

APPLICATION OF MACHINE LEARNING IN DIAGNOSING JAW FRACTURES THROUGH CT SCANNERS: PERFORMANCE ANALYSIS WITH DIFFERENT PARAMETERS

Tran Tuan Anh^{1*}, Nguyen The Huy¹, Nguyen Thi Hoai Nhi¹, Tran Van Dang²
Tran Hoang Anh³, Bui Duy Phuc⁴

¹Becamex International Hospital - Go Cat area, Binh Duong avenue, Thuan An city, Binh Duong province, Vietnam

²Medic Binh Duong General Hospital - 14A Nguyen An Ninh, Thu Dau Mot city, Binh Duong province, Vietnam

³Binh Duong General Hospital - 5 Pham Ngoc Thach, Thu Dau Mot city, Binh Duong province, Vietnam

⁴Binh Duong province Center for Disease Control - 209, Yersin, Thu Dau Mot city, Binh Duong province, Vietnam

Received: 12/02/2025

Revised: 26/3/2025; Accepted: 11/4/2025

ABSTRACT

Objective: Apply Teachable Machine to detect jaw fractures in CT images.

Subjects and methods: Retrospective study on 1341 images extracted from CT.

Results: Among 746 images with jaw fracture injuries, correct identification occurred at a rate of 91.8% with parameter settings of 50-16-0.001, which decreased gradually to 82.4% when parameter levels were increased to 150:64:0.003. In the mixed set of 1341 images (with and without jaw fractures), the correct identification rate for images with jaw fractures was 87.3% at parameter levels of 50:16:0.001, decreasing to 78.7% when parameters were increased to 150:64:0.003. This demonstrates a correlation between the adjustment of parameter groups such as Epochs, Batch size, and Learning rate to achieve optimal performance, significantly improving accuracy and general prediction ability on data, while avoiding overfitting.

Keywords: Artificial intelligence, Teachable Machine, jaw fractures.

*Corresponding author

Email: Tstrantuananh@gmail.com **Phone:** (+84) 915713171 **https://doi.org/10.52163/yhc.v66iCD4.2337**

ỨNG DỤNG HỌC MÁY TRONG PHÁT HIỆN GỠ XƯƠNG HÀM QUA CT SCANNER: ĐÁNH GIÁ HIỆU SUẤT MÔ HÌNH VỚI CÁC THÔNG SỐ KHÁC NHAU

Trần Tuấn Anh^{1*}, Nguyễn Thế Huy¹, Nguyễn Thị Hoài Nhi¹, Trần Văn Đáng²
Trần Hoàng Anh³, Bùi Duy Phúc⁴

¹Bệnh viện Quốc tế Becamex - khu Gò Cát, đại lộ Bình Dương, TP Thuận An, tỉnh Bình Dương, Việt Nam

²Bệnh viện Đa khoa Medic Bình Dương - 14A Nguyễn An Ninh, TP Thủ Dầu Một, tỉnh Bình Dương, Việt Nam

³Bệnh viện Đa khoa tỉnh Bình Dương - 5 Phạm Ngọc Thạch, TP Thủ Dầu Một, tỉnh Bình Dương, Việt Nam

⁴Trung tâm Kiểm soát bệnh tật tỉnh Bình Dương - 209, Yersin, TP Thủ Dầu Một, tỉnh Bình Dương, Việt Nam

Ngày nhận bài: 12/02/2025

Ngày chỉnh sửa: 26/3/2025; Ngày duyệt đăng: 11/4/2025

TÓM TẮT

Mục tiêu: Ứng dụng học máy trong phát hiện gãy xương hàm trên ảnh CT scanner.

Đối tượng và phương pháp: Nghiên cứu hồi cứu trên 1341 hình ảnh CT.

Kết quả: Trong 746 ảnh có tổn thương gãy xương hàm, kết quả nhận dạng đúng chiếm tỷ lệ 91,8% ở mức tham số 50:16:0,001 và giảm dần xuống 82,4% khi tăng tham số lên mức 150:64:0,003. Trong 1341 ảnh hỗn hợp có gãy xương hàm và không có gãy xương, tỷ lệ nhận dạng đúng ảnh có gãy xương hàm chiếm 87,3% ở mức tham số 50:16:0,001 và giảm dần xuống 78,7% khi tăng tham số lên mức 150:64:0,003. Từ đó cho thấy có mối tương quan giữa việc điều chỉnh các nhóm thông số Epochs, Batch size và Learning rate để đạt được hiệu suất tốt nhất giúp cải thiện đáng kể độ chính xác và khả năng dự đoán tổng quát hóa trên dữ liệu, đồng thời tránh được tình trạng quá tải (overfitting).

Từ khóa: Trí tuệ nhân tạo, học máy, gãy xương hàm.

1. ĐẶT VẤN ĐỀ

Chấn thương hàm mặt nói chung và gãy xương hàm nói riêng thường rất đa dạng và phức tạp, không những ảnh hưởng tới tính mạng người bệnh mà còn để lại nhiều di chứng nặng nề và cả chức năng thẩm mỹ nếu không điều trị sớm [1]. Việc chẩn đoán sớm và đánh giá chính xác, đầy đủ tránh bỏ sót nhằm đưa ra phương pháp điều trị kịp thời là rất cần thiết, giúp bệnh nhân phục hồi hiệu quả, tránh biến chứng, di chứng nặng nề [2]. Hiện nay ở Việt Nam, để chẩn đoán chính xác gãy xương hàm do chấn thương gây ra thường sử dụng phim chụp cắt lớp vi tính (Computerized Tomography - CT) và chụp cộng hưởng từ (Magnetic Resonance Imaging) [3].

Ngày nay, với sự ra đời của trí tuệ nhân tạo (artificial intelligence - AI) và học máy (Teachable Machine) đã hỗ trợ trong rất nhiều lĩnh vực, trong đó có ngành y tế [4]. Gần đây đã có nhiều công trình nghiên cứu khác nhau trên thế giới như Gulshan V cùng đồng nghiệp (2016) đã phát triển thuật toán học sâu để phát hiện bệnh vồng mạc tiêu đường trong ảnh chụp đáy mắt [5]. Olczak J và cộng sự sử dụng AI để phân tích ảnh chụp X quang chấn thương chỉnh hình năm 2017 [6]. Tuy nhiên, số lượng nghiên cứu liên quan đến học máy dựa

trên dữ liệu CT để phát hiện gãy xương còn thấp. Chính vì thế, chúng tôi tiến hành nghiên cứu đề tài này nhằm xây dựng hệ dữ liệu hình ảnh từ máy CT scanner để tự động phát hiện gãy xương hàm và sử dụng chúng làm điểm khởi đầu cho một ngân hàng dữ liệu phong phú về chấn thương hàm về sau với ưu điểm là dễ dàng sử dụng, có thể cập nhật dữ liệu liên tục một cách dễ dàng, từ đó hỗ trợ bác sĩ giảm bớt áp lực công việc, giảm thời gian phân tích phim CT nhằm sớm đưa ra kết quả chính xác để kịp thời điều trị cho bệnh nhân. Tuy nhiên công nghệ AI không thể thay thế hoàn toàn được vai trò của bác sĩ, nhưng nhờ AI sẽ hỗ trợ bác sĩ đưa ra kết quả chính xác nhất trong thời gian ngắn nhất.

2. ĐỐI TƯỢNG, PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Đối tượng nghiên cứu

Đối tượng nghiên cứu là dữ liệu hình ảnh từ máy CT scanner của bệnh nhân được chẩn đoán gãy xương hàm.

- Tiêu chuẩn chọn: hình ảnh rõ nét, không mờ, không sai tư thế và được chẩn đoán là gãy xương hàm tại Bệnh viện Quốc tế Becamex từ tháng 9/2021 đến tháng 11/2024.

- Tiêu chuẩn loại trừ: không đạt tiêu chí chọn lựa.

*Tác giả liên hệ

Email: Tstrantuananh@gmail.com Điện thoại: (+84) 915713171 <https://doi.org/10.52163/yhc.v66iCD4.2337>

2.2. Phương pháp nghiên cứu

Thiết kế nghiên cứu: nghiên cứu mô tả hồi cứu số liệu.

Cỡ mẫu: 1341 ảnh trên 123 bệnh nhân, trong đó 71 bệnh nhân tương ứng với 746 lát/ảnh đã được chẩn đoán là có gãy xương hàm (các lát cắt Axial được chọn có độ dày mỗi lát cắt từ 0,5-2 mm tại vị trí phát hiện gãy xương tương đương), và 52 bệnh nhân (595 lát/ảnh) không gãy xương hàm để so sánh kiểm tra.

2.3. Xử lý và phân tích số liệu

- Chuyển đổi định dạng dữ liệu hình ảnh Dicom từ máy CT scanner, sau đó tách và chọn các lát cắt tại vị trí phát hiện gãy xương đưa vào ngân hàng dữ liệu của Teachable Machine. Dữ liệu đào tạo sẽ được điều chỉnh theo 3 nhóm thông số để so sánh mức độ chính xác:

+ Nhóm I: Epochs 50, Batch size 16, Learning rate 0,001.

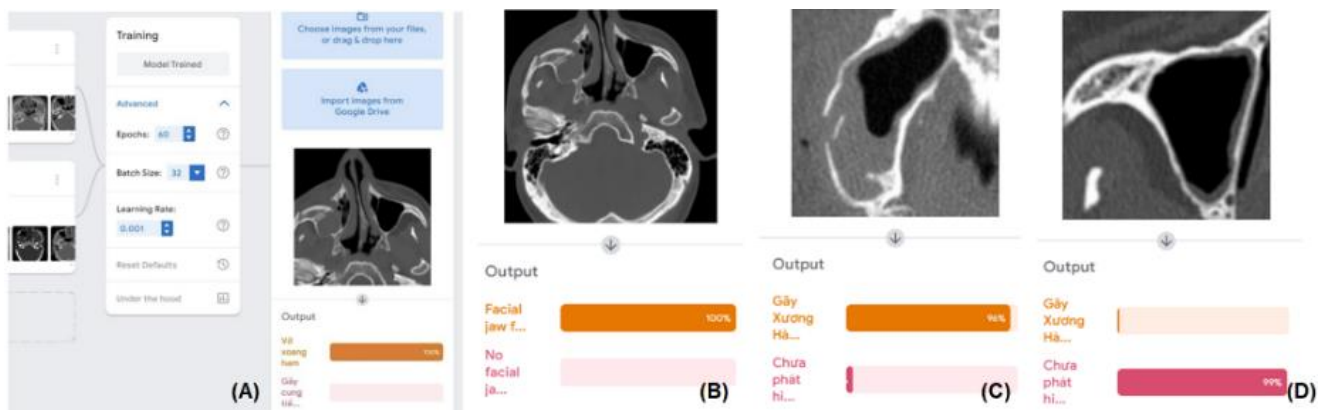
+ Nhóm II: Epochs 100, Batch size 32, Learning rate 0,002.

+ Nhóm III: Epochs 150, Batch size 64, Learning rate 0,003.

- Phân tích bằng phần mềm SPSS 22.0, thống kê mô tả sử dụng số lượng, tỷ lệ %. Thống kê phân tích dùng Chi-Square test.

2.4. Đạo đức nghiên cứu

Nghiên cứu đã được thông qua Hội đồng Nghiên cứu khoa học của Bệnh viện Quốc tế Becamex theo Quyết định số 804/QĐ-BIH ngày 29/5/2024.

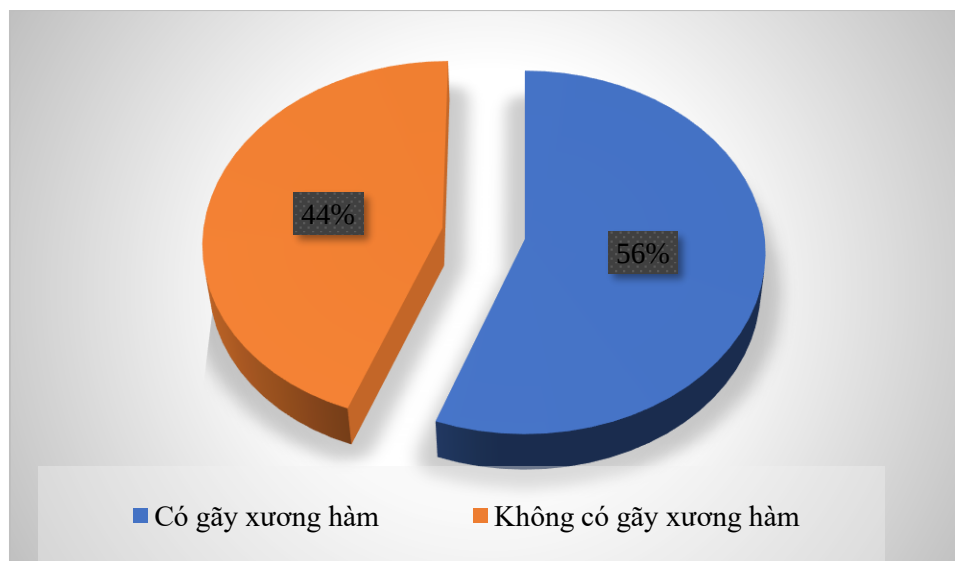


Hình 1. A: Giao diện trình huấn luyện trên Teachable Machine; B: Gãy xương hàm được nhận dạng 100%; C: Gãy xương hàm được nhận dạng 96%; D: Không gãy xương hàm được nhận dạng 99%

3. KẾT QUẢ NGHIÊN CỨU

3.1. Mẫu nghiên cứu

Trong nghiên cứu này, chúng tôi thực hiện với tổng cộng 1341 ảnh trên 123 bệnh nhân, trong đó 71 bệnh nhân tương ứng với 746 lát/ảnh đã được chẩn đoán là có gãy xương hàm (các lát cắt Axial được chọn có độ dày mỗi lát cắt từ 0,5-2 mm tại vị trí phát hiện gãy xương tương đương), và 52 bệnh nhân (595 lát/ảnh) không gãy xương hàm để so sánh kiểm tra.



Biểu đồ 1. Phân bố tỷ lệ dữ liệu ảnh nghiên cứu (n = 1341)

3.2. Tỷ lệ phát hiện có gãy xương hàm dựa trên hệ dữ liệu hình ảnh có gãy xương hàm với từng nhóm thông số đào tạo học máy

Bảng 1. Tỷ lệ phát hiện ảnh có gãy xương hàm theo các mức độ (n = 746)

Thông số máy học	Phát hiện gãy xương	Không phát hiện	p-value
Nhóm I (Epochs 50, Batch size 16, Learning rate 0,001) ⁽¹⁾	685 (91,8%)	61 (8,2%)	p ₁₋₃ < 0,05 p ₂₋₃ < 0,05 p ₁₋₂ > 0,05
Nhóm II (Epochs 100, Batch size 32, Learning rate 0,002) ⁽²⁾	688 (89,5%)	78 (10,5%)	
Nhóm III (Epochs 150, Batch size 64, Learning rate 0,003) ⁽³⁾	615 (82,4%)	131 (17,6%)	

Tăng thông số Epochs:Batch size:Learning rate từ 50:16:0,001 lên 150:64:0,003 làm giảm độ chính xác chẩn đoán từ 91,8% xuống 82,4% do overfitting. Kiểm định Chi-square test cho thấy sự khác biệt có ý nghĩa thống kê giữa 50:16:0,001 và 150:64:0,003, cũng như giữa 100:32:0,02 và 150:64:0,003 ($p < 0,05$), nhưng không có sự khác biệt giữa 50:16:0,001 và 100:32:0,02 ($p > 0,05$).

3.3. Tỷ lệ phát hiện có gãy xương hàm dựa trên hệ dữ liệu hình ảnh hỗn hợp với thông số đào tạo học máy

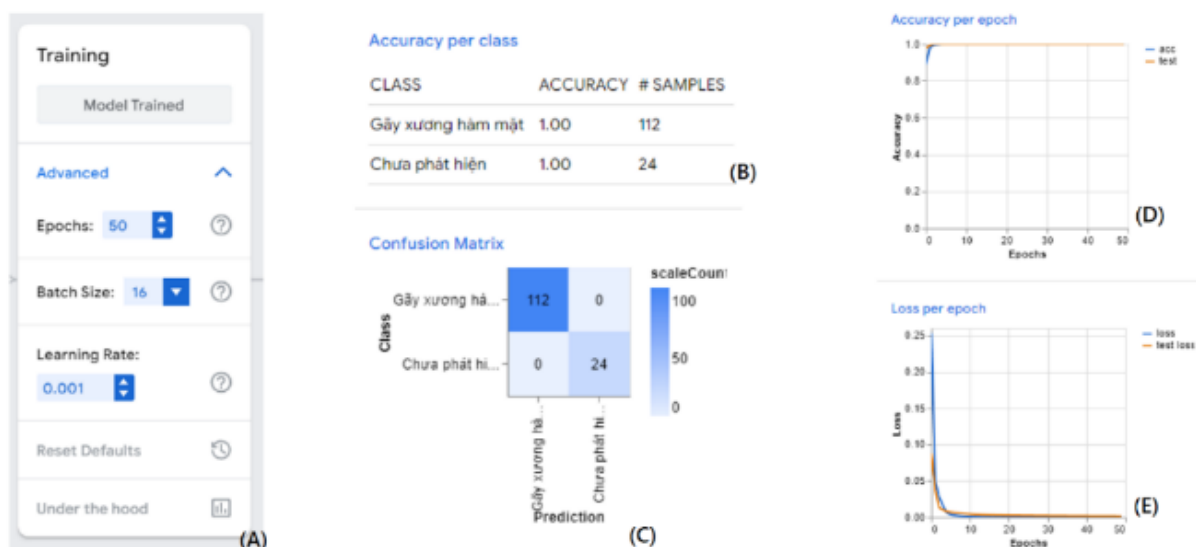
Bảng 2. Tỷ lệ phát hiện ảnh có gãy xương trong bộ dữ liệu ảnh hỗn hợp (n = 1341)

Thông số máy học	Phát hiện gãy xương	Không phát hiện	p-value
Mức độ 1 (Epochs 50, Batch size 16, Learning rate 0,001)	1171 (87,3%)	170 (12,7%)	< 0,05
Mức độ 2 (Epochs 100, Batch size 32; Learning rate 0,002)	1128 (84,1%)	213 (15,9%)	
Mức độ 3 (Epochs 150, Batch size 64, Learning rate 0,003)	1055 (78,7%)	286 (21,3%)	

Bảng 2 cho thấy độ chính xác phát hiện gãy xương hàm giảm khi tăng thông số từ Epochs 50:Batch size 16:Learning rate 0,001 (87,3%) lên Epochs 150:Batch size 64:Learning rate 0,002 (78,7%). Nguyên nhân có thể do giảm độ nhạy trong cập nhật trọng số, ảnh hưởng khả năng tổng quát hóa. Kiểm định Chi-Square test xác nhận sự khác biệt có ý nghĩa thống kê giữa các nhóm thông số ($p < 0,05$). Kết quả cho thấy việc lựa chọn thông số phù hợp là rất quan trọng để đạt được độ chính xác cao trong chẩn đoán hình ảnh y tế.

3.4. Kết quả chẩn đoán khi tối ưu hóa giữa các nhóm thông số cho máy học

3.4.1. Kết quả tối ưu hóa ở nhóm thông số Epochs 50:Batch size 16:Learning rate 0,001



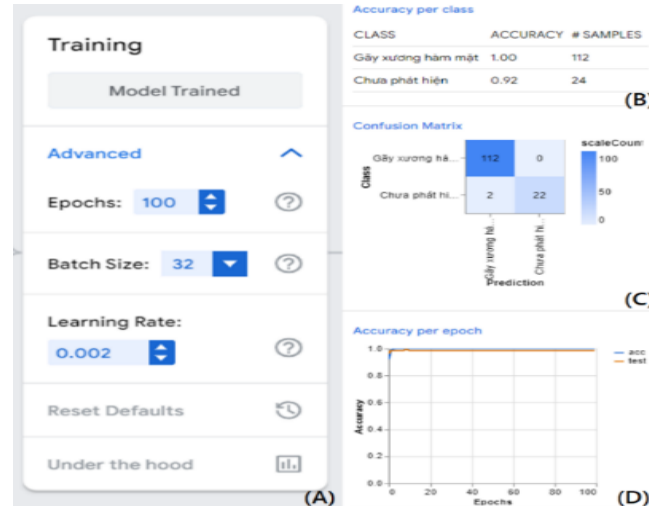
Hình 2. Độ chính xác ở thông số Epochs 50: Batch size 16: Learning rate 0,001

A: Bảng thông số huấn luyện máy học; B: Độ chính xác cho mỗi biến dự đoán; C: Ma trận phân loại các mẫu kiểm thử; D: Biểu đồ độ chính xác theo mỗi Epochs; E: Biểu đồ theo dõi mô hình mất qua mỗi Epoch

Trên hình B “Accuracy Per Class” cho thấy độ chính xác của mô hình theo từng lớp. Cả “Gãy xương hàm” và “Chưa phát hiện” đều đạt độ chính xác 1,00, chứng tỏ mô hình phân loại chính xác hoàn toàn. Hình C “Confusion

Matrix” thể hiện rằng không có sự nhầm lẫn giữa hai lớp trên. Hình D “Accuracy Per Epoch” hiển thị độ chính xác theo từng Epoch, với đồ thị ổn định ở mức cao, cho thấy mô hình học tốt hoặc dữ liệu không quá thách thức. Hình E “Loss Per Epoch” theo dõi tổn thất, cho thấy Loss giảm mạnh rồi ổn định, chứng tỏ mô hình học hiệu quả. Nhìn chung, kết quả đạt độ chính xác cao, cân bằng giữa hiệu quả học và tránh overfitting, nhưng có thể xem xét điều chỉnh tham số để cải thiện thêm.

3.4.2. Kết quả tối ưu hóa ở mức thông số Epochs 100:Batch size 32:Learning rate 0,002

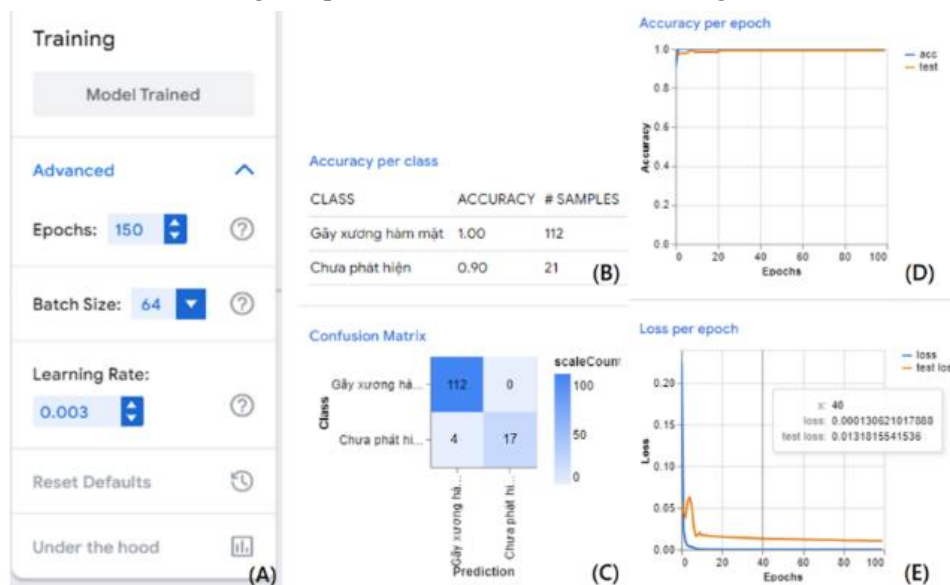


Hình 2. Độ chính xác ở thông số Epochs 100:Batch size 32:Learning rate 0,002

A: Bảng thông số huấn luyện máy học; B: Độ chính xác cho mỗi biến dự đoán; C: Ma trận phân loại các mẫu kiểm thử; D: Biểu đồ độ chính xác theo mỗi Epochs

Khi tăng thông số lên gấp đôi (100:32:0.02), độ chính xác giảm từ 1,00 xuống 0,92, cho thấy dấu hiệu overfitting. Máy học mất khả năng tổng quát hóa, gây nhầm lẫn giữa hai tình trạng (Confusion Matrix). Việc tăng mạnh Batch size và Epochs có thể không phải lựa chọn tối ưu cho nhận dạng này.

3.4.3. Kết quả tối ưu hóa ở mức thông số Epochs 150:Batch size 64:Learning rate 0,003



Hình 3. Độ chính xác ở thông số Epochs 150:Batch size 64:Learning rate 0,003

(A) Bảng thông số huấn luyện máy học; (B) Độ chính xác cho mỗi biến dự đoán; (C) Ma trận phân loại các mẫu kiểm thử; (D) Biểu đồ độ chính xác theo mỗi Epochs; (E) Biểu đồ theo dõi mô hình mất qua mỗi Epoch

Với thông số Epochs 150:Batch size 64:Learning rate 0,003, độ chính xác giảm so với 100:62:0.02, cho thấy ảnh hưởng của việc tăng tham số quá mức (thể hiện qua đường màu cam trong Loss Per Epoch). Learning rate

0,003 có thể khiến máy học quá nhanh, bỏ sót đặc điểm quan trọng. Tăng quá nhiều Epochs và Batch size cũng làm giảm hiệu quả huấn luyện do máy ghi nhớ dữ liệu thay vì tổng quát hóa.

4. BÀN LUẬN

Việc đánh giá hình ảnh CT bước đầu ở bệnh nhân chấn thương hàm mặt là cơ sở để lựa chọn phương pháp điều trị phù hợp và hướng họ đến các đơn vị chuyên môn cao hơn nếu cần thiết. Khi chấn thương xảy ra ở vùng đầu mặt vốn dĩ đã có độ phức tạp về mặt giải phẫu, nên khả năng không nhận ra hay bỏ sót các vị trí gãy xương có thể xảy ra sẽ làm tăng chi phí, thời gian cho hệ thống y tế và tạo gánh nặng cho bệnh nhân do phải nhập viện không cần thiết tại các chuyên khoa chưa phù hợp. Mặc dù có một số ứng dụng AI cũng đã được ứng dụng bước đầu trong lĩnh vực chỉnh hình nhưng chúng vẫn chưa được khám phá ở khu vực hàm mặt. Hệ thống chẩn đoán hình ảnh dựa trên AI sẽ cho phép giảm thiểu sai sót chẩn đoán bằng cách cung cấp cho bác sĩ lâm sàng và bác sĩ chẩn đoán hình ảnh một công cụ hỗ trợ để hướng tới các lựa chọn điều trị phù hợp. Tuy nhiên, hệ thống chẩn đoán hình ảnh dựa trên AI không thể thay thế công việc của bác sĩ mà sẽ trở thành một công nghệ hỗ trợ có giá trị để giảm rủi ro sai sót y tế, vận chuyển, nhập viện không cần thiết, giảm gánh nặng kinh tế cho xã hội và quản lý y tế công cộng.

Nghiên cứu của chúng tôi cho kết quả nhận dạng được gãy xương hàm chính xác đạt đến 91,8% thông qua điều chỉnh các nhóm tham số. Điều kiện tối ưu để máy học tốt nhất là Epochs 50:Batch size 16:Learning rate 0,001, kết quả này tương đồng so với kết quả nghiên cứu Szegedy C và cộng sự (2016) về việc ứng dụng bộ công cụ học máy Teachable Machine để nhận dạng lưỡi có dấu răng và lưỡi không có dấu răng, tỷ lệ nhận dạng chính xác 92,1% [7]. Warin K và cộng sự (2022) nghiên cứu phát triển các mô hình nhận dạng tự động để xác định gãy xương hàm dưới trong chụp X quang toàn cảnh bằng thuật toán mạng thần kinh chập (Convolutional Neural Network - CNN) bao gồm 1710 hình ảnh X quang, trong đó có 855 hình ảnh được xác định là gãy xương hàm dưới, kết quả cho thấy độ nhạy và độ đặc hiệu là 100%, hiệu suất phát hiện các tổn thương gãy xương đạt khoảng 90% [8]. So với nghiên cứu của Xiaoqiang Li và cộng sự (2018) khi sử dụng tính năng học tập đa phiên bản và thuật toán mạng thần kinh chập (CNN) nhận dạng hình ảnh lưỡi có dấu răng cho tỷ lệ nhận dạng đúng 72,7% [9]. So với các mô hình máy học khác như CNN dùng trong nhận dạng hình ảnh, kết quả học của Teachable Machine cao hơn, khẳng định tính khả thi khi sử dụng Teachable Machine trong việc nhận dạng gãy xương hàm và các tổn thương trên hình ảnh khác có thể bước đầu nghiên cứu ứng dụng trong một số lĩnh vực trong y học. Các nghiên cứu trong tương lai có thể tập trung vào việc phát hiện vết nứt nhỏ trên xương, cải thiện chất lượng và số lượng dữ liệu hình ảnh để phát hiện, chẳng hạn như hình ảnh các góc của xương bị gãy để cải thiện độ nhạy phát hiện của hệ thống. Hơn nữa, để nâng cao hiệu suất nhận

dạng, có thể đưa ra thêm một giai đoạn tiền xử lý ảnh CT scanner để loại bỏ vùng không cần thiết tập trung vào vùng gãy để dự đoán.

5. KẾT LUẬN

Nghiên cứu trên 746 ảnh CT gãy xương hàm bằng Teachable Machine cho thấy độ chính xác 91,8% ở mức tham số 50:16:0,001, giảm xuống 82,4% khi tăng tham số lên 150:64:0,003. Trên 1341 ảnh hỗn hợp, tỷ lệ nhận dạng đúng ảnh gãy xương là 87,3%, giảm còn 78,7% khi tăng tham số. Sự khác biệt giữa hai mức tham số có ý nghĩa thống kê ($p < 0,05$), cho thấy mối tương quan giữa Epochs, Batch size và Learning rate. Việc tối ưu hóa tham số giúp cải thiện độ chính xác, khả năng tổng quát hóa và tránh được tình trạng quá tải (overfitting).

TÀI LIỆU THAM KHẢO

- [1] Abosadegh M.M et al, Epidemiology of maxillofacial fractures at a teaching hospital in Malaysia: a retrospective study, BioMed research international, 2019.
- [2] Trần Văn Trường, Trương Mạnh Dũng, Tình hình chấn thương hàm mặt tại Viện Răng Hàm Mặt Hà Nội trong 11 năm (1988-1998), Tạp chí Y học Việt Nam, 1999, 10, tr. 71-80.
- [3] Kalmet P.H et al, Deep learning in fracture detection: a narrative review. Acta orthopaedica, 2020, 91 (2), p. 215-220.
- [4] Chamunyonga C et al, The impact of artificial intelligence and Teachable Machine in radiation therapy: considerations for future curriculum enhancement, Journal of Medical Imaging and Radiation Sciences, 2020, 51 (2): p. 214-220.
- [5] Gulshan V et al, Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs, Jama, 2016, 316 (22), p. 2402-2410.
- [6] Olczak J et al, Artificial intelligence for analyzing orthopedic trauma radiographs: deep learning algorithms-are they on par with humans for diagnosing fractures? Acta orthopaedica, 2017, 88 (6), p. 581-586.
- [7] Szegedy C et al, Rethinking the Inception Architecture for Computer Vision. In Proceedings of the IEEE conference on computer vision and pattern recognition, June 2016, DOI:10.1109/CVPR.2016.308.
- [8] Warin K et al, Assessment of deep convolutional neural network models for mandibular fracture detection in panoramic radiographs, International Journal of Oral and Maxillofacial Surgery, 2022, 51 (11), p. 1488-1494.
- [9] Li X et al, Tooth-marked tongue recognition using multiple instance learning and CNN features, IEEE transactions on cybernetics, 2018, 49 (2) p. 380-387.