

NGHIÊN CỨU ĐÁNH GIÁ ĐỘ CHÍNH XÁC CỦA THUẬT TOÁN EFFICIENTNET TRONG PHÁT HIỆN VÀ KHOANH VÙNG POLYP ĐẠI TRÀNG

Đào Việt Hằng^{1,2,3}, Nguyễn Thanh Tùng¹,
Lâm Ngọc Hoa¹, Nguyễn Phúc Bình¹, Đào Văn Long^{1,2,3},
Nguyễn Thị Thủy⁴, Đinh Việt Sang⁵

TÓM TẮT

Mục tiêu của nghiên cứu: đánh giá độ chính xác của thuật toán EfficientNet trong phát hiện polyp đại tràng và khảo sát các yếu tố liên quan đến tỷ lệ bỏ sót, nhận nhầm. **Phương pháp nghiên cứu:** Nghiên cứu mô tả cắt ngang. Thuật toán EfficientNet được kiểm chứng trên tập ảnh tĩnh bao gồm 4000 ảnh (2000 ảnh có polyp với 2111 tổn thương, 2000 ảnh không có polyp) bằng cách so sánh với phần gán nhãn chuẩn của chuyên gia. Độ chính xác được đánh giá bằng độ nhạy (Se), độ đặc hiệu (Sp), giá trị dự báo dương tính (PPV), giá trị dự báo âm tính (NPV). Mô hình hồi quy được sử dụng để khảo sát những yếu tố liên quan đến tỉ lệ bỏ sót và nhận nhầm. **Kết quả:** Se, PPV, Sp, NPV và độ chính xác lần lượt là 97,60%, 94,44%, 94,25%, 97,52% và 95,93%. Phân tích hồi quy đa biến cho thấy độ sạch, kích thước và số lượng vùng polyp có liên quan tới tỷ lệ bỏ sót, độ sạch và chẩn đoán có liên quan đến tỷ lệ nhận nhầm. **Kết luận:** Thuật toán EfficientNet có độ chính xác cao, có thể phát triển theo hướng đồng kiểm với bác sĩ nội soi trong quá trình soi và có thể tham gia huấn luyện, đào tạo y khoa với dữ liệu lớn.

Từ khóa: Trí tuệ nhân tạo, EfficientNet, nội soi đại tràng, phát hiện polyp, khoanh vùng polyp.

SUMMARY

EVALUATING THE ACCURACY OF EFFICIENTNET IN COLON POLYP DETECTION AND DELINEATION

Objective: to evaluate the accuracy of EfficientNet algorithm in detecting colon polyps and to determine factors associated with the rate of missed polyp and false detection. **Methods:** Cross-sectional study. EfficientNet algorithm was validated on a set of 4000 still image (2000 images containing 2111 polyps, 2000 images with no polyp) by comparing with the ground-truth delineated by experts. Accuracy was assessed by sensitivity (Se), specificity (Sp), positive predictive value (PPV) and negative predictive value (NPV). Regression models were used to determine factors

related to the rate of missed polyp and false detection. **Results:** Se, PPV, Sp, NPV and accuracy were 97.60%, 94.44%, 94.25%, 97.52% and 95.93%, respectively. Multivariate regression analysis showed that cleanliness, polyp's size and number of polyps on image were significantly associated with the missed rate; cleanliness and diagnosis were related to the false detection rate. **Conclusion:** EfficientNet algorithm had high accuracy, can be further developed using big data to support endoscopists during endoscopy or improve endoscopy and medical training.

Key words: Artificial intelligent, EfficientNet, colonoscopy, polyp detection, polyp localization.

I. ĐẶT VẤN ĐỀ

Ung thư đại trực tràng (UTĐTT) tại Việt Nam đứng hàng thứ năm cả về tỉ lệ mắc và là nguyên nhân gây tử vong do bệnh lý ác tính theo GLOBOCAN 2020 [2]. Khoảng 85% trường hợp UTĐTT phát triển từ tổn thương u tuyến đại tràng thông qua các thay đổi di truyền, vì vậy việc tầm soát, phát hiện và loại bỏ polyp đại trực tràng sẽ làm giảm tỷ lệ mắc UTĐTT [5]. Bằng chứng đã cho thấy rằng mỗi tỷ lệ phát hiện u tuyến tăng 1% dẫn đến giảm 3% nguy cơ UTĐTT [8]. Ứng dụng công nghệ trong đó có thuật toán trí tuệ nhân tạo (AI) là một hướng đi có nhiều hứa hẹn để làm giảm tỉ lệ bỏ sót polyp đại tràng.

Năm 2020, Viện Nghiên cứu và Đào tạo Tiêu hóa, Gan mật phối hợp với các chuyên gia công nghệ thông tin đã xây dựng thuật toán EfficientNet được huấn luyện từ 8.190 hình ảnh nội soi đại tràng có polyp được khoanh vùng, gán nhãn bởi chuyên gia nội soi và 4.000 hình ảnh nội soi đại tràng không có polyp. Đây là tập dữ liệu đầu tiên được xây dựng tại Việt Nam, bao gồm 76% ảnh chế độ WLI và 23% chế độ FICE với kích thước, hình thái polyp đa dạng. Kết quả ban đầu khi thử nghiệm trên tập ảnh gồm 1321 ảnh với 1543 tổn thương ghi nhận độ nhạy và giá trị dự báo dương tính là 96,39% và 94,6% [1]. Tuy nhiên tập ảnh thử nghiệm ban đầu đó chưa bao gồm nhóm ảnh không có tổn thương. Vì vậy chúng tôi tiến hành nghiên cứu tiếp theo để kiểm chứng độ chính xác của thuật toán với 2 mục tiêu:

1. Đánh giá độ chính xác của thuật toán EfficientNet trong phát hiện polyp đại trực tràng.
2. Mô tả một số yếu tố liên quan đến tỷ lệ

¹Viện Nghiên cứu và Đào tạo Tiêu hóa, Gan mật

²Bộ môn Nội tổng hợp, Trường Đại học Y Hà Nội

³Trung tâm Nội soi, Bệnh viện Đại học Y Hà Nội

⁴Học viện Nông nghiệp Việt Nam

⁵Trường Đại học Bách khoa Hà Nội

Chịu trách nhiệm chính: Đào Việt Hằng

Email: hangdao.fsh@gmail.com

Ngày nhận bài: 9.9.2021

Ngày phản biện khoa học: 2.11.2021

Ngày duyệt bài: 11.11.2021

phát hiện polyp đại trực tràng của thuật toán EfficientNet: độ sạch, vị trí polyp, hình dạng, kích thước polyp.

II. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

Thiết kế nghiên cứu: mô tả cắt ngang.

Nơi tiến hành nghiên cứu: Viện Nghiên cứu và Đào tạo Tiêu hóa, Gan mật.

Tiêu chuẩn lựa chọn. Hình ảnh nội soi được thu thập ở chế độ ánh sáng trắng sử dụng hệ thống máy nội soi độ phân giải cao của Fujifilm (6000, 7000 và LASEREO). Hình ảnh được lấy ở nhiều đoạn khác nhau của đại tràng, có độ phân giải ảnh tối thiểu 720 x1280 pixel và mức độ sạch theo thang điểm Boston [3] là ≥ 2 điểm.

Tiêu chuẩn loại trừ. Loại bỏ những hình ảnh quá tối, lóa, nhòe hoặc mờ, không xác định được rõ ranh giới và hình thái polyp hoặc có UTĐTT, bệnh lý đa polyp có tính chất gia đình, bệnh lý viêm ruột mạn tính kèm theo trên hình ảnh nội soi.

Quá trình chọn mẫu và cỡ mẫu. Quần thể nghiên cứu: Là hình ảnh nội soi đại tràng được thu thập tại Viện Nghiên cứu và Đào tạo Tiêu hóa Gan mật và Trung tâm Nội Soi – Bệnh viện Đại học Y Hà Nội từ tháng 01/2021 đến tháng

04/2021. Phương pháp: chọn mẫu thuận tiện. Tập ảnh để đánh giá độ chính xác của thuật toán EfficientNet bao gồm 2000 ảnh có polyp được gán nhãn và khoanh vùng và 2000 ảnh không có polyp 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 đánh giá mức độ sạch theo thang điểm Boston, vị trí, hình thái polyp theo phân loại Paris [1].

Đánh giá độ chính xác của thuật toán, khảo sát các yếu tố liên quan. Sử dụng 4000 hình ảnh 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 2.1). Trong quá trình thuật toán xử lý, chúng tôi nhận thấy có một số trường hợp xuất hiện dương tính giả trong tập ảnh polyp (nhận nhầm vùng không có tổn thương thành polyp), những trường hợp này được quy về các trường hợp AI nhận định sai.

Mô hình hồi quy logistic được sử dụng để khảo sát những yếu tố liên quan đến tỉ lệ bỏ sót và nhận nhầm với vị trí, độ sạch, kích thước, hình thái và bề mặt của polyp.

Bảng 2.1. Cách tính các chỉ số nghiên cứu

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

Trong đó: a: Số ảnh polyp AI nhận định đúng; b: Số ảnh không có polyp AI nhận định sai; c: Số ảnh polyp AI nhận định sai; d: Số ảnh không có polyp AI nhận định đúng.

III. KẾT QUẢ NGHIÊN CỨU

Từ tháng 03/2021 đến tháng 06/2021, nghiên cứu thu thập được 2000 hình ảnh tương ứng với 2111 vùng polyp đã được chuyên gia nhận định

và 2000 hình ảnh của niêm mạc đại trực tràng bình thường hoặc viêm, loét, chứa dụng cụ nội soi, có túi thừa. Các đặc điểm mô tả chi tiết trong các bảng dưới đây:

Bảng 3.1. Phân bố các đặc điểm vùng polyp trong tập polyp (n=2111)

Vị trí	n	%	Hình thái	n	%
Manh tràng	42	2,0	Ip	400	18,9
Đại tràng lên	344	16,3	Is	1454	68,9
Đại tràng ngang	416	19,7	IIa	246	11,7
Đại tràng xuống	388	18,4	IIC	11	0,5
Đại tràng sigma	581	27,5	Số polyp		
Trực tràng	340	16,1	1	1906	90,3
Độ sạch			>2	205	9,7
2	415	19,7	Bề mặt		
3	1696	80,3	Nhẵn	1953	92,5
Kích thước			Sùi, chia múi	134	6,3

≤5	816	38,7		Loét, xuất tiết	24	1,1
6-10mm	1064	50,4				
>10mm	231	10,9				

Trong tệp polyp, vị trí được thu thập nhiều nhất là đại tràng sigma, ít nhất là manh tràng. Độ sạch đại tràng phần lớn là BBPS=3, các vùng đại tràng phải có độ sạch kém hơn so với đại tràng ngang và đại tràng trái. Đa số polyp kích thước không quá 10mm. Bề mặt polyp chủ yếu là nhẵn, hình thái chủ yếu là Paris Is, không có polyp dạng IIb và có 11 vùng polyp dạng IIc.

Trong tệp không có polyp, vị trí nhiều nhất là

đại tràng lên (bao gồm cả đại tràng góc gan) với 20,8%, ít nhất là manh tràng với 12,5%. Độ sạch đại tràng Boston=3 chiếm phần lớn (67,3%), các đoạn đại tràng có Boston=2 chiếm đa số ở đại tràng phải và đại tràng xuống. Chẩn đoán bình thường chiếm 94,8%, các chẩn đoán khác (gồm viêm, loét, túi thừa) hoặc có dụng cụ nội soi trong ảnh chỉ chiếm 5,2%.

Bảng 3.2. Mô tả đặc điểm các vùng ảnh khoan nhằm và bỏ sót tổn thương

Lỗi khoan nhằm	Tệp polyp	Tệp không polyp	Lỗi bỏ sót	Số polyp	%
Bọt, nhầy, phân	4	22	Polyp nhỏ, ở xa	24	75
Nếp niêm mạc	11	48	Polyp trên nếp niêm mạc	5	15,6
Vùng lóa, mờ	1	21	Polyp ở vùng tối, lóa	3	9,4
Van bauhin	0	14	Tổng	32	100
Viêm, loét	0	4			
Chân cắt polyp	1	1	Lỗi khoan chưa chính xác		
Cuống polyp	1	0	Lỗi khoan thừa	39	78,0
Lỗ ruột thừa	0	3	Lỗi khoan thiếu	11	22,0
Túi thừa	0	1			
Nghi ngờ polyp	0	3			
Tổng	18	117			

Thuật toán AI đã mô tả 2214 vùng, trong đó AI phát hiện chính xác 2079 vùng polyp trong số 2111 vùng polyp thực. Có 50 vùng phát hiện đúng tổn thương nhưng khoan viên chưa chính xác (39 lỗi khoan thừa và 11 lỗi khoan thiếu), 32 lỗi bỏ sót (30 ảnh). Có 18 ảnh khoan nhằm 18 tổn thương trong tệp polyp và 115 ảnh khoan nhằm 117 tổn thương trong tệp không có polyp. Se, PPV, Sp, NPV và độ chính xác lần lượt là 97,60%, 94,44%, 94,25%, 97,52% và 95,93%.

Bảng 3.4. Mối liên quan giữa các yếu tố với tỷ lệ bỏ sót và nhận nhằm polyp

Đặc điểm polyp bị bỏ sót	OR	95%CI	p
Nằm trong vùng nhiều polyp	42,86	14,65 - 125,40	0,000
Kích thước ≤5mm	10,64	2,42-46,68	0,002
BBPS=2	2,34	1,026 - 5,33	0,043
Đặc điểm vùng ảnh bị nhận nhằm trong tệp không có polyp			
BBPS=2	1,61	1,09 - 2,38	0,016
Chẩn đoán (nhóm tham chiếu: Bình thường)			
Viêm, loét	3,19	1,46 - 6,70	0,004
Khác	2,26	0,87 - 5,86	0,095

Phân tích hồi quy đa biến cho thấy các polyp kích thước ≤5mm, các polyp nằm trong ảnh chứa nhiều hơn 1 polyp và các polyp nằm trong vùng có mức BBPS=2 có liên quan tới tỷ lệ bỏ sót. Các ảnh có độ sạch kém và có chẩn đoán viêm, loét làm tăng tỷ lệ nhận nhằm.

IV. BÀN LUẬN

Mục tiêu thứ nhất của nghiên cứu này là đánh giá thuật toán AI đã xây dựng với 2000 hình ảnh nội soi chứa polyp và 2000 hình ảnh nội soi không có polyp bao gồm hình ảnh bình thường, tổn thương viêm, loét, túi thừa, dụng cụ nội soi. Với sự xuất hiện đa dạng của các hình ảnh nội soi đại tràng sẽ là công cụ đánh giá thuật toán khách quan và chính xác hơn. Kết quả nghiên

cứ của chúng tôi tương tự khi đối chiếu với các nghiên cứu lớn trên thế giới như: Ozawa và cộng sự (1246 tổn thương với Se 92% và PPV 86%) [5], Pu Wang (5,541 ảnh có polyp, 21572 ảnh không có polyp, Se 94,38%; Sp 95,92%) [8]. Với độ chính xác cao, thuật toán hoàn toàn có thể phát triển theo hướng đồng kiểm với bác sĩ ngay trong cuộc nội soi hoặc hậu kiểm ngay 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 thuật toán AI còn có thể tham gia huấn luyện, đào tạo y khoa với dữ liệu lớn.

Trong số polyp bị bỏ sót, chúng tôi đặc biệt quan tâm đến vị trí và kích thước polyp bị bỏ sót. Theo Valarini [7], polyp ở vị trí đại tràng trái, trực tràng và tổn thương >10mm có liên quan đến loạn sản và ung thư biểu mô tuyến. Trong nghiên cứu, tỷ lệ bỏ sót ở đại tràng trái chỉ bằng khoảng 1/3 so với đại tràng phải ($p = 0,018$). Ngoài trừ 2 polyp kích thước trên 5 mm ở vùng tối và xa, các polyp bị bỏ sót còn lại đều có kích thước ≤ 5 mm với hình thái Paris là Is, IIa. Polyp ở đại tràng phải có tỷ lệ bỏ sót nhiều hơn phù hợp với nghiên cứu trước đó trên bác sĩ nội soi [6], có thể do các vị trí này có đặc điểm đi dây khó, độ sạch thấp và nhiều nếp gấp niêm mạc. Các ảnh có điểm BBPS=2 có tỷ lệ bỏ sót cao hơn 2 lần so với mức BBPS=3. Không có sự khác biệt về tỉ lệ bỏ sót giữa các polyp dạng gồi (Ip, Is) và dạng phẳng (IIa, IIc) với $p = 0,1$, tuy nhiên đối với các polyp có tổn thương loét, xuất tiết thì thuật toán AI phát hiện chính xác nhưng khi khoan vùng xảy ra tình trạng khoan thừa và khoan thiếu.

Ở tập không có polyp, đại tràng phải có tỷ lệ nhận nhầm là 10%, cao hơn có ý nghĩa so với đại tràng trái là 4,5% ($p = 0,001$), độ sạch BBPS=2 cũng có tỷ lệ nhận nhầm cao hơn so với mức BBPS=3 (7,3% so với 5%, $p = 0,033$). Trong đó, hơn một phần ba tổn thương loét đã bị nhận nhầm. Ở tập polyp, số dương tính giả thấp hơn rất nhiều tập không có polyp (18 so với 117), có thể lý do là những ảnh polyp thường được làm sạch khi bác sĩ nội soi tiến hành chụp và đoạn đại tràng cũng được bơm đủ căng. Mặt khác, lỗi nhận nhầm chủ yếu là nếp niêm mạc (59 trường hợp ở cả 2 tập), như vậy có thể suy ra nguyên nhân chính dẫn đến nhận nhầm là do đại tràng bơm chưa đủ căng để quan sát, bên cạnh chất lượng ảnh kém (độ sạch kém, mờ, lóa). Do đó, AI sẽ không thể bù đắp cho chất lượng nội soi đại tràng dưới mức tối ưu, mức độ làm sạch không tốt, thời gian rút dây soi vội vàng. Tất cả những yếu tố này vẫn là điều kiện tiên quyết để tối đa hóa độ chính xác của AI.

Trong hồi quy đa biến, yếu tố liên quan độc lập đến tỷ lệ bỏ sót của thuật toán là số polyp trong ảnh và kích thước polyp. Kết quả phù hợp với nghiên cứu của Thomas Lui [4], yếu tố liên quan chặt chẽ với tỷ lệ bỏ sót là số lượng u tuyến trên nội soi ban đầu. Các vùng dương tính

giả hay gặp ở vị trí đại tràng có độ sạch kém và các vùng có tổn thương viêm loét, những vùng này tương đối dễ nhận biết bằng mắt thường.

Tuy nhiên, nghiên cứu còn một số hạn chế, bao gồm sử dụng tập ảnh nội soi đại tràng tĩnh với ánh sáng trắng chưa có các chế độ ánh sáng đa dạng và chưa kiểm định trên các video để đánh giá khả năng nhận định trong thời gian thực của thuật toán. Cũng cần thêm các nghiên cứu so sánh với các bác sĩ ở nhiều mức độ kinh nghiệm, làm việc tại các cơ sở y tế có quy mô khác nhau.

V. KẾT LUẬN

Kết quả kiểm chứng Thuật toán EfficientNet trên tập ảnh tĩnh bao gồm 2000 ảnh có polyp với 2111 tổn thương và 2000 ảnh không có polyp ghi nhận Se, PPV, Sp, NPV và độ chính xác lần lượt là 97,80%, 94,45%, 94,25%, 97,72% và 96,03%. Kích thước và số lượng vùng polyp là 2 yếu tố liên quan đáng kể với tỷ lệ bỏ sót, độ sạch và chẩn đoán có liên quan đến tỷ lệ nhận nhầm.

TÀI LIỆU THAM KHẢO

1. **Đào Việt Hằng và cộng sự. (2020).** Kết quả bước đầu của ứng dụng trí tuệ nhân tạo trong phát hiện polyp đại tràng tại Việt Nam. Tạp chí Nghiên cứu Y học, 130 (6), 101-110.
2. **International Agency for Research on Cancer (2020).** The Globocan Cancer Observatory. Country profile: Vietnam.
3. **Lai E.J., Calderwood A.H., Doros G. và cộng sự. (2009).** The Boston Bowel Preparation Scale: A valid and reliable instrument for colonoscopy-oriented research. *Gastrointest Endosc*, 69(3 Pt 2), 620-625.
4. **Lui T.K.L., Hui C.K.Y. và cộng sự. (2021).** New insights on missed colonic lesions during colonoscopy through artificial intelligence-assisted real-time detection (with video). *Gastrointest Endosc*, 93(1), 193-200.
5. **Ozawa T., Ishihara S. và cộng sự. (2020).** Automated endoscopic detection and classification of colorectal polyps using convolutional neural networks. *Therap Adv Gastroenterol*, 13, 1756284820910659.
6. **Seo J.H., Lee B.-I. và cộng sự. (2020).** Adenoma miss rate of polypectomy-referring hospitals is high in Korea. *Korean J Intern Med*, 35(4), 881-888.
7. **Valarini S.B.M. và cộng sự. (2011).** Correlation between location, size and histologic type of colorectal polyps at the presence of dysplasia and adenocarcinoma. *Journal of Coloproctology (Rio de Janeiro)*, 31(3), 241-247.
8. **Wang P., Xiao X., Glissen Brown J.R. và cộng sự. (2018).** Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. *Nat Biomed Eng*, 2(10), 741-748.