính | Viettel AI |
| 5 | Nguyễn Minh Huy | Thành viên chính | Viettel AI |
| 6 | Đặng Khắc Toàn | Thành viên chính | Viettel AI |
| 7 | Trần Trung Kiên | Thành viên chính | Viettel AI |
| 8 | Ninh Quốc Cường | Thành viên chính | Viettel AI |
| 9 | Kỹ sư HRL 15 | Thành viên chính | Viettel AI |
| 10 | Kỹ sử HRL 15 | Thành viên chính | Viettel AI |
| 11 | Kỹ sử HRL 15 | Thành viên chính | Viettel AI |

**II. MỤC TIÊU, NỘI DUNG VÀ PHƯƠNG ÁN TỔ CHỨC THỰC HIỆN ĐỀ TÀI**

**13. Mục tiêu của đề tài**

**13.1. Về sản phẩm**

- Mục tiêu chung: Phát triển thành công “Nền tảng chỉ huy thông minh các phương tiện không người lái trên bộ và trên không”, bao gồm ba phân hệ: Phân hệ chỉ huy điều khiển tập trung, phân hệ điều khiển drone và phân hệ điều khiển robot mặt đất.

- Mục tiêu cụ thể:

* Phân hệ chỉ huy điều khiển tập trung gồm các chức năng cơ bản: Hiển thị bản đồ 3D của khu vực hoạt động, hiển thị trạng thái, vị trí của các đối tượng trong khu vực, lập kế hoạch nhiệm vụ, chỉ huy điều khiển đồng thời nhiều đối tượng theo các kịch bản, hỗ trợ các giao thức phổ biến để điều khiển robot và drone.
* Phân hệ điều khiển drone có khả năng tự động bay trong môi trường không có tín hiệu GNSS, với các module sau: Module điều khiển bay, các module dẫn đường, định vị, nhận thức môi trường bằng thị giác máy tính.
* Phân hệ điều khiển robot tự hành mặt đất di chuyển trên địa hình đồi núi có đường mòn cho phương tiện di chuyển, gồm các module sau: Module định vị dẫn đường bằng GNSS-RTK, module phát hiện và tránh vật cản bằng thị giác máy tính.
* Thử nghiệm thành công nền tảng trên drone và robot tự hành mặt đất theo các kịch bản đã xác định trong phần tính năng sản phẩm.

**13.2. Về công nghệ**

- Đối với phân hệ chỉ huy điều khiển tập trung:

* Làm chủ (cấp độ sáng tạo) công nghệ điều khiển bay và tối ưu bầy đàn áp dụng cho bài toán tìm kiếm cứu nạn hộ sử dụng drone.
* Làm chủ (cấp độ đổi mới) công nghệ sinh dữ liệu để phục vụ huấn luyện mô hình, mô phỏng hoạt động của robot, drone.

- Đối với phân hệ điều khiển drone:

* Làm chủ (cấp độ đổi mới) phần mềm điều khiển bay.
* Làm chủ (cấp độ sáng tạo) công nghệ định vị, dẫn đường bay bằng thị giác máy tính trong môi trường không có tín hiệu GNSS.
* Làm chủ (cấp độ sáng tạo) công nghệ thị giác máy tính giúp drone phát hiện, nhận dạng, bắt bám đối tượng, mục tiêu.

- Đối với phân hệ điều khiển robot mặt đất:

* Làm chủ (cấp độ sáng tạo) công nghệ nhận biết môi trường bằng thị giác máy tính phục vụ điều hướng dẫn đường cho robot mặt đất hoạt động trên địa hình đồi núi có đường mòn cho phương tiện di chuyển.
* Làm chủ (cấp độ đổi mới) công nghệ định vị robot sử dụng tín hiệu GNSS-RTK cho robot hoạt động ngoài trời.

**14. Tình trạng của đề tài**

☒ Mới ☐ Kế tiếp hướng nghiên cứu của chính nhóm tác giả

☐ Kế tiếp nghiên cứu của người khác

**15. Tổng quan tình hình nghiên cứu, luận giải về mục tiêu và những nội dung nghiên cứu của đề tài**

**15.1. Đánh giá tổng quan tình hình nghiên cứu thuộc lĩnh vực của đề tài**

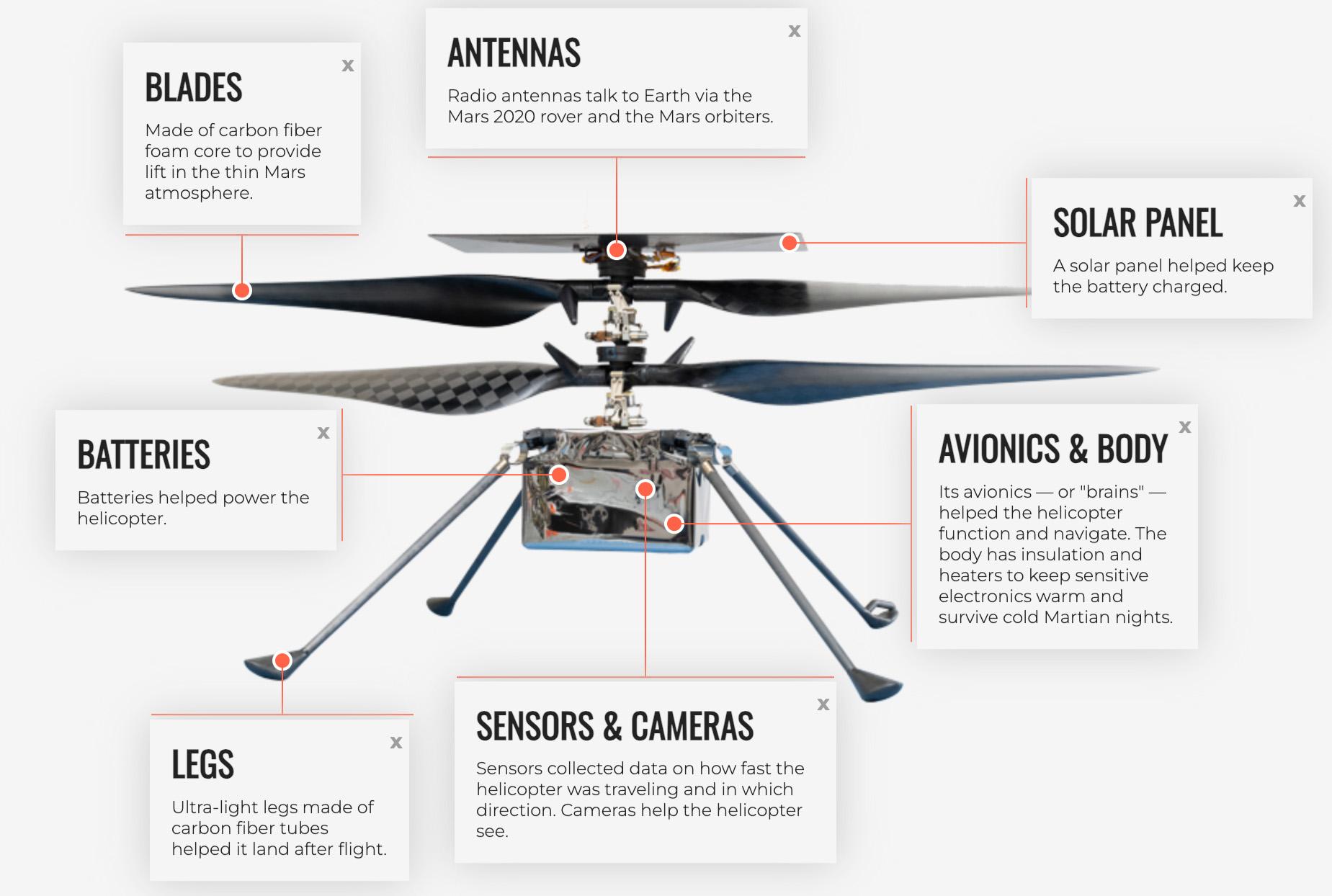
**A Ngoài nước**

Trong bối cảnh cuộc cách mạng công nghiệp 4.0 đang diễn ra mạnh mẽ, các phương tiện không người lái trên bộ và trên không ngày càng được ứng dụng rộng rãi trong nhiều lĩnh vực như quân sự, an ninh, giao thông thông minh, nông nghiệp chính xác và giám sát môi trường. Trên thế giới, nhiều tổ chức nghiên cứu, tập đoàn công nghệ và cơ quan quốc phòng đã đầu tư đáng kể vào việc phát triển các giải pháp điều khiển phương tiện không người lái. Dưới đây là một số khảo sát các nghiên cứu trên thế giới liên quan đến phạm vi của đề tài.

**15.1.1. Công nghệ định vị và xây dựng bản đồ dựa trên thị giác máy tính**

Trong những năm gần đây cùng với sự phát triển về hệ thống UAV, nhu cầu định vị và xây dựng bản đồ từ thiết bị bay không người lái cũng phát triển phục vụ cho nhiều lĩnh vực như nông nghiệp, quốc phòng cũng phát triển theo. Bài toán đặt ra yêu cầu hệ thống phải xác định vị trí của chính nó trong môi trường chưa biết, đồng thời xây dựng bản đồ của môi trường đó chỉ dựa trên các dữ liệu cảm biến đầu vào.

Trên thế giới hiện nay, nhiều sản phẩm thương mại đã ứng dụng thành công các công nghệ định vị và điều hướng tiên tiến để giúp các thiết bị bay không người lái hoạt động trong môi trường hạn chế hoặc không có tín hiệu GNSS. Nổi bật là Intel RealSense T265 [112], một camera tích hợp công nghệ Visual Inertial Odometry (VIO) gọn nhẹ, cho phép các nhà phát triển hoặc nhà sản xuất drone dễ dàng bổ sung khả năng định vị dựa vao camera và IMU với độ trễ thấp. Các hãng lớn như DJI cung đã thương mại hoa VIO trong nhiều sản phâm, từ Mavic [113] đến Matrice [114], nhằm giữ drone ổn định khi bay trong nhà hoặc ngoài khu đô thị nhiều vật cản. Một minh chứng ấn tượng cho khả năng VIO hoàn toàn tự chủ là dự án trực thăng thám hiểm sao hỏa Ingenuity của NASA [115]. Dự án này là nghiên cứu đầu tiên triển khai thành công thuật toán VIO trong điều kiện hành tinh khác. Do sao hỏa không có hệ thống định vị GPS, Ingenuity sử dụng VIO để ước lượng vị trí và hướng bay bằng cách kết hợp dữ liệu từ camera hướng xuống, IMU và laser altimeter. Phương pháp VIO của dự án dựa trên kỹ thuật MSCKF (Multi-State Constraint Kalman Filter), được tối ưu hóa cho phần cứng Snapdragon 801 và điều kiện ánh sáng, trọng lực thấp của sao Hỏa. Kết quả cho thấy VIO duy trì chính xác vị trí của trực thăng trong suốt các chuyến bay thử nghiệm, đóng vai trò then chốt giúp Ingenuity bay ổn định, hạ cánh chính xác và tự hành hoàn toàn trong hơn 70 lần bay thành công.



*Hình 15.1. Trực thăng sao hỏa Ingenuity của NASA*

Bên cạnh đó, công nghệ định vị Visual SLAM (VSLAM) được sử dụng để vừa định vị vừa xây dựng bản đồ 3D cũng đã được nhiều công ty tích hợp lên các sản phẩm của họ. Ví dụ như Skydio 2/X2 [116] trang bị sáu camera 4K cho khả năng nhận thức môi trường 360 độ va lập kế hoạch bay tự động nhờ thuật toan SLAM tích hợp. Autel EVO Max 4T [117] là một thiết bị bay không người lái công nghệ cao được trang bị hệ thống định vị tiên tiến dựa trên Visual SLAM, cho phép bay ổn định và tự hành trong môi trường không có GPS như trong nhà, rừng rậm hoặc đô thị phức tạp Với hệ thống cảm biến đa hướng, camera AI và mô-đun điều hướng thông minh, EVO Max 4T có khả năng tránh vật cản 360 độ, lập bản đồ 3D thời gian thực và hỗ trợ các nhiệm vụ trinh sát, cứu hộ và kiểm tra công nghiệp với độ chính xác cao. Đây là một trong những thiết bị bay không người lái thương mại đầu tiên tích hợp công nghệ thị giác máy tính vào hệ thống điều hướng lõi. Ngoài ra, các sản phẩm hoàn chỉnh, thị trường còn có SLAMcore SDK [118], cho phép các nhà phát triển tích hợp VSLAM vào drone hoặc robot mà không cần xây dựng thuật toán từ đầu. Với khả năng hoạt động trong môi trường không GPS, SDK giúp các thiết bị tự hành điều hướng ổn định trong nhà, nhà kho hoặc không gian phức tạp. SLAMcore nổi bật nhờ khả năng tích hợp dễ dàng trên phần cứng nhúng như NVIDIA Jetson và cung cấp dữ liệu bản đồ 2D/ 3D phục vụ điều hướng thông minh.

Một hướng công nghê mới nổi là Image Geo-Localization, sử dụng phương pháp so sánh ảnh mặt đất với ảnh vệ tinh để xác định tọa độ địa lý mà không cần GPS. Startup Mỹ Theseus [119], với hệ thống Micro Visual Positioning System (Micro VPS) có thể đặt trên bất kỳ drone nào thành “GPS giả lập”. Thay vì sử dụng GPS, Micro VPS xác định vị trí của robot hoặc thiết bị bay bằng cách so khớp ảnh thực tế với cơ sở dữ liệu bản đồ hình ảnh đã gán tọa độ địa lý, từ đó cho phép hoạt động ổn định trong khu đô thị phức tạp hoặc môi trường nhiễu GPS. Công nghệ này hướng tới các ứng dụng trong robot tự hành, drone và logistics thông minh, với độ chính xác có thể đạt đến centimet trong môi trường thực tế. Tương tự, OMNInav của OKSI [120] cũng cung cấp các giải pháp định vị kết hợp thị giác và IMU với khả năng so khớp bản đồ (map matching). Các giải pháp này có thể ứng dụng cho cả dân sự và quân sự, đặc biệt hữu ích trong các nhiệm vụ bay ngoài tầm nhìn (BVLOS) hoặc trong khu vực bị gây nhiễu GPS. Dưới đây là các nghiên cứu nổi bật về các công nghệ Visual SLAM, Visual Inertial Odometry và Image Geo-Localization.

Các công nghệ định vị cho drone hiện nay rất đa dạng và được lựa chọn tùy theo điều kiện môi trường và yêu cầu nhiệm vụ cụ thể. Một trong những phương pháp phổ biến là định vị bằng hệ thống vệ tinh định vị toàn cầu GNSS, sử dụng tín hiệu từ các hệ thống định vị toàn cầu như GPS (Hoa Kỳ), GLONASS (Nga), Galileo (EU) hoặc BeiDou (Trung Quốc) để xác định vị trí tuyệt đối của drone. Công nghệ này đặc biệt hiệu quả trong các môi trường ngoài trời thông thoáng, nhờ độ chính xác tương đối cao và chi phí hợp lý. Bên cạnh đó, IMU (Inertial Measurement Unit) là hệ thống gồm gia tốc kế và con quay hồi chuyển, cho phép đo gia tốc tuyến tính và vận tốc góc của drone, từ đó ước lượng chuyển động tương đối trong thời gian ngắn mà không phụ thuộc vào tín hiệu bên ngoài, tuy nhiên dễ tích lũy sai số. Trong khi đó, định vị bằng camera – hay còn gọi là Visual Odometry hoặc Visual SLAM – sử dụng hình ảnh từ camera để quan sát môi trường, tính toán chuyển động tương đối giữa các khung hình nhằm xác định vị trí drone và xây dựng bản đồ, phù hợp với môi trường trong nhà hoặc nơi tín hiệu GPS yếu. Ngoài ra, LiDAR SLAM sử dụng cảm biến LiDAR để thu thập dữ liệu không gian dưới dạng đám mây điểm (point cloud), từ đó dựng bản đồ 3D và định vị chính xác, đặc biệt thích hợp với các nhiệm vụ đòi hỏi độ chính xác cao hoặc trong môi trường phức tạp. Cuối cùng, kết hợp cảm biến (Sensor Fusion) là phương pháp tích hợp dữ liệu từ nhiều nguồn như GPS, IMU, camera và LiDAR nhằm tận dụng ưu điểm của từng loại cảm biến, với các thuật toán hợp nhất phổ biến như Kalman Filter hoặc Graph-based Sensor Fusion giúp xử lý và đồng bộ dữ liệu hiệu quả, mang lại khả năng định vị ổn định và chính xác hơn trong thực tế.

| **Công nghệ** | **Ưu điểm** | **Hạn chế** |
| --- | --- | --- |
| GNSS | Phổ biến, giá rẻ, hoạt động tốt ngoài trời | Không hoạt động trong nhà, dễ nhiễu, độ trễ cao |
| IMU | Phản hồi nhanh, không phụ thuộc tín hiệu bên ngoài | Tích lũy sai số theo thời gian, không tự định vị được |
| Visual SLAM | Chi phí thấp, hoạt động trong nhà, thu thập dữ liệu phong phú | Nhạy cảm ánh sáng, yêu cầu xử lý ảnh tốc độ cao |
| LiDAR SLAM | Độ chính xác cao, độc lập với điều kiện ánh sáng | Giá thành cao, xử lý dữ liệu lớn |
| Sensor Fusion | Ổn định, chính xác, thích nghi nhiều môi trường | Hệ thống phức tạp, yêu cầu đồng bộ và hiệu chỉnh cảm biến |

*Bảng 15.1. So sánh các công nghệ định vị phổ biến cho Drone*

Visual-Inertial Odometry (VIO) là một phương pháp ước lượng quỹ đạo chuyển động tương đối của thiết bị theo thời gian thực, thông qua việc tích hợp thông tin thị giác từ camera và dữ liệu quán tính từ cảm biến IMU (bao gồm gia tốc kế và con quay hồi chuyển). Không giống như VisualSLAM (hay còn gọi là VSLAM) – vốn tập trung vào việc xây dựng bản đồ và tái định vị toàn cục – VIO hướng đến cung cấp định vị liên tục, với độ trễ thấp và độ chính xác cao trong khoảng thời gian ngắn đến trung bình, mà không cần xây dựng bản đồ toàn cục.

Trong VIO, hệ thống tập trung so khớp giữa các chuỗi khung hình liên tiếp và kết hợp với dữ liệu IMU để ước lượng tư thế liên tục của thiệt bị. Có hai hướng chính là sử dụng bộ lọc như Kalman Filter (filter-based) và dựa trên tối ưu hóa phi tuyến.

**VIO dựa trên bộ lọc (filter-based):** Các phương pháp VIO dựa trên bộ lọc sử dụng các thuật toán như EKF (Extended Kalman Filter) hoặc MSCKF (Multi-State Constraint Kalman Filter) để cập nhật trạng thái theo thời gian thực. MSCKF [1] là một đại diện tiêu biểu cho nhóm này, trong đó thuật toán loại bỏ các đặc trưng hình ảnh khỏi không gian trạng thái của bộ lọc và thay vào đó sử dụng ràng buộc không gian từ nhiều khung hình quan sát để ước lượng chuyển động. Cách tiếp cận này giúp giảm đáng kể độ phức tạp tính toán so với các phương pháp EKF truyền thống phải theo dõi từng đặc trưng riêng biệt. Ngoài MSCKF, một số thuật toán nổi bật khác thuộc nhóm VIO dựa trên bộ lọc bao gồm ROVIO [2], trifocal-EKF[3] và SchurVINS [4] . Nhìn chung, các phương pháp dựa trên bộ lọc có hiệu suất tính toán cao, phù hợp với các hệ thống nhúng hoặc thiết bị có tài nguyên phần cứng hạn chế. Tuy nhiên, chúng thường nhạy với sai số tuyến tính hóa và có thể không duy trì được độ chính xác cao trong thời gian dài do không thực hiện tối ưu hóa toàn cục.

**VIO dựa trên tối ưu hóa (optimization-based):** Phương pháp này dựa trên tối ưu hóa phi tuyến (Nonlinear Optimization), bao gồm các kỹ thuật như tối ưu hóa đồ thị (graph optimization) và bundle adjustment [5]. Để đảm bảo tốc độ xử lý ổn định, các hệ thống VIO thường thực hiện tối ưu hóa trên một cửa sổ trượt (sliding window) chứa các trạng thái gần nhất. Để xử lý các ràng buộc từ IMU một cách hiệu quả, các thuật toán ứng dụng lý thuyết tiền tích phân (pre-integration) [6], giúp tích hợp dữ liệu IMU thu thập với tần số cao vào quá trình tối ưu hóa mà không tăng quá nhiều độ phức tạp tính toán. Các thuật toán VSLAM hiện đại như ORB-SLAM, SVO và Kimera đều sử dụng frontend là VIO để thực hiện bước ước lượng vị trí trước khi tối ưu và xây dựng bản đồ. Phương pháp này mang lại độ chính xác cao hơn nhờ khả năng mô hình hóa toàn diện các ràng buộc từ cả dữ liệu hình ảnh và cảm biến quán tính, tuy nhiên nó yêu cầu tài nguyên tính toán lớn hơn so với các phương pháp dựa trên bộ lọc.

Visual SLAM là một nhánh cụ thể của SLAM, trong đó camera đóng vai trò cảm biến chính để thu thập thông tin về môi trường. Camera có thể là monocular (một camera), stereo (một cặp camera) hoặc RGB-D (camera có thêm thông tin chiều sâu).

Ảnh có chứa văn bản, ảnh chụp màn hình, biểu đồ, Phông chữ

Nội dung do AI tạo ra có thể không chính xác.

*Hình 15.2. Sơ đồ thuật toán Visual SLAM*

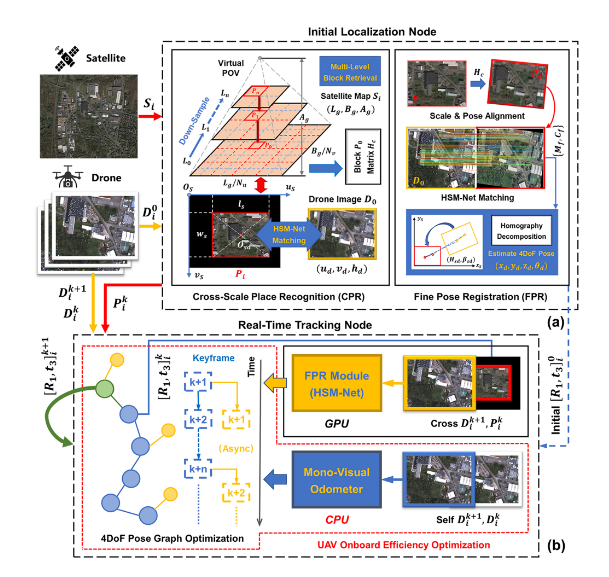
Một hệ thống Visual SLAM điển hình bao gồm nhiều thành phần hoạt động phối hợp theo trình tự, từ thu thập dữ liệu cho đến tái tạo bản đồ cuối cùng. Pipeline cơ bản của hệ thống Visual SLAM bao gồm:

* Data Preprocesing: Giai đoạn đầu tiên của Visual SLAM là thu thập dữ liệu từ cảm biến trong bài toán này là hình ảnh nhận được từ camera đồng thời xử lý tiền xử lý hình ảnh như hiệu chỉnh méo ảnh hoặc cân bằng ảnh sáng và đông bộ dữ liệu với các cảm biến phụ khác như IMU, GPS, encoder… nhằm hỗ trợ định vị chính xác và ổn định hơn.
* Visual Odometry (VO): Là thành phần đóng vai trò ước lượng chuyển động tương đối camera giữa các khung hình liên tiếp. VO giúp xây dựng một bản đồ cục bộ ban đầu và cung cấp chuỗi các pose. Do hoạt động trực tiếp trên dữ liệu hình ảnh đầu vào và đóng vai trò chính trong theo dõi liên tục, VO thường được gọi là frontend của hệ thống Visual SLAM.
* Optimization: Sau khi VO ước lượng được các chuyển động, backend sẽ tiếp nhận các pose này cùng với thông tin từ các thành phần khác như loop closing để thực hiện tối ưu hóa toàn cục. Quá trình này sử dụng các thuật toán tối ưu (như Graph Optimization, Bundle Adjustment) để hiệu chỉnh sai số tích lũy theo thời gian, nhằm tái cấu trúc quỹ đạo chuyển động chính xác hơn và xây dựng bản đồ nhất quán. Thành phần này được gọi là backend vì hoạt động phía sau VO trong chuỗi pipeline.
* Loop Closing: Loop closing có nhiệm vụ phát hiện việc robot đã quay lại một vị trí từng đi qua trước đó hay chưa. Khi một vòng lặp được phát hiện, hệ thống sẽ sử dụng thông tin tương ứng để cập nhật lại toàn bộ đồ thị pose, từ đó giảm thiểu lỗi tích lũy (drift) trong hệ thống. Loop closing đóng vai trò quan trọng trong việc đảm bảo tính toàn vẹn của bản đồ và độ chính xác của định vị lâu dài.
* Reconstruction: Dựa trên quỹ đạo camera đã được tối ưu hóa, hệ thống tiến hành xây dựng bản đồ phục vụ nhiệm vụ cụ thể. Tùy vào ứng dụng, bản đồ có thể là bản đồ điểm (point cloud), mô hình lưới 3D (mesh), hoặc bản đồ dạng mật độ cao (dense map). Mục tiêu của bước này là cung cấp không gian môi trường số hóa để phục vụ các tác vụ như dẫn đường, nhận dạng vật thể, hoặc tương tác thực tế tăng cường (AR).

Phân loại công nghệ Visual SLAM:

* **Phương pháp Visual SLAM truyền thống**: Các phương pháp Visual SLAM truyền thống đóng vai trò nền tảng trong việc xây dựng các hệ thống định vị và bản đồ sử dụng thông tin thị giác. Dựa trên cách thức xử lý dữ liệu hình ảnh, các phương pháp này có thể được phân loại thành ba nhóm chính: phương pháp trực tiếp (direct), phương pháp bán trực tiếp (semi-direct) và phương pháp dựa trên đặc trưng (feature-based). Nhóm đầu tiên là **phương pháp trực tiếp**, trong đó hệ thống tận dụng trực tiếp cường độ của các pixel trong ảnh để ước lượng chuyển động của camera và tối ưu hóa sai số photometric giữa các khung hình liên tiếp. Các phương pháp này không cần phát hiện hay mô tả đặc trưng điểm, giúp khai thác thông tin ảnh một cách đầy đủ hơn. Tiêu biểu cho nhóm này là LSD-SLAM [7], một hệ thống sử dụng ảnh từ camera đơn để xây dựng bản đồ bán mật độ (semi-dense) trong thời gian thực, và DSO (Direct Sparse Odometry)[8], một phiên bản tinh gọn hơn khi chỉ lựa chọn các điểm ảnh có độ tin cậy cao để tối ưu hóa, từ đó cải thiện hiệu suất và độ chính xác. Mặc dù đạt độ chính xác cao, các phương pháp trực tiếp thường nhạy cảm với ánh sáng và không có cơ chế đóng vòng mặc định. Nhóm thứ hai là **phương pháp bán trực tiếp (semi-direct)**, được phát triển nhằm kết hợp ưu điểm của hai phương pháp trên. Trong các hệ thống này, đặc trưng hình ảnh được phát hiện để làm điểm khởi tạo, sau đó quá trình theo dõi và tối ưu hóa chuyển động sử dụng thông tin photometric từ cường độ pixel. Cách tiếp cận này giúp đạt được sự cân bằng giữa tốc độ và độ chính xác, đồng thời tăng tính ổn định trong điều kiện thực tế. Đại diện tiêu biểu cho nhóm này là SVO (Semi-direct Visual Odometry) [9], một hệ thống có khả năng theo dõi chuyển động nhanh, thường được sử dụng trong các ứng dụng robot di động và drone có yêu cầu xử lý thời gian thực cao. Nhóm cuối cùng, cũng là nhóm phổ biến nhất trong các hệ thống Visual SLAM truyền thống, là **phương pháp dựa trên đặc trưng (feature-based hoặc keypoint-based)**. Các phương pháp này sử dụng các kỹ thuật xử lý ảnh cổ điển để phát hiện và mô tả các điểm đặc trưng như FAST, Harris, Shi-Tomasi và BRIEF, ORB, SIFT, SURF sau đó thực hiện so khớp đặc trưng giữa các khung hình để ước lượng chuyển động và tái dựng bản đồ. Nổi bật trong nhóm này là dòng hệ thống ORB-SLAM[10], bao gồm cả ORB-SLAM 2 [11] và ORB-SLAM 3 [12]. Các hệ thống này sử dụng đặc trưng ORB (Oriented FAST and Rotated BRIEF) để phát hiện và mô tả keypoint, kết hợp với cơ chế đóng vòng (loop closing) sử dụng Bag-of-Words nhằm giảm sai số tích lũy trong quỹ đạo di chuyển. ORB-SLAM2 mở rộng hỗ trợ cho camera stereo và RGB-D, trong khi ORB-SLAM3 tích hợp thêm dữ liệu IMU để hình thành hệ Visual-Inertial SLAM, giúp cải thiện đáng kể độ ổn định trong môi trường có chuyển động nhanh hoặc thay đổi bất thường. Ngoài ra VINS-Fusion [13], [14] là một hệ thống Visual-Inertial SLAM mã nguồn mở, dựa trên đặc trưng ảnh và kết hợp chặt chẽ với IMU. Hệ thống sử dụng tối ưu hóa phi tuyến theo cửa sổ trượt và hỗ trợ tích hợp thêm GPS hoặc các cảm biến đo tuyệt đối khác, phù hợp cho định vị trong cả môi trường trong nhà và ngoài trời. và Kimera [15] là hệ thống Visual-Inertial SLAM hiện đại với cấu trúc mô-đun, không chỉ cung cấp định vị chính xác mà còn hỗ trợ tái dựng bản đồ 3D dạng lưới và phân đoạn ngữ nghĩa. Kimera phù hợp cho các ứng dụng đòi hỏi nhận thức không gian như robot tự hành hoặc AR/VR.
* **Phương pháp kết hợp học sâu (hybrid method)**: Trong nhóm các phương pháp Visual SLAM kết hợp học sâu, cấu trúc tổng thể của hệ thống vẫn giữ theo pipeline truyền thống, bao gồm các thành phần chính như frontend, backend, loop closure và tái dựng bản đồ. Tuy nhiên, một hoặc nhiều thành phần, đặc biệt là các bước như phát hiện đặc trưng, mô tả đặc trưng, ghép đặc trưng, hoặc phân đoạn ảnh,được thay thế hoặc hỗ trợ bởi các mô hình học sâu, nhằm nâng cao độ chính xác, khả năng ổn định và tính khái quát của hệ thống.Một số hệ thống tiêu biểu cho hướng tiếp cận này đã được phát triển trong những năm gần đây. DynaSLAM [16] là một trong những hệ thống tiên phong, tích hợp mạng segmentation học sâu vào pipeline của ORB-SLAM để loại bỏ các vật thể động như người hoặc xe khỏi ảnh đầu vào, từ đó cải thiện độ chính xác trong môi trường thực. DF-SLAM [17] sử dụng mô tả đặc trưng học sâu thay cho các descriptor thủ công, giúp tăng tính ổn định khi xử lý ảnh trong điều kiện ánh sáng yếu hoặc phức tạp. Một hướng đi khác là thay thế hoàn toàn đặc trưng truyền thống bằng đặc trưng học sâu. Hệ thống SuperPoint kết hợp LighGlue [18] thực hiện cả phát hiện keypoint và ghép đặc trưng thông qua mạng nơ-ron học sâu, giúp nâng cao chất lượng matching và độ tin cậy của tracking. Tương tự, RWT-SLAM [19] sử dụng kiến trúc Transformer để ghép điểm ảnh trực tiếp mà không cần phát hiện keypoint, giảm độ trễ và tăng hiệu suất trong các cảnh biến đổi mạnh. Ngoài việc tăng cường phần đặc trưng, nhiều hệ thống còn khai thác học sâu để cải thiện khả năng hiểu ngữ nghĩa của môi trường. SP-SLAM [20] kết hợp phân đoạn ngữ nghĩa để hỗ trợ định vị trong môi trường phức tạp nhiều lớp đối tượng, trong khi OmniSLAM [21] mở rộng khả năng xử lý cho ảnh toàn cảnh 360 độ bằng cách tích hợp các thành phần học sâu chuyên xử lý ảnh dạng cầu (equirectangular images), phục vụ cho robot di chuyển tự do hoặc các hệ thống xe tự hành. Bên cạnh đó, Edge-SLAM [22] là một biến thể đáng chú ý, trong đó hệ thống không sử dụng keypoint truyền thống mà thay bằng đặc trưng biên (edge) được học sâu hỗ trợ, giúp cải thiện khả năng tracking trong môi trường ít texture. Mặc dù áp dụng học sâu ở phần frontend, các hệ thống này vẫn sử dụng các thuật toán tối ưu hóa truyền thống như pose graph optimization, bundle adjustment, hoặc sliding window optimization ở phần backend, nhằm đảm bảo độ chính xác hình học và khả năng hiệu chỉnh trôi (drift) trong thời gian dài. Chính sự kết hợp linh hoạt giữa đặc trưng học được và các kỹ thuật tối ưu hình học đã giúp các hệ thống hybrid này đạt được hiệu suất vượt trội trong nhiều môi trường thực tế.
* **Phương pháp End to end Visual SLAM**: Trái với nhóm hybrid, các phương pháp SLAM học sâu end to end hướng tới việc xây dựng toàn bộ hệ thống dưới dạng mạng nơ-ron thống nhất, trong đó các bước như trích xuất đặc trưng, theo dõi chuyển động, tái dựng bản đồ và tối ưu hóa quỹ đạo đều được tích hợp vào một pipeline duy nhất và được học trực tiếp từ dữ liệu ảnh đầu vào. Các hệ thống này không chia thành các module riêng biệt như frontend–backend truyền thống, mà thay vào đó, mạng học sâu được huấn luyện để thực hiện toàn bộ pipeline từ đầu đến cuối (end-to-end). Một đại diện nổi bật cho hướng tiếp cận này là DROID-SLAM [23]. Hệ thống này sử dụng kiến trúc gồm một mạng recurrent attention-based để thực hiện đồng thời matching, tracking, và pose optimization. Cụ thể, DROID-SLAM dynamic graph optimizer để liên tục cập nhật và tối ưu hóa các cặp khung hình có thể quan sát chéo lẫn nhau. Thay vì sử dụng đặc trưng cố định, mạng học được biểu diễn động qua các tầng attention để liên tục cập nhật đặc trưng không gian và quan hệ giữa các khung hình. Toàn bộ hệ thống được huấn luyện với loss dựa trên pose, reprojection error và matching consistency, cho phép nó hoạt động tốt mà không cần bước phát hiện keypoint truyền thống. Tiếp theo, Nice-SLAM [24] là một hệ thống end-to-end khác áp dụng mô hình học sâu để tái dựng không gian 3D dưới dạng NeRF-style (neural field). Cốt lõi của hệ thống là kết hợp giữa mạng NeRF để biểu diễn bản đồ 3D và một mạng camera tracking để ước lượng pose. Trong đó, bản đồ không được biểu diễn bằng điểm rời rạc hay lưới 3D, mà thay vào đó là một mạng MLP học biểu diễn liên tục của cảnh (neural implicit representation). Điều này giúp Nice-SLAM tái dựng bản đồ với độ mượt và liên kết không gian cao, đồng thời hỗ trợ render lại từ góc nhìn bất kỳ. TartanVO [25] là một hệ thống Visual Odometry học sâu, được huấn luyện trên tập dữ liệu tổng hợp TartanAir với mục tiêu tăng khả năng tổng quát hóa sang dữ liệu thực tế. Pipeline của TartanVO gồm các khối CNN trích xuất đặc trưng ảnh, sau đó một mạng regressor học trực tiếp pose tương đối giữa các khung hình, kết hợp với học supervision từ ground-truth trajectory hoặc photometric loss. Mặc dù không phải là hệ thống SLAM đầy đủ (chưa đóng vòng), TartanVO cho thấy hiệu quả vượt trội trong các môi trường indoor/outdoor với biến đổi phức tạp. Mặc dù các phương pháp học sâu đầu-cuối đạt được độ chính xác cao và có khả năng tự học được biểu diễn từ dữ liệu mà không cần đặc trưng thủ công, nhưng chúng vẫn tồn tại những hạn chế nhất định. Trong đó, độ trễ tính toán, yêu cầu tài nguyên phần cứng cao, và thiếu khả năng diễn giải (interpretability) là những thách thức lớn khiến việc áp dụng vào hệ thống thực tế như drone hoặc robot thời gian thực còn hạn chế. Ngoài ra, quá trình huấn luyện đòi hỏi tập dữ liệu lớn, đa dạng và có nhãn chất lượng cao (pose, depth), điều này đôi khi không khả thi trong các môi trường ứng dụng đặc thù.

Một hướng tiếp cận khác trong định vị bằng camera sử dụng hình ảnh là so khớp ảnh và ảnh địa hình tại khu vực Image-based geo-localization. Bài toán này xác định vị trí của drone dựa vào ảnh hiện tại (thường theo hướng từ trên cao nhìn xuống, hay còn gọi là Top-down hoặc Bird-eye-view) thu được và so sánh với một cơ sở dữ liệu ảnh đã gắn hệ tọa độ. Phương pháp này hỗ trợ định vị chính xác hơn giúp xác định, định vị trong hệ tọa độ toàn cầu khi tín hiệu GNSS yếu hoặc không có. Có 2 kỹ thuật chính để tiếp cận bài toán này:



*Hình 15.3. Cấu trúc tổng quan của mô hình GeoRVLF*

**Phương pháp dựa vào phân tích đặc trưng** (Detector based Feature Matching): Các phương pháp này tách riêng quá trình phát hiện đặc trưng và mô tả đặc trưng. Ban đầu, các đặc trưng thủ công như SIFT và ORB hoạt động tốt trong các bài toán VO/SLAM đồng nhất, nhưng kém ổn định trong điều kiện chênh lệch lớn về góc nhìn và tỷ lệ như drone–satellite matching. Để khắc phục, các mạng học đặc trưng như NetVLAD [26] được phát triển nhằm tăng khả năng khái quát trong cảnh phức tạp. Bên cạnh đó, các phương pháp học đặc trưng dày đặc như D2-Net [27] và SuperPoint [28] cho phép trích xuất đồng thời điểm đặc trưng và mô tả đặc trưng, giúp tăng tính ổn định dưới điều kiện ánh sáng thay đổi hoặc cảnh ít cấu trúc. Hiệu quả của các phương pháp này đã được xác nhận qua các thử nghiệm UAV [29], [30]. Các phương pháp hiện đại hơn không cần bước phát hiện riêng biệt mà học trích xuất và so khớp đặc trưng một cách đầu-cuối. Các phương pháp như NCNet và DualRC-Net [31] sử dụng CNN để tìm đặc trưng, kết hợp thuật toán MNN (mutual nearest neighbor) để xác định điểm tương ứng, tuy nhiên còn bị giới hạn bởi trường nhìn hẹp của tích chập. Gần đây, các mô hình như LoFTR [32] và ASTR [33] đã áp dụng cơ chế multi-head self-attention, giúp cải thiện việc mã hóa đặc trưng cục bộ, tương thích tốt với biến đổi tỷ lệ và cấu trúc cảnh — hoạt động tương tự như phương pháp template matching, nhưng ổn định hơn trong các cảnh địa lý đa dạng. Hoặc như GeoRVLF [34] sử dụng Mạng HSM-Net để so khớp ảnh cross-view và truy xuất vị trí toàn cục chính xác mà không cần DEM hay render ảnh trước.

**Phương pháp dựa vào mô hình transformer**:Các phương pháp sử dụng Transformer đang ngày càng thu hút sự chú ý trong bài toán định vị này nhờ khả năng mô hình hóa ngữ cảnh hiệu quả và tính khái quát cao hơn so với các kiến trúc CNN truyền thống. Điều này đặc biệt quan trọng trong các bối cảnh dị hình (heterogeneous scenes), nơi tồn tại sự khác biệt lớn về góc nhìn, tỷ lệ và điều kiện chiếu sáng. Các mô hình như TransGeo và GeoDTR [35] khai thác cơ chế self-attention của Vision Transformer (ViT) để xử lý chuỗi đặc trưng ở cấp độ patch, từ đó tăng cường khả năng nắm bắt quan hệ không gian giữa các vùng ảnh và cải thiện độ chính xác trong so khớp cross-view. Để nâng cao hiệu quả tính toán và đảm bảo tính ổn định trong các môi trường thực tế, SuperGlue [36] đã đề xuất cơ chế cross-attention để trực tiếp thiết lập mối liên kết giữa các điểm đặc trưng từ hai ảnh có góc nhìn khác nhau. Mô hình LoFTR sau đó kế thừa và mở rộng ý tưởng này bằng kiến trúc coarse-to-fine, giúp cải thiện đáng kể độ chính xác khi xử lý các cặp ảnh có sự chênh lệch lớn về tỷ lệ (cross-scale). Hiệu quả ứng dụng thực tiễn của các phương pháp dựa trên Transformer trong định vị UAV đã được xác thực qua các nghiên cứu thực nghiệm của Chen et al. [37] và Li et al. [38]. Ngoài ra, các mô hình nhẹ hơn như LightGlue [39] và ELoFTR [40] đã được phát triển nhằm phù hợp với việc triển khai trực tiếp trên các nền tảng UAV cỡ nhỏ, nhờ vào cấu trúc gọn nhẹ nhưng vẫn duy trì hiệu suất cao.

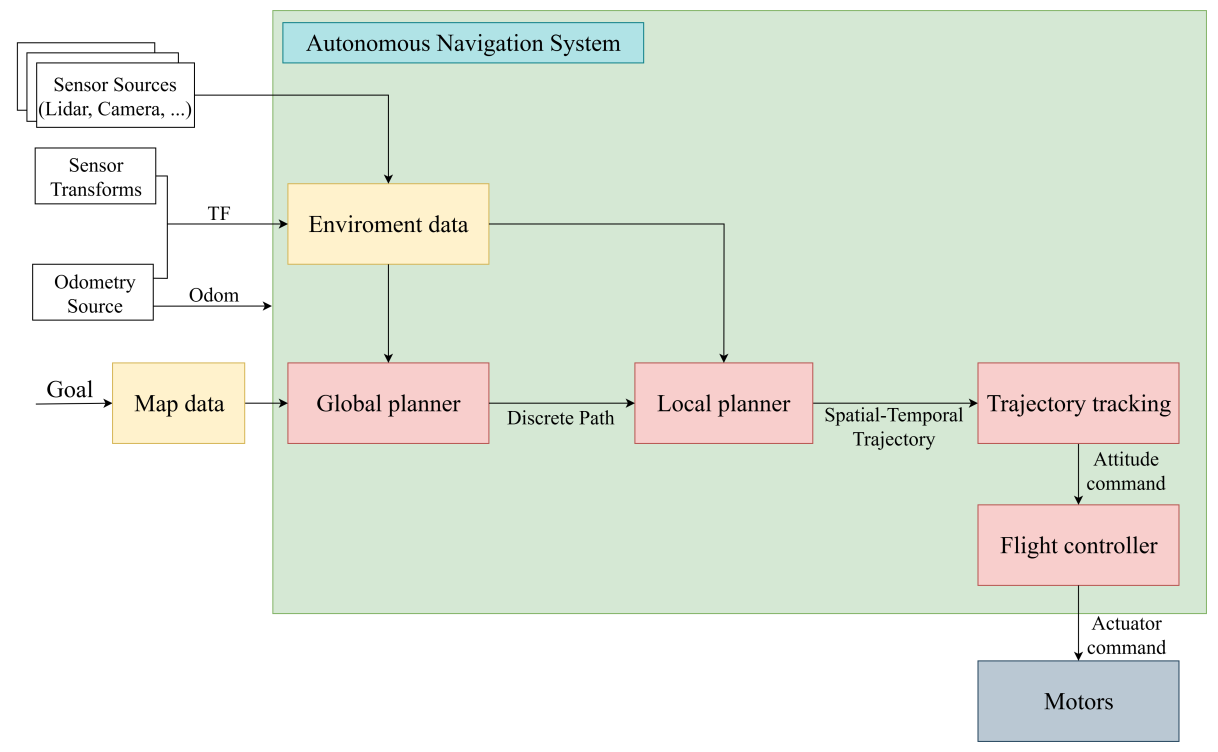
Dưới góc độ tổng quan, ba nhóm phương pháp SLAM truyền thống, hybrid và end-to-end thể hiện sự phát triển theo hướng tăng cường khả năng học và thích ứng của hệ thống. Các phương pháp truyền thống có ưu điểm về tính ổn định, dễ triển khai và phù hợp với thiết bị hạn chế tài nguyên. Nhóm hybrid tận dụng cấu trúc pipeline cũ nhưng bổ sung học sâu để tăng độ chính xác và khả năng xử lý môi trường phức tạp. Trong khi đó, các phương pháp end-to-end thể hiện tiềm năng lớn nhờ khả năng học trực tiếp toàn bộ pipeline, tuy nhiên vẫn gặp thách thức về tài nguyên và độ tin cậy trong ứng dụng thực tế.

Bảng 15.2. Bảng so sánh ưu nhược điểm của các giải pháp Visual SLAM

| **Phương pháp** | | **Ưu điểm** | **Nhược điểm** |
| --- | --- | --- | --- |
| Phương pháp truyền thống | Trực tiếp | Không cần phát hiện đặc trưng, tận dụng tối đa thông tin cường độ ảnh, hoạt động tốt trong môi trường ít đặc trưng. | Dễ bị ảnh hưởng bởi nhiễu, không ổn định khi thay đổi ánh sáng hoặc hình ảnh bị mờ. |
| Bán trực tiếp | Kết hợp lợi thế của phương pháp direct và feature-based, theo dõi nhanh, phù hợp với ứng dụng thời gian thực. | Phức tạp hơn trong xây dựng pipeline, nhạy với tham số và điều kiện ảnh hưởng đến phát hiện đặc trưng. |
| Feature-based | Độ ổn định cao, dễ triển khai, không cần dữ liệu huấn luyện, hoạt động tốt trên thiết bị hạn chế. | Phụ thuộc vào đặc trưng thủ công, kém hiệu quả trong điều kiện ánh sáng yếu, chuyển động nhanh hoặc cảnh trống. |
| Phương pháp Hybrid | | Tận dụng khả năng học đặc trưng mạnh mẽ từ deep learning, tăng độ chính xác trong môi trường phức tạp hoặc động học. | Cần mô hình học sâu được huấn luyện tốt, tăng chi phí tính toán, đôi khi khó chạy real-time trên thiết bị nhúng. |
| Phương pháp end-to-end | | Tự động học toàn bộ pipeline, độ chính xác cao, có tiềm năng khái quát tốt và tái dựng bản đồ chi tiết. | Yêu cầu tài nguyên cao, thiếu tính diễn giải, khó tích hợp với backend truyền thống và chưa tối ưu hóa real-time. |

**15.1.2. Công nghệ điều hướng UAV**

Trong lĩnh vực drone tự hành, điều hướng tự động là một thành phần cốt lõi trong hệ thống drone hiện đại, đóng vai trò quyết định đến khả năng hoạt động ổn định, an toàn và thông minh của phương tiện bay. Một hệ thống điều hướng hiệu quả không chỉ đảm bảo drone có thể bay đến điểm đích mà còn phải thích ứng tốt với môi trường không xác định, tránh chướng ngại vật ở tốc độ cao, đồng thời đáp ứng các tác vụ nâng cao như bám theo mục tiêu di động. Điều hướng thường gồm ba thành phần chính: lập kế hoạch toàn cục (global planning), lập kế hoạch cục bộ (local planning),và phần mềm điều khiển bay (flight control).



*Hình 15.4. Sơ đồ hệ thống điều hướng cho drone*

**Lập kế hoạch toàn cục (global planner):** là tầng cao nhất của hệ thống điều hướng, với mục tiêu xác định đường đi (path) khả thi từ vị trí xuất phát đến vị trí đích trên bản đồ đã biết trước (ví dụ bản đồ dạng grid hoặc bản đồ occupancy). Các thuật toán trong lớp này thường là các thuật toán tìm kiếm đồ thị. Đối với các dạng biểu diễn sử dụng occupancy-grid hoặc costmap, các thuật toán tìm đường ngắn nhất như A\* [41], Dijkstra [42] thường được sử dụng với chi phí tính toán thấp do sử dụng hàm heuristic để dẫn hướng tìm kiếm. Trong môi trường động, D\* lite [43] phù hợp với sinh đường tối ưu nhanh do cho phép cập nhật lại đường đi khi phát hiện chướng ngại vật. Còn đối với môi trường 3D rộng lớn, các thuật toán lấy mẫu ngẫu nhiên như RRT [44] (Rapidly-Exploring Random Tree) hay tìm kiếm đồ thị PRM [45] (Probabilistic RoadMap) đạt hiệu quả tốt. Các thuật toán này cho ra một chuỗi các điểm hoặc đoạn đường rời rạc, không liên tục theo thời gian và không tuân thủ động học của UAV. Do đó, kết quả từ global planning chỉ có thể được sử dụng làm định hướng ban đầu hoặc mục tiêu tạm thời cho tầng lập kế hoạch cục bộ.

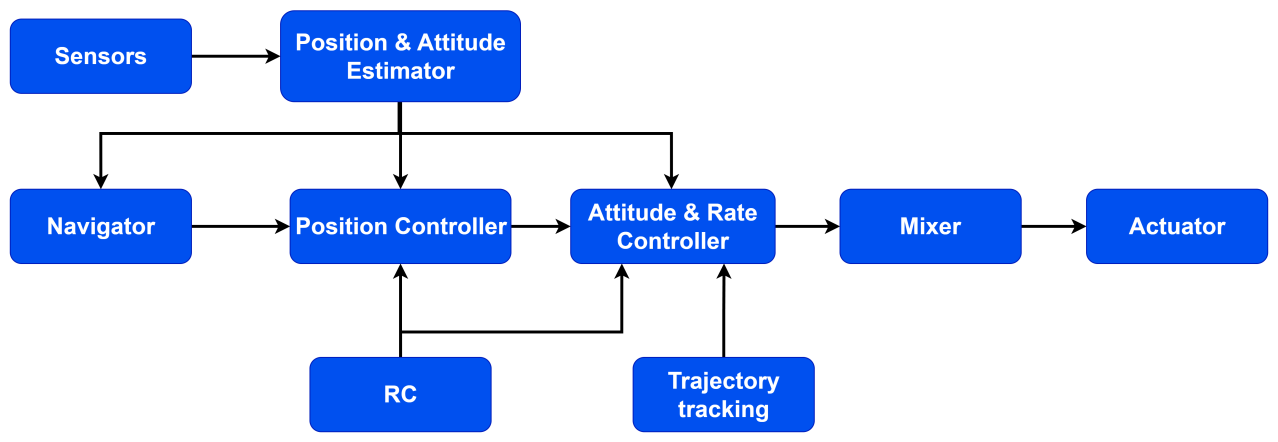
**Lập kế hoạch cục bộ (local planner):** là thành phần cốt lõi của điều hướng tự động, với nhiệm vụ sinh ra quỹ đạo khả thi trong không gian 3D, có xét đến ràng buộc động học của UAV và khả năng tránh vật cản theo thời gian thực. Nhiều thuật toán mới ra đời nhằm cải thiện tốc độ, độ an toàn và khả năng thực thi của quỹ đạo. Tiêu biểu có thể kể đến Fast-Planner [46] là một hệ thống kế hoạch quỹ đạo hiệu suất cao cho phép bay nhanh trong môi trường chưa biết, thuật toán kết hợp bước tìm kiếm động học (kinodynamic search) để tạo quỹ đạo ban đầu và bước tối ưu hoá B-spline dựa trên trường khoảng cách (EDF) để làm mượt và đảm bảo khả thi động lực học. Ngoài ra nhóm HKUST Aerial Robotics cũng giới thiệu EGO-Planner [47], một bộ lập kế hoạch cục bộ dựa trên gradient. Không giống như các thuật toán cổ điển dựa trên bản đồ đầy đủ (ESDF), EGO-Planner chỉ trích xuất thông tin chướng ngại vật cần thiết từ dữ liệu cảm biến khi cần và sử dụng một đường dẫn hướng (guiding path) để tính toán đạo hàm đẩy đường bay ra khỏi chướng ngại. Nhờ vậy, EGO-Planner rất nhẹ và nhanh, chỉ mất khoảng 1ms để lập một quỹ đạo mới, giảm đáng kể thời gian so với các phương pháp truyền thống. Ngoài ra, có nhiều giải thuật tối ưu quỹ đạo hướng đến thời gian tối ưu như nghiên cứu Minimum-Time Quadrotor Waypoint Flight [48] của Robotics and Perception Group (RPG) Đại học Zurich đã đề xuất một phương pháp tối ưu hoá cả về quỹ đạo và thời gian đồng thời, cho kết quả vượt qua khả năng của người điều khiển trong đua drone đạt 60 km/h. Trong khi đó, thuật toán FASTER [49] áp dụng chiến lược hai quỹ đạo song song: một quỹ đạo an toàn chỉ bay trong không gian đã biết và một quỹ đạo nhanh có thể đi xuyên vùng chưa biết để tiết kiệm thời gian. Drone chọn quỹ đạo theo mức độ chắc chắn của bản đồ và trạng thái môi trường, tạo ra khả năng bay linh hoạt và duy trì ở vận tốc cao mà vẫn đảm bảo an toàn. Gần đây, MaRS Lab, Đại học Hồng Kông đã phát triển SUPER [50] dựa trên FASTER, thiết kế cho bay tốc độ cao trong môi trường không biết trước. Điểm khác biệt chính của SUPER so với FASTER nằm ở phương pháp lưu trữ thông tin bản đồ và trình giải tối ưu quỹ đạo. SUPER duy trì bản đồ dạng PCM (Point Cloud Map) để lưu trữ thông tin từ cảm biến, trong khi FASTER sử dụng bản đồ OGM (Occupancy Grid Map) thuần túy. Với việc xử lý trực tiếp trên dữ liêụ point cloud kết hợp giải thụât IIM (Incremental Inflated Mapping) giúp cho SUPER giảm đáng kể thời gian cập nhật và tính toán inflation map so với phương pháp Raycasting của FASTER và hiệu quả hơn trong xử lý dữ liệu cảm biến ở khoảng cách xa. Bên cạnh đó, việc chuyển đổi bài toán tối ưu quỹ đạo về dạng không ràng buộc sử dụng MINCO [51] giúp cho SUPER nhanh chóng tìm được lời giải thay vì phương pháp MIQP trên FASTER. Hệ thống này đã chứng minh khả năng lập kế hoạch nhanh trong 5ms, với đảm bảo khả năng tránh vật cản khi di chuyển với vận tốc cao.

Bên cạnh các thuật toán dựa trên mô hình và cảm biến, các công nghệ học sâu (deep learning) đang dần chiếm vai trò quan trọng trong nghiên cứu điều hướng drone. Một số công trình nổi bật là “Champion-level drone racing using deep reinforcement learning*”* [52], sử dụng hệ thống Deep RL đã được huấn luyện để điều khiển drone đua với hiệu suất vượt cả phi công điều khiển drone chuyên nghiệp, “Learning Agile Flights Through Narrow Gaps with Varying Angles Using Onboard Sensing” [53] sử dụng Deep RL cho bài bay khó qua khe hẹp với góc tới lớn đạt được hiệu suất khá ấn tượng với tỉ lệ thành công lên tới 87.36 %. Tuy nhiên, dù đạt kết quả ấn tượng trong mô phỏng hoặc phòng lab, các hệ thống học sâu vẫn còn nhiều rào cản khi ứng dụng thực tế: khó kiểm chứng an toàn, cần dữ liệu huấn luyện lớn, khả năng tổng quát hạn chế và yêu cầu phần cứng mạnh. Do đó, hiện nay deep learning chủ yếu là công cụ bổ trợ trong điều hướng, như dự đoán môi trường, phát hiện vật cản hoặc tái lập quỹ đạo.

Trong tác vụ UAV bay theo dõi một mục tiêu di động, thuật toán bắt bám phải liên tục xác định vị trí mục tiêu từ cảm biến camera và LiDAR và điều chỉnh quỹ đạo bay để duy trì tầm nhìn. Có nhiều hướng tiếp cận như phương pháp thị giác truyền thống phát hiện đối tượng và tracking sử dụng bộ lọc Kalman hay tracker Siamese hay thuật toán tích hợp lập kế hoạch và điều khiển. Một ví dụ tiêu biểu thuật toán tích hợp lập kế hoạch và điều khiển là Elastic-Tracker [54]. Đây là một khung lập kế hoạch quỹ đạo không gian-thời gian đàn hồi nhằm theo dõi mục tiêu trong môi trường phức tạp. Elastic-Tracker kết hợp phát hiện mục tiêu và dự báo chuyển động, sau đó tìm đường dẫn tránh che khuất đảm bảo mục tiêu luôn trong tầm nhìn. Hệ thống xây dựng safe flight corridor thông minh dựa trên đường dẫn, và tối ưu hoá quỹ đạo cuối cùng sao cho các ràng buộc về va chạm và tầm nhìn được thỏa mãn đồng thời. Kết quả cho thấy Elastic-Tracker hoạt động tốt trong các tình huống khó như mục tiêu đổi hướng đột ngột, khuất tạm thời do chướng ngại vật và cần ít tài nguyên tính toán hơn các phương pháp trước đây. Điều này có nghĩa là Elastic-Tracker có thể chạy gần thời gian thực và đảm bảo an toàn. Khác với Elastic-Tracker, YOPOv2-Tracker [55] là một khung end-to-end sử dụng học sâu để kết hợp nhận thức và điều khiển. Thay vì chia tách nhiều giai đoạn, YOPOv2 huấn luyện một mạng nơ-ron nhận ảnh trực tiếp đầu vào và xuất ra lệnh điều khiển động cơ. Nó sử dụng cơ chế motion primitives dưới dạng các quỹ đạo mẫu để bao phủ không gian hành động có thể và dự đoán dịch chuyển quỹ đạo an toàn theo các chi phí về mượt, va chạm, tầm nhìn. Trong quá trình huấn luyện, YOPOv2 truyền ngược trực tiếp gradient từ chi phí quỹ đạo vào mạng, giúp học một chiến lược điều khiển thật nhanh và trực tiếp hơn so với thuần RL. Mô hình này đã được thử nghiệm thành công trên drone thực thử nghiệm trong rừng và trong nhà, chứng minh nó có thể xử lý được tình huống bay thật với độ trễ thấp và độ chính xác cao.

Bên cạnh hai ví dụ trên, còn có nhiều thuật toán khác cho theo dõi mục tiêu trên ảnh như SiamRPN [56], Deep-SORT [57], hay bộ lọc Kalman thường được dùng để tracking và ước lượng liên tục tọa độ mục tiêu, sau đó điều hướng UAV bám theo. Tuy nhiên, điểm yếu của những giải pháp tách rời này là khó đảm bảo tầm nhìn lâu dài và xử lý tốt va chạm – điều mà các bộ kế hoạch quỹ đạo như Elastic-Tracker hoặc hệ thống học tích hợp như YOPOv2-Tracker khắc phục được.

**Phần mềm điều khiển bay (flight controller):** Lớp phần mềm điều khiển bay là lớp thực thi thấp nhất trong hệ thống điều hướng cho drone. Lớp này được triển khai dưới dạng firmware trên các mạch điều khiển bay. Phần mềm điều khiển bay có nhiệm vụ xử lý dữ liệu từ các cảm biến trạng thái của drone như GNSS, IMU, barometer và nhận các tín hiệu điều khiển như lực đẩy (thrust), vận tốc góc (body rate) từ tầng điều khiển bám quỹ đạo (trajectory tracking) hoặc các lệnh điều khiển thủ công như điểm đặt từ người điều khiển và chuyển thành tín hiệu PWM/DSHOT ra ESC để điều khiển động cơ.



*Hình 15.5. Sơ đồ khối của phần mềm điều khiển bay*

Một số phần mềm điều khiển bay mã nguồn mở được sử dụng rộng rãi cho drone trên thế giới có thể kể đến như:

* PX4 [121]: PX4 là phần mềm điều khiển mã nguồn mở, được phát triển cho nhiều loại phương tiện không người lái mặt đất, trên không, PX4 thiết kế theo kiến trúc module linh hoạt, hỗ trợ cấu hình phức tạp và các tính năng tiên tiến cho cả các dự án UAV đơn giản và phức tạp. PX4 hỗ trợ giao thức MAVLink để giao tiếp với tầng điều khiển bám quỹ đạo và trạm điều khiển mặt đất (GCS), có hệ sinh thái phần cứng đa dạng, cộng đồng người dùng lớn, hỗ trợ cập nhật phần mềm liên tục. PX4 được phân phối theo giấy phép BSD 3 - Clause cho phép sử dụng, chỉnh sửa, phân phối lại cả trong các dự án thương mại mà không bắt buộc phải công bố mã nguồn đã chỉnh sửa.
* Ardupilot [122]: Ardupilot cũng là phần mềm điều khiển mã nguồn mở dành cho các phương tiện không người lái trên bộ, trên không, dưới nước với độ tin cậy cao, hệ sinh thái và cộng đồng phát triển lớn. Ardupilot được phân phối với giấy phép GPLv3 chặt chẽ hơn so với BSD 3 – Clause của PX4 với yêu cầu phải công bố mã nguồn nếu phân phối bản chỉnh sửa hoặc tích hợp vào các sản phẩm thương mại.
* Betaflight [123]: Betaflight là phần mềm điều khiển bay mã nguồn mở cho drone dạng đa cánh quạt và dạng cánh cố định. Betaflight tập trung vào tối ưu hóa thời gian phản hồi và ổn định bay nhằm đáp ứng nhu cầu bay biểu diễn và bay đua. Betaflight cũng được phân phối với giấy phép GPLv3 yêu cầu công bố mã nguồn khi phân phối bản chỉnh sửa hoặc tích hợp vào các sản phầm thương mại.

Từ phân tích trên, có thể thấy rằng PX4, ArduPilot và Betaflight đều là các nền tảng điều khiển bay mã nguồn mở với hệ sinh thái phong phú và được ứng dụng rộng rãi trong các hệ thống drone hiện nay. Tuy nhiên, sự khác biệt về kiến trúc phần mềm, khả năng mở rộng, và đặc biệt là giấy phép sử dụng khiến mỗi nền tảng phù hợp với những mục đích khác nhau. Betaflight chỉ phù hợp với các drone điều khiển thủ công với tốc độ cao cho các mục đích biểu diễn và đua. Ardupilot với hệ sinh thái và cộng đồng phát triển mạnh nhưng do giới hạn của giấy phép nên phù hợp với các dự án nghiên cứu hoặc mã nguồn mở. Còn PX4 với khả năng mở rộng linh hoạt, hệ sinh thái và cộng đồng lớn và giấy phép mở phù hợp cho các dự án phát triển phục vụ mục đích thương mại.

**15.1.3. Công nghệ điều khiển bầy đàn**

Bầy đàn Drones (hay các phương tiện không người lái) đang nổi lên như một công nghệ đột phá cho phép hệ thống tự động thông minh phân tán có khả năng tự cấu hình lại, theo yêu cầu, có tác động lớn đến nhiều lĩnh vực khoa học công nghệ và xã hội, bao gồm các hệ thống theo dõi, kiểm tra và vận chuyển. Trong bất kỳ ứng dụng nào, hệ thống bầy đàn luôn được kỳ vọng mang lại khả năng hoạt động hiệu quả hơn so với một phương tiện lớn đơn lẻ, đem lại tính linh hoạt đáng kể (khả năng thích ứng, mở rộng và bảo trì) và độ mạnh mẽ (độ tin cậy, khả năng sống sót và chịu lỗi) [58].

Các nghiên cứu chính đối với công nghệ điều khiển bầy đàn hiện nay chủ yếu tập trung vào các vấn đề: Lý thuyết phân tích hệ thống, phân tích động học, phân bổ tác vụ, tránh vật cản, ước lượng vị trí, khám phá môi trường và công nghệ giao tiếp. Cụ thể :

* Ước lượng vị trí: Fangcheng Zhu, Yunfan Ren và cộng sự [59] đã đưa ra thuật toán Swarm-LIO2 cho phép hệ thống bầy đàn có thể ước lượng vị trí dựa vào cơ chế phân tán sử dụng Lidar 3d, chỉ các thông tin với số chiều thấp, băng thông nhỏ được trao đổi từ đó giúp giảm thiểu sức nặng lên hệ thống hạ tầng. Bên cạnh đó Hao Xu, Peize Liu, Xinyi Chen, Shaojie Shen [60] cũng đưa ra thuật toán D2-SLAM, với việc sử dụng thông tin từ Sterio hoặc Omnidirectional Cameras cho phép đáp ứng nhiều nhu cầu vận hành khác nhau và khắc phục các thách thức về giới hạn trường nhìn đối với bầy đàn Drones
* Lập kế hoạch đường đi và điều hướng: Charbel Toumieh, Dario Floreano [61] đã đưa ra HDSM - Một motion planning framework cho phép bầy đàn Drone chuyển động tốc độ cao, hoạt động phân tán và đồng bộ trong môi trường nhiều vật cản không xác định.
* Phân chia nhiệm vụ: Ortools [62] do Google phát triển đã giúp giải quyết hiệu quả các bài toán phân chia tác vụ với trình giải VRP Solver với đa dạng các ràng buộc một cách nhanh chóng.
* Khám phá môi trường: Boyu Zhou, Hao Xu, Shaojie Shen [63] đã đưa ra thuật toán RACER sử dụng cơ chế phân tán cho phép bầy đàn Drones có thể khám phá môi trường hiệu quả bằng việc giải quyết bài toán CVRP (Capacitated Vehicle Routing Problem) kết hợp với phương pháp mở rộng bản đồ dựa trên các vùng biên giới (Frontier based Exploration). Trong khi đó Qianli Dong, Haobo Xi cùng cộng sự [64] đã cung cấp thuật toán MR-DTG giúp giảm thiểu một phần lớn lượng thông tin giao tiếp giữa các tác tử bằng cách sử dụng Multi Robot - Dynamic Topological Graph để lưu trữ thông tin các vùng đã và đang mở rộng, kết hợp Graph Voronoi Partion giúp cho việc phân bổ nhiệm vụ đến các tác tử một cách hợp lý.
* Giao tiếp: Arindam Majee, Rahul Saha và cộng sự [65] đã đưa ra các kiến trúc mạng phổ biến hoạt động cho hệ thống bầy đàn Drone bao gồm hệ thống phân tán và hệ thống tập trung và ưu nhược điểm của chúng.

Các ứng dụng của hệ thống bầy đàn Drone đa dạng trong cả lĩnh vực dân sự và quân sự. Trong đó:

* Dân sự:
* Trình diễn ánh sáng: Với sự kết hợp sử dụng 10500 Drones màn trình diễn ánh sáng tại thành phố Hồ Chí Minh vào kỷ niệm giải phóng miền Nam Việt Nam 30/4 -1/5 năm 2025, công ty Shenzhen DAMODA Intelligent Technology đã xác lập kỷ lục Guiness số lượng Drone điều khiển đồng thời một lúc trên thế giới.
* Thu thập thông tin địa hình: Các kỹ sư tại từ lab Autonomous Aerospace Systems đại học Michigan phát triển hệ thống swarm có thể tìm thấy, lập bản đồ và báo cáo hỏa hoạn giúp giảm thiệt hại và cứu hộ.
* Nông nghiệp: Dji Agricilture đã phát triển các dòng drone có dung tích 20-60 lít bồn chứa có khả năng phun từ 7-27 ha/ giờ gọn nhẹ, dễ vận hành giúp tăng năng xuất, rút ngắn thời gian làm việc và hoàn toàn tự động.
* Giao hàng: Amazon đã cho ra mắt dịch vụ Prime Air cho phép lựa chọn hơn 60000 mặt hàng có thể giao bằng Drone trong các khu vực được hỗ trợ. Bên cạnh đó, Meituan của Trung Quốc cũng phát triển hệ thống giao đồ ăn sử dụng Drone phục vụ cho các khu vực đông đúc với mật độ dân cư đông đúc một cách hiệu quả như ở Thâm Quyến, Hồng Kông.
* Quân sự:
* Tuần tra giám sát: Australia đã triển khai hệ thống các máy bay không người lái tiên tiến MQ4-C Triton phục vụ tuần tra, giám sát các vùng biển và biên giới phía bắc.
* Phối hợp tác chiến, tìm kiếm, tấn công mục tiêu: Các công ty khởi nghiệp Ukraine đã phát triển hệ thống tấn công nhiều Drone dựa vào AI cho lực luợng phòng thủ Ukraine trong việc tiêu diệt sinh lực phá hủy vũ khí của quân đội Nga với chi phí thấp.

Như vậy, với các nghiên cứu mới nhất của những năm gần đây, ta có thể thấy xu hướng phát triển và sử dụng hệ thống bầy đàn Drone ngày càng được phủ rộng khắp các lĩnh vực từ dân sự cho tới quân sự. Mức độ hiệu quả và khả năng thích ứng linh hoạt là điểm thành công nhất của công nghệ này.

**15.1.4. Hệ thống trung tâm điều hành và xử lý thông tin drone**

Trong những năm gần đây, việc ứng dụng máy bay không người lái đang ngày càng phát triển mạnh mẽ và trở thành xu thế công nghệ chủ đạo trong nhiều lĩnh vực, từ quân sự, an ninh cho đến dân sự, nông nghiệp và logistic. Cùng với sự phổ biến và đa dạng hóa về nhiệm vụ, kích thước và tính năng của drone, nhu cầu xây dựng các hệ thống điều hành và giám sát trung tâm để quản lý, điều phối và xử lý dữ liệu từ drone cũng trở nên cấp thiết hơn bao giờ hết. Điều này đã thúc đẩy nhiều tập đoàn công nghệ và quốc gia đầu tư nghiên cứu, sản xuất và triển khai các hệ thống trung tâm điều hành drone hiện đại, tích hợp công nghệ truyền thông, AI, điện toán đám mây và bản đồ số.

Những trung tâm này không chỉ thực hiện chức năng điều khiển bay theo thời gian thực, mà còn đóng vai trò thu thập, lưu trữ, phân tích và xử lý dữ liệu từ drone, từ đó phục vụ các mục đích như giám sát môi trường, cảnh báo thiên tai, bảo vệ an ninh lãnh thổ, kiểm tra công trình hạ tầng hay hỗ trợ các hoạt động cứu hộ – cứu nạn trong điều kiện khắc nghiệt. Tùy thuộc vào đặc thù ứng dụng, mỗi nhà sản xuất lại có chiến lược phát triển riêng, tập trung vào các chức năng chuyên biệt như giám sát biên giới, tuần tra vùng biển đảo, giao hàng tự động, giám sát giao thông, quản lý đô thị thông minh hoặc phục vụ sản xuất nông nghiệp công nghệ cao như phun thuốc, giám sát sâu bệnh, thu thập dữ liệu canh tác.

Với tốc độ phát triển nhanh của drone, các trung tâm điều hành đang dần trở thành một thành phần không thể thiếu trong hệ sinh thái UAV hiện đại. Dưới đây là một số công ty và tổ chức tiêu biểu trên thế giới đang phát triển và triển khai các hệ thống trung tâm điều hành và xử lý thông tin drone, đóng vai trò tiên phong trong việc kiến tạo hạ tầng vận hành drone một cách an toàn, hiệu quả và thông minh.

*Bảng 15.3. Một số trung tâm điều hành và xử lý thông tin Drone trên thế giới*

| **Stt** | **Tên sản phẩm** | **Lĩnh vực** | **Mô tả** | **Đối tượng quản lý** | **Tính năng** | **Tham khảo** |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | Helsing Altra | Quân sự | Nền tảng phần mềm tác chiến đa miền thông minh của Helsing, được thiết kế để kết nối tất cả các yếu tố trên chiến trường, từ máy bay không người lái (UAV), tình báo, trinh sát (ISR), pháo binh đến các phương tiện tác chiến khác. | UAV, pháo, các phương tiện tác chiến khác. | ◦ AI hỗ trợ phân tích đưa ra quyết định  ◦ Tự động tính toán điều chỉnh hỏa lực cho pháo binh ◦ Tổng hợp và trình bày hình ảnh chiến thuật (COP) thời gian thực từ nhiều nguồn  ◦ Tích hợp hiệu quả với các loại vũ khí, cảm biến và nền tảng bên thứ ba qua giao diện mở.  ◦ Xử lý tại biên các nhiệm vụ tấn công chính xác, trinh sát trong trường hợp bị gây nhiễu điện tử mạnh.  ◦ Dẫn đường, định vị và xác định mục tiêu không phụ thuộc GPS/GNSS | [68] |
| 2 | L3Harris-Amorphous | Quân sự | Nền tảng chỉ huy điều phối tự hành đa miền của L3Harris, được thiết kế để điều khiển hàng ngàn phương tiện tự hành trong các môi trường phức tạp. | UAV, UUV, USV, UGV. | ◦ Kiến trúc mở giúp dễ dàng tích hợp với các nền tảng và hệ thống khác nhau  ◦ Có khả năng mở rộng để điều khiển hàng ngàn phương tiện tự hành đồng thời  ◦ Quản lý, tích hợp và phân tích dữ liệu từ nhiều cảm biến khác nhau  ◦ Tự động hóa các tác vụ và giảm tải cho người điều hành ◦ Lên kế hoạch và phân công nhiệm vụ động, có thể điều chỉnh theo tình huống.  ◦ Cho phép chuyển giao quyền điều khiển giữa nhiều người vận hành hoặc các trung tâm chỉ huy khác nhau, hỗ trợ điều khiển phân tán và linh hoạt (Distributed Command and Control – C2) | [69] |
| 3 | Lorenz Hive | Dân sự | Cloud platform điều khiển UAV và UGV của Lorenz Technology | Drone, UGV | ◦ Xử lý AI tại biên cho phép phân tích data và detect tự động  ◦ Kết nối với cloud qua mạng 4G◦ Stream data với độ trễ thấp  ◦ Cloud service cho phép nhận data, điều khiển ở bất cứ đâu, dùng thiết bị gì  ◦ Report sự cố, xử lý từ video  ◦ Lưu log dữ liệu chuyến bay  ◦ Giới hạn vùng bay đảm bảo an toàn | [70] |
| 4 | VOTIX | Dân sự | Nền tảng điều khiển và quản lý UAV phù hợp với các nhiệm vụ khảo sát, giám sát, giao hàng, phản ứng nhanh trong tình huống nguy cấp (Drone as First Responder - DFR) | Drone | ◦ Hỗ trợ nhiều dòng drone như DJI, Autel, Parrot, ... và các drone sử dụng phần mềm điều khiển bay PX4 ◦ Logboook: Log lại dữ liệu nhiệm vụ, phân tích hiệu suất nhiệm vụ ◦ Tích hợp quản lý các trạm di động Drone-in-a-box ◦ Fleet Management: Quản lý các drone, cảm biến, trạm di động ◦ Streaming dữ liệu thời gian thực với độ trễ thấp ◦ Vận hành drone từ xa và bay ngoài tầm nhìn (BVLOS) thông qua internet  ◦ Sử dụng AI để phân tích hình ảnh và dữ liệu từ drone | [71] |
| 5 | DroneSense | Dân sự/ An ninh | Nền tảng quản lý và điều phối drone toàn diện thiết kế cho các nhiệm vụ giám sát tình huống, phản ứng nhanh trong các tình huống nguy cấp (Drone as First Responder - DFR) | Drone, camera an ninh. | ◦ Hỗ trợ đa dạng drone DJI (Matrice, Mavic), Parrot (Anafi USA), Autel Robotics  ◦ Stream data theo thời gian thực cung cấp góc nhìn toàn diện  ◦ Bản đồ hoạt động cung cấp nhiều tùy chỉnh theo layer để hiển thị tình huống, chỉ định các vùng đặc biệt  ◦ Fleet Management: Quản lý nhiều drone, bao gồm bảo trì, theo dõi và phân bổ nhiệm vụ.  ◦ Data Security: Đảm bảo an ninh dữ liệu và tuân thủ các quy định liên quan đến việc sử dụng drone trong các nhiệm vụ công cộng.  ◦ Có thể tích hợp với các hệ thống khác như CAD, RMS và các nền tảng GIS để cung cấp thông tin đầy đủ và chính xác. | [72] |
| 6 | Motorola CAPE | Dân sự | Phần mềm điều khiển và livestream drone được thiết kế cho các tình huống khẩn cấp và giám sát công cộng | Drone | ◦ Hỗ trợ đa dạng drone DJI, Parrot, Autel Robotics ◦ Livestream video HD: Phát trực tiếp video từ drone đến nhiều thiết bị di động, hỗ trợ ra quyết định nhanh chóng ◦ Điều khiển từ xa: Cho phép điều khiển drone từ xa qua mạng 4G/LTE hoặc 5G, hỗ trợ bay ngoài tầm nhìn (BVLOS) ◦ Quản lý dữ liệu an toàn: Lưu trữ video và dữ liệu chuyến bay trên đám mây, đảm bảo an ninh và dễ dàng truy cập. | [73] |
| 7 | UAVIA | Dân sự | Nền tảng quản lý, điều khiển drone chuyên nghiệp, cung cấp các công cụ để quản lý nhiều drone, lập kế hoạch nhiệm vụ và phân tích dữ liệu. Nền tảng này hỗ trợ các ứng dụng như giám sát xâm nhập, phát hiện tình huống. | Drone | ◦ Tạo và quản lý các kế hoạch bay cho từng drone, tích hợp khả năng phát hiện tình huống (xâm nhập bất hợp pháp, hỏa hoạn, người bị thương, ...)  ◦ Tự động tạo nhiệm vụ thu thập dữ liệu từ các cảm biến trong khu vực chỉ định, phân chia dữ liệu thu được dễ dàng quản lý  ◦ Chia sẻ dữ liệu nội bộ hoặc bên ngoài cho phép tham gia xem các nhiệm vụ đang diễn ra, hỗ trợ ra quyết định từ xa thông qua mạng internet | [74] |
| 8 | High Lander | Dân sự | Nền tảng điều phối drone chuyên nghiệp, cung cấp các công cụ lập kế hoạch nhiệm vụ và giám sát chuyến bay. Nền tảng này hỗ trợ các ứng dụng như giám sát hạ tầng, giao hàng và kiểm tra công nghiệp. | Drone | ◦ Tích hợp được nhiều loại drone thông qua SDK được cung cấp  ◦ Cung cấp hệ thống quản lý không phận Vega UTM cho phép giám sát, điều phối drone  ◦ Cung cấp khả năng thiết lập các nhiệm vụ tự động hoàn toàn cho drone  ◦ Quản lý và giám sát nhiệm vụ theo thời gian thực  ◦ Hỗ trợ các nhiệm vụ bay ngoài tầm nhìn (BVLOS) ◦ Không giới hạn số lượng nhiệm vụ quản lý đồng thời. | [75] |

Nhìn vào bảng so sánh các tính năng trên chúng ta có thể thấy rất ít các hệ thống chỉ huy điều khiển có khả năng đồng bộ, kiểm soát hiệu quả các phương tiện bay và robot mặt đất trong cùng một nền tảng tác chiến thống nhất. Phần lớn các hệ thống hiện tại chỉ tập trung vào một loại phương tiện, chủ yếu là UAV, nó là do sự khác biệt về dữ liệu cảm biến, yêu cầu điều khiển và địa hình hoạt động.

DroneSense là một nền tảng chỉ huy điều khiển UAV tiên tiến, được phát triển chuyên biệt cho các tổ chức công quyền tại Hoa Kỳ như cảnh sát, cứu hỏa, quản lý thiên tai. Hệ thống này được thiết kế để điều khiển thực địa giám sát từ xa và phân tích sau nghiệp vụ. Hệ thống DroneSense cho phép tích hợp với nhiều dòng drone đến từ các hãng lớn như DJI, Autel, Parrot, Skydio, Vantage Robotics, v.v., mang lại khả năng vận hành linh hoạt trong nhiều kịch bản ứng dụng khác nhau. Nền tảng này nổi bật với một loạt tính năng tiên tiến, trong đó có:

* Giám sát thời gian thực: Hệ thống hỗ trợ truyền video trực tiếp chất lượng HD với độ trễ cực thấp (< 0.5 giây), cho phép người điều hành nắm bắt tình hình tại hiện trường một cách tức thì và chính xác.
* Tích hợp bản đồ chiến thuật: Tương thích với các nền tảng bản đồ như ArcGIS, cùng với khả năng hiển thị lớp phủ radar thời tiết, vùng cháy rừng, đo khoảng cách, đánh dấu vị trí, và chú thích bản đồ – hỗ trợ lập kế hoạch và điều phối lực lượng hiệu quả.



*Hình 15.6. Hệ thống quản lý Dronesense*

* Quản lý nhật ký & thiết bị: Hệ thống lưu trữ và quản lý toàn bộ dữ liệu chuyến bay, theo dõi trạng thái hoạt động, lịch bảo trì của thiết bị bay, cũng như hồ sơ huấn luyện của phi công – giúp đảm bảo tuân thủ quy định và an toàn vận hành.
* Chỉ huy từ xa : Cho phép điều khiển drone từ xa qua giao diện web, sử dụng tay cầm như PS5 controller để thao tác điều khiển linh hoạt trong các tình huống khẩn cấp hoặc khoảng cách xa, đáp ứng nhu cầu vận hành từ trung tâm chỉ huy hoặc các đơn vị phối hợp.

**15.1.5. Công nghệ định vị và điều hướng dẫn đường cho robot mặt đất hoạt động trên địa hình đồi núi có đường mòn**

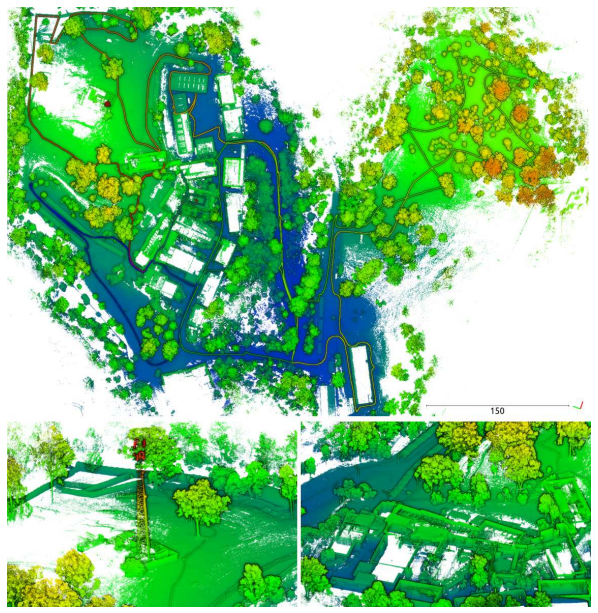
Trong những năm gần đây, xu hướng phát triển các hệ thống định vị chính xác cao cho robot mặt đất hoạt động ngoài trời, đặc biệt trong điều kiện rừng núi và địa hình phức tạp, đã được chú trọng mạnh mẽ trong nghiên cứu quốc tế. Trong số đó, việc ứng dụng công nghệ sử dụng GPS-RTK (Real-Time Kinematic) đang cho thấy hiệu quả vượt trội về mặt độ chính xác lẫn tính ứng dụng thực tế.

Nghiên cứu [99] đã đánh giá độ chính xác của hệ thống RTK-GNSS sử dụng thiết bị Trimble R12i trong môi trường rừng tự nhiên tại tỉnh Gangwon, Hàn Quốc - một khu vực có độ che phủ tán cây từ 30% đến 80%. Kết quả cho thấy hệ thống RTK vẫn giữ được độ chính xác định vị từ 0.26 đến 0.48 mét, ngay cả khi tín hiệu bị suy giảm do cây cối che phủ. Đây là minh chứng cho việc công nghệ định vị RTK hiện đại hoàn toàn có khả năng ứng dụng trong môi trường rừng dày và địa hình dốc, nơi tính hiệu GNSS thông thường thường bị hạn chế nghiêm trọng. Một hướng nghiên cứu thực tiễn và phù hợp với các ứng dụng robot mặt đất ngoài trời đến từ công trình của Liu và Nguyen [100], với mục tiêu phát triển hệ thống định vị chi phí thấp nhưng vẫn đạt độ chính xác trong môi trường nông nghiệp có cấu trúc không rõ ràng, địa hình phức tạp, mặt đất gồ ghề và tín hiệu GNSS thông thường thường xuyên bị suy giảm. Nhóm tác giả đã đề xuất so sánh hai phương pháp định vị: một là sử dụng mô hình Error-State Kalman Filter (ESKF) để tích hợp dữ liệu từ IMU giá rẻ, odometry bánh xe và mô-đun RTK-GNSS; hai là ứng dụng thuật toán LIO-SAM [101] kết hợp LiDAR và IMU để thực hiện SLAM. Cả hai hệ thống đều được triển khai trên một robot sáu bánh chạy ngoài thực địa, xử lý bằng máy tính nhúng Jetson Orin Nano. Kết quả cho thấy phương pháp sử dụng RTK-GNSS vẫn duy trì độ chính xác vị trí rất tốt trong thời gian thực và giảm chi phí đầu tư khoảng 36% so với hệ thống dựa trên LiDAR. Điều này cho thấy rằng việc kết hợp RTK-GNSS giá rẻ với các cảm biến khác là một lựa chọn khả thi, hiệu quả về mặt chi phí cũng như độ chính xác định vị cho robot hoạt động ngoài trời trong môi trường rừng núi, đồi dốc hoặc các khu vực không có bản đồ sẵn.

Một nghiên cứu tiêu biểu khác ứng dụng công nghệ định vị RTK-GNSS vào robot mặt đất ngoài trời được thực hiện bởi Moeller và cộng sự [102], với mục tiêu phát triển hệ thống dẫn đường tự hành cho robot nông nghiệp nhằm thực hiện nhiệm vụ loại bỏ các cây khoai tây bị nhiễm bệnh trên cánh đồng. Hệ thống sử dụng bộ điều khiển Pixhawk để điều khiển robot di chuyển theo các điểm đã định trước, trong khi thiết bị RTK-GNSS của Swift Navigation đảm nhận vai trò cung cấp dữ liệu định vị có độ chính xác ở mức centimet. Phương pháp này tận dụng tín hiệu GNSS được điều chỉnh thời gian thực (RTK) để dẫn đường chính xác đến từng cây khoai tây cần loại bỏ, đảm bảo robot di chuyển đúng quỹ đạo mong muốn, không gây ảnh hưởng đến những cây khỏe mạnh xung quanh. Kết quả thực nghiệm cho thấy hệ thống hoạt động hiệu quả, thực hiện thành công các tác vụ định vị và điều hướng ngoài trời với độ chính xác cao. Việc tích hợp mô-đun RTK-GNSS với nền tảng điều khiển phổ biến như Pixhawk có thể là một giải pháp thực tế cho các ứng dụng định vị chính xác trong môi trường nông nghiệp hoặc địa hình phức tạp đồng thời là cơ sở để mở rộng sang các nhiệm vụ tự động hóa phức tạp hơn cho robot mặt đất.

Bên cạnh những hệ thống định vị sử dụng GNSS mang lại độ chính xác cao trong môi trường ngoài trời, đặc biết các khu vực thoáng đãng, thì trong những tình huống đặc biệt tín hiệu vệ tinh suy giảm như dưới tán lá rừng rậm hoặc địa hình hiểm trở hay các khu vực hạn chế GNSS thì việc định vị bằng GNSS không còn hiệu quả. Chính vì vậy, các công nghệ định vị dựa trên SLAM sử dụng cảm biến LiDAR 3D đươc phát triển mạnh mẽ nhằm giúp Robot ngoài trời duy trì khả năng xác định vị trí và xây dựng bản đồ ngay cả khi không phụ thuộc vào tín hiệu GNSS. Một trong những thuật toán nổi bật về SLAM LiDAR 3D là LeGO-LOAM [103] và LIO-SAM [101]. LeGO-LOAM được thiết kế để hoạt động hiệu quả với dữ liệu cảm biến LiDAR 3D và IMU, sử dụng chiến lược phân loại điểm thành đất và không phải đất và áp dụng đồ thị tối ưu hóa để giảm trôi. LIO-SAM kế thừa cấu trúc của LeGO-LOAM nhưng cải tiến bằng cách tích hợp IMU theo cơ chế tightly-coupled va sử dụng factor graph để đồng xử lý dữ liệu LiDAR, IMU và nếu có thể, cả GNSS. Cả hai thuật toán đều cho kết quả tốt trong điều kiện thực địa, với khả năng giảm sai lệch tích lũy.

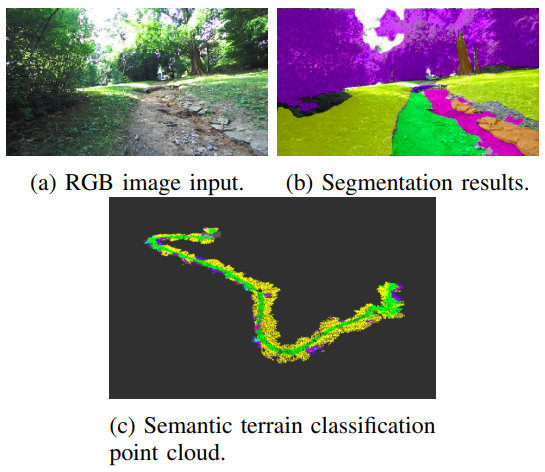
Trong các thuật toán LiDAR SLAM 3D hướng đến hiệu năng, FAST-LIO2 [104], Faster-LIO [105] là các hệ thống đáng chú ý nhất. Các thuật toán nay sử dụng bộ lọc Kalman trạng thái lỗi (iESKF) để kết hơp dữ liệu LiDAR và IMU theo cơ chế tightly-coupled, thay vì phục thuộc vào tối ưu đồ thị. Điều này giúp giảm độ trễ tính toán và tiết kiệm tài nguyên, rất phù hợp với các dòng robot UGV nhỏ hoặc hệ thống nhúng có giới hạn tính toán. Các thử nghiệm cho thấy hệ thống có thể duy trì sai số vị trí dưới một mét sau quãng đường vài km mà không cần GNSS, đặc biệt hiệu quả trong các môi trường khan hiếm đặc trưng như sa mạc, rừng khô hoặc đường mòn hoang dã.



*Hình 15.7. Wildcat [106] xây dựng bản đồ rộng lớn, phức tạp*

Một trong những hệ thống SLAM tiên tiến hiện nay là Windcat [106], được phát triển để hoạt động trong các điều kiện môi trường khắc nghiệt như hang động, đường hầm, hoặc địa hình phức tạp. Windcat sử dụng mô hình quỹ đạo thời gian liên tục để mô tả chuyển động của robot, và kết hợp dữ liệu từ LiDAR và IMU thông qua tối ưu hóa đồ thị. Hệ thống này được thử nghiệm thành công trong cuộc thi DARPA Subterranean Challenge và được đánh giá có độ ổn định vượt trội so với các thuật toán SLAM phổ biến khác. Wildcat phù hợp cho các ứng dụng đa robot hoặc các điều kiện cảm biến nhiễu động mạnh. Tổng quan các thuật toán SLAM sử dụng LiDAR 3D cho thấy sự đa dạng về phương pháp và hiệu quả ứng dụng trong nhiều môi trường ngoài trời khác nhau, từ khu vực bằng phẳng đến địa hình rừng núi phức tạp. Việc lựa chọn thuật toán phù hợp vào điều kiện cảm biến, tài nguyên tính toán và yêu cầu về độ chính xác hoặc thời gian thực của hệ thống. Đây cũng là nền tảng quan trọng để phát triển các hệ thống định vị tự hành cho robot mặt đất trong môi trường không có GPS hoặc tín vệ tinh không ổn định.

Bên cạnh đó, nhiều nghiên cứu quốc tế đã tập trung vào việc phát triển các phương pháp nhận diện đường mòn và điều hướng tự động cho Robot mặt đất hoạt động trong các môi trường khắc nghiệt, nơi địa hình phức tạp. Một trong những nghiên cứu tiêu biểu là GANav [74], đề xuất một phương pháp phân đoạn địa hình khả thi từ ảnh RGB bằng mạng attention nhẹ (group-wise attention). Hệ thống này được triển khai trên các robot như Jackal và Husky, giúp cải thiện độ chính xác trong việc xác định vùng đi được, đồng thời làm giảm đáng kể độ dao động trong quỹ đạo robot (4.6–13.9%) và tăng tỷ lệ hoàn thành nhiệm vụ thêm khoảng 10%. GANav cho thấy ưu điểm ở tốc độ xử lý thời gian thực và khả năng phân đoạn chính xác trên các địa hình gồ ghề. Tuy nhiên, hệ thống vẫn còn phụ thuộc nhiều vào điều kiện ánh sáng và chưa xử lý tốt các vật thể động hoặc môi trường có độ tương phản thấp.



*Hình 15.8. Kết quả phân đoạn địa hình từ ảnh RGB trong nghiên cứu [75]*

Tiếp theo là nghiên cứu của Reed et al. (2024) [75], hướng đến việc điều hướng robot trên các đường mòn thực tế bằng cách kết hợp giữa semantic segmentation từ ảnh RGB và dữ liệu hình học từ LiDAR. Nhóm nghiên cứu sử dụng mạng học sâu để phát hiện đường mòn và các vùng traversable, sau đó phân tích mặt đất bằng thuật toán phân cụm điểm LiDAR để phát hiện sườn dốc, đá lớn hoặc vật cản nguy hiểm. Kết quả được tích hợp vào bản đồ local costmap để lập kế hoạch đường đi một cách linh hoạt, vừa bám theo đường mòn vừa tránh chướng ngại vật. Ưu điểm của phương pháp này là tính chính xác cao trong môi trường thực địa (thử nghiệm tại các khu rừng thuộc West Virginia), nhưng đòi hỏi hệ thống cảm biến đồng bộ phức tạp và tiêu tốn tài nguyên tính toán.

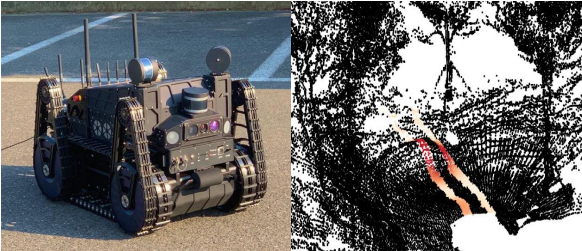
Trong hướng tiếp cận không cần bản đồ, Liang et al. (2023) [76] giới thiệu mô hình MTG (Mapless Trajectory Generator) sử dụng mạng học sâu với mô hình CVAE (Conditional Variational Autoencoder) kết hợp ràng buộc độ khả thi địa hình (traversability constraints). Hệ thống này sinh ra quỹ đạo điều hướng trực tiếp từ dữ liệu cảm biến mà không cần bản đồ tĩnh, nhờ vậy có khả năng ứng dụng tốt trong môi trường lạ hoặc chưa được khảo sát trước. Các thử nghiệm trên robot Husky và Spot cho thấy phương pháp này tăng diện tích vùng di chuyển an toàn lên 6% và giảm 89% số lần di chuyển ngoài ranh giới vùng an toàn. Tuy nhiên, mô hình MTG vẫn có hạn chế khi gặp điều kiện thời tiết xấu hoặc cảm biến đầu vào bị nhiễu.

Một nghiên cứu mang tính nền tảng khác là V-STRONG [77] – hệ thống học tự giám sát để nhận biết độ khả thi của địa hình mà không cần gán nhãn thủ công. Hệ thống sử dụng contrastive learning kết hợp với segmentation mask để huấn luyện mô hình dự đoán traversability một cách trực quan. V-STRONG không những đạt kết quả tốt trong huấn luyện thông thường, mà còn cho thấy khả năng zero-shot generalization, tức là vẫn hoạt động tốt trên các địa hình mới mà chưa từng gặp. Hệ thống được tích hợp vào controller để lập kế hoạch chuyển động thời gian thực. Dù vậy, phương pháp này phụ thuộc vào chất lượng dữ liệu thu thập ban đầu và đòi hỏi pipeline xử lý có khả năng học liên tục (online learning) để giữ hiệu quả dài hạn.

Về phía định vị, nghiên cứu TAIL [78] giới thiệu bộ dữ liệu SLAM đa modal gồm camera stereo, RGB-D, LiDAR quay, IMU và GPS RTK, được thu thập trên các môi trường địa hình biến dạng như cồn cát, đồi đất lún. Mục tiêu là xây dựng benchmark cho các thuật toán định vị trong điều kiện thiếu ổn định. Bộ dữ liệu này hỗ trợ kiểm thử các hệ thống định vị tự hành như ORB-SLAM, LIO-SAM, VINS-Mono,… đồng thời cho phép phân tích hiệu năng của hệ thống trong môi trường có độ nghiêng lớn, nhiễu cảm biến và thay đổi địa hình theo thời gian.

Một hướng nghiên cứu thú vị khác là sử dụng vision-language models để điều hướng trên địa hình phức tạp, điển hình là bài báo về điều hướng quadruped robot qua các cấu trúc địa hình như cầu, bậc thang và khe hẹp (2024) [79]. Mô hình được huấn luyện để “lý giải” mục tiêu chuyển động thông qua các biểu diễn hình ảnh và mô tả nhiệm vụ bằng ngôn ngữ tự nhiên. Điều này giúp hệ thống linh hoạt và có khả năng xử lý tác vụ đa dạng một cách khép kín. Tuy nhiên, việc sử dụng mô hình ngôn ngữ – thị giác yêu cầu nhiều tài nguyên (GPU, RAM) và không phù hợp với các robot cấp thấp hoặc hệ thống nhúng.

Các nghiên cứu như WayFAST [80] hay SCATE [81] cho thấy xu hướng mới trong việc học dự đoán traversability dựa trên dữ liệu thực tế từ robot hoạt động ngoài trời. Các mô hình này sử dụng thông tin hình học và hình ảnh để dự đoán xác suất vùng có thể đi qua và tích hợp vào hệ thống điều hướng cục bộ. Ưu điểm của hướng này là khả năng phản ứng nhanh và chính xác với môi trường thay đổi nhanh (như sạt lở, cây đổ). Tuy nhiên, các hệ thống này cần kiểm thử nghiêm ngặt trước khi triển khai thực địa để tránh dự đoán sai gây nguy hiểm.



*Hình 15.9. Mô hình robot và mountain trail driving data đề cập ở SCATE [81]*

Dù đã có nhiều nghiên cứu quốc tế đạt được những kết quả ấn tượng trong việc định vị và dẫn đường cho robot mặt đất trên địa hình đồi núi, việc tự chủ công nghệ trong lĩnh vực này vẫn mang tính chiến lược và cấp thiết, đặc biệt đối với các quốc gia đang phát triển hoặc có yêu cầu về an ninh – quốc phòng. Các hệ thống nhập khẩu hoặc mã nguồn mở thường không phù hợp hoàn toàn với đặc thù địa hình, khí hậu và yêu cầu nhiệm vụ tại Việt Nam, ví dụ như rừng rậm nhiệt đới, đồi núi đất đá, thời tiết ẩm ướt kéo dài, hoặc điều kiện ánh sáng thay đổi nhanh. Ngoài ra, nhiều công nghệ tiên tiến (deep learning-based planner, terrain-adaptive SLAM, self-supervised traversability learning) hiện vẫn được kiểm soát bởi các tập đoàn hoặc tổ chức nước ngoài, đi kèm với rào cản bản quyền, thiết bị phần cứng và rủi ro bảo mật. Do đó, việc tự nghiên cứu và làm chủ thuật toán định vị, dẫn đường không chỉ giúp tăng khả năng tùy biến hệ thống cho từng nhiệm vụ thực địa (như trinh sát, cứu hộ, tuần tra biên giới), mà còn đảm bảo tính bền vững lâu dài, độc lập công nghệ, và khả năng mở rộng cho các ứng dụng dân sự và quân sự trong tương lai.

**15.1.6. Công nghệ sinh dữ liệu để phục vụ huấn luyện mô hình, mô phỏng hoạt động của robot, drone**

Trong bối cảnh phát triển các mô hình Machine Learning cho robot và drone, việc thu thập dữ liệu thực thường phải đối mặt với nhiều thách thức: chi phí triển khai hệ thống drone hoặc robot cùng cảm biến cao, rủi ro hỏng hóc khi bay trong không gian hẹp hoặc thử nghiệm tải nặng, thiếu đa dạng kịch bản trong điều kiện ánh sáng và môi trường biến động, cũng như công đoạn gán nhãn tỉ mỉ và dễ phát sinh sai sót khi xác định pose hoặc phân đoạn semantic từng pixel. Lúc này, công nghệ sinh dữ liệu (Synthetic Data Generation) trở thành giải pháp chiến lược để khắc phục hạn chế của dữ liệu thực. Thay vì tốn kém thời gian, nhân lực và tài chính để thu thập, gán nhãn và xử lý dữ liệu từ môi trường thực - ví dụ như thu thập ảnh RGB, ảnh độ sâu, dữ liệu đám mây điểm LIDAR hay tín hiệu IMU - chúng ta có thể mô phỏng toàn bộ trong môi trường 3D. Mỗi khung hình thu được trong mô phỏng sẽ tự động gán nhãn chi tiết (bounding box, semantic mask, tọa độ pose 6-DOF), đảm bảo độ chính xác cao và nhất quán, đồng thời giảm thời gian hậu xử lý. Thông qua sinh dữ liệu tổng hợp, chúng ta không chỉ mở rộng quy mô dữ liệu lên hàng chục nghìn kịch bản khác nhau mà còn linh hoạt ngẫu nhiên hóa các tham số như kết cấu, vị trí, ánh sáng, đảm bảo tính đa dạng và độ phong phú cho tập dữ liệu huấn luyện.

Trong những năm gần đây, SDG đã trở thành một hướng nghiên cứu trọng tâm nhằm khắc phục các hạn chế của việc thu thập dữ liệu thực cho các hệ thống robot và drone. Các nghiên cứu quốc tế cho thấy SDG không chỉ giúp mở rộng quy mô dữ liệu mà còn tăng tính đa dạng của kịch bản, từ đó nâng cao khả năng khái quát và ổn định của các mô hình học máy. Một số công trình tiêu biểu bao gồm:

* **Dex-Net 3.0** [82]– bộ dữ liệu 6.7 triệu mẫu point cloud cho robot cánh tay pick-and-place, mô phỏng việc cầm nắm trên 1.500 mô hình 3D. Kết quả cho thấy chính sách học được có tỉ lệ thành công 84% trong thử nghiệm thực.
* **AirSim UAV Dataset** [83]– sử dụng Unreal Engine để tạo dữ liệu drone cho nhiệm vụ định vị và tránh chướng ngại, gồm 100.000 khung hình với nhiều kịch bản như thành phố, nông thôn và thời tiết khác nhau. Khi kết hợp với dữ liệu thực, mô hình SLAM tăng độ chính xác bản đồ lên 30%.
* **AI2-THOR Synthetic Benchmarks** [84]– cung cấp môi trường trong nhà tương tác cho các tác vụ đánh giá thị giác và điều hướng robot. Nhiều nghiên cứu đã dùng AI2-THOR để sinh dữ liệu semantic segmentation và navigation trajectories, giúp mô hình học sâu cải thiện kết quả định vị lên 25% so với không dùng SDG.

Các nghiên cứu này sử dụng 3 phương pháp chính trong SDG. Đầu tiên, **Domain Randomization** khởi xướng ý tưởng ngẫu nhiên hóa hoàn toàn các tham số hình ảnh—từ kết cấu đến ánh sáng, từ góc nhìn đến vị trí vật thể—nhằm buộc mô hình học tập tập trung vào đặc trưng hình học cốt lõi. Tiếp theo, **Procedural Generation** sử dụng các thuật toán quy tắc để tự động sinh ra cảnh 3D phức tạp, nhờ đó có thể mô phỏng môi trường với độ đa dạng cao nhưng vẫn cần thư viện asset phong phú và cấu hình quy tắc chặt chẽ để tránh tình trạng đơn điệu. Cuối cùng, **Photorealistic Rendering** áp dụng PBR kết hợp ray tracing, đem lại hình ảnh có độ chân thực rất cao, giúp giảm thiểu khoảng cách sim-to-real nhưng đòi hỏi hạ tầng GPU mạnh và chi phí thời gian tính toán lớn.

Bên cạnh đó, tích hợp động lực học chính xác bằng PhysX, Bullet hay MuJoCo với các engine đồ họa hiện đại cho phép mô phỏng đồng thời cả tương tác cơ học và quan sát thị giác, qua đó cung cấp dữ liệu đa giác độ cho các tác vụ học máy. Song song với đó, các phương pháp học tự chủ như fine-tuning trên dữ liệu thực và kỹ thuật domain adaptation (CycleGAN, DANN) ngày càng phổ biến, nhằm điều chỉnh mô hình được huấn luyện trên dữ liệu tổng hợp để hoạt động hiệu quả hơn trong môi trường thực. Kết quả từ các công trình như Synthetica, SynJet hay Dex-Net 3.0 đã chỉ ra rằng việc kết hợp khéo léo giữa các kỹ thuật sinh dữ liệu tổng hợp và phương pháp học sâu thích ứng miền có thể nâng cao đáng kể hiệu suất của hệ thống thị giác máy và điều khiển robot khi chuyển từ mô phỏng sang thực tế.

Để hỗ trợ SDG, bảy nền tảng mô phỏng chính thống được sử dụng rộng rãi trên thế giới là Gymnasium, MuJoCo, PyBullet, AirSim, NVIDIA Isaac, Genesis và AI2-THOR. Mỗi công cụ sở hữu những đặc điểm và ưu thế kỹ thuật riêng:

* **Gymnasium**: Gymnasium [85] là một nhánh duy trì bởi Farama Foundation của thư viện OpenAI Gym, cung cấp API chuẩn cho học tăng cường với thiết kế đơn giản, Pythonic. Nó bao gồm nhiều môi trường mẫu (CartPole, Pendulum, Box2D, MuJoCo, Atari), hỗ trợ tùy chỉnh phần thưởng, không gian quan sát/vận hành, vector hóa để tăng tốc đào tạo và tích hợp môi trường bên thứ ba. Gymnasium hoàn toàn tương thích với Gym cũ, cho phép chạy song song bất đồng bộ và là tiêu chuẩn phổ biến cho nghiên cứu, phát triển thuật toán và đánh giá hiệu năng.
* **MuJoCo**: MuJoCo [86] (Multi-Joint dynamics with Contact) là công cụ vật lý mã nguồn mở do DeepMind duy trì, chuyên mô phỏng động lực học đa khớp và tiếp xúc. Điểm nổi bật của MuJoCo là kết hợp hệ tọa độ tổng quát với phương pháp giải lực tiếp xúc tối ưu hóa lồi, cho phép mô phỏng tiếp xúc mềm, phân tích đảo ngược và xử lý va chạm ổn định, chính xác. Nó cung cấp pipeline tính toán có thể tùy biến (trường lực, trình điều khiển, va chạm), mô hình hình học linh hoạt và bộ truyền động tùy chỉnh. MuJoCo phù hợp cho nghiên cứu robot, cơ sinh học, hoạt hình và học máy, đặc biệt với hệ cơ xương phức tạp, và phiên bản mới nhất còn hỗ trợ tăng tốc qua mạng nơ-ron để nâng cao hiệu suất.
* **PyBullet**: PyBullet [87] là trình mô phỏng vật lý thời gian thực dựa trên Bullet với API Python, hỗ trợ đa nền tảng và tích hợp VR, ROS. Nó mô phỏng động lực học vật rắn, vật thể mềm, phát hiện va chạm, giải ràng buộc và trực quan hóa mạnh mẽ, kèm công cụ gỡ lỗi cho phép điều chỉnh thông số vật lý theo thời gian thực và xem quỹ đạo. Thư viện đi kèm nhiều mô hình robot, môi trường thử nghiệm, hỗ trợ học tăng cường và được cộng đồng phát triển tích cực duy trì.
* **AirSim**: AirSim [88] là trình mô phỏng UAV và ô tô mã nguồn mở của Microsoft, xây dựng trên Unreal Engine, cung cấp kết xuất môi trường chân thực và mô hình vật lý chính xác. Nó hỗ trợ đa dạng cảm biến (camera, LiDAR, IMU…), API phong phú cho điều khiển và thu thập dữ liệu, cùng khả năng ngẫu nhiên hóa miền và mô phỏng đa tác nhân. Công cụ này lý tưởng để tạo dữ liệu đào tạo chất lượng cao, phục vụ nghiên cứu thị giác máy tính, lái tự động, SLAM và điều hướng UAV, với khả năng tạo cảnh theo thủ tục, hệ thống thời tiết động và mô hình nhiễu thực tế của cảm biến.
* **NVIDIA Isaac**: NVIDIA Isaac [89] là nền tảng phát triển robot toàn diện gồm Isaac Sim, Isaac SDK và Isaac GYM, tận dụng GPU để mô phỏng và kết xuất vật lý chân thực. Nó hỗ trợ mô phỏng song song quy mô lớn, ngẫu nhiên hóa miền và tạo dữ liệu tổng hợp, đồng thời dễ dàng tích hợp với phần cứng thực. Với công cụ PhysX được tăng tốc GPU, Isaac cung cấp chuỗi công cụ đầy đủ cho nhận thức, điều hướng và vận hành robot, phù hợp cho ứng dụng công nghiệp, tự động hóa kho và robot cộng tác.
* **Genesis**: Genesis [90] là nền tảng mô phỏng vật lý tổng quát, triển khai hoàn toàn bằng Python với giao diện đầu cuối thân thiện. Nó mang lại hiệu suất vượt trội (nhanh hơn 10 - 80× so với các GPU-accelerated simulator hiện có) và kết xuất chân thực như ảnh. Hệ thống hỗ trợ nhiều solver tiên tiến để mô phỏng đa dạng vật liệu và hiện tượng, đồng thời có tính toán xúc giác. Điểm đột phá là tích hợp AI tạo sinh: người dùng có thể sinh kịch bản mô phỏng, dữ liệu nhiệm vụ và hàm phần thưởng chỉ qua mô tả ngôn ngữ tự nhiên. Kiến trúc song song tối ưu hóa cho phép mô phỏng thời gian thực quy mô lớn, giảm chi phí thu thập dữ liệu và tạo nền tảng nghiên cứu robot thống nhất, dễ mở rộng.
* **AI2-THOR:** AI2-THOR [91] là môi trường mô phỏng tương tác trong nhà, xây dựng trên Unity với kết xuất chân thực và vật lý chi tiết. Hầu hết các vật thể trong cảnh từ ngăn kéo, bếp, đến các đồ vật trang trí đều có thể thao tác: mở, đóng, di chuyển, nấu nướng, v.v. Nó hỗ trợ các nhiệm vụ điều hướng trực quan và thao tác đối tượng phức tạp, giúp nghiên cứu hiểu biết chức năng, mối quan hệ không gian và lập kế hoạch tác vụ trong môi trường nội thất. Phiên bản mới nhất bổ sung khả năng sinh cảnh nội thất theo thủ tục, tăng tính đa dạng và độ phức tạp cho thử nghiệm AI.

Sinh dữ liệu tổng hợp đã chứng minh giá trị vượt trội trong việc khắc phục những hạn chế của dữ liệu thực, đồng thời mang lại khả năng linh hoạt và hiệu suất cao. Việc kết hợp hợp lý giữa ba phương pháp Domain Randomization, Procedural Generation và Photorealistic Rendering, cùng với việc lựa chọn nền tảng mô phỏng phù hợp như Isaac Sim cho ứng dụng công nghiệp hay AirSim cho UAV, sẽ tối ưu hóa quy trình phát triển và thử nghiệm. Ngoài ra, tích hợp các kỹ thuật domain adaptation và học tự chủ sẽ giúp thu hẹp khoảng cách sim-to-real, đảm bảo mô hình hoạt động thực tế ổn định và chính xác. Trong tương lai, SDG tiếp tục mở ra nhiều hướng nghiên cứu mới và ứng dụng rộng rãi, góp phần thúc đẩy nhanh chóng quá trình triển khai robot và drone vào thực tiễn.

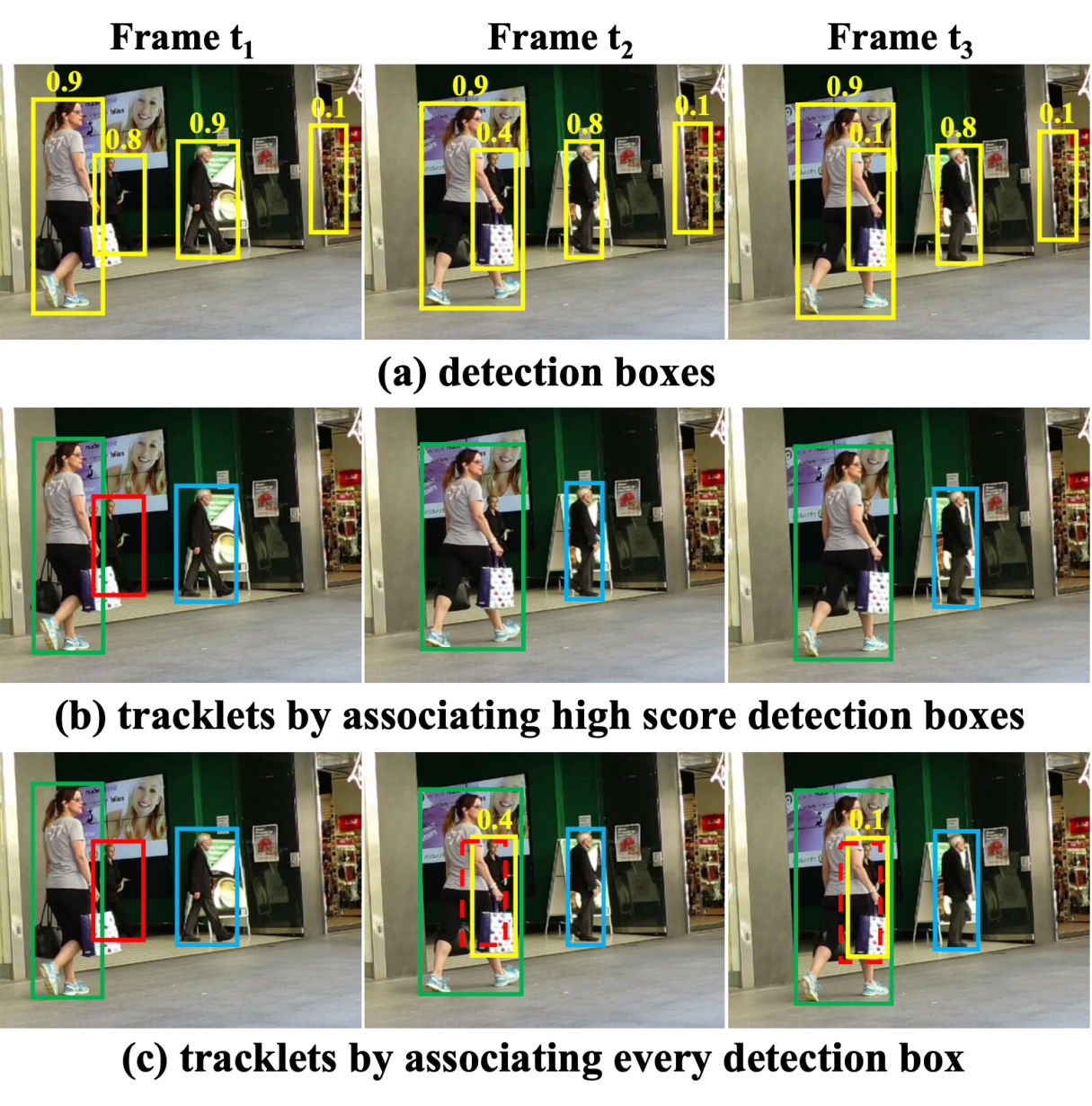
**15.1.7. Công nghệ nhận diện và bám bắt đối tượng**

Công nghệ nhận dạng và bắt bám đối tượng bằng camera trên thiết bị bay không người lái (UAV) đã trở thành lĩnh vực nghiên cứu trọng điểm, đóng vai trò quan trọng trong cả ứng dụng quân sự và dân sự. Trong quân sự, UAV tích hợp camera và thuật toán nhận dạng đối tượng được sử dụng để trinh sát, giám sát, chỉ thị mục tiêu, và thậm chí tấn công chính xác, nổi bật với khả năng tấn công tầm xa và chi phí thấp. Trong lĩnh vực dân sự, công nghệ này được ứng dụng rộng rãi trong giám sát môi trường, nông nghiệp thông minh, cứu hộ, và quản lý cơ sở hạ tầng, với ví dụ điển hình là tính năng ActiveTrack trên các dòng drone của DJI, được sử dụng để khảo sát địa hình và lập bản đồ với độ chính xác cao.



*Hình 15.10. ActiveTrack của DJI Mavic*

Các nghiên cứu gần đây đã đạt được những bước tiến đáng kể trong việc ứng dụng trí tuệ nhân tạo (AI) và học sâu (deep learning) vào hệ thống nhận dạng và bắt bám đối tượng trên UAV, tập trung vào các thuật toán tiên tiến để cải thiện độ chính xác và hiệu quả trong thời gian thực. Một trong những thuật toán nổi bật là YOLOv8 (You Only Look Once), được tối ưu hóa để nhận dạng đối tượng với tốc độ cao và độ chính xác vượt trội, phù hợp cho các ứng dụng UAV nhờ khả năng xử lý hình ảnh từ camera trong điều kiện môi trường phức tạp [92]. YOLOv8 sử dụng kiến trúc mạng nơ-ron sâu với các lớp tích chập (convolutional layers) để phát hiện đối tượng trong một lần quét duy nhất, giảm thiểu độ trễ tính toán, điều này đặc biệt quan trọng đối với các UAV có tài nguyên xử lý hạn chế. Một thuật toán khác, Deep SORT [93] (Simple Online and Realtime Tracking with a Deep Association Metric), được phát triển để theo dõi liên tục các đối tượng qua các khung hình video, tích hợp đặc trưng học sâu từ mạng ResNet để cải thiện khả năng nhận diện đối tượng trong môi trường động, như giám sát giao thông hoặc trinh sát quân sự. Deep SORT sử dụng bộ lọc Kalman và thuật toán ghép cặp Hungary để duy trì danh tính đối tượng, ngay cả khi đối tượng bị che khuất tạm thời. Trong khi đó, ByteTrack [94], một phương pháp mới hơn, đã được đề xuất để tối ưu hóa việc theo dõi các đối tượng có độ tin cậy thấp, đặc biệt trong các tình huống mà đối tượng bị che khuất hoặc ánh sáng yếu, như trong các ứng dụng giám sát mùa vụ tại Trung Quốc. ByteTrack tận dụng mọi hộp giới hạn (bounding box) được phát hiện, kể cả những hộp có điểm số thấp, để duy trì tính liên tục của quá trình theo dõi, cải thiện hiệu suất lên đến 10% so với các phương pháp trước đó trong các kịch bản phức tạp. Một nghiên cứu tiêu biểu, “Multi-Object Tracking Meets Moving UAV” [95], đã đề xuất mạng UAVMOT chuyên biệt cho việc theo dõi đa đối tượng trong video UAV. Nghiên cứu này giới thiệu mô-đun cập nhật đặc trưng ID để tăng cường khả năng liên kết đặc trưng đối tượng và mô-đun bộ lọc chuyển động thích ứng để xử lý chuyển động phức tạp trong góc nhìn UAV. Kết quả thực nghiệm trên các tập dữ liệu VisDrone2019 và UAVDT cho thấy UAVMOT đạt được cải thiện đáng kể với MOTA lần lượt là 36.1% và 46.4%, vượt trội so với các phương pháp tiên tiến khác trong việc theo dõi đa đối tượng từ UAV.



*Hình 15.11. Ví dụ kết quả của ByteTrack [96]*

Tuy nhiên, thách thức lớn nhất là đảm bảo độ chính xác và ổn định của hệ thống trong các điều kiện thời tiết bất lợi, nhiễu động khí quyển, và môi trường cùng với đối tượng phức tạp. Nhu cầu làm chủ công nghệ này ngày càng cấp thiết, đặc biệt trong bối cảnh cạnh tranh địa-chính trị và yêu cầu tự chủ quốc phòng. Việc phát triển các hệ thống nhận dạng và bắt bám đối tượng không chỉ nâng cao hiệu quả tác chiến mà còn thúc đẩy chuyển đổi số trong các lĩnh vực dân sự, từ nông nghiệp đến quản lý tài nguyên. Do đó, đầu tư vào nghiên cứu và phát triển công nghệ này là ưu tiên chiến lược để đảm bảo tính cạnh tranh và an ninh quốc gia.

**B. Trong nước**

Tại Việt Nam, lĩnh vực nghiên cứu và phát triển công nghệ UAV (thiết bị bay không người lái) đang có nhiều bước tiến rõ rệt trong cả môi trường học thuật và công nghiệp. Ban đầu, các nghiên cứu UAV chủ yếu được thực hiện dưới dạng đề tài khoa học tại các trường đại học, viện nghiên cứu như công trình nghiên cứu về áp dụng UAV trong việc khảo sát mỏ của Trường Đại học Mỏ Địa Chất [96], dự án về sử dụng thiết bị bay không người lái phục vụ giám sát cây nông nghiệp của Viện ứng dụng Công Nghệ và Đại học Quốc gia Hà Nội [97]. Các công trình nghiên cứu này tập trung vào ứng dụng UAV trong lập bản đồ 3D [98], khảo sát mỏ, quan trắc nông nghiệp và tích hợp cảm biến môi trường. Tuy nhiên, phần lớn vẫn dừng lại ở giai đoạn thử nghiệm trong phòng thí nghiệm, mang tính khảo sát và mô hình hoá, chưa có nhiều sản phẩm thương mại hóa thành công.

Gần đây, cùng với xu thế công nghiệp hóa – số hóa và nhu cầu thực tế từ thị trường, một số doanh nghiệp Việt Nam đã đầu tư mạnh vào nghiên cứu, sản xuất UAV, hướng đến xây dựng hệ sinh thái sản phẩm phục vụ trong nhiều lĩnh vực như an ninh, nông nghiệp, giao thông, môi trường. Có thể kể đến các đơn vị tiêu biểu như Realtime Robotics (RtR), CT UAV (CT Group), MAJ Vietnam, và Viettel High-Tech. Các sản phẩm của các công ty này không chỉ dừng lại ở phần cứng, mà còn tích hợp nhiều công nghệ tiên tiến như AI xử lý biên (edge-AI), khả năng điều khiển theo bầy đàn (swarm), camera đa cảm biến, cùng các nền tảng lập bản đồ thời gian thực và phần mềm điều phối bay thông minh.

* **Lĩnh vực UAV dân dụng, đa dụng**

Realtime Robotics (RtR) là một trong những công ty UAV tiên phong của Việt Nam với sản phẩm chủ lực là HERA – một drone AI có thiết kế gập gọn, tích hợp 4 camera AI, hỗ trợ tải trọng lên tới 15 kg. HERA có khả năng thay đổi payload linh hoạt, thời gian bay lên đến một giờ và tích hợp xử lý AI mạnh mẽ trên thiết bị để thực hiện nhận diện đối tượng, theo dõi và lập bản đồ môi trường theo thời gian thực. Đây là sản phẩm “Made in Vietnam” hiếm hoi đã có bằng sáng chế tại Mỹ và được xuất khẩu sang thị trường Bắc Mỹ.

Bên cạnh đó, CT UAV – một công ty thành viên của CT Group – đã phát triển một hệ sinh thái UAV đa năng với nhiều dòng sản phẩm phục vụ các mục tiêu khác nhau như UAV chở người, UAV cứu hộ, UAV phục vụ biểu diễn ánh sáng, giao thông đô thị, giám sát hạ tầng và nông nghiệp. CT UAV nổi bật với chiến lược tích hợp công nghệ swarm (điều khiển UAV theo bầy đàn) và khả năng bay xa tự động. Gần đây, công ty công bố hợp đồng xuất khẩu 5000 UAV vận tải không người lái đầu tiên cho Hàn Quốc, minh chứng cho việc nghiên cứu và thương mại hóa nhanh chóng trong lĩnh vực này.

MAJ Vietnam là đơn vị tập trung vào cả UAV dân sự và quân sự với các sản phẩm drone cánh cố định và drone đa cánh phục vụ giám sát, phun thuốc, đào tạo, và logistics. Công ty chủ động trong toàn bộ chuỗi sản xuất từ thiết kế, lắp ráp đến bảo trì.

* **Lĩnh vực UAV quân sự và phòng thủ**

Viettel High-Tech, một thành viên của Tập đoàn Viettel, là đơn vị tiên phong trong việc phát triển UAV quân sự tại Việt Nam. Các sản phẩm nổi bật bao gồm UAV VT-Patrol, VU-R70, và gần đây là loitering munition VU-C2 – dòng UAV có khả năng bay lượn và tấn công mục tiêu theo kiểu “tự sát”. Ngoài ra, Viettel còn phát triển hệ thống radar phát hiện UAV và hệ thống đối kháng điện tử để phòng thủ.

* **Các nghiên cứu của Viettel AI**

Trước đây nhóm đề tài đã thực hiện đề tài “Nghiên cứu phát triển nền tảng phần mềm cho robot tự hành thông minh”, mã số đề tài: 006-22-TĐ-RDP-DS, chủ nhiệm đề tài: Nguyễn Hòa Bình, thời gian thực hiện: tháng 08/2023 đến tháng 02/2024. Đề tài đã đặt nền móng cho việc xây dựng nền tảng phần mềm cho robot tự hành thông minh, tập trung làm chủ các công nghệ liên quan cho dòng robot tự hành mặt đất trong môi trường trong nhà, làm cơ sở cho các dòng robot tự hành ngoài trời. Trong đề tài trước, nhóm đề tài đã xây dựng và thử nghiệm thành công một số công nghệ có tiềm năng kế thừa và phát triển cho đề tài hiện tại, bao gồm: công nghệ bắt bám đối tượng theo thời gian thực, hệ thống định vị và thiết lập bản đồ sử dụng cảm biến LiDAR và công nghệ định vị VSLAM.

Cụ thể, nhóm đã xây dựng hệ thống bắt bám đối tượng sử dụng camera 2D và mạng nơ-ron kết hợp thuật toán ByteTrack [94], cho phép robot nhận diện và theo dõi người với tốc độ xử lý trên 30 FPS. Hệ thống đạt độ chính xác theo chỉ số MOTA lên đến 79.2% trên tập dữ liệu MOT17 [107], đảm bảo khả năng theo dõi ổn định. Bên cạnh đó, nhóm đề tài cũng đã triển khai hệ thống định vị và thiết lập bản đồ trong nhà sử dụng cảm biến LiDAR 2D, 3D với sai số định vị lặp 2 centimet và độ chính xác vượt trội hơn so với SLAM toolbox [108] và Google Cartographer [109] trên tập dữ liệu của MIT [110]. Đối với công nghệ định vị VSLAM, nhóm đề tài đã có bước đầu nghiên cứu và đánh giá về các thuật toán Visual SLAM nổi bật như ORB-SLAM3 [10] và cuVSLAM [111] của Nvidia Isaac. Các kết quả này không chỉ cho thấy khả năng mở rộng và phát triển cho bài toán robot mặt đất hoạt động ngoài trời mà còn mở ra hướng ứng dụng rộng hơn trong các yêu cần nhận biết mục tiêu bám bắt đối tượng cho máy bay không người lái.

**15.2. Luận giải về những nội dung cần nghiên cứu của đề tài**

**15.2.1. Cơ sở pháp lý**

Căn cứ Quyết định số 9415/QĐ- CNVTQĐ về việc ban hành “Quy chế Khoa học và Công nghệ của Công ty mẹ - Tập đoàn Công nghiệp - Viễn thông Quân đội” được Chủ tịch Tập đoàn phê duyệt ngày 09/11/2023;

Căn cứ Quyết định số 13382/QĐ-CNVTQĐ ngày 07/11/2024 của Chủ tịch Tập đoàn về việc sửa đổi một số điều của Quy chế KHCN của Công ty mẹ - Tập đoàn Công nghiệp - Viễn thông Quân đội ban hành kèm theo Quyết định số 9415/QĐ-CNVTQĐ ngày 09/11/2023;

Căn cứ theo Kế hoạch số 2369/KH-VAI ngày 05/12/2024 của Giám đốc Trung tâm Dịch vụ dữ liệu và Trí tuệ nhân tạo về Kế hoạch sản xuất kinh doanh và xây dựng đơn vị năm 2025;

**15.2.2. Tính cấp thiết**

Kinh tế tầng thấp (Low-Altitude Economy) là một loại hình kinh tế mới liên quan đến các hoạt động kinh tế diễn ra trong không phận tầm thấp dưới 1000m (có thể mở rộng đến dưới 3000m tùy thuộc vào đặc điểm từng khu vực). Với sự phát triển nhanh chóng của các phương tiện bay không người lái, phương tiện bay điện cất hạ cánh thẳng đứng (eVTOL) và Robot, cùng với sự trưởng thành của các công nghệ nền tảng như 5G, AI và IoT, việc thu nhập và xử lý dữ liệu thời gian thực, điều phối thông minh, vận hành chính xác và phản ứng nhanh đã trở nên khả thi. Nhờ đó, nền kinh tế tầm thấp đang hình thành và phát triển mạnh mẽ, trở thành xu thế tất yếu trên phạm vi toàn cầu.

1. **Tính cấp thiết về việc làm chủ công nghệ**

Trong bối cảnh thị trường UxV ngày càng đa dạng về công nghệ và xuất xứ, tình trạng thiếu đồng bộ về tiêu chuẩn kỹ thuật và kết nối đang đặt ra nhiều thách thức trong quản lý và điều phối. Bên cạnh đó, các phương tiện hiện đại còn có khả năng thu thập và truyền tải dữ liệu theo thời gian thực, nếu bị chiếm quyền điều khiển sẽ có nguy cơ trở thành công cụ tấn công, phá hoại đe dọa an ninh quốc gia. Vì vậy, cần thiết phải làm chủ hệ thống chỉ huy phương tiện không người lái này, để đảm bảo an toàn an ninh trong nền kinh tế tầm thấp. Các công nghệ lõi chính trong hệ thống điều khiển và nền tảng chi huy thông minh các phương tiện không người lái trên bộ và trên không bao gồm:

1. Công nghệ AI thị giác máy tính trong định vị Drone: Trong nhiều tình huống bay, drone có thể mất tín hiệu RTK/GPS do nhiễu sóng hoặc bị che khuất, dẫn đến việc định vị tuyệt đối trở nên không chính xác và thiếu tin cậy. Vì vậy, công nghệ Visual SLAM được ứng dụng để định vị và xây dựng bản đồ một cách độc lập, không phụ thuộc vào RTK/GPS, giúp drone duy trì khả năng tự hành trong mọi điều kiện. Việc nghiên cứu và làm chủ công nghệ này kết hợp với các tính năng AI hỗ trợ không chỉ mở rộng phạm vi ứng dụng mà còn giúp drone hoạt động hiệu quả trong nhiều môi trường phức tạp và đa dạng. Qua đó, tính năng định vị và xây dựng bản đồ ngày càng được hoàn thiện, đáp ứng tốt hơn các yêu cầu thực tế trong sản phẩm và ứng dụng.
2. Công nghệ điều hướng tự động cho Drone: Trong nội dung này nhóm đề tài tập trung nghiên cứu và phát triển hệ thống điều hướng tự động (Autonomous Navigation System - ANS) dành cho phương tiện bay không người lái Drone hoạt động trong môi trường có vật cản động, địa hình không đồng đều. Mục tiêu của nghiên cứu này là xây dựng một hệ thống điều hướng tự động thông minh, có khả năng lập kế hoạch đường bay an toàn và chính xác, tránh vật cản theo thời gian thực, có thể mở rộng cho các nhiệm vụ nâng cao như theo dõi mục tiêu động, giám sát, dò tìm cứu nạn.
3. Công nghệ điều khiển bầy đàn cho Drone: Nội dung này tập trung vào phát triển các thuật toán tối ưu vận hành đồng thời nhiều Drone theo cách tự động, an toàn và hiệu quả. Ứng dụng trực tiếp trong lĩnh vực quân sự, quốc phòng, nơi việc sử dụng đồng thời nhiều Drone phục vụ các nhiệm vụ như trinh sát diện rộng, định vị mục tiêu, gây nhiễu điện tư, yểm trợ hỏa lực, dò tìm mìn, bom, vật thể nguy hiểm, tấn công chính xác theo đội hình. Việc phát triển công nghệ bầy đàn trong nước mang ý nghĩa chiến lược, nhằm làm chủ công nghệ lõi, giảm sự phụ thuộc vào các giải pháp nước ngoài, đặc biệt trong bối cảnh an ninh mạng và bảo mật quân sự ngày càng quan trọng. Do đó, nhóm đề tài hướng đến việc xây dựng từ đầu một framework điều khiển drone bầy đàn đảm bảo tính linh hoạt, bảo mật và hiệu năng.
4. Hệ thống trung tâm chỉ huy thông minh cho các phương tiện không người lái: Hệ thống này xây dựng khung kiến trúc trung tâm chỉ huy chuyên biệt cho hệ thống điều phối, giám sát và lưu trữ dữ liệu drone quy mô lớn. Hệ thống còn đảm bảo quản lý các luồng camera với độ trễ thấp nhất, truyền tải dữ liệu video, hình ảnh xuống thiết bị và lưu dữ liệu vào ổ cứng. Hệ thống cũng xây dựng các màn hình giám sát, điều phối, quản lý đặt lịch bay theo lịch trình đã định sẵn. Ngoài ra hệ thống này còn lưu lại lịch sử chuyến bay, lịch sử tác động của người dùng và các api có thể cho bên khác tích hợp vào. Với các tính năng đấy nhóm đề tài sẽ xây dựng một hệ thống ưu tiên viêc tối ưu hiệu năng, độ chính xác từ đó có thể dễ dàng nâng cấp số lượng thiết bị lớn nhất hệ thống có thể quản lý.
5. Công nghệ định vị kết hợp GPS-RTK và điều hướng dẫn đường cho Robot mặt đất hoạt động trên địa hình đồi núi có đường mòn: Trong bối cảnh phát triển các hệ thống Robot tự hành phục vụ quốc phòng - an ninh, nông nghiệp, giám sát, việc định vị chính xác và điều hướng an toàn cho Robot mặt đất tại các khu vực đồi núi có địa hình phức tạp là một yêu cầu cấp thiết. Đặc thù của địa hình đồi núi Việt Nam rộng lớn, đường mòn nhỏ, tiềm ẩn nguy cơ trượt dốc hoặc lạc đường. Do đó, việc nghiên cứu và phát triển công nghệ định vị chính xác kết hợp GPS-RTK và các cảm biến hỗ trợ như IMU, LiDAR,... là nền tảng cốt lõi để đảm bảo Robot hoạt động ổn định, an toàn và thông minh trong các điều kiện khắc nghiệt. Việc làm chủ công nghệ này không chỉ giúp tăng cường năng lực nội địa trong việc chế tạo Robot mà còn giảm sự phụ thuộc vào các giải pháp thương mại nước ngoài, vốn thường thiếu khả năng tùy biến theo bài toán đặc thù, chi phí cao, tiềm ẩn rủi ro về bảo mật.
6. Công nghệ sinh dữ liệu để phụ vụ huấn luyện mô hình, mô phỏng hoạt động của Robot và Drone: Cùng với phát triển nhanh chóng của tự động hóa và trí tuệ nhân tạo, việc huấn luyện các mô hình thị giác máy (computer vision) và điều khiển (control) cho robot, drone đòi hỏi khối lượng dữ liệu lớn với độ đa dạng cao. Công nghệ sinh dữ liệu (Synthetic Data Generation – SDG) đã trở thành giải pháp then chốt để khắc phục các giới hạn về chi phí, thời gian và an toàn khi thu thập dữ liệu thực tế. Bảng dưới đây so sánh chi tiết hiệu suất của các nền tảng mô phỏng theo nhiều tính năng kỹ thuật khác nhau:

*Bảng 15.4. Bảng so sánh chi tiết hiệu suất của các nền tảng mô phỏng*

| **Tính năng** | **Gymnasium** | **MuJoCo** | **PyBullet** | **AirSim** | **Isaac Sim** | **Genesis** | **AI2-THOR** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Hỗ trợ tăng tốc GPU | Không | Không | Không | Không | Có | Có | Có |
| Hỗ trợ URDF | Không | Có | Có | Không | Có | Có | Không |
| Hỗ trợ động lực học | Đơn giản | Hoàn chỉnh | Hoàn chỉnh | Hoàn chỉnh | Hoàn chỉnh | Hoàn chỉnh | Đơn giản |
| Hỗ trợ cảm biến | Phản hồi trạng thái | Hoàn chỉnh | Hoàn chỉnh | Hoàn chỉnh | Hoàn chỉnh | Hoàn chỉnh | Thị giác |
| Hỗ trợ đa nền tảng | Có | Có | Có | Có | Có | Có | Có |
| Bảo trì liên tục | Có | Có | Có | Không | Có | Có | Có |
| Hỗ trợ mô phỏng chất lỏng | Không | Không | Có | Kh | Có | Có | Không |
| Multiple Intelligences | Không | Có | Có | Có | Có | Có | Có |
| Tích hợp ROS | Không | Có | Có | Có | Có | Có | Không |
| API Python | Có | Có | Có | Có | Có | Có | Có |
| Hỗ trợ RL | Gốc | Tốt | Tốt | Trung bình | Tốt | Tốt | Tốt |
| Chất lượng kết xuất | Thấp | Trung bình | Trung bình | Cao | Cao | Cao | Cao |
| Domain Randomization | Hạn chế | Có | Có | Có | Đầy đủ | Đầy đủ | Có |
| Độ khó triển khai | Thấp | Trung bình | Thấp | Cao | Cao | Thấp | Trung Bình |
| Learning curve | Thấp | Trung Bình | Thấp | Cao | Cao | Thấp | Trung bình |
| Cộng đồng hỗ trợ | Lớn | Lớn | Trung bình | Thấp | Lớn | Trung bình | Trung bình |

Khi so sánh các nền tảng mô phỏng phổ biến như Gazebo, Unity3D, PyBullet và AirSim, NVIDIA Isaac Sim nổi lên như một giải pháp toàn diện và tối ưu cho nhu cầu sinh dữ liệu của robot và drone. Isaac Sim tận dụng sức mạnh của ray tracing RTX và Physically-Based Rendering (PBR) trên nền tảng OpenUSD để cho ra hình ảnh rất chân thực, tương đương với thế giới thực. Nhờ khả năng mở rộng đa GPU, nền tảng này có thể chạy đồng thời hàng nghìn phiên bản mô phỏng, từ đó tăng tốc độ sinh dữ liệu lên nhiều lần so với các công cụ chỉ sử dụng CPU hoặc GPU đơn. Bên cạnh đó, Isaac Sim hỗ trợ đầy đủ các loại sensor như camera RGB, camera độ sâu (depth), LiDAR và IMU, đồng thời tích hợp chặt chẽ với Omniverse Replicator để tự động gán nhãn semantic mask, bounding box và tọa độ pose 6-DOF mà không cần can thiệp thủ công. Khả năng kết nối trực tiếp với ROS2 giúp nhóm phát triển dễ dàng tích hợp pipeline mô phỏng vào hệ sinh thái robot đang sử dụng, đồng thời tăng tính linh hoạt khi triển khai thực tế. Nhờ những ưu điểm trên, Isaac Sim trở thành lựa chọn hàng đầu để xây dựng pipeline sinh dữ liệu tổng hợp, đáp ứng cả yêu cầu về chất lượng hình ảnh, tốc độ và khả năng mở rộng.

1. Công nghệ nhận diện, bám bắt đối tượng bằng camera trên drone: Công nghệ này là thành phần then chốt trong các hệ thống tác chiến khong người lái, giám sát, cứu hộ khẩn cấp và theo dõi mục tiêu thời gian thực. Mục tiêu của công nghệ này là phát hiện , phân loại và theo dõi chính xác các đối tượng chuyển động trong môi trường phức tạp, bao gồm cả ban đề hoặc điều kiện thời tiết khắc nghiệt. Việc làm chủ công nghệ lõi này không chỉ giúp tăng tính tự chủ và phản ứng thông minh cho drone mà còn đảm bảo tính bảo mật, tránh phụ thuộc vào giải pháp thương mại nước ngoài. Ngoài ra, đây còn là nền tảng để phát triển các hệ thống điều phối đa drone theo bầy đàn, tăng khả năng bao phủ và độ tin cậy trong các nhiệm vụ giám sát chiến lược.

Qua đó có thể thấy rằng việc phát triển, hoàn thiện các công nghệ lõi điều khiển và nền tảng chi huy thông minh các phương tiện không người lái trên bộ và trên không không chỉ là bước tiến quan trọng trong việc hoàn thiện hệ sinh thái Robot, mà còn giúp tạo ra những sản phẩm đột phá, đáp ứng nhu cầu thực tiễn và mở rộng thị trường. Các công nghệ này không chỉ đáp ứng nhu cầu cấp thiết của thị trường mà còn mở ra cơ hội lớn để dẫn đầu trong lĩnh vực Robotics, mang lại giá trị vượt trội.

1. **Tính cấp thiết về nhu cầu thị trường**

Trong bối cảnh kinh tế tầm thấp (Low-Altitude Economy) đang hình thành vè phát triển mạnh mẽ trên phạm vi toàn cầu, nhu cầu về các nền tảng công nghệ lõi phục vụ cho phương tiện không người lái (UxV) ngày càng trở lên cấp bách. Các ứng dụng dân sự và quốc phòng như giám sát, tuần tra, tìm kiếm cứu nạn, cứu hộ thiên tai, ... đều đặt ra yêu cầu rất cao về khả năng tự hành, định vị chính xác, điều hướng an toàn và vận hành ổn định trong những điều kiện địa hình, thời tiết phức tạp.

Tuy nhiên, thực tiễn triển khai các hệ thống UxV hiện nay đang gặp nhiều hạn chế do phụ thuộc lớn vào các giải pháp điều khiển và nền tảng chỉ huy nước ngoài, vốn thường đóng gói, khó tùy biến và không đáp ứng được các yêu cầu đặc thù như địa hình đồi núi, khu vực mất tín hiệu GNSS, hoặc các nhiệm vụ phối hợp đa phương tiện. Do đó, nhu cầu cấp thiết đặt ra là phải làm chủ các công nghệ lõi điều khiển và nền tảng chỉ huy thông minh mang tính độc lập, linh hoạt và có thể tích hợp sâu vào các hệ thống Robot mặt đất và Drone hiện có.

Các ứng dụng chính của Drone tại Việt Nam: Công tác thăm dò, lập bản đồ, giám sát công trường/công trình; Ứng dụng trong nông nghiệp hiện đại sử dụng thiết bị bay không người lái để giúp người nông dân theo dõi mùa vụ, đánh giá tình trạng đất đai và tối ưu hóa thủy lợi. Bên cạnh drone, thị trường robot tự hành mặt đất ngoài trời (UGV – Unmanned Ground Vehicle) cũng đang chứng kiến sự tăng trưởng mạnh mẽ, đặc biệt trong các lĩnh vực quân sự, an ninh và công nghiệp.

Với bối cảnh toàn cầu đang chuyển đổi mạnh mẽ dưới tác động của công nghệ không người lái (UAV), nhu cầu đối với các hệ thống chỉ huy – điều khiển UAV tích hợp khả năng phân tích trí tuệ nhân tạo (AI-driven) đang trở thành một yếu tố thiết yếu. Các nền tảng này không chỉ cung cấp chức năng điều phối – giám sát công việc của UAV, mà còn hỗ trợ tổng hợp và xử lý dữ liệu bay, hình ảnh, cảm biến, và bản đồ số trong thời gian thực. Điều này đáp ứng yêu cầu nghiêm ngặt về ra quyết định nhanh, chính xác trong cả môi trường quốc phòng, an ninh và dân sự.

Ứng dụng của các hệ thống này trải dài từ giám sát biên giới, trinh sát chiến thuật, phòng cháy chữa cháy, cứu hộ, đến quản lý đám đông, giám sát đô thị, giao thông thông minh và quản lý hạ tầng cơ sở. Tại những quốc gia đang phát triển như Việt Nam, các giải pháp này không chỉ phải đảm bảo khả năng tùy biến cao theo đặc thù địa lý, điều kiện hạ tầng mà còn cần đảm bảo yếu tố bảo mật và tích hợp dễ dàng vào hệ thống thông tin liên lạc quốc gia.

Tóm lại, thị trường dành cho nền tảng chỉ huy và công nghệ lõi điều khiển cho các phương tiện không người lái đã sẵn sàng để bùng nổ, ứng dụng mạnh mẽ tại Việt Nam. Trong thực tế, VAI đã nhận được văn bản của Bộ tư lệnh Bộ đội Biên phòng với nhu cầu trang bị, điều khiển 450 drone trinh sát cỡ nhỏ phục vụ giám sát vùng biên.

1. **Liệt kê danh mục các công trình nghiên cứu, tài liệu có liên quan đến đề tài đã trích dẫn khi đánh giá tổng quan**
2. A. I. Mourikis and S. I. Roumeliotis, "A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation," Proceedings 2007 IEEE International Conference on Robotics and Automation, Rome, Italy, 2007, pp. 3565-3572, doi: 10.1109/ROBOT.2007.364024.
3. C. Jia, Y. Cao, J. Yang, Y. Rao, H. Fan and W. Yao, "Improving Visual- Inertial Odometry with Robust Outlier Rejection and Loop Closure," 2020 11th International Conference on Awareness Science and Technology (iCAST), Qingdao, China, 2020, pp. 1-6, doi: 10.1109/iCAST51195.2020.9319474.
4. Y. K. Yu, K. H. Wong, M. M. Y. Chang and S. H. Or, "Recursive Camera-Motion Estimation With the Trifocal Tensor," in IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 36, no. 5, pp. 1081-1090, Oct. 2006, doi: 10.1109/TSMCB.2006.874133.
5. Y. Fan, T. Zhao, and G. Wang, “SchurVINS: Schur Complement-Based Lightweight Visual Inertial Navigation System,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2024, pp. 17964–17973.
6. S. Agarwal, K. Mierle, and Others, “Ceres solver,” http://ceres-solver.org
7. T. Lupton and S. Sukkarieh, "Visual-Inertial-Aided Navigation for High-Dynamic Motion in Built Environments Without Initial Conditions," in IEEE Transactions on Robotics, vol. 28, no. 1, pp. 61-76, Feb. 2012
8. J. S. T. C. D. Engel, “LSD-SLAM: Large-Scale Direct Monocular SLAM,” trong ECCV 2014 — European Conference on Computer Vision, Cham, 2014.
9. J. a. K. V. a. C. D. Engel, “Direct Sparse Odometry,” IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 611-625, 2018.
10. C. a. P. M. a. S. D. Forster, “SVO: Fast semi-direct monocular visual odometry,” trong 2014 IEEE International Conference on Robotics and Automation (ICRA), Hong Kong, 2014.
11. R. a. M. J. M. M. a. T. J. D. Mur-Artal, “ORB-SLAM: A Versatile and Accurate Monocular SLAM System,” IEEE Transactions on Robotics, pp. 1147-1163, 2015.
12. R. a. T. J. D. Mur-Artal, “ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras,” IEEE Transactions on Robotics, pp. 1255-1262, 2017.
13. C. a. E. R. a. R. J. J. G. a. M. M. J. M. a. D. T. J. Campos, “ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual–Inertial, and Multimap SLAM,” IEEE Transactions on Robotics, pp. 1874-1890, 2021.
14. T. Q. a. J. P. a. S. C. a. S. Shen, “A General Optimization-based Framework for Local Odometry Estimation with Multiple Sensors,” arXiv:1901.03638, 2019.
15. T. Q. a. S. C. a. J. P. a. S. Shen, “A General Optimization-based Framework for Global Pose Estimation with Multiple Sensors,” arXiv:1901.03642, 2019.
16. A. a. A. M. a. C. Y. a. C. L. Rosinol, “Kimera: an Open-Source Library for Real-Time Metric-Semantic Localization and Mapping,” trong 2020 IEEE International Conference on Robotics and Automation (ICRA, Paris, France, 2020.
17. B. a. F. J. M. a. C. J. a. N. J. Bescos, “DynaSLAM: Tracking, Mapping, and Inpainting in Dynamic Scenes,” IEEE Robotics and Automation Letters, pp. 4076-4083, 2018.
18. R. K. a. J. S. a. X. L. a. Y. L. a. X. Liu, “DF-SLAM: A Deep-Learning Enhanced Visual SLAM System based on Deep Local Features,” CoRR, p. 1901.07223, 2019.
19. Z. Z. a. C. W. a. X. K. a. Z. L. a. X. D. a. Q. Li, “Light-SLAM: A Robust Deep-Learning Visual SLAM System Based on LightGlue under Challenging Lighting Conditions,” arXiv, p. 2407.02382, 2024.
20. Q. a. Z. X. a. D. R. a. X. Z. Peng, “Robust Visual SLAM for Weakly Textured Environments,” RWT-SLAM: Robust Visual SLAM for Weakly Textured Environments, pp. {913-919, 2024.
21. H. M. a. J. H. a. K. E. Cho, “SP-SLAM: Surfel-Point Simultaneous Localization and Mapping,” IEEE/ASME Transactions on Mechatronics, pp. 2568-2579, 2022.
22. C. W. a. H. S. a. Z. C. a. M. P. a. J. Lim, “OmniSLAM: Omnidirectional Localization and Dense Mapping for Wide-baseline Multi-camera Systems,” arXiv, p. 2003.08056, 2020.
23. A. J. a. H. Z. S. a. D. K. Ben Ali, “Edge-SLAM: Edge-Assisted Visual Simultaneous Localization and Mapping,” trong Proceedings of the 18th International Conference on Mobile Systems, Applications, and Services, New York, NY, USA, 2020.
24. Z. a. D. J. Teed, “DROID-SLAM: Deep Visual SLAM for Monocular, Stereo, and RGB-D Cameras,” Advances in Neural Information Processing Systems, pp. 16558--16569, 2021.
25. Z. a. P. S. a. L. V. a. X. W. a. B. H. a. C. Z. a. O. M. R. a. P. M. Zhu, “NICE-SLAM: Neural Implicit Scalable Encoding for SLAM,” trong Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022.
26. W. a. H. Y. a. S. S. Wang, “TartanVO: A Generalizable Learning-based VO,” trong Proceedings of the 2020 Conference on Robot Learning, 2021.
27. Li, Xinghui, et al. "Dual-resolution correspondence networks." Advances in Neural Information Processing Systems 33 (2020): 17346-17357.
28. Arandjelovic, R., Gronat, P., Torii, A., Pajdla, T., & Sivic, J. (2016). NetVLAD: CNN architecture for weakly supervised place recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 5297-5307).
29. Dusmanu, M., Rocco, I., Pajdla, T., Pollefeys, M., Sivic, J., Torii, A., & Sattler, T. (2019). D2-net: A trainable cnn for joint detection and description of local features. arXiv preprint arXiv:1905.03561.
30. Pogorzelski, T., & Zielinska, T. (2025). Comparative Studies of Descriptor-based Image Matching Techniques for UAV Applications. IEEE Access.
31. S. Chen, X. Wu, M. W. Mueller, and K. Sreenath, “Real-time geo- localization using satellite imagery and topography for unmanned aerial vehicles,” in Proc. 2021 IEEE/RSJ Int. Conf. Intell. Robots Syst., Sep. 2021, pp. 2275–2281
32. H. Hou et al., “UAV pose estimation in GNSS-denied environment as- sisted by satellite imagery deep learning features,” IEEE Access, vol. 9, pp. 6358–6367, 2021
33. J. Sun, Z. Shen, Y. Wang, H. Bao, and X. Zhou, “LoFTR: Detector-free local feature matching with transformers,” in Proc. 2021 IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Nashville, TN, USA, 2021, pp. 8918–8927
34. J. Yu, J. Chang, J. He, T. Zhang, J. Yu, and F. Wu, “Adaptive spot- guided transformer for consistent local feature matching,” in Proc. 2023 IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Jun. 2023, pp. 21898–21908
35. S. Zhu, M. Shah, and C. Chen, “TransGeo: Transformer is all you need for cross-view image geo-localization,” in Proc. 2022 IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2022, pp. 1162–1171
36. Zhang, Z., Chu, J., Song, T., Guo, J., Zhang, R., & Li, J. (2025). GeoRVLF: A Robust Drone-Satellite Visual Geo-Localization Framework for Small UAV Platforms. IEEE Robotics and Automation Letters.
37. P. - E. Sarlin, D. DeTone, T. Malisiewicz, and A. Rabinovich, “SuperGlue: Learning feature matching with graph neural networks,” in Proc. 2020 IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Jun. 2020, pp. 4937–4946.
38. C. Chen, M. He, J. Wang, and Z. Meng, “GeoCluster: Enhancing visual place recognition in spatial domain on aerial vehicle platforms,” IEEE Robot. Automat. Lett., vol. 9, no. 3, pp. 3013–3020, Mar. 2024.
39. L. Li et al., “Geo-localization with transformer-based 2D-3D match network,” IEEE Robot. Automat. Lett., vol. 8, no. 8, pp. 4855–4862, Aug. 2023.
40. P. Lindenberger, P. - E. Sarlin, and M. Pollefeys, “LightGlue: Local feature matching at light speed,” in Proc. 2023 IEEE/CVF Int. Conf. Comput. Vis., Oct. 2023, pp. 17581–17592.
41. Y. Wang, X. He, S. Peng, D. Tan, and X. Zhou, “Efficient LoFTR: Semi-dense local feature matching with sparse-like speed,” in Proc. 2024 IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2024, pp. 21666–21675.
42. J. Chen, M. Li, Z. Yuan and Q. Gu, "An Improved A\* Algorithm for UAV Path Planning Problems," 2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Chongqing, China, 2020, pp. 958-962, doi: 10.1109/ITNEC48623.2020.9084806.
43. Dhulkefl, Elaf & Durdu, Akif & Terzioğlu, Hakan. (2020). “DIJKSTRA ALGORITHM USING UAV PATH PLANNING”, Konya Journal of Engineering Sciences. 8. 92-105. 10.36306/konjes.822225.
44. Koenig S and Likhachev M (2002) Dˆ\* lite. Aaai/iaai 15
45. S.M.LaValle, "Rapidly-exploring random trees: A new tool for path planning,"  
    1998.
46. L.E. Kavraki, P. Svestka, J.C. Latombe, M.H. and Overmars, "Probabilistic roadmaps for path planning in high-dimensional configuration spaces," IEEE transactions on Robotics and Automation 12(4): 566–580, 1996.
47. Boyu Zhou, Fei Gao, Luqi Wang, Chuhao Liu and Shaojie Shen, “Robust and Efficient Quadrotor Trajectory Generation for Fast Autonomous Flight”, IEEE Robotics and Automation Letters (RA-L), 2019.
48. X. Zhou, Z. Wang, H. Ye, C. Xu and F. Gao, "EGO-Planner: An ESDF-Free Gradient-Based Local Planner for Quadrotors," IEEE Robotics and Automation Letters, vol. 6, no. 2, pp. 478-485, April 2021, doi: 10.1109/LRA.2020.3047728.
49. R. Penicka, D. Scaramuzza, Minimum-Time Quadrotor Waypoint Flight in Cluttered Environments, IEEE Robotics and Automation Letters, 2022.
50. Tordesillas, Jesus & Lopez, Brett & Everett, Michael & How, Jonathan. (2021). FASTER: Fast and Safe Trajectory Planner for Navigation in Unknown Environments. IEEE Transactions on Robotics. PP. 1-17. 10.1109/TRO.2021.3100142.
51. Lu, Guozheng & Ren, Yunfan & Zhu, Fangcheng & Li, Haotian & Xue, Ruize & Cai, Yixi & Ximin, Lyu & Zhang, Fu. (2025). Autonomous Tail-Sitter Flights in Unknown Environments. IEEE Transactions on Robotics. PP. 1-20. 10.1109/TRO.2025.3526102.
52. Z. Wang, X. Zhou, C. Xu and F. Gao, "Geometrically Constrained Trajectory Optimization for Multicopters," in IEEE Transactions on Robotics, vol. 38, no. 5, pp. 3259-3278, Oct. 2022, doi: 10.1109/TRO.2022.3160022.
53. Kaufmann, Elia & Bauersfeld, Leonard & Loquercio, Antonio & Mueller, Matthias & Koltun, Vladlen & Scaramuzza, Davide. (2023). Champion-level drone racing using deep reinforcement learning. Nature. 620. 982-987. 10.1038/s41586-023-06419-4.
54. Xie, Yuhan & Lu, Minghao & Peng, Rui & Lu, Peng. (2023). Learning Agile Flights Through Narrow Gaps with Varying Angles Using Onboard Sensing. IEEE Robotics and Automation Letters. PP. 1-8. 10.1109/LRA.2023.3295655.
55. Ji, Jialin & Pan, Neng & Xu, Chao & Gao, Fei. (2022). Elastic Tracker: A Spatio-temporal Trajectory Planner for Flexible Aerial Tracking. 47-53. 10.1109/ICRA46639.2022.9811688.
56. Lu, Junjie & Hui, Yulin & Zhang, Xuewei & Feng, Wencan & Shen, Hongming & Li, Zhiyu & Tian, Bailing. (2025). YOPOv2-Tracker: An End-to-End Agile Tracking and Navigation Framework from Perception to Action. 10.48550/arXiv.2505.06923.
57. Li, Bo & Yan, Junjie & Wu, Wei & Zheng, Zhu & Hu, Xiaolin. (2018). High Performance Visual Tracking with Siamese Region Proposal Network. 8971-8980. 10.1109/CVPR.2018.00935.
58. Wojke, Nicolai & Bewley, Alex & Paulus, Dietrich. (2017). Simple online and realtime tracking with a deep association metric. 3645-3649. 10.1109/ICIP.2017.8296962.
59. A Survey on Aerial Swarm Robotics - Soon-Jo Chung; Aditya Avinash Paranjape; Philip Dames; Shaojie Shen; Vijay Kumar. IEEE Transactions on Robotics, 2018
60. Swarm-LIO2: Decentralized, Efficient LiDAR-inertial Odometry for UAV Swarms - Fangcheng Zhu, Yunfan Ren, Longji Yin, Fanze Kong, Qingbo Liu, Ruize Xue, Wenyi Liu, Yixi Cai, Guozheng Lu, Haotian Li, Fu Zhang. IEEE Transactions on Robotics, 2025
61. D2SLAM: Decentralized and Distributed Collaborative Visual-Inertial SLAM System for Aerial Swarm - Hao Xu; Peize Liu; Xinyi Chen; Shaojie Shen. IEEE Transactions on Robotics, 2024
62. High-Speed Motion Planning for Aerial Swarms in Unknown and Cluttered Environments - Charbel Toumieh, Dario Floreano. IEEE Transactions on Robotics, 2024
63. OR-Tools' Vehicle Routing Solver: A Generic Constraint-Programming Solver with Heuristic Search for Routing Problems (VRPs). 2023
64. RACER: Rapid Collaborative Exploration with a Decentralized Multi-UAV System - Boyu Zhou, Hao Xu, Shaojie Shen. Transactions on Robotics, 2023
65. Fast and Communication-Efficient Multi-UAV Exploration Via Voronoi Partition on Dynamic Topological Graph - Qianli Dong, Haobo Xi, Shiyong Zhang, Qingchen Bi, Tianyi Li, Ziyu Wang, and Xuebo Zhang. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2024
66. Swarm UAVs Communication - Arindam Majee, Rahul Saha, Snehasish Roy, Srilekha Mandal, Sayan Chatterjee. arXiv:2405.00024, 2024
67. <https://helsing.ai/altra>
68. <https://www.l3harris.com/all-capabilities/amorphous>
69. <https://wpassets.porttechnology.org/wp-content/uploads/2021/02/03095430/63d481_e346f13be23242ea9aef81998a482c70.pdf>
70. <https://votix.com/>
71. <https://www.dronesense.com/>
72. <https://www.motorolasolutions.com/en_us/video-security-access-control/drones/cape-drone-software.html>
73. <https://uavia.eu/en/>
74. <https://www.highlander.io/>
75. T. Guan, M. R. Bhuyan, H. Dong, and M. G. Jadidi, “GANav: Efficient Terrain Segmentation for Robot Navigation,” Sensors, vol. 21, no. 6, p. 2055, Mar. 2021. [Online].
76. C. Reed, D. C. Conner, and A. Bobick, “Autonomous Hiking Trail Navigation via Semantic Segmentation and Geometric Analysis,” arXiv preprint, arXiv:2409.15671, 2024. [Online].
77. J. Liang, M. Liu, and J. Wang, “MTG: Mapless Trajectory Generator with Traversability Coverage for Outdoor Navigation,” arXiv preprint, arXiv:2309.08214, 2023. [Online].
78. S. Jung, Y. Seo, J. Ryu, and Y. Choi, “V-STRONG: Visual Self-Supervised Traversability Learning for Off-Road Navigation,” arXiv preprint, arXiv:2312.16016, 2023. [Online].
79. M. Gong et al., “TAIL: A Terrain-Aware Multi-Modal Dataset for SLAM in Complex Outdoor Environments,” arXiv preprint, arXiv:2403.16875, 2024. [Online].
80. X. Zhu et al., “General Quadruped Robot Navigation through Complex Terrains via Vision-Language Models,” arXiv preprint, arXiv:2407.16412, 2024. [Online].
81. J. R. Gasparino, F. Bonin-Font, and J. M. Azorin-Lopez, “WayFAST: A Framework for Learning and Prediction of Traversability from Vision and Terrain,” Autonomous Robots, vol. 47, pp. 1237–1254, 2023. [Online].
82. Y. Seo, S. Jung, J. Ryu, and Y. Choi, “SCATE: Self-supervised Costmap Adaptation for Traversability Estimation,” arXiv preprint, arXiv:2306.02511, 2023. [Online].
83. <https://arxiv.org/abs/1709.06670>
84. <https://www.zora.uzh.ch/id/eprint/257679/1/2003.05654.pdf>
85. <https://arxiv.org/abs/1712.05474>
86. <https://gymnasium.farama.org/index.html>
87. <https://mujoco.readthedocs.io/en/stable/overview.html>
88. <https://pybullet.org/wordpress/index.php/forum-2/>
89. <https://microsoft.github.io/AirSim/>
90. <https://developer.nvidia.com/isaac/sim>
91. <https://genesis-world.readthedocs.io/en-us/latest/>
92. <https://ai2thor.allenai.org/>
93. <https://docs.ultralytics.com/models/yolov8/>
94. [77]N. Wojke, A. Bewley, and D. Paulus, Simple Online and Realtime Tracking with a Deep Association Metric. 2017. [Online]. Available: <https://arxiv.org/abs/1703.07402>
95. Y. Zhang et al., ByteTrack: Multi-Object Tracking by Associating Every Detection Box. 2022. [Online]. Available: <https://arxiv.org/abs/2110.06864>
96. Shuai Liu, Xin Li, Huchuan Lu, You He, Multi-Object Tracking Meets Moving UAV Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 8876-8885
97. Nguyen, T. S., Quach, C. H., Dang, T. H. G., Vu, M. T., Vuong, Q. H., & Mai, A. T. (2021). Developing system of wireless sensor network and unmaned aerial vehicle for agriculture inspection. arXiv. <https://arxiv.org/abs/2107.01008>
98. Nguyen, L. Q., Dang, M. T., Bui, L. K., Bui, Q. N., & Tran, T. X. (2023). Application of unmanned aerial vehicles for surveying and mapping in mines: A review. In L. Q. Nguyen, L. K. Bui, X.-N. Bui, & H. T. Tran (Eds.), Advances in geospatial technology in mining and earth sciences (pp. 1–22). Springer International Publishing. <https://doi.org/10.1007/978-3-031-20463-0_1>
99. Nguyen Vu, G., Vu, L., Vu, C., Tong, S., Vu Huu, L., Lê Quang, T., Tong, A., Nguyen, H., Nguyen, T., Pham, H., & Nguyen, B. (2017, October). UAV photogrammetry for 3D mapping – A case study in Vietnam. In Proceedings of the 38th Asian Conference on Remote Sensing (ACRS), New Delhi, India.
100. CHO, Hyun-Min, et al. Assessment of the GNSS-RTK for application in precision forest operations. Remote Sensing, 2023, 16.1: 148.
101. LIU, Chongxiao; NGUYEN, Bao Kha. Low-Cost Real-Time Localisation for Agricultural Robots in Unstructured Farm Environments. Machines, 2024, 12.9:612
102. SHAN, Tixiao, et al. Lio-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping. In: 2020 IEEE/RSJ international conference on intelligent robots and systems (IROS). IEEE, 2020. p. 5135-5142
103. MOELLER, Ryan; DEEMYAD, Taher; SEBASTIAN, Anish. Autonomous navigation of an agricultural robot using RTK GPS and Pixhawk. In: 2020 Intermountain Engineering, Technology and Computing (IETC). IEEE, 2020. p.1-6
104. SHAN, Tixiao; ENGLOT, Brendan. Lego-loam: Lightweight and ground-optimized lidar odometry and mapping on variable terrain. In: 2018 IEEE/RSJ international conference on intelligent robots and systems (IROS). IEEE, 2018. p. 4758-4765.
105. XU, Wei, et al. Fast-lio2: Fast direct lidar-inertial odometry. IEEE Transactions on Robotics, 2022, 38.4: 2053-2073.
106. BAI, Chunge, et al. Faster-LIO: Lightweight tightly coupled LiDAR-inertial odometry using parallel sparse incremental voxels. IEEE Robotics and Automation Letters, 2022, 7.2: 4861-4868
107. RAMEZANI, Milad, et al. Wildcat: Online continuous-time 3d lidar-inertial slam. arXiv preprint arXiv:2205.12595, 2022.
108. CUI, Yutao, et al. Sportsmot: A large multi-object tracking dataset in multiple sports scenes. In: Proceedings of the IEEE/CVF international conference on computer vision. 2023. p. 9921-9931.
109. MACENSKI, Steve; JAMBRECIC, Ivona. SLAM Toolbox: SLAM for the dynamic world. Journal of Open Source Software, 2021, 6.61: 2783
110. HESS, Wolfgang, et al. Real-time loop closure in 2D LIDAR SLAM. In: 2016 IEEE international conference on robotics and automation (ICRA). IEEE, 2016. p. 1271-1278
111. <https://projects.csail.mit.edu/stata/>
112. GAO, Boyu; LANG, Haoxiang; REN, Jing. Stereo visual SLAM for autonomous vehicles: A review. In: 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, 2020. p. 1316-1322.
113. Intel Corporation, Intel RealSense Tracking Camera T265 – VIO integration with PX4. Accessed: Jul. 29, 2025. [Online]. Available: https://docs.px4.io/main/en/camera/camera\_intel\_realsense\_t265\_vio
114. DJI, Mavic Series. Accessed: Jul. 29, 2025. [Online]. Available: https://dji-vietnam.vn/mavic-series/
115. DJI, “Matrice 4 Series”. Accessed: Jul. 29, 2025. [Online]. Available: https://enterprise.dji.com/matrice‐4‐series
116. NASA, “Ingenuity Mars Helicopter”. Accessed: Jul. 29, 2025. [Online]. Available: https://science.nasa.gov/mission/mars-2020-perseverance/ingenuity-mars-helicopter/
117. Skydio, “Skydio 2+ Enterprise”. Accessed: Jul. 29, 2025. [Online]. Available: https://www.skydio.com/skydio-2-plus-enterprise
118. Autel Robotics, “EVO Max Series”. Accessed: Jul. 29, 2025. [Online]. Available: https://www.autelrobotics.com/productdetail/evo-max-series-old/
119. SLAMcore, “SLAMcore SDK”. Accessed: Jul. 29, 2025. [Online]. Available: https://www.slamcore.com/products/slamcore-sdk/
120. Theseus, “Theseus”. Accessed: Jul. 29, 2025. [Online]. Available: https://www.theseus.us/
121. OKSI, “OMNInav GPS‐Denied Navigation”. Accessed: Jul. 29, 2025. [Online]. Available: <https://oksi.ai/omninav-gps-denied-navigation/>
122. https://px4.io/
123. <https://ardupilot.org/>
124. <https://betaflight.com>

**17. Nội dung công việc và phương án thực hiện đề tài, công nghệ lõi của đề tài, tỷ lệ làm chủ sản phẩm, mô tả nền tảng sản phẩm**

**17.1. Các nội dung công việc và phương án thực hiện của đề tài**

Đề tài hướng đến mục tiêu phát triển thành công “Nền tảng chỉ huy thông minh các phương tiện không người lái trên bộ và trên không” bao gồm ba thành phần chính:

* Phương tiện không người lái (UxV): Bao gồm các phương tiện bay không người lái trên không và trên bộ. Trong phạm vi đề tài, nhóm đề tài hướng tới phát triển drone dạng đa rotor và robot tự hành mặt đất (UGV).
* Hệ thống điều khiển mặt đất: Bao gồm trạm điều khiển mặt đất (GCS) và trạm cất, hạ cánh thông minh cho drone. Hệ thống điều khiển mặt đất giao tiếp 2 chiều với các UxV thông qua kết nối 5G hoặc RF để nhận dữ liệu và gửi lệnh điều khiển.
* Nền tảng chỉ huy thông minh: Là thành phần trung tâm kết nối các UxV, hệ thống điều khiển mặt đất tới trung tâm điều hành và các hệ thống nghiệp vụ khác của bên thứ ba.



*Hình 17.1. Sơ đồ tổng quan nền tảng chỉ huy thông minh các phương tiện không người lái trên bộ và trên không*

***Nền tảng chỉ huy thông minh gồm các tính năng:***

* Tính năng quản lý, giám sát, hiển thị tình huống, hiển thị bản đồ (2D, 3D, bản đồ vệ tinh) khu vực tác chiến, vị trí phương tiện không người lái, vị trí các đối tượng, mục tiêu.
* Truyền dữ liệu (video, hình ảnh, các dữ liệu từ cảm biến khác) thời gian thực về hệ thống chỉ huy.
* Tính năng điều khiển, điều phối nhiều phương tiện không người lái.
* Tính năng tự động lập kế hoạch nhiệm vụ cho từng phương tiện.
* Tính năng tích hợp các module xử lý thông minh phân tích các dữ liệu được thu thập và truyền về bởi drone và robot.
* Tính năng hỗ trợ các giao thức phổ biến để điều khiển robot và drone.
* Tính năng lưu log, thống kê và xuất báo cáo.

***Drone bao gồm các tính năng:***

* Tính năng định vị, dẫn đường trong môi trường không có tín hiệu định vị toàn cầu GNSS.
* Tính năng ước lượng vị trí đối tượng trong không gian.
* Tính năng phát hiện, nhận dạng, bắt bám mục tiêu, tự động tránh va chạm

***Robot tự hành mặt đất (UGV) bao gồm các tính năng:***

* Tính năng định vị, dẫn đường sử dụng tín hiệu GNSS-RTK.
* Tính năng phát hiện, nhận dạng, tự động tránh vật cản.

**Nội dung 1:** **Nghiên cứu công nghệ AI thị giác máy tính trong định vị drone**

***Công việc 1.1:******Tiền xứ lý dữ liệu camera và IMU***

* Quá trình tiền xử lý bao gồm các bước nhằm đảm bảo chất lượng và độ chính xác của dữ liệu đầu vào từ hệ thống camera.
* Trước tiên, cần thực hiện đồng bộ tín hiệu giữa các camera trong hệ thống multi-camera để đảm bảo các khung hình được ghi nhận cùng thời điểm, phục vụ cho các thuật toán thị giác máy tính yêu cầu độ chính xác cao về thời gian.
* Tiếp theo là hiệu chỉnh các tham số nội tại của từng camera và loại bỏ biến dạng của hình ảnh do ống kính gây ra, sau đó hiệu chỉnh vị trí tương đối giữa các camera, nhằm chuẩn hóa thông tin hình ảnh, giảm thiểu sai số hình học và đảm bảo tính nhất quán giữa các nguồn dữ liệu.
* Bên cạnh đó, việc hiệu chuẩn IMU bao gồm loại bỏ nhiễu và tìm vị trí tương đối giữa imu và hệ thống multi- camera cho phép tích hợp dữ liệu IMU với hệ thống camera giúp nâng cao định vị và nhận biết chuyển động của camera.

***Công việc 1.2: Nghiên cứu và đánh giá hệ thống Visual SLAM***

* Thử nghiệm và cải tiến các thuật toán Visual SLAM trên các môi trường khác nhau, bao gồm cả môi trường thực tế và mô phỏng, là bước quan trọng nhằm đánh giá hiệu năng, độ ổn định và khả năng thích ứng của hệ thống trong các điều kiện hoạt động đa dạng.
* Thông qua các kịch bản mô phỏng có kiểm soát, thuật toán có thể được kiểm tra trong các tình huống đặc biệt như thay đổi ánh sáng đột ngột, hoặc chuyển động phức tạp. Trong khi đó, thử nghiệm thực tế ngoài hiện trường giúp xác thực khả năng hoạt động của thuật toán trong điều kiện môi trường thực, với các yếu tố nhiễu động và bất định không thể mô phỏng đầy đủ.
* Dựa trên kết quả thu được, tiến hành phân tích hiệu suất theo các chỉ tiêu như độ chính xác định vị, độ trôi (drift), tốc độ xử lý và khả năng tái lập bản đồ. Từ đó, có thể đề xuất các phương án hiệu chỉnh và cải tiến để nâng cao độ tin cậy, khả năng thời gian thực và độ bền vững khi triển khai trên các nền tảng thiết bị bay không người lái. Quá trình lặp lại giữa thử nghiệm và cải tiến đóng vai trò then chốt trong việc hoàn thiện hệ thống, hướng đến một giải pháp SLAM hiệu quả, ổn định và có tính ứng dụng cao trong các điều kiện hoạt động phức tạp.

***Công việc 1.3: Nghiên cứu và đánh giá hệ thống VIO (Visual Inertial Odometry)***

* Trong công việc này, nhóm sẽ tập trung khảo sát, triển khai thử nghiệm và đánh giá các giải pháp Visual-Inertial Odometry tiêu biểu, nhằm xác định hướng tiếp cận tối ưu cho hệ thống định vị trên drone trong điều kiện môi trường tín hiệu GNSS bị hạn chế hoặc không ổn định. Các giải pháp chúng tôi dự kiến lựa chọn tham khảo đại diện cho nhiều hướng tiếp cận khác nhau (tối ưu hóa, hợp nhất dữ liệu và tích hợp các giải pháp học sâu), giúp cung cấp cái nhìn toàn diện về hiệu năng, độ chính xác, độ ổn định và khả năng mở rộng. Cụ thể, nhóm đề tài dự kiến xem xét và thử nghiệm các giải pháp nổi bật sau:
* SVO2 (Semi-Direct Visual Odometry 2): Giải pháp bán trực tiếp, kết hợp ưu điểm của phương pháp trực tiếp (direct) và phương pháp dựa trên đặc trưng (feature-based). SVO2 không trích xuất và theo dõi tất cả các đặc trưng, mà chỉ sử dụng tập các điểm ảnh có giá trị thông tin cao. Điều này giúp giảm đáng kể chi phi tính toán và tăng tốc độ xử lý, phù hợp với nền tảng nhúng và drone có kích thước nhỏ.
* VINS-Mono/ VINS-Fusion: Đây là hệ thống tối ưu hóa dựa trên cửa sổ trượt (sliding-window optimization) với kỹ thuật IMU preintegration, có khả năng hoạt động trong môi trường phức tạp. Điểm mạnh của hai thuật toán này là khả năng khởi tạo nhanh, tái định vị và đóng vòng lặp để hạn chế trôi tích lũy. Nhóm đề tài sẽ tham khảo kiến trúc tối ưu hóa, tích hợp cơ chế khởi tạo tự động cho nhiều kịch bản bay khác nhau hoặc mở rộng cơ chế hợp nhất dữ liệu GNSS/ LiDAR đồng bộ để tăng độ tin cậy trong điều kiện ngoài trời và bay xa.
* Basalt: Hệ thống VIO tiên tiến, tập trung vào độ chính xác và tính ổn định lâu dài. Nó sử dụng kỹ thuật phục hồi phi tuyến để tái cấu trúc thông tin từ VIO thành các factor phục vụ cho tối ưu hóa toàn cục. Đây cũng là kỹ thuật mà nhóm đề tài có thể tham khảo và phát triển trên hệ thống định vị cho drone nhằm tối ưu hóa sử dụng bộ nhớ và CPU để phù hợp với yêu cầu tính toán hạn chế của drone.
* ROVIO (Robust Visual-Inertial Odometry): Áp dụng bộ lọc IEKF với cơ chế cập nhật quang học dựa trên các patch ảnh. Cách tiếp cận này không cần trích xuất đặc trưng rõ ràng mà tận dụng trực tiếp thông tin cường độ ảnh, giúp giảm độ trễ và tăng độ bền vững khi số lượng đặc trưng giảm mạnh. Ở hướng tiếp cận này, nhóm đề tài dự kiến tham khảo công nghệ xử lý ngoại lai (outlier), tinh chỉnh bộ tham số lọc để đảm bảo tracking ổn định ngay cả khi drone bay trong môi trường ánh sáng yếu hoặc ít đặc trưng.
* Learning-enhanced VIO: Đây là hướng tiếp cận có sử dụng các mô hình học máy, học sâu, bao gồm các giải pháp tiêu biểu như RAFT-VIO, LF-VIO hoặc DeepVIO. Hướng tiếp cận này giúp cải thiện khả năng theo dõi trong môi trường ít chi tiết hoặc khi hình ảnh bị mờ do rung lắc. Nhóm đề tài dự định thay thế mô-đun so khớp truyền thống bằng optical flow học sâu để cải thiện liên kết đặc trưng khi chất lượng ảnh suy giảm, duy trì phần tối ưu hóa truyền thống để đảm bảo tính minh bạch.
* Nhóm đề tài sẽ tiến hành triển khai, thử nghiệm các thuật toán đã nêu trên trên cùng một bộ dữ liệu chuẩn như EuRoC, TUM-VI và dữ liệu từ drone thử nghiệm thực tế. Tiến hành đánh giá ưu, nhược điểm của từng thuật toán, sau đó lựa chọn hướng tiếp cận phù hợp và phát triển nó để đáp ứng yêu cầu đặt ra của đề tài.
* Sau khi đã lựa chọn được phương pháp, nhóm đề tài tiếp tục kiểm tra hiệu năng của thuật toán qua các kịch bản mô phỏng và thực nghiệm, bao gồm điều kiện thay đổi ánh sáng, rung động cơ học, và chuyển động phức tạp để đánh giá khả năng ổn định và chống trôi (drift).
* Phân tích kết quả theo các tiêu chí như độ chính xác quỹ đạo, tốc độ xử lý, khả năng làm việc thời gian thực và độ tin cậy khi bay BVLOS; đồng thời đề xuất các cải tiến thích hợp.

***Công việc 1.4: Nghiên cứu và đánh giá hệ thống Image Geo‑Localization***

* Thu thập và xây dựng bộ dữ liệu: bộ dữ liệu phục vụ nghiên cứu được xây dựng từ hai nguồn chính là ảnh vệ tinh (satellite imagery) và ảnh thực thu từ drone hoặc camera mặt đất. Tất cả ảnh được gắn nhãn vị trí địa lý (geo-tagging) với độ chính xác cao, đảm bảo tính đồng bộ về thời gian và không gian. Dữ liệu sẽ được tiền xử lý nhằm chuẩn hóa về độ phân giải, định dạng, phối cảnh, cân bằng sáng và loại bỏ nhiễu, giúp giảm sai số trong quá trình so khớp. Ngoài ra để đảm bảo tính đa dạng và khả năng khái quát của mô hình, dữ liệu được thu thập trong nhiều điều kiện môi trường khác nhau, bao gồm cả sự thay đổi về thời thiết, ánh sáng, mùa vụ và góc nhìn, tạo điều kiện cho hệ thống học được các đặc trưng thị giác ổn định và bền vững.
* Công việc tiếp theo sẽ là tập trung phát triển thuật toán định vị UAV geo-localization. Hướng phát triển thuật toán mà nhóm đề tài dự kiến xây dựng tận dụng sức mạnh của Transformer và cơ chế Attention để giải quyết bài toán so khớp ảnh từ camera UAV với bản đồ vệ tinh độ phân giải cao. Giải pháp đề xuất bao gồm kiến trúc hai nhánh xử lý song song: nhánh thứ nhất tiếp nhận ảnh từ UAV (góc nhìn phối cảnh) và nhánh thứ hai xử lý ảnh từ bản đồ vệ tinh (góc nhìn trực giao). Mỗi nhánh sử dụng backbone dạng Vision Transformer (ViT) hoặc Swin Transformer để trích xuất đặc trưng không gian ngữ nghĩa, đồng thời áp dụng Self-Attention nhằm nắm bắt mối quan hê toàn cục giữa các vùng ảnh. Để giảm khoảng cách miền (domain gap) giữa ảnh UAV chụp và ảnh từ bản đồ vệ tinh, nhóm dự kiến tích hợp cơ chế Cross-Attention, cho phép đặc trưng ảnh từ UAV “tìm kiếm” và “liên kết” trực tiếp với đặc trưng tương ứng trên ảnh vệ tinh. Thông tin đa tỉ lệ được khai thác thông qua mô-đun feature pyramid, đảm bảo cả chi tiết vi mô và cấu trúc vĩ mô đều được đồng bộ hóa. Nếu có bản đồ vệ tinh độ chính xác cao cấp centimet đóng vai trò làm mốc neo tuyệt đối, cho phép hệ thống học căn chỉnh hình học chính xác. Sau bước so khớp đặc trưng, mô hình thực hiện định vị tinh qua các kỹ thuật alignment không gian và có thể bổ sung dữ liệu độ cao để cải thiện ước lượng tọa độ 3D và góc quay của máy bay không người lái. Nhóm đề tài dự kiến tham khảo các nghiên cứu đã được công bố có liên quan như TransGeo, SSPT (Single stream pyramid transformer network), MLPCAN (Multi-layer Local Pattern Cross-Attention Network), ...
* Phân tích hiệu suất dựa trên các tiêu chí như độ chính xác định vị, tỷ lệ nhận dạng đúng (top‑k accuracy) và tốc độ xử lý. Từ đó đề xuất các phương án hiệu chỉnh để nâng cao độ tin cậy, tính thời gian thực và khả năng triển khai trên các nền tảng drone. Thực hiện kiểm thử trên các tập dữ liệu và môi trường thực tế sau mỗi lần cải tiến, phát triển thuật toán, đến khi đạt được đầu ra như mong muốn.

***Công việc 1.5: Nghiên cứu kết hợp đa cảm biến nâng cao hiệu suất định vị***

* Nghiên cứu kết hợp các công nghệ định vị sử dụng camera với các cảm biến khác như RTK-GPS và LiDAR nhằm nâng cao độ chính xác, khả năng định vị tuyệt đối cũng như độ tin cậy của hệ thống trong các môi trường hoạt động phức tạp và đa dạng. RTK-GPS cung cấp dữ liệu định vị tuyệt đối với độ chính xác ở mức centimet, rất hữu ích trong các ứng dụng yêu cầu vị trí chính xác cao. Tuy nhiên, RTK-GPS phụ thuộc mạnh vào tín hiệu vệ tinh và có thể bị mất tín hiệu hoặc nhiễu trong môi trường bị che khuất như đô thị đông đúc, khu vực trong nhà hoặc dưới tán cây rậm rạp. Trong khi đó, LiDAR cho phép thu thập dữ liệu độ sâu và tái tạo cấu trúc không gian 3D chính xác, hoạt động ổn định trong cả điều kiện ánh sáng yếu hoặc thay đổi liên tục. Khi tích hợp với Visual SLAM/VIO hoặc image geo-localization , LiDAR giúp tăng cường khả năng nhận biết không gian, hỗ trợ xây dựng bản đồ chi tiết và cải thiện độ ổn định của định vị trong các môi trường phức tạp, nơi đặc trưng hình ảnh không rõ ràng hoặc dễ bị thay đổi.
* Việc kết hợp các nguồn cảm biến này đòi hỏi thực hiện hiệu chuẩn chéo (cross-calibration) giữa các thiết bị, đồng bộ dữ liệu theo thời gian, và phát triển các thuật toán hợp nhất dữ liệu (sensor fusion) thông minh. Hệ thống tích hợp sẽ tận dụng ưu điểm của từng loại cảm biến: RTK-GPS cung cấp chuẩn mốc vị trí tuyệt đối, LiDAR đảm bảo dữ liệu không gian chính xác, trong khi hệ thống định vị sử dụng camera đảm nhận vai trò dẫn đường và xây dựng bản đồ liên tục. Thông qua quá trình tích hợp và thử nghiệm, hướng tới xây dựng một hệ thống định vị và bản đồ cho drone kết hợp nhiều thông tin có độ chính xác cao, hoạt động ổn định trong thời gian thực.

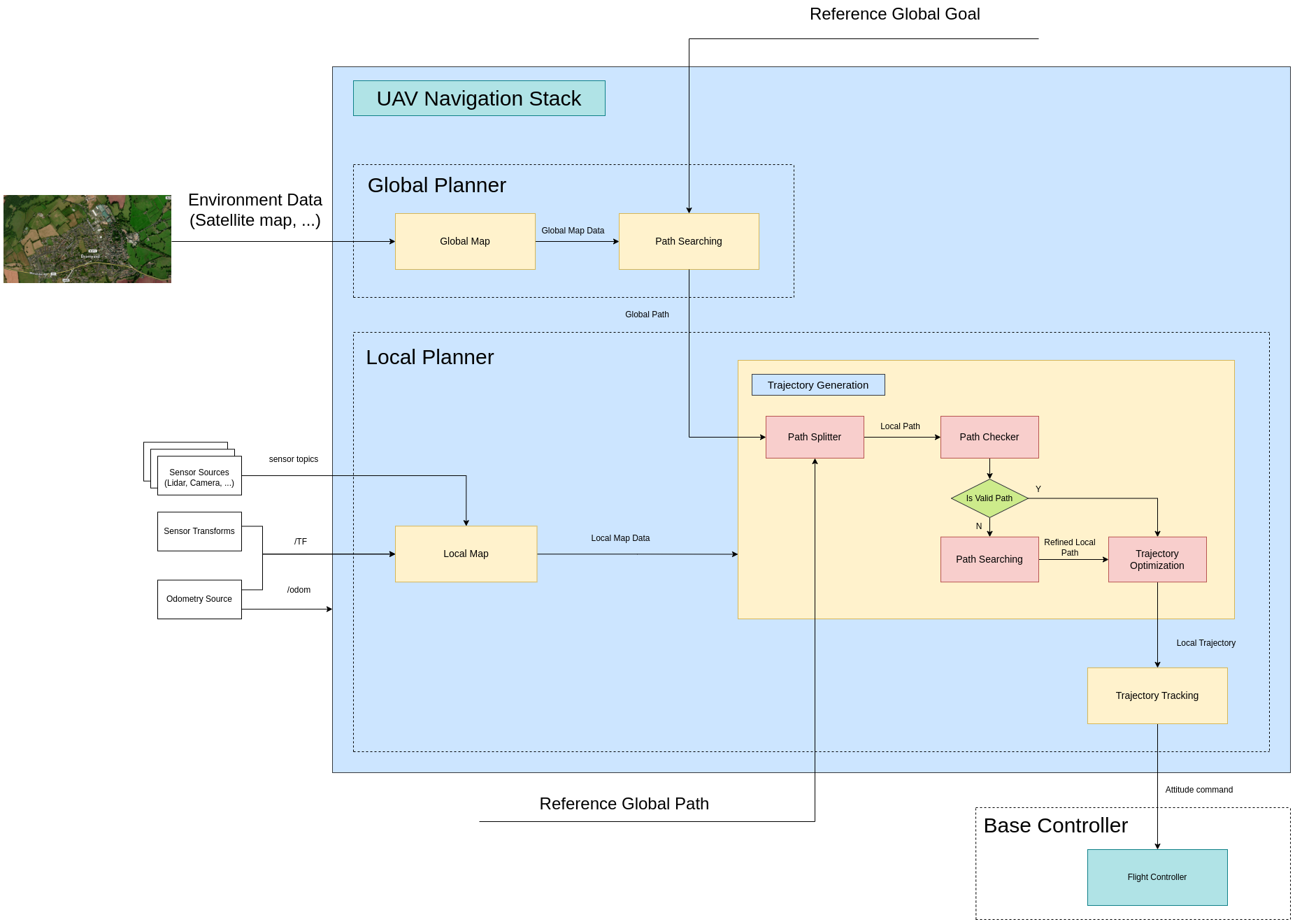
***Công việc 1.6: Tìm hiểu tập dữ liệu mở và xây dựng tiêu chí đánh giá***

* Khảo sát và thu nhập các bộ dữ liệu mã nguồn mở liên quan đến Visual SLAM, VIO và Image Geo-Localization, bao gồm cả dữ liệu ảnh, IMU, có thể bao gồm cả dữ liệu LiDAR và GNSS, để phục vụ nghiên cứu và thử nghiệm thuật toán.
* Xây dựng bộ tiêu chí đánh giá với các chỉ số đo lường chuẩn như độ chính xác định vị (APE/ RPE), mức độ trôi (drift), tỷ lệ nhận dạng đúng trong bài Image Geo-Localization và tốc độ xử lý thời gian thực.
* Đề xuất tiêu chuẩn đánh giá thống nhất cho các thuật toán trong nhiều kịch bản khác nhau, giúp so sánh hiệu năng khách quan và định hướng cải tiến hệ thống định vị trên UAV.

**Nội dung 2: Nghiên cứu công nghệ điều hướng tự động cho UAV**

***Công việc 2.1: Nghiên cứu, xây dựng kiến trúc mô-đun và các thuật toán cho hệ điều hướng tự động ANS***

* Thiết kế và xây dựng kiến trúc phân lớp (modular architecture) cho hệ điều hướng tự động ANS, gồm ba tầng chức năng chính: Lập kế hoạch toàn cục (tuỳ chọn), lập kế hoạch cục bộ, phần mềm điều khiển bay. Hình 17.2 là kiến trúc phần mềm do chúng tôi đề xuất.



*Hình 17.2. Sơ đồ tổng quan hệ thống điều hướng tự động cho UAV đề xuất*

***Công việc 2.2: Nghiên cứu, làm chủ phần mềm điều khiển bay***

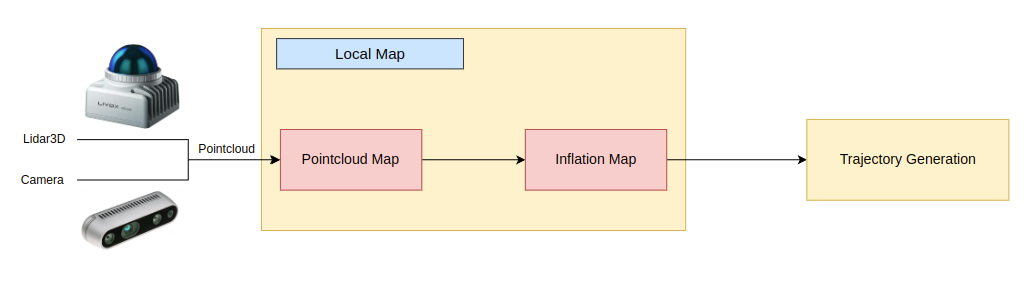
* Nghiên cứu, đánh giá các phần mềm điều khiển bay phổ biến như PX4, Ardupilot, Betaflight. Lựa chọn phần mềm phù hợp để phát triển hệ thống điều hướng tự động cho drone phục vụ mục đích dân sự và quân sự. Hình 17.3 là kiến trúc của phần mềm điều khiển bay do chúng tôi đề xuất.



*Hình 17.3. Sơ đồ tổng quan phần mềm điều khiển bay*

* Tuỳ chỉnh phần mềm bay, tối ưu tham số để tương thích với cấu hình phần cứng và hệ thống phần mềm của drone.

***Công việc 2.3: Nghiên cứu và phát triển thuật toán lập kế hoạch cục bộ theo thời gian thực***

* Phân tích, đánh giá các thuật toán tiên tiến trên thế giới.
* Nghiên cứu, phát triển kiến trúc, thuật toán lưu trữ thông tin bản đồ cục bộ từ dữ liệu cảm biến đảm bảo đáp ứng tốc độ cao, chi phí tính toán thấp. Ở bước này, chúng tôi tập trung vào xử lý tín hiệu pointcloud từ dữ 2 nguồn cảm biến chính là Lidar 3D và Camera để nhanh chóng trích xuất các vùng an toàn cho quá trình sinh quỹ đạo cục bộ. Dưới đây là kiến trúc phần mềm do chúng tôi đề xuất.  
  

*Hình 17.4. Luồng xử lý dữ liệu bản đồ cục bộ đề xuất*

* Nghiên cứu, phát triển thuật toán sinh quỹ đạo tối ưu phù hợp với yêu cầu môi trường hoạt động, phần cứng giới hạn. Chúng tôi tập trung vào giải quyết bài toán tối ưu thời gian và năng lượng hoạt động cũng như đảm duy trình tính an toàn trong quá trình bay (không va chạm với vật cản).
* Nghiên cứu, phát triển thuật toán điều khiển bám quỹ đạo. Mục đích của bộ điều khiển này là đưa ra tín hiệu điều khiển phù hợp để drone có thể di chuyển dọc theo quỹ đạo tối ưu đã được tính toán ở bước trước. Chúng tôi chọn bộ điều khiển dạng GTC (Geometric Tracking Control) vì khả năng đáp ứng ở tốc độ cao và tính đơn giản, hiệu quả của nó. Dưới đây là kiến trúc phần mềm do chúng tôi đề xuất.

A diagram of a company

AI-generated content may be incorrect.

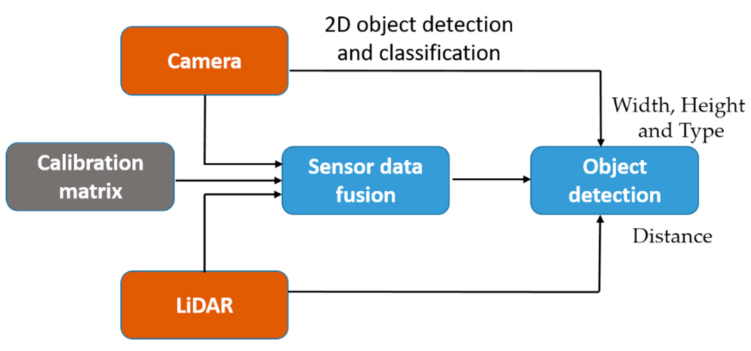
*Hình 17.4. Luồng xử lý của bộ điều khiển bám quỹ đạo đề xuất*

***Công việc 2.4: Nghiên cứu và phát triển thuật toán lập kế hoạch toàn cục***

* Nghiên cứu, phân tích các dạng biểu diễn bản đồ toàn cục cho drone (Statellite map, Topographic map, ...).
* Nghiên cứu, phát triển thuật toán sinh quỹ đạo toàn cục tối ưu dựa trên thông tin bản đồ được cung cấp.

***Công việc 2.5: Phát triển thuật toán điều khiển bắt, bám đối tượng***

* Phân tích, đánh giá các thuật toán tiên tiến trên thế giới.
* Nghiên cứu, phát triển thuật toán xác định vị trí và dự đoán di chuyển của đối tượng dựa trên dữ liệu cảm biến (Camera, Lidar3D, ...). Hình 17.5 là kiến trúc phần mềm do chúng tôi đề xuất.
* Nghiên cứu, phát triển thuật toán điều khiển bám, bắt đối tượng dựa trên thông tin về vị trí và quỹ đạo đự đoán. Các thông tin này được sử dụng làm đầu vào của bộ sinh quỹ đạo tối ưu (Trajectory Generation) bên cạnh thông tin từ bản đồ cục bộ (Local Map). Hình 17.6 là kiến trúc phần mềm do chúng tôi đề xuất.



*Hình 17.5. Luồng xử lý xác định đối tượng và ướng lượng vị trí*

A diagram of a computer

AI-generated content may be incorrect.

*Hình 17.6. Luồng xử lý của bộ điều khiển bám, bắt mục tiêu đề xuất*

***Công việc 2.6: Kiểm thử và hiệu chỉnh hệ thống điều hướng trên phần mềm mô phỏng***

* Xây dựng môi trường thử nghiệm bay trong các nền tảng mô phỏng như Gazebo, Isaacsim.
* Tiến hành kiểm thử và hiệu chỉnh các tham số của hệ thống điều hướng.

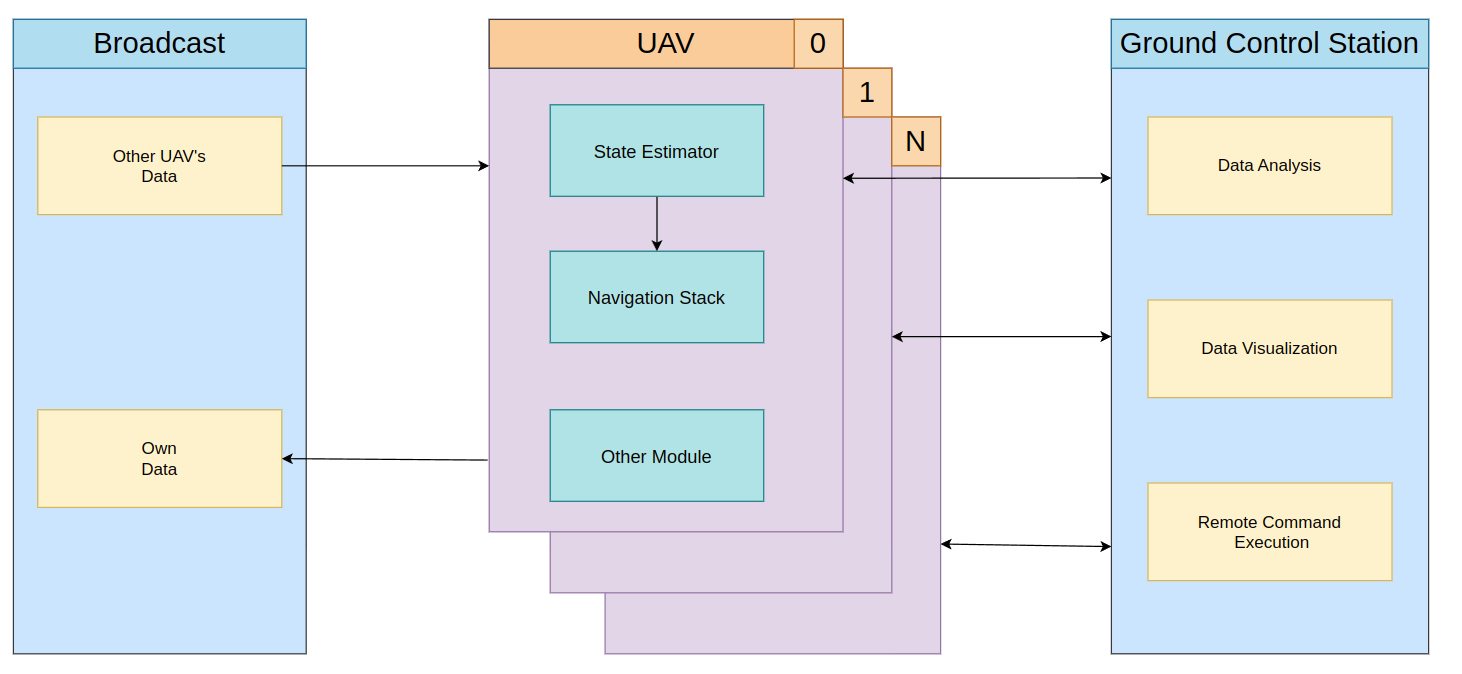
***Công việc 2.7: Triển khai thực nghiệm ngoài thực địa và hoàn thiện hệ thống***

* Triển khai thử nghiệm và hiệu chỉnh hệ thống trên thực tế

**Nội dung 3: Nghiên cứu công nghệ điều khiển bầy đàn cho drone**

***Công việc 3.1: Nghiên cứu, xây dựng kiến trúc giao tiếp cho hệ thống bầy đàn Drone***

* Phân tích, đánh giá các thuật toán tiên tiến trên thế giới.
* Tập trung vào xây dựng kiến trúc giao tiếp phân tán để tối ưu hóa khả năng chịu lỗi và thích nghi cho hệ thống bầy đàn Drone. Hệ thống được thiết kế nhằm mục đích hoạt động linh hoạt, hiệu quả bao gồm 3 thành phần chính: UAV, thành phần broadcast và thành phần giao tiếp với trạm điều khiển mặt đất (Ground Control Station). Các UAV sẽ chia sẻ dữ liệu hoạt động cần thiết với nhau qua broadcast (có thể là dịch vụ cloud hoặc MQTT broker), đồng thời cũng nhận kệnh và thực thi dưới sự kiểm soát của trạm điều khiển mặt đất. Dưới đây là kiến trúc phần mềm do chúng tôi đề xuất.



*Hình 17.7. Kiến trúc giao tiếp của hệ thống bầy đàn UAV đề xuất*

***Công việc 3.2: Nghiên cứu, phát triển thành phần tối ưu phân chia nhiệm vụ***

* Phân tích, đánh giá các thuật toán tiên tiến trên thế giới.
* Giải quyết bài toán phân bổ tác vụ và lập kế hoạch cho hệ thống drone dưới các ràng buộc giới hạn của mỗi Drone (các giới hạn về phần cứng: thời lượng pin, tải trọng tối đa). Các ràng buộc về thời gian hay các cặp điểm lấy-trả hàng (Pickup-Delivery) cũng cũng được đưa vào đối với nhiệm vụ giao hàng. Một trình giải VRP (Vehicle Routing Problem Solver) được triển khai để giải quyết bài toán phân bổ tác vụ có ràng buộc kèm với đó là đường dẫn mục tiêu và các hành động phù hợp để thực hiện nhiệm vụ được gửi tới các UAV. Dưới đây là kiến trúc phần mềm do chúng tôi đề xuất.

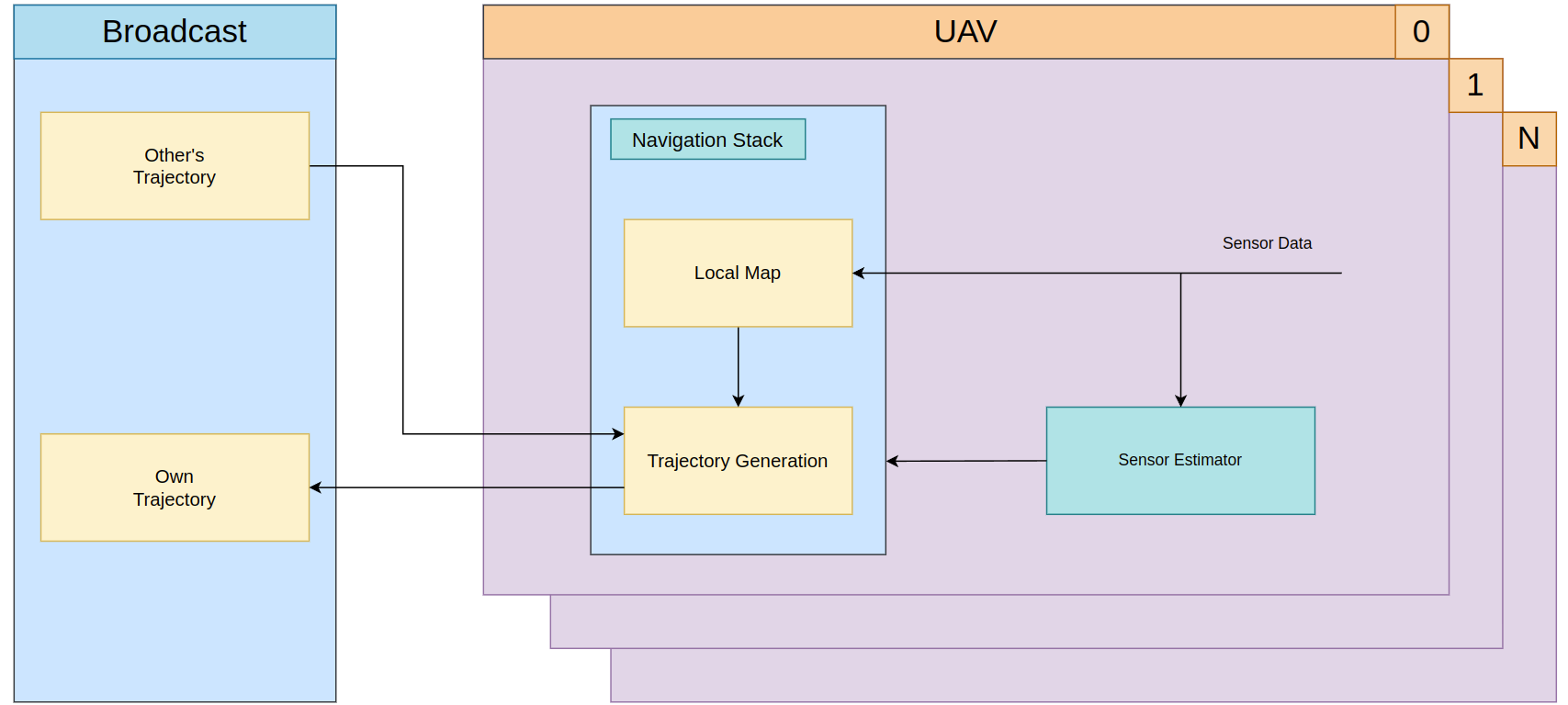
*A diagram of a diagram

AI-generated content may be incorrect.*

*Hình 17.8. Luồng xử lý bài toán phân bổ tác vụ do chúng tôi đề xuất*

***Công việc 3.3: Nghiên cứu, phát triển thành phần bay, giữ, di chuyển theo đội hình***

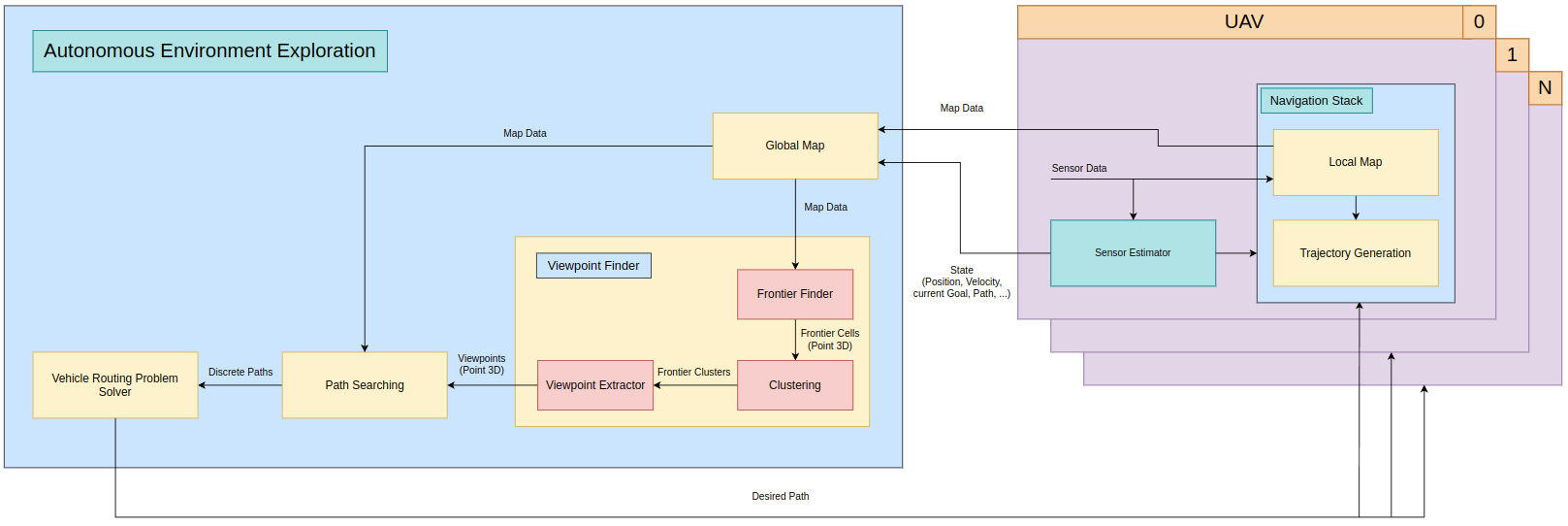
* Phân tích, đánh giá các thuật toán tiên tiến trên thế giới.
* Nghiên cứu và phát triển thuật toán điều khiển đội hình dưới dạng phân tán kết hợp với các phương pháp điều khiển mới như điều khiển thích nghi, điều khiển bền vững để tăng chất lượng điều khiển, khả năng chịu lỗi trong quá trình hoạt động cho hệ thống bầy đàn. Vị trí và quỹ đạo của các UAV sẽ được broadcast để các UAV khác có thể sử dụng cho quá trình sinh qũy đạo tối ưu. Dưới đây là kiến trúc phần mềm do chúng tôi đề xuất.



*Hình 17.9. Luồng xử lý điều khiển bay giữ đội hình phân tán đề xuất*

***Công việc 3.4: Nghiên cứu, phát triển thành phần phối hợp khám phá địa hình tự động***

* Phân tích, đánh giá các thuật toán tiên tiến trên thế giới.
* Nghiên cứu và phát triển thuật toán mở rộng và lập bản đồ một cách đồng thời hoàn toàn tự động dưới ràng buộc về khả năng giao tiếp của mỗi Drone trong bầy (băng thông giới hạn) nhằm giảm thiểu chi phí hạ tầng (bao gồm phần cứng cho Drone, hạ tầng truyền thông) cho các khu vực có phạm vi lớn. Luồng hoạt động được xử lý như sau: Các UAV thu thập dữ liệu cảm biến (Lidar3D, Camera), sau đó chuyển đến quá trình Volumetric Mapping để lưu trữ thông tin bản đồ toàn cục, tại đây các vùng biên (Frontier) mới mở rộng được trích xuất và thực hiện tính toán để tìm ra các điểm nhìn tối ưu (Optimial Viewpoint). Cuối cùng, kết hợp với trình giải tối ưu VRP (Vehicle Routing Problem Solver) để phân bổ các goal mục tiêu và đường dẫn kèm theo đến goal cho mỗi UAV. Quá trình lặp lại cho đến khi không còn Viewpoint nào được trích xuất (toàn bộ khu vực chỉ định đã được khám phá). Dưới đây là kiến trúc phần mềm do chúng tôi đề xuất.



*Hình 17.10. Luồng xử lý khám phá địa hình tự động cho bầy đàn UAV đề xuất*

**Nội dung 4: Nghiên cứu hệ thống trung tâm chỉ huy thông minh cho các phương tiện không người lái**

***Công việc 4.1: Xây dựng hệ thống kết nối và điều khiển đa robot***

* Nghiên cứu các SDK, API kết nối với drone của DJI, Authel để xây dựng giải pháp kết nối chung cho drone các hãng vào hệ thống.
* Thiết kế kiến trúc hệ thống điều phối và quản lý tập trung hoặc phân tán cho nhiều thiết bị không người lái cùng hoạt động.
* Phát triển giao thức truyền nhận dữ liệu giữa các giữa phương tiện không người lái với trung tâm điều khiển, sử dụng Wi-Fi, 4G/5G, hoặc Mesh Network.
* Tối ưu hệ thống xử lý dữ liệu (tại chỗ hoặc đám mây) để đảm bảo phản hồi nhanh và ổn định. Đảm bảo hệ thống hoạt động ổn định được với tối thiểu 1000 thiết bị.
* Kiểm thử và mô phỏng toàn bộ hệ thống trong các tình huống thực tế (môi trường đô thị, nhà máy, nông trại…).
* Lập lịch cho thiết bị không người lái hoạt động. Đảm bảo các hoạt động theo đúng trình tự và thời gian, đặc biệt quan trọng khi làm việc nhóm hoặc xử lý dữ liệu cảm biến đồng thời.

***Công việc 4.2: Xây dựng hệ thống lưu log và quản trị dữ liệu***

* Thiết kế kiến trúc lưu trữ dữ liệu (log, hình ảnh, video...) theo thời gian thực từ các thiết bị không người lái.
* Xây dựng cơ chế thu thập log tự động từ nhiều nguồn (thiết bị, hệ thống điều khiển, ứng dụng backend).
* Phân loại, chuẩn hóa và mã hóa dữ liệu log theo định dạng phù hợp (JSON, CSV, binary...).
* Tích hợp cơ sở dữ liệu hiệu năng cao (MongoDB, PostgreSQL) để lưu trữ và truy vấn nhanh.
* Phát triển dashboard giám sát log theo thời gian thực, cảnh báo bất thường, và thống kê hiệu suất hệ thống.

***Công việc 4.3: Xây dựng hệ thống hiển thị thông tin, đa luồng video stream, cho nhiều phương tiện không người lái***

* Thiết kế giao diện hiển thị tập trung để giám sát đồng thời nhiều phương tiện không người lái.
* Xây dựng hệ thống xử lý và truyền đa luồng video (multi-stream) theo thời gian thực với độ trễ thấp.
* Tích hợp bản đồ số (Mapbox, Leaflet, Google Maps...) để hiển thị vị trí, hướng di chuyển, và trạng thái từng thiết bị.
* Hiển thị dữ liệu cảm biến, thông số kỹ thuật, và cảnh báo theo thời gian thực trên giao diện người dùng.
* Phát triển cơ chế lựa chọn, phân loại và ưu tiên hiển thị video theo khu vực, loại nhiệm vụ hoặc mức độ cảnh báo.
* Đảm bảo khả năng mở rộng để theo dõi hàng chục đến hàng trăm thiết bị cùng lúc.
* Tối ưu hóa hiệu năng xử lý hình ảnh/video và bảo mật luồng truyền dữ liệu.

**Nội dung 5: Nghiên cứu công nghệ định vị kết hợp GPS-RTK và điều hướng dẫn đường cho robot mặt đất hoạt động trên địa hình đồi núi có đường mòn.**

***Công việc 5.1: Xây dựng và phát triển thuật toán định vị tích hợp nhiều cảm biến (Sensor Fusion)***

* Để đảm bảo khả năng định vị liên tục và chính xác cho Robot mặt đất trong môi trường địa hình đồi núi có đường mòn, thuật toán fusion đa cảm biến được thiết kế dựa trên việc kết hợp dữ liệu từ GPS-RTK (cho thông tin vị trí tuyệt đối), IMU (cho thông tin quán tính ngắn hạn) và một trong hai nguồn cảm biến hỗ trợ: camera hoặc LiDAR 3D (cho vị trí tương đối theo chuyển động thực tế). Mỗi loại cảm biến đảm nhiệm một vai trò quan trọng và được tích hợp vào một kiến trúc xử lý thông nhất để tạo ra chuỗi pose định vị chính xác theo thời gian thực.
* Cụ thể, tại các khu vực tín hiệu GNSS tốt, GPS-RTK đóng vai trò cung cấp vị trí tuyệt đối ở độ chính xác centimet. Dữ liệu đầu ra từ GPS được chuyển đổi về tọa độ địa phương (ENU hoặc map frame). Trong khi đó, IMU cung cấp dữ liệu gia tốc và vận tốc, dùng để dự đoán chuyển động ngắn hạn giữa hai lần cập nhật GPS, đặc biệt hữu ích khi Robot di chuyển trong vùng mất tín hiệu vệ tinh.
* Dự kiến nếu sử dụng camera (monocular hoặc stereo) thì thuật toán Visual-Inertial Odometry (VIO) sẽ được triển khai, có thể kế thừa kết quả từ nhóm công việc ở Nội dung 1 (Nghiên cứu công nghệ AI thị giác máy tính trong định vị Drone). VIO sử dụng các đặc trưng ảnh để tính toán chuyển động tương đối của Robot giữa các khung hình, đồng thời sử dụng IMU để bổ trợ định hướng và giảm trôi. Trường hợp sử dụng, LiDAR 3D, cần triển khai thuật toán LiDAR-Inertial Odometry (LIO) để tính toán chuyển động dựa trên sự thay đổi của đám mây điểm và cảm biến quán tính. Thuật toán LIO dự kiến kế thừa thuật toán của sản phẩm Robot giám sát (VSR).
* Toàn bộ dữ liệu GPS, IMU và VIO hoặc LIO sẽ được tích hợp thông qua một bộ Fusion. Có hai hướng đề xuất là sử dụng gói robot\_localization trong ROS hoặc xây dựng hệ thống tối ưu đồ thị (Graph-based Optimization) bằng các thư viện điển hình như GTSAM, iSAM.

***Công việc 5.2: Xử lý ngoại lệ khi mất GPS hoặc GPS bị nhiễu***

* Phát hiện trạng thái GPS chuyển từ FIXED sang FLOAT hoặc mất tín hiệu.
* Tự động chuyển sang chế độ định vị tương đối bằng VIO, IMU hoặc LIO.

***Công việc 5.3: Nghiên cứu và triển khai thuật toán điều hướng dẫn đường***

* Dự kiến thu thập dữ liệu môi trường đồi núi, tạo bộ dataset chuẩn hóa, gồm các nhãn như đường mòn, địa hình dốc, vạch biên an toàn, ...
* Phát triển mô-đun nhận dạng đường mòn và phân vùng môi trường: thiết kế mô hình thị giác máy tính hoặc có thể phát triển từ những mô hình pretrained như DeepLabV3+, YOLOv8-seg, đào tạo các mô hình đã chọn với bộ dữ liệu địa hình đồi núi để đảm bảo tính chuyên biệt.
* Kết hợp với bản đồ và hệ thống định vị: đồng bộ thông tin từ thị giác với vị trí GPS-RTK, gắn thông tin thị giác vào map local hoặc global, dự đoán vùng khả thi cho điều hướng trong local map, sử dụng kết quả phân đoạn để tạo costmap động cho planner.

***Công việc 5.4: Nghiên cứu và xây dựng mô-đun phát hiện và tránh vật cản bằng thị giác máy tính***

* Xác định loại vật cản và đặc điểm môi trường đồi núi, tìm kiếm các bộ dữ liệu ngoài trời phù hợp hoặc tự thu nhập dữ liệu và gán nhãn riêng.
* Lựa chọn mô hình phù hợp ví dụ các mô hình object detection (YOLOv5/v8, Faster R-CNN, ...), semantic/instance segmentation để phát hiện biên vật can rõ ràng hoặc phát triển từ kết quả kế thừa từ thuật toán phát hiện và tránh vật cản của các dòng Robot trước đó nhóm đề tài đã phát triển.
* Phát triển mô-đun dự đoán và ước lượng vị trí vật cản.
* Xây dựng bản đồ local obstacle map để Robot tự động tránh vật cản hoặc tạo đường mới khi gặp vật cản.

***Công việc 5.5: Kiểm tra hiệu năng và đánh giá thuật toán***

* Đánh giá sai số định vị thực, quỹ đạo với ground-truth (nếu có).
* Tính toán độ ổn định: trôi theo thời gian, jitter, sai lệch heading.
* Đánh giá độ chính xác của các mô hình nhận diện đường mòn trong khu vực đồi núi, phát hiện và tránh vật cản.

**Nội dung 6: Nghiên cứu công nghệ sinh dữ liệu để phục vụ huấn luyện mô hình, mô phỏng hoạt động của robot, drone**

***Công việc 6.1: Xây dựng và ngẫu nhiên hóa môi trường mô phỏng***

* Công việc đầu tiên tập trung vào việc tạo dựng một môi trường 3D mẫu và thực thi ngẫu nhiên hóa kịch bản (scenario randomization). Trên nền tảng Isaac Sim, nhóm sử dụng định dạng USD để mô hình hóa scene, tích hợp các mesh và vật liệu từ thư viện SimReady. Tiếp đó, thông qua API Python của Omniverse Replicator hoặc file cấu hình YAML, các tham số sau được thiết lập để sinh biến thể:
* Vị trí và hướng đặt vật thể (ví dụ: bàn, ghế trong drone indoor; chậu cây, thùng rác trong robot ngoài trời) được ngẫu nhiên hóa trong một vùng tọa độ xác định.
* Đặc tính bề mặt (texture) và ánh sáng (cường độ, góc chiếu) được phân bố ngẫu nhiên theo phân phối đồng đều hoặc Gauss.
* Góc nhìn và tọa độ của sensor camera được điều chỉnh thay đổi qua từng lần chạy để đảm bảo đa dạng dữ liệu về POV (point-of-view).
* Việc thực hiện ngẫu nhiên hóa kịch bản theo từng batch cho phép thu được hàng nghìn biến thể scene trong thời gian ngắn, đảm bảo bộ dữ liệu huấn luyện bao phủ rộng các tình huống thực tế và giảm thiểu nguy cơ overfitting.

***Công việc 6.2: Mô phỏng cảm biến, gán nhãn và xuất dữ liệu***

* Cấu hình sensor: Gắn camera RGB (độ phân giải 800×600 hoặc 1280×720), depth camera, LiDAR 16 kênh và IMU vào thực thể robot/drone trong scene. Mỗi sensor được cài đặt tần suất lấy mẫu và góc quét (đối với LiDAR).
* Ghi nhận dữ liệu: Chạy mô phỏng trong Isaac Sim, thu thập ảnh RGB, bản đồ độ sâu, point cloud, dữ liệu gia tốc và góc quay. Mỗi khung dữ liệu đồng bộ về thời gian và vị trí.
* Tự động gán nhãn: Sử dụng Semantics Schema Editor để đánh dấu semantic mask của các đối tượng mục tiêu, tính toán tọa độ pose 6-DOF và bounding box đã được xác thực. Quá trình này diễn ra song song với mô phỏng, đảm bảo nhãn chính xác và nhất quán.
* Xuất dữ liệu: Dữ liệu nhị phân (PNG cho ảnh, PLY cho point cloud) và file nhãn (JSON) được Replicator Writer lưu trữ có cấu trúc thư mục rõ ràng, kèm metadata mô tả tham số ngẫu nhiên hóa.

**Nội dung 7: Nghiên cứu công nghệ nhận diện bắt bám đối tượng bằng camera trên drone**

Để phát triển công nghệ nhận diện và bắt bám đối tượng bằng camera trên thiết bị bay không người lái (UAV), quá trình nghiên cứu cần được tổ chức chặt chẽ, bao gồm các công việc chính như phân tích yêu cầu, phát triển thuật toán, tích hợp phần cứng và phần mềm, thử nghiệm và tối ưu hóa hệ thống tuần tự theo mô tả chi tiết sau:

***Công việc 7.1. Phân tích yêu cầu và thu thập dữ liệu***

* Công việc đầu tiên là xác định các yêu cầu cụ thể cho hệ thống nhận diện và bắt bám đối tượng, bao gồm các kịch bản ứng dụng (quân sự như trinh sát, dân sự như giám sát nông nghiệp), điều kiện môi trường (ánh sáng, thời tiết), và độ chính xác cần đạt. Sẽ tiến hành khảo sát các tập dữ liệu hiện có như VisDrone2019 hoặc UAVDT để đánh giá tính phù hợp. Đồng thời, thu thập dữ liệu thực tế bằng cách sử dụng UAV được trang bị camera độ phân giải cao tại các môi trường thực tế như khu vực biên giới, nông thôn hoặc đô thị tại Việt Nam. Dữ liệu sẽ được gán nhãn để phục vụ huấn luyện các mô hình AI, đảm bảo bao quát các trường hợp như đối tượng di động, che khuất, hoặc ánh sáng yếu.

***Công việc 7.2. Phát triển và tối ưu hóa thuật toán AI***

* Giai đoạn này tập trung vào xây dựng và tối ưu hóa các thuật toán để nhận diện và bắt bám đối tượng. Các mô hình học sâu như YOLO, DETR sẽ được thử nghiệm để nhận diện đối tượng với tốc độ cao, kết hợp với Deep SORT hoặc ByteTrack để theo dõi liên tục qua các khung hình video. Công việc bao gồm: (1) huấn luyện mô hình trên các tập dữ liệu đã thu thập, sử dụng các nền tảng như TensorFlow hoặc PyTorch; (2) tối ưu hóa mô hình để giảm độ trễ và phù hợp với tài nguyên tính toán hạn chế trên UAV, chẳng hạn bằng cách áp dụng kỹ thuật cắt tỉa mô hình (model pruning) hoặc lượng tử hóa (quantization); (3) tích hợp các mô-đun xử lý chuyển động phức tạp như bộ lọc Kalman để tăng cường độ chính xác trong môi trường động. Các thuật toán sẽ được đánh giá qua các chỉ số như MOTA (Multiple Object Tracking Accuracy) và IDF1 (Identification F1 Score).

***Công việc 7.3. Tích hợp phần cứng và phần mềm***

* Sau khi phát triển thuật toán, cần tích hợp chúng vào hệ thống UAV. Công việc này bao gồm: (1) lựa chọn phần cứng phù hợp như camera độ phân giải cao và bo mạch tính toán nhúng như NVIDIA Jetson Board để xử lý AI thời gian thực; (2) phát triển phần mềm giao tiếp giữa camera, thuật toán AI, và hệ thống điều khiển UAV, đảm bảo đồng bộ dữ liệu hình ảnh và lệnh điều khiển; (3) thiết kế giao diện truyền dữ liệu không dây để gửi kết quả nhận diện về trạm điều khiển mặt đất. Các thử nghiệm mô phỏng sẽ được thực hiện để kiểm tra tính ổn định của hệ thống trước khi triển khai thực tế.

***Công việc 7.4. Thử nghiệm và tối ưu hóa hệ thống***

* Giai đoạn cuối bao gồm thử nghiệm hệ thống trong các điều kiện thực tế như môi trường nông thôn, đô thị, biên giới hoặc khu vực có thời tiết bất lợi (mưa, sương mù). Các kịch bản thử nghiệm sẽ bao gồm nhận diện và theo dõi các đối tượng như phương tiện, con người, hoặc công trình. Dữ liệu thu thập từ các bài kiểm tra sẽ được phân tích để đánh giá hiệu suất hệ thống qua các chỉ số như độ chính xác, độ trễ, và tỷ lệ lỗi. Dựa trên kết quả, hệ thống sẽ được tinh chỉnh, bao gồm điều chỉnh tham số thuật toán, nâng cấp phần cứng, hoặc bổ sung các kỹ thuật xử lý nhiễu. Cuối cùng, tài liệu hóa toàn bộ quy trình và kết quả sẽ được thực hiện để phục vụ triển khai thực tế và chuyển giao công nghệ.

**Nội dung 8: Thử nghiệm các công nghệ trên drone và robot tự hành mặt đất.**

***Công việc 8.1: Thử nghiệm các công nghệ trên drone***

* Việc thử nghiệm các công nghệ trên drone đang được triển khai nhằm đánh giá hiệu quả và độ ổn định của hệ thống trong môi trường thực tế. Đầu tiên, nền tảng chỉ huy thông minh phương tiện bay được đánh giá theo hai phương diện chính là quản lý thiết bị bay và điều phối hoạt động. Các yếu tố như số lượng thiết bị bay có thể quản lý, độ trễ khi truyền dữ liêu sẽ được kiểm tra bởi bộ testcase đã được chuẩn bị trước. Đối với phương diện điều phối hoạt động, một đội hình gồm 5 drone sẽ được triển khai. Đội hình thiết lập với khoảng cách giữa các vị trí được cố định sẵn, các drone vừa bay giữ nguyên đội hình đồng thời vừa thực hiện chức năng quét bản đồ. Diện tích địa hình khám phá xấp xỉ 3000 m2 để đảm bảo việc xác nhận khả năng giữ đội hình trong quá trình thực hiện nhiệm vụ. Trong quá trình bay, các drone phải đảm bảo độ lệch đội hình không vượt quá 2m khi có tín hiệu GNSS, tỉ lệ bao phủ không gian đạt tối thiểu 90% và hiệu suất phân công nhiệm vụ (tỷ lệ phương tiện hoàn thành nhiệm vụ) ≥ 80%. Thử nghiệm sẽ được lặp lại tối thiểu 30 lần để đảm bảo độ tin cậy của dữ liệu đo.
* Ngoài ra, việc thử nghiệm các chỉ tiêu liên quan đến độ chính xác hoạt động của nền tảng AI drone cũng là cần thiết. Khi có tín hiệu GNSS, nền tảng được thử nghiệm với quỹ đạo bay ngoài trời đã đặt sẵn, có khả năng bay với vận tốc tối đa 5 m/s vừa thực hiện bám bắt đối tượng như con người, phương tiện giao thông di chuyển với cùng vận tốc. Các đối tượng như con người, phương tiện giao thông di chuyển sẽ được chuẩn bị sẵn để kiểm tra độ chính xác bám bắt của mô hình AI, với mục tiêu độ chính xác bám bắt ≥ 70%. Môi trường thử nghiệm được thiết lập sao cho có các công trình hoặc cây cối tạo vật cản để thử nghiệm khả năng dừng hoặc tránh va chạm ở khoảng cách gần nhất 100m. Bên cạnh đó, các thử nghiệm liên quan đến công nghệ điều hướng, định vị bằng camera cũng được thực hiện ở các môi trường không có tín hiệu GNSS như môi trường trong nhà hay ở độ cao trên 100m. Các nội dung thử nghiệm sẽ được thực hiện lại tương tự như khi thử nghiệm với tín hiệu GNSS, với mục tiêu độ lệch định vị ≤ 10m.

***Công việc 8.2: Thử nghiệm các công nghệ trên robot tự hành mặt đất***

* Việc thử nghiệm robot tự hành ngoài trời nhằm đánh giá khả năng hoạt động của hệ thống trong điều kiện môi trường thực tế, với các chỉ tiêu kỹ thuật quan trọng. Robot được kiểm tra khả năng phát hiện, dừng hoặc tránh vật cản khi di chuyển trên địa hình phức tạp. Tốc độ di chuyển tối đa mà hệ thống vẫn đảm bảo khả năng phản ứng và xử lý vật cản được yêu cầu tối đa 1.3m/s. Đồng thời, độ sai số bám quỹ đạo theo tín hiệu định vị toàn cầu (GNSS) cần duy trì trong giới hạn ≤ 20cm để đảm bảo độ chính xác cao. Phương pháp thử nghiệm được áp dụng là xây dựng một địa hình đường đất mô phỏng điều kiện ngoài trời, đồng thời bố trí các vật cản như cọc gỗ, khối bê tông, hoặc cây nhân tạo trên đường đi. Robot sẽ được lập trình sẵn một quỹ đạo dự kiến, sau đó tiến hành chạy thực tế trên địa hình thử nghiệm. Dữ liệu vị trí thực tế của robot được ghi lại và so sánh với quỹ đạo ban đầu để đánh giá sai số. Kết quả thử nghiệm sẽ phản ánh rõ hiệu quả của hệ thống định vị, điều hướng và tránh vật cản. Qua đó, nhóm phát triển có thể tinh chỉnh thuật toán điều khiển và nâng cao hiệu suất hoạt động của robot.

**17.2. Các công nghệ lõi của đề tài**

1. **Các công nghệ lõi của đề tài sẽ làm chủ**

- Đối với phân hệ chỉ huy điều khiển tập trung

* Làm chủ (cấp độ sáng tạo) công nghệ điều khiển bay và tối ưu bầy đàn áp dụng cho bài toán tìm kiếm cứu nạn hộ sử dụng drone.
* Làm chủ (cấp độ đổi mới) công nghệ sinh dữ liệu để phục vụ huấn luyện mô hình, mô phỏng hoạt động của robot, drone.

- Đối với phân hệ điều khiển Drone

* Làm chủ (cấp độ đổi mới) phần mềm điều khiển bay.
* Làm chủ (cấp độ sáng tạo) công nghệ định vị, dẫn đường bay bằng thị giác máy tính trong môi trường không có tín hiệu GNSS.
* Làm chủ (cấp độ sáng tạo) công nghệ thị giác máy tính giúp drone phát hiện, nhận dạng, bắt bám đối tượng, mục tiêu.

- Đối với phân hệ điều khiển robot mặt đất

* Làm chủ (cấp độ sáng tạo) công nghệ nhận biết môi trường bằng thị giác máy tính phục vụ điều hướng dẫn đường cho robot mặt đất hoạt động trên địa hình đồi núi có đường mòn cho phương tiện di chuyển.
* Làm chủ (cấp độ đổi mới) công nghệ định vị robot sử dụng tín hiệu GNSS-RTK cho robot hoạt động ngoài trời.

1. **Các công nghệ lõi phụ thuộc vào đối tác**

- Công nghệ điều khiển bay, có thể sử dụng đến như phần mềm PX4, Betaflight, Adupilot.

**17.3. Khả năng làm chủ sản phẩm**

Các công nghệ lõi điều khiển và nền tảng chỉ huy thông minh các phương tiện không người lái trên bộ và trên không có các thành phần chính và tỷ lệ làm chủ của Viettel AI như sau:

| STT | Thành phần | Tỷ lệ làm chủ | Ghi chú |
| --- | --- | --- | --- |
| 1 | Công nghệ AI thị giác máy tính trong định vị drone | 100% |  |
| 2 | Công nghệ điều hướng tự động cho drone | 100% |  |
| 3 | Công nghệ điều khiển bầy đàn | 100% |  |
| 4 | Hệ thống trung tâm chỉ huy thông minh cho các phương tiện không người lái | 100% |  |
| 5 | Công nghệ định vị kết hợp GPS-RTK và điều hướng dẫn đường cho robot mặt đất hoạt động trên địa hình đồi núi có đường mòn | 100% |  |
| 6 | Công nghệ sinh dữ liệu để phục vụ huấn luyện mô hình, mô phỏng hoạt động của robot, drone | 70% | Do phụ thuộc phần mềm mô phỏng của bên thứ 3 |
| 7 | Nghiên cứu công nghệ nhận diện bắt bám đối tượng bằng camera trên drone | 100% |  |

**17.4. Nền tảng sản phẩm**

- Nền tảng phần mềm robot tự hành thông minh (Viettel AI Robot Platform)

- Nền tảng Nvidia Isaac Gr00t

- Nền tảng AI UxV cho drone: “Nền tảng chỉ huy thông minh các phương tiện không người lái trên bộ và trên không”, bao gồm ba phân hệ: Phân hệ chỉ huy điều khiển tập trung, phân hệ điều khiển drone và phân hệ điều khiển robot mặt đất.

**18. Phương án phối hợp với các tổ chức nghiên cứu và cơ sở sản xuất trong nước**

Không có

**19. Phương án hợp tác quốc tế** (nếu có)

Không có

**20. Phương án thuê chuyên gia** (nếu có)

Không có

**21. Tiến độ thực hiện**

| **TT** | **Các nội dung, công việc chủ yếu cần được thực hiện;  các mốc đánh giá chủ yếu** | **Kết quả phải đạt** | **Thời gian** (bắt đầu,  kết thúc) | **Cá nhân,  đơn vị  thực hiện\*** |
| --- | --- | --- | --- | --- |
| *(1)* | *(2)* | *(3)* | *(4)* | *(5)* |
| ***1*** | ***Nội dung 1: Nghiên cứu công nghệ V-SLAM, VIO trong định vị drone*** | Thuật toán định vị bằng thị giác máy trên drone. | T1-T15 | Kỹ sư HRL-15 |
|  | - Công việc 1: Tiền xử lý dữ liệu camera và IMU |  | T1 – T6 | Huynm12 |
|  | - Công việc 2: Nghiên cứu và đánh giá các hệ thống Visual SLAM |  | T1-T2 | Huynm12 |
|  | - Công việc 3: Nghiên cứu và đánh giá các hệ thống VIO |  | T2-T4 | Huynm12 |
|  | - Công việc 4: Nghiên cứu và đánh giá các hệ thống Image Geo-Localization |  | T4-T12 | Cuongnq23 |
|  | - Công việc 5: Nghiên cứu kết hợp đa cảm biến nâng cao hiệu suất định vị |  | T6-T12 | HRL15 |
|  | - Công việc 6: Tìm hiểu tập dữ liệu mở và nghiên cứu tiêu chí đánh giá |  | T1-T3 | Huynm12 |
| ***2*** | ***Nội dung 2: Nghiên cứu công nghệ điều hướng tự động cho drone*** | Thuật toán điều hướng tự động trên drone | T1-T12 | HRL15 |
|  | - Công việc 1: Nghiên cứu và xây dựng các mô-đun và thuật toán cho hệ thống điều hướng tự động ANS |  | T1-T2 | Datld5 |
|  | - Công nghệ 2: Nghiên cứu, làm chủ phần mềm điều khiển bay |  | T1-T6 | Datld5 |
|  | - Công việc 3: Nghiên cứu và phát triển thuật toán lập kế hoạch cục bộ theo thời gian thực |  | T3-T4 | Datld5 |
|  | - Công việc 4: Nghiên cứu, phát triển thuật toán điều khiển bám quỹ đạo |  | T4-T10 | Datld5 |
|  | - Công việc 5: Phát triển thuật toán điều khiển bắt, bám đối tượng |  | T3-T9 | HRL15 |
|  | - Công việc 6: Kiểm thử và hiệu chỉnh hệ thống điều hướng trên phần mềm mô phỏng |  | T9-T12 | Binhnh23 |
|  | - Công việc 7: Triển khai thực nghiệm ngoài thực địa và hoàn thiện hệ thống |  | T12-T15 | Binhnh23 |
| ***3*** | ***Nội dung 3: Nghiên cứu công nghệ điều khiển bầy đàn cho drone*** |  | T1-15 | Datld5 |
|  | - Công việc 1: Nghiên cứu, xây dựng kiến trúc giao tiếp cho hệ thống bầy đàn Drone |  | T1-T3 | Toandk2 |
|  | - Công việc 2: Nghiên cứu, phát triển thành phần tối ưu phân chia nhiệm vụ |  | T4-T6 | Datld5 |
|  | - Công việc 3: Nghiên cứu, phát triển thành phần bay, giữ, di chuyển theo đội hình |  | T7-T9 | Datld5 |
|  | - Công việc 4: Nghiên cứu, phát triển thành phần phối hợp khám phá địa hình tự động |  | T10-T15 | HRL15 |
| ***4*** | ***Nội dung 4: Nghiên cứu hệ thống trung tâm quản lý và xử lý thông tin cho drone*** |  | T1-T15 | Toandk2 |
|  | - Công việc 1: Xây dựng hệ thống kết nối và điều khiển đa robot |  | T1-T3 | Toandk2 |
|  | - Công việc 2: Xây dựng hệ thống lưu log và quản trị dữ liệu |  | T1-T3 | Toandk2 |
|  | - Công việc 3: Xây dựng hệ thống hiển thị thông tin, đa luồng video stream, cho nhiều phương tiện không người lái |  | T4-T6 | Toandk2 |
| ***5*** | ***Nội dung 5: Nghiên cứu công nghệ định vị kết hợp GPS-RTK và điều hướng dẫn đường cho robot mặt đất hoạt động trên địa hình đồi núi có đường mòn*** |  | T1-T15 | Huynm12, Kientt39 |
|  | - Công việc 1: Xây dựng và phát triển thuật toán định vị tích hợp nhiều cảm biến (Sensor Fusion) |  | T1-T6 | Kientt39 |
|  | - Công việc 2: Xử lý ngoại lệ khi mất GPS hoặc GPS bị nhiễu |  | T7-T9 | Tunglx12 |
|  | - Công việc 3: Nghiên cứu và triển khai thuật toán điều hướng dẫn đường |  | T10-T12 | Kientt39 |
|  | - Công việc 4: Nghiên cứu và xây dựng mô-đun phát hiện và tránh vật cản bằng thị giác máy tính |  | T12-T15 | Tunglx12 |
|  | - Công việc 5: Kiểm tra hiệu năng và đánh giá thuật toán |  | T15-T18 | Tunglx12 |
| ***6*** | ***Nội dung 6: Nghiên cứu công nghệ sinh dữ liệu để phục vụ huấn luyện mô hình, mô phỏng hoạt động của robot, drone*** |  | T1-T15 | HRL15 |
|  | - Công việc 1: Xây dựng và ngẫu nhiên hóa môi trường mô phỏng |  | T1-T6 | Datld5 |
|  | - Công việc 2: Mô phỏng cảm biến, gán nhãn và xuất dữ liệu |  | T7-T12 | Datld5 |
| ***7*** | ***Nội dung 7: Nghiên cứu công nghệ nhận diện bắt bám đối tượng bằng camera trên drone*** |  | T1-T15 | Cuongnq23 |
|  | - Công việc 1: Phân tích yêu cầu và thu thập dữ liệu |  | T1-T6 | Datld5 |
|  | - Công việc 2: Phát triển và tối ưu hóa thuật toán AI |  | T3-T6 | Cuongnq23 |
|  | - Công việc 3: Tích hợp phần cứng và phần mềm |  | T7-T9 | Kientt39 |
|  | - Công việc 4: Thử nghiệm và tối ưu hóa hệ thống |  | T9-T15 | Tunglx12 |
| ***8*** | ***Nội dung 8: Thử nghiệm các công nghệ trên Drone và Robot tự hành mặt đất*** |  |  | Binhnh23 |
|  | - Công việc 1: Thử nghiệm các công nghệ trên Drone | Thử nghiệm các tính năng trong bảng chỉ tiêu đề tài | T12-T15 | Binhnh23 |
|  | - Công việc 2: Thử nghiệm các công nghệ trên Robot tự hành mặt đất | Thử nghiệm các tính năng trong bảng chỉ tiêu đề tài | T15-T18 | Binhnh23 |
| **9** | **Đóng đề tài** | Các tài liệu báo cáo theo yêu cầu. | T18 | Binhnh23 |

**III. SẢN PHẨM KH&CN CỦA ĐỀ TÀI**

**22. Sản phẩm KH&CN chính của đề tài và yêu cầu chất lượng cần đạt**

**22.1 Dạng I:** Các sản phẩm mẫu, sản phẩm là hàng hóa, vật liệu, thiết bị, máy móc và các sản phẩm khác;

| **Số TT** | **Tên sản phẩm cụ thể** | **Sản phẩm tham khảo trên thế giới**  *(theo các tiêu chuẩn mới nhất)* | **Dự kiến số lượng/  quy mô sản phẩm  tạo ra** |
| --- | --- | --- | --- |
| (1) | (2) | (3) | (4) |
| 1 | Nền tảng cơ bản chỉ huy thông minh các phương tiện không người lái trên bộ và trên không | Dronsense, L3Harris – Amorphous, Skydio, Ffaxy. | 01 |

* *Bảng so sánh chỉ tiêu kỹ thuật của sản phẩm tương đương trong Phụ lục 2 kèm theo.*
* *Chi tiêu kỹ thuật của sản phẩm chi tiết trong Phụ lục 3 kèm theo.*

**Mức chất lượng các sản phẩm (Dạng I) so với các sản phẩm tương tự trong nước và nước ngoài** *(Làm rõ cơ sở khoa học và thực tiễn để xác định các chỉ tiêu về chất lượng cần đạt của các sản phẩm của đề tài)*

**Bảng 3.1. So sánh tính năng bay tự động không GNSS**

| **TT** | **Chỉ tiêu kỹ thuật** | **Đơn vị** | **Sản phẩm của Viettel** | **Sản phẩm của Skydio** | **Sản phẩm của FFaxy** | **So sánh, đánh giá**  **(Tốt hơn, kém hơn)** |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | Độ chính xác định vị | m | 10 | 0.1 | 50 | Tương đương |
| 2 | Độ cao hoạt động tối đa | m | 300 | 300 | 300 | Tương đương |
| 3 | Tốc độ bay tự động tối đa theo phương ngang | m/s | 5 | Không công bố | Không công bố | Tốt hơn |
| 4 | Tính năng định vị trong môi trường tối | Có/không | Có | Không | Có | Tương đương |
| 5 | Tính năng so khớp ảnh địa hình | Có/không | Có | Không | Không | Tốt hơn |
| 6 | Khoảng cách dừng, tránh va chạm gần nhất | m | 100 | Không công bố | Không công bố | Tốt hơn |

**Bảng 3.2. So sánh tính năng của hệ thống trung tâm điều hành và xử lý thông tin các phương tiện không người lái**

| **TT** | **Chỉ tiêu kỹ thuật** | **Đơn vị** | **Sản phẩm của Viettel** | **Sản phẩm của DroneSense** | **Sản phẩm của L3Harris-Amorphous** | **So sánh, đánh giá**  **(Tốt hơn, kém hơn)** |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | Lập lịch bay và khảo sát nghiệp vụ | Có/không | Có | Có | Có | Tương đương |
| 2 | Truyền hình ảnh trực tiếp và chia sẻ video | Có/không | Có | Có | Có | Tương đương |
| 3 | Giao diện điều khiển và quản lý (mobile/web) | Có/không | Có | Có | Có | Tương đương |
| 4 | Tích hợp được nhiều loại phương tiện không người lái | Có/không | Có | Không | Có | Tốt hơn |
| 5 | Lưu lịch sử, biểu đồ báo cáo thống kê | Có/không | Có | Có | Có | Tương đương |

**22.2 Dạng II:** Nguyên lý ứng dụng; Phương pháp; Tiêu chuẩn; Quy phạm; Phần mềm máy tính; Bản vẽ thiết kế; Quy trình công nghệ; Sơ đồ, bản đồ; Số liệu, Cơ sở dữ liệu; Báo cáo phân tích; Tài liệu dự báo *(phương pháp, quy trình, mô hình,...)*; Đề án, qui hoạch; Luận chứng kinh tế-kỹ thuật, Báo cáo nghiên cứu khả thi và các sản phẩm khác.

| **TT** | **Tên sản phẩm** | **ĐVT** | **Số lượng** | **Yêu cầu khoa học cần đạt** |
| --- | --- | --- | --- | --- |
| (1) | (2) | (3) | (4) | (5) |
| ***1*** | Tài liệu báo cáo kết quả thực hiện đề tài | Bộ | 01 | Tài liệu KHCN đầy đủ, chi tiết |
| 2 | Tài liệu tóm tắt kết quả thực hiện đề tài | Bộ | 01 | Tài liệu KHCN đầy đủ, chi tiết |
| 3 | Tài liệu sản phẩm |  |  |  |
|  | Tài liệu mô tả sản phẩm | Bộ | 01 | Tài liệu KHCN đầy đủ, chi tiết |
|  | Tài liệu hướng dẫn sử dụng sản phẩm. | Bộ | 01 | Tài liệu KHCN đầy đủ, chi tiết |
|  | Tài liệu hướng dẫn vận hành sản phẩm | Bộ | 01 | Tài liệu KHCN đầy đủ, chi tiết |
|  | Tài liệu bài đo, phương pháp nghiệm thu sản phẩm | Bộ | 01 | Tài liệu KHCN đầy đủ, chi tiết |
| 4 | Tài liệu thiết kế |  |  |  |
|  | Tài liệu thiết kế hệ thống | Bộ | 01 | Tài liệu KHCN đầy đủ, chi tiết |

**22.3 Dạng III:** Bài báo; Sách chuyên khảo; Sở hữu trí tuệ và các sản phẩm khác

| **TT** | **Dự kiến tạp chí công bố** | **Lĩnh vực** | **Xếp hạng tạp chí** | **Ghi chú** |
| --- | --- | --- | --- | --- |
| 1 | - | - | - | - |
| **TT** | **Dự kiến hội nghị, hội thảo công bố** | **Lĩnh vực** | **Xếp hạng hội nghị, hội thảo** | **Ghi chú** |
| 1 | ICRA | Robotics | A\* (CORE) |  |
| **TT** | **Dự kiến đăng ký sáng chế** | **Lĩnh vực** | **Trong nước/Quốc tế** | **Ghi chú** |
| 1 | Một thuật toán định vị drone bằng phương pháp so khớp với hình ảnh địa hình. | Robotics | Trong nước |  |

**Trình độ khoa học của sản phẩm (Dạng II & III) so với các sản phẩm tương tự hiện có** *(Làm rõ cơ sở khoa học và thực tiễn để xác định các yêu cầu khoa học cần đạt của các sản phẩm của đề tài)*

Dựa trên việc phân tích tình hình nghiên cứu trong và ngoài nước, cách tiếp cận để thực hiện các công nghệ lõi đã trình bày trong thuyết minh đề tài này và các chỉ tiêu kỹ thuật của các sản phẩm đầu ra, nhóm đề tài khẳng định trình độ khoa học của các sản phẩm đăng ký sau khi thực hiện đề tài tương đương với các sản phẩm tiên tiến nhất trên thế giới hiện nay. Ngoài ra, trình độ khoa học của sản phẩm cũng được thể hiện thông qua những kết quả nghiên cứu mới, cách làm mới được cộng đồng chuyên gia đánh giá trên bài báo khoa học quốc tế mà nhóm đề tài đăng ký.

**23. Khả năng ứng dụng và phương thức chuyển giao kết quả nghiên cứu**

**23.1. Khả năng về thị trường**

Về thị trường của thiết bị bay không người lái (Drone), theo Báo cáo Drone Market Size and Forecast 2023-2030 thị trường toàn cầu đạt 33,7 tỷ đô năm 2023, dự kiến sẽ tăng trưởng lên 54,6 tỷ đô vào năm 2030 với tốc độ tăng trưởng (CAGR) trong giai đoạn là 7,1%. Trong nước, theo một báo cáo của Tập đoàn IMARC đưa ra, quy mô thị trường máy bay không người lái Việt Nam đạt 63,8 triệu đô la Mỹ vào năm 2023 và kỳ vọng sẽ đạt 193,2 triệu đô la Mỹ vào năm 2032, CAGR đạt 13,1% trong giai đoạn 2024-2032.

Các ứng dụng chính của Drone tại Việt Nam: Công tác thăm dò, lập bản đồ, giám sát công trường/công trình; Ứng dụng trong nông nghiệp hiện đại sử dụng thiết bị bay không người lái để giúp người nông dân theo dõi mùa vụ, đánh giá tình trạng đất đai và tối ưu hóa thủy lợi.

Bên cạnh drone, thị trường robot tự hành mặt đất ngoài trời (UGV – Unmanned Ground Vehicle) cũng đang chứng kiến sự tăng trưởng mạnh mẽ, đặc biệt trong các lĩnh vực quân sự, an ninh và công nghiệp. Theo báo cáo của Omdia, quy mô thị trường UGV toàn cầu được dự đoán sẽ đạt 1,2 tỷ USD vào năm 2027, với tốc độ tăng trưởng CAGR 20%.

Trong bối cảnh toàn cầu đang chuyển đổi mạnh mẽ dưới tác động của công nghệ không người lái (UAV), nhu cầu đối với các hệ thống chỉ huy – điều khiển UAV tích hợp khả năng phân tích trí tuệ nhân tạo (AI-driven) đang trở thành một yếu tố thiết yếu. Các nền tảng này không chỉ cung cấp chức năng điều phối – giám sát công việc của UAV, mà còn hỗ trợ tổng hợp và xử lý dữ liệu bay, hình ảnh, cảm biến, và bản đồ số trong thời gian thực. Điều này đáp ứng yêu cầu nghiêm ngặt về ra quyết định nhanh, chính xác trong cả môi trường quốc phòng, an ninh và dân sự.

Ứng dụng của các hệ thống này trải dài từ giám sát biên giới, trinh sát chiến thuật, phòng cháy chữa cháy, cứu hộ, đến quản lý đám đông, giám sát đô thị, giao thông thông minh và quản lý hạ tầng cơ sở. Tại những quốc gia đang phát triển như Việt Nam, các giải pháp này không chỉ phải đảm bảo khả năng tùy biến cao theo đặc thù địa lý, điều kiện hạ tầng mà còn cần đảm bảo yếu tố bảo mật và tích hợp dễ dàng vào hệ thống thông tin liên lạc quốc gia.

Theo Verified Market Reports, thị trường phần mềm AI-driven cho drone đạt 1,5 tỷ USD vào năm 2024, và dự kiến tăng lên 5,7 tỷ USD vào năm 2033, với tốc độ tăng trưởng kép hàng năm (CAGR) đạt 16,7 % từ 2026 đến 2033. Con số này cho thấy tiềm năng rất lớn của nền tảng chỉ huy, khi chúng đóng vai trò then chốt trong hệ sinh thái AI UxV, được kỳ vọng mở rộng mạnh mẽ trong những năm tiếp theo.

Tóm lại, thị trường dành cho nền tảng chỉ huy và công nghệ lõi điều khiển cho các phương tiện không người lái đã sẵn sàng để bùng nổ, ứng dụng mạnh mẽ tại Việt Nam. Trong thực tế, VAI đã nhận được văn bản của Bộ tư lệnh Bộ đội Biên phòng với nhu cầu trang bị, điều khiển 450 drone trinh sát cỡ nhỏ phục vụ giám sát vùng biên.

*(Chi tiết về thị trường tại tài liệu MRD đính kèm)*

**23.2. Khả năng về ứng dụng các kết quả nghiên cứu vào sản xuất kinh doanh**

*1) Khả năng cạnh tranh về giá*

Mục tiêu của đề tài là xây dựng một sản phẩm với các tính năng tương đương hoặc đa dạng hơn so với các sản phẩm trong và ngoài nước, đồng thời đảm bảo chất lượng tốt hơn và giá cả thấp hơn so với đối thủ để có lợi thế cạnh tranh, tăng cường số lượng khách hàng sử dụng sản phẩm.

Đồng thời, đề tài cũng nhấn mạnh vào việc đạt chất lượng tốt hơn so với các sản phẩm tương tự và đối thủ cạnh tranh, bằng cách đảm bảo sự hoàn thiện, ổn định và đáp ứng được nhu cầu và mong muốn của khách hàng. Từ đó, đề tài hy vọng thu hút nhiều khách hàng sử dụng sản phẩm và đạt được doanh thu cao.

*2) Khả năng cạnh tranh về chất lượng sản phẩm*

Sản phẩm có tính năng tương đương với các hệ thống tương tự của nước ngoài, nhưng được làm chủ hoàn toàn về công nghệ, do đó đảm bảo tính bảo mật, an toàn an ninh khi sử dụng.

**23.3. Khả năng liên doanh liên kết với các doanh nghiệp trong quá trình nghiên cứu và triển khai ứng dụng sản phẩm**

**Phối hợp nghiên cứu với trường đại học hoặc học viện:** Theo hướng hợp tác với các chuyên gia trong các trường đại học có hướng nghiên cứu liên quan thông qua các buổi hội thảo chuyên đề.

**Phối hợp với khách hàng bay thử nghiệm drone:** Phối hợp với đơn vị BTL BĐBP để thử nghiệm drone, qua đó tối ưu sản phẩm theo nhu cầu của khách hàng.

**23.4. Mô tả phương thức chuyển giao**

Không có.

**24. Phạm vi và địa chỉ (dự kiến) ứng dụng các kết quả của đề tài**

**Bộ tư lệnh Bộ đội Biên phòng:** dự kiến ứng dụng nền tảng chỉ huy thông minh phương tiện không người lái của Viettel trong quản lý giám sát biên giới sử dụng UAV nhỏ gọn.

**25. Tác động và lợi ích mang lại của kết quả nghiên cứu**

* *Hiệu quả an ninh – quốc phòng:* Đề tài góp phần nâng cao năng lực tự chủ công nghệ, đặc biệt trong lĩnh vực chỉ huy điều khiển hệ thống phương tiện không người lái. Đầu ra của đề tài với khả năng bảo mật cao sẽ đáp ứng yêu cầu đặc thù của các nhiệm vụ quốc phòng như tuần tra trinh sát, tấn công cảm tử, cứu hộ/cứu cứu nạn chiến trường. Đồng thời cũng đảm bảo an ninh khi tham gia nền kinh tế tầm thấp.
* *Hiệu quả đối với tổ chức chủ trì:* Góp phần nâng tầm uy tín và năng lực nghiên cứu của VAI nói riêng và Tập đoàn nói chung trong lĩnh vực chỉ huy điều khiển phương tiện không người lái – một ngành khoa học công nghệ tiên phong, phù hợp với xu thế phát triển toàn cầu.
* *Hiệu quả đối với lĩnh vực KHCN*: Nghiên cứu và làm chủ các công nghệ AI tiên tiến trong điều khiển thiết bị không người lái thông minh, đặt nền móng phát triển các hệ thống AI tự hành cho ứng dụng quốc phòng và dân sự.
* *Hiệu quả về Kinh tế:* Dự kiến sẽ kinh doanh nền tảng chỉ huy thông minh các phương tiện không người lái theo dạng license. Giá bán dự kiến là 91.000.000 VNĐ/1 drone/ 1 năm, tương đương 70% giá cho license “Drone as First Responder Class 2” của DroneSense. Tổng chi phí đề tài dự toán 29,5 tỷ VNĐ. Tổng doanh thu từ nền tảng trong 3 năm 2027~ 2029 là 122,85 tỷ VNĐ, lợi nhuận sau thuế là 30,6 tỷ VNĐ. Như vậy, đề tài sẽ hoàn vốn sau 3 năm.

**26. Phương án trang bị thiết bị máy móc để thực hiện và xử lý tài sản được hình thành thông qua việc triển khai thực hiện đề tài**

**26.1. Phương án trang bị tài sản** *(xây dựng phương án, đánh giá và so sánh để lựa chọn phương án hợp lý, tiết kiệm và hiệu quả nhất, hạn chế tối đa mua mới)*

- Đề tài sẽ sử dụng tối đa các trang thiết bị dùng chung đã được trang bị từ những đề tài trước đó trong quá trình thực hiện đề tài.

- Đối với các tải sản phải mua mới nhằm đáp ứng nhu cầu nghiên cứu mới, nhóm đề tài đã liệt kê trong dự toán của đề tài. Các tài sản được mua mới được nhóm đề tài đề xuất dựa trên khối lượng công việc, yêu cầu của các nội dung nghiên cứu và sẽ lựa chọn sản phẩm uy tín, chất lượng từ nhà cung cấp cho chi phí tối ưu nhất theo đúng quy định hiện hành của Tập đoàn.

**26.2. Đề xuất phương án xử lý tài sản trang bị và tài sản là kết quả của quá trình triển khai thực hiện đề tài** *(tổ chức chủ trì đề xuất hình thức xử lý và đối tượng thụ hưởng)*

Các tài sản trang bị và tài sản là kết quả của quá trình triển khai thực hiện đề tài sẽ được đơn vị quản lý, theo dõi sử dụng và xử lý sau khi kết thúc đề tài như sau:

* Đối với các tải sản được trang bị trong quá trình triển khai thực hiện đề tài: Sau khi kết thúc sẽ được bàn giao lại cho tổ chức chủ trì quản lý, tái sử dụng cho các đề tài sau (nếu vẫn còn giá trị khấu hao).
* Đối với các tài sản là kết quả của quá trình triển khai thực hiện đề tài: Sau khi kết thúc sẽ được bàn giao lại cho tổ chức chủ trì quản lý và sử dụng cho việc kinh doanh sản phẩm dịch vụ hoặc sử dụng để làm nền tảng phát triển các công nghệ sản phẩm trong tương lai.

# IV. NHU CẦU KINH PHÍ THỰC HIỆN ĐỀ TÀI VÀ NGUỒN KINH PHÍ *(Giải trình chi tiết xem tại các phụ lục dự toán)*

*Đơn vị tính: VNĐ.*

| **STT** | **NỘI DUNG CÁC KHOẢN CHI** | **Phụ lục tại Dự toán** | **TIỀN HÀNG**  **(Trước thuế GTGT)** | **Thuế GTGT** | **THÀNH TIỀN (Đã bao gồm thuế GTGT)** | **Tỷ lệ (%)** |
| --- | --- | --- | --- | --- | --- | --- |
| **1** | **Chi nhân công** | **Phụ lục 1** | **10.321.157.283** | **0** | **10.321.157.283** | **35%** |
| **2** | **Chi nguyên liệu, vật liệu** | **Phụ lục 2** | **8.204.713.389** | **820.471.339** | **9.025.184.728** | **31%** |
| *2.1* | *Chi nguyên liệu, vật liệu phục vụ nghiên cứu* |  | 8.204.713.389 | 820.471.339 | 9.025.184.728 |  |
| *2.2* | *Chi nguyên liệu, vật liệu phục vụ chế thử sản phẩm* |  | 0 | 0 | 0 | 0 |
| *2.3* | *Chi nguyên liệu, vật liệu phục vụ chế thử công cụ, tài sản cố định* |  | 0 | 0 | 0 | 0 |
| *2.4* | *Chi nguyên liệu, vật liệu thử nghiệm* |  | 0 | 0 | 0 | 0 |
| **3** | **Chi công cụ dụng cụ và tài sản cố định** | **Phụ lục 3** | **806.294.800** | **80.629.480** | **886.924.280** | **3%** |
| *3.1* | *Chi công cụ dụng cụ* |  | 58.680.000 | 5.868.000 | 64.548.000 |  |
| *3.2* | *Chi tài sản cố định* |  | 747.614.800 | 74.761.480 | 822.376.280 |  |
| **4** | **Chi dịch vụ thuê ngoài** | **Phụ lục 4** | **4.190.483.695** | **211.238.696** | **4.401.722.391** | **15%** |
| **5** | **Chi khác phục vụ trực tiếp triển khai nhiệm vụ KH&CN** | **Phụ lục 5** | **2.003.434.139** | **56.300.000** | **2.059.734.139** | **7%** |
| **6** | **Chi quản lý nhiệm vụ KHCN** | **Phụ lục 6** | **1.288.279.165** | **58.431.976** | **1.346.711.141** | **5%** |
| *6.1* | *Chi quản lý chung nhiệm vụ KHCN* |  | 1.276.304.165 | 58.431.976 | 1.334.736.141 |  |
| *6.2* | *Chi đánh giá, nghiệm thu nhiệm vụ KHCN* |  | 11.975.000 | 0 | 11.975.000 |  |
| **7** | **Chi phí dự phòng** |  | **1.340.718.124** | **61.353.575** | **1.402.071.698** | **5%** |
| **TỔNG CỘNG:** | |  | **28.155.080.595** | **1.288.425.065** | **29.443.505.660** |  |
| **Tổng kinh phí thực hiện (đã bao gồm thuế GTGT): 0 đồng.**  *(Bằng chữ: Không đồng./.)* | | | | | | |

| **CHỦ NHIỆM ĐỀ TÀI**  *(Họ tên và chữ ký)* | *Hà Nội, ngày tháng năm 2025*  **ĐƠN VỊ CHỦ TRÌ ĐỀ TÀI**  *(Họ tên, chữ ký và đóng dấu)* |
| --- | --- |

**Phụ lục 01.**

**ĐÁNH GIÁ HIỆU QUẢ KINH TẾ VÀ KHẢ NĂNG THU HỒI CHI PHÍ NGHIÊN CỨU**

| **TT** | **Nội dung** | **Thành tiền** | **Số lượng** | **Tổng tiền** | **Tỷ trọng /doanh thu** |
| --- | --- | --- | --- | --- | --- |
| **A** | **Doanh thu dự kiến** |  |  | **122.850.000.000** |  |
| **B** | **Tổng chi phí tự sản xuất** |  |  | **84.583.505.660** |  |
| **I** | **Giá thành sản xuất** |  |  | **12.142.500.000** |  |
| 1 | Chi phí nhân công nội bộ (5% Doanh thu) |  | 1 | 6.142.500.000 | 5% |
| 2 | Chi phí hạ tầng |  | 1 | 6.000.000.000 |  |
| **II** | **Chi phí nghiên cứu phân bổ** |  |  | **29.443.505.660** |  |
| **III** | **Chi phí khác (chi phí bảo hành, bảo trì, quản lý doanh nghiệp, chi phí tài chính)** |  |  | **42.997.500.000** |  |
| 1 | Chi phí bán hàng | 9.277.500.000 | 1 | **12.285.000.000** | 10% |
| 2 | Chi phí bảo hành | 4.638.750.000 | 1 | **6.142.500.000** | 5% |
| 3 | Chi phí quản lý chung | 1.855.500.000 | 1 | **24.570.000.000** | 20% |
| **C** | **Lợi nhuận kế toán trước thuế** |  |  | **38.266.494.340** |  |
| **D** | **Tỷ suất lợi nhuận trước thuế (%) (Trên CP)** |  |  | **45,24%** |  |
|  | Thuế TNDN 20% |  |  | 7.653.298.868 |  |
| **E** | **Lợi nhuận kế toán sau thuế** |  |  | **30.613.195.472** | **24,92%** |

**Phụ lục 02.**

**BẢNG SO SÁNH VỚI CÁC SẢN PHẨM TRÊN THẾ GIỚI**

**Bảng 3.1. So sánh tính năng bay tự động không GNSS**

| **TT** | **Chỉ tiêu kỹ thuật** | **Đơn vị** | **Sản phẩm của Viettel** | **Sản phẩm của Skydio** | **Sản phẩm của FFaxy** | **So sánh, đánh giá**  **(Tốt hơn, kém hơn)** | **Ghi chú** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | Độ chính xác định vị | m | 10 | 0.1 | 50 | Tương đương |  |
| 2 | Độ cao hoạt động tối đa | m | 300 | 300 | 300 | Tương đương |  |
| 3 | Tốc độ bay tự động tối đa theo phương ngang | m/s | 5 | Không công bố | Không công bố | Tốt hơn |  |
| 4 | Tính năng định vị trong môi trường tối | Có/không | Có | Không | Có | Tương đương |  |
| 5 | Tính năng so khớp ảnh địa hình | Có/không | Có | Không | Không | Tốt hơn |  |
| 6 | Khoảng cách dừng, tránh va chạm gần nhất | m | 100 | Không công bố | Không công bố | Tốt hơn |  |

**Bảng 3.2. So sánh tính năng của hệ thống trung tâm điều hành và xử lý thông tin các phương tiện không người lái**

| **TT** | **Chỉ tiêu kỹ thuật** | **Đơn vị** | **Sản phẩm của Viettel** | **Sản phẩm của DroneSense** | **Sản phẩm của L3Harris-Amorphous** | **So sánh, đánh giá**  **(Tốt hơn, kém hơn)** | **Ghi chú** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | Lập lịch bay và khảo sát nghiệp vụ | Có/không | Có | Có | Có | Tương đương |  |
| 2 | Truyền hình ảnh trực tiếp và chia sẻ video | Có/không | Có | Có | Có | Tương đương |  |
| 3 | Giao diện điều khiển và quản lý (mobile/web) | Có/không | Có | Có | Có | Tương đương |  |
| 4 | Tích hợp được nhiều loại phương tiện không người lái | Có/không | Có | Không có | Có | Tốt hơn |  |
| 5 | Lưu lịch sử, biểu đồ báo cáo thống kê | Có/không | Có | Có | Có | Tương đương |  |

**Bảng 3.3. So sánh tính năng của điều khiển tự động robot ngoài trời**

| **STT** | **Chỉ tiêu/ tính năng** | **Sản phẩm Viettel** | **TS5.0 (Foxtech)** | **SuperDroid GPK‑32** | **Milrem THeMIS (UGV đa nhiệm)** | **Đánh giá** |
| --- | --- | --- | --- | --- | --- | --- |
| 3.1 | Khoảng cách dừng, tránh va chạm gần nhất | 5 | Không công bố | Không có khả năng | Không công bố | Tốt hơn |
| 3.2 | Hỗ trợ dải tốc độ robot (phụ thuộc vào phần cứng robot) | ≤1,3 | ≤ 0.7 | ≤ 1.4 | 5.5 | Kém hơn |
| 3.3 | Sai số bám quỹ đạo theo tín hiệu định vị toàn cầu | ≤20 | Không công bố | Không hỗ trợ | Không công bố | Tốt hơn |

**Đánh giá chung:** Đánh giá mức chất lượng dự kiến đạt được của các sản phẩm (Dạng I) so với các sản phẩm tương tự trong nước và nước ngoài.

**Phụ lục 03.**

**CHỈ TIÊU, TÍNH NĂNG CHIẾN KỸ THUẬT SẢN PHẨM**

**1. Chức năng nhiệm vụ**

Đề tài gồm các thành phần:

* Nền tảng chỉ huy phương tiện bay
* Phần mềm AI tại biên của drone
* Nền tảng phần mềm robot mặt đất

**2. Điều kiện hoạt động**

* Nền tảng chỉ huy thông minh phương tiện bay:
  + Tốc độ mạng: tối thiểu 20 Mbps
  + Thời tiết: Trời quang mây, tầm nhìn xa 10km, vận tốc gió tối đa 5 m/s
  + Tín hiệu GNSS: tín hiệu GNSS tốt, tối thiểu nhận được tín hiệu từ 25 vệ tinh
  + Độ rộng của môi trường khám phá: tối thiểu 3000 m2.
* Xử lý tại biên:
  + Khoảng cách tránh vật cản: cách xa tối thiểu 100m
  + Kích thước vật cản phát hiện tối thiểu 1x0.3m
  + Thời tiết: Trời quang mây, tầm nhìn xa 10km, vận tốc gió tối đa 10 m/s
  + Tín hiệu GNSS: tín hiệu GNSS tốt, tối thiểu nhận được tín hiệu từ 25 vệ tinh
  + Khoảng cách bám bắt đối tượng: tối đa 100m
  + Đối tượng bám bắt di chuyển liên tục trên đoạn đường 100m, không bị che khuất quá 50% diện tích
* Robot mặt đất:
  + Độ dốc địa hình tối đa: 30o
  + Kích thước vật cản phát hiện tối thiểu 0.2x0.2m
  + Tín hiệu GNSS: tín hiệu GNSS tốt, tối thiểu nhận được tín hiệu từ 25 vệ tinh

**3. Tính năng**

* Tính năng định vị, dẫn đường trong môi trường không có GPS.
* Tính năng bám bắt đối tượng di chuyển.
* Tính năng ước lượng vị trí đối tượng trong không gian.
* Tính năng điều khiển, điều phối bầy đàn
* Tính năng tránh va chạm trong khi thực hiện nhiệm vụ.

**4. Chỉ tiêu chiến kỹ thuật**

| **STT** | **Tên hạng mục** | **Đơn vị đo** | **Chỉ tiêu** | **Phương pháp đánh giá** |
| --- | --- | --- | --- | --- |
| **1** | **Chỉ tiêu cho nền tảng chỉ huy thông minh phương tiện bay** | | | |
| 1.1 | Số lượng drone tối đa có thể điều phối đồng thời | Chiếc | ≥5 | Thử nghiệm thực tế trong bài toán khám phá địa hình |
| 1.2 | Số lượng tối đa phương tiện mà kiến trúc nền tảng về mặt lý thuyết có khả năng quản lý điều khiển | Chiếc | ≥1000 | Lập testcase kiểm tra hệ thống quản lý phương tiện bay, số lượng, thông tin các phương tiện bay |
| 1.3 | Thời gian trễ khi truyền dữ liệu | s | ≤ 0,5 | Đo thời gian truyền dữ liệu ảnh chụp từ camera gửi tới hệ thống quản lý |
| 1.4 | Hiệu suất phân công nhiệm vụ (tỷ lệ phương tiện hoàn thành nhiệm vụ) | % | ≥ 80 | Định nghĩa chuỗi hành động của một nhiệm vụ bao gồm:  *Cất cánh từ điểm thứ nhất → bay tới điểm thứ hai → quay/ chụp dữ liệu → hạ cánh ở điểm thứ ba*  Đội hình drone tự động phân công và thực hiện nhiệm vụ. Tỉ lệ hoàn thành nhiệm vụ được tính bằng: |
| 1.5 | Độ lệch đội hình trung bình (khi có tín hiệu GNSS) | m | ≤ 2 | Đặt các vị trí đích của các drone trong đội hình di chuyển. So sánh dữ liệu vị trí đã thiết lập với vị trí thực tế của các drone để tính giá trị sai số trung bình |
| 1.6 | Tỷ lệ bao phủ không gian (trong bài toán khám phá địa hình) | % | ≥ 90 | Chia bản đồ cần khám phá thành các đơn vị nhỏ nhất (cell)  Thực hiện quét bản đồ bằng đội hình 5 drone và kiểm tra số lượng cell đã có thông tin sau khi quét. Tính tỉ lệ bằng công thức: |
| **2** | **Chỉ tiêu cho phần mềm tại biên của drone** | | | |
| 2.1 | Khoảng cách dừng, tránh va chạm gần nhất | m | 100 | Kiểm tra khả năng tránh va chạm trong phạm vi 100m ở trạng thái bay tự động, với các vật thể như cột, dây điện,... |
| 2.2 | Hỗ trợ dải tốc độ drone | m/s | ≤5 | Kiểm tra vận tốc lớn nhất 5m/s khi hoạt động ở trạng thái tự động |
| 2.3 | Hỗ trợ dải tốc độ đối tượng có thể bắt bám | m/s | ≤5 | Kiểm tra khả năng bám bắt đối tượng di chuyển ở vận tốc 5m/s |
| 2.4 | Sai số ước lượng vị trí đối tượng trong không gian (khoảng cách 50m) | m | ≤ 2 | Đặt vị trí đối tượng tại một điểm đã được xác định vị trí chính xác trong không gian.  So sánh vị trí của đối tượng tính toán bằng camera với vị trí của đối tượng tính toán bằng RTK |
| 2.5 | Độ cao lớn nhất có thể tự động bay không cần tín hiệu định vị toàn cầu GNSS | m | 300 | Thiết lập lộ trình bay tự động cho drone ở độ cao 300m.  So sánh độ lệch định vị của drone qua tín hiệu định vị toàn cầu. |
| 2.6 | Độ lệch định vị cục bộ (ở độ cao 100 m, không tín hiệu GNSS) | m | ≤10 | Thiết lập lộ trình bay tự động cho drone ở độ cao 100m.  So sánh độ lệch định vị của drone qua tín hiệu định vị toàn cầu. |
| 2.7 | Độ chính xác bắt bám đối tượng | % | ≥70% | Đo chỉ số MOTA  Trong đó:  - FP: Số lượng false positives (lỗi phát hiện sai đối tượng)  - FN: Số lượng false negatives (lỗi bỏ sót đối tượng)  - IDS: Số lần hoán đổi ID đối tượng (lỗi nhầm lẫn danh tính đối tượng)  - GT: Số lượng đối tượng thực tế (Ground Truth), tức là số lượng đối tượng trong dữ liệu gốc |
| **3** | **Chỉ tiêu cho phần mềm tại biên của robot** | | |  |
| 3.1 | Khoảng cách dừng, tránh va chạm gần nhất | m | 5 | Kiểm tra khả năng tránh va chạm trong phạm vi 5m bằng các cảm biến như lidar, camera, sonar |
| 3.2 | Hỗ trợ dải tốc độ robot (phụ thuộc vào phần cứng robot) | m/s | ≤1,3 | Kiểm tra khả năng định vị, điều hướng, lập bản đồ ở vận tốc lớn nhất 1.3m/s |
| 3.3 | Sai số bám quỹ đạo theo tín hiệu định vị toàn cầu | cm | ≤20 | So sánh giữa quỹ đạo thiết lập và quãng đường di chuyển thực tế để tính sai số bám quỹ đạo |

**5. Tiêu chuẩn áp dụng**

Quá trình nghiên cứu đề tài và thử nghiệm sản phẩm tuân theo quy định về việc cấp phép bay đối với thiết bị bay không người lái tại Việt Nam.