# NGHIÊN CỨU XÂY DỰNG HỆ THỐNG BẢO MẬT ĐA LỚP SỬ DỤNG THỊ GIÁC MÁY VÀ MẠNG NƠ RON HỌC SÂU

## A MULTI-LAYER SECURITY FRAMEWORK BASED ON COMPUTER VISION AND DEEP LEARNING

**Trần Ngọc Tiến[1,*], Nguyễn Quốc Trường[1], Trương Công Giang[2]**

DOI: https://doi.org/10.57001/huih5804.2025.290

**TÓM TẮT**

Nghiên cứu này để xuất một hệ thống bảo mật hai lớp kết hợp giữa công nghệ thị giác máy và mạng học sâu nhằm nâng cao khả năng phát hiện và ngăn chặn các hành vi xâm nhập bất hợp pháp trong các nhà máy sản xuất thông minh. Hệ thống sử dụng camera tiêu chuẩn để nhận dạng các điểm đặc trưng trên khuôn mặt, đặc biệt tập trung vào vùng mắt và miệng. Đây là những khu vực dễ quan sát sự thay đổi khi người dùng thực hiện các cử chỉ theo yêu cầu của hệ thống. Trên cơ sở đó, nhóm tác giả phát triển một thuật toán lai giữa trí tuệ nhân tạo và thị giác máy nhằm xây dựng cơ chế xác thực kép. Lớp bảo mật thứ nhất thực hiện nhận diện khuôn mặt bằng mô hình trí tuệ nhân tạo, trong khi lớp thứ hai kiểm tra tính sống dựa trên mô hình toán học phân tích chuyển động khuôn mặt. Kết quả thực nghiệm cho thấy hệ thống đạt độ chính xác nhận diện khoảng 95% và hoạt động ổn định trong điều kiện ánh sáng khác nhau, kể cả khi đối tượng đeo kính. Phương pháp đề xuất có tiềm năng ứng dụng cao trong các hệ thống nhà máy sản xuất công nghiệp đòi hỏi mức độ bảo mật nghiêm ngặt và khả năng chống giả mạo khuôn mặt thời gian thực.

***Từ khóa:*** *Nhận diện khuôn mặt, điểm mốc khuôn mặt, thị giác máy tính, trí tuệ nhân tạo.*

**ABSTRACT**

This study proposes a dual-layer security device that integrates computer vision technology and deep learning networks to enhance the detection and prevention of unauthorized intrusions in smart manufacturing environments. The system employs standard cameras to identify facial landmarks, with particular focus on the eye and mouth regions - areas where motion can be easily observed when users perform gestures as requested by the system. Based on this, the authors develop a hybrid algorithm combining artificial intelligence and computer vision to establish a two-step authentication mechanism. The first security layer performs facial recognition using an AI-based model, while the second layer conducts liveness detection through mathematical modeling of facial motion. Experimental results demonstrate that the system achieves approximately 95% recognition accuracy and maintains stability under varying lighting conditions, including scenarios where subjects wear glasses. The proposed method shows strong potential for application in industrial manufacturing systems that require high levels of security and real-time protection against facial spoofing.

***Keywords:*** *Face recognition, facial landmarks, computer vision, artificial intelligence.*

[1]Trường Cơ khí - Ô tô, Trường Đại học Công nghiệp Hà Nội

[2]Trường Cao đẳng Kỹ thuật - Công nghệ Vĩnh Phúc

[*]Email: tientn@haui.edu.vn

Ngày nhận bài: 06/5/2025

Ngày nhận bài sửa sau phản biện: 15/7/2025

Ngày chấp nhận đăng: 28/7/2025

## 1. GIỚI THIỆU

Sự phát triển nhanh chóng của trí tuệ nhân tạo (AI) trong thời gian gần đây đã thúc đẩy quá trình chuyển đổi các hệ thống bảo mật từ phương pháp truyền thống sang các giải pháp thông minh dựa trên học máy và thị giác máy. AI đang ngày càng đóng vai trò trung tâm trong việc tăng cường khả năng phòng thủ trước các mối đe dọa công nghệ với nhiều ứng dụng điển hình như: hệ thống phát

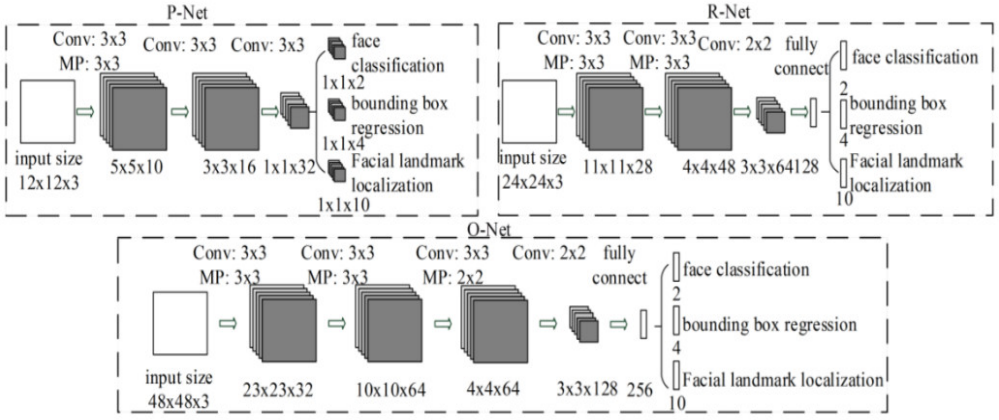
hiện xâm nhập (IDS) [1], nhận diện khuôn mặt và sinh trắc học [2], phân tích mã độc [3], phát hiện lừa đảo [4],… Hossain và cộng sự [5] đã đề xuất một hệ thống IDS sử dụng mạng LSTM để phát hiện các cuộc tấn công vào mạng CAN trong các phương tiện hiện đại, đạt độ chính xác lên đến 99,995%. Tuy nhiên, một hạn chế lớn của các hệ thống nhận diện khuôn mặt hiện nay là dễ bị đánh lừa bởi hình ảnh, video hoặc mặt nạ. Để giải quyết vấn đề này, nhiều nghiên cứu đã phát triển các kỹ thuật phát hiện sự sống, yêu cầu người dùng thực hiện các hành vi cụ thể để xác minh tính hợp lệ. Các mô hình học sâu như FaceNet [6], ArcFace [7] và DeepFace [8] đã đạt được những thành tựu đáng kể trong lĩnh vực nhận diện khuôn mặt, tuy nhiên vẫn tồn tại nhược điểm về tỷ lệ cảnh báo giả cao, đặc biệt khi phải phân biệt giữa hành vi hợp lệ và bất thường. Điều này thúc đẩy nhu cầu xây dựng các hệ thống bảo mật đa lớp, kết hợp giữa phân tích hành vi người dùng và xác thực bằng cử chỉ động.

Công nghệ thị giác máy cũng đóng vai trò quan trọng trong việc phân tích khuôn mặt, từ việc phát hiện điểm mốc, theo dõi biểu cảm đến phân tích trạng thái mắt và miệng. Các nghiên cứu như phát hiện nháy mắt [9], đánh giá liệt mặt [10] hay phát hiện buồn ngủ qua chuyển động của mắt [11] đã cho thấy tiềm năng của các tín hiệu sinh trắc học thời gian thực trong việc xác minh danh tính.

Nhằm khắc phục các hạn chế trên, nhiên cứu này đề xuất một hệ thống bảo mật đa lớp kết hợp giữa AI và thị giác máy. Lớp đầu tiên sử dụng mô hình AI (FaceNet + MTCNN) để nhận diện khuôn mặt, trong khi lớp thứ hai xác minh hành vi bằng cách yêu cầu người dùng thực hiện cử chỉ khuôn mặt ngẫu nhiên (như chớp mắt, mở miệng) và so sánh với mô hình đã học từ dữ liệu Multi-PIE. Hai lớp bảo mật này hoạt động song song theo thời gian thực để đảm bảo tính chính xác và an toàn của hệ thống. Chi tiết về phương pháp và kết quả thử nghiệm sẽ được trình bày trong các phần tiếp theo của bài báo.

## 2. VẬT LIỆU VÀ PHƯƠNG PHÁP

Trong hệ thống đề xuất, bước đầu tiên của quy trình nhận dạng là phát hiện chính xác vị trí khuôn mặt trong ảnh đầu vào. Chúng tôi sử dụng kiến trúc mạng MTCNN nhằm phát hiện khuôn mặt và các điểm mốc chính như mắt, mũi, miệng. MTCNN bao gồm ba mạng con hoạt động nối tiếp. P-Net: thực hiện quét ảnh đầu vào bằng cửa sổ trượt kích thước 12×12×3 để đề xuất các vùng nghi ngờ chứa khuôn mặt. R-Net: sàng lọc lại các vùng được đề xuất từ P-Net, loại bỏ các vùng không chứa khuôn mặt. O-Net: thực hiện xác định chính xác các vị trí mốc trên khuôn mặt, bao gồm các điểm chính (hai mắt, mũi, hai khóe miệng). Quá trình này giúp tạo ra tập hợp các vùng chứa khuôn mặt và tọa độ điểm mốc phục vụ cho bước nhận diện và phân tích tiếp theo. Hình 1 mô tả kiến trúc của MTCNN.




Hình 1. Kiến trúc của P-Net, R-Net và O-Net

Sau khi phát hiện khuôn mặt, hệ thống sử dụng mạng FaceNet để trích xuất đặc trưng khuôn mặt dưới dạng véc tơ 128 chiều. Mỗi véc tơ này biểu diễn duy nhất một khuôn mặt trong không gian nhúng, cho phép đo khoảng cách Euclid giữa các cá thể. Quy trình huấn luyện FaceNet sử dụng chiến lược triplet loss, bao gồm ba loại ảnh: Anchor (ảnh gốc cần nhận dạng), Positive (ảnh khác của cùng một người với anchor), Negative (ảnh của người khác). Hàm mất mát đảm bảo rằng khoảng cách giữa anchor và positive nhỏ hơn khoảng cách giữa anchor và negative một giá trị ngưỡng α nhất định:

$$\forall f(x_i^a), f(x_i^p), f(x_i^n) \in T: \left\| \begin{matrix} f(x_i^a) \\ -f(x_i^p) \end{matrix} \right\|_2^2 + \alpha < \left\| \begin{matrix} f(x_i^a) \\ -f(x_i^n) \end{matrix} \right\|_2^2 \quad (1)$$

Việc này giúp hệ thống tăng cường khả năng phân biệt giữa các khuôn mặt khác nhau, đồng thời giảm độ phụ thuộc vào điều kiện chiếu sáng hay biểu cảm. Hàm mất mát được biểu diễn bởi phương trình (2):

$$L = \sum_i^N \left[ ||f(x_i^a) - f(x_i^p)||_2^2 - ||f(x_i^a) - f(x_i^n)||_2^2 \right] + \alpha \quad (2)$$

Quá trình này được tiếp tục cho đến khi tất cả các hình ảnh của cùng một khuôn mặt người ở gần nhau và các

hình ảnh của các khuôn mặt người khác nhau ở xa nhau. Để đạt được hiệu quả đào tạo cao cho mô hình, chúng ta cần điều chỉnh mạng sao cho Anchor và Negative cách xa nhau nhất có thể, trong khi Anchor và Positive ở gần nhau nhất có thể:

$$\begin{cases} \| f(x_i^a) - f(x_i^p) \| \rightarrow \min \\ \| f(x_i^a) - f(x_i^n) \| \rightarrow \max \end{cases} \quad (3)$$

Để xác định sự sống và tránh các hành vi giả mạo bằng ảnh tĩnh hoặc mặt nạ, hệ thống sử dụng mô hình AAM (Active Appearance Model) và AOM (Active Orientation Model) để mô hình hóa các đặc điểm động của khuôn mặt. Mô hình AAM bao gồm hai thành phần chính là hình dạng và ngoại hình. Hình dạng biểu diễn bằng lưới tam giác với các điểm mốc cố định còn ngoại hình biểu diễn kết cấu bên trong khuôn mặt thông qua các biến đổi tuyến tính. Về mặt toán học, hình dạng được định nghĩa là lưới tam giác có các đỉnh là tọa độ của các điểm đặc trưng trên khuôn mặt được cho bởi:

$$s = (x_1, y_1, x_2, y_2, ..., x_n, y_n)^T \quad (4)$$

Trong đó, $x_i$, $y_i$ (i = 1…n) là tọa độ của các đỉnh, n là số đỉnh trong lưới và các điểm chuẩn tương ứng. Ngoại hình của khuôn mặt được biến đổi bằng cách kết hợp tuyến tính các vectơ biến đổi hình ảnh. Hình dạng s được biểu thị dưới dạng hình dạng cơ sở $s_0$ cộng với kết hợp tuyến tính của k hình dạng riêng $s_i$:

$$s = s_0 + \Phi_s c \quad (5)$$

Trong đó $\Phi_s = (s_1, s_2, ..., s_k)$ và các hệ số $c = (c_1, c_2, ..., c_k)$ lần lượt là các hình dạng riêng và các tham số hình dạng. Sự xuất hiện của các AAM mô hình độc lập được xác định trong lưới cơ sở $s_0$. Sự xuất hiện mô hình được học bằng cách làm cong các hình ảnh đào tạo. Một sự xuất hiện z(x) có thể được biểu diễn như một sự xuất hiện cơ sở $z_0(x)$ cộng với một tổ hợp tuyến tính của v hình ảnh xuất hiện $\Phi_z(x)$:

$$z(x) = z_0(x) + \Phi_z(x)\lambda \quad (6)$$

Trong đó $\Phi_z(x) = (z_1(x), z_2(x), ..., z_v(x))$, các hệ số $\lambda = (\lambda_1, \lambda_2, ..., \lambda_v)$ là các tham số về ngoại hình. Các giá trị $s_0$ và s xác định một độ cong từ $s_0$ đến s. Độ cong biểu thị $W(x;c)$. Tập x là một điểm ảnh trong $s_0$, khi điểm ảnh trong ảnh I là $W(x;c)$. Ảnh có cường độ $I(W(x;c))$. Mô hình AAM ước tính các giá trị tối ưu:

$$\{\tilde{c}, \tilde{\lambda}\} = \underset{\{c,\lambda\}}{\operatorname{argmin}} \| I(W(x;c)) - z_0(x) - \Phi_z(x)\lambda \|_2^2 \quad (7)$$

Việc giải hàm tối ưu trong phương trình (7) được coi là khó do tồn tại nhiều cực tiểu cục bộ. Do đó, để giải quyết vấn đề này, mô hình biến thể AOM được đề xuất. AOM được mô hình hóa bằng sự tương quan của hình ảnh thực nghiệm với mô hình hình ảnh đã học. Các ước tính tối ưu được đưa ra bởi:

$$\{\tilde{c}, \tilde{\lambda}\} = \underset{\{c,\lambda\}}{\operatorname{argmax}} \frac{z(x)[c]^T \Phi_z \lambda}{\| \Phi_z \lambda \|_2} \quad (8)$$

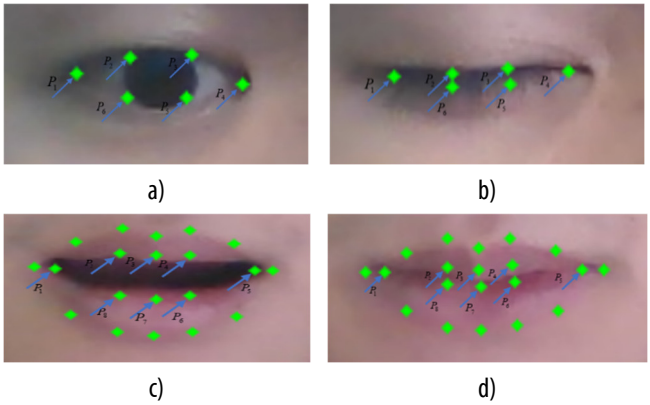
Phương pháp đề xuất được sử dụng để tạo chú thích cho điểm mốc trên khuôn mặt được trình bày bên trên. Các tham số mô hình được học và hiệu chỉnh bằng cách sử dụng tập dữ liệu Multi-PIE, bao gồm hơn 750.000 hình ảnh của 337 đối tượng với nhiều điều kiện chiếu sáng, góc nhìn và biểu cảm khác nhau. Đối với mỗi khung hình video theo thời gian thực, các điểm mốc của mắt được phát hiện. Tỷ lệ mắt (EAR) giữa chiều cao và chiều rộng của mắt được đưa ra bởi:

$$EAR = \frac{\| P_2 - P_6 \|_2 + \| P_3 - P_5 \|_2}{2 \| P_1 - P_4 \|_2} \quad (9)$$

Tỷ lệ cạnh miệng (MAR) được đưa ra bởi công thức:

$$MAR = \frac{\| P_2 - P_8 \|_2 + \| P_3 - P_7 \|_2 + + \| P_4 - P_6 \|_2}{2 \| P_1 - P_5 \|_2} \quad (10)$$

Trong đó, $P_j$ (j = 1…8) là các vị trí mốc 2D. $\| . \|_2$ biểu thị không gian véc tơ chuẩn. Hình 2 mô tả các điểm mốc đặc trưng để tính toán khoảng cách của mắt và miệng.



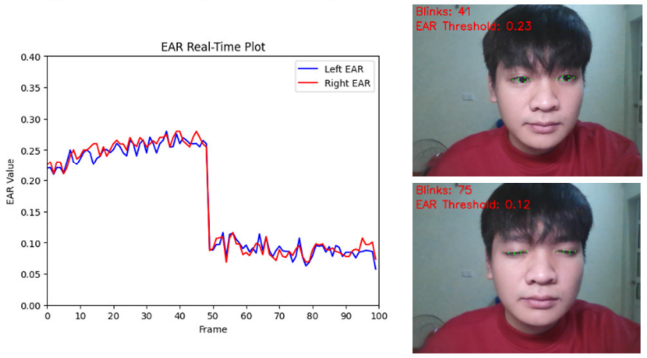
a) b) c) d)

Hình 2. Các điểm mốc tại mắt và miệng: a) Các điểm mốc trong trạng thái mở mắt; b) Các điểm mốc trong trạng thái nhắm mắt; c) Các điểm mốc trong trạng thái mở miệng; d) Các điểm mốc trong trạng thái đóng miệng

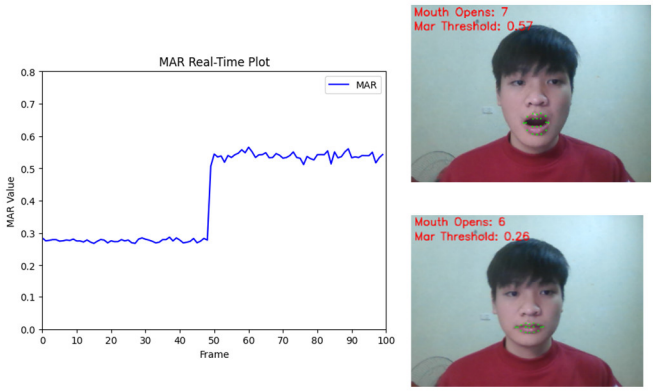
## 3. KẾT QUẢ VÀ THẢO LUẬN

Để đánh giá hiệu quả của hệ thống bảo mật hai lớp được đề xuất, chúng tôi tiến hành thử nghiệm trong các điều kiện khác nhau về ánh sáng và trạng thái người dùng (đeo kính hoặc không). Đồng thời, các cử chỉ khuôn mặt (nháy mắt, mở miệng) được theo dõi và phân tích theo thời gian thực nhằm xác định trạng thái của người dùng trong từng tình huống. Các thử nghiệm được thực hiện trong ba điều kiện chiếu sáng: đủ sáng, thiếu sáng và ánh

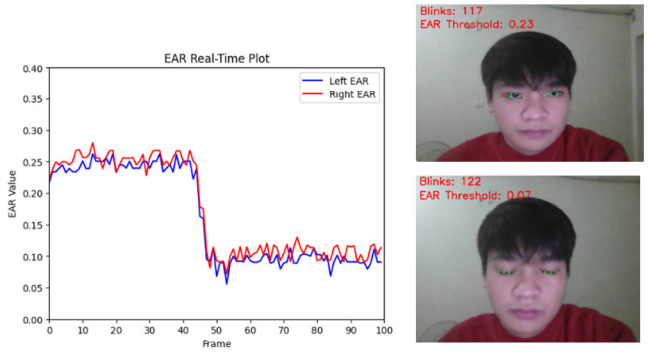
sáng yếu. Kết quả cho thấy tỷ lệ EAR và MAR có sự thay đổi rõ rệt giữa trạng thái mở và đóng của mắt, cũng như trạng thái mở miệng. Trong điều kiện đủ sáng, EAR dao động từ 0,24 đến 0,28 (mắt mở) và giảm xuống 0,07 - 0,12 khi mắt nhắm. MAR trong trạng thái miệng mở đạt khoảng 0,55, trong khi khi miệng đóng dao động quanh 0,28. Trong điều kiện thiếu sáng, giá trị EAR vẫn phân biệt được trạng thái mắt mở/nhắm, nhưng độ lệch tăng lên, cho thấy ảnh hưởng của ánh sáng yếu tới khả năng xác định chính xác. Trong điều kiện ánh sáng yếu, độ ổn định của hệ thống giảm đáng kể. EAR giữa mắt trái và phải có sự sai lệch lớn, ảnh hưởng đến độ chính xác khi phát hiện hành vi. Hình 3 ÷ 8 mô tả tỷ lệ EAR và MAR trong điều kiện cường độ ánh sáng từ đủ sáng, thiếu sáng, ánh sáng yếu.



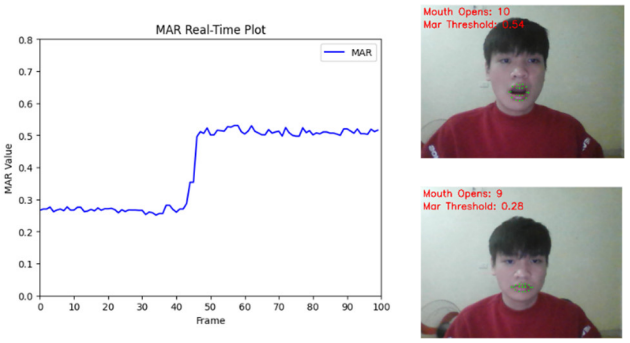

Hình 3. Tỷ lệ mắt trong điều kiện đủ sáng



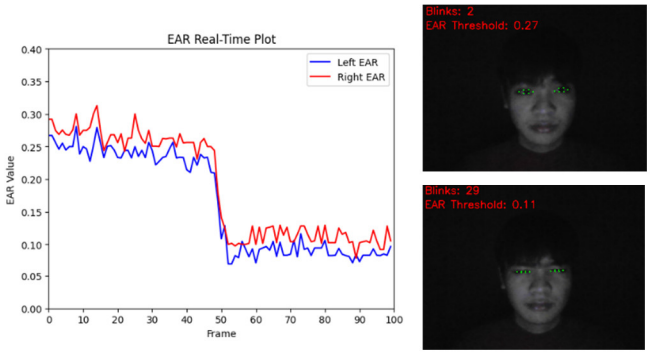

Hình 4. Tỷ lệ miệng trong điều kiện đủ sáng



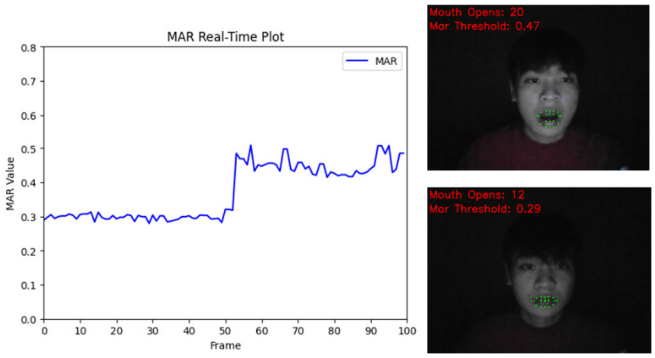

Hình 5. Tỷ lệ mắt trong điều kiện thiếu ánh sáng



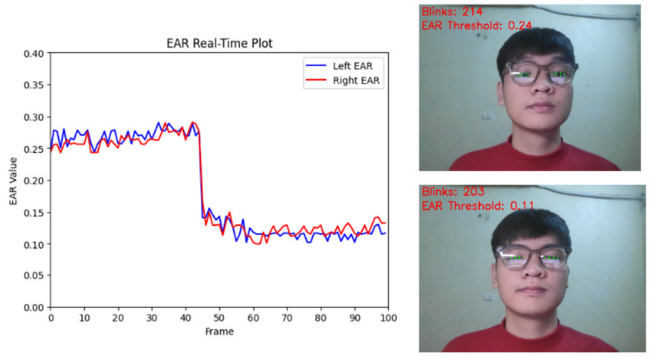

Hình 6. Tỷ lệ miệng trong điều kiện thiếu ánh sáng




Hình 7. Tỷ lệ mắt trong điều kiện thiếu ánh sáng yếu




Hình 8. Tỷ lệ miệng trong điều kiện thiếu ánh sáng yếu




Hình 9. Tỷ lệ mắt khi đeo kính trong trong điều kiện đủ sáng

Thử nghiệm tiếp theo được thực hiện với các đối tượng đeo kính (hình 9). Kết quả cho thấy rằng tỷ lệ EAR dao động từ 0,25 - 0,30 (mắt mở) và 0,12 - 0,15 (mắt nhắm), cho thấy hệ thống vẫn có thể nhận diện chính xác các cử chỉ mắt ngay cả khi người dùng đeo kính. Trong khí

đó tỷ lệ MAR không bị ảnh hưởng đáng kể bởi việc đeo kính, chứng minh độ ổn định của hệ thống trong điều kiện thay đổi phụ kiện cá nhân. Điều này cho thấy mô hình đề xuất có khả năng kháng nhiễu phụ kiện tốt, một yếu tố quan trọng trong các ứng dụng thực tế.

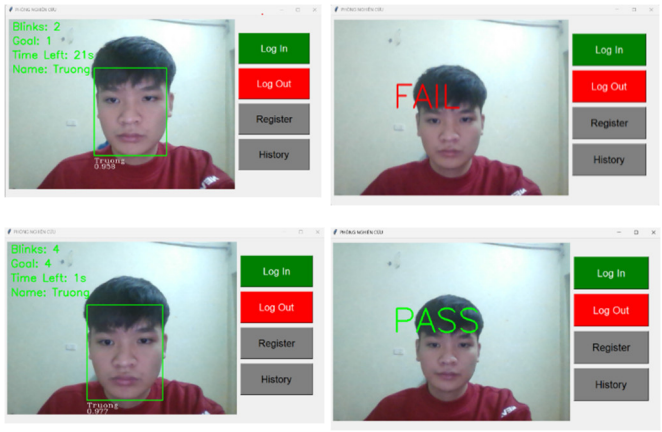
Lớp bảo mật thứ hai được đánh giá thông qua ba kịch bản kiểm thử:

(1) Cử chỉ nhắm/mở mắt: Hệ thống yêu cầu người dùng thực hiện số lần nháy mắt ngẫu nhiên trong khoảng thời gian xác định. Nếu số lần thực hiện khớp với yêu cầu, xác thực được chấp nhận; ngược lại, hệ thống từ chối truy cập.

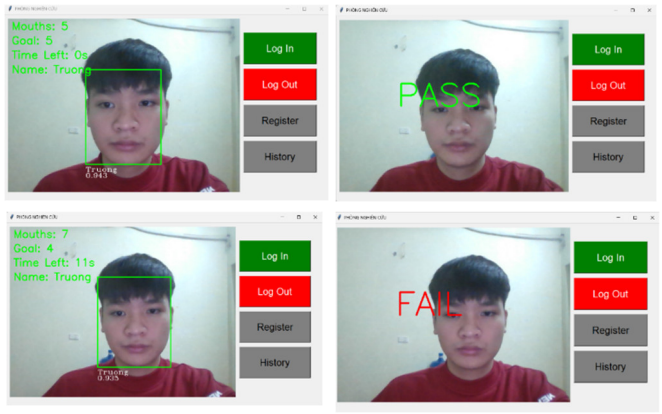
(2) Cử chỉ mở/đóng miệng: Áp dụng nguyên tắc tương tự như kiểm tra mắt, kiểm tra số lần mở miệng so với lệnh hệ thống.

(3) Kết hợp mắt và miệng: Người dùng phải thực hiện đồng thời cả hai loại cử chỉ. Đây là bài kiểm tra phức tạp nhất, yêu cầu hệ thống đồng bộ xử lý đa điểm mốc.

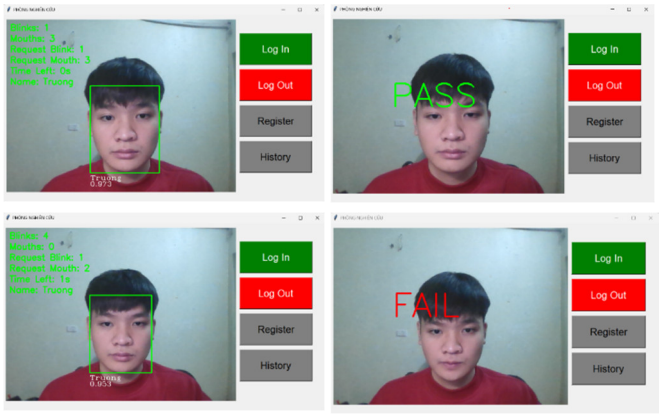
Kết quả được thể hiện trên hình 10 ÷ 12 cho thấy hệ thống hoạt động hiệu quả trong cả ba kịch bản. Nếu người dùng sử dụng ảnh tĩnh hoặc video, họ không thể thực hiện đúng các cử chỉ động như hệ thống yêu cầu, từ đó giúp loại bỏ các hành vi giả mạo.




Hình 10. Thực hiện xác thực danh tính và cử chỉ đóng mở mắt




Hình 11. Thực hiện xác thực danh tính và cử chỉ đóng mở miệng




Hình 12. Thực hiện xác thực danh tính và cử chỉ đóng kết hợp mắt và miệng

## 4. KẾT LUẬN

Nghiên cứu này đã đề xuất một hệ thống bảo mật hai lớp dựa trên sự kết hợp giữa trí tuệ nhân tạo (AI) và công nghệ thị giác máy, nhằm nâng cao khả năng nhận diện khuôn mặt và phát hiện hành vi giả mạo trong các hệ thống xác thực sinh trắc học. Cấu trúc hệ thống gồm: Lớp bảo mật thứ nhất sử dụng mạng học sâu (FaceNet kết hợp MTCNN) để nhận diện khuôn mặt dựa trên đặc trưng không gian nhúng; Lớp bảo mật thứ hai sử dụng mô hình toán học (AAM, AOM) để theo dõi và xác minh các cử chỉ khuôn mặt theo thời gian thực như chớp mắt và mở miệng. Kết quả thực nghiệm cho thấy hệ thống có khả năng nhận diện khuôn mặt với độ chính xác khoảng 95%, hoạt động ổn định trong nhiều điều kiện môi trường và hình thức của người dùng (kể cả khi đeo kính). Thuật toán xác thực cử chỉ giúp phát hiện các hành vi giả mạo bằng ảnh tĩnh hoặc mặt nạ, vượt qua hạn chế phổ biến của các hệ thống AI đơn tầng. Các chỉ số sinh trắc học như EAR và MAR được chứng minh là hiệu quả trong việc theo dõi trạng thái mắt và miệng, hỗ trợ đánh giá tính sống của đối tượng. Bên cạnh những kết quả khả quan, hệ thống vẫn tồn tại một số hạn chế nhất định, đặc biệt khi vận hành trong điều kiện ánh sáng yếu hoặc khi điểm mốc khuôn mặt bị che khuất. Trong các nghiên cứu tương lai, nhóm tác giả dự kiến sẽ tích hợp thêm các phương pháp xử lý ảnh tăng cường, kỹ thuật học tăng cường và tối ưu hóa thời gian phản hồi nhằm nâng cao tính chính xác và hiệu năng hệ thống trong môi trường thực tế phức tạp hơn.

### TÀI LIỆU THAM KHẢO

[1]. Abdulganiyu O. H., Ait Tchakoucht T., Saheed Y. K., "A systematic literature review for network intrusion detection system (IDS)," *International Journal of Information Security*, 22(5), 1125-1162, 2023.

[2]. Kaur P., Krishan K., Sharma S. K., Kanchan T., "Facial-recognition algorithms: A literature review," *Medicine, Science and the Law*, 60(2), 131-139, 2020.

[3]. Venkatasubramanian M., Lashkari A. H., Hakak S., "*IoT malware analysis using federated learning: A comprehensive survey*," *IEEE Access*, 11, 5004-5018, 2023.

[4]. Ali A., Abd Razak S., Othman S. H., Eisa T. A. E., Al-Dhaqm A., Nasser M., Saif A., "Financial fraud detection based on machine learning: A systematic literature review," *Applied Sciences,* 12(19), 9637, 2022.

[5]. Hossain M. D., Inoue H., Ochiai H., Fall D., Kadobayashi Y., "LSTM-based intrusion detection system for in-vehicle can bus communications," *IEEE Access*, 8, 185489-185502, 2020.

[6]. Schroff F., Kalenichenko D., Philbin J., "Facenet: A unified embedding for face recognition and clustering," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 815-823, 2015.

[7]. Deng J., Guo J., Xue N., Zafeiriou S., "Arcface: Additive angular margin loss for deep face recognition," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition,* 4690-4699, 2019.

[8]. Taigman Y., Yang M., Ranzato M. A., Wolf L., "Deepface: Closing the gap to human-level performance in face verification," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1701-1708, 2014.

[9]. Tran T. N., Van Nghia T., Tho N. Q., Minh N. D., "Real-time eye ratio tracking for drowsiness detection based on machine vision and facial landmarks," *Journal of Science & Technology*, *Hanoi University of Industry*, 58 (6A), 42-46, 2022.

[10]. Guarin D.L., Yunusova Y., Taati B., Dusseldorp J.R., Mohan S., Tavares J., van Veen M.M., Fortier E., Hadlock T.A., Jowett N., "Toward an automatic system for computer-aided assessment in facial palsy," *Facial Plast. Surg. Aesthetic Med.,* 22, 42 49, 2020.

[11]. Florez R., Palomino-Quispe F., Alvarez A. B., Coaquira-Castillo R. J., Herrera-Levano J. C., "A Real-Time Embedded System for Driver Drowsiness Detection Based on Visual Analysis of the Eyes and Mouth Using Convolutional Neural Network and Mouth Aspect Ratio," *Sensors,* 24(19), 6261, 2024.

**AUTHORS INFORMATION**

**Tran Ngoc Tien[1], Nguyen Quuc Truong[1], Truong Cong Giang[2]**

[1]School of Mechanical and Automotive Engineering, Hanoi University of Industry, Vietnam

[2]Vinh Phuc Technology - Economic College, Vietnam