



NON-CONTACT RESPIRATORY RATE ESTIMATION USING DENSE OPTICAL FLOW AND EDGE COMPUTING ON RASPBERRY PI 5

Hua Manh Truong, Doan Manh Tu, Viet Thanh Pham*

Faculty of Electronics Technology, Industrial University of Ho Chi Minh City, No. 12 Nguyen Van Bao Street, Ho Chi Minh City, Vietnam

ARTICLE INFO

TYPE: Research Article

Received: 25/12/2025

Revised: 12/03/2026

Accepted: 25/05/2026

Published online: 15/06/2026

<https://doi.org/10.47869/tcsj.77.5.10>

* *Corresponding author*

Email: phamvietthanh@iuh.edu.vn; Tel: 0915575666

Abstract. The development of non-contact health monitoring systems is becoming increasingly essential in remote healthcare and early intervention support. In this paper, we introduce HealthCam, a respiratory rate monitoring system using a Raspberry Pi 5 combined with a 3.5-inch LCD and a webcam to capture real-time video. The HealthCam system identifies the shoulder–chest ROI using MediaPipe Pose, extracts respiratory signals through landmark displacement and optical flow, applies bandpass filtering (0.1–2.0 Hz), and uses multi-method peak detection, including adaptive thresholding, Hilbert phase, and derivative-based methods, to calculate BPM. By integrating edge computing, Firebase data synchronization to an authenticated web dashboard, and a dual-chatbot system (Typebot for surveys and Groq API for AI consultation), users can observe four charts: respiratory rate per minute, breath count over the measurement period, average respiratory rate per minute, and average breathing cycle. HealthCam supports time-based filtering (day, week, month, and 1–2 years) and user profile storage. Validation over 50 measurement sessions showed a Pearson correlation coefficient of $r = 0.96$ compared with the manual counting method, demonstrating stability under different conditions. The proposed solution offers advantages in low cost, ease of deployment, real-time operation, and AI integration to support early detection and personalized respiratory management, providing high practical applicability.

Keywords: non-contact monitoring, Raspberry Pi, respiratory rate, MediaPipe, Firebase, AI chatbot.



ƯỚC LƯỢNG NHỊP THỞ KHÔNG TIẾP XÚC SỬ DỤNG LUỒNG QUANG HỌC VÀ ĐIỆN TOÁN BIÊN TRÊN RASPBERRY PI 5

Hứa Mạnh Trường, Đoàn Mạnh Tú, Phạm Việt Thành*

Khoa Công nghệ Điện tử, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh, Số 12 Nguyễn Văn Bảo, Thành phố Hồ Chí Minh, Việt Nam

THÔNG TIN BÀI BÁO

CHUYÊN MỤC: Công trình khoa học

Ngày nhận bài: 25/12/2025

Ngày nhận bài sửa: 12/03/2026

Ngày chấp nhận đăng: 25/05/2026

Ngày xuất bản Online: 15/06/2026

<https://doi.org/10.47869/tcsj.77.5.10>

* Tác giả liên hệ

Email: phamvietthanh@iuh.edu.vn; Tel: 0915575666

Tóm tắt. Việc phát triển hệ thống giám sát sức khỏe không tiếp xúc đang trở nên cấp thiết trong chăm sóc y tế từ xa và hỗ trợ can thiệp sớm. Trong bài báo này, chúng tôi giới thiệu một hệ thống HealthCam giám sát nhịp thở sử dụng Raspberry Pi 5 kết hợp webcam để thu video thời gian thực. Hệ thống HealthCam xác định ROI vai-ngực qua MediaPipe Pose, trích xuất tín hiệu hô hấp bằng dịch chuyển landmark và luồng quang học, áp dụng lọc bandpass và phát hiện đỉnh đa phương pháp để tính BPM. Thông qua việc kết hợp điện toán biên, dữ liệu đồng bộ Firebase lên dashboard web với xác thực, và chatbot kép, người sử dụng có thể quan sát bốn biểu đồ: nhịp thở mỗi phút, số nhịp thở theo thời gian đo, trung bình nhịp thở mỗi phút, và chu kỳ thở trung bình. HealthCam hỗ trợ lọc theo thời gian và lưu trữ hồ sơ người dùng. Việc kiểm chứng trên 50 phiên đo cho thấy hệ số tương quan Pearson $r = 0.96$ so với phương pháp đếm thủ công và độ ổn định trong các điều kiện khác nhau. Giải pháp đề xuất có lợi thế về chi phí thấp, dễ triển khai, khả năng chạy thời gian thực cùng với việc tích hợp AI hỗ trợ phát hiện sớm và quản lý hô hấp cá nhân hóa mang đến tính ứng dụng thực tiễn cao.

Từ khóa: giám sát không tiếp xúc, Raspberry Pi, nhịp thở, MediaPipe, Firebase, chatbot AI.

1. ĐẶT VẤN ĐỀ

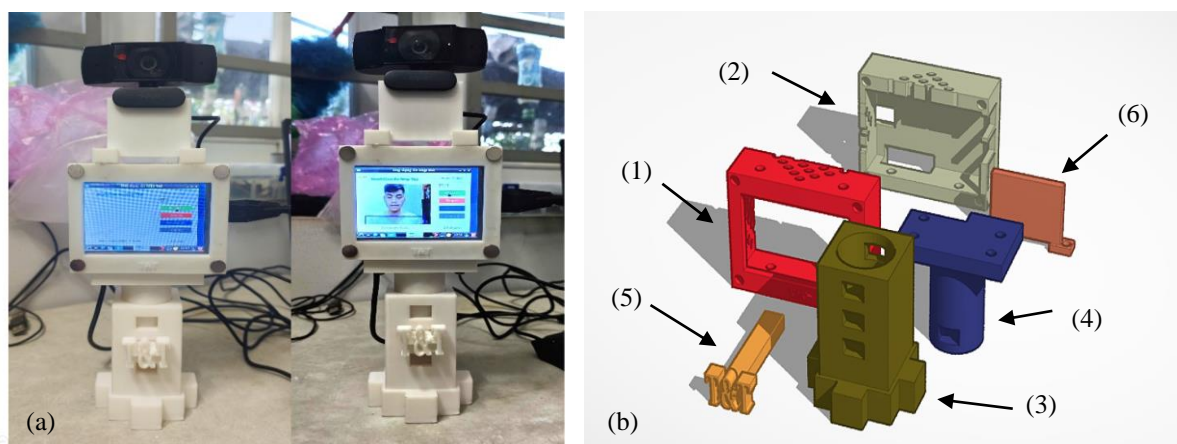
Các phương pháp truyền thống như dây đai đo biến dạng lồng ngực, cảm biến áp lực hoặc thiết bị đo lưu lượng khí tuy đạt độ chính xác cao nhưng thường gây khó chịu, hạn chế chuyển động tự nhiên và khó áp dụng trong giám sát dài hạn tại cộng đồng [1]. Trong những năm gần đây, các nghiên cứu đo nhịp thở không tiếp xúc dựa trên thị giác máy tính đã nhận được nhiều sự quan tâm, đặc biệt là các phương pháp sử dụng luồng quang học (optical flow) hoặc phân tích chuyển động vùng ngực–vai từ dữ liệu khung xương (pose landmarks). Ví dụ, một số nghiên cứu đã khai thác chuyển động vùng thân để ước lượng nhịp thở [2], trong khi các phương pháp phân tích tín hiệu hình ảnh cho giám sát hô hấp cũng được đề xuất [3]. Tuy nhiên, các phương pháp hiện tại vẫn tồn tại một số hạn chế như nhạy với chuyển động toàn thân, phụ thuộc mạnh vào chất lượng vùng quan tâm (ROI – Region of Interest, vùng quan tâm), hoặc thiếu khả năng vận hành trực tiếp trên thiết bị nhúng chi phí thấp. Các công trình gần đây đã phát triển nhiều phương pháp đo nhịp thở không tiếp xúc dựa trên camera, bao gồm sử dụng RGB camera để trích xuất nhịp thở [4], phân tích hệ thống giám sát từ xa các chỉ số như nhịp thở, nhịp tim, huyết áp và SpO2 [5], camera độ sâu (depth camera) để theo dõi nhịp thở robust [6], fusion multispectral video cho nhịp thở và apnea [7], cũng như giám sát nhịp thở trong môi trường gây mê khó khăn bằng depth camera [8]. Ngoài ra, các nghiên cứu tổng quan và ứng dụng về đo lường sinh lý không tiếp xúc bằng camera [9,10] và trong môi trường nhiễu [11] đã chỉ ra tiềm năng lớn của công nghệ này cho giám sát các chỉ số sự sống. Tuy nhiên, ít công trình xây dựng hệ thống hoàn chỉnh trên thiết bị nhúng chi phí thấp với tích hợp edge computing, đồng bộ đám mây và AI hỗ trợ.

Bên cạnh đó, phần lớn các công trình trước đây chủ yếu tập trung vào thuật toán xử lý tín hiệu mà chưa xây dựng được một hệ thống hoàn chỉnh bao gồm thu thập dữ liệu tại biên (edge computing), đồng bộ đám mây, hiển thị dashboard (bảng điều khiển) trực quan và tích hợp trợ lý AI nhằm hỗ trợ giải thích kết quả cho người dùng cuối [12]. Điều này làm giảm khả năng ứng dụng thực tế trong các môi trường như gia đình, ký túc xá hoặc các nghiên cứu thực địa. Xuất phát từ các hạn chế trên, nghiên cứu này đề xuất HealthCam – một hệ thống giám sát nhịp thở không tiếp xúc sử dụng Raspberry Pi 5 kết hợp webcam và màn hình cảm ứng 3.5-inch. Hệ thống sử dụng MediaPipe Pose để xác định vùng vai-ngực, trích xuất tín hiệu hô hấp thông qua dịch chuyển landmark và phân tích dense optical flow, sau đó áp dụng các bước xử lý tín hiệu và phát hiện chu kỳ thở để ước lượng nhịp thở theo thời gian thực. Dữ liệu đo được đồng bộ lên nền tảng Firebase để hiển thị trên dashboard và hỗ trợ tư vấn thông qua chatbot AI. Hệ thống HealthCam được thiết kế dựa trên bốn tiêu chí chính: (1) độ chính xác và ổn định trong môi trường thực tế, (2) khả năng vận hành thời gian thực trên thiết bị nhúng chi phí thấp, (3) bảo mật dữ liệu người dùng thông qua xác thực và phân quyền, và (4) khả năng mở rộng cho các ứng dụng chăm sóc sức khỏe cá nhân hóa. Các nội dung tiếp theo của bài báo trình bày chi tiết thiết kế phần cứng, pipeline xử lý tín hiệu, kiến trúc hệ thống và đánh giá thực nghiệm.

2. PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Phần cứng

Hệ thống HealthCam được triển khai trên nền tảng nhúng Raspberry Pi 5, cho phép xử lý video thời gian thực với mức tiêu thụ điện năng thấp (5–7 W), phù hợp cho các môi trường hạn chế nguồn cấp [13]. Thiết bị chạy Raspberry Pi OS Lite 64-bit nhằm tối ưu tài nguyên cho tác vụ thị giác máy và mô hình học sâu nhẹ. Màn hình LCD cảm ứng 3.5-inch (480 × 320 pixel, giao tiếp SPI/I2C) được tích hợp trực tiếp và giao diện điều khiển được xây dựng bằng Tkinter, cho phép vận hành hoàn toàn qua cảm ứng mà không cần thiết bị ngoại vi [14]. Hệ thống sử dụng webcam USB chuẩn V4L2 với tốc độ 30 FPS và độ trễ dưới 50 ms, phù hợp cho nhận dạng chuyển động vùng ngực–bụng trong bài toán đo nhịp thở không tiếp xúc. Dữ liệu được truyền qua Wi-Fi 6 hoặc Ethernet Gigabit lên Firebase Realtime Database và hỗ trợ tích hợp Groq API cho chức năng tư vấn. Thiết bị dùng nguồn USB-C 5 V–5 A, lưu trữ bằng thẻ microSD 64 GB, và vỏ được in 3D bằng PLA với chân đế điều chỉnh góc camera. Hệ thống được tối ưu để duy trì tốc độ xử lý thực tế từ 27–32 FPS trên Raspberry Pi 5, đáp ứng yêu cầu thời gian thực cho bài toán giám sát nhịp thở không tiếp xúc. Thiết bị có khối lượng dưới 350 g, kích thước 12 × 8 × 6 cm, có thể hoạt động độc lập mà không cần máy tính đi kèm, phù hợp triển khai tại gia đình, trạm y tế cơ sở và các nghiên cứu thực địa (Hình 1).



Hình 1. Cấu hình phần cứng và cấu trúc cơ khí của hệ thống HealthCam. (a) Hệ thống HealthCam thực tế gồm Raspberry Pi 5, webcam USB, màn hình LCD cảm ứng 3.5-inch và vỏ thiết bị in 3D với chân đế điều chỉnh góc quan sát. (b) Mô hình 3D các khối cấu thành thiết bị: (1) khối vỏ trước, (2) khối vỏ sau, (3) trụ đế đứng, (4) khối nâng hạ chiều cao, (5) thanh cố định độ cao, và (6) khối gá camera phía trên.

2.2. Phần mềm

2.2.1. Xử lý tín hiệu hô hấp thời gian thực trên thiết bị nhúng

Hệ thống được triển khai trên Raspberry Pi 5 chạy Raspberry Pi OS 64-bit, sử dụng Python cùng các thư viện OpenCV, MediaPipe, SciPy và NumPy. Kiến trúc đa luồng hỗ trợ truyền dữ liệu qua hàng đợi giúp duy trì tốc độ xử lý thực tế từ 27–32 FPS và độ trễ dưới 40 ms mỗi khung hình trên Raspberry Pi 5. Hình 2 trình bày lưu đồ giải thuật đo nhịp thở, trong đó quá trình bắt

đầu bằng việc nhận dạng các mốc giải phẫu của MediaPipe Pose và xác định tâm vùng quan tâm (ROI) từ trung điểm hai vai theo công thức:

$$c = \frac{L_{11} + L_{22}}{2} \quad (1)$$

Chiều rộng ROI được thiết lập bằng 1,6 lần khoảng cách hai vai:

$$W = 1,6 \parallel L_{11} - L_{22} \parallel \quad (2)$$

Chiều cao ROI được xác định theo khoảng cách vai – hông nhằm bảo đảm vùng quan sát luôn bao phủ dao động ngực-bụng trong nhiều tư thế khác nhau. Sau khi xác lập ROI, hai nguồn tín hiệu hô hấp độc lập được trích xuất song song. Tín hiệu thứ nhất dựa trên dao động dọc của hai vai, trong đó vị trí theo trục y được tính bằng trung bình hai vai:

$$y_{\text{shoulder}}(t) = \frac{y_{11}(t) + y_{12}(t)}{2} \quad (3)$$

Để giảm nhiễu do chuyển động toàn thân, hệ thống sử dụng thêm cơ chế kiểm tra độ tin cậy pose (pose confidence) và độ ổn định landmark. Khi phát hiện biến thiên lớn bất thường vượt quá ngưỡng MAD trong cửa sổ ngắn, các khung hình tương ứng sẽ bị loại khỏi quá trình tính chu kỳ thở. Ngoài ra, để xử lý các tình huống mất landmark tạm thời do quay người, cúi xuống, hoặc vật cản che khuất vùng ngực (ví dụ: tay che, áo khoác dày di chuyển), hệ thống áp dụng cơ chế temporal persistence cho ROI. Cụ thể, nếu landmark vai trái/phải ($left_{\text{shoulder}}$ và $right_{\text{shoulder}}$) có visibility < 0,5 hoặc không được phát hiện (*results.pose_landmarks* is None hoặc landmark.visibility thấp), hệ thống giữ nguyên tọa độ ROI (x1, y1, x2, y2) từ khung hình gần nhất có pose hợp lệ. Cơ chế này được giới hạn chặt chẽ để tránh trôi ROI (ROI drift) khi người dùng thay đổi tư thế mạnh: ROI chỉ được duy trì tối đa 12 khung hình liên tiếp (tương đương khoảng 0.4 giây ở tốc độ xử lý thực tế 27–32 FPS). Nếu mất landmark vượt quá ngưỡng này, phiên đo được đánh dấu là không hợp lệ (invalid session), hệ thống tạm dừng tính toán BPM, hiển thị cảnh báo trên giao diện Tkinter ("Vui lòng ngồi thẳng, giữ vùng ngực rõ ràng trước camera") và yêu cầu người dùng điều chỉnh tư thế để tiếp tục đo. Cơ chế temporal persistence giúp duy trì tính liên tục tín hiệu trong các gián đoạn ngắn (< 0,5s), đồng thời ngăn chặn tích lũy lỗi dài hạn. Trong thực nghiệm trên 50 phiên đo ban đầu, tỷ lệ phiên bị invalid do mất ROI kéo dài chỉ chiếm khoảng 7,7% (tương ứng với tỷ lệ ROI hợp lệ 92,3%), chủ yếu từ các trường hợp quay người đột ngột hoặc che khuất hoàn toàn vùng ngực.

Bên cạnh đó, để xử lý các chuyển động toàn thân dù nhỏ nhưng gây biến thiên tín hiệu lớn (ví dụ: ngồi không yên, cử tay nhẹ, rung nhẹ cơ thể), hệ thống áp dụng cơ chế "MAD-based outlier rejection". Cơ chế này hoạt động trên cửa sổ trượt ngắn (5–10 khung hình): tính median absolute deviation (MAD) của tín hiệu kết hợp; bất kỳ biến thiên nào vượt quá ngưỡng 3*MAD sẽ được coi là outlier và khung hình tương ứng bị loại khỏi quá trình phát hiện đỉnh/chu kỳ hô hấp. Phương pháp này lọc được nhiễu lớn mà không ảnh hưởng đến biên độ hô hấp nhỏ (thường < 5 pixel/frame). Kết hợp với việc trừ vị trí mũi (công thức (4)) để loại bỏ dao động đầu, hệ thống giảm đáng kể ảnh hưởng của chuyển động không liên quan. Trong thực nghiệm trên 50 phiên đo, cơ chế này giúp giảm sai số BPM trung bình 12–18% trong các phiên có chuyển động nhỏ, so với phiên bản không áp dụng "outlier rejection". Độ bền vững được chứng minh qua tỷ

lệ $IQR_{norm} < 0,5$ ở 81% phiên hợp lệ, ngay cả khi có nhiều chuyển động nhẹ. Dao động đầu được loại bỏ bằng cách trừ vị trí mũi, tạo thành tín hiệu landmark:

$$s_1(t) = y_{shoulder}(t) - y_0(t) \quad (4)$$

Tín hiệu thứ hai mô tả chuyển dịch thẳng đứng của ROI giữa hai khung hình liên tiếp:

$$s_2(t) = y_{ROI}(t) - y_{ROI}(t - 1) \quad (5)$$

Hai tín hiệu này được kết hợp tuyến tính với trọng số hiệu chỉnh nhằm tăng độ nhạy với dao động nhỏ của thành ngực:

$$s(t) = w_1 s_1(t) + w_2 s_2(t), \quad w_1 + w_2 = 1 \quad (6)$$

Trong thực nghiệm, các trọng số được thiết lập cố định với $w_1 = 0,7$ và $w_2 = 0,3$ nhằm ưu tiên tín hiệu “landmark displacement” vốn ổn định hơn trong điều kiện ánh sáng yếu, trong khi optical flow giúp tăng độ nhạy với các dao động nhỏ của thành ngực. Việc lựa chọn này dựa trên thử nghiệm sơ bộ trên nhiều phiên đo và cho thấy độ ổn định chu kỳ cao hơn so với các cấu hình trọng số khác. Các giá trị này được chọn dựa trên đánh giá độ ổn định chu kỳ hô hấp thông qua chỉ số IQR_{norm} (tỷ số giữa IQR và median của khoảng cách giữa các đỉnh phát hiện), với cấu hình 0,7/0,3 cho kết quả IQR_{norm} trung bình thấp nhất ($< 0,5$ ở khoảng 81% phiên đo hợp lệ). Hệ thống không sử dụng trọng số thích nghi theo điều kiện đo (ví dụ: dựa trên độ sáng hoặc độ biến thiên tín hiệu) nhằm giảm độ phức tạp tính toán và đảm bảo tốc độ xử lý thời gian thực ổn định trên Raspberry Pi 5 (27–32 FPS). Tín hiệu được lưu trong cửa sổ 20–60 giây để xử lý theo lô. Pipeline tiền xử lý bắt đầu bằng việc loại bỏ xu hướng tuyến tính:

$$s_d(t) = s(t) - \hat{s}_{linear}(t) \quad (7)$$

Sau đó tín hiệu được lọc median cửa sổ ba mẫu:

$$s_m(t) = \text{median}\{s_d(t - 1), s_d(t), s_d(t + 1)\} \quad (8)$$

Tiếp theo, làm mượt bằng bộ lọc EMA:

$$s_e(t) = \alpha s_m(t) + (1 - \alpha) s_e(t - 1), \quad \alpha = 0.05 \quad (9)$$

Tín hiệu sau đó được nội suy và tái lấy mẫu đều tại 30 Hz:

$$s_r(t) = \text{Resample}(s_e(t), 30 \text{ Hz}) \quad (10)$$

Bộ lọc thông dải Butterworth bậc ba 0,1–2 Hz được áp dụng bằng kỹ thuật lọc hai chiều nhằm duy trì pha:

$$s_b(t) = \text{filtfilt}(B, A, s_r(t)) \quad (11)$$

Trong các công thức trên, $y_{ROI}(t)$ biểu diễn chuyển dịch trung bình theo trục dọc của vùng ROI giữa hai khung hình liên tiếp; $\hat{s}_{linear}(t)$ là thành phần xu hướng tuyến tính được ước lượng bằng hồi quy tuyến tính trong cửa sổ trượt; α là hệ số làm mượt của bộ lọc EMA; θ là ngưỡng biên độ được xác định dựa trên độ lệch tuyệt đối trung vị; và k là hệ số điều chỉnh được chọn thực nghiệm trong khoảng 0,8 – 1,2 để tối ưu độ ổn định chu kỳ. MAD (median absolute deviation – độ lệch tuyệt đối trung vị) được sử dụng để phát hiện các biến thiên bất thường của tín hiệu và phục vụ chuẩn hóa robust trong công thức (12). Δt biểu diễn khoảng thời gian giữa

hai chu kỳ hô hấp liên tiếp được phát hiện từ tín hiệu đã lọc. IQR_norm là chỉ số đánh giá độ ổn định chu kỳ, được tính theo tỉ số giữa $IQR(\Delta t)$ và $median(\Delta t)$, trong đó giá trị nhỏ hơn thể hiện tín hiệu hô hấp ổn định hơn. Cuối cùng, chuẩn hóa robust bằng median và MAD được thực hiện như sau:

$$s_{norm}(t) = \frac{s_b(t) - median(s_b)}{1,4826 \cdot median(|s_b(t) - median(s_b)|)} \quad (12)$$

Giai đoạn phát hiện chu kỳ thở và ước lượng tần số được triển khai qua ba nhánh độc lập. Nhánh thứ nhất sử dụng phát hiện đỉnh thích nghi với khoảng cách tối thiểu được xác định động dựa trên:

$$s_{min} = k \cdot IQR(\Delta t) \quad (13)$$

Trong nhánh Hilbert, tín hiệu giải tích được xây dựng:

$$z(t) = s_{norm}(t) + j H \{s_{norm}(t)\} \quad (14)$$

Pha được trích xuất từ:

$$\phi(t) = \arg(z(t)) \quad (15)$$

và đỉnh được phát hiện tại các điểm thỏa mãn:

$$\frac{d\phi(t)}{dt} = 0 \quad \wedge \quad \cos(\phi(t)) < 0 \quad (16)$$

Nhánh thứ ba sử dụng zero-crossing của đạo hàm:

$$s'(t) = \frac{d}{dt} s_{norm}(t) \quad (17)$$

Chu kỳ được xác định tại thời điểm:

$$s'(t) = 0, |s_{norm}(t)| > \theta \quad (18)$$

Ba nhánh được đánh giá bằng độ ổn định chu kỳ:

$$IQR_{norm} = \frac{IQR(\Delta t)}{median(\Delta t)} \quad (19)$$

Nhánh có độ ổn định cao nhất được chọn để tính tần số hô hấp theo:

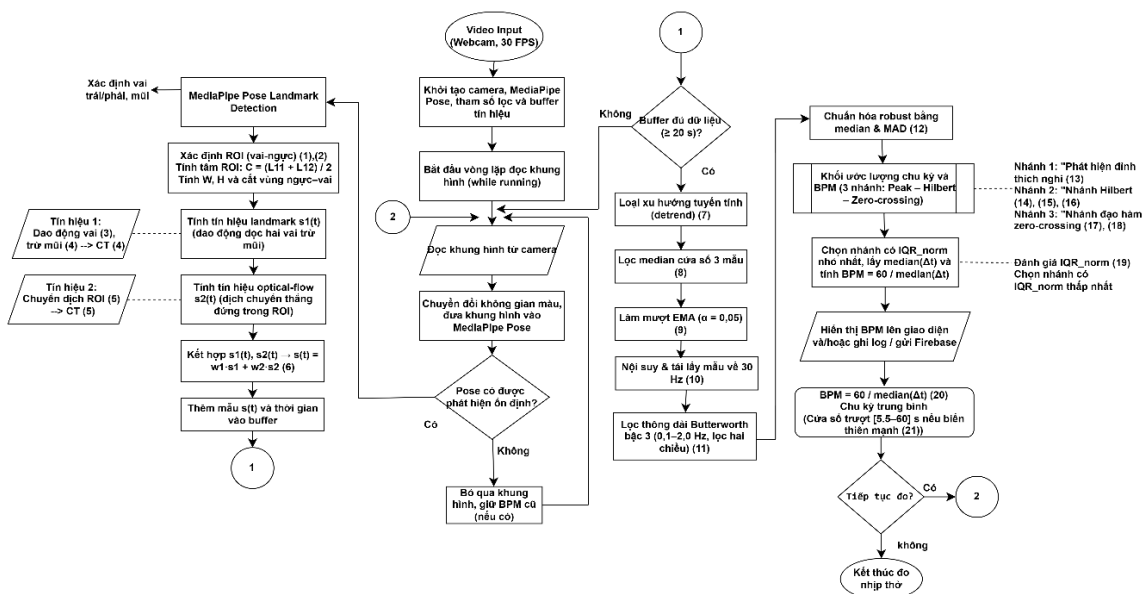
$$BPM = \frac{60}{median(\Delta t)} \quad (20)$$

Khi tín hiệu biến thiên mạnh, hệ thống áp dụng cửa sổ trượt trong khoảng:

$$T_{window} \in [5, 60] \text{ s} \quad (21)$$

Toàn bộ pipeline từ trích xuất tín hiệu, tiền xử lý, phát hiện chu kỳ đến ước lượng BPM được thực thi theo thời gian thực trong lớp phần mềm chuyên dụng, bảo đảm độ trễ xử lý luôn duy trì dưới 40 ms mỗi khung hình. Các bước xử lý và các công thức tương ứng trong pipeline được minh họa trong Hình 2. Cụ thể, giai đoạn xác định vùng quan tâm (ROI) sử dụng các công

thức (1) - (2); giai đoạn trích xuất tín hiệu hô hấp từ dữ liệu landmark và luồng quang học tương ứng với các công thức (3) - (5); việc kết hợp hai nguồn tín hiệu được thực hiện theo công thức (6); giai đoạn tiền xử lý tín hiệu bao gồm các bước loại bỏ xu hướng, lọc trung vị, làm mượt EMA, nội suy – tái lấy mẫu và lọc thông dải theo các công thức (7) - (12); cuối cùng, việc phát hiện chu kỳ thở và ước lượng nhịp thở được thực hiện theo các công thức (13) - (21).



Hình 2. Lưu đồ thuật toán ước lượng nhịp thở của hệ thống HealthCam.

2.2.2. Ứng dụng trên thiết bị

Giao diện ứng dụng được phát triển trên nền Tkinter và CustomTkinter, tối ưu cho màn hình cảm ứng 3,5-inch của Raspberry Pi 5 nhằm bảo đảm khả năng thao tác thuận tiện trong môi trường thực địa. Hệ thống tích hợp cơ chế đăng nhập/đăng ký bằng Firebase Authentication và hỗ trợ chế độ “offline fallback”, cho phép tiếp tục đo và lưu trữ dữ liệu tại chỗ khi mất kết nối mạng; toàn bộ dữ liệu cục bộ sẽ được tự động đồng bộ lên máy chủ ngay khi kết nối được khôi phục. Trong chế độ đo, giao diện hiển thị luồng video thời gian thực với lớp phủ ROI giúp định vị chính xác vùng ngực-vai, đồng thời cập nhật giá trị BPM liên tục theo thời gian thực. Một biểu đồ matplotlib nhúng trong Tkinter trình bày tín hiệu hô hấp hoặc biến thiên nhịp thở trong 60 giây gần nhất, hỗ trợ người dùng theo dõi xu hướng một cách trực quan. Người dùng có thể tùy chọn thời lượng ghi đo từ 20–60 giây; sau khi hoàn tất, dữ liệu được lưu dưới dạng CSV và tự động đồng bộ lên Firebase Realtime Database khi tài khoản đã được xác thực.

Các tiến trình nền được quản lý thông qua cơ chế đa luồng kết hợp với nút Start/Stop và bộ xử lý sự kiện đóng ứng dụng, giúp giải phóng an toàn camera và pipeline pose estimation, tránh hiện tượng treo hệ thống hoặc rò rỉ tài nguyên. Bên cạnh đó, giao diện được thiết kế với bố cục đơn giản, phông chữ lớn và các nút chức năng rõ ràng, phù hợp cho người dùng không chuyên kỹ thuật và thao tác nhanh trong điều kiện ánh sáng hoặc không gian hạn chế. Giao diện thực tế của ứng dụng được trình bày trong Hình 3.



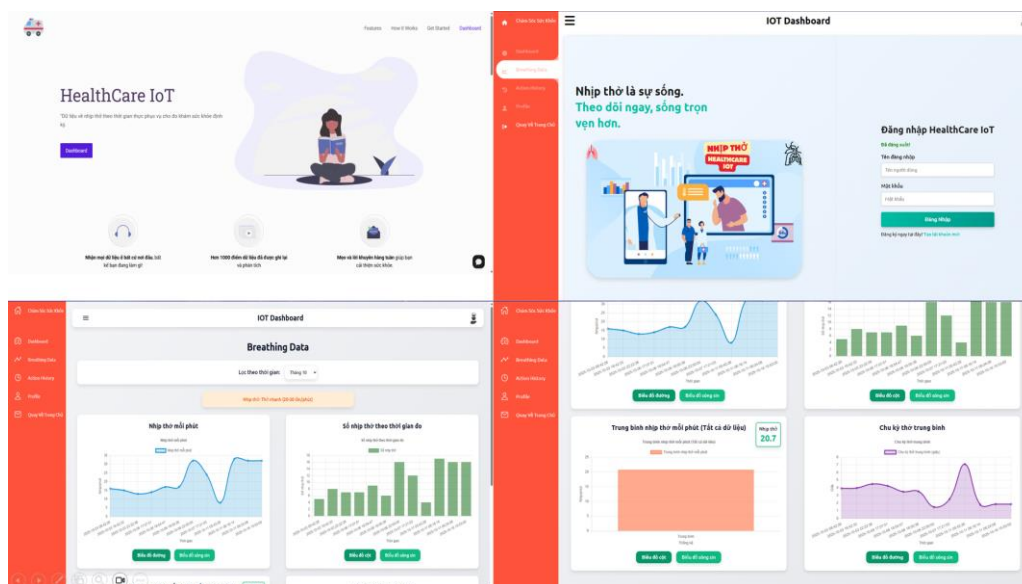
Hình 3. Giao diện ứng dụng HealthCam trên màn hình cảm ứng.

2.3. Hệ thống đám mây và tư vấn AI

2.3.1. Kiến trúc web và cơ sở dữ liệu

Hệ thống triển khai kiến trúc web trên nền tảng Firebase với mô hình serverless, không yêu cầu máy chủ chuyên dụng. Backend sử dụng Firebase Authentication để xác thực người dùng, Firebase Realtime Database để lưu trữ dữ liệu đo thời gian thực, và Cloud Functions để xử lý tác vụ phía server như thống kê theo ngày/tuần/tháng hoặc tạo cảnh báo bất thường. Frontend được xây dựng bằng HTML5 và Bootstrap 5, tích hợp Chart.js 4.4 để trực quan hóa dữ liệu và DataTables để hiển thị bảng thông tin với tính năng tìm kiếm, phân trang. Ứng dụng được host trên Firebase Hosting, cho phép kết nối trực tiếp với Authentication và Realtime Database mà không cần cấu hình hạ tầng phức tạp.

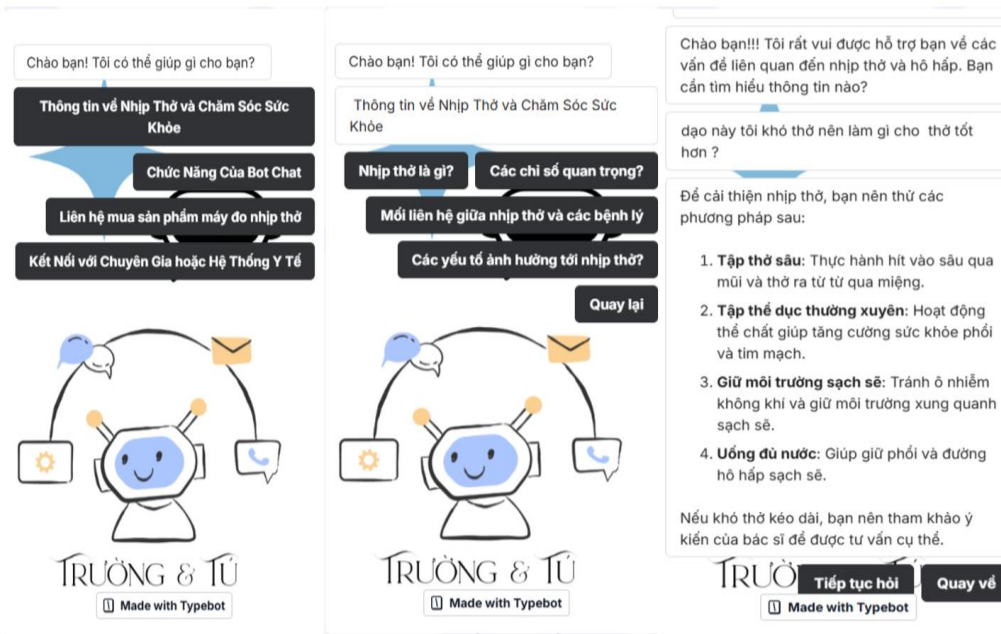
Cơ sở dữ liệu tổ chức dạng cây: nhánh /data_nhiptho/{uid}/{timestamp} lưu thông tin từng phiên đo (BPM, chu kỳ trung bình, số nhịp, thời gian đo); nhánh /users/{uid}/profile lưu hồ sơ người dùng để hỗ trợ phân tích và tư vấn chatbot cá nhân hóa. Bảo mật được đảm bảo qua Firebase Security Rules với điều kiện request.auth.uid == uid nhằm bảo vệ dữ liệu sức khỏe [15]. Dashboard cung cấp bốn nhóm biểu đồ chính: nhịp thở theo thời gian, phân bố BPM theo ngày/tuần/tháng, số nhịp thở theo phiên đo, và xu hướng chu kỳ thở trung bình (Hình 4).



Hình 4. Giao diện website và dashboard web cung cấp 4 nhóm biểu đồ chính.

2.3.2. Chatbot tư vấn sức khỏe thông minh

Hệ thống HealthCam tích hợp chatbot hai lớp để kết hợp tư vấn y khoa chuẩn hóa với phản hồi AI thời gian thực, cá nhân hóa dựa trên dữ liệu đo và hồ sơ người dùng. Lớp thứ nhất sử dụng Typebot (self-hosted trên Railway/Vercel), cung cấp kịch bản hội thoại cố định đã biên soạn và kiểm duyệt để đảm bảo tính chính xác lâm sàng [16]. Nội dung bao gồm: giải thích sinh lý nhịp thở, ngưỡng nhịp thở bình thường theo độ tuổi (theo WHO và Bộ Y tế), mối liên hệ giữa rối loạn nhịp thở với bệnh lý hô hấp–tim mạch, các yếu tố nguy cơ (thuốc lá, béo phì, tư thế ngủ), và các bài tập thở cải thiện chức năng hô hấp. Lớp thứ hai sử dụng mô hình Llama-3.3-70b-versatile (Meta Llama 3.3 70B) trên Groq Cloud để xử lý câu hỏi mở với độ trễ thấp. Truy vấn được gửi qua endpoint <https://api.groq.com/openai/v1/chat/completions>, tuân thủ đầy đủ các tham số API (model, messages, temperature, max_tokens, stream, ...). Thử nghiệm cho thấy thời gian phản hồi trung bình 0,9–1,4 giây, nhanh hơn đáng kể so với Llama 3.1 70B nhờ tối ưu hóa của Groq. System prompt được thiết kế đồng bộ với Typebot, nhắc lại ngưỡng BPM sinh lý, gợi ý thăm khám khi có dấu hiệu bất thường, và sử dụng ngôn ngữ tự nhiên, dễ hiểu cho người lớn tuổi. Giao diện hợp nhất của hai lớp chatbot được minh họa tại Hình 5 [17,18].



Hình 5. Giao diện thực tế của chatbot HealthCam với hai lớp Typebot và Groq AI Advisor hoạt động song song trên nền web.

3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Điều kiện và đối tượng thử nghiệm

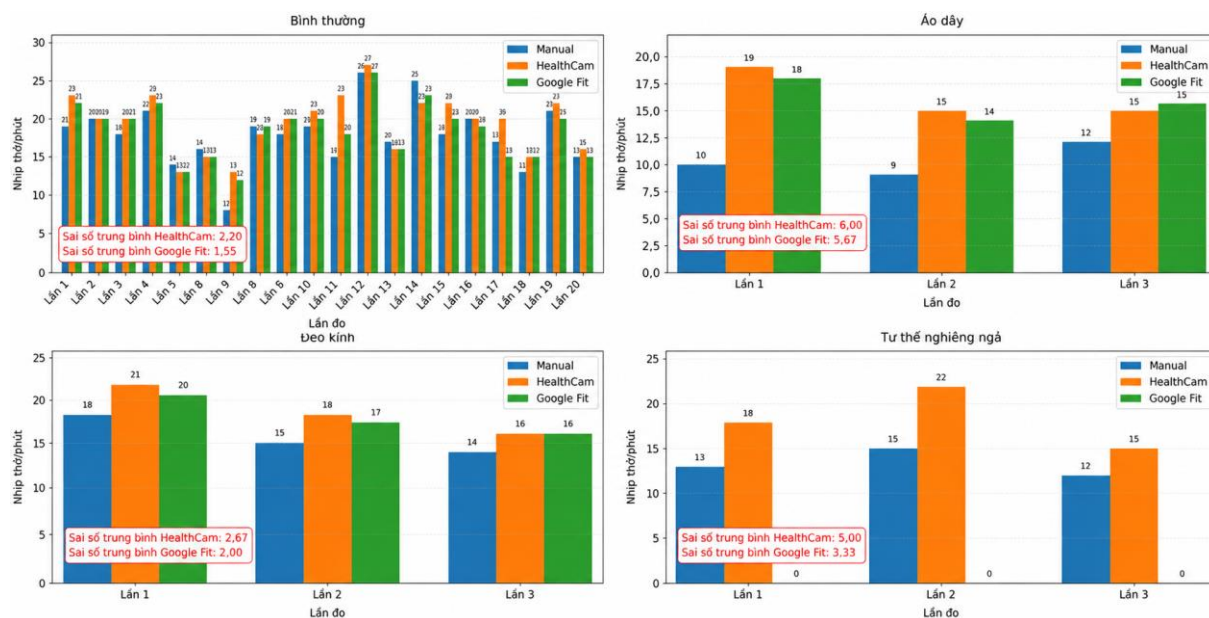
Thử nghiệm được tiến hành từ tháng 9 đến tháng 11 năm 2025 với tổng cộng 50 phiên đo thu thập từ 40 tình nguyện viên gồm sinh viên, người trưởng thành và một số người cao tuổi trong gia đình nhóm nghiên cứu. Do đặc thù môi trường đào tạo, phần lớn đối tượng là sinh viên nam trong độ tuổi 21–23, điều này được xem là một hạn chế của nghiên cứu và sẽ được mở rộng trong các nghiên cứu tiếp theo. Sau bước kiểm tra chất lượng dữ liệu, các phiên đo gặp lỗi nhận diện khuôn mặt từ Google Fit hoặc có chuyển động lớn gây mất ổn định tín hiệu được loại bỏ khỏi phân tích tương quan. Tập dữ liệu cuối cùng dùng cho đánh giá định lượng gồm 26 phiên đo hợp lệ và được sử dụng để xây dựng các biểu đồ trong Hình 6.

Các điều kiện thử nghiệm bao gồm nhiều loại trang phục (áo phông, áo sơ mi, áo khoác dày), tư thế ngồi thẳng hoặc nghiêng nhẹ (15–30°), khoảng cách camera từ 0,6–1,2 m và các mức chiếu sáng khác nhau như phòng học (150–400 lux), phòng ký túc xá buổi tối (40–80 lux) và ánh sáng tự nhiên gần cửa sổ. Ground truth được xác định bằng phương pháp đếm thủ công bởi hai người quan sát độc lập sử dụng đồng hồ bấm giây, kết hợp tham chiếu từ Google Fit dựa trên camera điện thoại. Các phiên đo có sai khác lớn giữa hai người đếm hoặc Google Fit không trả kết quả được loại khỏi phân tích tương quan để bảo đảm tính nhất quán dữ liệu. Do mẫu thử nghiệm chủ yếu là sinh viên nam khỏe mạnh (20–23 tuổi), tập dữ liệu chưa đại diện đầy đủ cho các nhóm có đặc trưng hô hấp khác biệt như người cao tuổi hoặc bệnh nhân mắc bệnh lý hô hấp mãn tính (COPD, hen suyễn). Điều này có thể ảnh hưởng đến khả năng tổng quát hóa kết quả và được thừa nhận ở phần 3.4. Ngoài ra, do hạn chế thiết bị y tế chuyên dụng,

ground truth vẫn có thể mang tính chủ quan do phụ thuộc vào đếm thủ công và Google Fit (không chuyên biệt cho đo nhịp thở). Sai số tiềm ẩn từ nguồn tham chiếu này có thể ảnh hưởng đến đánh giá độ chính xác của hệ thống HealthCam; hạn chế này được nêu ở phần 3.4 và sẽ được cải thiện trong các nghiên cứu tiếp theo bằng thiết bị chuẩn như đai đo lồng ngực hoặc capnography.

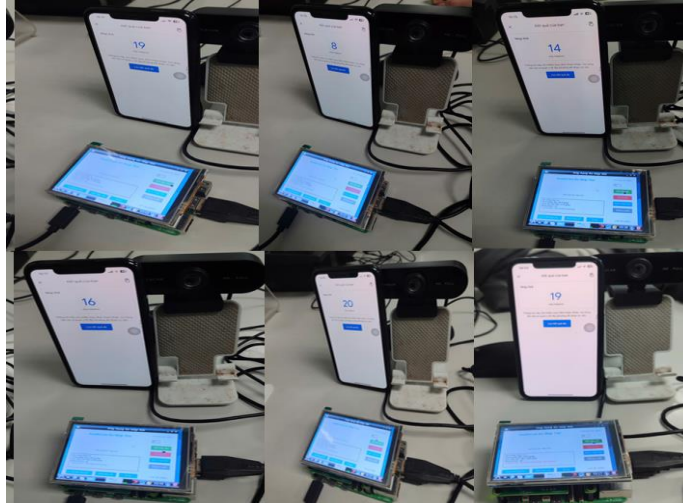
3.2. Kết quả chính

Tập dữ liệu phân tích cuối cùng gồm 26 phiên đo từ sinh viên Trường Đại học Công Nghiệp TP.HCM (độ tuổi chủ yếu 20–23, bao gồm cả nam và nữ). Thời lượng đo linh hoạt theo thiết kế ứng dụng (20, 25, 30, 35, 40, 45, 50, 55 hoặc tối đa 60 giây), với phần lớn phiên sử dụng 30 giây. Mỗi phiên thu thập đồng thời ba nguồn dữ liệu: Google Fit (tham chiếu gián tiếp qua camera điện thoại), hệ thống HealthCam (xử lý video không tiếp xúc), và đếm thủ công bởi nhóm nghiên cứu. Mẫu thử nghiệm có sự đa dạng cao về đặc điểm ngoại hình và điều kiện đo, với 4 trường hợp mặc áo dày/hoodie, 3 trường hợp đeo kính gây phản xạ ánh sáng nhẹ, và 2 trường hợp ngồi nghiêng khoảng 15–30° so với trục camera. Trong 26 phiên đo được sử dụng cho phân tích định lượng, hệ thống đạt tỷ lệ ROI hợp lệ 100%. Khi xét toàn bộ 50 phiên đo ban đầu, tỷ lệ phát hiện ROI trung bình đạt 92,3%, chủ yếu bị ảnh hưởng bởi các trường hợp quay người hoặc che khuất vùng ngực. Trong 50 phiên đo ban đầu, có 2 phiên Google Fit không trả kết quả do lỗi nhận diện khuôn mặt; các phiên này được loại khỏi phân tích tương quan nhưng vẫn được sử dụng để đánh giá độ ổn định ROI và tốc độ xử lý hệ thống. Biểu đồ tương quan giữa phương pháp đếm tay, HealthCam và Google Fit được trình bày tại Hình 6, kèm phân bố sai số của từng phương pháp. Hệ thống hoạt động ổn định với các chỉ số chính: tỷ lệ phát hiện ROI đạt 92,3%; tốc độ xử lý duy trì 27–32 FPS (độ phân giải 640×480) trên Raspberry Pi 5; độ trễ hiển thị BPM dưới 1,1 giây; 81% phiên đo có $IQR_{norm} < 0,5$ sau chuẩn hóa robust bằng median–MAD; thời gian đồng bộ dữ liệu lên Firebase trung bình 0,87 giây/gói. Khảo sát người dùng cho mức hài lòng trung bình là 4,8/5 điểm. Một số phiên đo đại diện được minh họa tại Hình 7.



Hình 6. Biểu đồ tương quan giữa đếm thủ công so với HealthCam và Google Fit, với phân bố sai số.

Hệ số tương quan Pearson được tính giữa giá trị BPM ước lượng từ HealthCam và phương pháp đếm thủ công trên tập dữ liệu 26 phiên đo hợp lệ.



Hình 7. Một số phiên đo thực tế trên sinh viên (áo dày, ánh sáng yếu, đeo kính, bình thường).

3.3. Thảo luận kết quả

Kết quả thực nghiệm cho thấy phương pháp đo nhịp thở không tiếp xúc của HealthCam, kết hợp chuyển dịch landmark và optical flow, đạt độ chính xác cao trong môi trường tự nhiên mà không cần cảm biến đeo, giảm thiểu can thiệp vào hành vi người dùng và phù hợp cho sử dụng hàng ngày (lớp học, ký túc xá, phòng trọ). Sai số thấp nhất khi đối tượng mặc trang phục mỏng, ngồi ổn định, ánh sáng đồng đều và vùng ngực quan sát rõ. Sai số tăng khi có áo dày che khuất, phản xạ kính nhiễu ROI, chuyển động toàn thân hoặc ánh sáng thấp (<35 lux) làm giảm chất lượng optical flow. Tuy nhiên, nhờ chuẩn hóa robust bằng median–MAD và cơ chế chọn nhánh phát hiện chu kỳ tối ưu dựa trên trên IQR_{norm} , hệ thống vẫn duy trì độ tin cậy cao và kiểm soát sai số hiệu quả trong phần lớn phiên đo. So với các phương pháp optical flow thuần túy hoặc rPPG phụ thuộc ánh sáng, phương pháp kết hợp landmark displacement và dense optical flow của HealthCam cho thấy khả năng chống nhiễu tốt hơn trước các dao động nhỏ của cơ thể. Tuy nhiên, khi xuất hiện chuyển động toàn thân mạnh hoặc ánh sáng dưới 35 lux, sai số có xu hướng tăng, cho thấy cần bổ sung các kỹ thuật ổn định hình ảnh trong các nghiên cứu tiếp theo.

3.4. Hạn chế của hệ thống

Mặc dù đạt được các kết quả tích cực và có tiềm năng ứng dụng rộng rãi, hệ thống vẫn tồn tại một số hạn chế. Đầu tiên, độ chính xác giảm đáng kể khi ánh sáng yếu hơn 35 lux, khiến optical flow suy giảm chất lượng và làm sai lệch biên độ dao động hô hấp. Thứ hai, các chuyển động mạnh của phần thân trên, đặc biệt là khi người đo cử động tay hoặc thay đổi tư thế nhanh, có thể gây mất ổn định tín hiệu và tạo ra sai số trong phát hiện chu kỳ. Thứ ba, số lượng mẫu thử nghiệm (50 phiên từ 40 tình nguyện viên) chưa đủ lớn và chủ yếu tập trung vào nhóm sinh viên trẻ khỏe mạnh (độ tuổi 20–23, chủ yếu nam). Điều này dẫn đến hạn chế trong khả năng

tổng quát hóa kết quả đối với các nhóm có đặc trưng hô hấp khác biệt, chẳng hạn như người cao tuổi (nhịp thở chậm/nhông hơn), bệnh nhân mắc bệnh lý hô hấp mãn tính (COPD, hen suyễn), hoặc trẻ em. Việc sử dụng ground truth dựa trên đếm thủ công và Google Fit thay vì thiết bị y tế chuyên dụng (như đai đo lồng ngực hoặc capnography) cũng góp phần vào sai số tiềm ẩn. Các hạn chế này sẽ được khắc phục trong các nghiên cứu tiếp theo bằng cách: (i) mở rộng mẫu thử nghiệm lên ít nhất 100–150 phiên với sự tham gia của người cao tuổi và bệnh nhân hô hấp thực tế; (ii) sử dụng thiết bị tham chiếu y tế chuẩn (đai đo lồng ngực hoặc capnography) để đánh giá độ chính xác lâm sàng; (iii) kiểm tra thêm trong môi trường thực địa đa dạng (phòng bệnh, xe buýt, khu vực đông người). Cuối cùng, hệ thống chưa được đánh giá trong các môi trường có nhiều nền động mạnh, chẳng hạn như khu vực đông người, phương tiện giao thông hoặc phòng tập thể dục, nơi mà chuyển động nền có thể ảnh hưởng đến thuật toán phát hiện ROI và optical flow.

4. KẾT LUẬN

Nghiên cứu trình bày một phương pháp đo nhịp thở không tiếp xúc dựa trên hình ảnh, kết hợp chuyển dịch landmark và optical flow, vận hành thời gian thực trên nền tảng nhúng chi phí thấp. Kết quả thực nghiệm cho thấy hệ thống đạt sai số trung bình 1,19 nhịp/phút (6,94%) và hệ số tương quan $r = 0,96$ so với phương pháp đếm thủ công, đồng thời duy trì tốc độ xử lý 27-32 FPS và độ trễ hiển thị dưới 1,1 giây. Kết quả này chứng minh tính khả thi của phương pháp cho các ứng dụng giám sát nhịp thở thụ động và các hệ thống IoT, thiết bị nhúng hoặc nền tảng giám sát không tiếp xúc trong nghiên cứu tín hiệu y sinh. Hệ thống HealthCam có thể triển khai tại gia đình, ký túc xá hoặc trạm y tế cơ sở, cũng như hỗ trợ các chương trình sàng lọc hô hấp cộng đồng hậu COVID-19. Trong tương lai, hệ thống có thể được mở rộng với chức năng đo nhịp tim, dự đoán nguy cơ bệnh lý bằng mô hình TensorFlow Lite và ứng dụng giám sát sức khỏe tài xế giao thông.

TÀI LIỆU THAM KHẢO

- [1]. C. Massaroni, A. Nicolò, D.L. Presti, M. Sacchetti, S. Silvestri, E. Schena, Contact-based methods for measuring respiratory rate, *Sensors*, 19 (2019) 908. <https://doi.org/10.3390/s19040908>
- [2]. M. Martinez, R. Stiefelhagen, Breath rate monitoring during sleep using near-IR imagery and PCA, *Proc. 19th ACM Int. Conf. Multimodal Interaction (ICMI)*, 2017, Glasgow, UK, 423-427. <https://doi.org/10.1145/3136755.3136804>
- [3]. L. Iozza, J. Lázaro, L. Cerina, D. Silvestri, L. Mainardi, P. Laguna, E. Gil, Monitoring breathing rate by fusing the physiological impact of respiration on video-photoplethysmogram with head movements, *Physiological Measurement*, 40 (2019) 094002. <https://doi.org/10.1088/1361-6579/ab4102>
- [4]. M. Mateu-Mateus, F. Guede-Fernández, N. Rodríguez-Ibáñez, M.A. García-González, J. Ramos-Castro, M. Fernández-Chimeno, A non-contact camera-based method for respiratory rhythm extraction, *Biomedical Signal Processing and Control*, 66 (2021) 102443. <https://doi.org/10.1016/j.bspc.2021.102443>
- [5]. R. Choudhary, B. Nigam, A systematic analysis: Remote monitoring of heart rate, blood pressure, respiratory rate, and oxygen saturation based on images and without physical touch, *Proc. International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*, 2022, 2349-2356. <https://doi.org/10.1109/ICAC3N56670.2022.10074569>

- [6]. P.S. Addison, A. Antunes, D. Montgomery, P. Smit, U.R. Borg, Robust Non-Contact Monitoring of Respiratory Rate using a Depth Camera, *Journal of Clinical Monitoring and Computing*, 37 (2023) 1003-1010. <https://doi.org/10.1007/s10877-023-01003-7>
- [7]. G. Scebba, G. Da Poian, W. Karlen, Multispectral Video Fusion for Non-Contact Monitoring of Respiratory Rate and Apnea, *IEEE Transactions on Biomedical Engineering*, 68 (2021) 350-359. <https://doi.org/10.1109/TBME.2020.2993649>
- [8]. D.B. MacLeod, P. Smit, A. Antunes, D. Montgomery, P.S. Addison, Non-contact (touchless) monitoring of respiratory rate in a challenging anesthesia setting using a depth camera, *Journal of Clinical Monitoring and Computing*, 40 (2025) 291-300. <https://doi.org/10.1007/s10877-025-01319-6>
- [9]. D. Shao, C. Liu, F. Tsow, Noncontact Physiological Measurement Using a Camera: A Technical Review and Future Directions, *ACS Sensors*, 6 (2021) 321-334. <https://doi.org/10.1021/acssensors.0c02042>
- [10]. D. McDuff, Camera Measurement of Physiological Vital Signs, *ACM Computing Surveys*, 55 (2023) 1-40. <https://doi.org/10.1145/3558518>
- [11]. B. Jalil, V. Lionetti, L. Valcarenghi, E-Health monitoring using camera: Measurement of vital parameters in a noisy environment, *Proc. IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, 2023, 32-37. <https://doi.org/10.1109/PerComWorkshops56833.2023.10150268>
- [12]. J. Wang, C. Shan, L. Liu, Z. Hou, Camera-based physiological measurement: Recent advances and future prospects, *Neurocomputing*, 575 (2024) 127282. <https://doi.org/10.1016/j.neucom.2024.127282>
- [13]. Raspberry Pi Foundation. Raspberry Pi 5 Product Brief. 2024. <https://pip.raspberrypi.com/documents/RP-008348-DS-3-raspberry-pi-5-product-brief.pdf>
- [14]. Waveshare Electronics. 3.5inch HDMI LCD User Manual. https://www.waveshare.com/wiki/3.5inch_HDMI_LCD
- [15]. Google Firebase. Understand Firebase Realtime Database Security Rules. <https://firebase.google.com/docs/database/security>
- [16]. Typebot. Typebot Documentation – Open-source conversational apps builder. Available: <https://docs.typebot.io/>
- [17]. Groq Inc. Model Card: Llama-3.3-70B-Versatile. <https://console.groq.com/docs/model/llama-3.3-70b-versatile>
- [18]. H. Nakamura, G. Lin, Y. Mitsukura, A Novel Approach for Pulse Signals Extraction and Heart Rate Estimation With Hilbert Space Orthogonalization Using Event-Based Camera, *IEEE Access*, 13 (2025) 112340-112353. <https://doi.org/10.1109/ACCESS.2025.3580578>