

ỦY BAN NHÂN DÂN
THÀNH PHỐ HỒ CHÍ MINH
SỞ KHOA HỌC VÀ CÔNG NGHỆ

CƠ QUAN CHỦ QUẢN
CƠ QUAN CHỦ TRÌ NHIỆM VỤ

CHƯƠNG TRÌNH KHOA HỌC VÀ CÔNG NGHỆ CẤP THÀNH PHỐ

BÁO CÁO TÓM TẮT
KẾT QUẢ NHIỆM VỤ NGHIÊN CỨU KHOA HỌC CÔNG NGHỆ
(TÊN NHIỆM VỤ)

Cơ quan chủ trì nhiệm vụ: THPT chuyên Lê Hồng Phong
Chủ nhiệm nhiệm vụ: Đỗ Quốc Anh Triết

Thành phố Hồ Chí Minh - 2023

Hình 1. Báo cáo tóm tắt kết quả nhiệm vụ

Phần mềm hệ thống tích hợp học sâu để phân đoạn và tái tạo cấu trúc tim 3 chiều cho ứng dụng y học

Tác giả 1: Nguyễn Lê Quốc Bảo

Tác giả 2: Lê Tuấn Hy

GVHD: Đỗ Quốc Anh Triết

29/09/2023

Mục lục

A. Mở đầu.....	3
B. Điều tra khảo sát	4
1. Khó và chuyên sâu.	4
2. Dị tật / dị dạng.....	4
3. Phần mềm nước ngoài.....	4
C. Thực hiện nghiên cứu	4
1. Nhiệm vụ nghiên cứu.....	4
2. Giả thuyết khoa học	5
3. Quy trình nghiên cứu	5
4. Cơ sở lý thuyết	6
4.1. Chuẩn bị và tìm hiểu dữ liệu.....	6
4.2 Tiền xử lý dữ liệu.....	7
4.3 Xây dựng kiến trúc và huấn luyện mô hình.....	8
4.4 Đánh giá mô hình (Metrics).....	10
D. Kết quả nghiên cứu (sản phẩm).....	12
E. Mục tiêu tương lai.....	13
F. Trích Dẫn	13

A. Mở đầu

1. Tóm tắt dự án

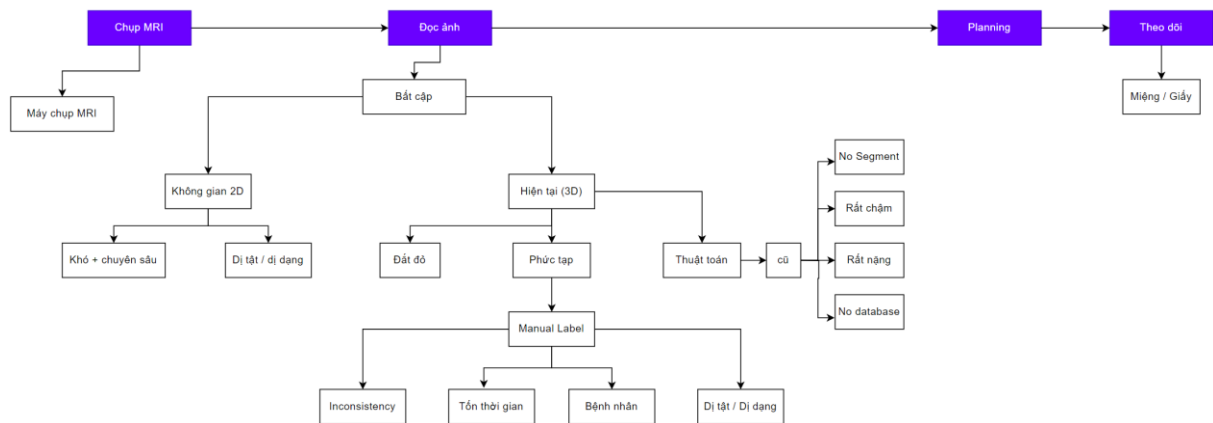
Trong ngành y khoa tim mạch, ảnh chụp cắt lớp tim (MRI hoặc CT scan) gây nhiều khó khăn trong việc đọc hiểu và phân tích chính xác trong giai đoạn tiền phẫu thuật[1]. Để giải quyết vấn đề này, chúng tôi phát triển một phần mềm hệ thống trên nền tảng trực tuyến (website) tích hợp các công cụ phân tích nhiều loại ảnh cắt lớp, và dựng mô hình tim 3D dựa trên khối ảnh cắt lớp 2D. Chúng tôi nghiên cứu về các mô hình học sâu Unet, VGG, Resnet và các hàm mất mát IoU, Dice Coefficient, Jaccard cho công việc phân khu (segmentation) hình ảnh 2D tim thành 7 phần khác nhau. Chúng tôi áp dụng thuật toán Marching cubes để trích xuất đặc trưng kết quả phân khu, tính toán khoảng cách tọa độ lưới và tái tạo mô hình tim 3D trong không gian 3 chiều, kết nối kính VR tăng tính tương tác. Để tăng cường sự quản lý hồ sơ bệnh án và theo dõi thay đổi trong cấu trúc tim, chúng tôi tích hợp database SQLAlchemy với Python Backend cho nền tảng ứng dụng. Trong quá trình thử nghiệm, mô hình phân khu đa lớp (multiclass model) sử dụng kiến trúc Unet kết hợp attention và contour shape-aware vượt xa kiến trúc truyền thống khác, đạt độ chính xác 0.9455 và độ mất mát 0.1414 trên bộ dữ liệu nghiên cứu. Mô hình được biên dịch bằng Tensorflow.js giảm độ nặng, tiết kiệm tài nguyên máy tính trong một lần chạy, tăng tốc độ cho ra kết quả nhanh hơn 20% trên nền tảng website. Nhờ sự gọn nhẹ này, ứng dụng của chúng tôi có thể hoạt động trong môi trường trình duyệt để xử lý dữ liệu người dùng cục bộ trên máy (client-side) mà không cần xử lý phía máy chủ (server-side) để đảm bảo các vấn đề liên quan đến quyền riêng tư và bảo mật.

Từ khóa: mô hình học sâu, phân khu cắt lớp, thuật toán tái tạo 3D, tim mạch

2. Giải nghĩa

Ảnh chụp cắt lớp: Chụp cắt lớp vi tính hay còn gọi là chụp CT là phương pháp sử dụng tia X và hệ thống, phần mềm tái tạo hình ảnh kỹ thuật số để tạo ra những hình ảnh thể hiện rõ cấu trúc, bộ phận, nội quan theo lát cắt ngang của cơ thể. Ảnh chụp này bao gồm các phương pháp MRI (Magnetic Resonance Imaging) và CT (Computed Tomography). Kết quả trả về bao gồm các file dưới định dạng: .dcm, .nii.gz, và .nii.[2]

B. Điều tra khảo sát



Hình 1: Biểu đồ quá trình của một ca mổ tim và những khó khăn bắt cặp

Phân đoạn mô hình tim có thể hỗ trợ phát hiện và ngăn ngừa sớm các bệnh tim như bệnh động mạch vành, đột quỵ, bệnh cơ tim, hở van tim,... []

1. Khó và chuyên sâu.

Việc đọc và hiểu dữ liệu từ ảnh chụp 2D MRI trắng đen là vẫn tương đối khó với bác sĩ, đặc biệt những bác sĩ chưa có nhiều năm kinh nghiệm. Với tỉ lệ xấp xỉ, trong 100 người học thì chỉ có 1-2 người thành công (BS Hữu, bệnh viện Quận 1, TPHCM, 2023). Điều này đồng nghĩa với việc thiếu hụt nhân lực để đáp ứng cho hơn 8000 ca phẫu thuật tim mỗi năm (TS.BS Nguyễn Anh Dũng [3]).

2. Dị tật / dị dạng.

Hiện nay các ca phẫu thuật tim thường rơi vào trường hợp tim bị dị tật, dị dạng bẩm sinh phức tạp (e.g bé Hoàng Lê Khánh Thy – ca ghép tim kỳ lạ []). Với các dị tật thì cấu trúc tim có thể bị thay đổi so với cấu trúc bình thường, có thể tăng nguy cơ sai sót. [4]

3. Phần mềm nước ngoài

Hiện tại Việt Nam chưa có phần mềm nào để giải quyết vấn đề trên. Nhưng trên thế giới, các bệnh viện đã có áp dụng phần mềm cùng mục tiêu tương tự. Nhưng các phần mềm đó khá đắt đỏ (B.S Hữu). Ví dụ phần mềm phổ biến ITK-snap bắt buộc phải thực hiện phân đoạn bằng tay (manual segmentation) tốn thời gian, thiếu nhất quán, tiếng Anh khó hiểu, công cụ quá nhiều gây rối, không có database.

C. Thực hiện nghiên cứu

1. Nhiệm vụ nghiên cứu

- Mô hình học sâu phân khu tim cần đạt độ chính xác xấp xỉ 98-99%, độ sai sót Dice Coefficient (mục 4.22 Metrics) chỉ chấp nhận trên 0.9 cho cả tập dữ liệu huấn luyện (MM-WHS, mục 4.1) và tập dữ liệu thực tế (MM, mục 4.1)

- Mô hình phải gọn nhẹ để load được trên website và chạy được trên laptop ít tài nguyên, không có GPU.
- Thuật toán tái tạo 3D cần vừa chính xác, nhanh, nhẹ, cho phép người dùng có thể tương tác dễ dàng bằng chuột hoặc bằng mắt kính VR kết nối. Không gian 3D cần có chức năng hiển thị/ẩn từng khu riêng biệt trong tim.
- Website cần có một quy trình, phương thức để kết nối với các bác sĩ chuyên khoa đầu ngành xác nhận tính chính xác và an toàn trước khi một quyết định quan trọng nào được đưa ra dựa trên kết quả của ứng dụng trước ca phẫu thuật.
- Website phải hoàn toàn bằng tiếng Việt, hạn chế các tác vụ thừa, có công cụ khó hiểu và tăng cường tính tự động hóa.
- Hệ cơ sở dữ liệu cần tổ chức logic, hợp lý, bảo mật thông tin, an toàn và dễ dàng truy cập khi có tài khoản hợp lệ.

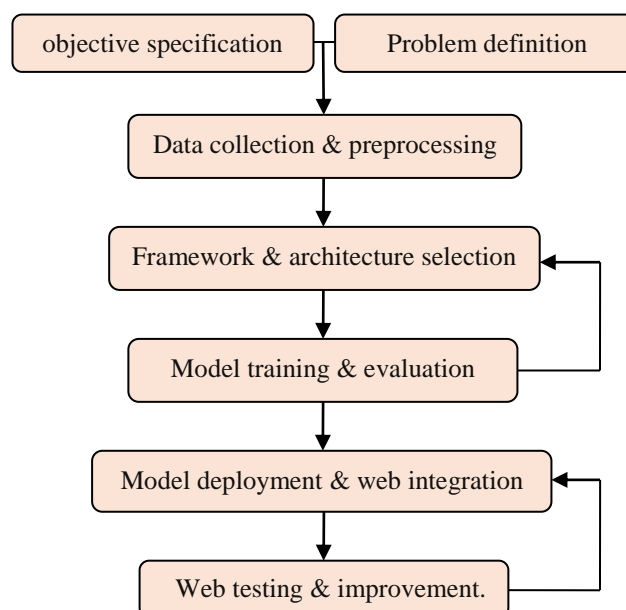
2. Giả thuyết khoa học

Chúng tôi đặt ra giả thuyết sau khi dự án được hoàn thiện:

- Áp dụng cho việc ***hướng dẫn, giảng dạy*** cho sinh viên:
 - o Sinh viên tiếp thu nhanh chóng giải phẫu học, được trải nghiệm với những ca dị tật thật tế khi giảng viên tương tác và hướng dẫn trực tiếp cho sinh viên trên ứng dụng.
 - o Tăng khả năng có nhiều bác sĩ giỏi, giải quyết thiếu nhân lực.
- Áp dụng cho công tác ***lên kế hoạch tiền phẫu thuật***:
 - o Giảm thiểu rủi ro sai sót, giảm thời gian phân tích hình ảnh và tăng tính hiệu quả, tính cộng tác cho nhiều bác sĩ cộng tác trong một ca phẫu thuật.

3. Quy trình nghiên cứu

Chúng tôi áp dụng quy trình “*Deep learning algorithm research for web-based application*” [6]

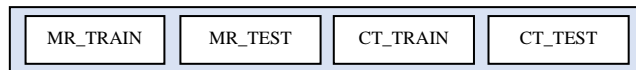


4. Cơ sở lý thuyết

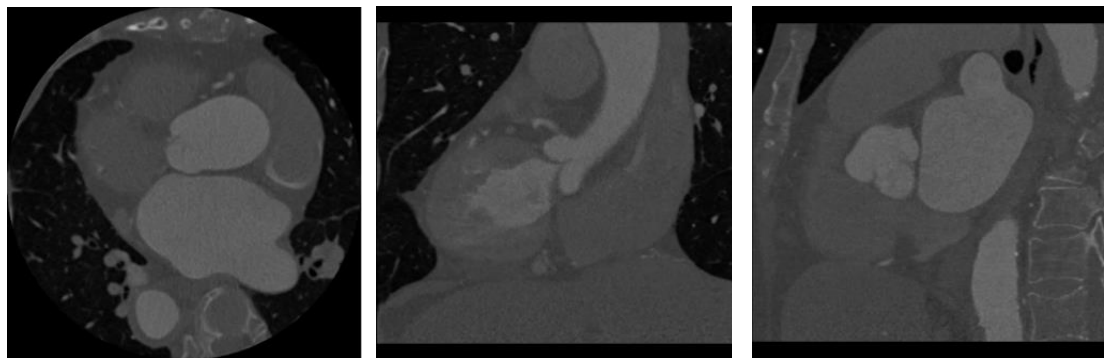
4.1. Chuẩn bị và tìm hiểu dữ liệu

Trong quá trình thử nghiệm, chúng tôi sử dụng 2 bộ dữ liệu (dataset) sau:

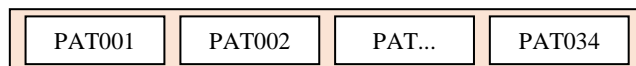
- MM-WHS (Multi-Modality Whole Heart Segmentation): Bộ dữ liệu có 120 bộ ảnh bao gồm 60 bộ ảnh CT/CTA và 60 bộ ảnh MRI [], lưu dưới dạng “.nii.gz” . Bộ dữ liệu được lưu trữ theo cấu trúc:



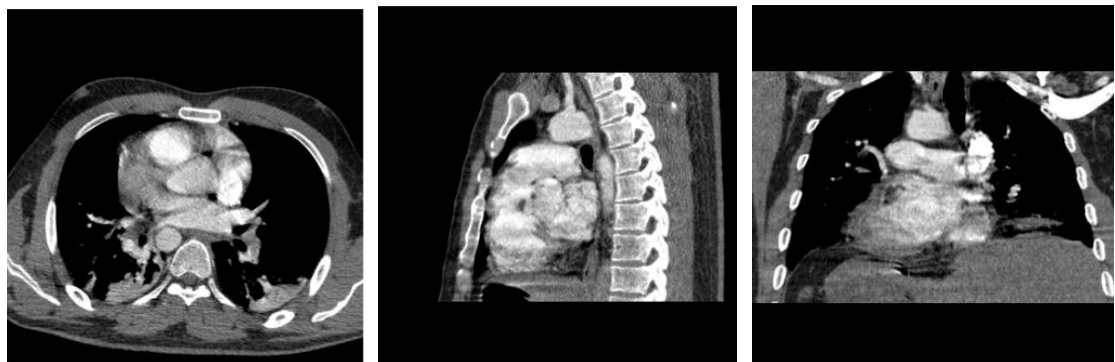
Bộ dataset này đã loại bỏ slice thừa, cắt tròn ở góc nhìn chính để tập trung vào các vùng cần phân khu:



- Bộ dataset của Mojtaba Masoudi (MM): ảnh chụp cắt lớp lòng ngực của 34 bệnh nhân. Trong mỗi bộ ảnh có khoảng từ 180-210 lát cắt lưu dưới dạng “.dcm”[10]. Bộ dữ liệu được lưu trữ theo cấu trúc:



Bộ dataset này sẽ đóng vai trò như thử nghiệm thực tế, vì hình ảnh không được xử lý trước, ảnh chụp toàn cảnh lòng ngực và chứa nhiều hình thừa gây nhiễu:



Như vậy để model của chúng tôi hoạt động tốt ở thực tế thì đòi hỏi phải có quy trình tiền xử lý dữ liệu thô để mô hình cho ra kết quả giống bộ dataset huấn luyện nhất.

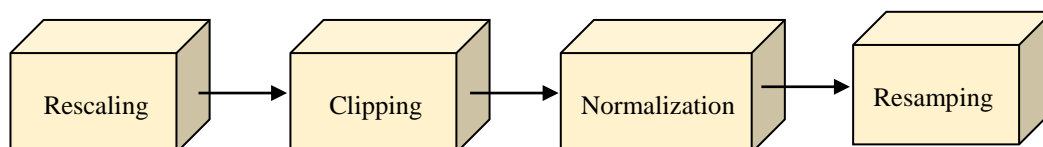
Lưu ý: bộ ảnh (hay còn gọi slices) là tập hợp nhiều ảnh 2D trong một lần chụp cắt lớp. Mỗi bộ ảnh tương ứng với một bệnh nhân.

4.2 Tiền xử lý dữ liệu

Đơn vị Hounsfield (HU) là một đơn vị đo được sử dụng trong các quét CT để định lượng mật độ phóng xạ của các mô. Nhờ đơn vị này, chúng tôi có thể thực hiện tiền xử lý dữ liệu và loại bỏ một số thành phần không liên quan đến tim như phổi, xương.

Mô		Đơn vị Hounsfield (HU)
Không khí		-1000
Phổi		-500 đến -700
Mỡ		-100 đến -50
Nước		0
Máu		+30 đến +45
Cơ bắp		+10 đến +40
Gan		+40 đến +60
Xương	Xương bột	+700
	Xương sụn	+3000

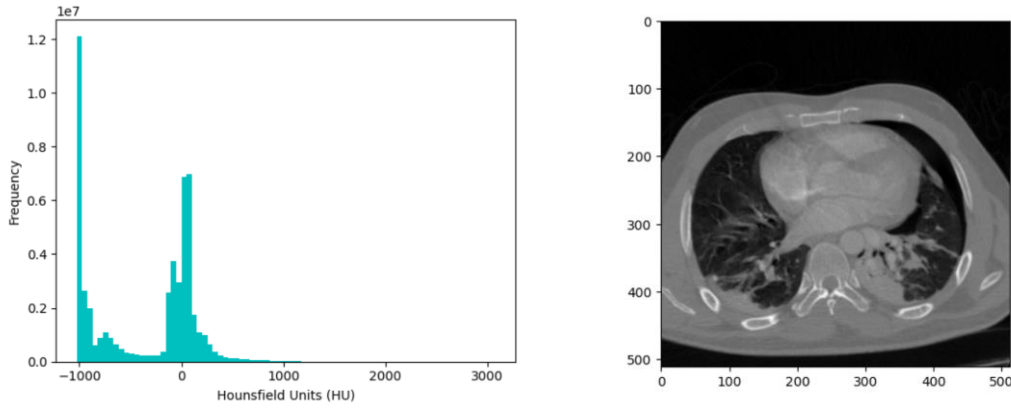
Chúng tôi thực hiện 4 bước tiền xử lý dữ liệu cho pixel values theo chuẩn quy trình 3D Volume Preprocessing[12]:



Rescaling: Các máy chụp cắt lớp khác nhau có thể dẫn đến các phạm vi giá trị pixel khác nhau. Nên chúng tôi cần tiêu chuẩn hóa các giá trị này về thang độ Hounsfield unit.

Gọi tập hợp các lát cắt (slices) là $S = \{x_1, \dots, x_t, \dots, x_N\} = \{x^t\}_{t=1}^N$. Trong đó $x_t \in R^{w \times h}$ là một lát cắt (2D numpy array) có chiều dài w và chiều cao h .

Chúng tôi dùng công thức $x_t = x_t \times \text{slope} + \text{intercept}$ để chuẩn hóa mật độ phóng xạ về một đơn vị Hounsfield thống nhất.



Hình 2: Giá trị pixel của ảnh khi chuyển về thang độ HU

Clipping: Các thành phần khác trong tấm ảnh có thể gây nhiễu cho mô hình trong quá trình huấn luyện nên chúng tôi loại bỏ một số thành phần không liên quan đến tim như phổi, xương bằng phương pháp **filtering** với **numpy array** và cắt tròn góc nhìn chính:

Normalization: Các mô hình học sâu hay mạng neuron nhân tạo chỉ có thể học tốt và hội tụ nhanh khi các giá trị xử lý nằm trong khoảng nhỏ, từ (0,1) hoặc từ (-1,1).

Nên chúng tôi normalize bộ data từ scale (-1000, 1000) về (0,1) với công thức:

$$X_{norm} = a + \frac{(X - X_{min})(b - a)}{X_{max} - X_{min}}$$

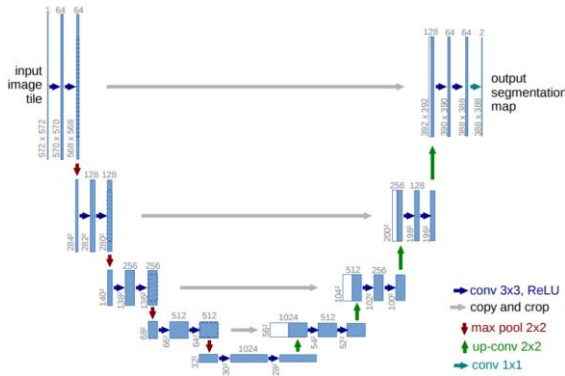
Trong đó **(a,b) = (0,1)** nên công thức sau cùng sẽ là:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

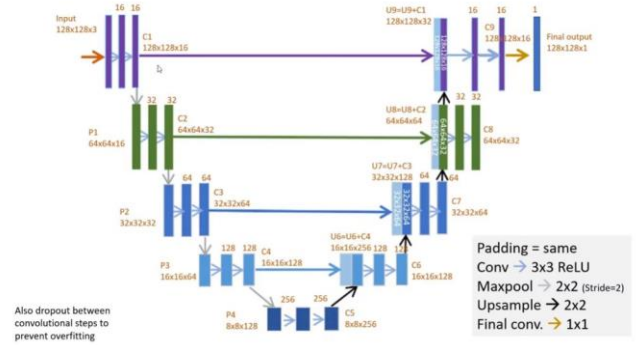
Resampling: Mô hình chỉ nhận đầu vào với cùng một kích thước. Tuy nhiên để thử nghiệm bộ data khác với kích thước khác (do máy chụp) thì cần đưa về cùng một kích thước (cùng voxel size). Sử dụng module **spice.ndimage** trong thư viện **spicy**.

4.3 Xây dựng kiến trúc và huấn luyện mô hình

Chúng tôi bắt đầu quá trình thử nghiệm với mô hình thị giác máy tính (computer vision) truyền thống Unet với kiến trúc bao gồm 2 phần: encoder và decoder tạo thành hình chữ U đặc trưng. [12] Tuy nhiên vì tài nguyên ổ nhớ có hạn nên chúng tôi đã hiệu chỉnh các thông số của kiến trúc: giảm kích thước đầu vào (572, 572) thành (128,128).



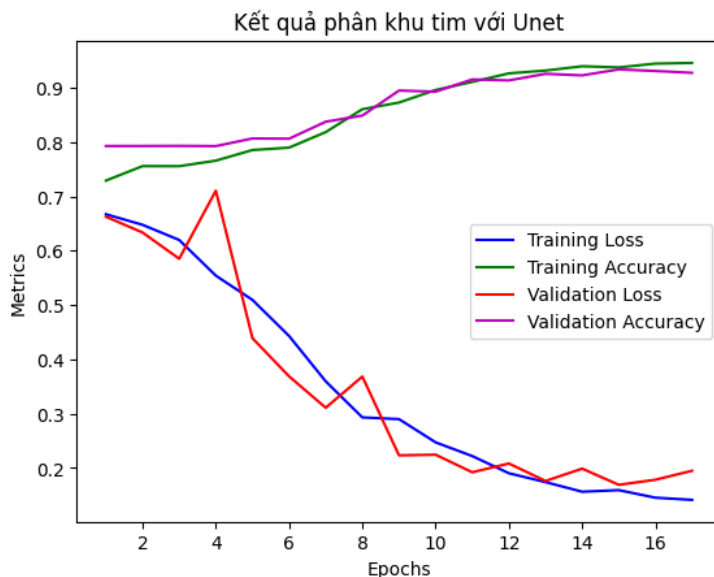
Kiến trúc Unet cũ



Kiến trúc Unet hiệu chỉnh

Tuy nhiên cả hai kiến trúc trên đều gặp chung một vấn đề: không nhận diện được chính xác các cạnh của các khu khi hình dạng liên tục thay đổi kích thước, nhầm lẫn các khu (classes) với nhau. Vì thế chúng tôi tiếp cận một phương pháp hiện đại hơn: Unet tích hợp attention Sự thay đổi này giúp cho mô hình bỏ qua các thành phần không quan trọng, nhận diện sự thay đổi hình thù của khu tốt hơn (shape-aware). [13].

Thông số tham số của mô hình	
Total params:	1941105 (7.40 MB)
Trainable params:	1941105 (7.40 MB)
Non-trainable params:	0 (0.00 Byte)

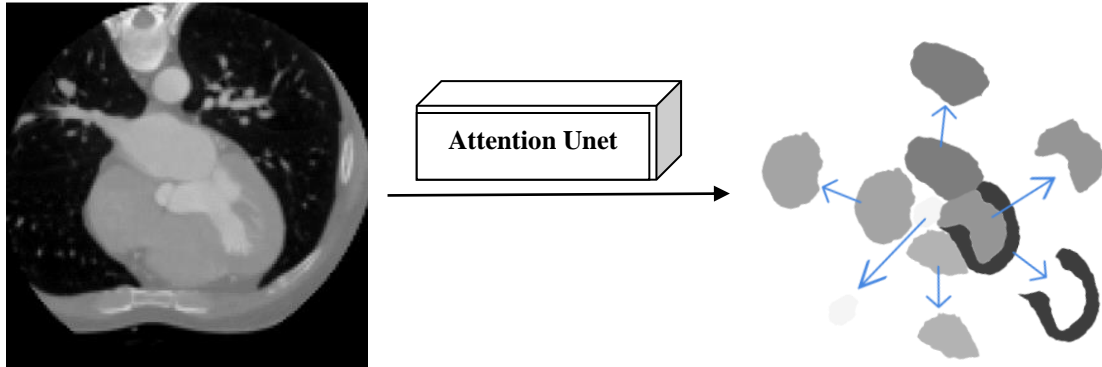


Chúng tôi cho “fit” với 25 epochs và cài đặt early stopping với callback trong thư viện tensorflow. Vì vậy ở epoch 16, khi Validation Loss có dấu hiệu tăng (nghĩa là mô hình đang học thuộc lòng bộ dữ liệu) thì quá trình training kết thúc để đảm bảo tính tổng quát hóa (generalizability) cho mô hình.

Hình 3. Quá trình huấn luyện mô hình học sâu

4.4 Đánh giá mô hình (Metrics)

Chúng tôi dùng phương pháp phổ biến để đánh giá mô hình cho bài toán phân khu nhiều lớp cho dữ liệu y học (medical multiclass segmentation) là DSC (Dice Similarity Coefficient) [14]

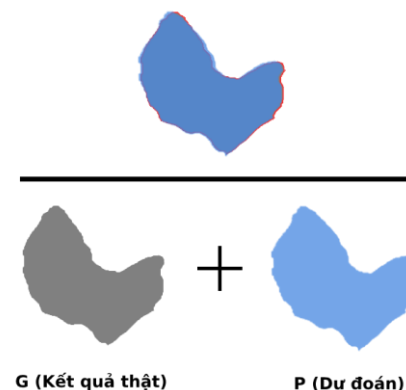


Dice Coefficient: Sau khi có kết quả dự đoán như hình trên thì Chúng tôi tính Dice trên từng class (từng vùng một). Bên dưới là ví dụ tính Dice cho tâm thất trái: so sánh hai hình ảnh trên từng pixel: kết quả thật (G - ground truth) và dự đoán của mô hình (P - prediction).

$$Dice = \frac{2|P \cap G|}{|P| + |G|}$$

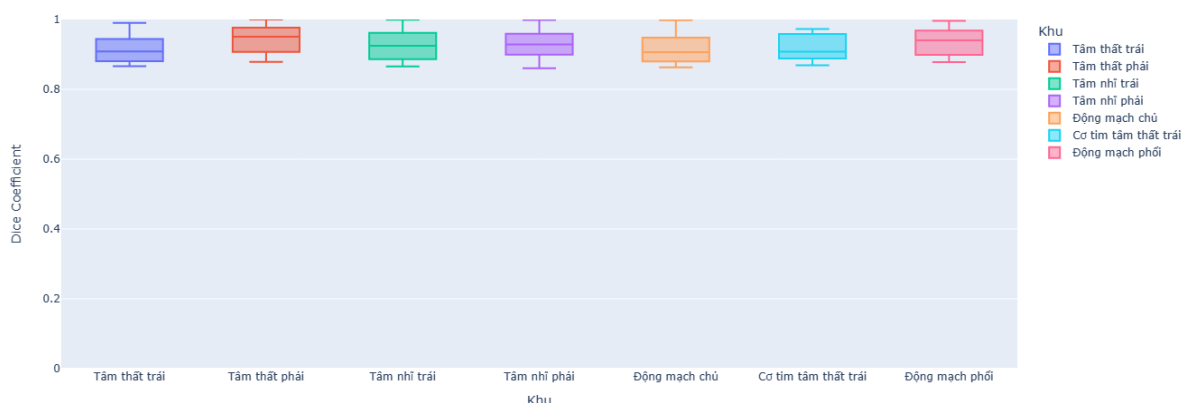
Chia phần giao nhau (intersection)
cho phần tổng (Union)

Ở phần code chúng tôi cộng thêm 1 vào tử và mẫu để tránh kết quả trả về 0 hoặc chia cho mẫu 0.



Dice Coefficient **tiến về 0** cho thấy không có sự trùng lặp giữa hai bộ (nghĩa là kết quả phân đoạn **P** và kết quả thật **G** hoàn toàn không khớp nhau).

Dice Coefficient **tiến về 1** cho thấy sự khớp hoàn hảo (tức là kết quả phân đoạn **P** và kết quả thật **G** khớp hoàn hảo).



Hình 4. Biểu đồ Boxplot cho Dice Coefficient 7 khu

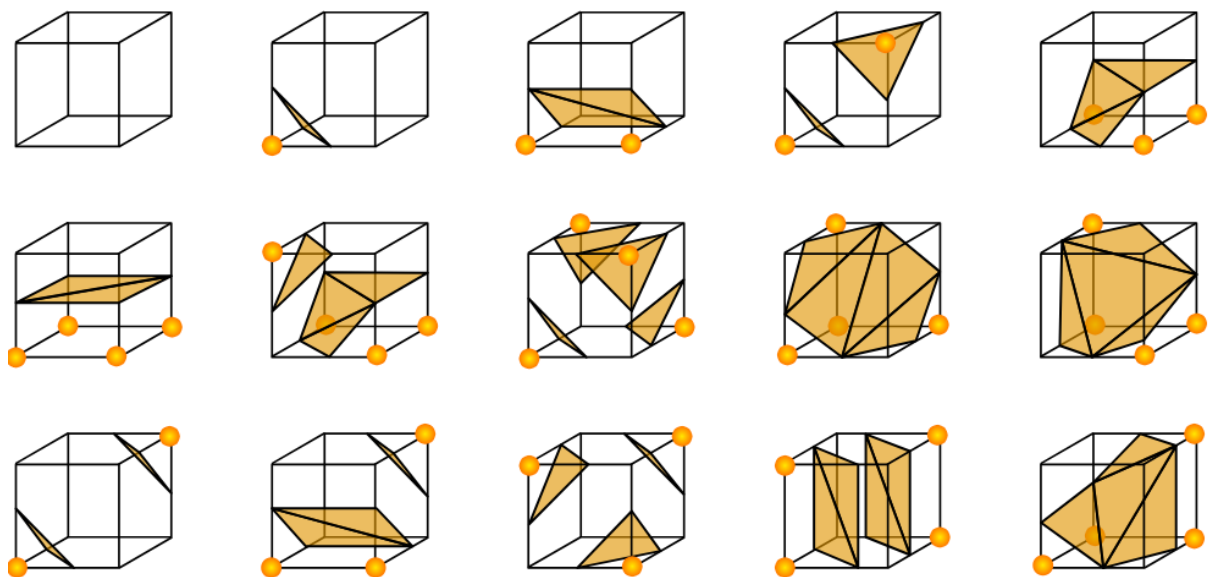
Khu	Dice Coefficient cho 20 bệnh nhân / 20 slices								lower_fence	upper_fence
	count	mean	std	min	25%	50%	75%	max		
Cơ tim tâm thất trái	20	0.93302	0.042186	0.864674	0.901512	0.933976	0.96989	0.991049	0.798946	1.072456
Tâm nhĩ phải	20	0.929527	0.04093	0.86328	0.892338	0.93238	0.954225	0.999991	0.799508	1.047055
Tâm nhĩ trái	20	0.925144	0.038813	0.865499	0.893871	0.917029	0.958943	0.999968	0.796262	1.056552
Tâm thất phải	20	0.940627	0.036126	0.864289	0.918125	0.93834	0.963854	0.999686	0.849532	1.032447
Tâm thất trái	20	0.929294	0.032247	0.880099	0.895661	0.927914	0.960857	0.977849	0.797867	1.058651
Động mạch chủ	20	0.919959	0.042516	0.865959	0.884811	0.913828	0.954293	0.994028	0.780589	1.058515
Động mạch phổi	20	0.934415	0.041747	0.86063	0.902986	0.927965	0.9632	0.994484	0.812665	1.05352

Hình 5. Bảng thống kê Dice coefficient của model Attention Unet cho 20 slices

4.5 Thuật toán tái tạo 3D

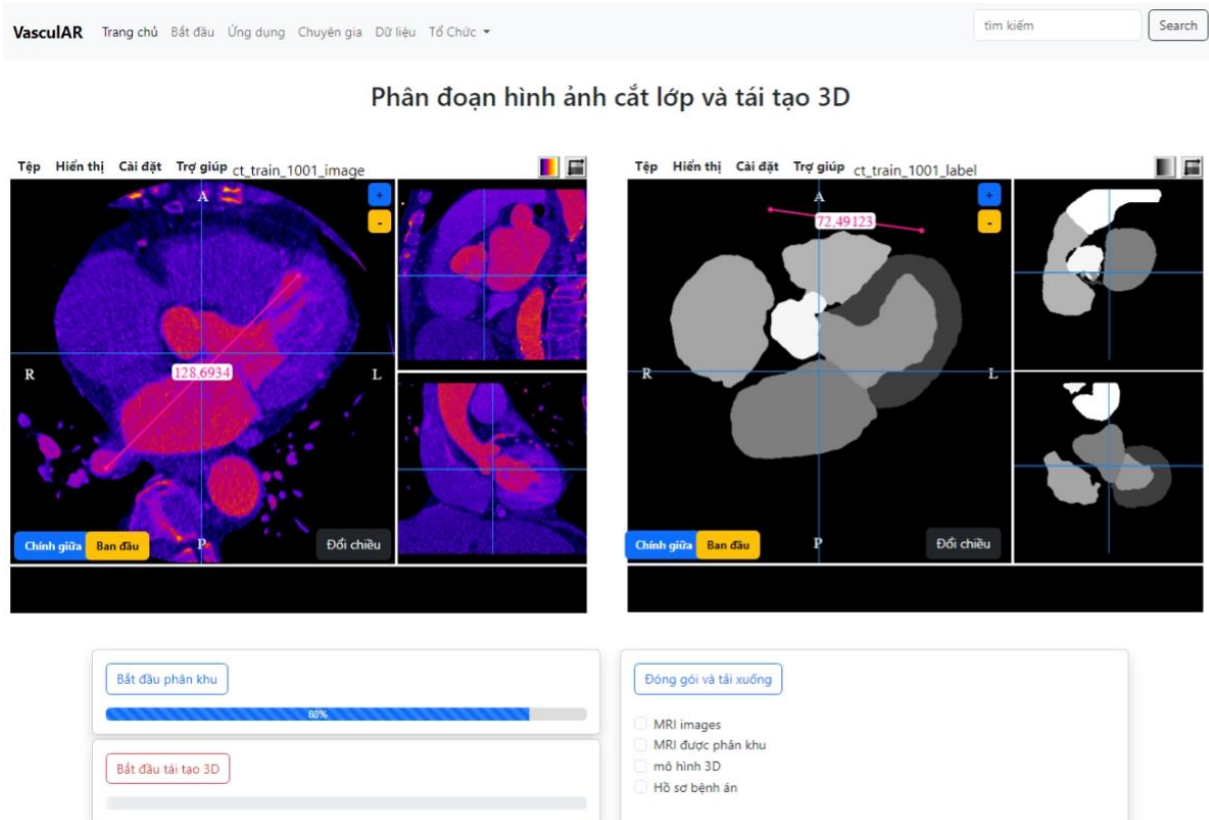
Nhóm nghiên cứu thuật toán **Marching Cubes** được xuất bản trong cuốn kỷ yếu SIGGRAPH bởi William E. Lorensen and Harvey E. Cline [15].

Về mặt toán học, thuật toán sử dụng “Chia để trị” (Divide and Conquer)[] bằng cách lặp qua, tìm kiếm các khu vực vượt qua một mức độ quan tâm nhất định. (level of interest) Nếu các vùng như vậy được tìm thấy, tam giác được tạo và thêm vào lưới đầu ra. Kết quả cuối cùng là một tập hợp các đỉnh và một tập hợp các mặt tam giác. Các mặt tam giác nhau được liên kết với nhau trong không gian 3D tạo thành mạng lưới (meshing extraction) [16].

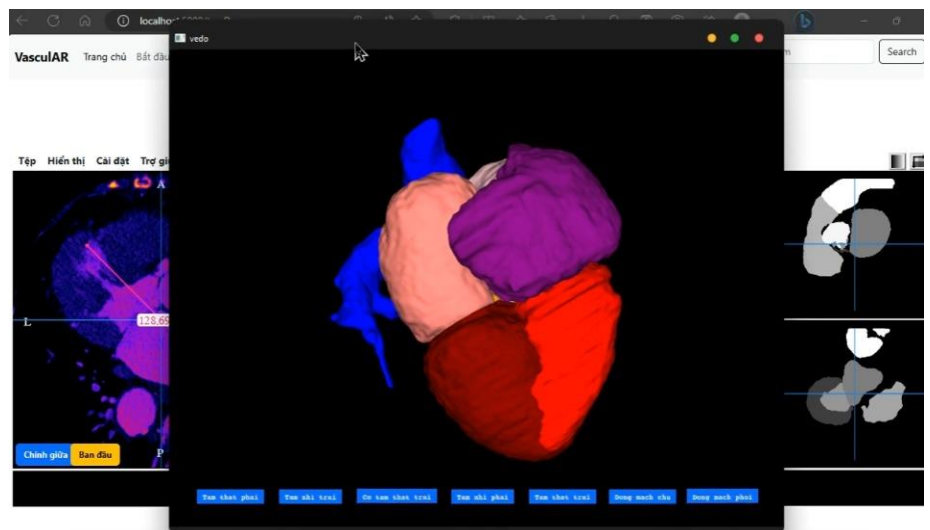
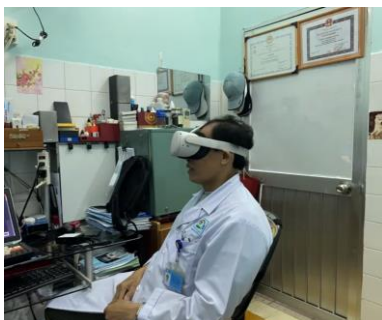


D. Kết quả nghiên cứu (sản phẩm)

Nền tảng website tiếng Việt giao diện dễ dùng tích hợp các chức năng để hỗ trợ bác sĩ đọc ảnh chụp cắt lớp dễ dàng với phổ màu, thước đo, hệ tọa độ. Tích hợp được model deep learning Unet lên browser để thực hiện phân khu



Tái tạo cấu trúc tim 3D trong không gian 3 chiều với các nút hiển thị/ẩn từng khu riêng biệt, sẵn sàng kết nối với mắt kính thực tế ảo VR.



Hệ cơ sở dữ liệu (database) cho phép cơ sở y tế đăng ký tài khoản và lưu trữ hồ sơ bệnh án.

Kết nối các bác sĩ chuyên gia tim mạch đầu ngành hỗ trợ kiểm chứng độ an toàn của kết quả trả về, từ đó đảm bảo thêm độ an toàn của kết quả cuối cùng

(Đang trong quá trình thực hiện)

E. Mục tiêu tương lai

Nhóm nghiên cứu quyết định tiến hành các mục tiêu trong tương lai sau đây:

- Nâng cấp, hiệu chỉnh kiến trúc mô hình học sâu để phân khu và tái tạo các cấu trúc chi tiết như mạch máu, van tim,.. đồng thời tăng tính chính xác (tiệm cận 98-100%.) cho mô hình học sâu và thuật toán dựng 3D.
- Phối hợp với các đơn vị y tế để tìm kiếm dữ liệu ảnh chụp cắt lớp tim của người Việt Nam để hỗ trợ nhóm đối tượng bệnh nhân trong nước một cách tốt nhất.
- Nâng cấp hệ thống cơ sở dữ liệu để đáp ứng nhu cầu quản trị tập dữ liệu lớn (Big Data) cũng như thực hiện tăng cường hệ thống bảo mật, đảm bảo thông tin an toàn.

F. Trích Dẫn

- [1] Cardiac CT: Indications and Limitations Susanna Prat-Gonzalez, Javier Sanz, Mario J. Garcia Journal of Nuclear Medicine Technology Mar 2008, 36 (1) 18-24; DOI: 10.2967/jnmt.107.042424
- [2] Vì sao cần chụp cắt lớp vi tính, Trần Công Trình - Bác sĩ Chẩn đoán hình ảnh - Khoa Chẩn đoán hình ảnh - Bệnh viện Đa khoa Quốc tế Vinmec Central Park.
- [3] Những điều cần biết về phẫu thuật tim, TS.BS Nguyễn Anh Dũng – Trưởng khoa Phẫu thuật tim – Lồng ngực, Mạch máu – Trung tâm Tim mạch, Bệnh viện Đa khoa Tâm Anh, TP.HCM

- [4] Phẫu thuật tim ở Việt Nam đang bước lên đẳng cấp mới, PGS-TS Nguyễn Quang Tuấn - Giám đốc BV Tim Hà Nội
- [5] User-Guided {3D} Active Contour Segmentation of Anatomical Structures: Significantly Improved Efficiency and Reliability, Paul A. Yushkevich, 2006
- [6] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press
- [7] Xiahai Zhuang: Multivariate mixture model for myocardial segmentation combining *multi-source* images. IEEE Transactions on Pattern Analysis and Machine Intelligence 41(12): 2933-2946, 2019
- [8] Xiahai Zhuang and Juan Shen: Multi-scale patch and *multi-modality* atlases for whole heart segmentation of MRI, Medical Image Analysis 31: 77-87, 2016
- [9] X Luo & X Zhuang: X-Metric: An N-Dimensional Information-Theoretic Framework for Groupwise Registration and Deep Combined Computing. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022-11 online published (IF: 24.314)
- [10] Masoudi, Mojtaba (2018). Patient01-16. figshare. Dataset. <https://doi.org/10.6084/m9.figshare.6265292.v3>
- [11] CT-scan images preprocessing and segmentation to improve bioprosthesis leaflets morphological analysis, P. Haigron, Q. Wang, S. Esneault, R. Madeleine, J.F. Heautot, A. Leguerrier, J.-Ph. Verhoye
- [12] . Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional net works for biomedical image segmentation," in International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2015, pp. 234–241.
- [13] Attention U-Net: Learning Where to Look for the Pancreas, arXiv:1804.03999v3 [cs.CV] 20 May 2018
- [14] Continuous Dice Coefficient: a Method for Evaluating Probabilistic Segmentations > cs > arXiv:1906.11031
- [15] Lorensen, William E.; Cline, Harvey E. (1 August 1987). "Marching cubes: A high resolution 3D surface construction algorithm". *ACM SIGGRAPH Computer Graphics*. **21** (4): 163–169.
- [16] "System and method for the display of surface structures contained within the interior region of a solid body". 5 June 1985