

PHÁT TRIỂN THUẬT TOÁN TRÍ TUỆ NHÂN TẠO TRONG PHÁT HIỆN VIÊM THỰC QUẢN TRÀO NGƯỢC TRÊN TẬP ẢNH NỘI SOI

Bùi Trí Thức¹, Lâm Ngọc Hoa², Vũ Thị Ly³ và Đào Việt Hằng^{1,2,4,✉}

¹Trường Đại học Y Hà Nội

²Viện Nghiên cứu và Đào tạo Tiêu hóa - Gan mật

³Đại học kỹ thuật Lê Quý Đôn

⁴Bệnh viện Đại học Y Hà Nội

Mục tiêu: Đánh giá độ chính xác của thuật toán AI trong phát hiện tổn thương viêm thực quản trào ngược, khảo sát yếu tố liên quan đến việc bỏ sót, nhận nhầm. Phương pháp nghiên cứu: Mô tả cắt ngang. Thuật toán được kiểm chứng trên tập ảnh tĩnh bao gồm 1000 ảnh với các chế độ ánh sáng bằng cách so sánh phần khoanh vùng chuẩn của chuyên gia. Độ chính xác được đánh giá bằng độ nhạy, độ đặc hiệu, giá trị dự báo dương tính, giá trị dự báo âm tính. So sánh các tỷ lệ - xét mối liên quan được sử dụng để khảo sát những yếu tố liên quan đến tỉ lệ bỏ sót và nhận nhầm. Kết quả: Độ chính xác là 81,7%. Số lượng và kích thước tổn thương liên quan đến tỉ lệ bỏ sót. Tổn thương kèm theo, độ sạch liên quan đến tỉ lệ nhận nhầm. Kết luận: Thuật toán Yolov8 có độ chính xác tốt, có tiềm năng phát triển theo hướng đồng kiểm với bác sĩ nội soi trong khí soi, hậu kiểm sau nội soi, tham gia đào tạo y khoa với dữ liệu lớn.

Từ khóa: Trí tuệ nhân tạo, nội soi tiêu hóa trên, viêm thực quản trào ngược, khoanh vùng tổn thương.

I. ĐẶT VẤN ĐỀ

Viêm thực quản trào ngược (VTQTN) được định nghĩa là tình trạng viêm niêm mạc thực quản gặp trong bệnh trào ngược dạ dày thực quản (GERD) khi các chất trong dạ dày trào ngược lên thực quản gây ra các triệu chứng và biến chứng khó chịu.¹ Đây là bệnh phổ biến trong cộng đồng và thường gặp trong thực hành lâm sàng hàng ngày. Bệnh có thể gây ra các triệu chứng ảnh hưởng đến chất lượng cuộc sống người bệnh và có thể gây ra một số biến chứng như: loét thực quản, hẹp thực quản, Barrett thực quản và ung thư thực quản.^{2,3} Đồng thuận SEA đã chỉ ra rằng tỷ lệ mắc GERD tại Đông Nam Á trong đó có Việt Nam ước tính là 6% - 18,3%.⁴ Nội soi đường tiêu hóa trên (NSTHT)

được chỉ định giúp đánh giá mức độ tổn thương niêm mạc thực quản, các biến chứng nếu có và được đưa vào các tiêu chuẩn để chẩn đoán xác định GERD.⁵ Trong những năm gần đây việc ứng dụng công nghệ thông tin, điển hình là trí tuệ nhân tạo – AI (artificial intelligence) trong nội soi tiêu hóa là một bước tiến đáng kể trong việc phát hiện và hỗ trợ chẩn đoán đối với các bác sĩ. Sự phát triển của trí tuệ nhân tạo (AI) có thể giúp giải quyết vấn đề chẩn đoán sai hoặc bỏ sót tổn thương do lỗi của người dùng, đặc biệt là đối với các bác sĩ nội soi mới vào nghề.⁶ AI trong tương lai có thể được kì vọng đóng vai trò song song hoặc hậu kiểm hỗ trợ trong quá trình bác sĩ nội soi bằng các tập ảnh thu được hoặc các video thời gian thực.⁶ Hơn nữa, AI có thể hỗ trợ các bác sĩ tránh bỏ sót tổn thương nhất là trong hoàn cảnh áp lực các ca nội soi ngày càng đông, giúp giảm chi phí y tế trong nhiều lần khám.⁷ Một nghiên cứu của Chi-Chih-Wang về sử dụng mô hình học sâu để phân loại GERD

Tác giả liên hệ: Đào Việt Hằng

Trường Đại học Y Hà Nội

Email: hangdao.fsh@gmail.com

Ngày nhận: 23/08/2023

Ngày được chấp nhận: 17/09/2023

trên hình ảnh nội soi cho độ chính xác 87,9% cao hơn hẳn so với kết quả của bác sĩ ít kinh nghiệm (75,0% và 65,6%).⁸ Tuy nhiên, chưa có nghiên cứu nào về ứng dụng trí tuệ nhân tạo đặc biệt là xây dựng các thuật toán học máy trong nội soi đường tiêu hóa trên để đánh giá mức độ khả thi trong việc chẩn đoán viêm thực quản trào ngược. Chính vì vậy nghiên cứu này được tiến hành với mục tiêu: *Đánh giá độ chính xác và một số yếu tố liên quan của thuật toán Yolo v8 trong phát hiện viêm thực quản trào ngược trên hình ảnh chụp nội soi thực quản dạ dày.*

II. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP

1. Đối tượng

Địa điểm và thời gian nghiên cứu

Hình ảnh nội soi thực quản được thu thập tại Viện Nghiên cứu và Đào tạo Tiêu hóa Gan mật từ tháng 01/2023 đến tháng 07/2023.

Thuật toán sử dụng

Thuật toán Yolo v8 do nhóm công nghệ thông tin của Đại học kỹ thuật Lê Quý Đôn xây dựng. Với kết quả thu được từ thuật toán là các tổn thương được khoanh vùng dưới dạng ô phát hiện tổn thương (bounding box).

Tiêu chuẩn lựa chọn

Hình ảnh nội soi được thu thập ở các chế độ ánh sáng (WLI, FICE, BLI, LCI) sử dụng hệ thống máy nội soi độ phân giải cao của Fujifilm (6000, 7000, LASEREO). Hình ảnh được lấy ở thực quản đoạn 1/3 dưới có đường Z, có độ phân giải ảnh tối thiểu 720 x1280 pixel. Đường Z là đường chuyển tiếp từ biểu mô vảy thực quản sang biểu mô trụ dạ dày.⁸

Tiêu chuẩn loại trừ

Những hình ảnh không xác định được rõ tổn thương viêm thực quản do nguyên nhân khác như: Viêm thực quản do thuốc, tăng bạch cầu ái toan, phóng xạ và viêm thực quản nhiễm trùng; giãn tĩnh mạch thực quản, ung thư thực quản.

2. Phương pháp

Thiết kế nghiên cứu

Mô tả cắt ngang.

Phương pháp chọn mẫu

Tập ảnh để đánh giá độ chính xác của thuật toán được thu thập theo 2 nhóm gồm nhóm ảnh có tổn thương viêm thực quản trào ngược (và nhóm ảnh không có tổn thương viêm thực quản trào ngược được gán nhãn bởi bác sĩ nội soi tiêu hóa có kinh nghiệm trên 5 năm. Tất cả các ảnh đều được ghi nhận thông tin theo thang điểm đánh giá mức độ sạch đường tiêu hóa trên (Mucosal Visibility Score - MVS) theo 4 mức độ:⁹

1) là không có nhầy bọt dính và nhìn rõ niêm mạc;

2) có một lớp nhầy mỏng nhưng không ảnh hưởng tới quan sát niêm mạc;

3) có một số chất nhầy bọt che khuất một phần niêm mạc (tổn thương niêm mạc nhỏ có thể bị bỏ sót nếu không rửa);

4) chất nhầy, bọt nhiều che khuất tầm nhìn của niêm mạc. Hình thái tổn thương viêm thực quản trào ngược theo phân loại Los Angeles.¹⁰

Gán nhãn và khoanh vùng: Ảnh nội soi thu được đều được nhận định và phân loại trên phần mềm Excel dựa trên các đặc điểm: chế độ ánh sáng, kích thước tổn thương, phân loại viêm thực quản trào ngược theo LA, chia tệp ban đầu các tổn thương. 3 chuyên gia nội soi tham gia khoanh vùng tổn thương với các tiêu chuẩn như: Có kinh nghiệm nội soi đường tiêu hóa trên trên 5 năm; là những người tham gia xây dựng, giảng dạy chương trình đào tạo nội soi tiêu hóa trên ở các trường đại học, bệnh viện; sau đó kết quả khoanh vùng dưới dạng ô phát hiện tổn thương được cả 3 chuyên gia thống nhất.

Xử lý số liệu

Kết quả hình ảnh sau khi thuật toán xử lý

thu được dùng để để đánh giá độ nhạy (Se), độ đặc hiệu (Sp), độ chính xác, giá trị tiên đoán âm tính (NPV), giá trị tiên đoán dương tính của thuật toán (PPV), với giá trị chuẩn là ý kiến của chuyên gia (Bảng 1).¹¹ Tất cả các trường hợp có xuất hiện ít nhất 01 âm tính giả (bỏ sót) hoặc 01 dương tính giả trong tệp ảnh viêm thực

quản trào ngược (nhận nhầm vùng không có tổn thương thành viêm thực quản trào ngược) được quy về các trường hợp AI nhận định sai. Phương pháp so sánh các tỷ lệ - nhận xét mối liên quan được sử dụng để khảo sát những yếu tố liên quan đến tỉ lệ bỏ sót và nhận nhầm.

Bảng 1. Cách tính các chỉ số nghiên cứu theo số ảnh¹¹

Chỉ số	Công thức	Định nghĩa
Độ nhạy (Se)	$a/(a+c)$	Số ảnh có viêm thực quản được AI nhận định đúng/ tổng số ảnh có viêm thực quản
Độ đặc hiệu (Sp)	$d/(b+d)$	Số ảnh không có viêm thực quản được AI nhận định đúng/ tổng số ảnh không có viêm thực quản
Giá trị dự đoán dương (PPV)	$a/(a+b)$	Xác suất ảnh có viêm thực quản được AI nhận định đúng
Giá trị dự đoán âm (NPV)	$d/(c+d)$	Xác suất ảnh không có viêm thực quản được AI nhận định đúng
Độ chính xác	$(a+d)/(a+b+c+d)$	Độ chính xác của AI

Trong đó:

- a: Số ảnh viêm thực quản AI nhận định đúng;
- b: Số ảnh không có viêm thực quản AI nhận định sai;
- c: Số ảnh viêm thực quản AI nhận định sai;
- d: Số ảnh không có viêm thực quản AI nhận định đúng.

III. KẾT QUẢ

Từ tháng 1/2023 đến tháng 07/2023 nghiên cứu thu thập được 500 hình ảnh với 693 vùng tổn thương viêm thực quản trào ngược đã được chuyên gia nhận định và 500 hình ảnh của niêm

mạc đường Z không có viêm thực quản trào ngược. Các đặc điểm mô tả chi tiết trong các bảng sau đây:

Bảng 2. Phân bố các đặc điểm vùng viêm thực quản trong tập viêm thực quản (n = 500)

Chế độ ánh sáng	n	%	VTQ theo phân độ LA	n	%
WLI	177	35,4	Độ A	416	83,2
FICE	66	13,2			
BLI	117	23,4	Độ B	72	14,4
LCI	140	28			
Tổn thương kèm theo			Độ C	10	2
Barrett	24	4,8	Độ D	2	0,4
Dị sản niêm mạc dạ dày	13	2,6	Kích thước trượt		
Loét	8	1,6	Kích thước < 5mm	416	83,2
Số trượt trong 1 hình ảnh			Kích thước ≥ 5mm	84	16,8
1	350	70			
≥ 2	150	30			

Trong tập ảnh viêm thực quản, chế độ ánh sáng hay được dùng nhất trong phát hiện tổn thương là chế độ WLI, ít nhất là FICE. Đa số tổn thương là viêm thực quản trào ngược độ A với kích thước trượt < 5mm. Tổn thương thực quản độ D có 2 trường hợp. Tổn thương đi kèm

chủ yếu là Barrett thực quản đoạn ngắn.

Trong tập không có viêm thực quản, chế độ ánh sáng hay được dùng nhất là WLI, các tổn thương kèm theo thường là dị sản niêm mạc dạ dày trên thực quản. Chẩn đoán thực quản bình thường chiếm 96,6%.

Bảng 3. Mô tả đặc điểm các vùng ảnh khoan nhằm và bỏ sót tổn thương theo số tổn thương

Lỗi khoan nhằm	Tập VTQTN	Tập không VTQTN	Lỗi bỏ sót	Số tổn thương	%
Bọt, nhầy,	0	1	Viêm thực quản nhỏ, ở xa	20	16,8
Nếp niêm mạc	8	16	Viêm thực quản trên nếp niêm mạc	36	30,2
Vùng lóa, mờ	8	9	Viêm thực quản vùng tối, lóa	63	53
Niêm mạc thực quản	0	5	Tổng	119	100
Loét	2	0	Lỗi khoan chưa chính xác		
Barrett	8	4	Lỗi khoan thừa	31	75,6
Dị sản niêm mạc dạ dày	4	2	Lỗi khoan thiếu	10	24,4
Tổng	29	37			

Trong tập ảnh viêm thực quản thuật toán AI đã mô tả 693 vùng, trong đó AI phát hiện chính xác 497 (chiếm 71,7%) vùng có tổn thương viêm thực quản. Có 41 vùng phát hiện đúng tổn thương nhưng khoan viên chưa chính xác (31

lỗi khoan thừa và 10 lỗi khoan thiếu), 119 lỗi bỏ sót. Tỷ lệ khoan nhằm tổn thương trong 2 tập viêm thực quản trào ngược và không có viêm thực quản trào ngược lần lượt là 4,2% (29 ảnh) và 5,3% (27 ảnh).

Bảng 4. Kết quả nhận định của AI theo số ảnh

	Số ảnh do chuyên gia nhận định	
	Có viêm thực quản trào ngược	Không có viêm thực quản trào ngược
Nhận định đúng	354	463
Nhận định sai	146	37
Tổng	500	500

Như vậy: Se, PPV, Sp, NPV và độ chính xác lần lượt là 70,8%, 90,5%, 92,6%, 76% và 81,7%.

Bảng 5. Mối liên quan giữa các yếu tố với tỷ lệ bỏ sót và nhận nhầm viêm thực quản

Đặc điểm VTQ bị bỏ sót	OR	95%CI	p
Nằm trong vùng nhiều tổn thương VTQ	2,07	1,38 - 3,12	0,001
Kích thước $\leq 5\text{mm}$	1,98	1,18 - 4,16	0,01
Điểm sạch thực quản MVS = 2	2,93	1,63 - 5,27	0,001
Chế độ ánh sáng (nhóm tham chiếu): WLI	0,91	0,61 - 1,36	0,31
Đặc điểm vùng ảnh bị nhận nhầm trong tập không có VTQ			
Điểm sạch thực quản MVS = 2	1,53	1,58 - 12,97	0,002
Chẩn đoán không có VTQ: Dị sản, Barrett	15,94	11,33 - 22,41	0,001
Chế độ ánh sáng (nhóm tham chiếu): WLI	1,2	0,6 - 2,41	0,26

Các yếu tố liên quan đến tỉ lệ bỏ sót tổn thương bao gồm các tổn thương trợt niêm mạc kích thước $\leq 5\text{mm}$, các tổn thương nằm trong ảnh chứa nhiều hơn 1 trợt và các trợt nằm trong vùng có mức MVS = 2. Các ảnh có độ sạch kém và có chẩn đoán dị sản dạ dày, Barrett làm tăng tỷ lệ nhận nhầm. Chế độ ánh sáng sử dụng không liên quan đến việc nhận nhầm và bỏ sót tổn thương.

IV. BÀN LUẬN

Chúng tôi sử dụng thuật toán Yolo v8 là mô hình phát hiện đối tượng phổ biến được biết đến với tốc độ nhanh và độ chính xác cao. Thuật toán cho phép sử dụng hình ảnh ở độ phân giải cao hơn giúp phát hiện các đối tượng nhỏ hơn và có độ chính xác tổng thể cao hơn.

Các nghiên cứu hiện nay trên thế giới xây

dựng để phát hiện viêm thực quản trào ngược còn ít, chủ yếu là các thuật toán để phát hiện các tổn thương thực quản như ung thư biểu mô tế bào vảy, Barrett.¹² Trong nghiên cứu, thuật toán của chúng tôi xây dựng có độ chính xác 82,7%. Kết quả này thấp hơn so với nghiên cứu của Hailong Ge và cộng sự tại Trung Quốc (91- 95%) và tương đương với kết quả nghiên cứu của Chi-Chih-Wang (87,9%) với thuật toán GERD-VGGNet.^{8,11} Sự khác biệt trên là do nghiên cứu của Hailong Ge kéo dài trong 5 năm và sử dụng bộ dữ liệu lớn về hình ảnh nội soi của HyperKvasir để huấn luyện nhiều thuật toán được áp dụng như EfficientNet, MobileNet, ResNet. Trong xây dựng các thuật toán AI, dữ liệu đầu vào là vô cùng quan trọng, các nghiên cứu trên được huấn luyện bởi các bộ ảnh đầu vào là các tập ảnh tiêu chuẩn và thời gian đủ lớn để quá trình đào tạo diễn ra tốt nhất. Trong khi, nghiên cứu của chúng tôi mới được thực hiện trong 1 năm và bộ ảnh dùng để huấn luyện được chúng tôi tự thu thập, cỡ mẫu huấn luyện chưa đủ lớn và quá trình đào tạo thuật toán vẫn đang diễn ra. Chúng tôi cũng là nhóm đầu tiên sử dụng thuật toán Yolo v8 để phát triển vấn đề chẩn đoán viêm thực quản trào ngược này. Một yếu tố nữa là chế độ ánh sáng thu thập ban đầu. Các nghiên cứu của Hailong Ge và Chi-Chih-Wang thu thập 2 chế độ ánh sáng cơ bản là ánh sáng trắng và ánh sáng dải tần hẹp (NBI), trong khi đó chúng tôi thu thập bộ hình ảnh với 4 chế độ ánh sáng khác nhau (WLI, FICE, BLI, LCI). Điều này cũng có thể làm đa dạng quá trình nhận biết của thuật toán, tuy nhiên cũng cần thời gian để quá trình huấn luyện thuật toán được hoàn thiện hơn. Với độ chính xác trên cùng với việc sử dụng các hình ảnh có độ phân giải cao, thuật toán trong tương lai có thể phát triển theo hướng đồng kiểm với bác sĩ trong cuộc nội soi hoặc hậu kiểm sau nội soi, khắc phục được các nhược điểm liên quan đến con người như khả năng tập trung có hạn

của bác sĩ và áp lực số lượng ca nội soi. Ngoài ra, việc phát triển thuật toán AI còn có thể tham gia đào tạo y khoa với dữ liệu lớn.

Trong số các tổn thương viêm thực quản bị bỏ sót, chúng tôi đặc biệt quan tâm đến chế độ ánh sáng và kích thước nốt bị bỏ sót. Phân loại LA cung cấp các định nghĩa về nốt niêm mạc chỉ được chẩn đoán bởi ánh sáng trắng. Chế độ ánh sáng trắng (WLI) được sử dụng để phát hiện ra những tổn thương viêm thực quản nhiều nhất với 35,4%. Sự khác biệt lớn cũng có thể liên quan đến sự thay đổi trong ước tính kích thước của tổn thương, nghĩa là khó xác định liệu tổn thương có dài hơn hay nhỏ hơn 5 mm hay không¹³. Trong tập này, những tổn thương nốt niêm mạc < 5mm bị bỏ sót chiếm tới 65,1%. Các ảnh có điểm sạch thực quản MVS = 2 cũng làm tăng tỷ lệ bỏ sót.

Ở tập không có viêm thực quản, những tổn thương kèm theo như dị sản dạ dày, Barrett thực quản cũng làm ảnh hưởng đến tỷ lệ nhận nhầm, điểm MVS = 2 cũng có tỷ lệ nhận nhầm cao hơn so với mức MVS = 3 với $p = 0,002$, có thể lý do là những tổn thương đi kèm thường sát đường Z và màu sắc tổn thương gần giống nốt niêm mạc nên dễ bị nhận nhầm. Với lỗi nhận nhầm chủ yếu là nốt niêm mạc (16 ảnh) cho thấy việc quan sát chi tiết đường Z còn khó khăn ở những bệnh nhân được an thần quá mức vì đường Z không giãn ra. Ngoài ra, kỹ thuật được sử dụng bởi bác sĩ nội soi đôi khi có thể dẫn đến hình ảnh kém bên cạnh chất lượng ảnh kém (độ sạch kém, mờ, lóa).¹⁴ Do đó, AI sẽ không thể bù đắp cho chất lượng nội soi dạ dày dưới mức tối ưu, mức độ sạch. Những yếu tố này vẫn là điều kiện chính để tối đa hóa độ chính xác của AI.

Nghiên cứu của chúng tôi có một vài hạn chế. Đây là một nghiên cứu tiến tại một trung tâm, sử dụng hình ảnh tĩnh do đó cần thêm bộ ảnh dữ liệu tại các trung tâm khác với các dòng

máy nội soi với chế độ hình ảnh và độ phân giải khác nhau để tăng tính đa dạng bộ dữ liệu thu thập. Thuật toán Yolo v8 mặc dù có tốc độ xử lý và độ chính xác cao tuy nhiên để hoạt động tối ưu cần phải được đào tạo trên một tập dữ liệu đủ lớn và đang dạng hơn nữa. Các kết quả ban đầu này đã gợi ý cần thêm các nghiên cứu dựa trên video thời gian thực để đưa ra những đánh giá về vai trò của AI trong phát hiện tổn thương đường tiêu hóa trên trong tương lai, đặc biệt với nhóm tổn thương không phải ác tính.

V. KẾT LUẬN

Kết quả kiểm chứng Thuật toán Yolo v8 trên tập ảnh tĩnh ghi nhận độ chính xác là 81,7%. Kích thước, số lượng vùng tổn thương và độ sạch là yếu tố liên quan đáng kể tới tỉ lệ bỏ sót, độ sạch và chẩn đoán có liên quan đến tỉ lệ nhận nhầm. Kết luận của nghiên cứu cho thấy AI có thể có tiềm năng trong việc hỗ trợ chẩn đoán viêm thực quản trào ngược. Tuy nhiên, cần tiếp tục nghiên cứu để cải thiện độ chính xác và ứng dụng thực tế của phương pháp AI trong chẩn đoán viêm thực quản trào ngược.

VI. KIẾN NGHỊ

Dựa trên kết quả của nghiên cứu, chúng tôi kiến nghị rằng các nghiên cứu tiếp theo nên được thực hiện với cỡ mẫu lớn hơn để đánh giá độ tin cậy của phương pháp chẩn đoán bằng trí tuệ nhân tạo (AI). Bên cạnh đó, cần thực hiện các nghiên cứu liên quan đến áp dụng AI trong việc chẩn đoán viêm thực quản trào ngược. Nghiên cứu này đã mở ra nhiều cơ hội trong việc áp dụng công nghệ trí tuệ nhân tạo vào lĩnh vực y học và tạo tiền đề cho các nghiên cứu sâu hơn về phát hiện và chẩn đoán tổn thương đường tiêu hóa.

TÀI LIỆU THAM KHẢO

1. Azer SA, Reddivari AKR. Reflux Esophagitis. In: *StatPearls*. StatPearls Publishing;

2023. Accessed August 8, 2023. <http://www.ncbi.nlm.nih.gov/books/NBK554462/>.

2. Tack J, Becher A, Mulligan C, Johnson DA. Systematic review: the burden of disruptive gastro-oesophageal reflux disease on health-related quality of life. *Aliment Pharmacol Ther*. 2012; 35(11): 1257-1266.

3. Herbella FA, Patti MG. Gastroesophageal reflux disease: From pathophysiology to treatment. *World J Gastroenterol*. 2010; 16(30): 3745-3749.

4. Goh KL, Lee YY, Leelakusolvong S, et al. Consensus statements and recommendations on the management of mild-to-moderate gastroesophageal reflux disease in the Southeast Asian region. *JGH Open*. 2021; 5(8): 855-863. doi:10.1002/jgh3.12602.

5. Gyawali CP, Kahrilas PJ, Savarino E, et al. Modern diagnosis of GERD: the Lyon Consensus. *Gut*. 2018; 67(7): 1351-1362.

6. Bjørsum-Meyer T, Koulaouzidis A, Baatrup G. Comment on "Artificial intelligence in gastroenterology: A state-of-the-art review." *World J Gastroenterol*. 2022; 28(16): 1722-1724.

7. Sharma P, Hassan C. Artificial Intelligence and Deep Learning for Upper Gastrointestinal Neoplasia. *Gastroenterology*. 2022; 162(4): 1056-1066. doi:10.1053/j.gastro.2021.11.040.

8. Chiu YC, Chen WL, Yang TW, Tsai MC, Tseng MH. A Deep Learning Model for Classification of Endoscopic Gastroesophageal Reflux Disease. *Int J Environ Res Public Health*. 2021; 18(5): 2428.

9. Basford PJ, Brown J, Gadeke L, et al. A randomized controlled trial of pre-procedure simethicone and N-acetylcysteine to improve mucosal visibility during gastroscopy – NICEVIS. *Endosc Int Open*. 2016; 4(11): E1197-E1202. doi:10.1055/s-0042-117631.

10. Sami S, Rangunath K. The Los Angeles

Classification of Gastroesophageal Reflux Disease. *Video Journal and Encyclopedia of GI Endoscopy*. 2013; 1:103-104. doi:10.1016/S2212-0971(13)70046-3.

11. Ge H, Zhou X, Wang Y, et al. Development and Validation of Deep Learning Models for the Multiclassification of Reflux Esophagitis Based on the Los Angeles Classification. *Journal of Healthcare Engineering*. 2023; 2023: e7023731. doi:10.1155/2023/7023731.

12. Visaggi P, Barberio B, Gregori D, et al. Systematic review with meta-analysis: artificial intelligence in the diagnosis of oesophageal

diseases. *Aliment Pharmacol Ther*. 2022; 55(5): 528-540.

13. Kusano M, Ino K, Yamada T, et al. Interobserver and intraobserver variation in endoscopic assessment of GERD using the “Los Angeles” classification. *Gastrointestinal Endoscopy*. 1999; 49(6): 700-704. doi:10.1016/S0016-5107(99)70285-3.

14. Hoshihara Y, Kogure T. What are longitudinal vessels? Endoscopic observation and clinical significance of longitudinal vessels in the lower esophagus. *Esophagus*. 2006; 3(4): 145-150. doi:10.1007/s10388-006-0096-2.

Summary

APPLICATION OF ARTIFICIAL INTELLIGENCE IN DETECTING REFLUX ESOPHAGITIS ON ENDOSCOPY IMAGES

Objectives: The purpose of this study was to evaluate the accuracy of an AI algorithm in detecting reflux esophagitis and to investigate factors related to missed lesions and errors. **Methods:** A cross-sectional descriptive study was conducted. The algorithm was tested on a dataset consisting of 1000 endoscopic images with various lighting modes, and the results were compared to the standard expert labeling. Accuracy was assessed using sensitivity, specificity, positive predictive value, and negative predictive value. Regression models were employed to investigate factors influencing the rate of missed and misdiagnosed lesions. **Result:** The algorithm achieved an accuracy rate of 81.7%. The analysis revealed that the number and size of lesions in the images were significantly associated with the rate of missed detections, while accompanying damage and image cleanliness levels were related to misdiagnosed lesions. **Conclusion:** The Yolo v8 algorithm demonstrates high accuracy and holds potential for further development, including real-time co-assistance during endoscopy, post-endoscopic examination, and medical training using large datasets.

Keywords: Artificial intelligence, upper gastrointestinal endoscopy, reflux esophagitis, lesion localization.