

# ỨNG DỤNG TRÍ TUỆ NHÂN TẠO ĐỂ PHÁT HIỆN BẤT THƯỜNG TRONG GIÁM SÁT RỪNG

Vũ Quang Vinh\*, Trần Anh Đạt†, Trần Tiến Công\*

\* Học Viện Công Nghệ Bưu Chính Viễn Thông

† Đại Học Thủy Lợi

**Tóm tắt**—Giám sát rừng là việc giám sát các chủ thể và thiết chế, ra quyết định vận hành, phản ứng một cách kịp thời và phù hợp với vai trò, trách nhiệm của các bên liên quan trong công tác bảo tồn, phát triển rừng. Việt Nam với tỉ lệ diện tích rừng hơn 40%, nhiệm vụ giám sát và bảo vệ rừng càng đặc biệt quan trọng. Mặc dù đã có các ứng dụng công nghệ thông tin hỗ trợ, công tác giám sát vẫn còn phụ thuộc nhiều vào năng lực chuyên môn của lực lượng Kiểm lâm quốc gia. Trong bài báo này, chúng tôi nghiên cứu ứng dụng giúp phát hiện và cảnh báo sớm cháy rừng, một loại bất thường trong giám sát rừng, bằng công nghệ trí tuệ nhân tạo (AI), trong đó hướng tiếp cận là phát hiện đối tượng trong video quay từ các camera tầm cao đặt tại khu vực rừng. Cơ chế của hệ thống là sử dụng kỹ thuật học máy trong các lĩnh vực xử lý ảnh và thị giác máy tính như phát hiện đối tượng, phân vùng ảnh. Qua thực nghiệm, hệ thống phát hiện và cảnh báo cháy rừng đạt kết quả vượt trội với độ đo mAP@0.50 lên tới 0.89 trên tập thử nghiệm và độ trễ thấp 10ms.

**Từ khóa**—Giám sát rừng, phát hiện khói cháy, phát hiện đối tượng, bảo vệ rừng

## I. GIỚI THIỆU

Rừng đóng một vai trò quan trọng trong sự cân bằng sinh thái của trái đất. Tuy nhiên, những tài nguyên thiên nhiên này đang bị đe dọa bởi thiên tai như hỏa hoạn, nạn xây dựng trái phép, chặt phá rừng... Những năm gần đây, thảm họa cháy rừng diễn ra thường xuyên hơn do tình trạng khô

hạn và tần suất các đợt nắng nóng do hiện tượng nóng lên toàn cầu ngày càng tăng. Điển hình như vào đầu tháng 8 năm 2023, một loạt vụ cháy rừng đã bùng phát ở bang Hawaii của Hoa Kỳ, chủ yếu là trên đảo Maui làm ít nhất 115 người tử vong, 1.294 ha khu vực bị ảnh hưởng. Thiệt hại về tài sản do vụ cháy gây ra ước tính khoảng 5,52 tỷ USD.

Cháy rừng không những gây thiệt hại nghiêm trọng tức thời mà công tác khắc phục hậu quả rất tốn kém thời gian và chi phí. Nhận thức được tác động nghiêm trọng của cháy rừng đối với an toàn nhân loại, nền kinh tế, hệ sinh thái và cơ sở hạ tầng, các công nghệ, kỹ thuật phát hiện và cảnh báo sớm cháy rừng liên tục được nghiên cứu, phát triển. Trước đây, cháy rừng hầu hết được phát hiện và ngăn chặn bởi con người. Ngày nay, công tác giám sát bảo vệ rừng hiệu quả hơn nhiều nhờ sự trợ giúp từ các ứng dụng khoa học và công nghệ cao. Trong đó, hướng tiếp cận phát hiện cháy rừng bằng trí tuệ nhân tạo với các kỹ thuật thị giác máy tính được đặc biệt quan tâm những năm gần đây [1], [2].

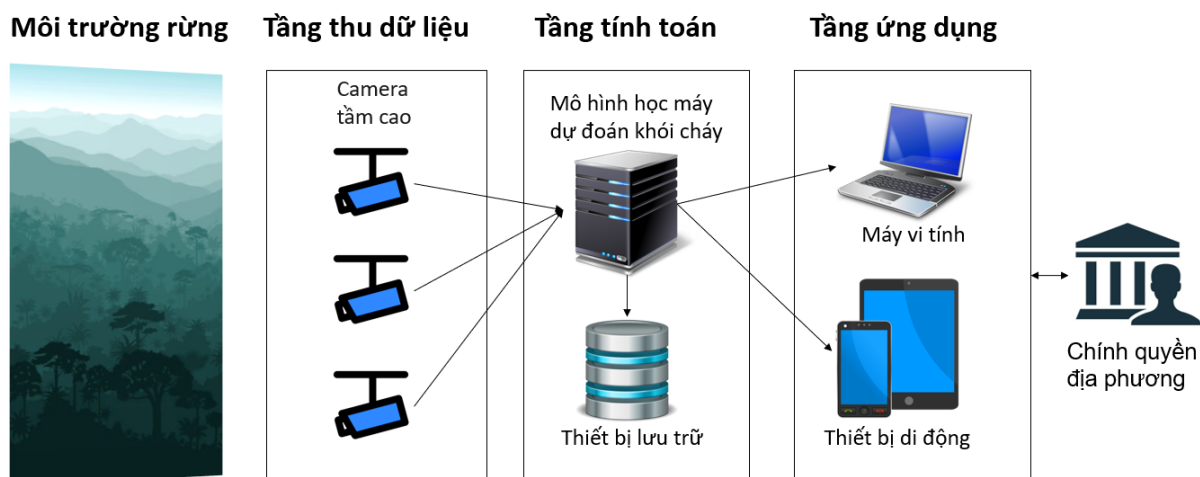
Trong khu rừng chúng tôi thực nghiệm được quản lý bởi đơn vị bảo vệ rừng địa phương, hơn 30 camera tầm cao được lắp đặt có độ phân giải cao, có thể xoay 360 độ. Mỗi camera bao quát được khu vực xung quanh với bán kính khoảng 10km được chia thành 12 presets, mỗi preset tương ứng với khoảng 30 độ và được lên lịch thay đổi preset sau mỗi 10 phút. Các camera này được kết nối với máy chủ đặt tại cơ quan để người giám sát có thể theo dõi trực tuyến 24/7. Đơn vị mong muốn xây dựng một hệ thống có khả năng giám sát tự động để khi xảy ra thảm họa cháy rừng, hệ thống sẽ kịp thời gửi hình ảnh phát hiện khói cháy này

Tác giả liên hệ: Trần Tiến Công,

Email: congtt@ptit.edu.vn

Đến tòa soạn: 10/2023, chỉnh sửa: 11/2023,

chấp nhận đăng: 12/2023.



Hình 1: Hệ thống phát hiện và cảnh báo cháy rừng.

đến màn hình TV, điện thoại thông minh, máy tính của từng cá nhân, tổ chức chuyên trách. Sau đó, lực lượng chức năng xác nhận cháy rừng và lên phương án xử lý. Hệ thống phát hiện cháy rừng này cần phát hiện nhanh chóng, chính xác trong trường hợp tài nguyên tính toán của đơn vị bị hạn chế mà vẫn đảm bảo hoạt động tốt với hàng chục camera thời gian thực.

Hình 1 minh họa sơ đồ hệ thống bao gồm ba lớp, lần lượt là Tầng thu dữ liệu, Tầng tính toán và Tầng ứng dụng. Cụ thể, hệ thống triển khai các camera tầm cao rải rác trong rừng để thu thập các luồng dữ liệu hình ảnh bao quát mọi hành vi của các chủ thể. Dữ liệu hình ảnh sẽ được truyền đến tầng tính toán ở gồm máy chủ đặt tại cơ quan. Tại tầng này, chúng tôi đề xuất mô hình phát hiện khói cháy dựa trên học sâu xử lý các đặc trưng không gian-thời gian của các luồng dữ liệu ảnh thời gian thực. Vị trí xuất hiện khói cháy sau đó được định vị và chuyển ngay đến lực lượng thường trực thuộc chính quyền địa phương thông qua tầng ứng dụng. Vì hệ thống cần xử lý nhiều video phát trực tuyến thời gian thực, các mô hình, kỹ thuật học máy sử dụng cần đảm bảo các tiêu chí nhanh, gọn nhẹ, tốn ít tài nguyên tính toán mà vẫn đảm bảo độ chính xác cao.

Phần còn lại của bài báo được tổ chức như sau. Trong phần II, chúng tôi trình bày những nghiên cứu liên quan tới nội dung của bài báo. Trong phần III, chúng tôi trình bày phương pháp đề xuất. Phần II đưa ra những kết quả thực nghiệm. Cuối cùng, chúng tôi kết luận bài báo trong phần V.

## II. CÁC NGHIÊN CỨU LIÊN QUAN

Để phát hiện sớm cháy rừng, có thể sử dụng nhiều loại tín hiệu khác nhau như nhiệt độ/độ ẩm bằng cảm biến [3], [4], [5], khói cháy nhận dạng qua thiết bị thị giác máy [6], [7], [8] hay thậm chí bằng tiếng chim hót trong trạng thái hoảng loạn khi phát hiện đám cháy tại môi trường sống của chúng [1]. Nhờ sự phát triển của máy ảnh kỹ thuật số có độ phân giải cao, khói cháy dựa trên quang học được xem là tín hiệu cảnh báo cháy rừng nghiên cứu trọng tâm trong suốt quá trình phát triển thị giác máy tính. Phát hiện khói bằng thị giác máy tính thực sự khó đối với cả kỹ thuật học máy cơ bản cũng như kỹ thuật học sâu sử dụng mạng nơ-ron. Nhiều nghiên cứu chỉ tập trung vào việc trích xuất các đặc trưng hình thái của khói cháy bằng mô hình màu, thiết kế thủ công [9], dựa trên CNN [10],... với ưu điểm độ trễ thấp, tuy nhiên, kết quả thực tế cho độ chính xác không cao vì những đặc trưng này rất giống với các vật thể tương tự khói (mây, sương mù, tường nhà,... Hình 2). Một số nghiên cứu khác trích xuất thêm các đặc trưng thời gian bằng LSTM [11], 3D-PFCN [12] cho kết quả thực tế tốt hơn nhưng đòi hỏi một lượng lớn tài nguyên tính toán dẫn đến độ trễ xử lý cao.

Trong nghiên cứu [13], nhóm tác giả áp dụng multitask learning để huấn luyện mô hình học song song đặc trưng chuyển động và đặc trưng hình thái của khói cháy. Tuy nhiên, việc kết hợp này khiến cho không gian dữ liệu đầu vào tăng lên đáng kể, độ chính xác thực nghiệm phụ thuộc



Hình 2: Minh họa mô hình phát hiện sai các đối tượng tương tự khói.

nhiều vào mức độ tổng quát của dữ liệu thu được so với phân phối dữ liệu thực tế.

Hầu hết các nghiên cứu trước đây không có tính thực tiễn cao do chỉ tập trung xử lý một hoặc một vài luồng dữ liệu ảnh, hoặc thậm chí chỉ trên từng hình ảnh. Chưa có hệ thống nào được thiết kế để có thể xử lý đồng thời nhiều luồng dữ liệu hình ảnh thời gian thực với lượng tài nguyên tính toán hạn chế. Đây cũng là mục tiêu nghiên cứu chính của bài báo này.

### III. PHƯƠNG PHÁP ĐỀ XUẤT

#### A. Định nghĩa bài toán

Trong bài báo này, nhóm nghiên cứu đề xuất một hệ thống có khả năng cảnh báo sớm thảm họa cháy rừng bằng việc ứng dụng trí tuệ nhân tạo trong công tác phát hiện khói cháy qua các camera tầm cao. Cụ thể, yêu cầu của hệ thống cảnh báo cháy rừng bao gồm:

- Phát hiện khói cháy xuất hiện tại các khu vực rừng qua camera tầm cao;
- Định vị vị trí, khu vực xuất hiện khói cháy;
- Cảnh báo phát hiện khói cháy đến lực lượng thường trực qua các thiết bị thông minh;

Hệ thống thiết kế đảm bảo các tiêu chí:

- Phát hiện nhanh chóng khu vực xảy ra khói cháy theo thời gian thực, đảm bảo cảnh báo kịp thời đến chính quyền địa phương.
- Độ phức tạp tính toán thấp, đảm bảo hoạt động ổn định với lượng tài nguyên tính toán hạn chế.

Để thực hiện được những mục tiêu này, nhóm nghiên cứu áp dụng hai công nghệ chính, bao gồm

#### Thông báo từ Camera Lâm Đồng Inbox x



LAM DONG NOTIFY <mobifone.deeplearning@gmail.com>  
to me, liethoa018, viettien992

Phát hiện có smoke từ Camera mã số: 9\_lamdong1



Hình 3: Minh họa cảnh báo khói cháy qua email.

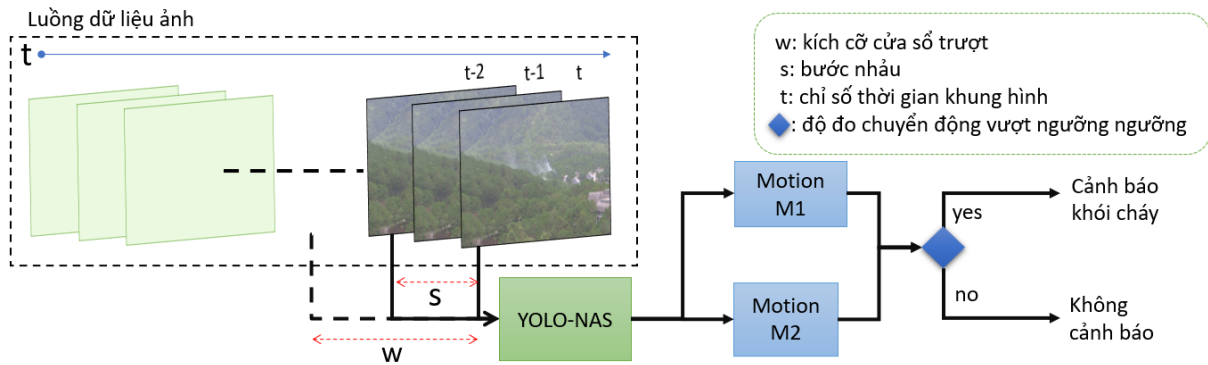
mô hình phát hiện đối tượng thời gian thực để phát hiện khói cháy (khối YOLO-NAS trong Hình 4) và mô hình phân vùng ảnh, các kỹ thuật xử lý ảnh giúp giảm cảnh báo sai False Positive (khối Motion M1, M2 trong Hình 4). Đầu ra của mô hình tổng là hình ảnh khói cháy đã khoanh vùng được cảnh báo qua email cùng với các thông tin về camera, độ tự tin phát hiện và diện tích vùng khói cháy (Hình 3).

#### B. Phát hiện đối tượng khói cháy

Với yêu cầu từ định nghĩa bài toán, nhóm nghiên cứu xác định yêu cầu chức năng của mô hình học máy phát hiện khói cháy bao gồm phát hiện khói cháy và khu vực xảy ra khói cháy. Để thực hiện mục tiêu này, nhóm nghiên cứu áp dụng mô hình học sâu phát hiện đối tượng.

Phát hiện đối tượng là một tác vụ thị giác máy tính nhằm mục đích xác định vị trí và phân loại các đối tượng trong hình ảnh hoặc video. Mục tiêu chính là xây dựng một mô hình có khả năng xác định sự hiện diện của các đối tượng cụ thể và vẽ các hộp giới hạn (bbox - bounding box) xung quanh chúng để phân định chính xác vị trí của chúng. Bbox là một vùng hình chữ nhật bao bọc vị trí của một đối tượng, được biểu thị bằng bốn tọa độ (x, y) của góc trên bên trái và (chiều rộng, chiều cao) của hình hộp.

Cụ thể, nhóm nghiên cứu sử dụng mô hình YOLO-NAS là một mô hình phát hiện đối tượng theo thời gian thực cải tiến, sản phẩm trí tuệ của Deci.ai tận dụng những tiến bộ gần đây nhất trong



Hình 4: Mô hình học máy phát hiện khói cháy

công nghệ học sâu. YOLO-NAS đánh dấu một cột mốc quan trọng trong họ YOLO, giải quyết các hạn chế chính của các phiên bản YOLO trước đây và đẩy khả năng phát hiện đối tượng theo thời gian thực lên một tầm cao mới. Ký hiệu NAS liên quan đến thuật toán tìm kiếm kiến trúc mạng tối ưu (Neural Architecture Search) do YOLO-NAS ứng dụng thuật toán này để tìm kiếm kiến trúc mạng YOLO. Ngoài ra mô hình YOLO-NAS khai thác sức mạnh của cơ chế tập trung (attention) và tham số hóa lại thời gian suy luận để tăng cường khả năng phát hiện đối tượng. Nhìn chung, YOLO-NAS cho kết quả tốt với độ đo mAP cao và thời gian suy diễn thấp.

Các thuật toán phát hiện đối tượng phổ biến chỉ sử dụng đặc trưng hình thái, phát hiện hiệu quả với các đối tượng tĩnh nhưng tỏ ra kém hiệu quả với các đối tượng cần thêm đặc trưng chuyển động theo thời gian để phân biệt. điển hình như khói cháy rất khó phát hiện nếu không có các đặc trưng chuyển động theo thời gian do về hình thái, khói tương đối giống với mây, sương mù hoặc khói bụi,... Xử lý thêm đặc trưng chuyển động giúp tăng khả năng phân biệt đối tượng khói với những đối tượng tĩnh nhưng cần thêm chi phí tính toán. Nhìn chung, cân bằng mức tiêu thụ tài nguyên máy tính và phát hiện khói cháy chính xác là một thách thức. Do đó, nhóm nghiên cứu nỗ lực cải thiện độ chính xác phát hiện khói cháy với ràng buộc hạn chế tăng độ phức tạp tính toán bằng việc sử dụng phương pháp trích chọn đặc trưng hiệu quả. Cụ thể, đặc trưng phát hiện khói cháy được chia thành đặc trưng không gian (hình thái) được trích chọn bằng mạng backbone của YOLO-NAS và đặc trưng thời gian (chuyển động) trích chọn bằng các kỹ thuật học máy và xử lý ảnh sẽ được

trình bày trong phần sau.

### C. Giảm tỉ lệ phát hiện nhầm khói cháy

Xem xét đặc trưng hình thái, một số đối tượng tương tự khói cháy có thể bị phát hiện nhầm. Tuy nhiên, mạng học sâu vẫn có thể phân biệt được nhóm đối tượng này nếu được học tăng cường mẫu đối nghịch. Nhóm nghiên cứu sử dụng một kỹ thuật học tương phản (contrastive learning) được sử dụng phổ biến trong lĩnh vực học máy là hard negative samples. Cụ thể, ngoài huấn luyện mô hình với dữ liệu khói cháy, chúng tôi xây dựng thêm dữ liệu đối nghịch là các đối tượng giống khói cháy mà mô hình học sâu phát hiện sai khi thực nghiệm hệ thống tại khu vực rừng.

Dựa trên thực tế quan sát thấy đối tượng khói thường chuyển động mạnh về hình thái, nhóm nghiên cứu thực hiện trích chọn đặc trưng chuyển động như sau: giả sử khung hình hiện tại ở thời gian  $t$ , chúng tôi lựa chọn các khung hình trước đó bởi một cửa sổ trượt kích cỡ  $w$  và cơ chế bước nhảy  $s$  để tính độ đo chuyển động. Cụ thể, đầu tiên  $w$  khung hình liên tiếp  $[t, t-1, \dots, t-(w-1)]$  được chọn, sau đó bỏ qua lần lượt  $s-1$  khung hình liên tục sẽ còn lại  $[t, t-s, \dots, t-(w-1)]$ . Tại khung hình  $t$  với mỗi đối tượng khói cháy  $O$ , chúng tôi tính hai độ đo chuyển động:

- Motion M1 tính độ đo chuyển động tích lũy theo cơ chế bbox-to-bbox như (1) bên dưới:

$$M_1 = \sum_{n=0}^l |O_{t-n*s} - O_{t-(n+1)*s}| \quad (1)$$

$$|O_i - O_j| = \sum_{x,y,c}^{H_o, W_o, C_o} |O_{i,x,y,c} - O_{j,x,y,c}| \quad (2)$$

Với  $l = \frac{w-1}{s} - 1$ ,  $O_i$  là ảnh của đối tượng  $\mathbb{R}^{h \times w \times c}$  từ khung hình  $i$ ;  $H_o, W_o, C_o$  là kích cỡ ảnh của đối tượng  $O$ ;  $O_{n,x,y,c}$  là giá trị điểm ảnh của  $O_n$ . Công thức (1) chỉ tính được khi mọi  $O$  từ cùng đối tượng và kích cỡ. Vì mô hình phát hiện nhiều đối tượng tại mỗi khung hình nên cần phân biệt các phát hiện trước đó là thuộc về đối tượng nào trong khung hình hiện tại dựa vào độ đo  $IoU = \frac{Area-of-Intersection}{Area-of-Union}$  như sau: với mỗi đối tượng phát hiện ở khung hình  $t$ , chúng tôi lặp tìm kiếm đối tượng gần nhất trước đó cho đến khi không có đối tượng trùng khớp hoặc gặp khung hình cuối trong các khung hình được lựa chọn. Sau đó đưa các ảnh đối tượng được cắt ra về cùng một kích cỡ và áp dụng công thức (1).

- Motion M2 tính độ đo chuyển động tính lũy dựa trên cơ chế Lastest-Bbox. khác với M1, M2 chỉ dùng Bbox tại thời điểm  $t$ , tất cả  $O$  được cắt theo Bbox này (khối motion measure bên dưới Hình 4). Phần còn lại và công thức tính tương tự M1.

Việc kiểm tra đối tượng chuyển động như trên có thể loại bỏ các đối tượng bất động về hình thái và vị trí tương tự như khói (bức tường trong Hình 2) nhưng không thể loại bỏ các đối tượng chuyển động (các đám mây trong Hình 2). May mắn thay, vì những đối tượng này nằm trên bầu trời nên có thể dễ dàng loại bỏ. Chúng tôi sử dụng một mô hình nhỏ gọn để phân vùng ảnh thành hai phần bầu trời và mặt đất, sau đó loại bỏ đối tượng nếu tỉ lệ đối tượng nằm trên phần bầu trời vượt quá ngưỡng xác định.

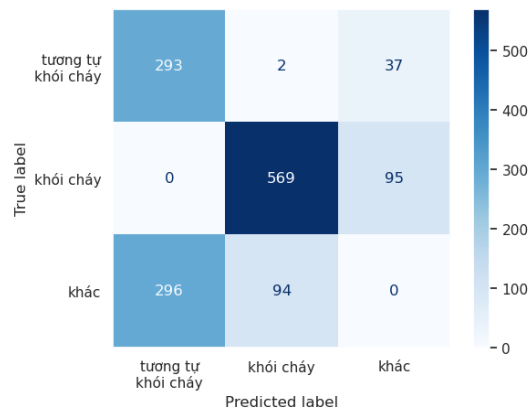
#### IV. KẾT QUẢ THỰC NGHIỆM

Để đánh giá mô hình học sâu YOLO-NAS được ứng dụng trong bài toán này, chúng tôi thử nghiệm hiệu năng của mô hình trên tập dữ liệu đầu vào là bộ dữ liệu được thu thập trực tiếp tại khu vực rừng thực nghiệm. Tập dữ liệu 9397 ảnh bao gồm 6709 ảnh chứa đối tượng khói cháy và 2688 ảnh chứa đối tượng tương tự khói cháy. Hình 5 thể hiện một số mẫu hình ảnh và nhãn của tập dữ liệu phát hiện khói cháy.

Các thông số thực nghiệm được liệt kê như sau. Tham số inputSize được đặt là 640, nghĩa là ảnh



Hình 5: Minh họa dữ liệu huấn luyện.



Hình 6: Độ chính xác mô hình phát hiện khói thể hiện qua confusion matrix với confidence score 0.5 và iou threshold 0.5.

đầu vào được resize thành kích thước 640x640 pixels. Tập dữ liệu phát hiện khói được chia thành ba tập bao gồm train, validation, test set với tỉ lệ tương ứng 0.7:0.2:0.1. Đối với tập huấn luyện mô hình, chúng tôi sử dụng các kỹ thuật tăng cường ảnh phổ biến trong thị giác máy tính giúp tăng số lượng ảnh lên gấp 3 lần tương đương với gần 20k ảnh. Tập dữ liệu phân vùng ảnh trời gồm tương tự được chia thành ba tập với tỉ lệ 0.8:0.1:0.1. Đối với tập dữ liệu này, chúng tôi chỉ áp dụng các kỹ thuật tăng cường ảnh phù hợp do ràng buộc đặc thù bài toán phân vùng ảnh trời. Mô hình được huấn luyện từ đầu và không sử dụng pretrain

Hình 6 mô tả độ chính xác đạt được của mô hình đề xuất bằng confusion matrix. Chỉ những tiên đoán với confidence score và iou lớn hơn ngưỡng 0.5 được thống kê. Mẫu tương tự khói cháy là những đối tượng không phải khói cháy bị tiên đoán nhầm khi mô hình được huấn luyện mà không sử dụng kỹ thuật hard negative samples. Kết quả cho thấy rằng, việc ứng dụng hard negative samples gần như loại bỏ được các đối tượng này với chỉ 2 lần sai sót. Tuy nhiên kỹ thuật này cũng gây ảnh hưởng tiêu cực làm cho mô hình khó phát hiện các đối tượng thực sự là khói nhưng đặc trưng hình thái gần với đối tượng khói cháy. Nhìn chung, chỉ việc ứng dụng hard

negative samples có thể giúp mô hình phân biệt được những đối tượng tương tự khói. Cụ thể, mô hình phát hiện khói cháy đạt mAP@0.50 lên đến 0.89. Đối với mô hình phân vùng ảnh trời cũng cho kết quả cao với độ chính xác lên đến 99%.

Chúng tôi thử nghiệm khả năng xử lý của mô hình học máy tổng quát với thông tin tài nguyên tính toán sử dụng gồm: Vi xử lý (CPU) 11th Gen Intel® Core™ i9-11900 @ 2.50GHz, Dung lượng bộ nhớ RAM: 32GB, Ổ cứng: SSD 512GB, Card đồ họa: NVIDIA Quadro RTX 4000 8GB, Hệ điều hành: Ubuntu 22.04.1 LTS. Kết quả thực nghiệm cho thấy chỉ với lượng tài nguyên tính toán hạn chế như trên, hệ thống có thể hoạt động ổn định với tối đa 34 luồng camera 4 fps. Với một hiệu năng cao như vậy, chúng tôi tin rằng việc triển khai ứng dụng thực tế là rất khả thi.

## V. KẾT LUẬN

Trong bài báo này chúng tôi đã trình bày một hệ thống cảnh báo cháy rừng dựa trên phát hiện khói cháy qua hình ảnh thu được từ camera tầm cao đặt tại khu vực rừng. Hệ thống dựa chủ yếu vào mô hình phát hiện đối tượng YOLO-NAS giúp phát hiện khói cháy và các kỹ thuật xử lý ảnh cơ bản giúp giảm cảnh báo sai. Qua thực nghiệm với các độ đo là độ chính xác, mAP@.50 và độ trễ đã được chứng minh hệ thống có khả năng phát hiện cháy rừng nhanh và chính xác với mức tiêu thụ tài nguyên thấp. Cụ thể, hệ thống đã đạt được mAP@0.50 0.89 với YOLO-NAS trên bộ dữ liệu phát hiện khói cháy thử, độ chính xác 99% với mô hình phân vùng ảnh trên bộ dữ liệu phân vùng ảnh trời. Ngoài ra mô hình học sâu có độ trễ thấp xấp xỉ 10ms đảm bảo phát hiện khói cháy trong thời gian thực.

Chúng tôi có kế hoạch tích hợp thêm những mô hình học sâu để phát hiện các bất thường khác như công trình xây dựng trái phép, khu vực đất trống đồi trọc, và các hành vi phá hoại tài sản rừng giúp giám sát và bảo vệ rừng triệt để hơn.

## LỜI CẢM ƠN

Nghiên cứu này được hỗ trợ bởi đề tài cấp Bộ Thông tin và Truyền thông, mã số ĐT.28/23.

## TÀI LIỆU THAM KHẢO

- [1] S. D. H. Permana, G. Saputra, B. Arifitama, W. Caesarendra, R. Rahim *et al.*, "Classification of bird sounds as an early warning method of forest fires using convolutional neural network (cnn) algorithm," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 7, pp. 4345–4357, 2022.
- [2] F. Abid, "A survey of machine learning algorithms based forest fires prediction and detection systems," *Fire technology*, vol. 57, no. 2, pp. 559–590, 2021.
- [3] V. Dubey, P. Kumar, and N. Chauhan, "Forest fire detection system using iot and artificial neural network," in *International Conference on Innovative Computing and Communications: Proceedings of ICICC 2018, Volume 1*. Springer, 2019, pp. 323–337.
- [4] A. Sharma, P. K. Singh, and Y. Kumar, "An integrated fire detection system using iot and image processing technique for smart cities," *Sustainable Cities and Society*, vol. 61, p. 102332, 2020.
- [5] W. Liu, X. Wang, Y. Song, R. Cao, L. Wang, Z. Yan, and G. Shan, "Self-powered forest fire alarm system based on impedance matching effect between triboelectric nanogenerator and thermosensitive sensor," *Nano Energy*, vol. 73, p. 104843, 2020.
- [6] Z. Jiao, Y. Zhang, J. Xin, L. Mu, Y. Yi, H. Liu, and D. Liu, "A deep learning based forest fire detection approach using uav and yolov3," in *2019 1st International conference on industrial artificial intelligence (IAI)*. IEEE, 2019, pp. 1–5.
- [7] F. Bu and M. S. Gharajeh, "Intelligent and vision-based fire detection systems: A survey," *Image and vision computing*, vol. 91, p. 103803, 2019.
- [8] S. Sudhakar, V. Vijayakumar, C. S. Kumar, V. Priya, L. Ravi, and V. Subramaniaswamy, "Unmanned aerial vehicle (uav) based forest fire detection and monitoring for reducing false alarms in forest-fires," *Computer Communications*, vol. 149, pp. 1–16, 2020.
- [9] Y. Peng and Y. Wang, "Real-time forest smoke detection using hand-designed features and deep learning," *Computers and Electronics in Agriculture*, vol. 167, p. 105029, 2019.
- [10] S. Saponara, A. Elhanashi, and A. Gagliardi, "Implementing a real-time, ai-based, people detection and social distancing measuring system for covid-19," *Journal of Real-Time Image Processing*, pp. 1–11, 2021.
- [11] Y. Cao, F. Yang, Q. Tang, and X. Lu, "An attention enhanced bidirectional lstm for early forest fire smoke recognition," *IEEE Access*, vol. 7, pp. 154 732–154 742, 2019.
- [12] X. Li, Z. Chen, Q. J. Wu, and C. Liu, "3d parallel fully convolutional networks for real-time video wildfire smoke detection," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 1, pp. 89–103, 2018.
- [13] Y. Hu and X. Lu, "Real-time video fire smoke detection by utilizing spatial-temporal convnet features," *Multimedia Tools and Applications*, vol. 77, pp. 29 283–29 301, 2018.
- [1] S. D. H. Permana, G. Saputra, B. Arifitama, W. Caesarendra, R. Rahim *et al.*, "Classification of bird sounds as an early warning method of forest fires using con-

## ANOMALY DETECTION IN FOREST MONITORING USING ARTIFICIAL INTELLIGENCE

**Abstract:** Forest monitoring is the monitoring of subjects and institutions, making operational decisions, and responding in a timely and appropriate manner to the roles and responsibilities of relevant parties in forest protection. Due to the fact that Vietnam has a forest area ratio of more than 40%, the task of monitoring and protecting forests is important. Although there are supporting applications applying modern technologies, the monitoring task still depends heavily on the professional capacity of the national forest ranger force. In this article, we study an application to help detect provide early notifications of forest fires, a type of anomaly in forest monitoring, by applying the object detection approach to videos recorded from high-altitude cameras located within the forest area. Through intensive experiments, our forest fire detection system achieves outstanding results with the measuring metric mAP@0.50 of up to 0.89 on the test set with low latency.

**Keywords:** Forest monitoring, forest protection, fire smoke detection, object detection



**Vũ Quang Vinh** sinh ra ở Nam Định, Việt Nam, vào năm 2000, là sinh viên của Học viện Công nghệ Bưu chính Viễn thông Hà Nội. Tại Học viện, Vũ Quang Vinh học kỹ sư chuyên ngành Hệ thống thông tin, lĩnh vực nghiên cứu của Vũ Quang Vinh bao gồm xử lý ảnh, thị giác máy tính, xử lý ngôn ngữ tự nhiên và học máy/học sâu.



**Trần Anh Đạt** nhận bằng kỹ sư An toàn thông tin tại Học viện Công nghệ Bưu chính Viễn thông năm 2020 và bằng thạc sỹ Khoa học Máy tính tại Học viện Công nghệ Bưu chính Viễn thông, Việt Nam năm 2022. Hiện tại, Thạc sỹ Đạt đang là nghiên cứu sinh ngành Khoa học Máy tính tại Viện Hàn lâm Khoa học và Công nghệ Việt Nam. Từ năm 2023, thạc sỹ Đạt là giảng viên bộ môn Trí tuệ nhân tạo, Đại học Thủy lợi. Hướng nghiên cứu của thạc sỹ Đạt bao gồm nhận dạng ảnh và học sâu.



**Trần Tiến Công** nhận bằng tiến sĩ khoa học máy tính tại Đại học Dankook, Yongin, Hàn Quốc, vào năm 2021. Trước đó TS. Trần Tiến Công đã nhận bằng Thạc sĩ khoa học máy tính vào năm 2014 và Cử nhân về mạng và truyền thông vào năm 2009 từ Đại học Quốc gia Hà Nội, Việt Nam. Từ tháng 9 năm 2021, Trần Tiến Công làm việc tại Khoa Công nghệ Thông tin, Học viện Công nghệ Bưu chính Viễn thông, Hà Nội, Việt Nam với vai trò là giảng viên. Lĩnh vực nghiên cứu của TS. Trần Tiến Công bao gồm phân tích mạng xã hội, khai thác dữ liệu và học máy.