

PHÁT TRIỂN PHƯƠNG PHÁP XÁC ĐỊNH VẬT CẢN CHO DRONE SỬ DỤNG YOLOV8

Nguyễn Phương Nam, Hứa Ngọc Nhi và Nguyễn Đình Tú

Khoa Kỹ thuật Cơ khí, Trường Đại học Kỹ thuật – Công nghệ Cần Thơ

Email:ndtu@ctu.edu.vn

Thông tin chung:

Ngày nhận bài: 10.01.2024

Ngày nhận bài sửa: 15.4.2024

Ngày duyệt đăng: 17.4.2024

Từ khóa:

Drone, học sâu, xử lý ảnh, IoTs, YOLOv8

TÓM TẮT

Hiện nay, ứng dụng của drone trong lĩnh vực Internet vạn vật (IoTs) ngày càng trở nên rộng rãi vì tính thuận tiện cho việc thực hiện nhiệm vụ giám sát. Drone có thể được tích hợp với nhiều cảm biến và được sử dụng trong nhiều lĩnh vực khác nhau. Tuy nhiên, để hoàn thành một hay nhiều tác vụ thì việc điều khiển drone chính xác và thông minh là cần thiết. Vì thế, nghiên cứu đề xuất một phương pháp điều khiển drone mới để xác định và tránh vật cản sử dụng kết hợp giữa giải thuật xử lý ảnh thời gian thực và phương pháp học sâu dựa trên cấu trúc của YOLOv8 với độ chính xác trên 90%.

1. ĐẶT VẤN ĐỀ

Trong thời đại công nghệ 4.0, thiết bị bay không người lái (drone) đang được ứng dụng rộng rãi từ lĩnh vực quân sự đến dân dụng. Nhờ vào chi phí thấp và khả năng cơ động cao nên drone đã trở thành công cụ phổ biến trong nhiều lĩnh vực như nông nghiệp (Mogili và cộng sự, 2018), dịch vụ giao nhận (Lee, 2017), giám sát (Ding và cộng sự, 2018), làm phim (Fleureau và cộng sự, 2016) và nhiếp ảnh (Puttock và cộng sự, 2015).

Để điều khiển được drone thực hiện các tác vụ nêu trên, thì việc thiết kế mô hình tương tác người – robot (HRI) rất cần thiết. Hiện nay, HRI đang chuyển sang xu hướng sử dụng các động tác/ngôn ngữ cơ thể để điều khiển các thiết bị/máy móc (Shyh-Wei Chen và cộng sự, 2022), (Farhat Naseer và cộng sự, 2022). Nhiều ứng dụng đã được triển khai thực tế dựa trên sự tương tác này (Salvine và cộng sự, 2011). Hơn nữa, việc sử dụng HRI vào việc điều khiển drone cũng thu hút được nhiều sự quan tâm gần đây trong nhiều lĩnh vực khác nhau như: chụp ảnh trên không và trinh sát các

khu vực, địa điểm xảy ra thảm họa (Erat và cộng sự, 2018). Đối với phương pháp điều khiển theo các mô hình HRI, drone được điều khiển thông qua bộ điều khiển từ xa hoặc thiết bị cầm tay sử dụng các nút ấn định hướng vật lý/cảm ứng hoặc sử dụng động tác/ngôn ngữ cơ thể để di chuyển theo 4 hướng: tiến, lùi, trái, và phải. Ngoài ra, một số ứng dụng cho phép điều khiển drone bằng núm xoay một cách thuận tiện. Tuy nhiên, đối với phương pháp truyền thống này đòi hỏi người vận hành cần tuân thủ các thao tác chính xác và phức tạp.

Trong quá trình hoạt động và thực hiện các nhiệm vụ chỉ định sẵn, drone phải tránh va chạm với các vật cản trong môi trường hoạt động. Để khắc phục vấn đề này, người vận hành drone cần được đào tạo và thực nghiệm nhiều lần. Vì vậy, việc đề xuất giải pháp tự động cho drone thực hiện các tác vụ có thể giảm bớt yêu cầu giám sát và điều khiển liên tục từ người vận hành, tiết kiệm thời gian và chi phí. Nhiều nghiên cứu đã dựa vào cảm biến LiDAR (Ramasamy và cộng sự, 2016), (Ferrick và cộng sự, 2012), cảm biến siêu âm

(Gageik và cộng sự, 2015), và stereo camera (Iacono và cộng sự, 2018) để xây dựng phương pháp tránh vật cản cho drone. Ngoài ra, Alvarez và cộng sự (2016) đề xuất phương pháp tránh vật cản sử dụng duy nhất một camera cho drone. Tuy nhiên, việc sử dụng duy nhất một camera khi không có được thông tin chiều sâu của ảnh sẽ rất khó để xác định chính xác vật cản và phương pháp này có độ chính xác thấp hơn phương pháp sử dụng stereo camera.

Với sự phát triển của phương pháp học sâu - deep learning (DL), thì việc ứng dụng phương pháp học sâu vào các lĩnh vực như nhận dạng đối tượng, giọng nói, định vị và điều khiển drone càng trở nên dễ dàng hơn, đạt được nhiều kết quả chính xác (Fraga-Lamas và cộng sự, 2019). Cụ thể, mạng nơ-ron nhân tạo tích chập - convolutional neural network (CNN) đã được ứng dụng thành công vào các lĩnh vực thị giác máy tính như nhận dạng đối tượng và định vị đối tượng. Phương pháp này cho thấy tính khả thi khi áp dụng vào tác vụ tránh vật cản trên drone. Với sự hỗ trợ của phương pháp học sâu - mạng nơ-ron tích chập, tính chính xác của việc xác định vị trí của vật

cản sẽ được cải thiện trong quá trình nhận dạng đối tượng với duy nhất một camera trên drone.

Nghiên cứu này đề xuất một phương pháp điều khiển kết hợp với giải thuật tránh vật cản sử dụng mạng nơ-ron tích chập được xây dựng tối ưu trên YOLOv8 (You Only Look Once version 8). Phiên bản YOLOv8 đã hỗ trợ các tác vụ về thị giác máy tính bao gồm nhận dạng đối tượng (detection), phân đoạn đối tượng (segmentation), theo dõi (tracking), ước tính cử động của đối tượng (pose estimation), và phân loại (classification). Mục tiêu của nghiên cứu hướng tới cải thiện độ chính xác trong việc điều khiển xác định vật cản của drone khi sử dụng duy nhất một camera.

2. PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Mini drone DJI Tello

Nghiên cứu sử dụng mini drone DJI Tello. Cụ thể, mini drone DJI Tello sử dụng vi xử lý Intel 14 – lõi, với trọng lượng chỉ 80g bao gồm pin Li-Po – 1100 mAh, 3,8 V, thể hiện ở Hình 1. Thời gian hoạt động tối đa cho mini drone DJI Tello lên đến 13 phút và có thể điều khiển bằng máy tính thông qua kết nối Wi-Fi.



Hình 2. Mini drone DJI Tello

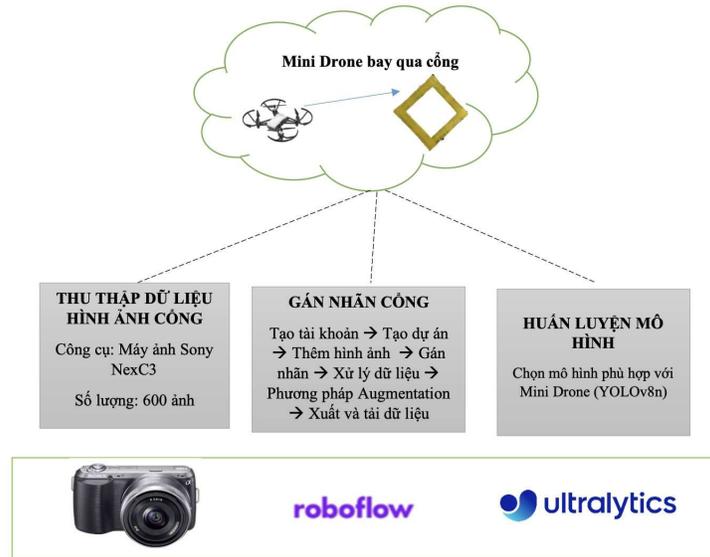
Nguồn: Ryze Robotics, Tello – Feel the fun, <https://www.ryzerobotics.com/tello>

2.2. Điều khiển mini drone xác định và di chuyển qua vật cản

2.2.1. Thiết lập mini drone tello di chuyển qua vật cản

Việc thiết lập giúp mini drone nhận dạng được vật cản và di chuyển qua các vật cản

khác mang ý nghĩa quan trọng trong thời đại chuyên đổi kỹ thuật số, đóng góp sự sáng tạo và đổi mới trong nghiên cứu và khoa học. Các bước phương pháp thực hiện việc giúp mini drone bay qua vật cản được trình bày trong Hình 2.



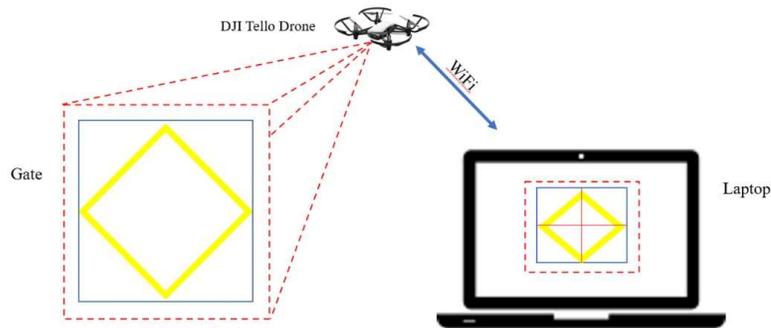
Hình 2. Các bước thiết lập mini drone nhận dạng và di chuyển qua vật cản

Hình 2 trình bày các bước thiết lập ban đầu từ quá trình thu thập dữ liệu đến quá trình huấn luyện mô hình YOLOv8. Cụ thể, số lượng ảnh của đối tượng ban đầu là 600 ảnh, tác giả đã sử dụng công cụ Roboflow để thực hiện các tác vụ gán nhãn, tăng cường số lượng hình ảnh gốc với phương pháp Augmentation – số lượng ảnh sau khi áp dụng phương pháp Augmentation là 1000 ảnh. Việc áp dụng phương pháp Augmentation sẽ làm tăng tính đa dạng trong dữ liệu và giúp mô hình máy học trong việc tổng quát hóa dữ liệu, hạn chế được trường hợp quá khớp (overfitting). Mô hình YOLOv8n được sử dụng để huấn luyện trong nghiên cứu này được hỗ trợ trên

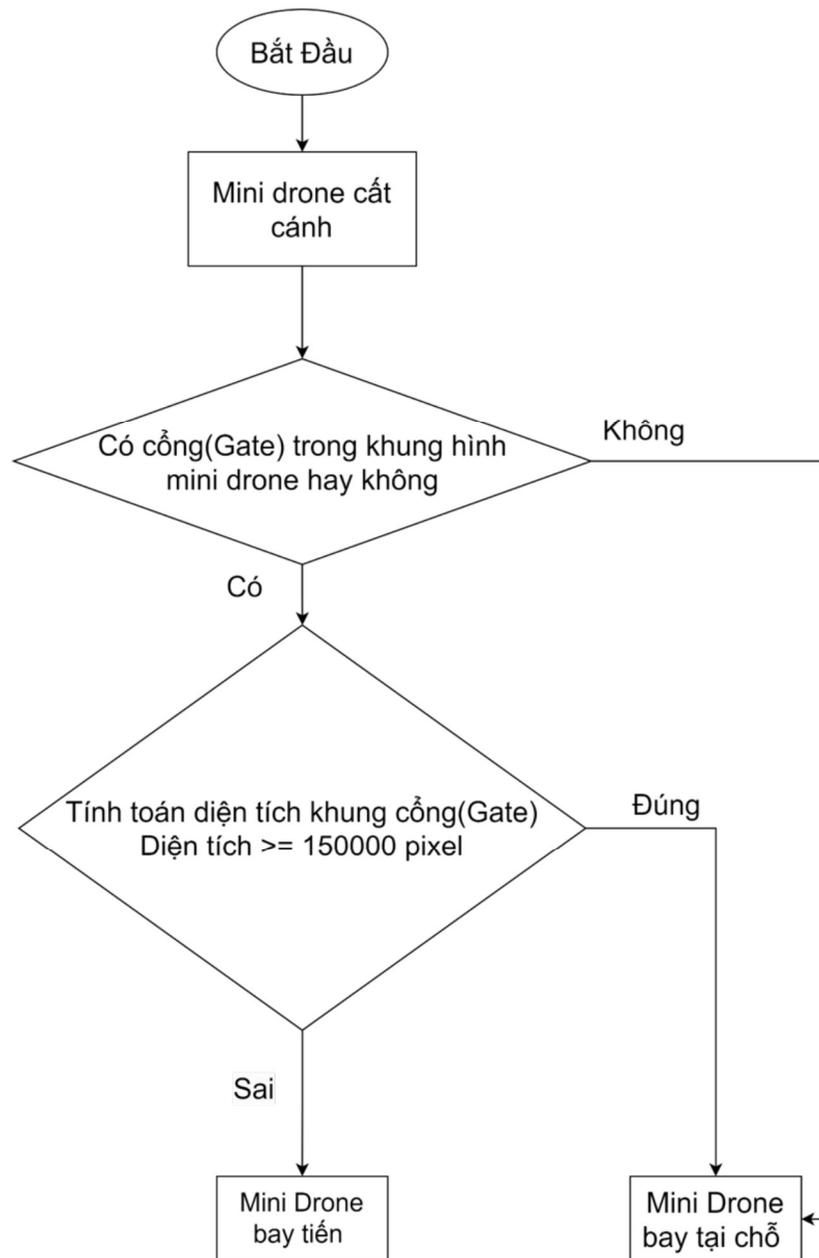
Ultralytics HUB (Nguồn: Ultralytics: ultralytics [Online]. Available from: <https://github.com/ultralytics/ultralytics>). Việc lựa chọn mô hình YOLOv8n cho nghiên cứu là vì tính phù hợp với cấu hình máy tính và phần cứng đề xuất.

2.2.2. Các phương pháp tính toán

Kiến trúc của hệ thống điều khiển mini drone được mô tả trong Hình 3. Camera được trang bị trên DJI Tello sẽ kết nối với máy tính, thông qua chuẩn Wi-Fi. Sau đó, dữ liệu được đồng bộ sẽ gửi đến mini drone, đồng nghĩa với việc drone đã nhận được tín hiệu và tiến hành tích hợp, thực hiện các lệnh mà người vận hành yêu cầu.



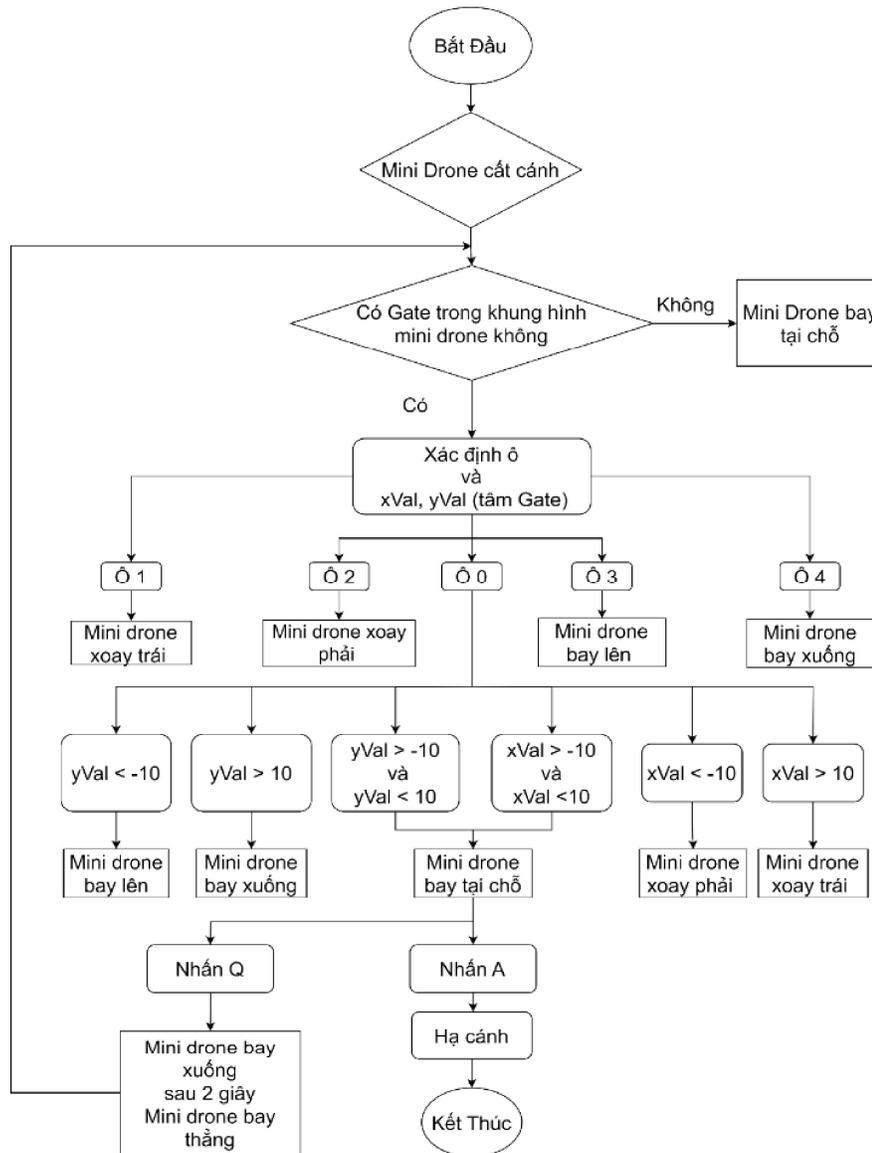
Hình 3. Sơ đồ điều khiển drone sử dụng tính năng theo dõi vật cản



Hình 4. Lưu đồ quá trình tính toán giữ khoảng cách với vật cản

Lưu đồ tính toán giữ khoảng cách với vật cản được thể hiện ở Hình 4. Quá trình giữ khoảng cách với vật cản này diễn ra song hành cùng lưu đồ điều khiển mini drone di chuyển qua vật cản (Hình 5). Đầu tiên, khung hình trên camera mini drone (kích thước là

480×640 pixel) được chia thành 9 ô tương ứng với số thứ tự 0 đến 4 và chỉ khi vật cản nằm ở vị trí ô 0 thì lúc này mini drone mới thực hiện điều chỉnh xác định theo tâm của vật cản bằng tham số PID (Proportional-Integral-Derivative).

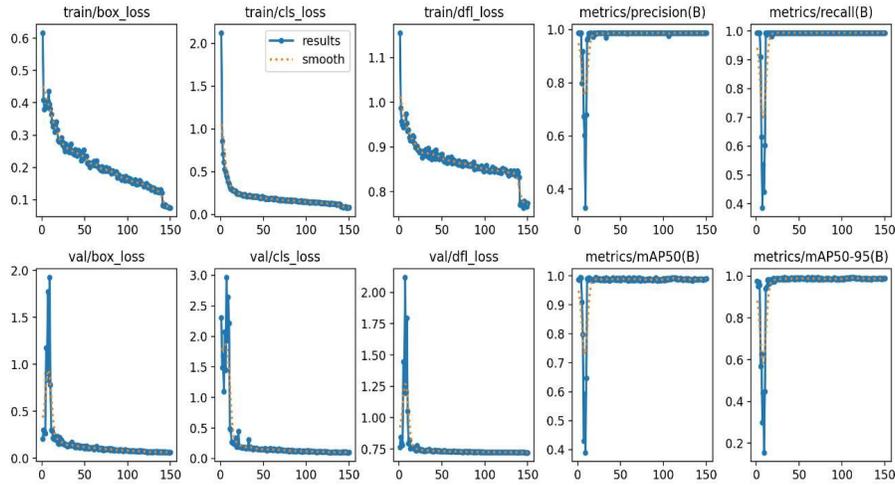


Hình 5. Lưu đồ quá trình điều khiển mini drone di chuyển qua vật cản

3. KẾT QUẢ VÀ THẢO LUẬN

Việc đánh giá hiệu suất của của mô hình YOLOv8n sử dụng các thuật ngữ tiêu chuẩn hoặc phép đo lường sau: mAP50 (Mean Average Precision at 50) đo lường độ chính xác trung bình của mô hình trong việc phát hiện đối tượng với độ tin cậy trên 50%. mAP50 - 95 (Mean Average Precision from 50 to 95): mAP50 - 95 là một độ đo tương tự, nhưng nó tính toán giá

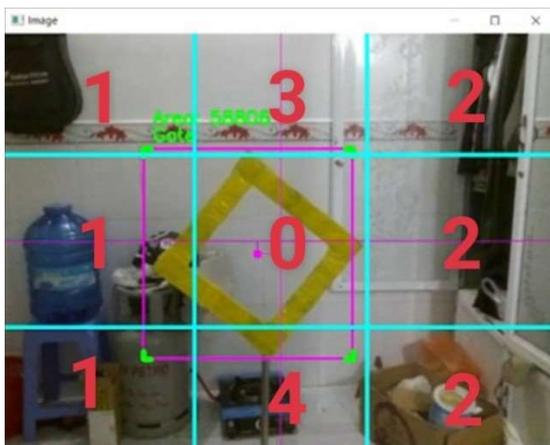
trị mAP trong khoảng từ 50% đến 95% độ tin cậy của mô hình. Precision (độ chính xác) đo lường tỷ lệ số lượng dự đoán đúng (true positives) trên tổng số lượng dự đoán tính được. Thông số này chỉ ra khả năng của mô hình trong việc không gây ra sai dự đoán dương (false positive). Recall (độ nhớ lại): Recall đo lường tỷ lệ số lượng dự đoán đúng (true positives) trên tổng số lượng thực tế tính được.



Hình 6. Giá trị của hàm mất mát trong các tập dữ liệu huấn luyện (train) và tập dữ liệu kiểm chứng (val) trong 150 vòng lặp

Các giá trị Precision và Recall thường được sử dụng để đánh giá hiệu suất của mô hình phát hiện đối tượng ở các ngưỡng độ tin cậy khác nhau. Mô hình YOLO thường sẽ tạo ra nhiều dự đoán ở các ngưỡng khác nhau và đánh giá Precision và Recall tại mỗi ngưỡng này để xác định ngưỡng tốt nhất cho từng ứng dụng cụ thể. Sau 150 vòng lặp, các giá trị của Precision, Recall, mAP50, và mAP50-95 lần lượt là 0,988; 0,994; 0,988 và 0,988 (Hình 6).

Hình 7 trình bày kết quả nhận dạng vật cản sử dụng mô hình YOLOv8. Việc chia khung hình thành nhiều ô nhằm đảm bảo vật cản nằm giữa khung hình hỗ trợ cho việc mini drone di chuyển chính xác. Khi mini drone ở vị trí như Hình 7 thì bộ điều khiển PID được sử dụng trong quá trình di chuyển của mini drone. Bên cạnh đó, hướng di chuyển của drone (Hình 8) bao gồm: trái, phải, lên, xuống cũng được sử dụng dựa trên thông tin từ việc đối vật cản và sự hiệu chỉnh của hệ thống PID.

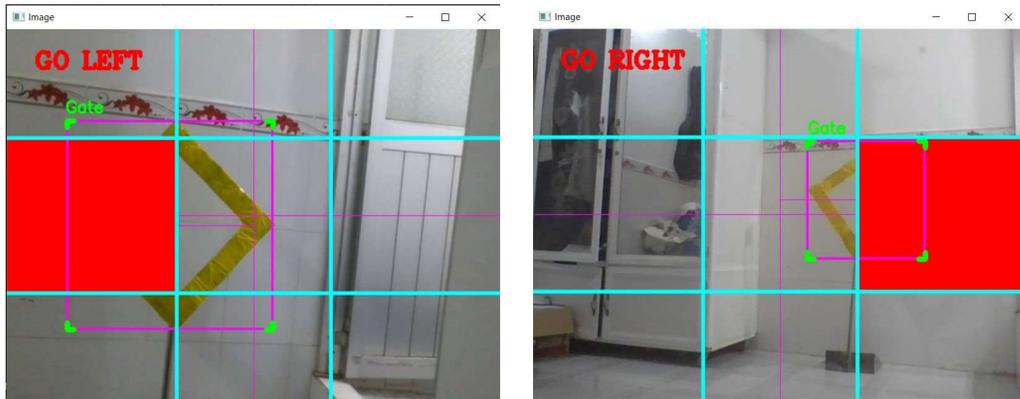


Hình 7. Các vùng ô đã được đánh dấu theo thứ tự (camera mini drone Tello)

Bảng 1. Giá trị PID đã được thiết lập trên mini drone

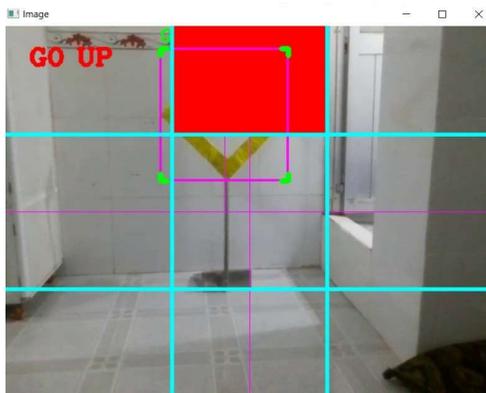
Trục	Tỷ lệ (P)	Tích phân (I)	Vi phân (D)
x	0,22	0	0,1
y	0,27	0	0,1

Các giá trị P, I và D là các hệ số trong bộ điều khiển PID, mỗi giá trị ứng với phần tương ứng của bộ điều khiển PID: tỷ lệ (Proportional), tích phân (Integral), và vi phân (Derivative). Bảng 1 thể hiện giá trị của bộ điều khiển PID được sử dụng trong quá trình điều khiển và điều chỉnh mini drone theo các trục x, y.



(a)

(b)



(c)



(d)

Hình 8. Hình ảnh vị trí tâm của vật cản trên camera DJI Tello (a) vị trí ô 1; (b) vị trí ô 2; (c) vị trí ô 3; (d) vị trí ô 4

Bảng 2. Kết quả thực nghiệm quá trình phát hiện vị trí tâm của vật cản sử dụng với camera trên DJI Tello

Số lần thử nghiệm	Vị trí ô	Phát hiện đúng	Phát hiện sai	Tỷ lệ chính xác
50	0	45	5	90%
50	1	45	5	90%
50	2	46	4	92%
50	3	49	1	98%
50	4	49	1	98%

Tuy nhiên nghiên cứu đã đạt được các kết quả nhất định, bao gồm phát triển thuật toán và hệ thống điều khiển cho mini drone để bay qua vật cản được xác định trước. Mini drone đã có khả năng phát hiện vật cản và thực hiện các động tác điều khiển để bay qua một cách an toàn. Bảng 2 trình bày kết quả drone phát hiện vị trí tâm của vật cản như tiêu đề của bảng chứ không trình bày kết quả drone bay qua vật cản một cách an toàn? Kết quả phát hiện vật cản và bay qua vật cản an toàn hoàn toàn có thể khác nhau??? Ngoài ra, mini drone đã xác định chính xác tọa độ và hướng của vật cản. Tuy nhiên các yếu tố khác vẫn cần được xem xét, các trường hợp cụ thể như: điều kiện thời tiết, ánh sáng môi trường đều có thể gây ảnh hưởng trực tiếp đến khả năng nhận dạng hình ảnh của vật cản.

4. KẾT LUẬN VÀ HÀM Ý QUẢN TRỊ

Nghiên cứu đã thành công trong việc thiết kế phương pháp di chuyển cho mini drone khi gặp vật cản sử dụng duy nhất 1 camera trên DJI Tello. Việc phát triển hệ thống điều khiển thành công sử dụng mini drone DJI Tello kết hợp mô hình YOLOv8 cho thấy kết quả nhận dạng vật cản với độ chính xác trên 90%, góp phần cung cấp giải pháp hiệu quả cho các nhiệm vụ phức tạp cho người vận hành khi gặp vật cản hoặc cần vượt qua chướng ngại vật một cách tự động. Đối với việc mở rộng trong tương lai, nghiên cứu này có thể mở rộng tích hợp thêm phương pháp học tăng cường để cải thiện khả năng tự động của mini drone nhằm mục tiêu thích nghi với môi trường và tình huống thay đổi.

Tài liệu tham khảo

Alvarez, H., Paz, L. M., Sturm, D. and Cremers, D. (2016), "Collision Avoidance for quadrotors with a monocular camera", *Experimental Robotics*, Springer, Cham, pp. 195-209.

Chen, S. W., Lai, Y. C., Tsai, C. T., Liu, C. H. and Tu, J. F. (2022), "Development of

intelligent drone remote control system based on internet of things", *Sensors and Materials*, Vol. 34, No. 7, pp. 2581-2589.

Ding, G., Wu, Q., Zhang, L., Lin, Y., Tsiftsis, T. A. and Yao, Y. (2018), "An amateur drone surveillance system based on the cognitive internet of things", *IEEE Communications Magazine*, Vol. 56, No. 1, pp. 29-35.

Erat, O., Isop, W., Kalkofen, A. D. and Schmalstieg, D. (2018), "Drone – augmented human vision: Exocentric control for drones exploring hidden areas", *IEEE Transactions on Visualization and Computer Graphics*, Vol. 24, No. 4, pp. 1437-1446.

Ferrick, A., Fish, J., Venator, E. and Lee, G. S. (2012), "Uav obstacle avoidance using image processing techniques", *IEEE international conference on technologies for practical robot applications (TePRA)*, Massachusetts, USA, IEEE, pp. 73-78.

Fleureau, J., Galvane, Q., Tariolle, F. L. and Guillotel, P. (2016), "Generic drone control platform for autonomous capture of cinema scenes", *Proceedings of the 2nd workshop on micro aerial vehicle networks, systems, and applications for civilian use*, Singapore, Association for Computing Machinery, pp. 35-40.

Fraga-Lamas, P., Ramos, L., Mondejar-Guerra, V. and Fernandez-Carames, T. M. (2019), "A review on IoT deep learning UAV systems for autonomous obstacle detection and collision avoidance", *Remote Sensing*, Vol. 11, No. 18, 2144.

Gageik, N., Benz, P. and Montenegro, S. (2015), "Obstacle detection and collision avoidance for a UAV with complementary low-cost sensors", *IEEE Access*, Vol. 3, pp. 599-609.

Iacono, M. and Sgorbissa, A. (2018), "Path following and obstacle avoidance for an autonomous UAV using a depth camera",

Robotics and Autonomous Systems, Vol. 106, pp. 38-46.

Salvine, P., Nicolescu, M. and Ishiguro, H. (2011), “Benefits of human – robot interactions”, IEEE Robotics & Automation Magazine, Vol. 18, No. 4, pp. 98-99.

Lee, J. (2017), “Optimization of a modular drone delivery system”, Annual IEEE international systems conference (SysCon), Montreal, Canada, IEEE, pp. 1-8.

Mogili, U. R. and Deepak, B. B. V. L. (2018), “Review on application of drone systems in precision agriculture”, Procedia Computer Science, Vol. 133, pp. 502-509.

Naseer, F., Ullah, G., Siddiqui, M. A., Khan, M. J. and Hong, K. S. (2022), “Deep learning-based unmanned aerial vehicle control with hand gesture and computer vision”, Republic of Korea, IEEE, pp. 1-6.

Puttock, A. K., Cunliffe, A. M., Anderson, K. and Brazier, R. E. (2015), “Aerial photography collected with a multicopter drone reveals impact of eurasian beaver reintroduction on ecosystem structure”, Journal of Unmanned Vehicle Systems, Vol. 3, No. 3, pp. 123-130.

Ramasamy, S., Sabatini, R., Gardi, A. and Liu, J. (2016), “LIDAR obstacle warning and avoidance system for unmanned aerial vehicle sense-and-avoid”, Aerospace Science and Technology, Vol. 55, pp. 344-358.

Ryze Robotics, Tello – Feel the fun, <https://www.ryzerobotics.com/tello>, 2023 (accessed 15 July 2023).

Ultralytics: ultralytics [Online]. Available from: <https://github.com/ultralytics/ultralytics> (accessed 01 January 2024).

DEVELOPING AN OBSTACLE AVOIDANCE METHOD FOR DRONE USING YOLOV8

ABSTRACT

Recently, the use of drones in the Internet of Things (IoTs) is becoming increasingly widespread due to their convenience in performing surveillance tasks. Drones can be integrated with various sensors and utilized in many different fields. However, precise and intelligent drone control is essential to successfully complete one or multiple tasks. Therefore, this study proposes a new method for drones to determine and avoid obstacle, using a combination of real-time image processing algorithms and deep learning methods based on the structure of YOLOv8 to achieve an accuracy of over 90%.

Keywords: Drones, deep learning, image processing, IoTs, YOLOv8