

PHÁT HIỆN SỎI THẬN SỬ DỤNG VISION TRANSFORMER

DETECTION OF KIDNEY STONES USING VISION TRANSFORMER

HÀ THỊ ĐOAN TRANG

Trường Đại học Sư phạm Kỹ thuật Vĩnh Long,

Tác giả liên hệ: tranghtd@vlute.edu.vn

Nhận bài(Received): 17/8/2023; Phản biện (Reviewed):22/8/2023; Chấp nhận (Accepted):27/9/2023

TÓM TẮT

Sỏi thận là nguyên nhân chính gây bệnh tật và có tỷ lệ mắc bệnh ngày càng gia tăng trên thế giới; sỏi thận dễ muộn sẽ gây ra biến chứng nhiễm khuẩn tiết niệu, giãn đài bể thận và suy thận không hồi phục; gây ra thiệt hại về kinh tế hoặc thậm chí là tính mạng con người nếu không được điều trị kịp thời. Do đó, việc phát hiện và điều trị sớm bệnh sỏi thận là điều vô cùng cần thiết. Mặt khác, hiện nay, Vision Transformer (ViT) nổi lên như một giải pháp thay thế cạnh tranh so với các mạng thần kinh tích chập trong lĩnh vực thị giác máy tính, và được sử dụng rộng rãi trong các tác vụ nhận dạng hình ảnh khác nhau. Trong bài báo này, chúng tôi đề xuất phương pháp phát hiện sỏi thận sử dụng Vision Transformer. Kết quả nghiên cứu cho thấy, phương pháp đề xuất đạt độ chính xác lên đến 93% trên tập dữ liệu hình ảnh CT trong quá trình huấn luyện

Từ khóa: Vision Transformer, Sỏi thận, Deep learning, Coronal CT

ABSTRACT

Kidney stones are the primary cause of illness, and the incidence of this condition is increasing worldwide. Delayed kidney stone diagnosis can lead to complications such as urinary tract infections, kidney obstruction, and irreversible kidney damage, potentially resulting in significant economic and even life-threatening consequences if not treated promptly. Therefore, early detection and treatment of kidney stones are essential. Furthermore, Vision Transformer (ViT) has emerged as a competitive alternative to traditional convolutional neural networks in the field of computer vision and is widely used in various image recognition tasks. In this paper, we propose a method for kidney stone detection using Vision Transformer. Research results demonstrate that the proposed method achieves an accuracy rate of up to 93% on CT image datasets during the training process.

Keywords: Vision Transformer, Kidney Stone, Deep learning, Coronal CT

1. MỞ ĐẦU

như một hệ thống lọc của cơ thể con người

1.1. Giới thiệu bài toán

[2]. Sỏi thận là tình trạng các chất cặn bã

Thận là cơ quan phức tạp đóng vai trò

không mong muốn lắng đọng trong thận,

cản trở hoạt động bình thường của hệ tiết niệu và trong một số trường hợp, làm tắc nghẽn dòng nước tiểu, gây đau đón khó chịu. Khoảng 12% dân số trên toàn thế giới mắc bệnh sỏi thận và tỷ lệ mắc bệnh đã tăng lên trong vài thập kỷ qua [6 – 7]. Hằng năm, có khoảng 1,3 triệu lượt khám và cấp cứu liên quan đến sỏi thận và gây ra thiệt hại kinh tế lên đến 5 tỷ USD [8 – 9]. Dựa trên dữ liệu thống kê, mỗi năm tỷ lệ sỏi thận tăng 16% ở cả trẻ em và người lớn. [10]. Việc xây dựng một hệ thống hỗ trợ việc chẩn đoán sỏi thận nhằm tạo nhiều thuận lợi hơn cho các bác sĩ trong quá trình phân tích và chẩn đoán thông qua hình ảnh Coronal CT để có giải pháp điều trị kịp thời vô cùng cần thiết.

1.2. Những nghiên cứu liên quan

Fitri và cộng sự. [3] đã đề xuất một mô hình dựa trên CNN để phân loại sỏi tiết niệu từ chụp cắt lớp vi tính (micro-CT). Mô hình CNN được đào tạo đạt được độ chính xác thử nghiệm là 0,9959 với sai số phân loại là 1,2%. Alper Caglayan et al. [4] đề xuất mô hình phát hiện sỏi thận sử dụng mô hình học sâu ResNet sử dụng tập dữ liệu ảnh chụp CT của 455 bệnh nhân (trong đó 405 người được chẩn đoán mắc bệnh sỏi thận và 50 người thì không) với độ chính xác đạt được trên ảnh chụp coronal CT là 97% cho giai đoạn huấn luyện. Irudayaraj et al [5] đã đề xuất bốn phương pháp học sâu cụ thể đã được sử dụng để phát hiện xem có sỏi thận hay không bao gồm VGG16, ResNet50V2, MobileNetV2, InceptionNetV3. Kết quả, mạng thần kinh InceptionNet tạo ra kết quả phân loại tốt nhất về độ chính xác là 0,862.

1.3. Đặc điểm của sỏi thận

Sỏi thận (còn gọi là sỏi tiết niệu) là

những chất lắng đọng cứng làm từ khoáng chất và muối hình thành bên trong thận [1]. Sỏi thận thường không gây ra triệu chứng cho đến khi nó di chuyển trong thận hoặc đi vào một trong các niệu quản. Niệu quản là ống nối thận và bàng quang [1]. Đáng ngạc nhiên là một số bệnh nhân không có triệu chứng (không có triệu chứng), trong khi những người khác có thể bị đau dữ dội (đau sườn, được gọi là cơn đau quặn thận), hoặc tiểu ra máu (tiểu ra máu). Nếu không được điều trị, nó có thể dẫn đến nhiễm trùng và gây sốt. Nó cũng có thể dẫn đến suy giảm chức năng thận. Khoảng 1 trong 17 người trưởng thành sẽ có người bị mắc bệnh sỏi thận [11].

1.4. Mạng Vision Transformer

Vision Transformer [3] là một mạng nơ-ron được sử dụng để xử lý và phân loại hình ảnh dựa trên cơ chế tự chú ý; cho phép xử lý hình ảnh bằng cách chia ảnh thành các patch nhỏ. Điều này giúp mô hình nhận biết các đặc trưng không gian trong ảnh và mối quan hệ giữa các patch; quá trình xử lý thông tin được thực hiện dựa trên các phép chuyển đổi tuyến tính và cơ chế tự chú ý mà không yêu cầu các phép tích chập như các mạng CNN, các mô hình ViT được đánh giá là vượt trội hơn so với CNN gần 4 lần về hiệu quả tính toán và độ chính xác. Ngoài ra, Vision Transformer có thể mở rộng để xử lý ảnh có kích thước lớn hơn bằng cách tăng số lượng patch [4].

1.5. Các độ đo đánh giá mô hình

Accuracy [16] được đánh giá dựa trên các giá trị như true positives (TP), false positives (FP), true negatives (TN) và false negatives (FN) được mô tả theo phương trình (1). Hàm categorical cross-entropy loss $L(c,p)$ [17] với c là nhãn dữ liệu, p là

giá trị thực tế và p là giá trị dự đoán. Hàm $L(c,p)$ được mô tả theo phương trình (2).

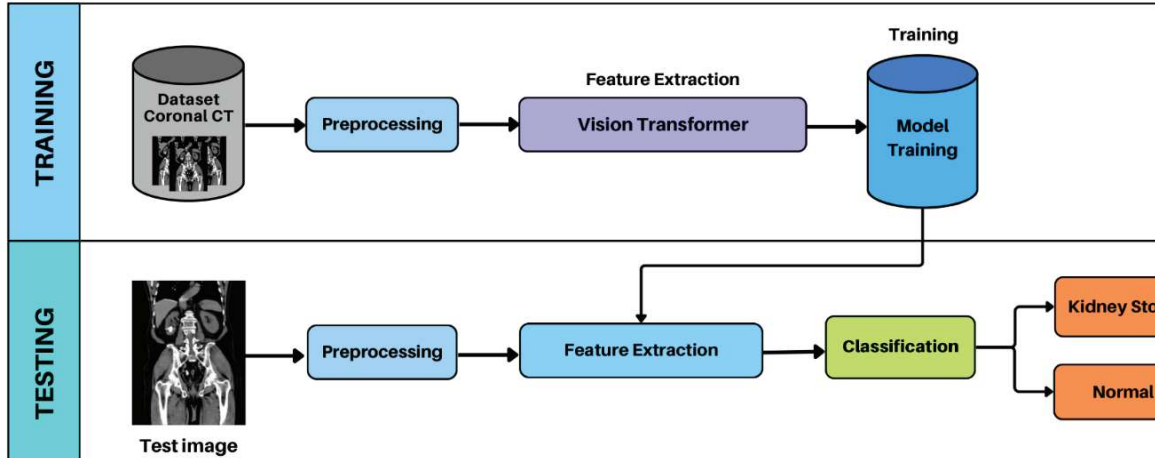
$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\text{Loss} = \sum_{i=1}^K y_i \log(\hat{y}_i) \quad (2)$$

2. KẾT QUẢ NGHIÊN CỨU

2.1. Phương pháp đề xuất

Trong nghiên cứu này, chúng tôi sử dụng phương pháp học chuyển giao dựa trên mạng Vision Transformer để phát hiện sỏi thận. Mô hình đề xuất gồm 2 giai đoạn là huấn luyện và kiểm thử. Chi tiết của các giai đoạn được thể hiện ở Hình 1:



Hình 1. Mô hình đề xuất phát hiện sỏi thận

2.1.1. Đào tạo mô hình

a) Tiền xử lý và chuẩn bị dữ liệu:

Chúng tôi thực hiện chuẩn hoá tất cả hình ảnh thành kích thước 224×224 pixel, đây là kích thước phù hợp để tạo bản vá từ hình ảnh đầu vào. Sau đó, chúng tôi tiến hành tăng cường dữ liệu bằng cách thay đổi góc quay hình ảnh (xoay ảnh) và lật ảnh.

b) Rút trích đặc trưng và huấn luyện:

Hình ảnh đầu vào sau khi được tiền xử lý sẽ được cắt thành các patch đưa qua tầng liner projection với kích thước cho mỗi patch là 16×16 do tính ổn định của nó đối với sự suy giảm hiệu suất và tính phức tạp tính toán [4], mỗi patch sẽ được gán vị trí và nhãn cụ thể đảm bảo ảnh sau khi rút trích các đặc trưng sẽ không bị xáo trộn. Sau đó đưa qua

transform encoder để tiến hành rút trích đặc trưng mã hóa, phân loại và khoanh vùng lành tính và ác tính. Kết quả của quá trình này sẽ là ảnh được phân loại thuộc vào nhãn nào.

2.1.2 Kiểm thử

Ở giai đoạn kiểm thử, hình ảnh đầu vào sẽ được chuẩn hoá thành kích thước 224×224 , sau đó tiến hành đưa qua mô hình đã được huấn luyện để tiến hành dự đoán. Kết quả kiểm thử sẽ đưa ra một trong ba kết luận là Có sỏi (Stone) hoặc Không sỏi (Normal)

2.2 Kết quả thực nghiệm

2.2.1. Tham số huấn luyện

Trong nghiên cứu này, chúng tôi thực hiện huấn luyện với các tham số trong **Bảng 1**

Bảng 1. Các tham số huấn luyện

STT	Kiến trúc mạng	Mạng rút trích đặc trưng và phân loại	Số bước huấn luyện	Tốc độ học	Kích thước ảnh
1	Vision Transformer	Default	500	0.001	224 224

2.2.2. Môi trường cài đặt và tập dữ liệu thực nghiệm

a. *Môi trường cài đặt:* Phương pháp đề xuất được cài đặt trên nền tảng Kaggle với thư viện hỗ trợ đào tạo mô hình mạng sử dụng là Tensorflow và Keras.

b. Tập dữ liệu

Tập dữ liệu sử dụng được lấy từ Một nghiên cứu tại trường Đại học Firat, Thổ Nhĩ Kỳ đã phối hợp với y, bác sĩ bệnh viện Thành phố Elazığ Fethi Sekin đề xuất một tập dữ liệu hình ảnh CT cho bệnh nhân nhập viện vì bệnh sỏi thận [5]. Cụ thể:

- 790 hình ảnh cho bệnh nhân được

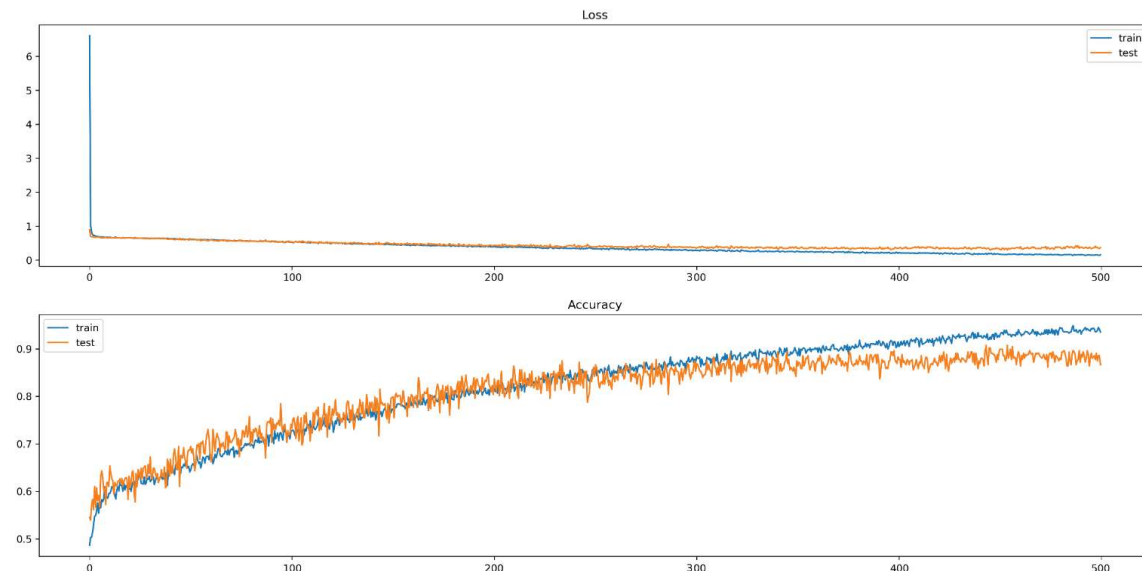
chẩn đoán sỏi thận (Kidney Stone)

- 1009 hình ảnh của các đối tượng bình thường (Normal) đã thu được

Trong nghiên cứu này, chúng tôi sử dụng 80% cho quá trình huấn luyện và 20% cho kiểm thử.

2.2.3. Các kết quả thu được

Hình 2 biểu diễn giá trị mất mát (loss) và độ chính xác (accuracy) của mô hình đề xuất trong giai đoạn huấn luyện. Giá trị train loss và val loss lần lượt là 0.2358 và 0.3615. Giá trị train accuracy và val accuracy lần lượt là 93% trên tập huấn luyện và 88% trên tập kiểm thử với thời gian huấn luyện là 936 phút.

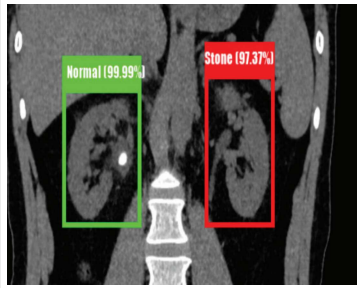
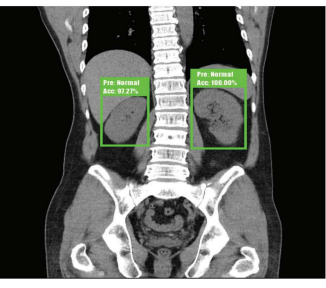
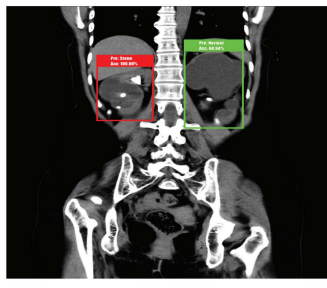


Hình 2. Độ đo Loss và Accuracy của mô hình đề xuất trong giai đoạn huấn luyện

c) Một số hình ảnh kết quả phân loại

Bảng 2 trình bày một số kết quả thực nghiệm của mô hình đề xuất trong 3 trường hợp lần lượt là Stone – Normal, Normal và Stone. Trong các trường hợp, mô hình đề

xuất đều cho kết quả dự đoán chính xác. Qua một số kết quả thực nghiệm, ta có thể thấy rằng, mô hình đề xuất mang lại kết quả chính xác tốt hơn so với các kịch bản còn lại.

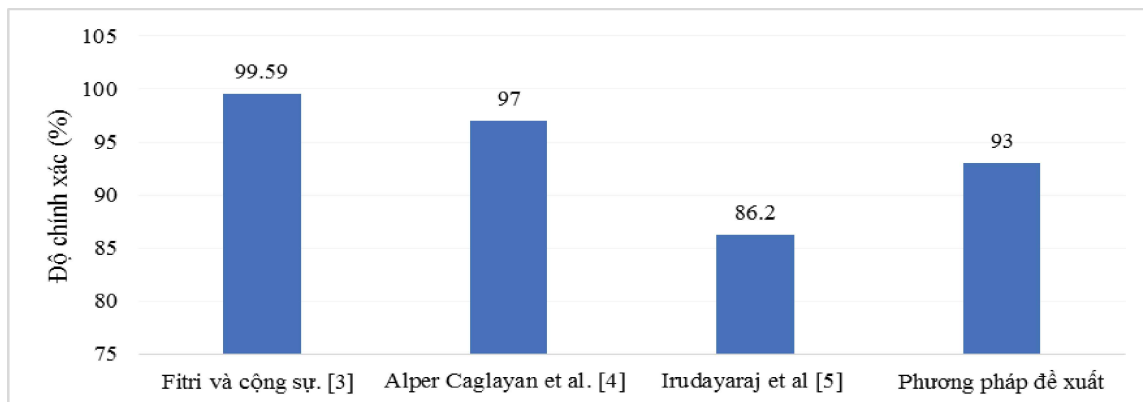
	Stone – Normal	Normal	Stone
Predicted			

Bảng 2. Một số kết quả trên tập dữ liệu kiểm thử

d) Một số kết quả so sánh với các công trình nghiên cứu liên quan

Hình 4 trình bày kết quả so sánh của phương pháp đề xuất với các công trình nghiên cứu có liên quan. Nhìn chung,

mô hình đề xuất đạt độ chính xác còn hạn chế so với một số công trình nghiên cứu liên quan, trong tương lai sẽ thực hiện cải thiện độ chính xác cho phương pháp đề xuất.



Hình 4. Một số kết quả so sánh với các công trình nghiên cứu liên quan

3. KẾT LUẬN

Trên toàn cầu, tỷ lệ mắc sỏi thận đã tăng lên theo thời gian. Nếu không được điều trị tốt hơn, sỏi trong thận có thể dẫn đến tắc nghẽn niệu quản, nhiễm trùng lặp đi lặp lại ở đường tiết niệu, tiểu buốt và suy thận vĩnh viễn nếu không được điều trị kịp thời. Trong nghiên cứu này, tôi cung cấp một kết quả đánh giá mô hình đề xuất để góp phần trong

việc tìm kiếm và lựa chọn được phương pháp phù hợp để có thể xây dựng hệ thống phát hiện và khoanh vùng một cách chính xác nhằm hỗ trợ các bác sĩ có thể đưa ra chuẩn đoán một cách nhanh chóng và sớm điều trị kịp thời cho bệnh nhân. Kết quả nghiên cứu cho thấy, phương pháp đề xuất đạt độ chính xác lên đến 93% trên tập dữ liệu Coronal CT trong quá trình huấn luyện.

TÀI LIỆU THAM KHẢO

- [1] Kai Han, Yunhe Wang, Hanting Chen, Xinghao Chen, Jianyuan Guo, Zhenhua Liu, Yehui Tang, An Xiao, Chunjing Xu, Yixing Xu, Zhaohui Yang, Yiman Zhang, and Dacheng Tao Fellow, “A Survey on Vision Transformer”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Jan. 2023, pp. 87-110, vol. 45

- [2] Thurman JM, “Complement and the Kidney: An Overview. *Adv Chronic Kidney Dis*”, 2020 Mar;27(2):86-94. doi: 10.1053/j.ackd.2019.10.003. PMID: 32553250; PMCID: PMC7310605.
- [3] Fitri LA, Haryanto F, Arimura H, YunHao C, Ninomiya K, Nakano R, Haekal M, Warty Y, Fauzi U, “Automated classification of urinary stones based on microcomputed tomography images using convolutional neural network”, *Phys Med*. 2020 Oct;78:201-208. doi: 10.1016/j.ejmp.2020.09.007. Epub 2020 Oct 8. PMID: 33039971.
- [4] Caglayan A, Horsanali MO, Kocadurdu K, Ismailoglu E, Guneyli S, “Deep learning model-assisted detection of kidney stones on computed tomography”, *Int Braz J Urol*. 2022 Sep-Oct;48(5):830-839. doi: 10.1590/S1677-5538.IBJU.2022.0132. PMID: 35838509; PMCID: PMC9388181.
- [5] Irudayaraj, Alexander Albert, “Kidney Stone Detection using Deep Learning Methodologies”, Masters thesis, Dublin, National College of Ireland, 2022
- [6] Alelign T, Petros B. Kidney Stone Disease: An Update on Current Concepts. *Adv Urol*. 2018 Feb 4;2018:3068365. doi: 10.1155/2018/3068365. PMID: 29515627; PMCID: PMC5817324.
- [7] Romero V, Akpınar H, Assimos DG. Kidney stones: a global picture of prevalence, incidence, and associated risk factors. *Rev Urol*. 2010 Spring;12(2-3):e86-96. PMID: 20811557; PMCID: PMC2931286.
- [8] Hyams ES, Matlaga BR. Economic impact of urinary stones. *Transl Androl Urol*. 2014 Sep;3(3):278-83. doi: 10.3978/j.issn.2223-4683.2014.07.02. PMID: 26816777; PMCID: PMC4708578.
- [9] Strohmaier WL. Economics of stone disease/treatment. *Arab J Urol*. 2012 Sep;10(3):273-8. doi: 10.1016/j.aju.2012.02.002. Epub 2012 Mar 22. PMID: 26558036; PMCID: PMC4442936.
- [10] WebMD Editorial Contributors, “Kidney Stones”, December 16, 2022
- [11] Zeng, G, Mai, Z, Xia, S, Wang, Z, Zhang, K, Wang, L, “Prevalence of kidney stones in China: an ultrasonography based cross-sectional study”, *BJU Int*. (2017) 120:109–16. doi: 10.1111/bju.13828