ỨNG DỤNG TRÍ TUỆ NHÂN TẠO TRONG CHẨN ĐOÁN VÀ ĐIỀU TRỊ Y HỌC CỔ TRUYỀN: TỔNG QUAN CHUYÊN ĐỀ

Bùi Phạm Minh Mẫn[1], Nguyễn Văn Huy[1]

TÓM TẮT

Y học cổ truyền (YHCT) đang đối mặt với thách thức lớn về tính chủ quan trong chẩn đoán; trí tuệ nhân tạo (AI) cùng các tiến bộ về học sâu và mô hình ngôn ngữ lớn (LLM) trở thành công cụ then chốt để số hóa và chuẩn hóa thực hành lâm sàng. Bài báo tổng quan chuyên đề các nghiên cứu đột phá giai đoạn 2022-2026 nhằm đánh giá hiệu năng AI trong lĩnh vực này. Kết quả cho thấy thiệt chẩn ứng dụng AI đạt AUC từ 0,88 đến 0,92; cảm biến mạch nhiều kênh đạt độ chính xác 97,8%. Các mô hình ngôn ngữ chuyên biệt cho YHCT cho thấy khả năng phân loại chứng hậu vượt trội các LLM đa năng như ChatGPT-4. Hệ thống hỗ trợ kê đơn CIPS tạo giá trị kinh tế-y tế thực, tiết kiệm 77 triệu Đài tệ/năm và phát hiện 3% đơn thuốc rủi ro. AI đã đạt thành tựu định lượng quan trọng, tuy nhiên cần thêm thử nghiệm lâm sàng ngẫu nhiên có đối chứng để khẳng định giá trị trong chẩn đoán và điều trị YHCT thực tiễn.

Từ khóa: y học cổ truyền, trí tuệ nhân tạo, mô hình ngôn ngữ lớn, thiệt chẩn, mạch chẩn, biện chứng luận trị.

ABSTRACT:

Traditional Medicine (TM) faces significant challenges due to the subjectivity of clinical diagnosis. Artificial intelligence (AI), alongside recent advances in deep learning and large language models (LLMs), has become a pivotal tool for digitalizing and standardizing clinical practice. This article provides a narrative review of breakthrough studies from 2022 to 2026 across reputable databases to evaluate the practical performance of AI in this field. Results indicate that AI-applied tongue diagnosis achieves impressive area under the ROC curve (AUC) scores ranging from 0.88 to 0.92, while multichannel pulse sensors reach recognition accuracies of up to 97.8%. Notably, specialized TCM language models demonstrate syndrome classification capabilities superior to general-purpose LLMs like ChatGPT-4. Furthermore, the Clinical Intelligent Prescription System (CIPS) has generated real medical and economic value, saving 77 million New Taiwan Dollars annually while detecting 3% of risky prescriptions. While AI has achieved important quantitative milestones, randomized controlled trials are still required to establish its genuine value in practical clinical diagnosis and treatment.

Keywords: traditional medicine, artificial intelligence, large language model, tongue diagnosis, pulse diagnosis, syndrome differentiation.

I. ĐẶT VẤN ĐỀ

Y học cổ truyền (YHCT) có lịch sử hàng nghìn năm và vẫn giữ vị trí quan trọng trong chăm

1. Trường Đại học Khoa học Sức khỏe – Đại học Quốc gia TPHCM
Tác giả liên hệ: TS. Bùi Phạm Minh Mẫn
Email: bpmman@uhs.vnu.edu.vn ĐT: 84916080803
Ngày nhận bài: 10.4.2026
Ngày chấp nhận đăng: 30.6.2026.

sóc sức khỏe ban đầu tại nhiều quốc gia châu Á. Một thách thức cố hữu chưa được giải quyết: tứ chẩn (vọng, văn, vấn, thiết) phụ thuộc nhiều vào kinh nghiệm cá nhân của thầy thuốc. Cùng một bệnh nhân, hai thầy thuốc khác nhau có thể đưa ra hai chứng hậu khác nhau và kê hai phương án điều trị khác biệt. Việc chuẩn hóa thực hành, đào tạo thế hệ kế tiếp và kiểm chứng hiệu quả điều trị bằng phương pháp khoa học hiện đại đều gặp trở ngại.

Trí tuệ nhân tạo (AI), nhất là học sâu (deep learning) và mô hình ngôn ngữ lớn (large language models, LLM), đang định hình lại cách số hóa và định lượng hóa YHCT. Học sâu trích xuất được những đặc trưng tinh tế từ ảnh lưỡi, tín hiệu mạch hay văn bản bệnh án vốn khó định lượng bằng mắt thường. LLM học từ kho y văn cổ truyền đồ sộ, hỗ trợ thầy thuốc tra cứu, suy luận và đề xuất điều trị. Sự kết hợp hai công nghệ vừa giữ được tinh thần tứ chẩn vừa mang lại tính khách quan và khả năng nhân rộng.

Trong bài tổng quan chuyên đề này, chúng tôi tổng hợp các công trình đột phá tiêu biểu giai đoạn 2022-2026 trên các cơ sở dữ liệu quốc tế gồm PubMed, Scopus, IEEE Xplore, ScienceDirect và arXiv, có kết quả định lượng rõ ràng. Cách tiếp cận theo hướng tổng quan chuyên đề, ưu tiên phân tích các công trình đại diện thay vì liệt kê toàn bộ, nhằm cung cấp cái nhìn về các xu hướng quan trọng và những rào cản còn tồn tại trong lĩnh vực này.

II. NỘI DUNG TỔNG QUAN

2.1. Các mô hình hỗ trợ chẩn đoán lâm sàng

Chẩn đoán lâm sàng là cánh cửa đầu tiên AI bước vào YHCT, với ba hướng tương ứng ba phương pháp tứ chẩn dễ số hóa nhất: thiệt chẩn qua ảnh lưỡi, mạch chẩn qua tín hiệu mạch, biện chứng luận trị qua văn bản bệnh án.

Năm 2023, Yuan và cộng sự công bố trên eClinicalMedicine kết quả từ một nghiên cứu đoàn hệ tiến cứu đa trung tâm quy mô lớn. Sau giai đoạn thăm dò trên 328 bệnh nhân ung thư dạ dày và 304 đối chứng, nhóm nghiên cứu mở rộng triển khai tại 10 trung tâm y tế ở Trung Quốc với 937 bệnh nhân mới chẩn đoán cùng 1.911 đối chứng không ung thư. Nghiên cứu ứng dụng đồng thời ba kiến trúc mạng thần kinh học sâu gồm APINet, TransFG và DeepLabV3+ để phân tích dữ liệu hình ảnh lưỡi kết hợp đặc trưng hệ vi sinh rêu lưỡi thu được qua kỹ thuật giải trình tự 16S rDNA. Các mô hình AI đạt diện tích dưới đường cong ROC (AUC) từ 0,88 đến 0,92 trong nội kiểm và duy trì hiệu năng ổn định 0,83-0,88 khi ngoại kiểm với 7 trung tâm khác. Đáng chú ý, độ chính xác của AI cao hơn có ý nghĩa thống kê so với bảng tổ hợp 8 dấu ấn sinh học huyết thanh sàng lọc kinh điển (CEA, AFP, CA19-9...), qua đó lần đầu khẳng định các đặc trưng lưỡi trong YHCT mang giá trị sinh học có thể đo lường khoa học.[1]

Năm 2025, nghiên cứu TongueNet của Yang và cộng sự tạo đột phá kỹ thuật khi tích hợp đồng thời hình ảnh lưỡi và mô tả triệu chứng bằng văn bản. Mô hình áp dụng cơ chế chú ý đa tỉ lệ EMA để khai thác sâu các đặc trưng bệnh lý quan trọng, đồng thời dùng mạng Kolmogorov-Arnold (KAN) thay thế MLP truyền thống để nâng cao khả năng biểu diễn dữ liệu phức tạp. Với 32,1 triệu tham số, AI đạt độ chính xác 89,12% (AUC 0,83) trong xác định tính chất bệnh và 86,47% (AUC 0,81) cho vị trí tổn thương.[2] Song song đó, Chen và cộng sự công bố mô hình U2Net ứng dụng học bán giám sát (semi-supervised learning) để tự động phân đoạn ảnh lưỡi -

giải pháp then chốt giúp tháo gỡ "nút thắt" về chi phí và công sức chuẩn bị dữ liệu khi xây dựng các hệ thống AI y học quy mô lớn.[3]

Mạch chẩn khó số hóa hơn vì tín hiệu mang tính động học và biến thiên giữa các cá thể. Tiến bộ về cảm biến vi cơ điện tử kết hợp AI đang mở rộng triển vọng. Liu và cộng sự (2023) trên ACS Nano giới thiệu mảng cảm biến mềm 3x9 với 27 kênh, độ nhạy 35,02 kPa^{-1}, đạt độ chính xác 97,8% trong nhận dạng các dạng mạch.[4] Năm 2024, Zhao và cộng sự phát triển vòng đeo tay thông minh tích hợp các túi khí siêu nhỏ để tự động hóa việc "bắt mạch" tại ba vị trí Thốn, Quan, Xích; thiết bị mô phỏng chính xác lực ấn từ ngón tay bác sĩ và có thể dự đoán huyết áp người dùng.[5] Năm 2025, Qiao và cộng sự sử dụng thiết bị cảm biến siêu nhỏ (MEMS/CMOS) kết hợp thuật toán GBDT; hệ thống nhận diện thai kỳ và bệnh mạch vành đạt độ chính xác 80% nếu chỉ dùng dữ liệu mạch, tăng đến 91% khi bổ sung các thông số lâm sàng như tuổi và nhiệt độ.[6]

Bảng 1. Hiệu năng các hệ thống số hóa Thiệt chẩn và Mạch chẩn

Lĩnh vực	Tác giả (Năm)	Công nghệ / Mô hình	Mục tiêu chẩn đoán	Hiệu năng chi tiết
Thiệt chẩn	Yuan và cs. (2023)	APINet, TransFG, DeepLabV3+	Ung thư dạ dày	AUC: 0,88-0,92 (nội kiểm); 0,83-0,88 (ngoại kiểm)
	Yang và cs. (2025)	TongueNet (cơ chế chú ý EMA & mạng KAN)	Tính chất và vị trí tổn thương	Chính xác: 89,12% (tính chất); 86,47% (vị trí)
	Chen và cs. (2025)	U2Net-MT (học bán giám sát)	Phân đoạn ảnh lưỡi tự động	MAE: 0,022; chỉ số Dice: 0,967
Mạch chẩn	Liu và cs. (2023)	Mảng cảm biến mềm 27 kênh	Nhận dạng hình thái mạch	Độ chính xác: 97,8%; độ nhạy: 35,02 kPa^{-1}
	Zhao và cs. (2024)	Vòng đeo tay túi khí siêu nhỏ	Huyết áp và lực ấn	Mô phỏng lực ấn ngón tay; dự đoán huyết áp
	Qiao và cs. (2025)	Cảm biến MEMS/ CMOS + GBDT	Thai kỳ và bệnh mạch vành	Chính xác: 80% (mạch đơn thuần) đến 91% (kèm lâm sàng)

Biện chứng luận trị là lĩnh vực khó định lượng nhất vì phụ thuộc nhiều vào kinh nghiệm chủ quan của thầy thuốc. Năm 2024, mô hình TCM-BERT-CNN ra đời để giải quyết thách thức này bằng cách chẩn đoán đồng thời 21 loại thể chứng khác nhau. Dựa trên dữ liệu chuẩn từ sách giáo khoa và hướng dẫn lâm sàng, AI này đạt độ chính xác trên 92% ở mọi chỉ số đo lường - chứng minh khả năng tư duy chẩn đoán tổng thể và cân bằng, giúp đưa ra kết luận tin cậy ngay cả với

những ca bệnh phức tạp.[7] Năm 2025, mô hình TCMPCSD-BERT chứng minh triết lý "nhỏ mà chuyên" vượt trội hơn việc chạy đua số lượng tham số khổng lồ. Dù kích thước gọn nhẹ, AI này vẫn vượt qua các mô hình đa năng như ChatGPT-4 và Kimi trong việc phân loại chính xác các thể bệnh ung thư tụy (đạt 91,9%); việc tinh chỉnh sâu trên dữ liệu chuyên ngành chất lượng cao giúp AI hoạt động ổn định, chính xác và hiệu quả thực tiễn hơn, trong khi chi phí vận hành thấp hơn nhiều lần.[8]

2.2. Kỷ nguyên của Mô hình ngôn ngữ lớn

Trước năm 2022, AI chủ yếu gồm các mô hình đơn nhiệm, chỉ làm tốt một việc cụ thể. Từ năm 2023, các Mô hình ngôn ngữ lớn (LLM) trở thành "bộ não" trung tâm, giúp hàng loạt AI chuyên biệt cho YHCT ra đời và phát triển vượt bậc.

Ra đời năm 2024, TCM-GPT là mô hình AI mã nguồn mở quy mô lớn đầu tiên chuyên biệt cho YHCT với 7 tỷ tham số. Thay vì tốn kém huấn luyện lại từ đầu, mô hình được bổ sung thêm 1 tỷ token chuyên môn (TCM-Corpus-1B) để nhanh chóng am hiểu kiến thức ngành; nhờ đó AI đạt kết quả thi và độ chính xác trong chẩn đoán vượt trội các mô hình cũ từ 12% đến 17%.[9] Cùng năm 2024, mô hình Lingdan được thiết lập, tập trung vào việc luyện AI cách tư duy và suy luận logic theo đúng kiến thức YHCT thay vì chỉ ghi nhớ máy móc. Phiên bản Lingdan-PR đã được thử nghiệm trực tiếp trên dữ liệu bệnh án thật tại Bệnh viện Quảng An Môn và cho kết quả kê đơn chính xác hơn các công cụ cũ tới 18,39% - một trong những AI đầu tiên có bằng chứng rõ ràng về hiệu quả hỗ trợ bác sĩ trong thực tế điều trị.[10] Tháng 8 năm 2025, mô hình ShizhenGPT ra đời, đánh dấu bước ngoặt khi lần đầu tiên một AI có "nhiều giác quan" dành riêng cho YHCT. Thay vì chỉ đọc văn bản, mô hình này (dựa trên nền tảng Qwen2.5 với 7B và 32B tham số) có thể thực hiện trọn vẹn quy trình tứ chẩn: từ nhìn ảnh lưỡi, nghe tiếng ho đến phân tích sóng mạch và mùi bệnh lý. Tuy nhiên, vì đây là bản nghiên cứu sơ khởi (preprint) chưa qua hội đồng chuyên môn thẩm định, các kết quả ấn tượng vẫn cần được kiểm chứng độc lập trước khi áp dụng thực tế.[11]

Trong năm 2025, Liu và cộng sự đã so sánh 7 mô hình AI nổi tiếng (gồm GPT-4o, Gemini 1.5 Flash, LLaMA 3.2, Claude 3.5 Sonnet, ERNIE 3.5, Qwen 2.5 Max và Doubao 1.5 Pro) với các bác sĩ châm cứu thực thụ trên một ca bệnh xơ cứng cột bên teo cơ. Kết quả đánh giá từ 28 chuyên gia quốc tế cho thấy GPT-4o, Qwen 2.5 Max và Doubao 1.5 Pro đưa ra chẩn đoán và chọn huyệt chính xác gần như bác sĩ chuyên khoa. Tuy nhiên, các AI này vẫn còn hạn chế trong những tình huống đòi hỏi tư duy biện chứng chuyên sâu và xử lý nhiều bước phức tạp.[12]

Bảng 2. Hiệu năng các mô hình ngôn ngữ lớn chuyên biệt trong YHCT

Mô hình (Năm)	Kiến trúc/ Quy mô	Nguồn dữ liệu huấn luyện	Tác vụ trọng tâm	Kết quả đối chứng (Bench-mark)
TCM-GPT (2024)	BLOOM (7B tham số)	1 tỷ token chuyên ngành (TCM-Corpus-1B)	Chẩn đoán và thi cử	Độ chính xác tăng 12%-17% so với các mô hình nền tảng cũ.
Lingdan-PR (2024)	Baichuan2 (13B tham số)	Y văn cổ, dược điển & bệnh án thực tế	Đề xuất bài thuốc	F1-score cải thiện 18,39% so với các phương pháp học máy truyền thống.
TC-MPCSD-BERT (2025)	BERT (tinh chỉnh sâu)	Bệnh án ung thư tụy (y văn và lâm sàng)	Biện chứng luận trị	Độ chính xác 91,9%, vượt trội rõ rệt so với ChatGPT-4 và Kimi.
OpenTCM (2025)	GraphRAG + LLM	68 sách phụ khoa cổ (3,73 triệu chữ Hán)	Truy xuất và hỏi đáp	Trích xuất tri thức 98,55%; điểm chuyên gia chẩn đoán 4,045/5.
ShizhenGPT (2025)	Qwen2.5 (7B-32B)	Đa phương thức (ảnh, sóng mạch, âm thanh)	Tứ chẩn tổng hợp	Dẫn đầu các tác vụ thị giác YHCT; điểm thi hành nghề YHCT 78,1/100.
TCM-BERT-CNN (2024)	BERT + CNN	6.000 mẫu chuẩn từ SGK & hướng dẫn LS	Phân loại chứng hậu	Độ chính xác trên 92% đối với 21 loại thể chứng đồng thời.

2.3. Hệ thống hỗ trợ ra quyết định và triển khai thực tiễn

Chẩn đoán chính xác là bước khởi đầu; trọng tâm tiếp theo là việc thiết lập phác đồ điều trị, mà cốt lõi trong YHCT là nghệ thuật phối ngũ thuốc theo nguyên tắc Quân-Thần-Tá-Sứ. Đây là địa hạt mà AI bắt đầu chuyển mình mạnh mẽ từ các mô hình nghiên cứu sang ứng dụng lâm sàng thực tế.

Năm 2025, He và cộng sự giới thiệu OpenTCM, hệ thống kết hợp LLM với đồ thị tri thức chuyên sâu thông qua kiến trúc GraphRAG. Đồ thị được xây dựng từ 68 cuốn sách phụ khoa cổ (3,73 triệu chữ Hán) trong Chinese Medical Classics Database, tích hợp hơn 48.000 thực thể và 152.000 mối quan hệ. Hệ thống đạt độ chính xác trích xuất tri thức 98,55% và điểm chuyên gia chẩn đoán đạt 4,045/5. Dù là bản tiền in tập trung vào phụ khoa, OpenTCM khẳng định xu hướng dùng tri thức cấu trúc để giảm thiểu ảo giác (hallucination) và tăng cường tính giải thích cho AI trong YHCT.[13]

Về hệ đề xuất bài thuốc, PresRecRF của Yang và cộng sự năm 2024 trên Phytomedicine hợp nhất hai nguồn thông tin: ngữ nghĩa quy mô lớn của YHCT học từ LLM huấn luyện trên y văn cổ truyền và tri thức phân tử của các vị thuốc. Kiến trúc hợp nhất biểu diễn (representation fusion) tạo ra một biểu diễn tổng hợp cho từng bệnh nhân, cải thiện rõ rệt độ chính xác đề xuất so với các mô hình so sánh chuẩn truyền thống.[14] PrescDRL của Yang và cộng sự cùng năm 2024 trên Chinese Medicine áp dụng học tăng cường sâu (deep reinforcement learning) để lập kế hoạch điều trị tuần tự cho bệnh đái tháo đường mạn tính. PrescDRL không đề xuất một đơn thuốc tĩnh; mô hình mô phỏng quá trình điều chỉnh đơn theo từng giai đoạn bệnh, với độ chính xác tăng 40,5% và độ bao phủ tăng 63% so với mô hình chuẩn trong dự đoán đơn thuốc.[15]

Một minh chứng thực tế tiêu biểu về hiệu quả của AI trong YHCT là Hệ thống Đơn thuốc Thông minh Trung Hoa (CIPS), được cấp bằng sáng chế M617562 và triển khai tại Bệnh viện Đại học Y học Trung Hoa. Về mặt an toàn lâm sàng, nghiên cứu của Lin và cộng sự (2023) trên Heliyon cho thấy CIPS phát hiện 3% đơn thuốc sai liều lượng, giúp ngăn chặn nguy cơ nhiễm độc tiềm tàng cho hơn 170.000 đơn thuốc mỗi tháng trên toàn Đài Loan.[16] Về mặt vận hành, Huang và cộng sự (2023) trên BMC Health Services Research ghi nhận CIPS giúp tối ưu hóa đơn thuốc bằng cách giảm số vị thuốc trung bình (từ 8,19 xuống 7,37), rút ngắn thời gian phát thuốc xuống còn 1,63 phút/đơn, từ đó mỗi dược sĩ tiết kiệm 3,75 giờ làm việc mỗi tháng. Tổng tiết kiệm chi phí mỗi dược sĩ mỗi năm là 20.005 Đài tệ và toàn Đài Loan ước tính tiết kiệm 77 triệu Đài tệ mỗi năm, tương đương khoảng 2,7 triệu USD. CIPS khẳng định AI không chỉ là lý thuyết mà còn tạo ra giá trị y tế và tài chính trực tiếp khi tích hợp vào quy trình bệnh viện.[17]

Bảng 3. Hiệu quả vận hành và giá trị kinh tế-y tế của hệ thống hỗ trợ kê đơn CIPS

Phân loại	Chỉ số đánh giá	Kết quả định lượng
An toàn lâm sàng	Phát hiện sai sót liều lượng	3% tổng số đơn thuốc được cảnh báo liều không chính xác
	Quy mô ngăn ngừa độc tính	Ngừa nguy cơ nhiễm độc cho > 170.000 đơn/tháng (toàn Đài Loan)
Tối ưu đơn thuốc	Số vị thuốc trung bình/đơn	Giảm từ 8,19 xuống 7,37 vị
	Tỷ lệ giảm hạng mục thuốc	Giảm 9,31% nhờ thuật toán loại bỏ phối ngũ trùng lặp
Hiệu suất vận hành	Thời gian phát thuốc	Rút ngắn từ 1,79 xuống 1,63 phút/đơn
	Tiết kiệm nhân lực	3,75 giờ làm việc/tháng cho mỗi dược sĩ
Giá trị kinh tế	Tiết kiệm trên mỗi dược sĩ	20.005 Đài tệ/năm (gồm nhân công và giảm thất thoát thuốc)
	Tổng tiết kiệm toàn hệ thống	77 triệu Đài tệ/năm (~2,7 triệu USD) cho toàn Đài Loan

2.4. Bàn luận về xu hướng và rào cản

Các công trình đã phân tích cho thấy ba xu hướng. Thứ nhất, xu hướng chuyển đổi đa phương thức: TongueNet kết hợp ảnh lưỡi với văn bản triệu chứng; ShizhenGPT bao quát cả vọng văn vấn thiết; PresRecRF kết hợp ngữ nghĩa với tri thức phân tử các vị thuốc - hướng đi này phản ánh đúng tinh thần tứ chẩn truyền thống. Thứ hai, LLM chuyên biệt đang thay thế dần học máy truyền thống. Trước năm 2022, các mô hình chuyên cho từng nhiệm vụ chiếm ưu thế; từ 2023, các LLM chuyên biệt cho YHCT như TCM-GPT, Lingdan, ShizhenGPT trở thành nền tảng đa nhiệm. Tuy vậy, công trình của Ba và cộng sự (2025) cho thấy LLM lớn không phải lúc nào cũng tốt hơn: một mô hình BERT chuyên biệt nhỏ vẫn

vượt được ChatGPT-4 trong phân loại chứng hậu chuyên sâu. Thứ ba, khoảng cách giữa nghiên cứu và lâm sàng đang nới rộng: mỗi năm hàng trăm công trình báo cáo độ chính xác trên 90%, nhưng số mô hình thực sự được triển khai tại bệnh viện và tạo giá trị đo lường được vẫn rất hạn chế - CIPS tại Đài Loan là một trong những ngoại lệ hiếm hoi.

Dù sở hữu tiềm năng phát triển to lớn, việc ứng dụng AI vào YHCT hiện đang đối mặt với ba rào cản lớn. Thứ nhất là trở ngại ở tính hệ thống và chuẩn hóa dữ liệu, do YHCT có nhiều trường phái khác nhau giữa các nước; ngoài ra tên thuốc, cách phân loại chứng hậu, chứng trạng và cách ghi bệnh án không có chuẩn thống nhất, gây khó khăn cho việc huấn luyện các mô hình AI quy mô lớn. Thứ hai là thách thức về hành lang pháp lý: trong khi Hoa Kỳ và Châu Âu đã có những bước tiến cụ thể với hướng dẫn Predetermined Change Control Plan của FDA (12/2024) hay Đạo luật AI (EU AI Act) có hiệu lực từ tháng 8/2024, thì phần lớn các nước Châu Á, bao gồm Việt Nam, vẫn đang thiếu hụt khung pháp lý chuyên biệt để quản lý và định hướng cho các thiết bị y tế tích hợp AI trong YHCT. Cuối cùng là rào cản về mức độ tin cậy lâm sàng; trong khi các thầy thuốc Y học hiện đại phần nào chấp nhận tin vào AI cho các nhiệm vụ hẹp như phân loại ảnh X-quang ngực, thì trong lĩnh vực YHCT lại khác - quy trình chẩn đoán mang tính cá thể hóa cao, biện chứng luận trị đa tầng, khiến việc tin tưởng hoàn toàn vào các kết luận từ AI vẫn là một bài toán cần thêm nhiều minh chứng thực tiễn về tính giải thích được và sự phù hợp chặt chẽ với y văn cổ điển.

III. KẾT LUẬN

AI đã tạo nên những thành tựu định lượng quan trọng trong YHCT giai đoạn 2022 đến đầu năm 2026. Thiệt chẩn đạt AUC 0,88-0,92 trong xác thực nội bộ và 0,83-0,88 trong xác thực ngoài. Mạch chẩn đạt độ chính xác 97,8% với mảng cảm biến nhiều kênh và 80-91% với thiết bị MEMS/CMOS. Biện chứng luận trị có mô hình BERT chuyên biệt vượt trội ChatGPT-4 trong phân loại chứng hậu ung thư tụy. Các LLM chuyên biệt như TCM-GPT, Lingdan, ShizhenGPT đã chuyển lĩnh vực sang giai đoạn đa nhiệm và đa phương thức. Hệ thống CIPS tại Đài Loan đã chứng minh giá trị thực tiễn với việc phát hiện 3% đơn thuốc rủi ro và tiết kiệm 77 triệu Đài tệ Mới mỗi năm.

Phần lớn các con số nêu trên đến từ đánh giá hồi cứu hoặc thiết kế quan sát. Ngoài công trình của Yuan và cộng sự (2023) cùng hai báo cáo về CIPS, lĩnh vực AI y học cổ truyền vẫn thiếu các thử nghiệm lâm sàng ngẫu nhiên có đối chứng đa trung tâm chứng minh AI thực sự hiệu quả trong chẩn đoán và điều trị trên bệnh nhân.

TÀI LIỆU THAM KHẢO

1. **Yuan L, Yang L, Zhang S, et al.** Development of a tongue image-based machine learning tool for the diagnosis ofgastric cancer: a prospective multicentre clinical cohort study. eClinicalMedicine. 2023;57:101834.

2. **Yang L, Dong Q, Lin D, Lü X.** TongueNet: a multi-modal fusion and multi-label classification model for traditional Chinese Medicine tongue diagnosis. Front Physiol. 2025;16:1527751.

3. **Chen T, Chen Y, Zhou Z, Zhu Y, He L, Zhang J.** Deep learning-based automated tongue analysis system forassisted Chinese medicine diagnosis. Front Physiol. 2025;16:1559389.

4. **Liu T, Gou GY, Gao F, et al.** Multichannel flexible pulse perception array for intelligent disease diagnosis system. ACS Nano. 2023;17(6):5673-5685.

5. **Zhao Y, Sun Q, Mei S, et al.** Wearable

multichannel-active pressurized pulse sensing platform. Microsyst Nanoeng. 2024;10:77.

6. **Qiao M, Yan Z, Tan C, et al.** Integrating traditional Chinese pulse diagnosis with machine learning: novel approaches for pregnancy and coronary heart disease identification. Sci Rep. 2025;15:35857.

7. **Chen Z, Zhang D, Liu C, et al**. Traditional Chinese medicine diagnostic prediction model for holistic syndromedifferentiation based on deep learning. Integr Med Res. 2024;13(1):101019.

8. **Ba H, Du H, Cheng C, Zhang Y, Ruan L, Chen Z.** Identification of syndrome types in patients with pancreaticcancer from free text in electronic medical records: model development and validation. JMIR Form Res. 2025;9:e70602.

9. **Yang G, Liu X, Shi J, Wang Z, Wang G. TCM-GPT:** efficient pre-training of large language models for domainadaptation in traditional Chinese medicine. Comput Methods Programs Biomed Update. 2024;6:100158.

10. **Hua R, Dong X, Wei Y, et al.** Lingdan: enhancing encoding of traditional Chinese medicine knowledge for clinical reasoning tasks with large language models. J Am Med Inform Assoc. 2024;31(9):2019-2029.

11. **Chen J, Cai Z, Liu Z, et al.** ShizhenGPT: towards multimodal LLMs for traditional Chinese medicine. arXiv. Preprint posted online August 20, 2025. arXiv:2508.14706.

12. **Liu Y, Yuan Y, Yan K, et al.** Evaluating the role of large language models in traditional Chinese medicine diagnosis and treatment recommendations. NPJ Digit Med. 2025;8(1):466.

13. **He J, Guo Y, Lam LK, et al.** OpenTCM: a GraphRAG-empowered LLM-based system for traditional Chinesemedicine knowledge retrieval and diagnosis. arXiv. Preprint posted online April 28, 2025. arXiv:2504.20118.

14. **Yang K, Dong X, Zhang S, et al.** PresRecRF: herbal prescription recommendation via the representation fusion of large TCM semantics and molecular knowledge. Phytomedicine. 2024;135:156116.

15. **Yang K, Yu Z, Su X, et al.** PrescDRL: deep reinforcement learning for herbal prescription planning in treatment of chronic diseases. Chin Med. 2024;19:144.

16. **Lin DY, Huang WT, Lin YC, et al.** Prescription system to calculate precise doses of Chinese herbal medicine to avoid toxic effects. Heliyon. 2023;9(6):e16612.

17. **Huang TY, Huang WT, Lin YC, et al.** Chinese Intelligence Prescription System improves prescription accuracy while decreasing labor and drug costs. BMC Health Serv Res. 2023;23(1):514.