

NGHIÊN CỨU MÔ HÌNH HỌC SÂU FASTER R-CNN ĐỂ PHÁT HIỆN VÀ PHÂN LOẠI CÁC TỔN THƯƠNG KHU TRÚ THƯỜNG GẶP Ở GAN TRÊN ẢNH CHỤP CẮT LỚP VI TÍNH

*Nguyễn Hoàng Thái**, *Phù Trí Nghĩa*, *Dương Quốc Thanh*,
Nguyễn Thanh Hùng, *Mai Quốc Trường*, *Trần Thị Bích Phương*,
Cao Tấn Phát, *Huỳnh Đăng Khoa*
Trường Đại học Y Dược Cần Thơ
* Email: nhthai@ctump.edu.vn

TÓM TẮT

Đặt vấn đề: Sự cần thiết của việc phát hiện, phân loại sớm tổn thương gan và việc nghiên cứu ứng dụng mô hình học sâu vào xử lý ảnh y khoa. **Mục tiêu nghiên cứu:** Thu thập dữ liệu, xây dựng, huấn luyện mô hình Faster R-CNN để phát hiện, phân loại các tổn thương khu trú thường gặp ở gan; Kiểm thử, đánh giá hiệu quả mô hình theo tiêu chí về thời gian, độ chính xác. **Đối tượng và phương pháp nghiên cứu:** Bộ dữ liệu ảnh chụp cắt lớp vi tính tiêm thuốc cản quang vùng bụng có tổn thương gồm nang gan, u mạch máu, ung thư tế bào gan nguyên phát; Áp dụng mô hình Faster R-CNN để phát hiện, phân loại tổn thương. **Kết quả:** Bộ dữ liệu thu thập tại Bệnh viện Trường Đại học Y Dược Cần Thơ gồm 51 người bệnh có tổn thương khu trú thường gặp ở gan, với 2828 ảnh, 2836 vùng tổn thương được xác định bởi bác sĩ chẩn đoán hình ảnh. Trong đó, 11 người bệnh thuộc nhóm nang gan (440 vùng), 18 người bệnh thuộc nhóm u mạch máu (648 vùng), 21 người bệnh thuộc nhóm ung thư tế bào gan nguyên phát (1748 vùng) và 01 người bệnh có cả nang và u mạch máu; Mô hình Faster R-CNN cho kết quả độ chính xác mAP là 94%, thời gian huấn luyện 583 phút và thời gian xử lý 0,13 giây. **Kết luận:** Bộ dữ liệu đã thu thập là nền tảng cho các nghiên cứu tiếp theo; Mô hình Faster R-CNN có thời gian huấn luyện ngắn, thời gian xử lý nhanh, độ chính xác cao, phù hợp và có thể áp dụng để triển khai các ứng dụng thực tế.

Từ khóa: Máy học, Học sâu, Faster R-CNN, Chụp cắt lớp vi tính, Bệnh gan.

ABSTRACT

STUDYING DEEP LEARNING MODELS TO DETECT AND CLASSIFY COMMON FOCAL LIVER LESIONS ON CT IMAGES

*Nguyen Hoang Thai**, *Phu Tri Nghia*, *Duong Quoc Thanh*,
Nguyen Thanh Hung, *Mai Quoc Truong*, *Tran Thi Bich*
Phuong, *Cao Tan Phat*, *Huynh Dang Khoa*
Can Tho University of Medicine and Pharmacy

Background: The necessary of early detection, classification of liver lesions and the researching for the application of deep learning models to the field of medical image. **Objectives:** Collect dataset, build, train the deep learning model Faster R-CNN to detect and classify common focal liver lesions on CT images; Test and evaluate the effectiveness of this model according to the criterias of time and accuracy. **Materials and methods:** The abdominal contrast-enhanced CT image dataset with liver lesions including cyst, hemangioma, and hepatocellular carcinoma; Applying Faster R-CNN model to the detection and classification of lesions. **Results:** The dataset had been collected at Can Tho University of Medicine and Pharmacy Hosital included 51 patients who had common focal liver lesions, with 2828 images, 2836 lesion areas that were identified by radiologists, of which 11 patients belonged to the liver cyst group (440 lesion areas), 18 patients belonged to the hemangioma group (648 lesion areas), 21 patients belonged to the hepatocellular carcinoma group (174 lesion areas) and 01 patient had both cyst and hemangioma; Faster R-CNN model resulted in mAP accuracy of 94%, training time of 583 minutes and processing time of 0.13

seconds. **Conclusion:** The collected data set is the foundation for further studies; The Faster R-CNN model had short training time, fast processing time and high accuracy, it was suitable and could be applied to deploy real-life applications.

Keywords: Machine learning, Deep learning, Faster R-CNN, CT Scan, Liver disease.

I. ĐẶT VẤN ĐỀ

Việc phát hiện và phân loại sớm tổn thương gan nhằm đưa ra giải pháp điều trị đúng đắn là vô cùng cần thiết. Để đánh giá chi tiết các tổn thương gan thì kỹ thuật chụp cắt lớp vi tính có tiêm thuốc cản quang là một lựa chọn hoàn hảo. Tuy nhiên, kết quả chẩn đoán phụ thuộc nhiều vào tính tỉ mỉ, nhãn quan và kinh nghiệm của bác sĩ. Xuất phát từ thực tiễn, nghiên cứu ứng dụng mô hình học sâu để phát hiện và phân loại các tổn thương khu trú thường gặp ở gan trên ảnh chụp cắt lớp vi tính cần được triển khai thực hiện.

Các đề tài ứng dụng máy học, đặc biệt là kỹ thuật học sâu để xử lý ảnh chụp cắt lớp vi tính có tiêm thuốc cản quang liên quan đến các tổn thương gan đã được nhiều nhà khoa học tham gia nghiên cứu. Năm 2017, nhóm tác giả Changjian Sun đề xuất mô hình mạng học sâu kết nối đầy đủ đa kênh để vẽ đường biên bao quanh tổn thương [9]. Năm 2018, nhóm tác giả Sang-gil Lee sử dụng mô hình Grouped Single Shot MultiBox Detector để phát hiện các tổn thương gan [5]. Năm 2018, nhóm tác giả Dong Liang đề xuất mô hình Residual Convolutional Neural Networks để trích xuất đặc trưng và sử dụng thuật toán máy học véc tơ hỗ trợ để phân loại các tổn thương [6]. Đầu năm 2021, nhóm tác giả Jiarong Zhou sử dụng ba mô hình Faster R-CNN với mạng rút trích đặc trưng FPN riêng biệt để phân loại các tổn thương gan [10].

Các đề tài trên đều đạt được các kết quả khả quan, tuy nhiên vẫn chưa thể thực hiện được việc phân loại tổn thương hoặc chưa tận dụng được cả bốn thì chụp làm dữ liệu đầu vào. Từ những hạn chế trên, đề tài đề xuất giải pháp xây dựng bốn mô hình học sâu tương ứng cho bốn thì chụp, sử dụng mô hình học sâu hai bước là Faster R-CNN [8] với mạng rút trích đặc trưng ResNet-101 [4], áp dụng vào bộ dữ liệu ảnh cắt lớp vi tính có tiêm thuốc cản quang vùng bụng thu thập tại Bệnh viện Trường Đại học Y Dược Cần Thơ để phát hiện và phân lớp ba loại tổn thương khu trú thường gặp ở gan bao gồm nang (Cyst), u mạch máu (Hemangioma) và ung thư tế bào gan nguyên phát (HCC), với các mục tiêu: (1) Thu thập tập dữ liệu, xây dựng và huấn luyện mô hình học sâu Faster R-CNN để phát hiện và phân loại các tổn thương khu trú thường gặp ở gan trên ảnh chụp cắt lớp vi tính. (2) Kiểm thử và đánh giá hiệu quả của mô hình Faster R-CNN theo các tiêu chí về thời gian và độ chính xác.

II. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Đối tượng nghiên cứu

Đề tài tập trung vào việc phát hiện và phân loại các tổn thương khu trú thường gặp ở gan. Các hình ảnh chụp cắt lớp vi tính có tiêm thuốc cản quang vùng bụng liên quan đến ba loại tổn thương gồm nang gan, u mạch máu và ung thư tế bào gan nguyên phát được thu thập tại Bệnh viện Trường Đại học Y Dược Cần Thơ làm dữ liệu cho quá trình huấn luyện và kiểm thử trên mô hình học sâu là Faster R-CNN với mạng rút trích đặc trưng ResNet-101.

2.2. Phương pháp nghiên cứu

Bài toán có đầu vào là ảnh chụp cắt lớp vi tính có tiêm thuốc cản quang vùng bụng tương ứng với các thì chưa tiêm thuốc cản quang, thì động mạch, thì tĩnh mạch và thì muộn.

Sau quá trình xử lý, kết quả đầu ra là các ảnh tương ứng trên đó đã thực hiện khoanh vùng và đưa ra dự đoán phân lớp thuộc một trong ba loại tổn thương khu trú thường gặp ở gan gồm nang, u mạch máu và ung thư tế bào gan nguyên phát.

Đây là bài toán thuộc về lĩnh vực nhận dạng đối tượng (Object Detection [1]) trên ảnh. Để giải quyết bài toán thì các mô hình học sâu như R-CNN, Fast R-CNN, Faster R-CNN, R-FCN, YOLO, SSD... đều có tính khả thi và có thể mang lại hiệu quả cao. Với đặc thù bài toán chuyên ngành y khoa cần độ chính xác cao, đề tài lựa chọn mô hình học sâu hai bước là Faster R-CNN với mạng trích xuất đặc trưng ResNet-101, áp dụng kỹ thuật làm giàu dữ liệu (Transfer learning) với bộ trọng số đã được huấn luyện trên tập dữ liệu lớn COCO Dataset [7] giúp giải quyết được vấn đề tập dữ liệu nhỏ đồng thời rút ngắn thời gian trong việc huấn luyện các mô hình. Sau đó sử dụng bộ trọng số đã được học tiếp tục huấn luyện và kiểm thử trên tập dữ liệu đã thu thập. Từ đó tiến hành đánh giá, so sánh hiệu quả về độ chính xác và thời gian của mô hình.

Quy trình đề xuất

Quy trình đề xuất giải quyết bài toán gồm ba giai đoạn chính, đó là giai đoạn xây dựng bộ dữ liệu, giai đoạn huấn luyện và cuối cùng là giai đoạn kiểm thử đánh giá. Chi tiết được mô tả như sau:

- Giai đoạn xây dựng bộ dữ liệu:

+ Tập dữ liệu ảnh DICOM chụp cắt lớp vi tính có tiêm thuốc cản quang bụng được thu thập tại Bệnh viện Trường Đại học Y Dược Cần Thơ. Mỗi tập tin gồm bốn tập ảnh tương ứng với thì chưa tiêm thuốc cản quang, thì động mạch, thì tĩnh mạch và thì muộn.

+ Giai đoạn tiền xử lý dữ liệu tiến hành chuyển đổi ảnh DICOM về các ảnh định dạng JPG có kích thước 512x512 nhằm tạo dữ liệu đầu vào phù hợp cho mô hình. Dữ liệu liên quan đến thông tin người bệnh đã được xóa hoàn toàn trong quá trình tiền xử lý để đảm bảo thông tin cá nhân được bảo mật. Công cụ sử dụng để đọc ảnh DICOM và chuyển đổi định dạng ảnh sang JPG là RadiAnt DICOM Viewer.

+ Giai đoạn khoanh vùng, đánh nhãn phân loại tổn thương được hỗ trợ từ chuyên gia là bác sĩ chẩn đoán hình ảnh. Quá trình khoanh vùng và đánh nhãn phân loại trên tập dữ liệu ảnh JPG được thực hiện bằng công cụ LabelImg. Các vùng tổn thương được gán nhãn thành một trong ba lớp là nang (nhãn là NAN), u mạch máu (nhãn HEM), ung thư tế bào gan nguyên phát (nhãn HCC).

+ Tập dữ liệu gồm các hình ảnh thu được dưới định dạng JPG đã qua các bước tiền xử lý và khoanh vùng, đánh nhãn được chia thành tập dữ liệu dùng cho huấn luyện và tập dữ liệu dùng cho kiểm thử. Các tập dữ liệu này được sử dụng làm đầu vào cho các mô hình học sâu.

- Giai đoạn huấn luyện mô hình:

+ Tiến hành xây dựng bốn mô hình tương ứng cho bốn thì chụp. Giai đoạn huấn luyện được thực hiện trên tập dữ liệu huấn luyện. Kết quả thu được là các mô hình đã được huấn luyện. Một số tham số huấn luyện chính bao gồm:

+ Sử dụng phương pháp làm giàu dữ liệu (Transfer learning) kế thừa bộ trọng số đã được huấn luyện trên tập dữ liệu lớn COCO.

+ Số lượng phân lớp (num classes) trong đề tài là ba phân lớp, tương ứng với ba loại tổn thương khu trú thường gặp ở gan gồm nang (ký hiệu nhãn là NAN), u mạch máu (nhãn HEM) và ung thư tế bào gan nguyên phát (nhãn HCC).

+ Số bước học (num steps) là 100.000, con số này được lựa chọn sau quá trình theo

đôi biểu đồ chỉ số Loss giảm dần đến khi không còn giảm được nữa và có dấu hiệu ổn định, khi đó mô hình huấn luyện đã đạt được kết quả tốt.

+ Tham số mAP@IoU được dùng để đánh giá khả năng phân vùng bằng cách đo độ chồng lấp phù hợp giữa vùng dự đoán và vùng nhãn dữ liệu gốc. Nếu độ chồng lấp lớn hơn hoặc bằng 0,5 được xem là dự đoán chính xác và ngược lại nếu nhỏ hơn 0,5 thì dự đoán được xem là không chính xác.

- Giai đoạn kiểm thử:

Sau khi thu được các mô hình đã được huấn luyện ở giai đoạn huấn luyện, giai đoạn kiểm thử được thực hiện cho bốn thì tương ứng với tập dữ liệu kiểm thử. Dữ liệu cũng được chia làm bốn tập tương ứng với từng thì. Kết quả kiểm thử sẽ được ghi nhận với các thông số về độ chính xác và thời gian huấn luyện và xử lý.

Môi trường cài đặt

Tensorflow [3] là một nền tảng nguồn mở được phát triển bởi Google dành cho việc nghiên cứu về máy học và học sâu. Tensorflow có một hệ sinh thái toàn diện, linh hoạt gồm các công cụ, thư viện và tài nguyên cộng đồng cho phép các nhà nghiên cứu thúc đẩy công nghệ tiên tiến trong máy học và các nhà phát triển dễ dàng xây dựng và triển khai các ứng dụng máy học.

Google Colab (Colaboratory) [2] là một dịch vụ đám mây miễn phí, hỗ trợ cung cấp GPU và TPU là các thành phần xử lý đồ họa hỗ trợ cho nhu cầu tính toán lớn. Đây là một công cụ lý tưởng cho nghiên cứu và triển khai ứng dụng máy học, học sâu vì đã được cài đặt sẵn những thư viện rất phổ biến như PyTorch, TensorFlow, Keras và OpenCV.

Các mô hình được xây dựng trên môi trường trực tuyến Google Colab với cấu hình mạnh phù hợp cho nhu cầu tính toán lớn, đi kèm nền tảng TensorFlow, ngôn ngữ Python.

III. KẾT QUẢ NGHIÊN CỨU

3.1. Kết quả thu thập tập dữ liệu, xây dựng và huấn luyện mô hình

Tập dữ liệu thu thập tại Bệnh viện Trường Đại học Y Dược Cần Thơ gồm hình ảnh chụp cắt lớp vi tính vùng bụng có tiêm thuốc cản quang của 51 người bệnh với chẩn đoán có tổn thương khu trú thường gặp ở gan. Số lượng người bệnh và các loại tổn thương được mô tả ở bảng bên dưới.

Bảng 1. Số lượng người bệnh tương ứng với các loại tổn thương

	Nang	U mạch máu	HCC
Số lượng người bệnh	51		
	11	18	21
	01		

Nhận xét: Số lượng người bệnh có tổn thương HCC chiếm số lượng cao nhất là 21 và có 01 người bệnh vừa có tổn thương nang gan vừa có u mạch máu.

Quá trình tiền xử lý dữ liệu chuyển đổi định dạng từ DICOM sang JPG được thực hiện bằng công cụ Radiant Viewer. Sau đó quá trình khoanh vùng đánh nhãn được thực hiện bằng công cụ LabelImg với sự hỗ trợ của chuyên gia là bác sĩ chẩn đoán hình ảnh. Kết quả số lượng vùng tổn thương tương ứng với từng phân loại và tập dữ liệu huấn luyện, kiểm thử được mô tả như bảng bên dưới.

Bảng 2. Số lượng vùng tổn thương

Số lượng vùng tổn thương	Nang	U mạch máu	HCC	Tổng cộng
Tập huấn luyện (37 người bệnh)	316	504	1188	2008
Tập kiểm thử (14 người bệnh)	124	144	560	828
Tổng cộng	440	648	1748	2836

Nhận xét: Tổng số lượng vùng tổn thương là 2836 vùng. Trong đó, 2008 vùng tổn thương nằm ở tập dữ liệu huấn luyện và 828 vùng còn lại dùng cho tập kiểm thử. Số lượng vùng tổn thương nang gan là thấp nhất với 440 vùng và cao nhất là HCC với 1748 vùng.

Tương ứng bốn thì chụp, bốn mô hình Faster R-CNN với mạng rút trích đặc trưng ResNet-101 đã được xây dựng trên môi trường Google Colab, hệ điều hành Ubuntu với cấu hình với bộ vi xử lý, RAM, bộ xử lý đồ họa GPU mạnh hỗ trợ rất nhiều cho nhu cầu tính toán lớn với bài toán xử lý ảnh đầu vào kích thước 512x512, đi kèm nền tảng TensorFlow GPU 1.15 và ngôn ngữ Python 3.7. Tập dữ liệu huấn luyện và kiểm thử được sử dụng làm đầu vào tương ứng cho các mô hình đã xây dựng. Chi tiết thông số cấu hình hệ thống được mô tả như bảng bên dưới.

Bảng 3. Thông số cấu hình và môi trường cài đặt.

Môi trường	Google Colab
CPU	Intel® Xeon® 2.00 GHz
RAM	32 GB
GPU	Nvidia Tesla P100 12GB RAM
Hệ điều hành	Ubuntu
Framework	Tensorflow GPU 1.15
Ngôn ngữ lập trình	Python 3.7

Nhận xét: Môi trường thực nghiệm trực tuyến với cấu hình mạnh hỗ trợ hiệu quả cho quá trình huấn luyện và kiểm thử với dữ liệu đầu vào ảnh kích thước 512x512.

3.2. Kết quả kiểm thử và đánh giá hiệu quả của mô hình

Quá trình huấn luyện với 100.000 bước học và kiểm thử mô hình Faster R-CNN trên tập dữ liệu đã thu thập, đề tài thu được một số kết quả thực nghiệm về độ đo giá trị mất mát (chỉ số Loss), thời gian huấn luyện, thời gian xử lý, kết quả phân loại tổn thương với độ đo chính xác được sử dụng là AP và mAP.

Bảng 4. Tổng hợp kết quả thực nghiệm trên mô hình Faster R-CNN

Chỉ số Loss trung bình	Thời gian trung bình		AP trung bình			mAP trung bình
	Huấn luyện	Xử lý	Nang gan	U mạch máu	HCC	
0,024122	583 phút	0,13 giây	85,1%	98,3%	98,5%	94%

Nhận xét: Chỉ số Loss của mô hình Faster R-CNN với mạng rút trích đặc trưng ResNet-101 là 0,024122, chỉ số Loss thấp cho thấy mô hình rất phù hợp và có khả năng rút trích đặc trưng hiệu quả, từ đó cho độ chính xác mAP trung bình cao là 94%. Thời gian huấn luyện trung bình ngắn với 583 phút, tốc độ xử lý nhận dạng phân loại tổn thương của mô hình nhanh với 0,13 giây.

IV. BÀN LUẬN

Do số lượng tổn thương gan ngày càng tăng và việc chẩn đoán các tổn thương này bằng cách sử dụng kỹ thuật chụp cắt lớp vi tính có tiêm thuốc cản quang vùng bụng phụ thuộc nhiều vào tính tỉ mỉ và kinh nghiệm của bác sĩ, việc ứng dụng kỹ thuật học sâu vào

lĩnh vực xử lý ảnh y khoa hỗ trợ chẩn đoán ngày càng trở nên quan trọng nhằm tăng cường sự hiệu quả và giảm thiểu thời gian công sức. Trong nghiên cứu này, chúng tôi đã xây dựng bốn mô hình học sâu tương ứng cho cả bốn thì chụp cắt lớp vi tính có tiêm thuốc cản quang đa pha, sử dụng mô hình học sâu hai bước là Faster R-CNN với mạng rút trích đặc trưng ResNet-101, áp dụng vào bộ dữ liệu ảnh cắt lớp vi tính có tiêm thuốc cản quang vùng bụng thu thập tại Bệnh viện Trường Đại học Y Dược Cần Thơ để phát hiện và phân lớp ba loại tổn thương khu trú thường gặp ở gan bao gồm nang (Cyst), u mạch máu (Hemangioma) và ung thư tế bào gan nguyên phát (HCC) và đạt được kết quả tốt.

Ưu điểm của các hệ thống học sâu là chúng có thể liên tục nhận ra, trích xuất và tìm hiểu các đặc điểm phân cấp khác nhau. Điều này hỗ trợ các bác sĩ trong việc phát hiện và phân loại các tổn thương, đặc biệt là những tổn thương với kích thước nhỏ dễ bị bỏ sót. Một số nghiên cứu trước đây đã chỉ ra rằng các mô hình mạng học sâu có thể được sử dụng để phát hiện và khoanh vùng tổn thương gan khu trú trong hình ảnh chụp cắt lớp vi tính, điển hình với nghiên cứu của tác giả Changjian Sun cùng các cộng sự với việc sử dụng mô hình mạng học sâu kết nối đầy đủ đa kênh [9], hay nghiên cứu của nhóm tác giả Sang-gil Lee với mô hình Grouped Single Shot MultiBox Detector [5]. Tuy nhiên, các đề tài này chỉ dừng lại ở việc phát hiện mà chưa thể thực hiện nhiệm vụ phân loại tổn thương.

Việc áp dụng các mô hình máy học, học sâu vào xử lý ảnh y khoa thường có dữ liệu đầu vào là một hình ảnh duy nhất, quá trình xử lý và nhận dạng sẽ cho ra kết quả dự đoán trên ảnh này. Do đó, các quá trình xử lý máy học, học sâu thường được thực hiện trên ảnh X quang, siêu âm, ảnh chụp các tổn thương da...Tiên thêm một bước, các đề tài của nhóm tác giả Dong Liang với mô hình Residual Convolutional Neural Networks và thuật toán máy học véc tơ hỗ trợ [6], hay đề tài của nhóm tác giả Jiarong Zhou cùng cộng sự với ba mô hình Faster R-CNN cùng mạng rút trích đặc trưng FPN riêng biệt [10], đã thực hiện thành công việc phân loại các tổn thương khu trú thường gặp ở gan trên ảnh cắt lớp vi tính đa pha có tiêm thuốc cản quang với đầu vào là các ảnh ở các thì chụp tương ứng với độ chính xác phát hiện và phân loại tương đối cao. Tuy nhiên, các đề tài vẫn còn hạn chế khi chỉ sử dụng ba trên bốn thì chụp cắt lớp vi tính có tiêm thuốc cản quang. Khắc phục điểm yếu này, đề tài đề xuất giải pháp xây dựng bốn mô hình học sâu tương ứng để sử dụng đầy đủ cả bốn thì chụp của quá trình chụp cắt lớp vi tính có tiêm thuốc cản quang làm dữ liệu đầu vào là một điểm rất mới so với các nghiên cứu trước đây. Bốn mô hình tiến hành quá trình học các thuộc tính của tổn thương trong giai đoạn huấn luyện bằng cách trích xuất các đặc trưng tại mỗi thì chụp thông qua bốn ảnh đầu vào. Điều này góp phần mang đến hiệu quả cao hơn so với các nghiên cứu trước đây. Với kết quả kiểm thử thực nghiệm, có thể thấy mô hình Faster R-CNN có thời gian huấn luyện ngắn (583 phút) và tốc độ xử lý rất nhanh (0,13 giây). Bên cạnh đó, mô hình có khả năng rút trích đặc trưng rất hiệu quả khi chỉ số Loss đạt giá trị thấp, từ đó độ chính xác thu được cao (94%), điều này đến từ việc mạng rút trích đặc trưng Resnet-101 hoạt động rất hiệu quả với loại dữ liệu ảnh chụp cắt lớp vi tính có tiêm thuốc cản quang đa pha.

Với kết quả thu được về mặt thời gian và độ chính xác, mô hình Faster R-CNN cho thấy mình rất phù hợp với bài toán xử lý ảnh y khoa, đặc biệt là bài toán phát hiện và phân loại các tổn thương khu trú thường gặp ở gan trên ảnh chụp cắt lớp vi tính. Tuy nhiên, vẫn còn nhiều mô hình học sâu khác bao gồm các mô hình thuộc nhóm học một bước (SSD, YOLO...) với thời gian xử lý nhanh hơn nhưng độ chính xác thấp hơn các mô hình học hai bước như R-FCN... các mô hình này đã và đang chứng minh rất tốt tính hiệu quả trong lĩnh vực xử lý ảnh y khoa, do đó cần mở rộng nghiên cứu đánh giá trên các mô hình học sâu

khác, lựa chọn phương pháp tối ưu với sự cân bằng giữa hai yếu tố là thời gian và độ chính xác nhằm tìm ra mô hình phù hợp cho giai đoạn xây dựng ứng dụng thực tế có hiệu quả cao.

Đặc tính hình ảnh chụp cắt lớp vi tính vùng bụng có tiêm thuốc cản quang bao gồm bốn thì chụp với số lượng ảnh nhiều và kích thước lớn là một thách thức. Quá trình thu thập và xây dựng tập dữ liệu tiêu tốn nhiều thời gian, công sức và có sự đóng góp rất lớn từ bác sĩ chẩn đoán hình ảnh. Tập dữ liệu gồm 51 người bệnh, trong đó có 11 người bệnh thuộc nhóm nang gan (440 vùng), 18 người bệnh thuộc nhóm u mạch máu (648 vùng), 21 người bệnh thuộc nhóm ung thư tế bào gan nguyên phát (1748 vùng) và 01 người bệnh có cả nang và u mạch máu. Tại thời điểm nghiên cứu, chưa thể tìm được bất kỳ bộ dữ liệu chụp cắt lớp vi tính có tiêm thuốc cản quang vùng bụng có đánh nhãn tổn thương gan nào được chia sẻ công cộng để tiến hành các nghiên cứu ứng dụng các mô hình học sâu vào xử lý và so sánh đối chiếu các kết quả trên cùng tập dữ liệu. Tập dữ liệu đã thu thập được là nền tảng cho phép mở rộng hướng phát triển các nghiên cứu tiếp theo. Tuy nhiên, cần liên tục bổ sung tăng cường cho bộ dữ liệu, không chỉ tăng cường số lượng hình ảnh mà còn cung cấp nhiều phân loại tổn thương gan hơn kể cả tổn thương khu trú và lan tỏa, ngoài ra còn phát triển bổ sung tổn thương ở các cơ quan khác.

Ngoài ra có thể phát triển bài toán này trên môi trường tiếp cận dữ liệu lớn như Apache Spark khi bộ dữ liệu ngày càng được tăng cường và bổ sung. Khi đó các giải pháp tính toán thông thường sẽ không còn đạt hiệu quả cao với bộ dữ liệu cực lớn.

V. KẾT LUẬN

Đề tài đã hoàn thiện bộ cơ sở dữ liệu ảnh chụp cắt lớp vi tính có tiêm thuốc cản quang vùng bụng thu thập tại Bệnh viện Trường Đại học Y Dược Cần Thơ gồm 51 trường hợp có một trong các chẩn đoán tổn thương khu trú thường gặp ở gan bao gồm nang, u mạch máu và ung thư tế bào gan nguyên phát, với 2828 ảnh cùng 2836 vùng tổn thương được xác định bởi chuyên gia là bác sĩ chuyên ngành chẩn đoán hình ảnh. Áp dụng mô hình học sâu Faster R-CNN với mạng rút trích đặc trưng ResNet-101 để phát hiện và phân loại tổn thương khu trú thường gặp ở gan với kết quả độ chính xác mAP là 94%, thời gian huấn luyện 583 phút và tốc độ xử lý 0,13 giây. Mô hình Faster R-CNN cho thấy sự phù hợp và có thể được áp dụng để triển khai các ứng dụng thực tế trong tương lai.

TÀI LIỆU THAM KHẢO

1. Amit, Y., P. Felzenszwalb, and R. Girshick, *Object detection*. Computer Vision: A Reference Guide, 2020: pp. 1-9.
2. Bisong, E., *Google colab*, in *Building Machine Learning and Deep Learning Models on Google Cloud Platform*. 2019, Springer. pp. 59-64.
3. Dillon, J.V., et al., *Tensorflow distributions*. arXiv preprint arXiv:1711.10604, 2017.
4. He, K., et al. *Deep residual learning for image recognition*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
5. Lee, S.-g., et al. *Liver lesion detection from weakly-labeled multi-phase CT volumes with a grouped single shot multibox detector*. in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2018. Springer.
6. Liang, D., et al. *Residual convolutional neural networks with global and local pathways for classification of focal liver lesions*. in *Pacific Rim International Conference on Artificial Intelligence*. 2018. Springer.
7. Lin, T.-Y., et al. *Microsoft coco: Common objects in context*. in *European conference on*

- computer vision*. 2014. Springer.
8. Ren, S., *et al.*, *Faster r-cnn: Towards real-time object detection with region proposal networks*. arXiv preprint arXiv:1506.01497, 2015.
 9. Sun, C., *et al.*, *Automatic segmentation of liver tumors from multiphase contrast-enhanced CT images based on FCNs*. *Artificial intelligence in medicine*, 2017. 83: pp. 58-66.
 10. Zhou, J., *et al.*, *Automatic Detection and Classification of Focal Liver Lesions Based on Deep Convolutional Neural Networks: A Preliminary Study*. *Frontiers in oncology*, 2021. 10: pp. 3261.
- (Ngày nhận bài: 13/10/2022 – Ngày duyệt đăng: 18/02/2023)
-