

THIẾT KẾ, CHẾ TẠO HỆ THỐNG CẢNH BÁO SỚM ĐÁM CHÁY TRONG TÒA NHÀ CAO TẦNG SỬ DỤNG CÔNG NGHỆ HỌC MÁY

DESIGN AND IMPLEMENTATION OF FIRE EARLY WARNING SYSTEM IN IN-BUILDING INFRASTRUCTURE BASED ON DEEP LEARNING

Phạm Ngọc Pha¹, Nguyễn Trọng Hiếu¹, Nguyễn Việt Thắng², Nguyễn Trường Sơn², Quách Công Hoàng², Phạm Minh Triển^{2,*}

TÓM TẮT

Phát hiện và cảnh báo cháy kịp thời góp phần quan trọng trong việc đảm bảo an toàn, giảm thiểu thiệt hại cho con người. Hiện nay, cách tiếp cận chủ yếu được sử dụng nhằm giải quyết vấn đề trên là sử dụng các cảm biến nhiệt độ, cảm biến khói. Phương pháp này có hạn chế là khi hệ thống phát hiện ra đám cháy thì tình trạng cháy đã lớn và lan rộng. Trong bài báo này chúng tôi đề xuất một hệ thống phát hiện đám khói nhằm cảnh báo sớm đám cháy được xây dựng dựa trên nền tảng công nghệ học máy. Quá trình xây dựng, thử nghiệm cho thấy tính khả thi của hệ thống trong việc giải quyết vấn đề phát hiện sớm và cảnh báo đám cháy. Kết quả thử nghiệm đã chứng tỏ rằng hệ thống đề xuất có thể đáp ứng tốt mục tiêu cảnh báo sớm các đám cháy và phù hợp để triển khai với các hệ thống máy tính hiện tại.

Từ khóa: Học máy, xử lý ảnh, mạng cảm biến, cảnh báo cháy.

ABSTRACT

A fire detection system is essential for people's safety. During the past few years, many approaches based on smoke sensors, humidity sensors were proposed. However, because of the limitation of the capability of sensors, those proposed methods are infeasible for early-warning systems. The purpose of this paper is to present a new fire early-warning system based on deep learning. The proposed system is designed by using sensors and surveillance cameras to detect smoke. The system is tested on our dataset and reality indoor environment. Experiments show that the new approach is successfully applied to various scenarios and significant for improving the accuracy of fire smoke detection.

Keywords: Deep learning, smoke detection, sensors, image processing.

¹Viện Nghiên cứu sáng chế và Khai thác công nghệ, Bộ Khoa học và Công nghệ

²Trường Đại học Công nghệ, Đại học Quốc gia Hà Nội

*Email: trienpm@vnu.edu.vn

Ngày nhận bài: 10/3/2020

Ngày nhận bài sửa sau phản biện: 15/4/2020

Ngày chấp nhận đăng: 24/4/2020

1. GIỚI THIỆU

Hiện nay, trên thế giới cũng như ở Việt Nam, mỗi năm có hàng nghìn vụ cháy, gây thiệt hại lớn về người và kinh tế. Thống kê thiệt hại trong 9 tháng đầu năm 2017 cả nước xảy ra hơn 3.000 vụ cháy, làm chết 75 người, bị thương 143

người, thiệt hại về tài sản là hơn 1.500 tỷ đồng [1]. Vì vậy, việc tìm kiếm và phát triển những phương pháp phát hiện sớm các khu vực sắp cháy, cháy nhỏ một cách chính xác, kịp thời thực sự cấp thiết. Hiện tại, đã có nhiều nghiên cứu đề xuất những biện pháp phát hiện và cảnh báo cháy như dùng các đầu báo cháy nhiệt, đầu báo cháy khói và đầu báo cháy lửa. Hệ thống phát hiện và cảnh báo cháy dựa trên đầu báo cháy truyền thống đã phát huy hiệu quả trong nhiều tình huống, cảnh báo cháy chính xác khi nhiệt độ, khói lan tỏa tới đầu báo cháy và đạt ngưỡng hoạt động của thiết bị. Tuy vậy, phương pháp này vẫn có hạn chế là các đầu báo cháy chỉ làm việc khi nhiệt độ, khói đã lan truyền tới đầu cảm biến và đạt ngưỡng hoạt động, khi đó thường đám cháy đã phát triển lớn. Vì vậy, hệ thống chỉ hiệu quả trong không gian nhỏ và kín (như trong tòa nhà), phát hiện cháy khi đám cháy bùng phát không nhanh, còn với vùng giám sát có không gian mở như các hành lang, phòng không kín, ảnh hưởng gió thì hệ thống hoạt động kém hiệu quả. Việc phát hiện và cảnh báo cháy sử dụng đầu dò cũng còn nhiều hạn chế do phụ thuộc hoàn toàn vào thời gian lan truyền nhiệt và khói kể từ khi bắt đầu xảy ra hiện tượng cháy cho đến khi khói hoặc nhiệt độ lan tỏa tới đầu dò.

Những năm gần đây, một hướng mở ra trong nghiên cứu cảnh báo cháy là sử dụng kỹ thuật xử lý ảnh nhận dạng thông qua hệ thống quan sát camera. Những nghiên cứu ứng dụng kỹ thuật xử lý ảnh, video vào các hệ thống camera cảnh giới với mục đích phát hiện và cảnh báo cháy phát triển mạnh và đã có những kết quả nhất định. Phần lớn các giải pháp được đề xuất cho bài toán phát hiện ngọn lửa sử dụng kỹ thuật xử lý ảnh, video số hiện đều dựa trên những tính chất có thể quan sát được của ngọn lửa như màu sắc, sự thay đổi về vị trí các điểm ảnh của ngọn lửa theo thời gian. Tuy nhiên, hầu hết các nghiên cứu mới chỉ dừng lại ở việc phân tích đặc trưng của ngọn lửa, kết quả còn ở mức thử nghiệm, độ chính xác chưa cao; nghiên cứu mới dừng lại ở vật liệu chống cháy và phương pháp chữa cháy.

Trên thế giới việc áp dụng công nghệ xử lý nhận dạng ảnh tiến triển từ giai đoạn đầu dùng các phương pháp nhận diện sử dụng màu sắc và trích xuất đặc trưng [2 - 10], cách tiếp cận mới hơn là optical flow sử dụng dữ liệu chuỗi

ảnh (Red Green Blue) để xác định vùng chuyển động [11, 12, 13]. Gần đây, cách tiếp cận ứng dụng học máy ngày càng phổ biến hơn. Trong [14], nhóm tác giả sử dụng mạng CNN, hệ thống có thể phát hiện cháy cả trong nhà và ngoài trời. Dựa trên giải thuật gốc, nhóm đề xuất cơ chế thích nghi ưu tiên đối với camera và giải thuật lựa chọn kênh động cho các camera. Với nhóm sử dụng học máy [15], dữ liệu từ các bộ thu thập dữ liệu từ cảm biến khói, nhiệt độ, ngọn lửa được đưa qua bộ học máy để phát hiện ra vùng nguy hiểm. Với việc ứng dụng mạng nơ-ron các đặc trưng của ngọn lửa mô phỏng trong phòng thí nghiệm được sử dụng làm lối vào [16]; hình ảnh chứa ngọn lửa từ ảnh động và các đặc trưng của ảnh đưa vào ANN thu được từ ảnh từ camera CCD. Có thể thấy, các hướng nghiên cứu trên quốc tế đều hướng đến ứng dụng học máy, các doanh nghiệp cũng đã bước đầu ứng dụng công nghệ này vào sản phẩm của mình. Tuy nhiên, hầu hết các giải pháp này có giá thành khá cao.

Trong bài báo này, chúng tôi trình bày mô hình của hệ thống phát hiện và cảnh báo cháy sớm và cách thức triển khai, hoạt động của hệ thống. Hệ thống được nghiên cứu với mong muốn khắc phục được những nhược điểm đang tồn tại trong bài toán phát hiện, cảnh báo cháy. Các phần còn lại của bài báo được tổ chức theo ba mục chính, phần hai trình bày các nghiên cứu liên quan đến hệ thống, phần ba và bốn mô tả thiết kế các phương thức triển khai, cài đặt hệ thống. Phần năm và sáu đưa ra các thông số đánh giá và kết luận của bài báo.

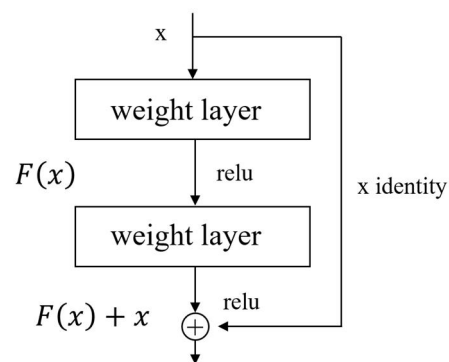
2. CÁC NGHIÊN CỨU LIÊN QUAN

Trên thế giới việc áp dụng công nghệ xử lý nhận dạng ảnh tiến triển từ giai đoạn đầu dùng các phương pháp nhận diện sử dụng màu sắc và trích xuất đặc trưng [2 - 10], cách tiếp cận mới hơn là optical flow sử dụng dữ liệu chuỗi ảnh (Red Green Blue) để xác định vùng chuyển động [11, 12, 13]; tiếp theo là việc ứng dụng các công cụ học máy để nâng cao khả năng phán đoán [14, 15, 16]. Gần đây, cách tiếp cận ứng dụng học máy ngày càng phổ biến hơn. Trong [14], nhóm tác giả sử dụng mạng CNN, hệ thống có thể phát hiện cháy cả trong nhà và ngoài trời. Dựa trên giải thuật gốc, nhóm đề xuất cơ chế thích nghi ưu tiên đối với camera và giải thuật lựa chọn kênh động cho các camera. Với nhóm sử dụng học máy [15], dữ liệu từ các bộ thu thập dữ liệu từ cảm biến khói, nhiệt độ, ngọn lửa được đưa qua bộ học máy để phát hiện ra vùng nguy hiểm. Với việc ứng dụng mạng nơ-ron các đặc trưng của ngọn lửa mô phỏng trong phòng thí nghiệm được sử dụng làm lối vào [16]; hình ảnh chứa ngọn lửa từ ảnh động và các đặc trưng của ảnh đưa vào ANN thu được từ ảnh từ camera CCD.

Phân lớp ảnh (Image Classification) là nhiệm vụ gán nhận một ảnh đầu vào với một tập nhãn trong một danh sách cố định [17]. Đây là nhiệm vụ cơ bản và đơn giản nhất trong bài toán xử lý ảnh và nó cũng có tính ứng dụng rất cao. Rất nhiều nhiệm vụ trong xử lý ảnh cũng có thể quy về bài toán phân lớp ảnh. Nhiệm vụ phân lớp ảnh được thực hiện rất thành công bởi mạng nơ-ron xoắn hay CNNs.

Những mạng phổ biến nhất hiện nay có thể kể đến như: ResNet [18], InceptionNet [19], MobileNet [20]... Chúng tôi đã thử nghiệm các mô hình kể trên trong bài toán nhận dạng đám cháy. Dựa theo kết quả thực nghiệm, chúng tôi lựa chọn mô hình mạng ResNets-50 do mô hình mạng này dễ huấn luyện và số trọng số không quá lớn, phù hợp với độ phức tạp của bài toán nhận dạng đám cháy. Một số mạng mới nhất hiện nay như DenseNets hay NasNets, mặc dù có kết quả chính xác rất cao nhưng không được sử dụng bởi những mô hình này rất khó huấn luyện. Kể cả khi chúng ta chỉ thực hiện điều chỉnh tham số, nếu không có một hệ tính toán đủ mạnh thì khó lòng có thể đánh giá được độ tin cậy của kết quả huấn luyện.

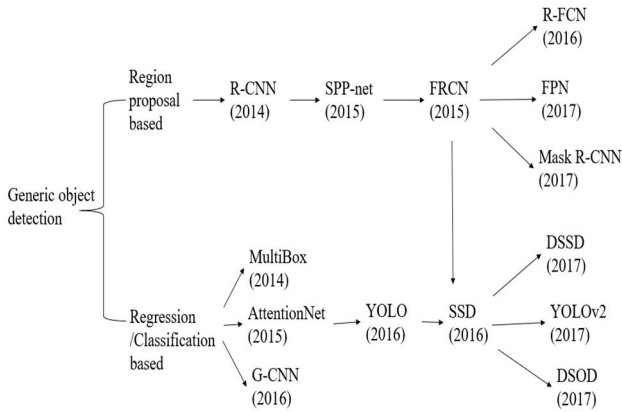
ResNet được phát triển bởi microsoft năm 2015 [18]. ResNet đã chiến thắng ImageNet ILSVRC competition 2015 với tỷ lệ lỗi là 3,57% và ResNet có cấu trúc gần giống VGG với nhiều lớp làm cho mô hình sâu hơn. Không giống VGG, ResNet có số lớp sâu hơn như 34, 55, 101 và 151. Resnet giải quyết được vấn đề của học sâu truyền thống, đó là có thể dễ dàng huấn luyện mô hình với hàng trăm lớp. Để hiểu ResNet chúng ta cần hiểu vấn đề khi xếp chồng nhiều lớp khi huấn luyện. Vấn đề đầu tiên khi tăng độ sâu của mô hình hơn thì đạo hàm sẽ bị suy biến. Vấn đề này có thể giải quyết bằng cách thêm Batch Normalization [21], nó giúp chuẩn hóa đầu ra giúp các hệ số trở nên cân bằng hơn không quá nhỏ hoặc quá lớn nên sẽ giúp mô hình dễ hội tụ hơn. Vấn đề thứ hai là mất mát thông tin khi mô hình quá sâu do khi mô hình sâu hơn độ chính xác bắt đầu bão hòa thậm chí là giảm. Resnet được ra đời để giải quyết vấn đề này. ResNet có kiến trúc gồm nhiều khối kết nối dư. Ý tưởng chính là dùng cạnh kết nối một lớp với một lớp từ trước đó nhiều lớp (bỏ qua các lớp ở giữa). Mô hình đặc trưng này được minh họa trong hình 1.



Hình 1. Mô hình kết nối các tầng của ResNet [17]

Bài toán xác định vật thể đề cập đến khả năng của hệ thống máy tính và phần mềm để định vị các đối tượng trong một hình ảnh và xác định từng đối tượng. Bài toán này đã được sử dụng rộng rãi để phát hiện khuôn mặt, phát hiện xe, đếm số người đi bộ, hệ thống bảo mật và xe không người lái. Có nhiều cách để nhận diện đối tượng có thể được sử dụng cũng như trong nhiều lĩnh vực thực hành. Nhiệm vụ bài toán xác định vật thể đã có nhiều bước phát triển trong những năm gần đây. Phương pháp đầu tiên là

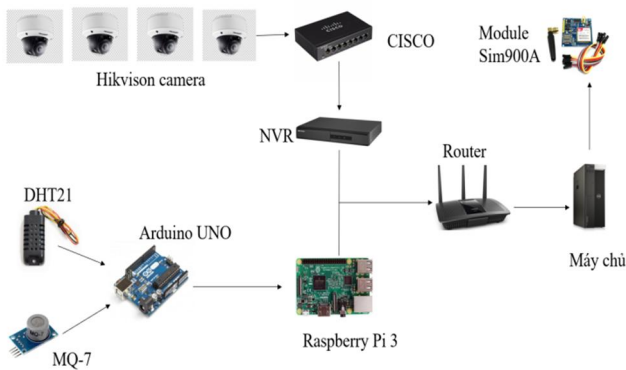
kiểu brute-force với một cửa sổ trượt quét tất cả các vùng trong ảnh để xác định vật thể. Cùng với học sâu và hàng loạt kĩ thuật mới được giới nghiên cứu đề xuất như Feature Pyramid [22], Region Proposal [23]... Bài toán này đã đạt được những tiến bộ vô cùng khả quan. Các hướng phát triển của bài toán nhận diện vật thể được mô tả trong hình 2. Hiện nay xác định vật thể được chia làm hai cách tiếp cận chính: Region proposal based và Regression Classification based; và các đại diện tiêu biểu của các phương pháp này là các mạng Faster R-CNN [23], YOLO [24] và SSD [25].



Hình 2. Quá trình phát triển của bài toán xác định đối tượng [16]

3. THIẾT KẾ VÀ TRIỂN KHAI HỆ THỐNG

3.1. Tổng quan hệ thống



Hình 3. Tổng thể kết nối của hệ thống

Hệ thống được xây dựng bao gồm hai phần chính là ghép nối thiết bị phần cứng và phần mềm triển khai hệ thống. Đầu tiên về hệ thống phần cứng, phần cứng được chia làm hai phần chính, thứ nhất là các cảm biến thu thập thông tin về môi trường và thứ hai là máy chủ xử lý các tác vụ như phát hiện cháy, đưa ra cảnh báo, thông báo. Trong hình 3 là mô hình kết nối của hệ thống, các thiết bị chính của hệ thống bao gồm:

- 01 máy chủ, cấu hình máy chủ bao gồm CPU Intel Xeon E5, GPU Nvidia GTX 2080Ti
- 04 camera an ninh, giám sát. Loại camera được sử dụng là IP Camera.
- 01 cảm biến nhiệt độ độ ẩm DHT21

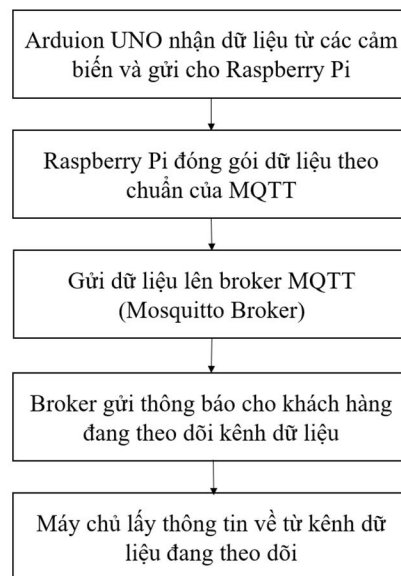
- 01 cảm biến khí CO MQ-7
- 01 mô-đun Sim900A
- 01 Arduino UNO
- 01 Raspberry Pi 3

3.2. Cảm biến nhiệt độ, độ ẩm, khí CO

Các cảm biến nhiệt độ, độ ẩm DHT21 và cảm biến khí CO MQ-7 được đặt ở những nơi có nguy cơ cao xảy ra cháy nổ như tủ điện, trạm biến áp... Các vị trí này cần sự giám sát đặc biệt nhằm phát hiện kịp thời những nguy cơ xuất hiện cháy nổ. Quá trình đọc và truyền thông tin cảm biến về máy chủ có hai bước chính:

• Arduino UNO có nhiệm vụ nhận thông tin từ DHT21 và MQ-7 sau đó gửi lên cho Raspberry Pi 3. Đây là một kit vi xử lý cơ bản và phổ biến cung cấp những giao thức cần thiết để kết nối tín hiệu tương tự và tín hiệu số đồng thời kết nối với máy tính nhưng bằng giao thức nối tiếp nhằm tạo ra một mô-đun kết nối đa cảm biến tiện dụng và nhỏ gọn.

• Raspberry Pi 3 là thành phần xử lý chính. Raspberry Pi 3 được xây dựng xoay quanh bộ xử lí SoC Broadcom BCM2835 (là chip xử lí mạnh mẽ có kích thước nhỏ hay được dùng trong điện thoại di động) bao gồm CPU, GPU, bộ xử lí âm thanh, hình ảnh và các tính năng khác... tất cả được tích hợp bên trong chip có điện năng thấp này. Mô-đun này sẽ nhận thông tin cảm biến từ Arduino UNO và hình ảnh từ Pi Camera. Sau đó sẽ gửi thông tin cảm biến lên máy chủ bằng giao thức MQTT [26].



Hình 4. Lưu đồ quá trình truyền nhận thông tin từ cảm biến

Trong hình 4 mô tả các bước trong quá trình truyền nhận thông tin từ cảm biến. Sau khi Raspberry Pi nhận được dữ liệu cảm biến từ Arduino UNO thông qua cổng USB, dữ liệu sẽ được đưa vào tin nhắn theo chuẩn của MQTT bao gồm các thuộc tính cơ bản như Topic, data, QoS (Qualities of service)... Sau đó tin nhắn này sẽ được gửi lên Broker MQTT (trong hệ thống này là Mosquitto Broker). Broker sau khi nhận được tin nhắn sẽ kiểm tra chủ đề trong

tin nhắn và thông báo tới máy chủ đang theo dõi chủ đề đó biết rằng có tin nhắn gửi lên để máy chủ lấy tin nhắn về.

3.3. Hệ thống camera giám sát

Mô hình đề xuất cho hệ thống cảnh báo cháy nổ bằng camera có nhiều điểm tương đồng với một hệ thống camera giám sát thông thường. Với điều kiện thực tế tại Việt Nam, chúng tôi phát triển hệ thống như một gói nâng cấp của hệ thống Camera IP giám sát thông dụng trên thị trường hiện nay. Hệ thống phần cứng của chúng tôi sử dụng đường mạng hữu tuyến bao gồm các thành phần chính sau:

- 04 Camera IP giám sát Hikvision model DS-2CD2121G0-I
- 01 Đầu ghi loại NVR Hikvision DS-7104NI-Q1/M
- 01 Router Gigabite Ethernet (thí dụ CISCO SG95-16)
- (Tùy chọn) Wireless bridge cho các khu vực không thể kéo dây

Hiện nay việc giải mã video là một công việc tương đối nhẹ nhàng với máy tính cá nhân và máy tính nhúng. Với hệ thống đề xuất, quá trình giải mã cần đảm bảo đáp ứng thời gian thực với nhiều luồng video phân giải cao. Hai giải pháp được chúng tôi sử dụng để giải quyết vấn đề này là FFMpeg và NVCodec:

- FFMPEG là một framework hàng đầu về đa phương tiện (xử lý âm thanh, hình ảnh). Nó có thể decode (giải mã), encode (mã hóa), transcode (chuyển mã), mux (ghép kênh), demux (phân kênh, tách kênh), stream (ví dụ như livestream trên youtube, facebook...), filter (lọc) và play (chạy, phát video) rất nhiều thứ mà con người hay máy móc tạo ra. FFMPEG hỗ trợ hầu hết các định dạng. Và nó khá là linh hoạt, có thể compile, run và chạy trên nhiều nền tảng như Linux, Mac OS X, Microsoft Windows, BSD, Solaris... và ở trên nhiều môi trường, kiến trúc khác nhau.

- CUDA là kiến trúc GPU đặc trưng của NVidia phát triển, CUDA không chỉ làm nhiệm vụ đồ họa mà còn thực hiện thêm các nhiệm vụ hỗ trợ tính toán thông dụng. Nhiệm vụ giải mã Video cũng là một trong những tác vụ được kiến trúc CUDA hỗ trợ qua bộ thư viện NVidia Video Codec SDK. Trải qua quá trình phát triển lâu dài, phiên bản 9.1 mới nhất hiện nay đã hỗ trợ giải mã chuẩn nén H265 ở độ phân giải lên tới 8K.

Việc sử dụng NVCodec trong quá trình giải mã giúp tăng hiệu năng xử lý của hệ thống: Giải nén được nhiều luồng video hơn với chi phí tính toán và điện năng thấp hơn, giảm bớt độ trễ trong quá trình giải mã video xuống mô đun xử lý ảnh.

3.4. Mô-đun Sim 900A

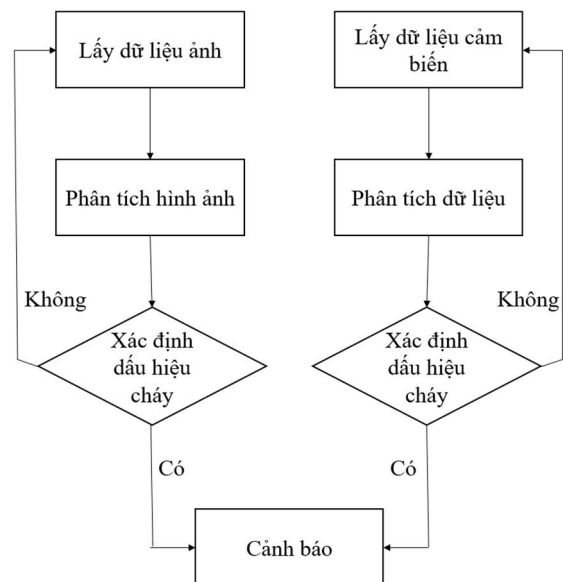
Mô-đun Sim900A là thành phần gửi thông báo cảnh báo khi phát hiện đám cháy. SIM900A là một mô-đun không dây nhỏ gọn và đáng tin cậy. Mô-đun có ngăn xếp TCP/IP nội bộ cho phép bạ kết nối với mạng toàn cầu thông qua GPRS. Nó được kết nối với máy chủ thông qua kết nối UART. Sau khi có cảnh báo cháy, nó sẽ gửi tin nhắn tới những số điện thoại được lưu sẵn trong mô-đun (có thể là

số của cơ quan cứu hỏa, cảnh sát...) một cách liên tục để đảm bảo có sự ứng phó kịp thời với sự cố cháy.

3.5. Phần mềm hệ thống cảnh báo cháy

Trong hình 5 mô tả tổng quan quá trình vận hành của hệ thống. Phần mềm máy chủ đánh giá mức độ cảnh báo cháy bao gồm các mô-đun chính sau đây:

- **Phân tích video:** mô đun có vai trò trích xuất các sự kiện từ các luồng video gửi tới trung tâm xử lý. Đây là luồng xử lý quan trọng của hệ thống do phải giải quyết một lượng thông tin lớn, độ tin cậy cao. Nếu được phân tích ngữ nghĩa đúng cách sẽ làm giảm các trường hợp cảnh báo nhầm.
- **Cảm biến môi trường:** mô đun có vai trò lưu trữ và biểu diễn các thông tin từ các nút cảm biến báo cháy truyền thống. Luồng thông tin này không chỉ giúp chúng ta quyết định cảnh báo cháy, mà còn hỗ trợ việc dự báo các khu vực có nguy cơ cháy cao.
- **Phân tích và quyết định:** đây là nơi phân tích đánh giá thông tin cảnh báo cho cả tòa nhà. Từ các điểm bất thường trên các nút cảm biến và camera trong tòa nhà, kết hợp với kinh nghiệm của các giải thuật học máy để đưa ra các mức cảnh báo phù hợp.
- **Hệ thống cảnh báo:** nhiệm vụ của hệ thống cảnh báo là sau khi nhận được kết quả phân tích môi trường từ dữ liệu thu thập sẽ đưa ra cảnh báo tùy theo kết quả nhận được. Hệ thống cảnh báo có thể gửi tin nhắn đến những số điện thoại trong danh sách và gửi email đến những địa chỉ thư đã được cài đặt.



Hình 5. Tổng quan phần mềm của hệ thống

Các bộ thư viện mã nguồn mở được sử dụng trong thiết kế phần mềm hệ thống bao gồm:

- **Phần mềm giao diện Qt:** Qt là framework tập hợp cung cấp nhiều công cụ, thư viện giúp việc lập trình để tạo nên các giao diện đồ họa người dùng, chủ yếu với các cửa sổ trở nên hữu hiệu hơn. Qt được viết bằng C++ và được

thiết kế để sử dụng trong C++. Hiện nay đã có thể dùng thư viện này với nhiều ngôn ngữ khác như Java hay Python...

- Thư viện xử lý ảnh OpenCV: OpenCV (Open Computer Vision) là một thư viện mã nguồn mở hàng đầu cho xử lý về thị giác máy tính, machine learning, xử lý ảnh. OpenCV được viết bằng C/C++, vì vậy có tốc độ tính toán rất nhanh, có thể sử dụng với các ứng dụng liên quan đến thời gian thực. OpenCV có các interface cho C/C++, Python Java vì vậy hỗ trợ được cho Window, Linux, MacOS lẫn Android, iOS

- Cơ sở dữ liệu MongoDB: MongoDB là một hệ thống quản lý cơ sở dữ liệu nguồn mở (DBMS) sử dụng mô hình hỗ trợ các dạng dữ liệu khác nhau. Đây là một trong nhiều công nghệ cơ sở dữ liệu phát sinh vào giữa những năm 2000 dưới biểu ngữ NoQuery để sử dụng trong các ứng dụng dữ liệu lớn và các công việc xử lý khác.

- Thư viện truyền nhận thông tin cảm biến Eclipse Mosquitto: Eclipse Mosquitto là một nhà môi giới thông báo mã nguồn mở (được cấp phép EPL / EDL) thực hiện các giao thức MQTT phiên bản 5.0, 3.1.1 và 3.1. Mosquitto rất nhẹ và phù hợp để sử dụng trên tất cả các thiết bị từ máy tính nhúng năng lượng thấp đến máy chủ đầy đủ.

- Công cụ tối ưu học máy TensorRT: NVIDIA TensorRT là một nền tảng cho suy luận học tập sâu hiệu suất cao. Nó bao gồm một trình tối ưu hóa suy luận học tập sâu và thời gian chạy mang lại độ trễ thấp và thông lượng cao cho các ứng dụng suy luận học tập sâu. Các ứng dụng dựa trên TensorRT thực hiện nhanh hơn tới 40 lần so với các nền tảng chỉ sử dụng CPU trong quá trình suy luận. Với TensorRT, có thể tối ưu hóa các mô hình mạng thần kinh được đào tạo trong tất cả các nền tảng chính, hiệu chỉnh cho độ chính xác thấp hơn với độ chính xác cao và cuối cùng triển khai đến các trung tâm dữ liệu, thiết bị nhúng.

4. TỐI ƯU KHẢ NĂNG NHẬN DẠNG ĐÁM CHÁY

Trong thực tế, khi hiện tượng cháy xảy ra sẽ xuất hiện khói và sự lan tỏa của khói nhanh hơn khi so sánh với sự lan tỏa của lửa. Với mục đích phát hiện sớm các đám cháy, hệ thống xử lý hình ảnh chỉ tập trung vào nhận dạng các đám khói, đặc biệt là các đám khói nhỏ, độ đậm đặc thấp. Bước đầu tiên trong quá trình nhận dạng này là thử nghiệm độ hiệu quả của học máy với việc phân loại ảnh có khói và ảnh không có khói. Các mạng học sâu phân loại ảnh phổ biến được thử nghiệm bao gồm ResNet-18, ResNet-50, mạng VGG và DenseNet. Các mạng kể trên về sau có nhiệm vụ trích xuất đặc trưng về đám khói, làm tiền đề cho việc khoanh vùng đám khói ở phần sau. Do dữ liệu hình ảnh về các đám khói trong thực tế khá ít, đặc biệt là khói với mật độ mỏng, nên để khắc phục điều này chúng tôi có sử dụng phần mềm giả lập hiệu ứng 3D Blender để tạo ngẫu nhiên ra một số mẫu đám khói và sau đó chèn các mẫu nhân tạo vào trong tập dữ liệu có sẵn. Với trường hợp là cảnh báo cháy trong nhà, khu dân cư, tập dữ liệu mở được sử dụng là SUN RGB-D [27]. Trong hình 6 là hình ảnh minh họa cho việc sinh ra dữ liệu giả lập nhằm khắc phục việc thiếu dữ liệu cho việc huấn luyện. Với việc kết hợp như vậy tổng số

mẫu dữ liệu sinh ra đa dạng và đảm bảo sự cân bằng dữ liệu. Cơ cấu tổ chức dữ liệu được mô tả trong bảng 1. Kết quả huấn luyện được mô tả trong bảng 2. Với những kết quả từ thực nghiệm chúng tôi quyết định chọn mô hình ResNet-50 làm mạng trích xuất đặc trưng cho việc xây dựng mạng xác định đám khói.



Hình 6. Minh họa chèn khói nhân tạo vào bộ dữ liệu SUN RGB-D

Bảng 1. Cơ cấu số lượng mẫu trong việc huấn luyện

Tập dữ liệu	Ảnh thường	Ảnh có khói	Tổng
Huấn luyện	1350	1580	2930
Thử nghiệm	870	835	1705
Kiểm tra	850	850	1700
Tổng	3070	3265	6335

Bảng 2. Độ chính xác của các mạng phân loại

Loại mạng	Độ chính xác
ResNet-18	95,40%
ResNet-50	98,85%
VGG	96,70%
DenseNet	97,15%

Xác định vật thể để cập đến khả năng của hệ thống máy tính và phần mềm để định vị các đối tượng trong một hình ảnh và xác định từng đối tượng. Việc áp dụng đột phá và nhanh chóng của deep learning vào năm 2012 đã đưa vào sự tồn tại các thuật toán và phương pháp phát hiện đối tượng hiện đại và chính xác cao như R-CNN, Fast-RCNN, Faster-RCNN, RetinaNet và nhanh hơn nhưng rất chính xác như SSD và YOLO. Qua quá trình tìm hiểu, thực nghiệm với các mạng đã nêu ra ở trên, chúng tôi nhận ra các mạng này có các ưu nhược điểm:

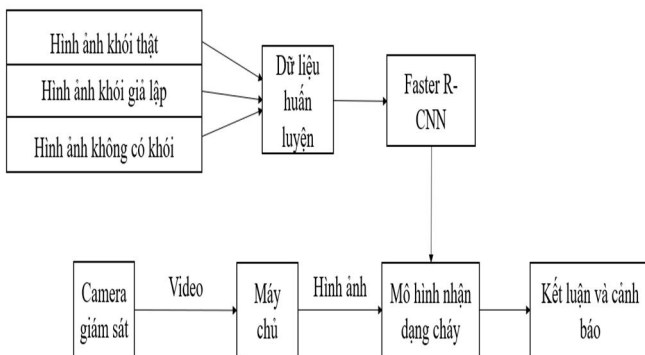
- Các mô hình mạng SSD về có tốc độ trung bình nhanh hơn nhưng không thể tốt hơn Faster R-CNN ở độ chính xác
- Faster R-CNN cần ít nhất 100ms để xử lý một ảnh
- Sử dụng bản đồ thuộc tính phân giải thấp giảm độ chính xác của tất cả các phương pháp

- Phân giải ảnh đầu vào ảnh hưởng đáng kể đến độ chính xác của hệ thống

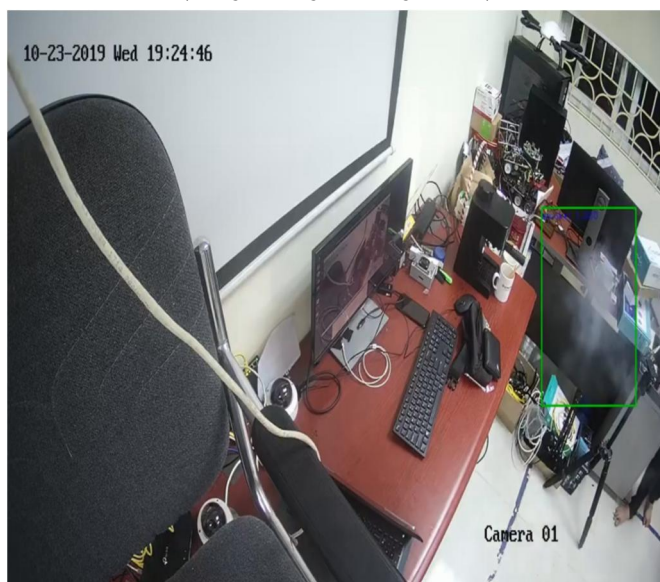
- Việc lựa chọn bộ trích xuất dữ liệu ảnh hưởng nhiều tới kết quả của Faster R-CNN nhưng ít ảnh hưởng tới SSD

Trong phần thực nghiệm, yếu tố chính xác được đặt lên hàng đầu, vì vậy mạng Faster R-CNN được chúng tôi lựa chọn để làm đối tượng thử nghiệm. Điểm cải thiện lớn nhất của Faster R-CNN so với các mạng thể hệ trước là Fast R-CNN là tốc độ. Thay vì sử dụng các phương pháp chậm chạp như Selective Search và Edge Boxes, Faster R-CNN sử dụng một mạng học sâu rất nhỏ để đề xuất các khu vực ưu tiên.

Sau khi thực hiện các thí nghiệm đã mô tả ở trên, chúng tôi quyết định xây dựng hệ thống nhận diện xác định đám khói bằng mô hình FasterR-CNN với mạng trích xuất đặc trưng là ResNet-50. Lưu đồ tổng thể của quá trình cải thiện khả năng phát hiện đám cháy bằng hình ảnh được mô tả trong hình 7. Sau 50 lần huấn luyện, mỗi lần huấn luyện có 6000 vòng lặp, độ chính xác chúng tôi đạt được khi thử nghiệm với bộ dữ liệu xây dựng ở trên là 89,7% và quá trình phát hiện duy trì tốc độ xử lý 0,45ms. Kết quả khi thử nghiệm hệ thống với một camera giám sát thực được hình họa trong hình 8.



Hình 7. Lưu đồ xây dựng hệ thống nhận dạng đám cháy



Hình 8. Kết quả thực nghiệm hệ thống với môi trường thực

5. KẾT LUẬN

Trong bài báo này chúng tôi đã mô tả hệ thống nhận dạng, cảnh báo sớm đám cháy thông qua phát hiện đám khói sử dụng công nghệ học máy. Hệ thống được xây dựng từ nhu cầu thực tế, tận dụng hệ thống phần cứng phổ thông như camera giám sát, cảm biến nhiệt độ độ ẩm... Công nghệ học máy được ứng dụng nhằm tăng khả năng cũng như độ chính xác cho hệ thống. Vấn đề thiếu hụt dữ liệu cho quá trình huấn luyện được khắc phục thông qua việc sử dụng hình ảnh khói nhân tạo. Các công nghệ, kỹ thuật cập nhật của học máy trong bài toán nhận diện, phân loại hình ảnh đã được chúng tôi thử nghiệm, đánh giá. Với thực trạng về cháy nổ hiện tại và qua các phân tích về các hệ thống đang sử dụng hệ thống hứa hẹn cải thiện được hiệu quả, giảm thiểu thiệt hại. Trong tương lai, hệ thống sẽ tiếp tục được phát triển kết hợp nhận diện khói và lửa đồng thời nhằm tăng khả năng phát hiện đám cháy trong những điều kiện khác nhau.

LỜI CẢM ƠN

Nghiên cứu này được hỗ trợ bởi Bộ Khoa học và Công nghệ trong đề tài "Khai thác sáng chế trong lĩnh vực xử lý và nhận dạng ảnh nhằm ứng dụng giám sát, cảnh báo cháy", hợp đồng số 01/2019/VSCCN-ĐTCB.

TÀI LIỆU THAM KHẢO

- [1]. <https://vov.vn/tin-24h/tu-dau-nam-2017-ca-nuoc-xay-ra-hon-3000-vu-chay-lam-chet-75-nguoi-678519.vov>.
- [2]. T. H. Chen, P. H. Wu, and Y. C. Chiou, 2004. *An early fire-detection method based on image processing*. Proc. - Int. Conf. Image Process. ICIP, vol. 3, pp. 1707–1710.
- [3]. J. Seebamrungsat, S. Praising, and P. Riyamongkol, 2014. *Fire detection in the buildings using image processing*. Proc. 2014 3rd ICT Int. Sr. Proj. Conf. ICT-ISPC 2014, pp. 95–98.
- [4]. B. H. Cho, J. W. Bae, and S. H. Jung, 2008. *Image processing-based fire detection system using statistic color model*. Proc. - ALPIT 2008, 7th Int. Conf. Adv. Lang. Process. Web Inf. Technol., pp. 245–250.
- [5]. W. Tjokorda Agung Budi and I. S. Suardi, 2011. *Fire alarm system based-on video processing*. Proc. 2011 Int. Conf. Electr. Eng. Informatics, ICEEI 2011.
- [6]. J. Shao, G. Wang, and W. Guo, 2012. *An image-based fire detection method using color analysis*. Proc. - 2012 Int. Conf. Comput. Sci. Inf. Process. CSIP 2012, pp. 1008–1011.
- [7]. J. Z. Zhou and G. Song, 2010. *Research on the technology of fire detection based on image processing in unmanned substation*. Proc. - 3rd Int. Conf. Intell. Networks Intell. Syst. ICINIS 2010, pp. 108–111.
- [8]. D. Pritam and J. H. Dewan, 2017. *Detection of fire using image processing techniques with LUV color space*. 2017 2nd Int. Conf. Conver. Technol. I2CT 2017, vol. 2017-Janua, pp. 1158–1162.
- [9]. S. Bayoumi, E. AlSobky, M. Almohsin, M. Altwaim, M. Alkaldi, and M. Alkahtani, 2013. *A real-time fire detection and notification system based on computer vision*. 2013 Int. Conf. IT Conver. Secur. ICITCS 2013.

- [10]. T. Celik and K. K. Ma, 2008. *Computer vision based fire detection in color images*. SMCia/08 - Proc. 2008 IEEE Conf. Soft Comput. Ind. Appl., pp. 258–263.
- [11]. M. Mueller, P. Karasev, I. Kolesov, and A. Tannenbaum, 2013. *Optical flow estimation for flame detection in videos*. IEEE Trans. Image Process., vol. 22, no. 7, pp. 2786–2797.
- [12]. Sunita Harlapur and Dr. K. R. Nataraj, 2015. *Fire Detection using Optical Flow Method in Videos*. Int. J. Eng. Res., vol. V4, no. 05, pp. 710–713.
- [13]. S. Rinsurongkawong, M. Ekpanyapong, and M. N. Dailey, 2012. *Fire detection for early fire alarm based on optical flow video processing*. 2012 9th Int. Conf. Electr. Eng. Comput. Telecommun. Inf. Technol. ECTI-CON 2012.
- [14]. K. Muhammad, J. Ahmad, and S. W. Baik, 2018. *Early fire detection using convolutional neural networks during surveillance for effective disaster management*. Neurocomputing, vol. 288, pp. 30–42.
- [15]. M. Sultan Mahmud, M. S. Islam, and M. A. Rahman, 2017. *Smart Fire Detection System with Early Notifications Using Machine Learning*. Int. J. Comput. Intell. Appl., vol. 16, no. 2, pp. 1–17.
- [16]. V. Tipsuwanporn, V. Krongratana, S. Gulpanich, and K. Thongnopakun, 2006. *Fire detection using neural network*. 2006 SICE-ICASE Int. Jt. Conf., pp. 5474–5477.
- [17]. Z. Q. Zhao, P. Zheng, S. T. Xu, and X. Wu, 2019. *Object Detection with Deep Learning: A Review*. IEEE Trans. Neural Networks Learn. Syst., vol. 30, no. 11, pp. 3212–3232.
- [18]. K. He, X. Zhang, S. Ren, and J. Sun, 2016. *Deep residual learning for image recognition*. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2016-Decem, pp. 770–778.
- [19]. C. Szegedy *et al.*, 2015. *Going deeper with convolutions*. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 07-12-June, pp. 1–9.
- [20]. A. G. Howard *et al.*, 2017. *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications*.
- [21]. S. Ioffe and C. Szegedy, 2015. *Batch normalization: Accelerating deep network training by reducing internal covariate shift*. 32nd Int. Conf. Mach. Learn. ICML 2015, vol. 1, pp. 448–456.
- [22]. T. Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, 2017. *Feature pyramid networks for object detection*. Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017, vol. 2017-Janua, pp. 936–944.
- [23]. S. Ren, K. He, R. Girshick, and J. Sun, 2017. *Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks*. IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137–1149.
- [24]. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, 2016. *You only look once: Unified, real-time object detection*. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2016-Decem, pp. 779–788.
- [25]. W. Liu *et al.*, 2016. *SSD: Single shot multibox detector*. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 9905 LNCS, pp. 21–37.
- [26]. D. Soni and A. Makwana, 2017. *A survey on mqtt: a protocol of internet of things (IoT)*. Int. Conf. Telecommun. Power Anal. Comput. Tech. (Ictpact - 2017), no. April, pp. 0–5.
- [27]. S. Song, S. P. Lichtenberg, and J. Xiao, 2015. *SUN RGB-D: A RGB-D scene understanding benchmark suite*. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 07-12-June, pp. 567–576.

AUTHORS INFORMATION

**Pham Ngoc Pha¹, Nguyen Trong Hieu¹, Nguyen Viet Thang²,
Nguyen Truong Son², Quach Cong Hoang², Pham Minh Trien²**

¹National Institute of Patent and Technology EXploitation, Ministry of Science and Technology

²VNU University of Engineering and Technology