EVALUATION OF THE EFFECTIVENESS OF A MEDICAL KNOWLEDGE GRAPH COMBINED WITH MULTIMODAL ARTIFICIAL INTELLIGENCE IN SUPPORTING THE EARLY DIAGNOSIS OF NASOPHARYNGEAL CANCER IN VIETNAM

Pham Thi Bich Dao[1*], Pham Huy Tan[2], Khong Van Dong[3], Tran Van Tam[2], Doan Tien Luu[2], Le Duy Chung[2], Le Minh Dat[2], Bui Thi Mai[2], Nguyen Dieu My[1], Nguyen Thi Hang[2], Tong Xuan Thang[1], Pham Tran Anh[1], Pham Thi Bich Thuy[4], Nguyen Thi Anh Dao[2], Nguyen Xuan Hoa[2], Nguyen Thi Quynh Trang[5], Phan Xuan Nam[6], Hoang Huu Truong[7], Dinh Tuan Anh[4], Tran Hai Yen[1], Nguyen Thi Ngoc Anh[8], Hoang Minh Cuong[8], Nguyen Van Dang[1], Nguyen Thi Thanh Huong[9], Nguyen Thanh Thuy[10], Bui Quang Tuyen[11], Pham Thi Phuong Hanh[12]

[1]Hanoi Medical University - 1 Ton That Tung, Kim Lien Ward, Hanoi City, Vietnam
[2]Hanoi Medical University Hospital - 1 Ton That Tung, Kim Lien Ward, Hanoi City, Vietnam
[3]Viettel Business Solutions Corporation, Viettel Military Industry and Telecoms Group - No. 1 Tran Huu Duc, Tu Liem Ward, Hanoi City, Vietnam
[4]National Otorhinorarynology Hospital of Vietnam - 78 Giai Phong, Kim Lien Ward, Hanoi City, Vietnam
[5]University of Medicine and Pharmacy, Vietnam National University, Hanoi - 144 Xuan Thuy, Cau Giay Ward, Hanoi City, Vietnam
[6]Quang Tri Provincial General Hospital - 266 Hung Vuong, Nam Dong Ha Ward, Quang Tri Province, Vietnam
[7]Thanh Hoa Provincial General Hospital - 181 Hai Thuong Lan Ong, Hac Thanh Ward, Thanh Hoa Province, Vietnam
[8]Thai Nguyen University of Medicine And Pharmacy - 284 Luong Ngoc Quyen, Phan Dinh Phung Ward, Thai Nguyen Province, Vietnam
[9]Institute of Geodesy and Cartography - No. 479 Hoang Quoc Viet, Nghia Do Ward, Hanoi City, Vietnam
[10]Spatial Data Institute - No. 15, Lane 77 Bang Liet, Hoang Liet Ward, Hoang Mai Dist, Hanoi City, Vietnam
[11]Viettel Academy, Military Industry-Telecoms Group – Thach Hoa, Thach That, Hanoi City, Vietnam
[12]Ninh Binh Department of Health - No. 18 Kim Dong, Hoa Lu Ward, Ninh Binh Province, Vietnam

Received: 19/07/2026
Revised: 19/08/2026; Accepted: 04/09/2026

ABSTRACT

Objective: To evaluate the effectiveness of an artificial intelligence system integrated with a medical knowledge graph for the early detection of nasopharyngeal carcinoma.

Subjects and Methods: This multicenter study included 1,200 cases, comprising 600 patients with nasopharyngeal carcinoma and 600 controls, collected from five healthcare institutions representing different levels of care in Vietnam. The artificial intelligence system was developed using a multimodal framework integrating medical imaging and clinical data, combined with a medical knowledge graph. Model performance was evaluated on an independent test set using accuracy, sensitivity, specificity, and area under the ROC curve.

Results: The male-to-female ratio was 2.5:1 in the nasopharyngeal carcinoma group, compared with 1.2:1 in the control group ($p < 0.05$). On the independent test set of 180 cases, the system achieved an accuracy of 94.5% (95% CI: 92.1–96.8), sensitivity of 96.2%, specificity of 92.8%, and an AUC of 0.97. The correct detection rate for early-stage disease increased from 60.0% to 85.0%, while the miss rate for lesions smaller than 1 cm decreased from 42.0% to 12.0%. Compared with the image-only model, the multimodal model integrated with a medical knowledge graph improved diagnostic accuracy from 82.4% to 94.5%.

Conclusion: The artificial intelligence system integrated with a medical knowledge graph enabled early detection of nasopharyngeal carcinoma, reduced missed small lesions, and improved diagnostic accuracy.

Keywords: Nasopharyngeal carcinoma; artificial intelligence; medical knowledge graph; early diagnosis; multimodal integration.

*Corresponding author
Email: nguyenanhdao9991@gmail.com **Phone**: (+84) 975167076 **DOI:** 10.52163/yhc.v67iCD13.6408

ĐÁNH GIÁ HIỆU QUẢ ỨNG DỤNG ĐỒ THỊ TRI THỨC Y KHOA KẾT HỢP TRÍ TUỆ NHÂN TẠO ĐA PHƯƠNG THỨC HỖ TRỢ CHẨN ĐOÁN SỚM UNG THƯ VÒM MŨI HỌNG TẠI VIỆT NAM

Phạm Thị Bích Đào[1*], Phạm Huy Tần[2], Khổng Văn Đông[3], Trần Văn Tâm[2], Đoàn Tiến Lưu[2], Lê Duy Chung[2], Lê Minh Đạt[2], Bùi Thị Mai[2], Nguyễn Diệu My[1], Nguyễn Thị Hằng[2], Tống Xuân Thắng[1], Phạm Trần Anh[1], Phạm Thị Bích Thủy[4], Nguyễn Thị Anh Đào[2], Nguyễn Xuân Hòa[2], Nguyễn Thị Quỳnh Trang[5], Phan Xuân Nam[6], Hoàng Hữu Trường[7], Đinh Tuấn Anh[4], Trần Hải Yến[1], Nguyễn Thị Ngọc Anh[8], Hoàng Minh Cương[8], Nguyễn Văn Đăng[1], Nguyễn Thị Thanh Hương[9], Nguyễn Thanh Thuỷ[10], Bùi Quang Tuyến[11], Phạm Thị Phương Hạnh[12]

[1]Trường Đại học Y Hà Nội - 1 Tôn Thất Tùng, P. Kim Liên, Tp. Hà Nội, Việt Nam
[2]Bệnh viện Đại học Y Hà Nội - 1 Tôn Thất Tùng, P. Kim Liên, Tp. Hà Nội, Việt Nam
[3]Tổng công ty Giải pháp doanh nghiệp Viettel, Tập đoàn Công nghiệp Viễn thông Quân đội - Số 1 Trần Hữu Dực, P. Từ Liêm, Tp. Hà Nội, Việt Nam
[4]Bệnh viện Tai Mũi Họng Trung ương - 78 Giải Phóng, P. Kim Liên, Tp. Hà Nội, Việt Nam
[5]Trường Đại học Y Dược, Đại học Quốc gia Hà Nội - 144 Xuân Thủy, P. Cầu Giấy, Tp. Hà Nội, Việt Nam
[6]Bệnh viện Đa khoa Tỉnh Quảng Trị - 266 Hùng Vương, P. Nam Đông Hà, Tỉnh Quảng Trị, Việt Nam
[7]Bệnh viện Đa khoa Tỉnh Thanh Hóa - 181 Hải Thượng Lãn Ông, P. Hạc Thành, Tỉnh Thanh Hóa, Việt Nam
[8]Trường Đại học Y Dược Thái Nguyên, Đại học Thái Nguyên - 284 Lương Ngọc Quyến, P. Phan Đình Phùng, Tỉnh Thái Nguyên, Việt Nam
[9]Viện Khoa học Đo đạc và Bản đồ - Số 479 Hoàng Quốc Việt, P. Nghĩa Đô, Tp. Hà Nội, Việt Nam
[10]Viện Dữ liệu không gian - Số 15, ngõ 77 Bằng Liệt, P. Hoàng Liệt, Q. Hoàng Mai, Tp. Hà Nội, Việt Nam
[11]Học viện Viettel, Tập đoàn Công nghiệp Viễn thông Quân đội - Thạch Hòa, Thạch Thất, Tp. Hà Nội, Việt Nam
[12]Sở Y tế Tỉnh Ninh Bình - Số 18 Kim Đồng, P. Hoa Lư, Tỉnh Ninh Bình, Việt Nam

Ngày nhận: 19/07/2026
Ngày sửa: 19/08/2026; Ngày đăng: 04/09/2026

TÓM TẮT

Mục tiêu: Đánh giá hiệu quả chẩn đoán sớm của hệ thống trí tuệ nhân tạo tích hợp Đồ thị tri thức y khoa phát hiện ung thư vòm mũi họng.

Đối tượng và phương pháp: Nghiên cứu đa trung tâm, thực hiện trên 1.200 trường hợp (600 ung thư vòm mũi họng, 600 đối chứng). Hệ thống trí tuệ nhân tạo được xây dựng tích hợp đa phương thức: dữ liệu hình ảnh, lâm sàng, có tập kiểm tra độc lập, và có logic đánh giá mô hình.

Kết quả: Tỷ lệ nam/nữ ở nhóm bệnh là 2,5/1, cao hơn so với 1,2/1 ở nhóm đối chứng (p < 0,05). Trên tập kiểm tra độc lập gồm 180 trường hợp, hệ thống đạt độ chính xác 94,5% (KTC 95%: 92,1–96,8), độ nhạy 96,2%, độ đặc hiệu 92,8% và AUC 0,97. Tỷ lệ phát hiện đúng ung thư vòm mũi họng giai đoạn sớm tăng từ 60,0% lên 85,0%, trong khi tỷ lệ bỏ sót tổn thương nhỏ dưới 1 cm giảm từ 42,0% xuống 12,0%. So với mô hình chỉ sử dụng hình ảnh, mô hình tích hợp đa phương thức kết hợp Đồ thị tri thức y khoa giúp tăng độ chính xác từ 82,4% lên 94,5%.

Kết luận: Hệ thống trí tuệ nhân tạo tích hợp Đồ thị tri thức y khoa phát hiện sớm và giảm bỏ sót tổn thương nhỏ, đồng thời nâng cao độ chính xác.

Từ khóa: Ung thư vòm mũi họng; trí tuệ nhân tạo; Đồ thị tri thức y khoa; chẩn đoán sớm; đa phương thức.

1. ĐẶT VẤN ĐỀ

Ung thư vòm mũi họng (UTVMH) là một trong những ung thư đầu cổ có phân bố dịch tễ đặc thù, gặp nhiều tại Đông Nam Á và miền nam Trung Quốc, trong đó Việt Nam là quốc gia có gánh nặng bệnh tật đáng kể. Theo GLOBOCAN 2022, mỗi năm Việt Nam ghi nhận khoảng 5.613 ca mắc mới UTVMH, cho thấy đây vẫn là vấn đề y tế cần được quan tâm trong chiến lược phát hiện sớm và kiểm soát ung thư [1, 2].

Một thách thức lớn trong thực hành lâm sàng là phần lớn người bệnh UTVMH vẫn được chẩn đoán ở giai đoạn muộn, khi khối u đã xâm lấn rộng tại chỗ hoặc có di căn hạch cổ, làm giảm cơ hội điều trị triệt căn và ảnh hưởng rõ đến tiên lượng sống [3]. Nguyên nhân chủ yếu xuất phát từ đặc điểm giải phẫu của vòm mũi họng nằm sâu, khó tiếp cận khi khám thông thường, đồng thời triệu chứng giai đoạn sớm như nghẹt mũi một bên, ù tai, chảy máu mũi ít hoặc cảm giác vướng họng thường không đặc hiệu, dễ bị bỏ qua hoặc nhầm với các bệnh lý tai mũi họng lành tính [3].

Trong những năm gần đây, trí tuệ nhân tạo (Artificial Intelligence - AI), đặc biệt là các mô hình học sâu như mạng nơ-ron tích chập (Convolutional Neural Network - CNN), đã cho thấy tiềm năng lớn trong phân tích hình ảnh y học và hỗ trợ chẩn đoán ung thư. Các nghiên cứu bước đầu trên UTVMH cho thấy CNN có thể giúp phân biệt tổn thương ác tính với tổn thương lành tính trên cộng hưởng từ, đồng thời hỗ trợ nhận diện các đặc điểm hình ảnh khó đối với người đọc ít kinh nghiệm [4]. Tuy nhiên, đa số các mô hình hiện nay vẫn chủ yếu dựa trên một nguồn dữ liệu đơn lẻ, thường là hình ảnh, nên chưa phản ánh đầy đủ bản chất đa chiều của chẩn đoán UTVMH trong thực hành lâm sàng [4, 5].

*Tác giả liên hệ
Email: nguyenanhdao9991@gmail.com **Điện thoại**: (+84) 975167076 **DOI:** 10.52163/yhc.v67iCD13.6408

Trên thực tế, chẩn đoán UTVMH không chỉ dựa vào hình ảnh mà còn đòi hỏi tích hợp đồng thời nhiều lớp thông tin, bao gồm đặc điểm lâm sàng, vị trí và hình thái tổn thương, hạch cổ, giai đoạn bệnh [3]. Các yếu tố này có giá trị quan trọng trong định hướng nghi ngờ, phân tầng nguy cơ và củng cố quyết định chẩn đoán, nhưng chưa được tích hợp một cách hệ thống trong phần lớn các mô hình AI hiện có. Bên cạnh đó, hạn chế về khả năng giải thích của các mô hình "hộp đen" cũng làm giảm mức độ tin cậy và cản trở ứng dụng rộng rãi trong môi trường lâm sàng [5, 6, 7].

Đồ thị tri thức y khoa (Medical Knowledge Graph - MKG) được xem là một hướng tiếp cận có tiềm năng để giải quyết những hạn chế nêu trên, nhờ khả năng biểu diễn và liên kết tri thức giữa các thực thể lâm sàng, hình ảnh, xét nghiệm và chẩn đoán trong cùng một cấu trúc có ngữ nghĩa [5, 6]. Việc tích hợp AI đa phương thức với MKG không chỉ kỳ vọng nâng cao độ chính xác chẩn đoán mà còn tăng tính giải thích, giúp hệ thống đưa ra quyết định gần hơn với tư duy lâm sàng của bác sĩ [6, 7]. Trong bối cảnh UTVMH là bệnh lý dễ bỏ sót ở giai đoạn sớm và thường đòi hỏi phối hợp nhiều nguồn dữ liệu để xác định chẩn đoán, cách tiếp cận này đặc biệt có ý nghĩa.

Mặc dù vậy, đến nay vẫn còn thiếu các nghiên cứu đánh giá một cách hệ thống hiệu quả của mô hình trí tuệ nhân tạo tích hợp Đồ thị tri thức y khoa trong hỗ trợ chẩn đoán UTVMH, nhất là ở các khía cạnh có ý nghĩa thực tiễn cao như phát hiện giai đoạn sớm, giảm bỏ sót tổn thương nhỏ và cải thiện hiệu quả chẩn đoán khi tích hợp đa phương thức. Đây chính là khoảng trống nghiên cứu quan trọng cần được làm rõ để tạo cơ sở cho việc ứng dụng AI vào thực hành tai mũi họng - ung bướu tại Việt Nam [3, 4, 6].

Xuất phát từ những lý do trên, nghiên cứu này được thực hiện nhằm đánh giá hiệu quả chẩn đoán của hệ thống trí tuệ nhân tạo tích hợp Đồ thị tri thức y khoa trong hỗ trợ phát hiện ung thư vòm mũi họng, đồng thời phân tích giá trị của mô hình trong phát hiện sớm và giảm bỏ sót tổn thương nhỏ thông qua tích hợp dữ liệu đa phương thức [2, 4, 6].

2. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Đối tượng nghiên cứu

Nghiên cứu được thực hiện trên 1.200 trường hợp, thu thập từ các cơ sở y tế tham gia, gồm 600 bệnh nhân ung thư vòm mũi họng và 600 trường hợp đối chứng.

- *Tiêu chuẩn lựa chọn:* Các trường hợp được đưa vào nghiên cứu khi đáp ứng đầy đủ các tiêu chuẩn sau:

+ Chẩn đoán xác định thuộc một trong hai nhóm nghiên cứu: ung thư vòm mũi họng hoặc tổn thương lành tính vùng vòm, có đầy đủ thông tin.

+ Có ít nhất một loại dữ liệu hình ảnh y học đạt tiêu chuẩn phân tích, gồm cộng hưởng từ (MRI), cắt lớp vi tính (CT) hoặc nội soi vòm mũi họng.

- *Tiêu chuẩn loại trừ*: Các trường hợp bị loại khỏi nghiên cứu khi có một trong các đặc điểm sau:

+ Chất lượng hình ảnh không đạt yêu cầu phân tích

+ Đã được điều trị trước đó bằng phẫu thuật, xạ trị, hóa trị hoặc can thiệp khác.

+ Hồ sơ trùng lặp, không thống nhất thông tin, không đủ căn cứ xác định chẩn đoán.

2.2. Phương pháp nghiên cứu và cỡ mẫu

Nghiên cứu được thiết kế theo phương pháp mô tả cắt ngang có phân tích, kết hợp đánh giá hiệu quả chẩn đoán của hệ thống trí tuệ nhân tạo.

Mẫu nghiên cứu được chọn theo phương pháp thuận tiện có chủ đích: toàn bộ các trường hợp đủ tiêu chuẩn lựa chọn, được ghi nhận liên tiếp tại 5 cơ sở y tế tham gia trong thời gian nghiên cứu, sau đó bắt cặp theo tỷ lệ 1:1 giữa nhóm ung thư vòm mũi họng và nhóm đối chứng nhằm hạn chế sai số do mất cân bằng cỡ mẫu giữa hai nhóm khi huấn luyện và đánh giá mô hình. Cỡ mẫu tối thiểu cho mỗi nhóm được ước tính theo công thức tính cỡ mẫu cho nghiên cứu giá trị của một xét nghiệm chẩn đoán, dựa trên độ nhạy kỳ vọng 90%, sai số cho phép d = 5% và độ tin cậy 95% (Z = 1,96), cho kết quả tối thiểu khoảng 138 trường hợp mỗi nhóm; cỡ mẫu thực tế 600 trường hợp/nhóm vượt yêu cầu tối thiểu, giúp tăng độ chính xác của các ước lượng hiệu năng mô hình và cho phép chia thành các tập huấn luyện, hiệu chỉnh, kiểm tra độc lập như mô tả ở mục 2.3.1. Cỡ mẫu nghiên cứu gồm 1.200 trường hợp, được phân bố cân bằng giữa hai nhóm.

2.3. Thiết kế nghiên cứu và chỉ số đánh giá

2.3.1. Thiết kế nghiên cứu

Hệ thống trí tuệ nhân tạo trong nghiên cứu được xây dựng theo hướng tích hợp đa phương thức, sử dụng hai nhóm dữ liệu chính gồm:

- Dữ liệu hình ảnh y học: MRI, CT và/hoặc nội soi vòm mũi họng;

- Dữ liệu lâm sàng: tuổi, giới và các thông tin hỗ trợ chẩn đoán.

Bộ dữ liệu được chuẩn hóa và chia thành các tập huấn luyện, hiệu chỉnh và kiểm tra độc lập theo tỷ lệ 70% - 15% - 15%, tương ứng với 840 trường hợp cho tập huấn luyện, 180 trường hợp cho tập hiệu chỉnh và 180 trường hợp cho tập kiểm tra độc lập, bảo đảm phân bố cân bằng giữa nhóm ung thư vòm mũi họng và nhóm đối chứng. Việc chia dữ liệu được thực hiện theo phương pháp ngẫu nhiên phân tầng/theo trung tâm, nhằm hạn chế sai lệch chọn mẫu và giảm nguy cơ rò rỉ dữ liệu giữa các tập. Dữ liệu hình ảnh và lâm sàng được thu thập hồi cứu từ hồ sơ bệnh án điện tử và hệ thống lưu trữ hình ảnh (PACS) của 5 cơ sở y tế tham gia. Chẩn đoán ung thư vòm mũi họng được xác định theo chuẩn vàng là kết quả mô bệnh học dương tính từ sinh thiết vòm; nhóm đối chứng được xác định khi mô bệnh học âm tính hoặc hình ảnh nội soi/MRI/CT bình thường kèm theo dõi lâm sàng tối thiểu 12 tháng không ghi nhận tổn thương ác tính. Các trường hợp thiếu dữ liệu ở biến đầu vào (ví dụ thiếu một trong các phương tiện chẩn đoán hình ảnh) được loại khỏi phân tích theo tiêu chuẩn loại trừ ở mục 2.1.2, không áp dụng suy diễn (imputation) để tránh sai lệch khi đánh giá hiệu năng mô hình. Dữ liệu sau khi ẩn danh được lưu trữ, xử lý trên hệ thống nội bộ có phân quyền truy cập và mã hóa khi lưu trữ/truyền tải, phù hợp với quy định về bảo mật thông tin y tế và đạo đức nghiên cứu nêu ở mục 2.7. Để đánh giá giá trị gia tăng của từng mức tích hợp dữ liệu, nghiên cứu tiến hành so sánh giữa các mô hình sau:

(1) mô hình chỉ sử dụng dữ liệu hình ảnh;

(2) mô hình đa phương thức tích hợp hình ảnh và dữ liệu lâm sàng;

(3) mô hình đa phương thức tích hợp hình ảnh, dữ liệu lâm sàng và Đồ thị tri thức y khoa.

Thiết kế này cho phép đánh giá không chỉ hiệu quả chẩn đoán chung của hệ thống AI mà còn làm rõ đóng góp riêng của tích hợp đa phương thức và của Đồ thị tri thức y khoa trong hỗ trợ chẩn đoán sớm ung thư vòm mũi họng.

2.3.2. Các chỉ số đánh giá

2.3.2.1. Các chỉ số đánh giá chính

Độ chính xác (Accuracy); Độ nhạy (Sensitivity); Độ đặc hiệu (Specificity); Diện tích dưới đường cong ROC (AUC).

2.3.2.2. Các chỉ số đánh giá bổ sung

Tỷ lệ phát hiện đúng ung thư vòm mũi họng giai đoạn sớm (giai đoạn I–II); Tỷ lệ bỏ sót tổn thương nhỏ (có kích thước < 1 cm); Mức cải thiện hiệu quả chẩn đoán khi tích hợp đa

phương thức, so sánh giữa các mô hình: chỉ sử dụng dữ liệu hình ảnh; mô hình đa phương thức tích hợp hình ảnh, lâm sàng và Đồ thị tri thức y khoa.

2.4. Địa điểm nghiên cứu

Dữ liệu nghiên cứu được thu thập từ 5 cơ sở y tế đại diện cho các tuyến chuyên môn khác nhau gồm: Bệnh viện Đại học Y Hà Nội, bệnh viện Tai mũi họng Trung ương, bệnh viện Đa khoa tỉnh Thanh Hóa, bệnh viện đa khoa tỉnh Quảng Trị, bệnh viện Trung ương Thái Nguyên.

2.5. Các biến số nghiên cứu

2.5.1. Nhóm biến đầu vào

- Đặc điểm nhân khẩu học: Tuổi; Giới.

- Dữ liệu hình ảnh: Loại phương tiện chẩn đoán hình ảnh sử dụng (MRI, CT, nội soi); Vị trí tổn thương; Kích thước tổn thương; Đặc điểm tổn thương nhỏ dưới 1 cm;Sự hiện diện của hạch nghi ngờ; Các dấu hiệu hình ảnh gợi ý ác tính.

2.5.2. Nhóm biến đầu ra và biến đánh giá

- Chẩn đoán cuối cùng: ung thư vòm mũi họng hoặc tổn thương lành tính;

- Kết quả dự đoán của hệ thống trí tuệ nhân tạo;

- Mức độ phù hợp giữa dự đoán của AI và chẩn đoán xác định;

- Giai đoạn bệnh, phân thành giai đoạn sớm (I–II) và giai đoạn khác;

- Tình trạng tổn thương nhỏ < 1 cm;

Nhóm mô hình phân tích: chỉ hình ảnh; mô hình tích hợp đa phương thức với Đồ thị tri thức y khoa.

2.6. Phân tích và xử lý số liệu

Dữ liệu được chuẩn hóa trước khi đưa vào phân tích và huấn luyện mô hình.

Các biến định tính được trình bày dưới dạng tần số và tỷ lệ phần trăm. Hiệu suất chẩn đoán của hệ thống AI được tính toán thông qua các chỉ số: độ chính xác; độ nhạy; độ đặc hiệu; diện tích dưới đường cong ROC.

Ngoài ra, nghiên cứu thực hiện: so sánh tỷ lệ phát hiện đúng giai đoạn sớm giữa các mô hình; so sánh tỷ lệ bỏ sót tổn thương nhỏ; so sánh hiệu quả của từng phương thức phân tích trong hệ thống AI. Các phép kiểm thống kê phù hợp được sử dụng để so sánh giữa các nhóm, với mức ý nghĩa thống kê $p < 0,05$.

Sai số nghiên cứu được hạn chế bằng cách:áp dụng thống nhất tiêu chuẩn lựa chọn và loại trừ; sử dụng tiêu chuẩn chẩn đoán xác định giống nhau cho toàn bộ mẫu; chuẩn hóa quy trình thu nhận và xử lý dữ liệu đầu vào; loại bỏ các trường hợp có dữ liệu không đầy đủ hoặc chất lượng hình ảnh không đạt yêu cầu; sử dụng tập kiểm tra độc lập để đánh giá khách quan hiệu năng của mô hình.

2.7. Đạo đức nghiên cứu

Nghiên cứu được thực hiện trên bộ dữ liệu đã được ẩn danh, không sử dụng thông tin nhận diện cá nhân của người bệnh. Việc thu thập, lưu trữ và xử lý dữ liệu tuân thủ các quy định về bảo mật thông tin và đạo đức trong nghiên cứu y sinh học.

3. KẾT QUẢ NGHIÊN CỨU

3.1. Đặc điểm mẫu nghiên cứu

Bảng 1. Đặc điểm chung của đối tượng nghiên cứu (n = 1.200)

Đặc điểm	Nhóm ung thư vòm mũi họng (n = 600)	Nhóm đối chứng (n = 600)	p
Nam/Nữ	2,5 / 1	1,2 / 1	< 0,05
Tuổi trung bình (năm)	48,6 ± 12,4	45,2 ± 11,8	> 0,05

Nhận xét: Tỷ lệ nam/nữ ở nhóm ung thư vòm mũi họng là 2,5/1, cao hơn so với nhóm đối chứng là 1,2/1 ($p < 0,05$). Tuổi trung bình ở hai nhóm lần lượt là 48,6 ± 12,4 và 45,2 ± 11,8, không có sự khác biệt có ý nghĩa thống kê ($p > 0,05$).

3.2. Hiệu quả chẩn đoán và giá trị ứng dụng của hệ thống trí tuệ nhân tạo

3.2.1. Hiệu suất chẩn đoán của hệ thống trí tuệ nhân tạo

Bảng 2. Hiệu suất chẩn đoán của hệ thống AI trên tập kiểm tra độc lập (n = 180)

Chỉ số	Giá trị	Khoảng tin cậy 95%
Độ chính xác (Accuracy)	94,5%	92,1 – 96,8
Độ nhạy (Sensitivity)	96,2%	94,0 – 98,4
Độ đặc hiệu (Specificity)	92,8%	90,5 – 95,1
Diện tích dưới đường cong (Area Under Curve)	0,97	0,95 – 0,99
Giá trị dự báo dương tính (PPV)*	92,6%	85,4 – 96,4
Giá trị dự báo âm tính (NPV)*	96,5%	90,2 – 98,8

Nhận xét: Trên tập kiểm tra độc lập gồm 180 trường hợp (phân bố cân bằng 90 ca ung thư vòm mũi họng và 90 ca đối chứng), hệ thống đạt độ chính xác 94,5% (KTC 95%: 92,1–96,8), độ nhạy 96,2% (KTC 95%: 94,0–98,4), độ đặc hiệu 92,8% (KTC 95%: 90,5–95,1) và AUC 0,97 (KTC 95%: 0,95–0,99). Giá trị dự báo dương tính và âm tính lần lượt là 92,6% (KTC 95%: 85,4–96,4) và 96,5% (KTC 95%: 90,2–98,8);

3.2.2. Hiệu quả phát hiện ung thư vòm mũi họng giai đoạn sớm

Bảng 3. So sánh khả năng phát hiện đúng ung thư vòm mũi họng giai đoạn sớm (I–II) trên nhóm bệnh giai đoạn sớm (N = 120)

Phương pháp	Số phát hiện đúng / tổng số	Tỷ lệ phát hiện đúng (%)	p
Lâm sàng	72/120	60,0	
AI + MKG	102/120	85,0	< 0,001

Nhận xét: Trên cùng một nhóm 120 trường hợp ung thư vòm mũi họng giai đoạn sớm, tỷ lệ phát hiện đúng khi chỉ dựa vào thăm khám lâm sàng thông thường (chưa có hỗ trợ AI) là 60,0%, trong khi tỷ lệ phát hiện đúng khi bác sĩ có hỗ trợ của hệ thống AI tích hợp MKG (đọc dựa trên hình ảnh, dữ liệu lâm sàng và MKG) đạt 85,0%, cao hơn 25,0 điểm phần trăm, khác biệt có ý nghĩa thống kê ($p < 0,001$.

3.2.3. Khả năng giảm bỏ sót tổn thương nhỏ

Bảng 4. So sánh tỷ lệ bỏ sót tổn thương nhỏ dưới 1 cm giữa bác sĩ ít kinh nghiệm và bác sĩ có hỗ trợ AI (N = 100)

Nhóm đánh giá	Số tổn thương bỏ sót / tổng số	Tỷ lệ bỏ sót (%)	p
Bác sĩ có kinh nghiệm < 5 năm	42/100	42,0	
Bác sĩ có AI hỗ trợ	12/100	12,0	< 0,001

Nhận xét: Trên 100 tổn thương nhỏ dưới 1 cm, tỷ lệ bỏ sót giảm từ 42,0% ở nhóm bác sĩ có kinh nghiệm dưới 5 năm xuống còn 12,0% khi có hỗ trợ từ hệ thống AI, tương ứng giảm 30,0 điểm phần trăm ($p < 0,001$).

3.2.4. Vai trò của tích hợp đa phương thức

Bảng 5. Hiệu quả theo từng mức độ tích hợp dữ liệu

Phương thức phân tích	Độ chính xác (%)	Giá trị p
Chỉ sử dụng hình ảnh	82,4%	
Đa phương thức (MKG)	94,5%	< 0,001

Nhận xét: Độ chính xác khi sử dụng hình ảnh 82,4%, đạt 94,5% khi tích hợp thêm dữ liệu lâm sàng, tương ứng tăng 12,1 phần trăm so với mô hình chỉ sử dụng hình ảnh.

4. BÀN LUẬN

Nghiên cứu cho thấy hệ thống trí tuệ nhân tạo đa phương thức tích hợp MKG đạt hiệu quả chẩn đoán cao trong hỗ trợ phát hiện ung thư vòm mũi họng, với độ chính xác 94,5%, độ nhạy 96,2%, độ đặc hiệu 92,8% và diện tích dưới đường cong đạt 0,97 trên tập kiểm tra độc lập. Kết quả này có ý nghĩa thực tiễn, bởi ung thư vòm mũi họng là bệnh lý ác tính có gánh nặng đáng kể tại Việt Nam, trong khi chẩn đoán sớm còn gặp nhiều khó khăn do vị trí giải phẫu của vòm mũi họng nằm sâu, khó tiếp cận khi khám thông thường, đồng thời triệu chứng ban đầu thường nghèo nàn, không đặc hiệu và dễ nhầm với các bệnh lý tai mũi họng lành tính. Mục tiêu ưu tiên không chỉ là phân loại đúng mà còn là giảm tối đa nguy cơ bỏ sót bệnh ngay từ lần tiếp cận đầu tiên [1, 2, 3, 4, 5].

Trong nghiên cứu, nam giới chiếm ưu thế (nam/nữ là 2,5/1), cao hơn rõ rệt so với nhóm đối chứng. Kết quả này phù hợp với đặc điểm dịch tễ của ung thư vòm mũi họng đã được ghi nhận ở các quốc gia có tỷ lệ lưu hành cao tại Đông Nam Á. Theo GLOBOCAN 2022, Việt Nam ghi nhận 5.613 ca mắc mới ung thư vòm mũi họng trong năm 2022, cho thấy đây vẫn là bệnh lý cần được ưu tiên trong chiến lược phát hiện sớm và kiểm soát ung thư. Điều này làm tăng tính phù hợp lâm sàng của hệ thống khi xem xét khả năng ứng dụng tại các cơ sở tai mũi họng và ung bướu trong nước [1, 2].

Một kết quả rất đáng chú ý của nghiên cứu là tỷ lệ phát hiện đúng ung thư vòm mũi họng giai đoạn sớm tăng từ 60,0% khi chỉ dựa vào thăm khám lâm sàng lên 85,0% khi có sự hỗ trợ của hệ thống trí tuệ nhân tạo kết hợp MKG. Giá trị lớn nhất của trí tuệ nhân tạo trong ung thư vòm mũi họng là khả năng hỗ trợ nhận diện bệnh sớm hơn, khi cơ hội điều trị triệt căn và tiên lượng sống còn tốt. Chen và cộng sự cho thấy mô hình học máy dựa trên dữ liệu hồ sơ khám chữa bệnh có thể hỗ trợ dự báo nguy cơ xuất hiện ung thư vòm mũi họng và rút ngắn giai đoạn tiền chẩn đoán [5]. Trong lĩnh vực nội soi, Xu và cộng sự ghi nhận việc kết hợp đồng thời hình ảnh ánh sáng trắng và nội soi dải ánh sáng hẹp làm tăng rõ hiệu quả nhận diện tổn thương so với cách tiếp cận chỉ sử dụng từng nguồn dữ liệu riêng lẻ [6]. Wang và cộng sự chi ra hệ thống trí tuệ nhân tạo dựa trên ảnh nội soi vòm có khả năng nhận diện ung thư vòm mũi họng với độ chính xác cao [7]. Các công trình gần đây của He, Yue và đặc biệt là nghiên cứu đa trung tâm quy mô quốc gia của Shi và cộng sự đều củng cố nhận định rằng trí tuệ nhân tạo có thể làm tăng năng lực phát hiện sớm tổn thương nghi ngờ, nhất là trong môi trường có sự chênh lệch về kinh nghiệm giữa những người đọc [8, 9, 10].

Kết quả giảm tỷ lệ bỏ sót tổn thương nhỏ dưới 1 cm từ 42,0% xuống 12,0% khi có sự hỗ trợ của hệ thống trí tuệ nhân tạo là một bằng chứng có giá trị ứng dụng rất cao. Trong thực hành lâm sàng, các tổn thương giai đoạn sớm ở vòm mũi họng thường nhỏ, kín đáo, có thể nằm ở hố Rosenmüller, dưới niêm mạc hoặc không tạo hình ảnh lồi sùi rõ rệt, do đó rất dễ bị bỏ qua, đặc biệt bởi bác sĩ ít kinh nghiệm. Wang và cộng sự từng nhấn mạnh rằng chẩn đoán ung thư vòm mũi họng là một thách thức, do có một tỷ lệ đáng kể tổn thương khó quan sát trên nội soi và thậm chí có thể bị bỏ sót ở lần đánh giá đầu tiên [4]. Vì vậy, một hệ thống có khả năng gợi ý vùng nghi ngờ, nhấn mạnh vị trí tổn thương và làm giảm tỷ lệ bỏ sót các tổn thương nhỏ không chỉ có ý nghĩa về mặt thống kê mà còn có tác động trực tiếp đến quyết định sinh thiết, chỉ định chụp cắt lớp vi tính hoặc cộng hưởng từ bổ sung và định hướng chuyển tuyến kịp thời. Ý nghĩa này càng rõ hơn trong bối cảnh tuyến đầu hoặc cơ sở không chuyên sâu, nơi kinh nghiệm đọc nội soi vòm có thể chưa đồng đều giữa các bác sĩ [3, 4, 8, 9].

Giá trị trung tâm của nghiên cứu nằm ở chỗ mô hình không dừng lại ở việc phân tích hình ảnh đơn thuần, mà cho thấy khi tích hợp thêm dữ liệu lâm sàng và MKG, độ chính xác tăng từ 82,4% lên 94,5%, tương ứng tăng 12,1 điểm phần trăm. Mức tăng này cho thấy bản chất của chẩn đoán ung thư vòm mũi họng là một quá trình suy luận đa chiều, trong đó hình ảnh chỉ là một thành phần. Trong thực hành thật, bác sĩ không đưa ra kết luận chỉ từ một ảnh nội soi hay một lát cắt cộng hưởng từ riêng lẻ, mà luôn kết hợp tuổi, giới, triệu chứng, vị trí tổn thương, hình thái niêm mạc, tình trạng hạch cổ và bối cảnh nguy cơ. Vì vậy, kết quả của nghiên cứu phù hợp với tư duy lâm sàng hơn so với mô hình chỉ dựa trên hình ảnh. Các nghiên cứu gần đây về học máy kết hợp nhiều nguồn dữ liệu trong y học cũng cho thấy mô hình tích hợp thường cho hiệu quả tốt hơn mô hình chỉ sử dụng một loại dữ liệu đơn lẻ [11]. Bên cạnh đó, các tổng quan gần đây về ứng dụng Đồ thị tri thức trong y học và trong phân tích hình ảnh y học đều cho thấy việc mã hóa quan hệ ngữ nghĩa giữa triệu chứng, dấu hiệu hình ảnh, xét nghiệm và chẩn đoán có thể giúp tăng năng lực suy luận và cải thiện khả năng giải thích của hệ thống [12, 13]. Về mặt triển khai, đây là ưu thế rất quan trọng, bởi trong môi trường y tế, một mô hình càng gần với cách lập luận của bác sĩ thì càng dễ được chấp nhận và ứng dụng [14].

So với các nghiên cứu trí tuệ nhân tạo trước đây trên ung thư vòm mũi họng, điểm khác biệt của công trình này là đã chuyển trọng tâm từ việc chứng minh hệ thống "đọc ảnh tốt" sang việc chứng minh hệ thống tích hợp tri thức có thể hỗ trợ quyết định sớm và có giá trị thực hành rõ hơn. Các mô hình nội soi trước đây như của Xu, Wang, He hay Yue chủ yếu tập trung tối ưu hóa khả năng phân biệt hình ảnh ác tính và lành tính [6, 7, 8, 9]. Nghiên cứu đa trung tâm của Shi đã tiến thêm một bước khi đánh giá khả năng hỗ trợ bác sĩ trên dữ liệu nội soi thực tế [10]. Tuy nhiên, nghiên cứu hiện tại bổ sung thêm một tầng giá trị mới, đó là sử dụng MKG như cơ chế liên kết giữa dữ liệu hình ảnh và dữ liệu lâm sàng, từ đó phản ánh tốt hơn quy trình tư duy chẩn đoán trong thực hành tai mũi họng - ung bướu. Nói cách khác, hệ thống trong nghiên cứu này không chỉ đóng vai trò phân loại hình ảnh, mà còn là một công cụ hỗ trợ chẩn đoán có khả năng gom nối và diễn giải thông tin theo hướng gần với suy luận lâm sàng hơn [11, 12, 13].

Kết quả nghiên cứu cũng cho thấy mô hình có tiềm năng triển khai trong hệ thống y tế Việt Nam khi dữ liệu được xây dựng từ nhiều tuyến chuyên môn khác nhau. Việc tích hợp dữ liệu hình ảnh, thông tin lâm sàng và MKG không chỉ giúp nâng cao hiệu quả chẩn đoán mà còn góp phần chuẩn hóa quy trình hỗ trợ ra quyết định giữa các cơ sở y tế. Đặc biệt, tại các bệnh viện tuyến tỉnh hoặc cơ sở chưa có nhiều chuyên gia về ung thư đầu cổ, hệ thống AI tích hợp MKG có thể hỗ trợ nhận diện sớm các trường hợp nghi ngờ, góp phần giảm sự khác biệt về chất lượng chẩn đoán giữa các tuyến và hỗ trợ chuyển tuyến kịp thời.

Từ góc nhìn quản lý hệ thống y tế tại Việt Nam, kết quả nghiên cứu cho thấy mô hình có tiềm năng triển khai trong thực hành lâm sàng nhờ được xây dựng và đánh giá

trên dữ liệu thu thập từ năm cơ sở y tế đại diện cho các tuyến chuyên môn khác nhau. Việc tích hợp dữ liệu hình ảnh, thông tin lâm sàng và MKG (MKG) không chỉ nâng cao hiệu quả chẩn đoán mà còn góp phần chuẩn hóa quá trình hỗ trợ ra quyết định giữa các cơ sở y tế, đặc biệt tại những bệnh viện chưa có nhiều chuyên gia về ung thư đầu cổ. Điều này có thể giúp giảm sự khác biệt về chất lượng chẩn đoán giữa các tuyến, hỗ trợ phát hiện sớm và chuyển tuyến kịp thời đối với các trường hợp nghi ngờ ung thư vòm mũi họng. Trong bối cảnh ung thư vòm mũi họng vẫn là bệnh lý có gánh nặng đáng kể tại Việt Nam và khu vực Đông Nam Á, việc ứng dụng một hệ thống AI tích hợp MKG có tiềm năng nâng cao chất lượng chẩn đoán sớm và tối ưu hóa sử dụng nguồn lực y tế [1, 2, 3].

Tuy vậy, vẫn cần nhìn nhận thận trọng một số hạn chế. Nghiên cứu hiện mới được đánh giá trên bộ dữ liệu của các cơ sở tham gia, chưa có ngoại kiểm (external validation) trên một quần thể hoàn toàn độc lập, do đó các kết luận về giá trị hỗ trợ chẩn đoán sớm cần được diễn giải ở mức thận trọng và mang tính gợi ý hơn là khẳng định, cho đến khi có nghiên cứu kiểm định ngoài xác nhận. Bên cạnh đó, thiết kế bệnh-chứng với tỷ lệ cân bằng 600 bệnh nhân ung thư vòm mũi họng/600 đối chứng giúp tối ưu việc huấn luyện và đánh giá mô hình, nhưng không phản ánh tỷ lệ hiện mắc thực tế của bệnh trong quần thể sàng lọc, vốn thấp hơn nhiều so với tỷ lệ 1:1 trong nghiên cứu. Vì vậy, độ chính xác, PPV và NPV báo cáo ở mục 3.2.1 phản ánh hiệu năng của mô hình trên phổ bệnh cân bằng của nghiên cứu, cần được hiệu chỉnh lại theo tỷ lệ hiện mắc thực tế trước khi ngoại suy trực tiếp cho quần thể sàng lọc cộng đồng hoặc tuyến khám ban đầu, nơi tỷ lệ ung thư vòm mũi họng thấp hơn đáng kể so với thiết kế của nghiên cứu này. Các biến lâm sàng tích hợp vào mô hình hiện còn tương đối cơ bản; nếu bổ sung thêm các dấu ấn sinh học như vi rút Epstein-Barr trong máu, mô tả nội soi có cấu trúc, đặc trưng cộng hưởng từ chi tiết hoặc dữ liệu bệnh án dạng văn bản tự do, hiệu quả của mô hình có thể còn được cải thiện thêm. Mặc dù MKG được kỳ vọng làm tăng khả năng giải thích, nghiên cứu hiện chưa định lượng riêng mức độ giải thích của hệ thống hoặc mức độ cải thiện niềm tin của người dùng cuối, đây sẽ là hướng cần tiếp tục đánh giá trong các nghiên cứu tiếp theo [11, 12, 13, 14].

5. KẾT LUẬN

Hệ thống trí tuệ nhân tạo tích hợp MKG đạt hiệu quả chẩn đoán cao với độ chính xác 94,5%, độ nhạy 96,2%, độ đặc hiệu 92,8% và AUC 0,97. Việc tích hợp đa phương thức giúp tăng tỷ lệ phát hiện đúng ung thư vòm mũi họng giai đoạn sớm từ 60,0% lên 85,0% và giảm tỷ lệ bỏ sót tổn thương nhỏ từ 42,0% xuống 12,0%, qua đó trả lời mục tiêu nghiên cứu về hiệu quả chẩn đoán sớm của hệ thống trên tập dữ liệu và thiết kế của nghiên cứu này.

6. KIẾN NGHỊ

Cần tiếp tục thực hiện các nghiên cứu kiểm định ngoài (external validation) trên quần thể độc lập và tại nhiều cơ sở y tế hơn nhằm đánh giá khả năng khái quát của mô hình. Đồng thời, cần triển khai các nghiên cứu tiền cứu, có hiệu chỉnh theo tỷ lệ hiện mắc thực tế của quần thể đích, để đánh giá tính khả thi và giá trị sử dụng của hệ thống trí tuệ nhân tạo tích hợp MKG trong hỗ trợ chẩn đoán sớm ung thư vòm mũi họng trước khi xem xét tích hợp vào hệ thống HIS/PACS phục vụ thực hành lâm sàng thường quy.

TÀI LIỆU THAM KHẢO

[1] International Agency for Research on Cancer. Viet Nam fact sheet. GLOBOCAN 2022, version 1.1; 2024.

[2] Jiang W, Zheng B, Wei H. Recent advances in early detection of nasopharyngeal carcinoma. Discov Oncol. 2024;15:365. doi:10.1007/s12672-024-01242-3.

[3] Yuan Y, Ye F, Wu JH, et al. Early screening of nasopharyngeal carcinoma. Head Neck. 2023;45(10):2700-2709. doi:10.1002/hed.27466.

[4] Wang KH, Austin SA, Chen SH, Sonne DC, Gurushanthaiah D. Nasopharyngeal carcinoma diagnostic challenge in a nonendemic setting: our experience with 101 patients. Perm J. 2017;21:16-180. doi:10.7812/TPP/16-180.

[5] Chen JW, Lin ST, Lin YC, Wang BS, Chien YN, Chiou HY. Early detection of nasopharyngeal carcinoma through machine-learning-driven prediction model in a population-based healthcare record database. Cancer Med. 2024;13(7):e7144. doi:10.1002/cam4.7144.

[6] Xu J, Wang J, Bian X, Zhu JQ, Tie CW, Liu X, et al. Deep learning for nasopharyngeal carcinoma identification using both white light and narrow-band imaging endoscopy. Laryngoscope. 2022;132(5):999-1007. doi:10.1002/lary.29894.

[7] Wang SX, et al. The detection of nasopharyngeal carcinomas using a deep convolutional neural network-based artificial intelligence system. Laryngoscope. 2024;134(1):127-135. doi:10.1002/lary.30781.

[8] He Z, Zhang K, Zhao N, et al. Deep learning for real-time detection of nasopharyngeal carcinoma during nasopharyngeal endoscopy. iScience. 2023;26:107463. doi:10.1016/j.isci.2023.107463.

[9] Yue Y, Zeng X, Lin H, et al. A deep learning based smartphone application for early detection of nasopharyngeal carcinoma using endoscopic images. npj Digit Med. 2024;7:384. doi:10.1038/s41746-024-01403-2.

[10] Shi Y, et al. Artificial intelligence-assisted detection of nasopharyngeal carcinoma on endoscopic images: a national, multicentre, model development and validation study. Lancet Digit Health. 2025;7(6):100869. doi:10.1016/j.landig.2025.03.001.

[11] Zhou HY, Yu Y, Wang C, et al. A transformer-based representation-learning model with unified processing of multimodal input for clinical diagnostics. Nat Biomed Eng. 2023;7:743-755. doi:10.1038/s41551-023-01045-x.

[12] Wang S, Lin M, Ghosal T, Ding Y, Peng Y. Knowledge graph applications in medical imaging analysis: a scoping review. Health Data Sci. 2022;2022:9841548. doi:10.34133/2022/9841548.

[13] Huang Q, et al. Knowledge graph based reasoning in medical image analysis: a scoping review. Comput Biol Med. 2024;182:109100. doi:10.1016/j.compbiomed.2024.109100.

[14] Sadeghi Z, Alizadehsani R, Cifci MA, et al. A review of explainable artificial intelligence in healthcare. Comput Electr Eng. 2024;122:109370. doi:10.1016/j.compeleceng.2024.109370.