

MÔ HÌNH LƯỚI PHÂN LOẠI CHO PHÁT HIỆN ĐỐI TƯỢNG TỪ HÌNH ẢNH CAMERA

Nguyễn Đăng Bình

Khoa Công nghệ Thông tin, Trường Đại học Khoa học, Đại học Huế

Email: ndbinh@husc.edu.vn

Ngày nhận bài: 11/6/2021; ngày hoàn thành phản biện: 16/6/2021; ngày duyệt đăng: 02/11/2021

TÓM TẮT

Mô hình lưới phân loại đã cho thấy là một lựa chọn đáng quan tâm để phát hiện đối tượng từ hình ảnh camera. Bằng cách áp dụng một bộ phân loại đơn cho mỗi vùng cụ thể trên hình ảnh. Mục tiêu của nghiên cứu này là để tăng hiệu năng của hệ thống phát hiện đối tượng trong khi vẫn giữ được tính ổn định, độ chính xác cũng như tốc độ trực tuyến của việc phát hiện đối tượng từ Camera. Đóng góp của nghiên cứu là mô hình lưới phân loại trên ý tưởng liên kết bộ phân loại ngoại tuyến với bộ phân loại trực tuyến trong một lưới theo cách tiếp cận thích nghi nhằm gia tăng sự ổn định trong phát hiện đối tượng khi môi trường và đối tượng thay đổi. Kết quả thực nghiệm cho thấy kết quả phân loại độ chính xác cao với sự hiện diện của các đối tượng không di chuyển, các đối tượng có kích thước và tư thế khác nhau trong môi trường phức tạp.

Từ khóa: phát hiện đối tượng, học trực tuyến, lưới phân loại.

1. MỞ ĐẦU

Với ngày càng gia tăng số lượng của camera giám sát cũng như nhu cầu của các hệ thống giám sát đồng bộ dữ liệu trực quan hình ảnh đang là một đòi hỏi thực tế ngày càng gia tăng. Một trong những bước đầu tiên trong nhiều ứng dụng trong hệ thống giám sát thông minh xác định các đối tượng, hướng tới giám sát trực quan; phát hiện đối tượng từ các camera trong môi trường thực. Để đảm bảo phát hiện đối tượng chính xác mà không cần sự can thiệp của con người chúng ta phát triển các cách tiếp cận khác nhau mà ở đó cho phép kết hợp với thông tin của khung cảnh cụ thể tại từng thời điểm khác nhau. Cách tiếp cận nổi bật phổ biến nhất là áp dụng một kỹ thuật cửa sổ trượt [1, 2, 3, 4, 5, 7, 8]. Mỗi vùng ảnh của một hình ảnh nhất định được thử nghiệm cho dù đó là phù hợp với một mô hình ước tính trước đó hay không, và cuối cùng tất cả các vùng ảnh phù hợp được thông báo kết quả. Thông thường, mục tiêu của phương pháp này là xây dựng một mô hình chung mà ở đó có thể áp dụng cho tất cả

các kịch bản có thể, và các bài toán phát hiện đối tượng khác nhau [7, 8, 12]. Các nghiên cứu này làm tốn kém nhiều công sức và mất nhiều thời gian để chuẩn bị mẫu. Thêm vào đó do các mẫu chuẩn bị từ trước nên khi áp dụng vào các khung cảnh mới để phát hiện đối tượng thì có thể không phát huy hiệu quả, muốn hiệu quả thì phải huấn luyện lại với các mẫu mới hoặc cập nhật bổ sung thích nghi các mẫu trong khung cảnh mới này. Điều này phải có mô hình mới, lưới phân loại với cách tiếp cận học máy mới trên cơ sở kết hợp học ngoại tuyến và trực tuyến.

Đóng góp của bài báo gồm: (1) Mô hình lưới phân loại dựa trên kết hợp học trực tuyến và ngoại tuyến với chiến lược cập nhật có hiệu quả nhằm gia tăng sự ổn định trong phát hiện đối tượng khi môi trường và đối tượng thay đổi; (2) Xây dựng mô phỏng thực nghiệm với các bộ dữ liệu benchmark được dùng rộng rãi so sánh trong các nghiên cứu được công bố công khai..

Phần còn lại của bài báo được tổ chức như sau: Trong phần 2, đề cập đến các vấn đề nghiên cứu liên quan. Trong phần 3, giới thiệu mô hình lưới bộ phân loại dựa trên liên kết học ngoại tuyến và trực tuyến dùng cho phát hiện đối tượng. Phần 4 đưa ra đánh giá thực nghiệm của phương pháp đề xuất. Cuối cùng, kết luận ở mục 5.

2. ĐÁNH GIÁ MỘT SỐ NGHIÊN CỨU LIÊN QUAN

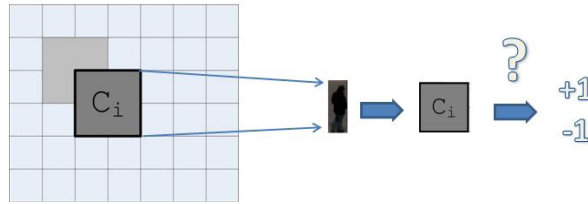
Để cải thiện sức mạnh bộ phân loại và tiếp tục giảm số lượng mẫu huấn luyện của một bộ phân loại thích nghi sử dụng thuật toán học trực tuyến có thể được áp dụng [11]. Vì vậy, hệ thống có thể thích nghi sự các thay đổi của môi trường (ví dụ, thay đổi các điều kiện của ánh sáng) và những biến đổi mà không cần phải xử lý bởi các mô hình ban đầu. Trong thực tế, theo cách này sự phức tạp của bài toán là giảm và bộ phân loại có thể được huấn luyện hiệu quả hơn. Các hệ thống thích nghi có một nhược điểm: dữ liệu mới chưa được gán nhãn sẽ được đưa vào một mô hình đã được xây dựng. Cách tiếp cận này điển hình là tự huấn luyện [14], huấn luyện đồng thời [4, 13], học bán giám sát [8]. Các phương pháp bán giám sát, thường được sử dụng bởi kết hợp những thông tin cho trước và khai phá các mẫu mới từ dữ liệu có sẵn để hình thành nên một bộ phân loại. Phương pháp tự huấn luyện hay huấn luyện đồng bộ thường gặp những hạn chế về các ràng buộc lý thuyết không thể đảm bảo trên thực tế hoặc là dựa vào phản hồi của các bộ phân loại hiện hành, cả hai kết quả phân loại không đáng tin cậy. Các bộ phân loại hữu hiệu hơn tránh những vấn đề trên có thể được huấn luyện sử dụng lưới các bộ phân loại [6, 9, 10]. Ngược lại với kỹ thuật cửa sổ trượt, ở đó một bộ phân loại được lượng hóa với các vị trí khác nhau trên ảnh, ý tưởng chính của lưới bộ phân loại là huấn luyện các bộ phân loại riêng biệt cho mỗi vị trí khác nhau của hình ảnh. Như vậy, sự phức tạp của nhiệm vụ phân loại đã được xử lý bởi một bộ phân loại đơn vì vậy độ phức tạp được giảm đáng kể. Mỗi bộ phân loại là chỉ có thể phân biệt đối tượng cần phát hiện từ nền ảnh tại một vị trí cụ thể trong ảnh.

Bằng cách sử dụng hệ thống các bộ phân loại trực tuyến có thể thích nghi với sự thay đổi của các điều kiện môi trường, làm giảm hơn nữa sự phức tạp của các bộ phân loại.

3. MÔ HÌNH LƯỚI PHÂN LOẠI

3.2. Lưới phân loại

Ý tưởng chính của lưới phân loại là khai thác kiến thức của bộ phân loại sẵn có cho trước và ràng với camera là cố định. Bằng cách sử dụng thông tin này, toàn bộ nhiệm vụ phát hiện đối tượng có thể đơn giản hóa để lấy mẫu từ đầu vào của hình ảnh sử dụng cố định lưới lồng vào nhau (cả vị trí và tỉ lệ), trong đó mỗi phần tử của lưới $i=1..N$ tương ứng với một bộ phân loại C_i . Điều này được minh họa trong Hình 1. Như vậy, nhiệm vụ phân loại được thực hiện theo bộ phân loại C_i đơn giản hóa sự phân biệt với nền ảnh với phần tử lưới cụ thể từ đối tượng cần được quan tâm. Hơn nữa, các camera cố định cho phép ước tính mặt phẳng của ảnh nền, trong đó tiếp tục giúp giảm số lượng các bộ phân loại trong lưới phân loại. Do đơn giản hóa này các bộ phân loại ít phức tạp hơn có thể được áp dụng. Đặc biệt, biểu diễn dựa trên lưới là rất thích hợp cho các bộ phân loại trực tuyến gọn nhỏ, có thể được đánh giá lượng hóa và cập nhật rất hiệu quả.

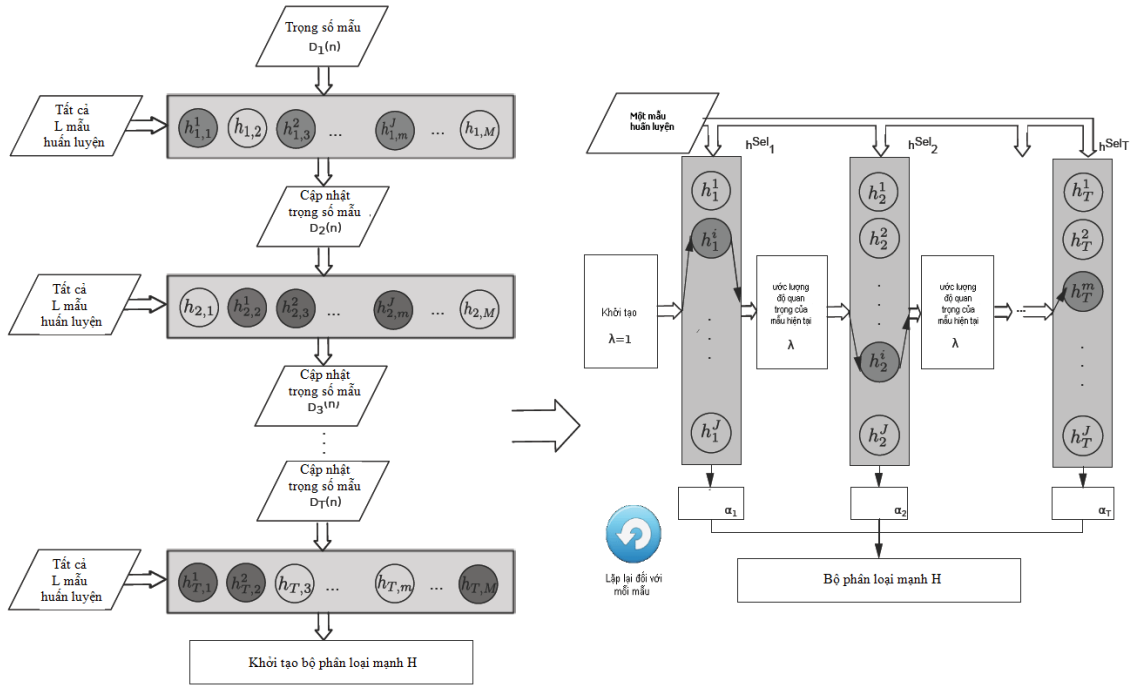


Hình 1. Ý tưởng chính của lưới phân loại theo nguyên tắc chia để trị. Ảnh được chia thành các vùng với lưới chồng lấp cao về cả vị trí và tỉ lệ, trong đó mỗi phần tử lưới có một bộ phân loại riêng của nó và chịu trách nhiệm phát hiện đối tượng trong vùng đó.

Tại mỗi thời điểm t chiến lược cập nhật cố định được sử dụng cho cả mẫu dương và mẫu âm đối với bộ phân loại C_{i-1} . Cho một tập biểu diễn mẫu dương được gán nhãn X^+ , sau đó sử dụng $\langle x, +1 \rangle$, $x \in X^+$. Để cập nhật bộ phân loại một mẫu dương đúng theo định nghĩa. Xác suất mà một đối tượng xuất hiện trong mẫu x_i được cho bởi $p(x_i, object) = \frac{\#p_i}{\Delta t}$; ở đó $\#p_i$ là số đối tượng xuất hiện trong một vùng cụ thể trong một khoảng thời gian Δt . Do đó, cập nhật mẫu âm với mẫu hiện tại tương ứng cho bộ phân loại tương ứng với lưới đã thiết lập $\langle x_{i,t}, -1 \rangle$, $x \in X^+$ là chính xác với phần lớn thời gian với xác suất $p(x_i = object)$. Với các tiếp cận này, xác suất của cập nhật sai cho một mẫu cụ thể là thật sự rất thấp.

3.2. Bộ phân loại dựa trên kết hợp học ngoại tuyến và trực tuyến

Kết hợp Boosting ngoại tuyến để lựa chọn đặc trưng với Boosting trực tuyến để lựa chọn đặc tính cho phép kết hợp thông tin từ các dữ liệu cho trước với thông tin mới mà những thông tin này không có sẵn khi huấn luyện bộ phân loại ngoại tuyến. Kết hợp này được minh họa trong Hình 2.



Hình 2. Mô hình kết hợp học ngoại tuyến và trực tuyến: mô tả hoạt động của thuật toán Boosting ngoại tuyến và Boosting trực tuyến lựa chọn đặc trưng.

Thuật toán 1: Huấn luyện ngoại tuyến

Vào: Tập mẫu huấn luyện

$$S = \{(x_1, y_1), \dots, (x_L, y_L)\} \mid y_i \in \{-1, +1\}.$$

Ra: $H(x) = \text{sign}(\sum_{t=1}^T \alpha_t h_t(x)).$

Phương pháp:

1. Khởi tạo các trọng số $D_1(i) = \frac{1}{L}$;
2. **for** $t = 1, 2, \dots, T$ **do**
3. Đối với mỗi đặc trưng j huấn luyện một bộ phân loại yếu $h_j: X \rightarrow Y$ với lỗi đối với phân bố D_t

$$4. e_j^{\text{off-line}} = \sum_{i=1}^L D_t(n).I(h_j(x_i) \neq y_i) ;$$

5. Chọn J bộ phân loại yếu tốt nhất để khởi tạo chọn bộ chọn t với các đặc trưng phù hợp

$$6. \text{Chọn } \alpha_t = \ln \frac{1 - e_t^{\text{off-line}}}{e_t^{\text{off-line}}} ;$$

7. Cập nhật phân bố các trọng số

$$D_{t+1(i)} = \frac{D_t(i)}{Z_t} \begin{cases} \exp(-\alpha_t) & h(x_i) = y_i \\ \exp(\alpha_t) & h(x_i) \neq y_i \end{cases}$$

8. **end for**

Thuật toán 2: Huấn luyện trực tuyến kết hợp với huấn luyện ngoại tuyến

Vào: - Mẫu huấn luyện $\langle x, y \rangle, y \in \{-1, +1\}$;

- Khởi tạo $\lambda_{t,j}^{corr} = \lambda_{t,j}^{wrong} = 1$;

- Khởi tạo trọng số $\lambda = 1$.

Ra: $H(x) = \text{sign}(\sum_{t=1}^T \alpha_t h_t(x))$.

Phương pháp:

1. **for** $t = 1, 2, \dots, T$ **do**

2. **for** $j = 1, 2, \dots, J$ **do**

3. $h_{t,j} = \text{update}(h_{t,j} \langle x, y \rangle, \lambda)$;

4. **if** $h_{t,j}(x) = y$ **then** $\lambda_{t,j}^{corr} = \lambda_{t,j}^{corr} + \lambda$;

5. **else** $\lambda_{t,j}^{wrong} = \lambda_{t,j}^{wrong} + \lambda$;

6. $e_{t,j} = \frac{1}{2}(e_{t,j}^{off-line} + e_{t,j}^{on-line})$;

7. **end for**

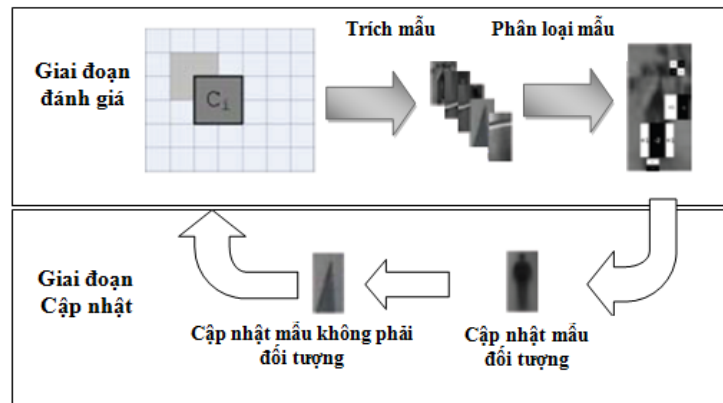
8. $j^+ = \arg \min_j (e_{t,j})$; $e_t = e_{t,j^+}$; $h_t = h_{t,j^+}$;

9. $\alpha_t = \frac{1}{2} \ln \left(\frac{1 - e_t}{e_t} \right)$; $\lambda = \lambda \times \begin{cases} \frac{1}{2(1 - e_t)} & h_t(x_i) = y_i \\ \frac{1}{2e_t} & h_t(x_i) \neq y_i \end{cases}$

10. **end for**

3.3. Áp dụng mô hình kết hợp học ngoại tuyến và học trực tuyến cho lưới phân loại

Với ý tưởng xây dựng một hệ thống phát hiện đối tượng dựa trên lưới các bộ phân loại. Cụ thể, áp dụng phương pháp học liên kết Boosting trực tuyến với học ngoại tuyến để huấn luyện bộ phân loại với việc sử dụng quy tắc cập nhật cố định có xác minh dựa vào bộ phân loại đối tượng đã được huấn luyện trước. Dựa vào đó, hệ thống phát hiện đối tượng dựa trên lưới các bộ phân loại bao gồm giai đoạn đánh giá và cập nhật được mô tả trong Hình 3.



Hình 3. Lưới các bộ phân loại. Mỗi phần tử lưới là một bộ phân loại độc lập, có kích thước cố định. Để thích ứng với sự thay đổi của khung cảnh, từng bộ phân loại được cập nhật bằng quy tắc cập nhật kết hợp với việc xác minh các mẫu sử dụng một bộ phân loại đã được huấn luyện trước.

Để khai thác các thông tin có sẵn cho trước, bài báo đề xuất cách tiếp cận liên kết Boosting ngoại tuyến và trực tuyến với nhau. Boosting ngoại tuyến dùng để khởi tạo các bộ phân loại với các đặc trưng phù hợp với từng bài toán cụ thể. Quá trình lựa chọn đặc trưng cho phép xử lý các tình huống thay đổi một cách hiệu quả bằng cách

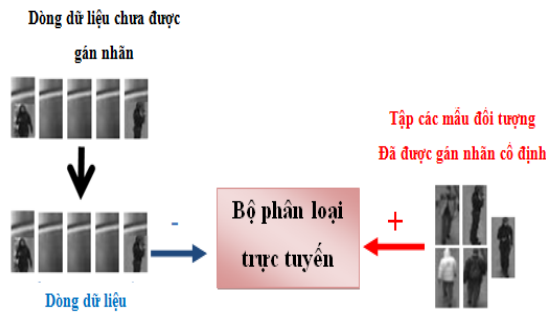
chuyển đổi giữa các đặc trưng khác nhau và lựa chọn các đặc trưng phù hợp nhất cho bài toán thực tế. Vì vậy, phát triển thuật toán Boosting ngoại tuyến [2] thành thuật toán Boosting ngoại tuyến có lựa chọn đặc trưng được mô tả trong Thuật toán 1, thay vì chỉ lựa chọn một bộ phân loại yếu tốt nhất tại mỗi vòng lặp thì chúng ta chọn J bộ phân loại yếu tốt nhất tại mỗi vòng lặp. Hơn thế nữa, để giữ cho các thông tin từ Boosting ngoại tuyến lựa chọn đặc trưng, bài báo cải tiến thuật toán Boosting trực tuyến [3] cho lựa chọn đặc trưng như trong Thuật toán 2. Ở đó việc tính lỗi đã được sửa đổi tính toán lại một lỗi kết hợp dựa trên cả lỗi ngoại tuyến cũng như lỗi trực tuyến $e_{t,j} = \frac{1}{2}(e_{t,j}^{off-line} + e_{t,j}^{on-line})$. Bằng cách sử dụng Thuật toán 2, một bộ học trực tuyến, một hệ thống dựa trên lưới phân loại có khả năng phát hiện thích nghi cao nhưng ổn định với những thay đổi của đối tượng và môi trường và những hạn chế của Roth và các cộng sự [6] được giải quyết.

Để xây dựng bộ phát hiện nói trên cần thực hiện các bước sau:

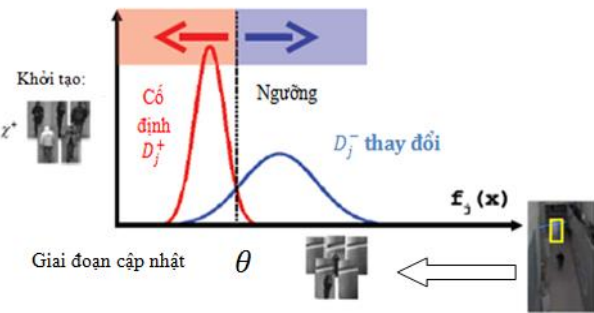
Huấn luyện bộ phân loại ngoại tuyến trước: Cho một tập cố định X^+ các mẫu đối tượng cần huấn luyện. Bước đầu tiên ta huấn luyện một bộ phân loại sử dụng thuật toán boosting ngoại tuyến. Với các đặc tính f_j được chọn ta có thể ước tính được phân phối D_t^+ tương ứng cũng như lỗi $e_+^{off-line}$ trong quá trình huấn luyện, và lỗi này được giữ cố định trong suốt quá trình huấn luyện trực tuyến sau này.

Cập nhật bộ phân loại trực tuyến: Phân phối D_t^- các mẫu không phải đối tượng được cập nhật thông qua các vùng ảnh hiện tại trong khi phân phối D_t^+ được giữ cố định. Dựa vào hai phân phối D_t^+ và D_t^- ta có thể xây dựng được mô hình phân biệt đối tượng/ không đối tượng cho mỗi phân loại yếu tương ứng với mỗi đặc tính. Hình 5. thể hiện sự phân phối D_t^+ và D_t^- .

Lựa chọn đặc tính trực tuyến: trong quá trình lựa chọn các đặc tính, các lỗi huấn luyện cũng được tính toán. Như đã trình bày ở phần trước, bộ phân loại yếu có lỗi nhỏ nhất sẽ được chọn. Cuối cùng, ta được một phân loại mạnh $H(x)$. Ta sử dụng $H(x)$ này để đánh giá trên các khung hình mới, nếu các $H(x)$ trả về kết quả lớn hơn ngưỡng cho trước (ví dụ: lớn hơn 0) thì chỉ ra đó là đối tượng, ngược lại thì không phải đối tượng.

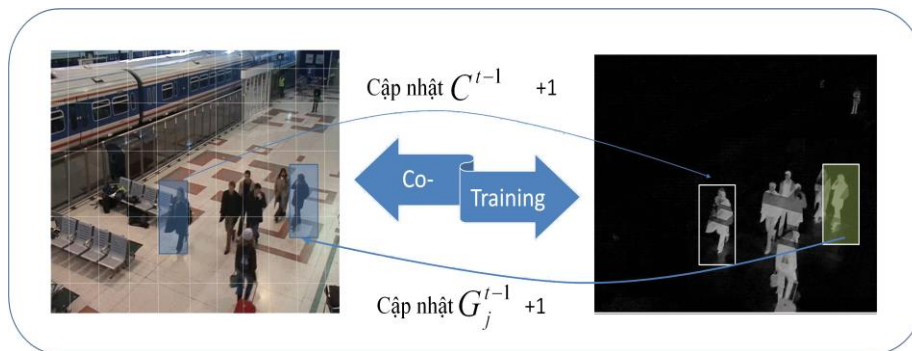


Hình 4. Quy tắc cập nhật.

Hình 5. Với mỗi đặc tính f_j , ngưỡng θ được tính toán dựa trên hai phân phối D_j^+ và D_j^- .

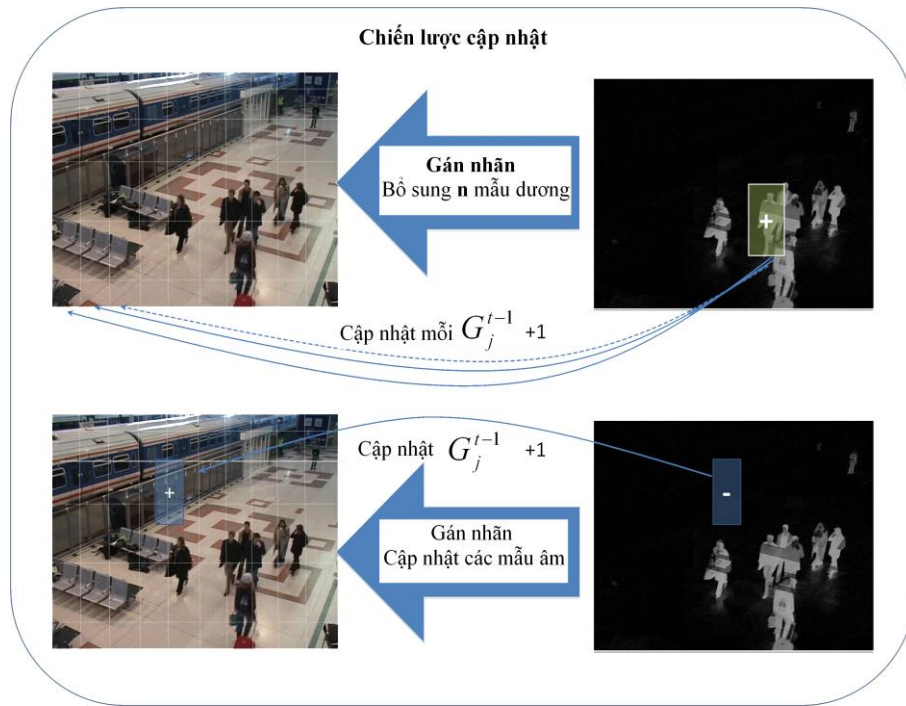
Phát hiện đối tượng và chiến lược cập nhật

Trong giai đoạn đầu, hệ thống được huấn luyện một cách đồng bộ như thể hiện trong Hình 4. Cho lưới có n bộ phân loại G_j hoạt động trên các vùng ảnh X_j và một bộ phân loại C được khởi động cùng theo phương thức trượt cửa sổ trên nền trừ ảnh nền B . Để bắt đầu huấn luyện đồng bộ, bộ phân loại G_j cũng như bộ phân loại được khởi động cùng với bộ phân loại được huấn luyện ngoại tuyến (Thuật toán 1). Các bộ phân loại trong lưới G_j và bộ phân loại C thao tác trên nền trừ ảnh đồng thời với nhau. Một sự phân lớp đáng tin cậy của bộ phân loại G_j được sử dụng để cập nhật bộ phân loại C với nền trừ đại diện tại vị trí j . Ngược lại, một sự phân lớp có độ tin cậy C tại vị trí j được tạo ra một mẫu cập nhật cho bộ phân loại G_j . Thông tin cho trước của bộ phân loại ngoại tuyến đã thu giữ các thông tin chung gây ra một số lượng nhỏ các bản cập nhật đủ khả năng thích nghi với các bộ phân loại trong khung cảnh hình ảnh mới.



Hình 6. Giai đoạn khởi tạo lưới: lưới phân loại ở phía bên trái là được đồng huấn luyện với một bộ phân loại độc lập hoạt động trên hình ảnh loại bỏ nền ở bên phải.

Bộ phân loại C được sử dụng như là sinh ra dữ liệu mẫu dương để cho cập nhật cũng như mẫu âm cập nhật cho các bộ phân loại trong lưới (Thuật toán 2). Cập nhật đối tượng dương được lan rộng cho tất cả các bộ phân loại trong lưới trong khi đó cập nhật các mẫu âm được thực hiện đối với mỗi bộ phân loại cụ thể trong lưới.



Hình 7. Hình ảnh minh họa giai đoạn phát hiện và cập nhật đồng thời các bộ phân loại trên lưới.

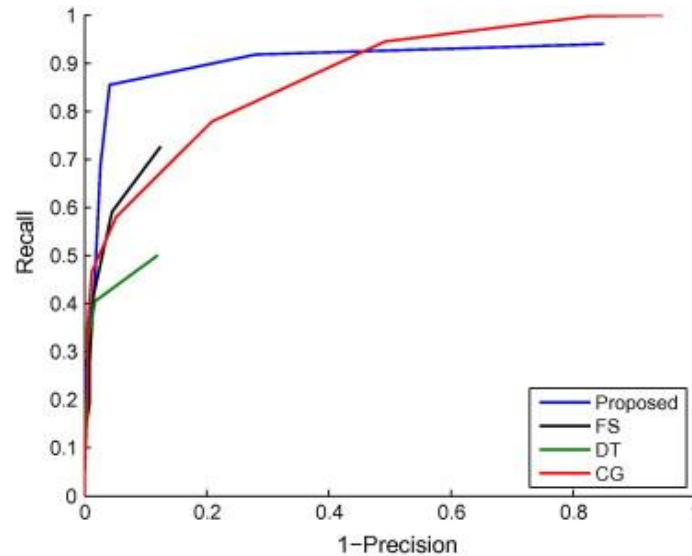
4. THỰC NGHIỆM VÀ KẾT QUẢ

Để minh chứng những điểm mạnh của cách tiếp cận đề xuất, nghiên cứu thực hiện 2 thực nghiệm, đối tượng người đi bộ. Nghiên cứu lựa chọn một số bộ dữ liệu được công bố công khai dùng cho lượng hóa kết quả nghiên cứu để tiến hành thực nghiệm. Từ các thí nghiệm những lợi ích của phương pháp đề xuất là rõ ràng. Đối với các thực nghiệm về phát hiện người đi bộ, Mỗi bộ phân loại có 20 bộ chọn, trong đó mỗi bộ chọn gồm 10 bộ phân loại yếu. Để tăng sự vững chắc của các cập nhật mẫu âm, hệ thống thu thập chồng lấp bốn vùng hình ảnh nền hoạt động trong bốn khoản thời gian khác nhau.

4.1 Thực nghiệm 1: Bộ dữ liệu PETS

Trong thực nghiệm này, nghiên cứu sử dụng bộ dữ liệu PETS (<http://www.cvg.reading.ac.uk/PETS2006/data.html>) công bố công khai số liệu năm 2006 bao gồm 308 khung hình (720×576 pixel), trong đó có 1.714 người đi bộ. Nghiên cứu này so sánh cách tiếp cận với các phương pháp tiên tiến khác, cụ thể là mô hình đối tượng biến dạng của Felzenszwalb và cộng sự năm 2008 [7] và Biểu đồ của hướng tiếp cận Gradients của Dalal và Triggs năm 2005 [5]. Cả hai phương pháp sử dụng cố định bộ phân loại đã được huấn luyện ngoại tuyến và được dựa trên kỹ thuật cửa sổ trượt. Ngoài ra, bài báo so sánh phương pháp tiếp cận cho lưới phân loại phương pháp

của Roth và các cộng sự [6]. Phương pháp tiếp cận lưới phân loại sử dụng thông tin nền để tạo ra lưới, loại bỏ tất cả đối tượng dương phát hiện sai cho các cửa sổ trượt dựa các bộ phát hiện đó nhỏ hơn 75% hoặc lớn hơn 125% kích thước groundtruth để đảm bảo một so sánh công bằng.



Hình 8. RPC: Recall-precision cho PETS2006 chuỗi các bộ phát hiện tiên tiến khác nhau so với các tiếp cận phương pháp đề xuất.

Các kết quả được thể hiện trong Hình 8, có thể thấy rằng cách tiếp cận phương pháp đề xuất bộ phân loại ổn định hơn các bộ phát hiện đối tượng trên cùng bộ dữ liệu cũng như cách tiếp cận phân loại lưới ban đầu, có thể được coi là một cơ sở cho phương pháp đề xuất. Ngoài ra, trong Bảng 1, cung cấp thông tin recall, độ chính xác cho các giá trị F-Measure tốt nhất. Kết quả minh họa phát hiện đối tượng được thể hiện trong Hình 9.

Bảng 1. So sánh Recall và Precision

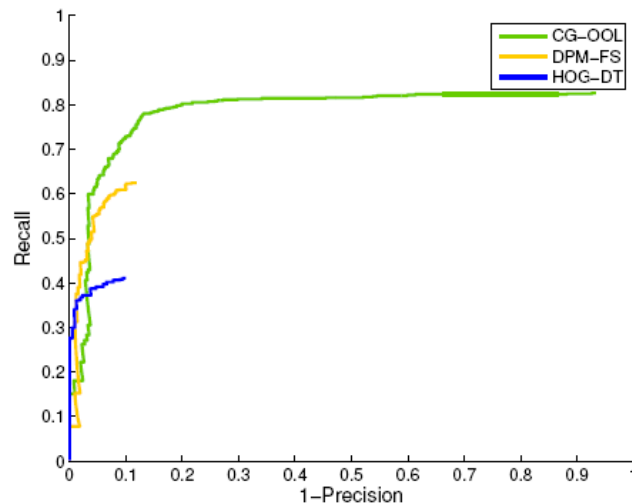
Phương pháp	So Sánh		
	<i>Recall</i>	<i>Precision</i>	<i>F-Measure</i>
Felzenszwalb và cộng sự (FS) [7]	0.73	0.88	0.79
Dalal và Triggs (DT) [5]	0.50	0.88	0.64
Roth và các cộng sự (CG) [6]	0.78	0.79	0.78
Phương pháp đề xuất (Proposed)	0.86	0.96	0.90



Hình 9. Minh họa kết quả phát hiện của phương pháp tiếp cận trên bộ dữ liệu PETS.

4.2 Thực nghiệm 2: Bộ dữ liệu Caviar

Bộ dữ liệu Caviar hiển thị một hành lang trong một trung tâm mua sắm từ hai góc khác nhau. Góc đầu tiên bên hông hành lang, góc thứ hai nhìn trực diện bản. (<https://groups.inf.ed.ac.uk/vision/CAVIAR/CAVIARDATA1/>) Vì nghiên cứu bài báo quan tâm đến quá trình phát hiện người với tỷ lệ thay đổi nên tập trung vào bộ dữ liệu đầu tiên. Dữ liệu có dạng MPEG hoặc JPEG và có độ phân giải là 384×288 . Đối với thực nghiệm này lựa chọn một trong tập dữ liệu khá phức tạp để đánh giá là ShopAssistant2cor vì nó có chứa một số lượng lớn người đi bộ (1265). Có 370 khung hình với kích thước hình ảnh 384×128 . Để tiến hành thực nghiệm với phương pháp tiếp cận dựa trên lưới các bộ phân loại trên bộ dữ liệu Caviar, các tham số sau được khởi tạo: Kích thước vùng ảnh: 32×64 . Số các bộ chọn dùng để huấn luyện trực tuyến cho một bộ phân loại là: 10. Số các bộ phân loại yếu của một bộ chọn là 20.



Hình 10. RPC: Recall-precision cho Caviar Dataset.

Kết quả của chuỗi dữ liệu Caviar được thể hiện trong Hình 10 và Bảng 2. Một lần nữa nó có thể thấy rằng các lưới phát hiện thích nghi tốt hơn bộ phát hiện đối tượng chung [5, 7], đặc biệt là Recall. Kết quả minh họa phát hiện đối tượng được thể hiện trong Hình 11.

Bảng 2. So sánh Recall và Precision

Phương pháp	So Sánh		
	Recall	Precision	F-Measure
Felzenszwalb và cộng sự (DPM-FS) [7]	0.62	0.90	0.74
Dalal và Triggs (HOG-DT) [5]	0.41	0.91	0.57
Roth và các cộng sự (CG-OOL) [6]	0.78	0.87	0.82
Phương pháp đề xuất	0.92	0.93	0.92



Hình 11. Minh họa kết quả phát hiện đối tượng người đi bộ trên bộ dữ liệu Caviar.

5. KẾT LUẬN

Trong bài báo này, phương pháp lưới phân loại cho bài toán phát hiện đối tượng trong ảnh từ camera được đề xuất; Ở đó với các bộ phân loại trên lưới được huấn luyện kết hợp cả học ngoại tuyến và học trực tuyến. Nghiên cứu kết hợp với một bộ phân loại đã được huấn luyện trước một cách cẩn thận dùng để xác minh kiểm tra các mẫu trước khi cập nhật. Tiếp tục giữ các mẫu dương đại diện cố định và tạo ra một tập mẫu âm ước tính từ mô hình nền. Thực nghiệm mô hình lưới phân loại được tiến hành trên hai bộ dữ liệu PETS2006 và CAVIAR. Các kết quả thực nghiệm, được đánh giá và so sánh với các phương pháp khác trên từng bộ dữ liệu đều cho thấy phương pháp đề xuất cho bài toán phát hiện đối tượng có độ chính xác cao, thời gian đảm bảo có thể thực hiện trực tuyến, thích nghi với nhiều môi trường và vấn đề thất lạc đối tượng trong hệ thống phát hiện đối tượng được giải quyết. Hướng tiếp cận tiếp theo

của bài báo này là phát triển thành mô hình lưới phân loại với nhiều đối tượng được phát hiện khác nhau nhằm đáp ứng các bài toán trong thực tế.

TÀI LIỆU THAM KHẢO

- [1] Agarwal S., Awan A., Roth D. (2004). Learning to detect objects in images via a sparse, part-based representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 26 (11), 1475–1490.
- [2] Freund Y. and Schapire R. (1999). A short introduction to boosting. *Journal of Japanese Society for Artificial Intelligence*, 771–780.
- [3] Grabner H. and Bischof H. (2006). On-line boosting and vision. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 260–267.
- [4] Blum A., Mitchell T. (1998). Combining labeled and unlabeled data with co-training”. In: *Proc. Conf. on Computational Learning Theory*, 92–100.
- [5] Dalal N., Triggs B. (2005). Histograms of oriented gradients for human detection. In: *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 886–893.
- [6] Roth P. M., Sternig S., Grabner H., Bischof H. (2009). Classifier grids for robust adaptive object detection. In: *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2727–2734.
- [7] Felzenszwalb P., McAllester D., Ramanan D. (2008). A discriminatively trained, multiscale, deformable part model. In: *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 1–8.
- [8] Goldberg A. B., Li M., Zhu X. Online manifold regularization: A new learning setting and empirical study. In: *Proc. European Conf. on Machine Learning and Knowledge Discovery in Databases*, 393–407.
- [9] Grabner H., Roth P. M., Bischof H. (2007). Is pedestrian detection really a hard task In: *Proc. Tenth IEEE International Workshop on PETS*, 1–9.
- [10] Stalder S., Grabner H., Gool L. (2009). Exploring context to learn scene specific object detectors”. In: *Proc. IEEE Int’l Workshop on Performance Evaluation of Tracking and Surveillance*, 63–70.
- [11] Javed O., Ali S., Shah M. (2005). Online detection and classification of moving objects using progressively improving detectors. In: *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 696–701.
- [12] Leibe B., Leonardis A., Schiele B. (2008), Robust object detection with interleaved categorization and segmentation. *International Journal of Computer Vision*, 77 (1–3), 259–289.
- [13] Levin A., Viola P., Freund Y. (2003) (2003). Unsupervised improvement of visual detectors using co-training”, In: *Proc. ICCV*, 626–633.
- [14] Li L. J., Wang G., Fei-Fei L. (2007). Optimol: automatic online picture collection via incremental model learning. In: *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 1–8.

GRID CLASSIFIER MODEL FOR OBJECT DETECTION FROM CAMERA IMAGES

Nguyen Dang Binh

Faculty of Information Technology, University of Sciences, Hue University

Email: ndbinh@husc.edu.vn

ABSTRACT

The grid classifier approach has proven to be a viable option for detecting objects in camera images. Each specific region on the image is classified using a single classifier. The goal of this work is to improve the object detection system's performance while maintaining the stability, accuracy, and online speed of object identification via a camera. This paper's contribution is a grid classifier model based on the idea of combining the off-line and on-line classifiers in a grid in an adaptive method to improve object detection stability when the environment and object change. Experimental results show high accuracy classification results in the presence of non-moving objects, objects of different sizes and postures in complex environments.

Keywords: object detection, on-line learning, grid classifiers.



Nguyễn Đăng Bình Sinh ngày 08/11/1974 tại Thừa Thiên Huế. Năm 1996, ông tốt nghiệp Đại học ngành Toán - Tin tại Trường Đại học Sư phạm, Đại học Huế. Ông nhận bằng thạc sĩ Công nghệ thông tin tại Trường Đại học Bách Khoa Hà Nội năm 2022; nhận học vị Tiến sĩ ngành Công nghệ thông tin tại Viện Công nghệ Kyushu, Nhật Bản, và hoàn thành nghiên cứu Sau tiến sĩ tại Viện Thị giác và Đồ họa máy tính năm 2008 tại Đại học Công nghệ Graz, Cộng hòa Áo. Hiện ông công tác tại khoa Công nghệ Thông tin, Trường Đại học Khoa học, Đại học Huế.

Lĩnh vực nghiên cứu: Học máy, Thị giác máy tính, Nhận dạng và Xử lý ảnh số.

