

# PHÁT HIỆN KHUÔN MẶT NGƯỜI TRONG ẢNH BẰNG KỸ THUẬT PHÂN LỚP NHỊ PHÂN

Huỳnh Cao Tuấn\*, Đỗ Sĩ Trường\*, Nguyễn Thanh Bình\*, Lâm Thành Hiển\*

\*Trường Đại học Lạc Hồng

\*Học viện Công nghệ Bưu chính Viễn thông

**Tóm tắt:** Phát hiện khuôn mặt người trong ảnh là một kỹ thuật ước lượng các thông số cho các mô hình dự đoán nhằm tìm ra vị trí khuôn mặt. Đây là một khâu quan trọng trong một hệ thống phân tích biểu cảm khuôn mặt hoặc các hệ thống giám sát an ninh.... Bài báo này đề xuất phương pháp phát hiện khuôn mặt người trong ảnh theo hướng tiếp cận phân lớp nhị phân. Kết quả thực nghiệm cho thấy phương pháp được đề xuất cho độ chính xác phát hiện khuôn mặt khá tốt và có tốc độ xử lý nhanh hơn, đây là cơ sở để có thể phát triển những ứng dụng xử lý khuôn mặt yêu cầu thời gian thực trên dữ liệu video.

**Từ khóa:** Phát hiện khuôn mặt, ước lượng biểu cảm, cử chỉ khuôn mặt

## I. GIỚI THIỆU

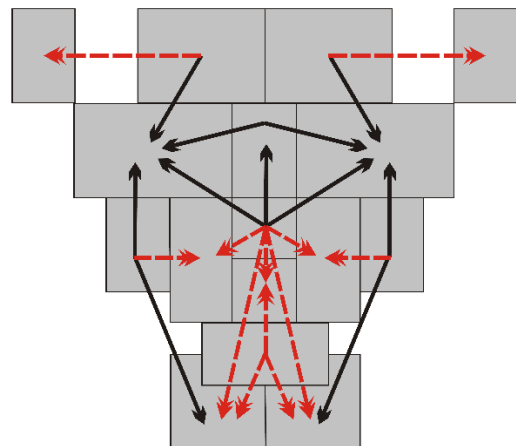
Việc phát hiện khuôn mặt người trong ảnh đã được nghiên cứu nhiều và có nhiều phương pháp khác nhau [1] [2] [3] [4]; chẳng hạn như phương pháp phát hiện và định vị khuôn mặt từ ảnh có nền phức tạp nhằm tìm ra các cạnh, sau đó loại bỏ bớt và nhóm các cạnh lại sao cho chỉ còn lại một biên bao quanh khuôn mặt phân biệt vùng đầu và nền [5]. Độ chính xác của giải thuật này đạt đến 80% với 48 ảnh có nền phức tạp.

Nghiên cứu của Cootes [6] dùng các đốm và vạch sọc để phát hiện khuôn mặt. Tiếp cận này dùng thuật toán Laplace để biến đổi hình ảnh và xác định các đốm, sau đó tìm các hình tam giác ứng với các thành phần của khuôn mặt. Họ dùng các vạch sọc để ghi nhận nét mặt, đồng thời dùng hai hình tam giác để mô tả quan hệ giữa các đốm, 2 đốm tối và 3 đốm sáng để tả 2 mắt, 2 gò má và mũi. Khuôn mặt được phát hiện nếu các vạch sọc bao quanh các thành phần.

Hay các nghiên cứu của Guggisberg [7] đã dùng phương pháp định vị các đặc trưng khuôn mặt cho ảnh xám. Ý tưởng là: sử dụng các bộ lọc để làm nổi các biên, tiếp theo là làm nổi bật các vùng có hình dạng rõ ràng. Lúc này biểu đồ mức xám (Histogram) của ảnh có một đỉnh nổi bật, dựa vào đỉnh này, xác định các giá trị ngưỡng phù hợp để tạo ra 2 ảnh nhị phân tương ứng. Vùng ứng viên của khuôn mặt là vùng có các phân liên hệ giữa hai ảnh nhị phân này. Nhược điểm của phương pháp này là dùng các phép toán hình thái học nên khó có thể kết hợp với các đặc trưng khác để định vị một khuôn mặt.

Vào năm 2013, Li và Chung [8] đã đề xuất một cách tiếp cận để định vị khuôn mặt trong ảnh có nền phức tạp bằng cách sử dụng 5 đặc trưng là 2 lỗ mũi, 2 mắt, đoạn giữa mũi và môi để thể hiện một khuôn mặt chuẩn. Sử dụng mô hình Gauss để mô hình tập các khoảng cách, và với mỗi loại đặc trưng thì ta tính khoảng cách giữa chúng. Vì các đặc trưng không thể xuất hiện ngẫu nhiên nên ta dùng mô hình xác suất để định vị chúng qua khoảng cách.

Còn Su và Zheng [9] đã dựa vào sự thay đổi mức độ sáng của các vùng khác nhau của khuôn mặt (như hai má, hai mắt và trán). Sau đó, xác định các cặp tỷ số của một số vùng cho ta một lượng bất biến khá rõ ràng. Sự thay đổi độ sáng của các vùng trên khuôn mặt tạo ra các cặp quan hệ sáng-tối giữa các vùng nhỏ đều được ghi nhận lại. Một khuôn mặt được định vị nếu nó thỏa mãn tất cả các cặp sáng-tối và mối quan hệ tương quan giữa các cặp đó với nhau.



Hình 1.1 Mẫu khuôn mặt trong phương pháp định vị

Tác giả liên hệ: Huỳnh Cao Tuấn,

Email: caotuan@lhu.edu.vn

Đến tòa soạn: 10/2020, chỉnh sửa: 11/2020, chấp nhận đăng: 12/2020.

Hình 1.1 là một mẫu khuôn mặt có kích thước theo tỷ lệ 14x16 pixel được sử dụng trong phương pháp định vị khuôn mặt của Schneiderman và Kanade [10], gồm 16 vùng trên khuôn mặt và 23 quan hệ được thể hiện thông qua các mũi tên. Có 11 quan hệ thiết yếu (mũi tên đen) và 12 quan hệ xác thực (mũi tên đỏ). Các tác giả Wu và Toan cùng cộng sự [2] [11] sử dụng đặc trưng Haar trong thuật toán AdaBoost để xác định một thành phần nằm trong một vùng chứa nó.

## II. ĐỀ XUẤT KỸ THUẬT PHÁT HIỆN MẶT NGƯỜI

### 1. Phương pháp tiếp cận

Việc phát hiện khuôn mặt được thực hiện dựa trên ý tưởng cơ bản là thực hiện phân lớp nhị phân với mỗi vùng ảnh quan tâm, từ đó kết luận vùng ảnh đó là khuôn mặt hay không. Việc xét duyệt được thực hiện bởi một chuỗi các bộ phân lớp nhị phân và một vùng ảnh được chấp nhận nếu nó được chấp nhận bởi toàn bộ các bộ phân lớp nhị phân trong chuỗi. Các bộ phân lớp nhị phân được xây dựng trên cơ sở sử dụng cây quyết định, trong đó việc đánh giá tại mỗi nút là một bộ phân lớp nhị phân con. Các bộ phân lớp nhị phân con này được xây dựng dưới nhiều dạng khác nhau, ý tưởng cho điều này là để có thể lựa chọn những cách thức phân lớp đơn giản và nhanh chóng tại một số cây đầu tiên, giúp nhanh chóng loại những vùng nền, đồng thời thiết kế như vậy cho phép việc nghiên cứu thử nghiệm với những cách thức phân lớp khác nhau tại mỗi nút.

Việc phát hiện mặt người được tiến hành thông qua hai bước chính.

- Bước thứ nhất là học; tức là từ một tập các ảnh khuôn mặt và các ảnh không phải khuôn mặt để tạo ra một bộ dữ liệu học mô hình khuôn mặt mẫu.
- Bước thứ hai là phát hiện; việc phát hiện khuôn mặt là việc dò tìm xem một vùng bất kỳ trong ảnh có thỏa mãn các đặc điểm của mô hình đã được xây dựng hay không.

### 2. Kỹ thuật xác định mô hình gương mặt mẫu

Mô hình được xây dựng dựa trên tập dữ liệu huấn luyện như sau:

$$\{(I_s, v_s, w_s) : s = 1, 2, \dots, S\} \quad (2.1)$$

Trong đó,  $v_s$  là nhãn đúng của ảnh  $I_s$ , và  $w_s$  là trọng số tương ứng.

Cụ thể, trong trường hợp này, vấn đề bài toán đưa ra là thực hiện phân lớp nhị phân, các nhãn có giá trị tương ứng là +1 và -1. Giá trị trọng số  $w_s$  cho phép đánh dấu mức độ quan trọng khác nhau của mỗi mẫu đầu vào trong tập huấn luyện. Quá trình xây dựng cây được thực hiện tại từng nút trên cơ sở lựa chọn bộ phân lớp nhị phân con có khả năng phân lớp tốt nhất bộ dữ liệu huấn luyện, tức là đạt giá trị cực tiểu cho hàm mục tiêu; cụ thể là thực hiện tìm sai số bình phương nhỏ nhất ứng với việc phân chia tập huấn luyện được xác định tại nút đó. Hàm mục tiêu có dạng:

$$WMSE(I, v, w) = \sum_{(I, v, w) \in C_0} w \cdot (v - \bar{v}_0)^2 + \sum_{(I, v, w) \in C_1} w \cdot (v - \bar{v}_1)^2 \quad (2.2)$$

trong đó:  $C_0$  và  $C_1$  là hai cụm của tập huấn luyện tương ứng là kết quả phân chia của hai giá trị 0 và 1. Các tham số  $\bar{v}_0$  và  $\bar{v}_1$  là trung bình các giá trị nhãn trong  $C_0$  và  $C_1$ .

Nói cách khác, tại mỗi nút trong cây, ta xét duyệt các bộ phân lớp nhị phân con có thể, tùy từng dạng phân lớp mà ta học hoặc lựa chọn tham số, hoặc thậm chí là không có tham số. Như vậy, từ tập dữ liệu ban đầu, tại mỗi bước học nút khi xây dựng cây, tập dữ liệu huấn luyện tương ứng được chia đôi. Cơ sở của thuật toán đề xuất nằm ở việc học từng nút. Thuật toán học cây có thể mô tả như sau:

---

#### Algorithm 2.1: Thuật toán huấn luyện mẫu

---

**Input:** Tập dữ liệu huấn luyện

$$U = \{(I_s, v_s, w_s) : s = 1, 2, \dots, S\}$$

**Output:** Tập các nút

$$T = \{N_0, N_1, \dots\}$$

**Process:**

1.  $T := \emptyset$
2.  $Idx0 = \{0, 1, 2, \dots, S-1\}$  // các trọng số tương ứng với các mẫu trong dữ liệu huấn luyện.
3.  $Stack := \emptyset$
4. **push** ( $Stack, (N_0, Idx_0)$ )
5. **while** ( $Stack \neq \emptyset$ )
6.  $\{N_i, Idx_i\} := \mathbf{pop}(Stack)$  // Nếu nút đã đạt đủ độ cao, không thực hiện tính toán chia đôi nữa
7.     **if** ( $N_i.level \geq MAX\_DEPTH$ )
8.         **Continue**
9.     **else**
10.          $min\_err := MAX\_VAL$
11.          $best\_bincls := \mathbf{null}$
12.         **for all** bincls of BCS
13.              $e := WMSE(bincls, U, Idx_i)$
14.             **if** ( $e < minerr$ )
15.                  $best\_bincls := bincls$
16.              $min\_err := e$
17.             **endif**
18.             **endfor**
19.         **setupNode** ( $N_i, best\_bincls, U, Idx_i$ )
20.          $\{Idx_{i*2+1}, Idx_{i*2+2}\} := \mathbf{SplitDataSet}(U, Idx_i, best\_bincls)$
21.         **push** ( $Stack, (N_{i*2+1}, Idx_{i*2+1})$ )

```

22.         push( Stack, (Ni+2+2, Idxi+2+2) )
23.     endif
24. endwhile
    
```

Trong đó: Giá trị MAX\_DEPTH là tham số giới hạn chiều sâu của cây sẽ học, đây là tham số đặt bằng tay, hiện tại đặt mặc định là 5; MAX\_VAL là giá trị khởi tạo cho sai số, dùng để tìm bộ phân lớp con có giá trị sai số nhỏ nhất.

Theo đó, với mỗi kết quả phân lớp trên tập học của nút, ta nhận được giá trị hàm mục tiêu, và kết quả lựa chọn cho nút là bộ tương ứng đạt giá trị nhỏ nhất cho hàm mục tiêu. Như vậy, ta có thể phân tích, thời gian học bằng tổng thời gian học từng mức của cây. Thời gian học mỗi mức của cây là tổng thời gian học từng nút và tổng số lượng dữ liệu huấn luyện của các nút tại mỗi mức bằng tập dữ liệu học ban đầu. Thời gian học một nút bằng tổng thời gian học từng bộ phân lớp con có thể. Trong quá trình tối ưu hóa cho một nút trong cây, vì không gian các bộ phân lớp có thể phải xét là khá lớn, bởi vậy, một tập con được sinh ngẫu nhiên đã được sử dụng thay thế. Do đó, các cây được xây dựng kết hợp dựa trên tiếp cận boosting với lựa chọn là GentleBoost.

Trên cơ sở cấu trúc phân lớp được xây dựng như vậy, việc còn lại là thiết kế các bộ phân lớp nhị phân con.

**a) Phép so sánh cường độ điểm ảnh**

Một phép so sánh cường độ điểm ảnh trên ảnh I được định nghĩa như sau:

$$B(I: l_1, l_2) = \begin{cases} 0, & I(l_1) \leq I(l_2) \\ 1, & otherwise \end{cases} \quad (2.3)$$

trong đó:  $I(l_i)$  là giá trị cường độ điểm ảnh trong ảnh I tại vị trí  $l_i$ . Trong kỹ thuật này, tọa độ  $l_1$  và  $l_2$  được xác định trong phạm vi  $[-1, +1] \times [-1, +1]$ . Các giá trị tọa độ  $l_1$  và  $l_2$  được xác định như vậy cho phép việc triển khai xác định tập các vị trí không phụ thuộc vào kích thước ảnh mẫu. Do đó, với mỗi vùng ảnh cần xác định, giá trị tọa độ cần được thay đổi theo tỉ lệ phù hợp.

44	43	67	71
36	41	76	71
37	30	85	95
81	77	94	94

Hình 2.2 Ví dụ của phép so sánh cường độ

Để hình dung phép so sánh cường độ ảnh, ta xét ví dụ như trong Hình 2.2; theo đó, ảnh được xét là ảnh xám,

tức là các giá trị nằm trong đoạn  $[0,255]$ . Cụ thể, trong trường hợp này, ta xét vị trí trên tọa độ ảnh. Một số kết quả so sánh cường độ được chỉ ra trong Bảng 2.1. So sánh cường độ điểm ảnh là một trong những cách phân lớp đơn giản nhất và không cần đến tham số. Việc tính toán nó còn đơn giản hơn đặc trưng Haar, chưa kể đến nó không yêu cầu đến dữ liệu được tính trước như trường hợp ảnh tích phân cho đặc trưng Haar. Về mặt mở rộng, nó có khả năng tùy biến cho các góc nghiêng khác nhau của đối tượng một cách dễ dàng chỉ với một phép biến đổi tọa độ trong không gian hai chiều đối với 2 vị trí cần so sánh.

Để xét duyệt tại mỗi nút, một tập con các phép so sánh được sử dụng. Theo đó, hai vị trí trong phép so sánh được sinh ngẫu nhiên theo phân phối đều trong phạm vi  $[-1, +1] \times [-1, +1]$ . Nếu cây cần xây dựng có chiều sâu D, ta phải thử B phép so sánh tại mỗi nút và tập huấn luyện chứa S mẫu, khi đó thời gian huấn luyện cây là  $O(D \cdot B \cdot S)$ .

Bảng 2.1 Một số kết quả ví dụ so sánh cường độ ảnh

$l_1$	$l_2$	$I(l_1)$	$I(l_2)$	$B(I: l_1, l_2)$
(0, 0)	(1, 0)	44	43	1
(0, 0)	(2, 0)	44	67	0
(0, 0)	(0, 1)	44	36	1
(0, 0)	(1, 1)	44	41	1
(0, 0)	(2, 1)	44	76	0
(2, 0)	(3, 0)	67	71	0
(2, 1)	(3, 1)	76	71	1
(1, 2)	(2, 3)	30	94	0
(0, 3)	(3, 1)	81	71	1
(1, 1)	(3, 3)	41	94	0

**b) Giá trị trung bình cục bộ**

$$B(I: R(x, y, w, h), \delta) = \begin{cases} 1, & \frac{1}{w * h} \sum_{x \leq i < x+w} \sum_{y \leq j < y+h} I(i, j) < \delta \\ 0, & otherwise \end{cases} \quad (2.4)$$

trong đó:  $I(i, j)$  là giá trị cường độ điểm ảnh trong ảnh I tại vị trí  $(i, j)$ . Kết quả phân lớp dựa trên việc so sánh giá trị trung bình các điểm ảnh trong vùng hình chữ nhật  $R(x, y, w, h)$  với giá trị ngưỡng  $\delta$ .

Việc tính giá trị trung bình cục bộ vốn cần xét duyệt các giá trị trong vùng, tuy nhiên bằng cách kế thừa sử dụng kỹ thuật ảnh tích phân, việc tính toán có thời gian là  $O(1)$ . Tiếp đến, việc sử dụng giá trị trung bình cục bộ cần đến một tham số ngưỡng  $\delta$ . Tham số này được lựa chọn trong quá trình học. Về cơ bản, các vị trí ngưỡng

được xét duyệt tương ứng với vị trí sát cạnh của từng giá trị tính được từ các mẫu trên trục số. Tương tự như trên, nếu cây cần xây dựng có chiều sâu  $D$ , ta phải thử  $B$  phép so sánh tại mỗi nút và tập huấn luyện chứa  $S$  mẫu, khi đó thời gian huấn luyện cây là  $O(D \cdot B \cdot S^2)$ .

### 3. Phát hiện khuôn mặt trong ảnh

Không xét đến các thông tin có thể biết trước nhằm giảm thiểu không gian tìm kiếm, chẳng hạn như vùng ảnh nền, ảnh chuyển động trong ngữ cảnh dùng video, ảnh đầu vào được thực hiện xét duyệt tại mỗi vị trí và kích thước có thể. Như đề xuất đã nêu ở phần trên, một khuôn mặt được chấp nhận nếu nó khớp với mô hình đã xây dựng, trong đó tại mỗi nút là một bộ phân lớp nhị phân con và phương pháp lựa chọn là phép so sánh giá trị trung bình cục bộ của điểm ảnh.

Xét việc tính toán trên từng cây và giả sử rằng cây đang xét có chiều sâu  $D$ . Khi đó, thời gian phân lớp với cây là  $O(D \cdot K)$  với  $K$  là thời gian thực hiện tính toán tại bộ phân lớp nhị phân con. Trong các trường hợp bộ phân lớp nhị phân con được sử dụng là so sánh điểm ảnh và đánh giá giá trị trung bình cục bộ, thời gian tính toán đều là  $O(I)$ ; do đó, thời gian phân lớp với cây là  $O(D)$ .

### III. ĐÁNH GIÁ THỬ NGHIỆM

Việc thử nghiệm phương pháp được đề xuất trong bài báo này được tiến hành theo 2 bước: học và chạy thử nghiệm phát hiện. Pha học được tiến hành với cơ sở dữ liệu khuôn mặt GENKI-SZSL nằm trong bộ MPLab GENKI của Đại học California, San Diego [12] bao gồm 3.500 ảnh khuôn mặt. Việc huấn luyện cũng sử dụng Bộ dữ liệu Face negative gồm 3.019 ảnh được cung cấp trong một hướng dẫn sử dụng bộ Haartraining của OpenCV. Thời gian huấn luyện cây là khoảng 30 phút trên máy tính core i7 có RAM 8GB.

Việc thử nghiệm được tiến hành với cơ sở dữ liệu khuôn mặt của Viện Công nghệ California, bao gồm 450 ảnh khuôn mặt, có độ phân giải 896x592 và định dạng jpeg. Các khuôn mặt trong ảnh được thu thập dưới nhiều điều kiện khác nhau, về ánh sáng, biểu cảm mặt và nền. Phương pháp cũng được so sánh với kết quả phát hiện sử dụng thuật toán Haar Adaboost được cung cấp bởi thư viện OpenCV, cụ thể là sử dụng mẫu haarcascade frontalface alt\_tree với các tham số mặc định. Tác giả sử dụng mô hình đã huấn luyện để tìm kiếm và định vị khuôn mặt. Sau đó, tác giả sử dụng lớp Haar Feature-based Cascade Classifiers từ thư viện OpenCV để rút trích các đặc trưng Haar của tất cả các ảnh khuôn mặt (không phải ảnh gốc) và lưu vào cơ sở dữ liệu Haar. Sau 5 lần kiểm thử thì tính trung bình cộng để xác định tổng giá trị lỗi.

Về mặt lý thuyết, phương pháp đề xuất và haar adaboost đều dùng tiếp cận boosting. Tuy nhiên, việc xây dựng các bộ phân lớp yếu thì có những tiếp cận khác nhau. Trong haar adaboost, các bộ phân lớp yếu được xây dựng từ các đặc trưng haar; trong phương pháp đề xuất các bộ phân lớp yếu được tổ chức bằng mô hình cây quyết định và có khả năng tùy biến việc thiết kế các bộ phân lớp nhị phân con. Một bộ phân lớp nhị phân con được sử dụng là phép so sánh cường độ điểm ảnh cũng đã đưa ra phân tích so sánh với đặc trưng Haar. Cụ thể: so sánh cường độ điểm ảnh là một trong những cách phân lớp đơn giản nhất và không cần đến tham số. Việc

tính toán nó còn đơn giản hơn đặc trưng Haar, chưa kể đến nó không yêu cầu đến dữ liệu được tính trước như trường hợp ảnh tích phân cho đặc trưng Haar. Về mặt mở rộng, nó có khả năng tùy biến cho các góc nghiêng khác nhau của đối tượng một cách dễ dàng chỉ với một phép biến đổi tọa độ trong không gian hai chiều đối với 2 vị trí cần so sánh. Kết quả thử nghiệm được tóm tắt như trong Bảng 3.1.

Bảng 3.1 So sánh kết quả định vị khuôn mặt

	<i>Haar Adaboost</i>	<i>PP đề xuất</i>
Số mẫu phát hiện được	427/450	416/450
Thời gian xử lý trung bình ảnh (giây/ảnh)	0,109713	0,018982

Trong 450 ảnh, thuật toán đề xuất trong bài báo phát hiện đúng 416 ảnh - tương đương 92,44%. Thấp hơn so với 427 ảnh của phương pháp Haar Adaboost đạt 94,89%. Tuy nhiên thời gian xử lý thì nhanh hơn xấp xỉ 5,7 lần.

### IV. KẾT LUẬN

Phát hiện được khuôn mặt người chính là bước đầu tiên quan trọng của một hệ thống rút trích cử chỉ và biểu diễn biểu cảm. Trong bài báo này, tác giả đã trình bày một số phương pháp tiếp cận và đề xuất một kỹ thuật phát hiện khuôn mặt người trong ảnh dựa trên ý tưởng cơ bản là thực hiện phân lớp nhị phân với mỗi vùng ảnh quan tâm, từ đó kết luận vùng ảnh đó có khuôn mặt hay không.

Như vậy, kết quả 92,44% cho thấy phương pháp được đề xuất cho độ chính xác phát hiện khuôn mặt khá tốt. Mặc dù thuật toán Haar Adaboost đạt giá trị 94,89% tốt hơn so với 92,44% của phương pháp đề xuất trong bài báo nhưng bù lại thì nó có tốc độ xử lý nhanh hơn, đây là cơ sở để có thể phát triển những ứng dụng xử lý khuôn mặt yêu cầu thời gian thực trên dữ liệu video.

### TÀI LIỆU THAM KHẢO

- [1] Han C.C., Liao H.Y.M., Yu G.J., Chen L.H., "Fast face detection via morphology-based pre-processing," *Pattern Recognition*, vol. 33, pp. 1701-1712, 2000.
- [2] Wu Y., Ai X., "Face detection in color images using Adaboost algorithm based on skin color information," *Proceedings of the First International Workshop on Knowledge Discovery and Data Mining*, pp. 339-342, 2008.
- [3] Hien L.T., Toan D.N., Toan H.M., "Detecting Human Face with Ridge-Valley-Normal Model," *International Journal of Computer Science and Software Engineering (IJCSSE)*, vol. 4, pp. 107-113, 2015.

- [4] Hien L.T., Toan D.N., "An algorithm to detect driver's drowsiness based on nodding behavior," *International Journal of Soft Computing, Mathematics and Control*, vol. 5(1), pp. 1-8, 2016.
- [5] Hien L.T., Toan D.N., Lang T.V., "Detection of human head direction based on facial normal algorithm," *International Journal of Electronics Communication and Computer Engineering*, vol. 6(1), pp. 110-114, 2015.
- [6] Cootes, T.F.; Edwards, G.J.; Taylor, C.J., "Active appearance models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23(6), p. 681-685, 2001.
- [7] Guggisberg A.G., Mathis J., Schnider A., Hess C.W., "Why do we yawn?," *Neuroscience & Biobehavioral Reviews*, 34, pp. 1267-1276, 2010.
- [8] Li G., Chung W.Y., "Detection of driver drowsiness using wavelet analysis of heart rate variability and a support vector machine classifier," *Sensors*, vol. 13(12), pp. 16494-16511, 2013.
- [9] Su H., Zheng G., "A partial least squares regression-based fusion model for predicting the trend in drowsiness," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 38(5), pp. 1085-1092, 2008.
- [10] Schneiderman H., Kanade T., "Probabilistic modeling of local appearance and spatial relationships for object detection," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 40-50, 1998.
- [11] Toan H.M., Toan D.N., Hien L.T., Lang T.V., "Modeling the human face and its application for detection of driver drowsiness," *International Journal of Computer Science and Telecommunications*, vol. 3(11), pp. 56-59, 2012.
- [12] MPLab, "The MPLAB GENKI database, GENKI-4K subset," *Machine Perception Laboratory*, 2009.

**Keywords:** Face detection, expression estimation, face gestures



**Th.S Huỳnh Cao Tuấn**, Nhận học vị Thạc sỹ năm 2012. Hiện công tác tại Trường Đại học Lạc Hồng. Lĩnh vực nghiên cứu: Công nghệ robot, khai phá dữ liệu, xử lý ảnh, học máy.  
Email: caotuan@lhu.edu.vn



**ThS. Đỗ Sĩ Trường**, Nhận học vị thạc sỹ năm 2012. Hiện công tác tại trường Đại học Lạc hồng. Lĩnh vực nghiên cứu: Khai phá dữ liệu, xử lý ảnh, ngôn ngữ tự nhiên  
Email: truongds@lhu.edu.vn



**TS. Nguyễn Thanh Bình**, Nhận học vị Tiến sỹ năm 1995. Hiện công tác tại Học viện Bưu chính Viễn Thông. Lĩnh vực nghiên cứu: Kỹ thuật phát thanh và truyền hình, xử lý âm thanh, xử lý hình ảnh.  
Email: thanhbinh68@gmail.com



**TS. Lâm Thành Hiển**, Nhận học vị Tiến sỹ năm 2017. Hiện công tác tại trường Đại học Lạc hồng. Lĩnh vực nghiên cứu: Khai phá dữ liệu, xử lý ảnh  
Email: lthien@lhu.edu.vn

## IN THE FACE DETECTION TECHNOLOGY PHOTOS BY BINARY CLASSIFIERS

**Summary:** Human face detection in an image is a parameter estimation technique for predictive models in order to find out the face position. This is an important stage in a facial expression analysis system or security surveillance systems.... This article proposes a method to detect human faces in images according to the binary classification approach. Experimental results show that the proposed method has quite good face detection accuracy and faster processing speed, which is the basis for developing time-demanding face-processing applications. real on video data.