

SỞ GIÁO DỤC VÀ ĐÀO TẠO TP. HCM
CUỘC THI NGHIÊN CỨU KHOA HỌC KỸ THUẬT CẤP THÀNH PHỐ
HỌC SINH TRUNG HỌC NĂM HỌC 2023-2024

DỰ ÁN

PHẦN MỀM TÍCH HỢP HỌC SÂU ĐỂ PHÂN VÙNG VÀ TÁI TẠO CẤU TRÚC
TIM 3 CHIỀU CHO ỨNG DỤNG THỰC HÀNH Y KHOA

LĨNH VỰC: PHẦN MỀM HỆ THỐNG MÃ DỰ ÁN: 21_1033_02

Giáo viên hướng dẫn: Đỗ Quốc Anh Triết

Tác giả 1: Nguyễn Lê Quốc Bảo
Tác giả 2: Đỗ Xuân Huy

LỜI CẢM ƠN

Cảm ơn bác sĩ Lê Văn Nghĩa (trưởng khoa tim mạch bệnh viện Chợ Rẫy) đã chỉ dẫn những vấn đề quan trọng liên quan tới tim và ảnh chụp cắt lớp.

Cảm ơn bác sĩ Lê Văn Phước (Trưởng khoa Chẩn đoán hình ảnh Bệnh viện Chợ Rẫy) đã thực hiện, tư vấn và hướng dẫn phân vùng bằng tay trên bộ ảnh cắt lớp lòng ngực thô.

Cảm ơn bác sĩ Trần Văn Hữu (BS chuyên khoa 2, trưởng khoa ngoại tổng hợp bệnh viện Quận 1) đã tư vấn hướng phát triển và ứng dụng thực tế cho dự án.

Cảm ơn bác sĩ Lê Thị Phương Nga BS khoa nội, bệnh viện Nguyễn Trãi, TPHCM đã hỗ trợ kiến thức về giải phẫu lồng ngực, giải phẫu tim người.

Cảm ơn anh Đoàn Văn Tuấn (cựu sinh viên trường Đại học Kỹ Thuật Y Tế Hải Dương, khoa chuẩn đoán hình ảnh) đã giúp giới thiệu form khảo sát đến cho các bạn sinh viên khác.

TÓM TẮT

Trong ngành kỹ thuật chuẩn đoán hình ảnh, chuyên khoa tim mạch, việc đọc hiểu và phân tích chính xác ảnh chụp cắt lớp tim lồng ngực gây nhiều khó khăn cho sinh viên Việt Nam do sự hạn chế về thực hành, kỹ thuật. Để giải quyết vấn đề này, nhóm phát triển một phần mềm hệ thống có tên **VAS** trên nền tảng trực tuyến (website) tích hợp các công cụ phân tích nhiều loại ảnh cắt lớp, và tái tạo cấu trúc tim 3D hoàn toàn tự động. Các mô hình học sâu (Deep Learning) đã được nghiên cứu như Unet, VGG, Resnet và các hàm mất mát IoU, Dice Coefficient, Jaccard cho công việc phân vùng (segmentation) hình ảnh cắt lớp tim. Các kiến trúc hình học đã được áp dụng như phương trình đường tròn cho quá trình xử lý dữ liệu và hình học không gian cho thuật toán Marching Cubes để trích xuất đặc trưng kết quả phân vùng, 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. Database SQL Alchemy với Python Backend được tích hợp cho nền tảng ứng dụng để hỗ trợ việc lưu trữ ghi chú, phân tích, kế hoạch tiền phẫu thuật cho chuyên viên chẩn đoán. Trong quá trình thử nghiệm, mô hình phân vùng đa lớp (multiclass model) sử dụng kiến trúc Unet kết hợp attention 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 biến đổi bằng Tensorflow.js giảm độ nặng, tiết kiệm tài nguyên máy tính trong một lần chạy, tăng tốc độ cho ra kết quả trên nền tảng website nhanh hơn 20% so với các phần mềm khác.

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

MỞ ĐẦU

1. Lý do chọn đề tài

Theo Tổ chức Y tế Thế giới, bệnh tim mạch là nguyên nhân gây tử vong hàng đầu trên toàn cầu (Mendis et al., 2011) [1]. Theo thống kê, Việt Nam có khoảng 200.000 người tử vong vì bệnh tim mạch. Số ca tử vong do các bệnh tim mạch cao hơn cả tử vong do ung thư, hen phế quản và đái tháo đường cộng lại [2]. Trong ca phẫu thuật tim, công tác chuẩn đoán hình ảnh đóng vai trò quan trọng trong việc lập kế hoạch tiền phẫu thuật. Một sai sót trong việc đọc hiểu các ảnh cắt lớp có thể dẫn đến di chứng khó lường, đặc biệt với những ca dị tật, di dạng bẩm sinh (e.g bé Hoàng Lê Khánh Thy – ca ghép tim kỳ lạ [3]). Thế nhưng theo điều tra khảo sát, việc học khoa chuẩn đoán hình ảnh vẫn có nhiều khó khăn cho sinh viên và giảng viên trong công việc giảng dạy. Cụ thể, sinh viên khoa chuẩn đoán hình ảnh ít được tiếp xúc, thực hành với việc đọc và phân tích các hình ảnh cắt lớp thực tế, đa số là tập đọc và học qua sách vở với các tình huống có sẵn. Tuy nhiên thực tế cho thấy tim có thể có nhiều biến chứng bất thường gây khó khăn cho kỹ thuật

viên mới ra trường tiếp xúc những ca phân tích hình ảnh cho việc lên kế hoạch tiền phẫu thuật tim. Sinh viên ở Việt Nam hiện nay phần lớn tải các phần mềm nước ngoài để hỗ trợ việc tự nghiên cứu thêm. Tuy nhiên, các phần mềm đó đắt đỏ, chưa tự động hóa hoàn toàn, yêu cầu laptop cấu hình cao và các dữ liệu không tập trung vào nhóm đối tượng người Việt Nam. Để giải quyết nhu cầu này, nhóm nghiên cứu đã chọn đề tài.

2. Mục tiêu nghiên cứu

- Huấn luyện mô hình học sâu phân vùng tim cần đạt độ chính xác xấp xỉ 98-99%, chỉ số Dice Coefficient Score chỉ chấp nhận trên 0.9.
- Lập trình 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 vùng riêng biệt trong tim.
- Lập trình website hoàn toàn bằng tiếng Việt, hạn chế các tác vụ thừa, có công cụ khó hiểu và tăng cường tính tự động hóa.
- Lập trình thêm cơ sở dữ liệu lưu trữ lại ghi chú của người dùng khi phân tích một file ảnh chụp cắt lớp cụ thể.

3. Giả thuyết khoa học

- Nếu nghiên cứu và sản phẩm được hoàn thiện sinh viên sẽ nâng cao khả năng đọc hiểu ảnh cắt lớp, tăng trải nghiệm phân tích dữ liệu thô thực tế, tăng khả năng hình dung, tiết kiệm thời gian. Giảng viên có thể áp dụng nền tảng ứng dụng này trong việc giảng dạy chuẩn đoán hình ảnh trực quan hơn cho sinh viên trong không gian ảo. Điều này giúp cải thiện số lượng và chất lượng bác sĩ chẩn đoán hình ảnh, phục vụ tốt công tác lập kế hoạch tiền phẫu thuật tim, đáp ứng nhu cầu cho hơn 8000 ca phẫu thuật tim tại Việt Nam. [4]

4. Nhiệm vụ nghiên cứu

- Nghiên cứu về các định dạng file, đơn vị Hounsfield, xử lý nhiễu.
- Nghiên cứu về các vùng trong cấu trúc tim và các bệnh lý liên quan đến tim.
- Nghiên cứu các mô hình học sâu tốt cho việc chuẩn đoán hình ảnh y học.
- Nghiên cứu thuật toán hình học không gian tái tạo cấu trúc tim 3 chiều.
- Nghiên cứu môi trường thực tế ảo (VR) cách liên kết với ứng dụng.

5. Câu hỏi nghiên cứu

C1: Làm sao để huấn luyện mô hình học sâu đạt kết quả tốt nhất.

C2: Làm sao để xây dựng một website tự động hóa đầy đủ các tính năng, công cụ.

C3: Làm sao để xây dựng một cơ sở dữ liệu an toàn, bảo mật, lưu trữ xử lý dữ liệu lớn.

6. Phương pháp nghiên cứu

- Phương pháp quan sát và điều tra khảo sát.
- Phương pháp thống kê số liệu và đặt giả thuyết.
- Phương pháp thực nghiệm và quy trình thiết kế ứng dụng.

7. Tính mới và tính sáng tạo của đề tài

- Nền tảng ứng dụng đầu tiên chạy trên website tại Việt Nam sử dụng các mô hình học sâu ứng dụng cho nhiệm vụ phân tích chuyên sâu về ảnh cắt lớp tim mạch, phân vùng và tái tạo cấu trúc tim 3D một cách tự động hóa hoàn toàn.
- Cùng bác sĩ, chuyên gia chẩn đoán hình ảnh tạo ra một bộ data mới với số vùng nhiều và chi tiết nhất hiện nay (12 vùng), trên thế giới chỉ tối đa 7 vùng. [5]
- Kết nối môi trường thực tế ảo vào trong việc giảng dạy.

CHƯƠNG 1: TỔNG QUAN

1.1 Các công trình nghiên cứu phần mềm ứng dụng trên thế giới và Việt Nam

1.1.1 Tại Việt Nam

Theo khảo sát tổng quan, tại Việt Nam đã có dự án liên quan đến phần mềm hỗ trợ đọc phim - VinDir của Vingroup. Phần mềm thực hiện khoanh vùng những tổn thương (vd: gai xương, gai cột sống) trong phim chụp (X-quang, CT, MRI) giúp cho bác sĩ. Phần mềm tập trung chủ yếu vào phần phân tích xương, cột sống, sọ não (đạt độ chính xác trung bình 90%). [6]

Tuy nhiên, Phần mềm Vindir không tái tạo cấu trúc nội quan 3D, không tập trung về phân tích hình ảnh chụp cắt lớp tim. Một bất cập khá lớn là sự đồng thuận giữa bác sĩ và phần mềm là 87%. Điều này nghĩa là bác sĩ hoàn toàn có thể nhìn ra những tổn thương đó và 13% bất đồng còn lại giữa “máy” và bác sĩ vẫn đặt ra một vấn đề khá lớn. [6]

1.1.2 Trên thế giới

Theo khảo sát của nhóm với sinh viên khoa chuẩn đoán hình ảnh: tại Việt Nam chưa có phần mềm nào có chức năng phân vùng cấu trúc và tái tạo mô hình tim 3D một cách tự động (automatic) trên nền tảng website, được truy cập dễ dàng với laptop cá nhân. Sinh viên thường lựa chọn sử dụng phần mềm “lậu” (crack) Radiantview với mức giá 4-5 triệu cho 5 năm sử dụng. Chi tiết về các phần mềm hiện có trên thế giới được mô tả chi tiết ở bảng sau:

Phần mềm trên thế giới	Mô tả hạn chế	Kết quả	Giá tiền
RadiantViewer	Hoàn toàn không sử dụng trí tuệ nhân tạo trong việc tái tạo cấu trúc 3D. Phần mềm sử dụng phương pháp phân loại đơn vị Hounsfield để phân biệt xương, mô, cơ.	Kết quả hiển thị tái tạo cấu trúc 3D là một khối đặc, không thể xem và phân tích từng vùng riêng biệt, bị gây nhiễu bởi các phần không quan trọng khác như khung xương.	179 EUR 4.733.393 VNĐ cho giới hạn 5 năm sử dụng.
ITK-snap	Hoàn toàn phải phân vùng các thành phần cấu trúc tim bằng tay (bút vẽ) → Không tự động hóa	Người sử dụng phải bỏ ra ít nhất 1-2 ngày để tự tay phân chia vùng → mất thời gian, thiếu chính xác khi người sử dụng là sinh viên ít kinh nghiệm.	Miễn phí
Materialise	Bán tự động hóa (semi-auto segmentation). Ở phần mềm này, người dùng vẫn phải thực hiện phân vùng bằng tay một phần.	Người sử dụng vẫn phải mất thời gian để tự tay phân chia vùng → mất thời gian, thiếu chính xác khi người sử dụng là sinh viên ít kinh nghiệm.	5,091,240 VNĐ/tháng Hoặc 4,068,120 VNĐ/tháng cho một năm sử dụng

Các phần mềm nước ngoài trên đều có giao diện tiếng Anh với các thuật ngữ chuyên ngành gây khó hiểu, các công cụ dày đặc phức tạp vì chưa hoàn toàn tự động hóa các bước xử lý. Vì thế, VAS được phát triển với các tính năng nổi trội hơn như sau.

Tính năng	RadiantViewer	ITK-snap	Materialise	VAS
Sử dụng mô hình học sâu	Không	Không	Có	Có
Tự động hoàn toàn	Có	Không	Không	Có
Phân biệt từng vùng riêng biệt	Không	Có	Có	Có
Tiếng Việt dễ hiểu, dễ sử dụng, tiết kiệm thời gian	Không	Không	Không	Có

Miễn phí cho sinh viên	Không	Có	Không	Có
Nền tảng website, không trực tiếp tải về, không tốn nhiều tài nguyên máy tính.	Không	Không	Không	Có
Kết nối thiết bị thực tế ảo (kính VR)	Có	Không	Không	Có

1.1.3 Thống kê điều tra khảo sát

Nhóm thực hiện thống kê khảo sát về việc giảng dạy và học tập của sinh viên trong khoa chuẩn đoán hình ảnh ở trường đại học Kỹ Thuật Y Tế Hải Dương và đại học Y Dược TPHCM. Kết quả hiện nay có 15 người điền form khảo sát cho thống kê như sau:

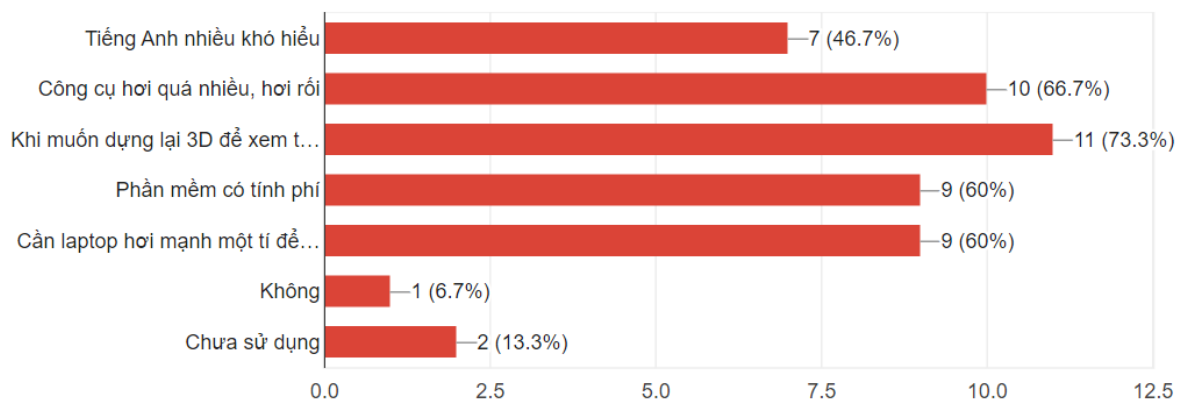
Khó khăn chung:

- Tài liệu sách giấy, bài giảng điện tử không đủ đáp ứng nhu cầu học thực tiễn.
- Thời gian học ngắn, thời gian thực tập ở các máy ở bệnh viện không đủ.
- Kiến thức, ảnh chụp học từ sách vở rập khuôn, không trang bị được cho người học xử lý những ca bất thường trong thực tế như dị tật, di dạng.
- Khó khăn trong việc hình dung trong việc đọc các tấm ảnh cắt ngang 2 chiều.
- Các phần mềm hỗ trợ thì phức tạp, đòi hỏi nhiều thao tác, yêu cầu kỹ thuật cao.

Thống kê:

- 53.3% người trả lời cho rằng mức độ khó của việc đọc hiểu ảnh cắt lớp lồng người người là mức độ 4 (khó), còn lại 46.7% cho là mức độ 5 (rất khó).
- 40% người trả lời chưa sử dụng các phần mềm hỗ trợ và 80% đã sử dụng phần hỗ trợ.
- Trong 80% đó, có 40% người sử dụng RadiantViewer, 46.7% người đến bệnh viện thực hành, 13.3% sử dụng ITK-SNAP, 0% sử dụng Materialise.
- Trong 46.7% người đến bệnh viện thực hành có 85.7% cho rằng số lần đi thực hành không được nhiều và 100% người cho rằng mỗi lần đi rất đông sinh viên nên ít có cơ hội thực hành.
- Đánh giá của người điền đơn về phần mềm hiện họ đang sử dụng:

46.7% thấy tiếng Anh khó hiểu, 66.7% thấy công cụ nhiều và rối, 73.3% cho rằng phân vùng bằng tay (không tự động) mất thời gian, 60% sử dụng phần mềm tính phí, 60% cho rằng để sử dụng cần laptop mạnh.



- 100% người điền đơn cho rằng ứng dụng tiếng Việt, tự động hóa hoàn toàn, tích hợp các công cụ phân tích ảnh cắt lớp, chức năng tái tạo cấu trúc 3D tự động, nhanh và nhẹ, sử dụng trên website không cần tải, miễn phí cho người Việt là hữu ích.

* Xem chi tiết kết quả, biểu đồ, thông số thống kê của cuộc khảo sát trên ở [link này](#).

CHƯƠNG 2: CƠ SỞ LÝ THUYẾT

2.1. Bộ dữ liệu

2.1.1 Bộ dữ liệu MM-WHS (Multi-Modality Whole Heart Segmentation):

Nguồn gốc: Dữ liệu CT/CTA tim được thu thập từ hai máy chụp CT 64 lát cắt hiện đại (Philips Medical Systems, Hà Lan) sử dụng quy trình chụp CT tiêu chuẩn tại Thượng Hải, Trung Quốc. Dữ liệu MRI tim được lấy từ hai bệnh viện ở London, Vương quốc Anh. Một bộ dữ liệu được thu thập từ Bệnh viện St. Thomas trên máy quét Philips 1.5T (Philips Healthcare, Best, The Netherlands) và từ bệnh viện Royal Brompton trên máy chụp Siemens Magnetom Avanto 1.5T (Siemens Medical Systems, Erlangen, Đức). [5]

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:



Nhược điểm: Bộ dữ liệu đã được xử lý trước quá kỹ: loại bỏ các thành phần không liên quan như phổi, bo tròn góc nhìn. Điều này có nghĩa mô hình huấn luyện trên tập dữ liệu này sẽ không đảm bảo tốt với dữ liệu thô của người Việt Nam. Ngoài ra, bộ dữ liệu này được phân vùng gán nhãn bằng tay bằng phần mềm ITK-SNAP (Yushkevich et al., 2006) chỉ có 7 vùng cơ bản: tâm thất trái/phải, tâm nhĩ trái/phải, động mạch chủ, động mạch phổi, cơ tim của tâm thất trái. [5]

2.1.2 Bộ dữ liệu thô của người Việt Nam.

Hiện tại, nhóm tác giả đang cùng BS. Lê Văn Nghĩa (trưởng khoa tim mạch BV Chợ Rẫy) và BS. Lê Văn Phước (trưởng khoa chuẩn đoán hình ảnh BV Chợ Rẫy) để thực hiện gán nhãn, phân vùng cho bộ ảnh dữ liệu thô của người Việt Nam với mức độ chi tiết hơn (thêm mạch máu, van ba lá, van hai lá,...).

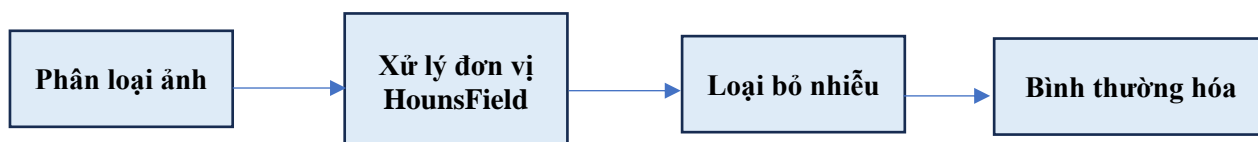


Bộ ảnh thô được bác sĩ Nghĩa đồng thuận giám sát cho công việc phân vùng
Nhận thấy rằng nếu huấn luyện trên tập dữ liệu MM-WHS và đưa đầu vào là bộ ảnh thô như hình trên thì mô hình học sâu sẽ không thực thi được tốt. Vì vậy, nhóm đề xuất các

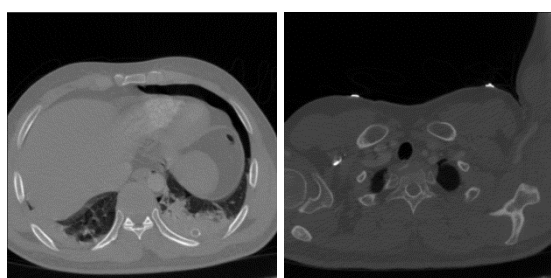
phương pháp bên dưới để đưa tấm ảnh thô về giống với ảnh được huấn luyện nhất. Sau khi bộ dữ liệu trên hoàn thiện, nhóm sẽ tiến hành huấn luyện lại trên bộ dữ liệu mới.

2.2 Tiền xử lý dữ liệu

Hầu hết các phần mềm hiện nay không thể tự động hoàn toàn vì yêu cầu người dùng thực hiện một số thao tác trên ảnh thô như cắt khung ảnh cần thiết, bỏ qua các ảnh không liên quan. Vì vậy, yêu cầu tiên quyết cho một quy trình tự động hóa hoàn toàn là đưa quá trình xử lý được các loại dữ liệu thô trở nên tự động trước khi đưa ảnh vào mô hình học sâu. Để làm được việc này, nhóm thực một quy trình tự động như sau:



2.2.1 Phân loại ảnh



Các bộ ảnh thô chụp từ các máy khác nhau sẽ có số lát cắt khác nhau. Ví dụ bộ ảnh A có 200 lát cắt và đến lát cắt thứ 10 thì mới có hình ảnh tim, trong khi bộ ảnh B có 210 lát cắt và đến lát cắt thứ 15 mới có hình ảnh tim.

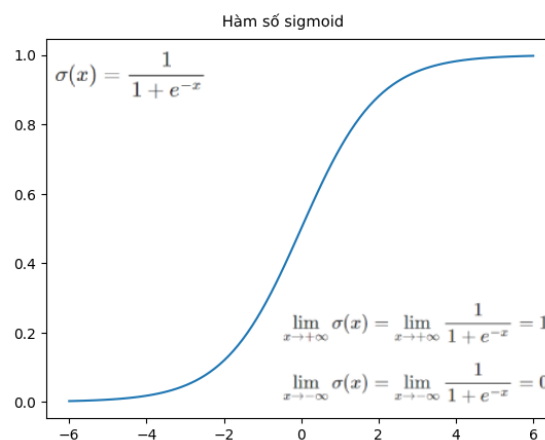
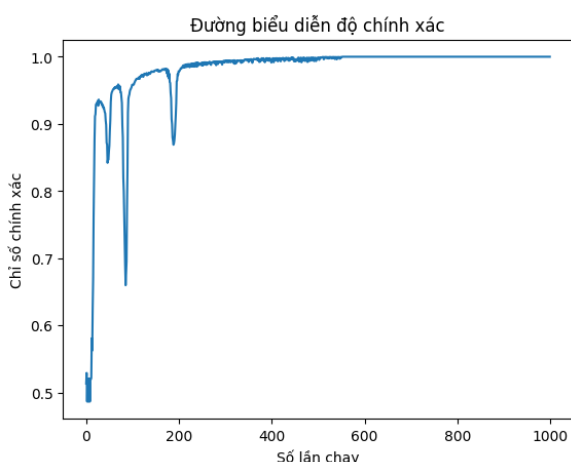
(Ảnh chụp không cho thấy tim nên không cần thiết cho việc phân vùng và tái tạo 3D nên ta cần loại bỏ)

Việc xác định đúng ở lát cắt nào thì xuất hiện hình ảnh tim để ta phân vùng rất quan trọng trong khâu tự động hóa. Vì vậy, nhóm sử dụng thuật toán Logistic Regression cho bài toán phân loại ảnh thành 2 phần này. Ta cần tối ưu các hệ số w trong phương trình tuyến tính (toán 10) sau:

$$y = \sigma(w_1 \times x_1 + w_2 \times x_2 + \dots + w_n \times x_n + b)$$

$$y = \sigma(W \times X + b)$$

Trong đó: X là vector ảnh đầu vào có n phần tử và σ là hàm sigmoid. Quan sát giới hạn lim của đồ thị (toán 11) ta thấy sigmoid để biến đổi kết quả đầu ra vào $(0,1)$. Ảnh không cần thiết nếu $y = 0$ và cần thiết khi $y = 1$. [7]



2.2.2 Xử lý đơn vị Hounsfield

Đây là một đơn vị thể hiện sự tương phản giữa các mô cơ quan trong tấm ảnh cắt lớp. Nhờ đơn vị này, ta có thể thực hiện tiền xử lý dữ liệu và loại bỏ một số thành phần

không liên quan đến tim như phổi, xương. Nhóm áp dụng phương pháp tiêu chuẩn hóa đơn vị Hounsfield. [8]

Mô		Đơn vị Hounsfield
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 bọt	+700
	Xương sụn	+3000

Gọi tập hợp các lát cắt là

$$\mathcal{S} = \{\mathbf{x}_1, \dots, \mathbf{x}_t, \dots, \mathbf{x}_N\} = \{\mathbf{x}^t\}_{t=1}^N$$

Trong đó $\mathbf{x}_t \in \mathbb{R}^{w \times h}$ là một lát cắt có chiều dài w và chiều cao h .

Nhóm dùng công thức tuyến tính [8]:

$$\mathbf{x}_t = \mathbf{x}_t \times \text{slope} + \text{intercept}$$

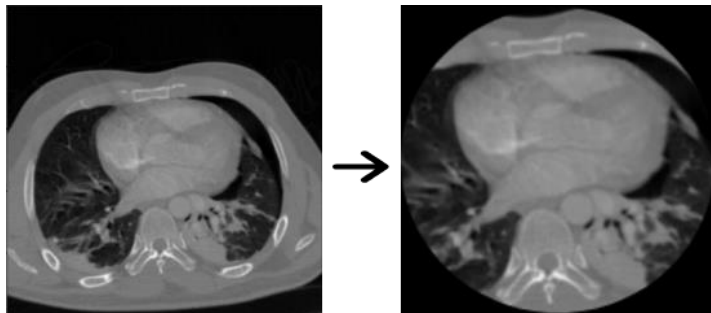
để chuẩn hóa mật độ phóng xạ về một đơn vị Hounsfield thống nhất.

2.2.3 Loại bỏ nhiễu

Vì để giúp cho mô hình học sâu tránh bị “phân tâm” đến những bộ phận khác như phổi, xương,... thì nhóm thực hiện việc bo tròn phần tim như bộ dữ liệu MM-WHS. Nhóm áp dụng phương trình đường tròn ở chương trình toán 10 để thực hiện tác vụ này một cách tự động:

$$(x - a)^2 + (y - b)^2 = R^2$$

Ở phương trình này ta thấy nếu một điểm ảnh A có tọa độ (x,y) nằm trên đường tròn thì khoảng cách Euclide (toán 8) của A đến tâm đường tròn $I(a,b)$ bằng với bán kính R .

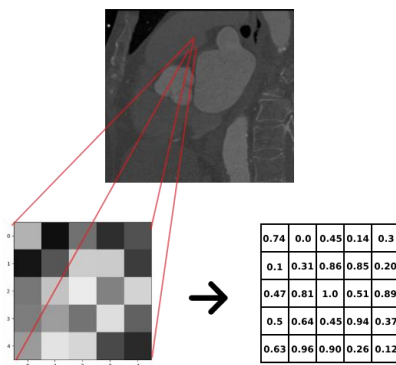


Vậy để khoanh tròn hình tim thì ta chỉ cần lặp qua hết các điểm ảnh trên một lát cắt 2D và xác định khoảng cách D từ điểm đó đến tâm đường tròn I : s

Nếu $D > R \rightarrow$ đổi giá trị thành nền đen, còn $D \leq R \rightarrow$ giữ nguyên.

2.2.4 Bình thường hóa (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 intensity normalization [9]:



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

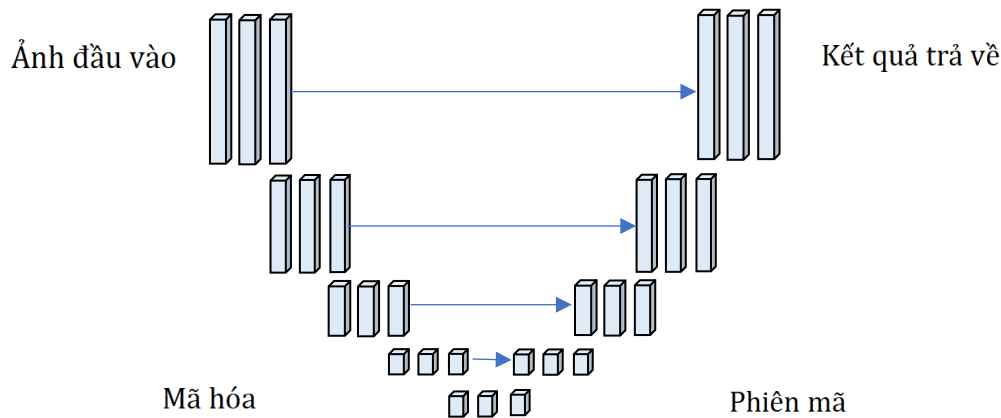
Thay $a = 0$ và $b = 1$ ta được:

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

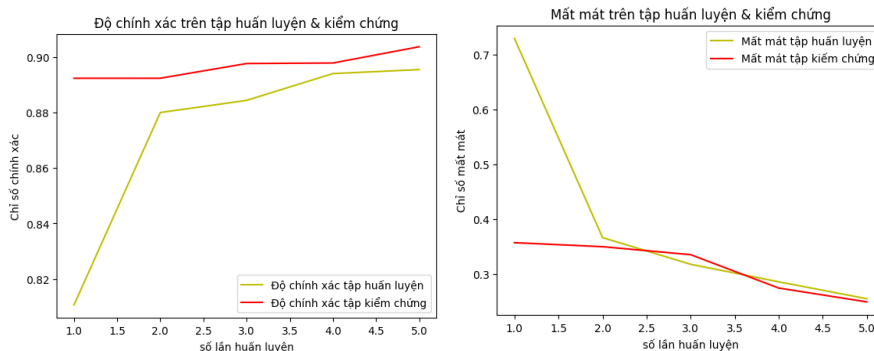
2.3 Huấn luyện

2.3.1 Xây dựng và thử nghiệm với các kiến trúc học sâu

Nhóm bắt đầu quá trình thử nghiệm với mô hình thị giác máy tính truyền thống Unet được chứng minh hiệu năng chính xác cao với kiến trúc gồm 2 phần : Mã hóa (encoder) và Phiên mã (Decoder) tạo thành hình chữ U đặc trưng [10].

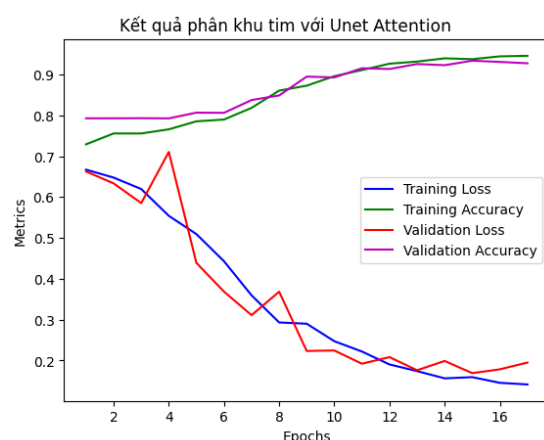


Tuy nhiên kiến trúc trên đều gặp chung một vấn đề: không nhận diện được chính xác các cạnh của các vùng khi hình dạng liên tục thay đổi kích thước, nhầm lẫn các vùng (classes) với nhau.



Kết quả cho thấy độ chính xác chỉ xấp xỉ 90% và độ mất mát xấp xỉ 0.25.

Vì thế Nhóm 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 vùng tốt hơn. [11].



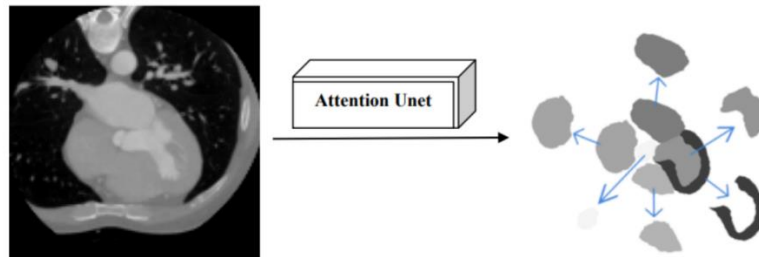
Chúng tôi cho “fit” với 25 epochs và cài đặt early stopping với callback trong thư viện tensorflow. Vì vậy ở epoch 16, khi Validation Loss có dấu hiệu tăng (nghĩa là mô hình đang học thuộc lòng bộ dữ liệu) thì quá trình training kết thúc để đảm bảo tính tổng quát hóa (generalizability) cho mô hình.

Hình 3. Quá trình huấn luyện mô hình học sâu

Thông số tham số của mô hình	
Tổng tham số:	1940936 (7.40 MB)
Tham số huấn luyện:	1940936 (7.40 MB)
Tham số không huấn luyện:	0 (0.00 Byte)

2.3.2 Đánh giá mô hình (Metrics)

Nhóm dùng phương pháp phổ biến để đánh giá mô hình cho bài toán phân vùng nhiều lớp cho dữ liệu y học (medical multiclass segmentation) là DSC (Dice Similarity Coefficient) [12].

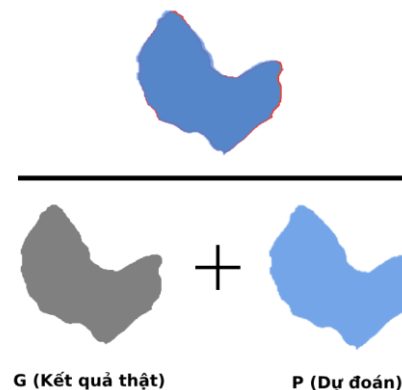


Dice Coefficient: Sau khi có kết quả dự đoán như hình trên thì Chúng tôi tính Dice trên từng class (từng vùng một). Bên dưới là ví dụ tính Dice cho tâm thất trái: so sánh hai hình ảnh trên từng pixel: kết quả thật (G - ground truth) và dự đoán của mô hình (P - prediction).

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

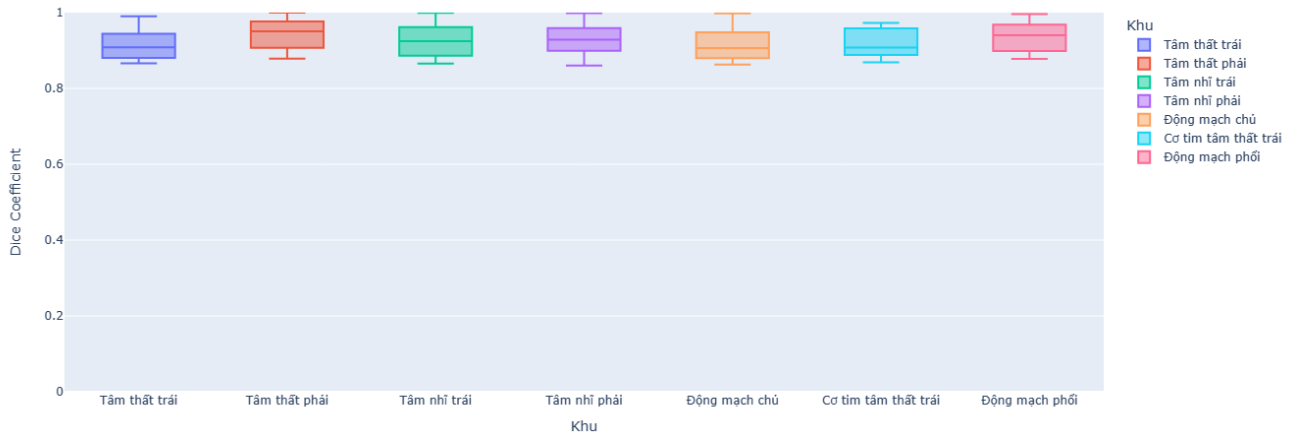
Chia phần giao nhau (intersection)
cho phần tổng (Union)

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 vùng

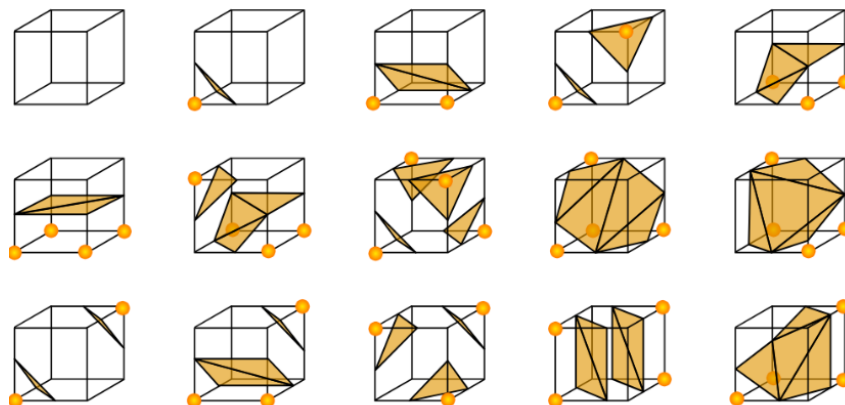
Vùng	Dice Coefficient cho 20 bệnh nhân / 20 slices								lower_fence	upper_fence
	count	mean	std	min	25%	50%	75%	max		
Cơ tim tâm thất trái	20	0.93302	0.042186	0.864674	0.901512	0.933976	0.96989	0.991049	0.798946	1.072456
Tâm nhĩ phải	20	0.929527	0.04093	0.86328	0.892338	0.93238	0.954225	0.999991	0.799508	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

2.3.3 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 [13].

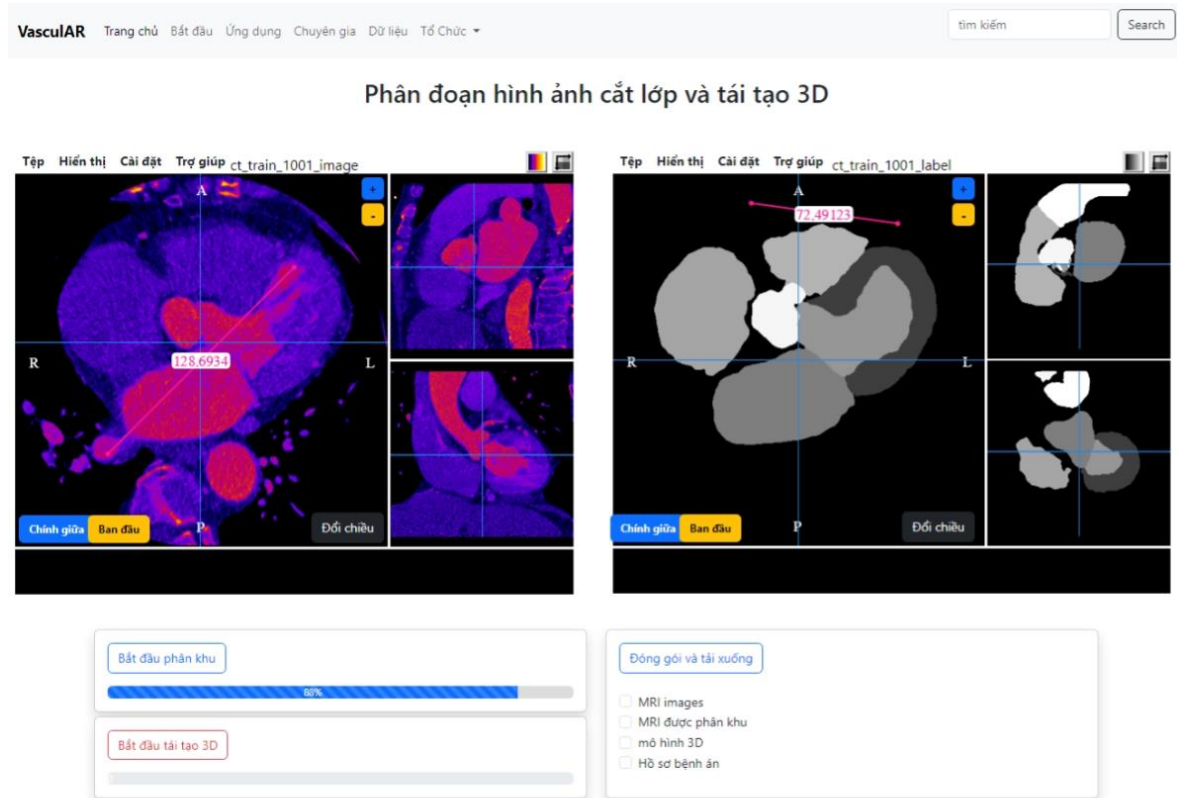
Về mặt toán học, thuật toán sử dụng “Chia để trị” (Divide and Conquer) bằng cách lặp qua, tìm kiếm các khu vực vượt qua một mức độ quan tâm nhất định. (level of interest) Nếu các vùng như vậy được tìm thấy, tam giác được tạo và thêm vào lưới đầu ra. Kết quả cuối cùng là một tập hợp các đỉnh và một tập 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) [14].



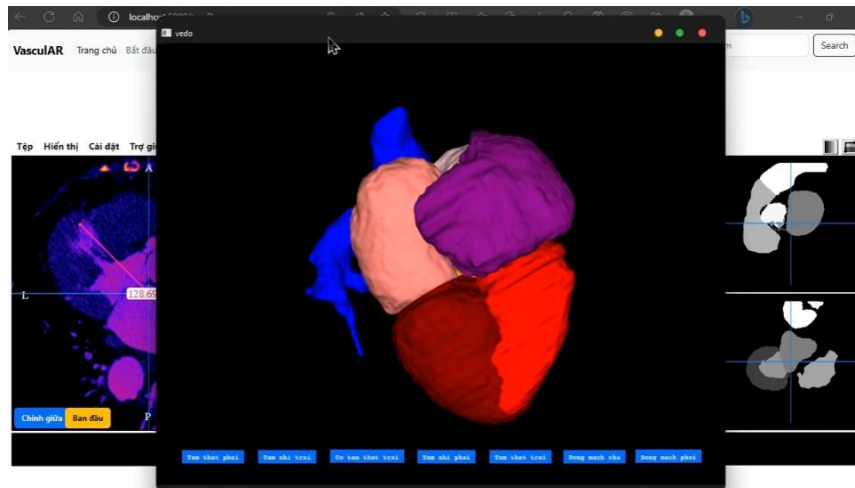
CHƯƠNG 3: KẾT QUẢ NGHIÊN CỨU & MỤC TIÊU TƯƠNG LAI

3.1 Sản phẩm đã hoàn thành (Xem thêm tại [link này](#))

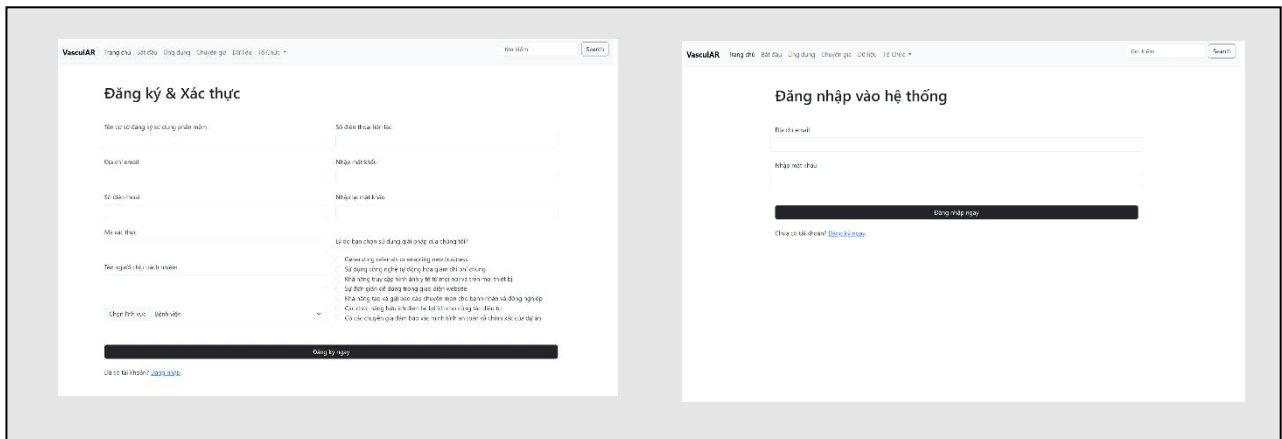
3.1.1 Nền tảng website tiếng Việt giao diện dễ dùng tích hợp các chức năng để hỗ trợ người dùng đọc và phân ảnh chụp cắt lớp dễ dàng với phổ màu, thước đo, hệ tọa độ. Tích hợp được model deep learning Unet lên browser để thực hiện phân vùng.



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



3.1.3 Hệ cơ sở dữ liệu (database) cho phép cơ sở y tế đăng ký tài khoản và lưu trữ các ghi chú của người dùng cho từng ca chụp cắt lớp.



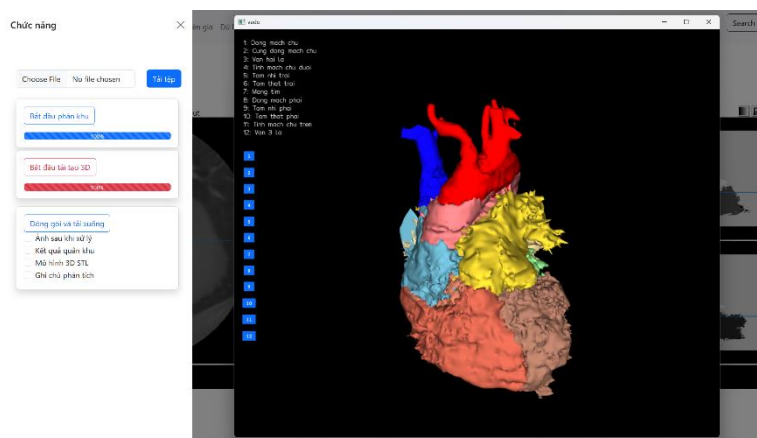
3.2 Công việc đang thực hiện

3.2.1 Bộ dữ liệu mới chi tiết hơn

Với sự cố vấn của các bác sĩ về các bệnh lý liên quan đến nhiều bộ phận khác của trái tim như bệnh đông mạch vành, hở van tim,... nhóm quyết định sẽ cùng các bác sĩ Lê Văn Nghĩa - trưởng khoa tim mạch bệnh viện Chợ Rẫy, bác sĩ Lê Văn Phước - trưởng khoa Chẩn đoán hình ảnh Bệnh viện Chợ Rẫy, và bác sĩ Lê Thị Phương Nga thực hiện việc phân vùng trên bộ dữ liệu thô chụp từ máy chụp của bệnh viện. Hiện tại, bộ dữ liệu của bệnh nhân đầu tiên (được xử lý ẩn danh bảo mật) được thực hiện phân vùng với 12 vùng chi tiết (thay vì chỉ 7 vùng cơ bản như bộ data MM-WHS).

1. Tâm nhĩ trái (Left Atrium)	7. Màng tim (Pericardium)
2. Tâm nhĩ phải (Right Atrium)	8. Cung động mạch chủ (Arch of Aorta)
3. Tâm thất trái (Left Ventricle)	9. Động mạch chủ (Aorta)
4. Tâm thất phải (Right Ventricle)	10. Tĩnh mạch chủ trên (Superior Vena Cava)
5. Van 2 lá (Mitral Valve)	11. Tĩnh mạch chủ dưới (Inferior Vena Cava)
6. Van 3 lá (Tricuspid Valve)	12. Động mạch phổi (Pulmonary artery)

Hình ảnh mẫu tái tạo cấu trúc tim 3D với 12 vùng từ bộ dữ liệu của bệnh nhân thứ nhất. (Xem thêm tại [link này](#))



3.3 Mục tiêu tương lai

Nhóm đang khảo sát về tính khả quan về việc chế tạo thiết bị trình chiếu mô hình 3D trong không gian thực tế (Hologram) có thể tương tác bằng chuột hoặc kết nối găng tay cảm ứng để giảng viên có thể vừa truyền đạt vừa thực hành trong không gian thực cho sinh viên xem, nâng cao chất lượng truyền đạt và giảng dạy.

Tài liệu tham khảo

- [1] Xiaohai Zhuang et al., "Evaluation of Algorithms for Multi-Modality Whole Heart Segmentation: An Open-Access Grand Challenge," *Medical Image Analysis* 58 (2020): 101537
- [2] Dân Trí, 7/10/2023, Mỗi năm, Việt Nam có 200.000 người tử vong vì bệnh tim mạch
- [3] Tuấn, N. Q. (n.d.). Phẫu thuật tim ở Việt Nam đang bước lên đẳng cấp mới. BV Tim Hà Nội [Heart surgery in Vietnam is stepping up to a new level]. BV Tim Hà Nội
- [4] Dũng, N. A. (n.d.). Những điều cần biết về phẫu thuật tim [What you need to know about heart surgery]. Trung tâm Tim mạch, Bệnh viện Đa khoa Tâm Anh, TP.HCM.
- [5] X. Zhuang and J. Shen, "Multi-scale patch and multi-modality atlases for whole heart segmentation of MRI," *Medical Image Analysis*, vol. 31, pp. 77-87, 2016.
- [6] Vindir, Đọc phim cùng người máy, ngày 6 tháng 9 năm 2023 <https://vindr.ai/doc-phim-cung-nguoi-may.html>
- [7] Vũ Hữu Tiệp, Logistic Regression, Machine Learning cơ bản, ngày 21 tháng 1 năm 2017
- [8] P. Haigron et al., "CT-scan images preprocessing and segmentation to improve bioprosthesis leaflets morphological analysis".
- [9] "Intensity Normalization—A Critical Pre-processing Step for Efficient Brain Tumor Segmentation in MR Images" by S. Poornachandra, C. Naveena & Manjunath Aradhya.
- [10] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Springer LNCS Vol.9351: pp. 234–241, 2015.
- [11] "Attention U-Net: Learning Where to Look for the Pancreas," arXiv:1804.03999v3 [cs.CV], May 20, 2018.
- [12] "Continuous Dice Coefficient: a Method for Evaluating Probabilistic Segmentations," arXiv:1906.11031.
- [13] W.E Lorensen and H.E Cline, "Marching cubes: A high resolution 3D surface construction algorithm," *ACM SIGGRAPH Computer Graphics*, vol.21(4), pp:163–169, August 1987.
- [14] "System and method for the display of surface structures contained within the interior region of a solid body", Patent filed on June 5th, 1985.