

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC BÁCH KHOA
KHOA KHOA HỌC VÀ KỸ THUẬT MÁY TÍNH



Xử lý ảnh số và thị giác máy tính (CO3057)

Báo cáo topic mở rộng

Medical Image Segmentation

GVHD: Võ Thanh Hùng

Sinh viên: Nguyễn Minh Quang - 2153724
Nguyễn Tường Khang - 2152112
Trần Hữu Vinh - 2153976

Thành phố Hồ Chí Minh, Tháng 11/2024

Mục lục

1	Giới thiệu	2
1.1	Giới thiệu đề tài	2
1.2	Mục đích	2
2	Cơ sở lý thuyết	2
2.1	Phân đoạn hình ảnh (image segmentation)	2
2.2	Những thách thức trong nhiệm vụ phân đoạn (segmentation)	2
2.3	Quá trình phân đoạn	3
3	Các công trình liên quan	3
4	Phương pháp thực hiện	4
4.1	Dataset và tiền xử lý	4
4.2	Kiến trúc của mô hình	5
4.3	Up-convolution và kênh	6
4.4	Skip Connection	7
4.5	Hàm mất mát	7
5	Đánh giá	8
6	Kết luận	9
	Tài liệu tham khảo	10

1 Giới thiệu

1.1 Giới thiệu đề tài

Phân đoạn hình ảnh y tế là một trong những ứng dụng quan trọng trong lĩnh vực chẩn đoán y tế và nghiên cứu thông qua khả năng phân tích chính xác cấu trúc và vị điểm trong hình ảnh. So với phương pháp phân loại hình ảnh truyền thống đánh giá ý nghĩa của một ảnh toàn vẹn chỉ bằng một nhãn, thì các yêu cầu trong y tế hình này cần phải chỉ rõ từng nhãn của từng pixel của ảnh. Trong đó, U-Net [4], cấu trúc mạng tích chập (convolutional) được phát minh bởi Ronneberger, Fischer và Brox, có khả năng dẫn nhãn đến mức độ chuẩn xác từng pixel này bằng cách sử dụng mạng tích chập đầy đủ (fully convolutional network) mà có thể vượt qua thử thách dẫn nhãn trên tập dữ liệu hạn hẹp thông qua việc tự mở rộng dữ liệu và có cấu trúc thiết kế đặc biệt.

1.2 Mục đích

Cho đến nay, ung thư đã luôn là lý do hàng đầu trong số tử vong trên thế giới với số ca mới và ca tử vong tăng liên tục hằng năm. Theo như số liệu 2022 từ GLOBOCAN [2], được xác nhận bởi Tổ chức Y tế Thế giới (WHO), một trong sáu ca tử vong trên thế giới sẽ là do ung thư (16.8%). Trong năm đó, khoảng 20 triệu ca được chẩn đoán là bị ung thư, và 9.7 triệu trong số đó đã qua đời do ung thư.

Vì vậy khả năng chẩn đoán và điều trị sớm là một trong những yếu tố quan trọng trong việc cải thiện tiên lượng của bệnh nhân, đặc biệt đối với các loại ung thư có tốc độ tiến triển nhanh. Trong bối cảnh này, hình ảnh y tế đóng một vai trò then chốt trong quá trình phát hiện và đánh giá ung thư. Các phương thức chuẩn đoán ảnh như PET/CT/MTI đã trở thành công cụ không thể thiếu trong việc xác định vị trí, kích thước và giai đoạn của khối u. Tuy nhiên việc con người phân tích và xử lý hình ảnh y tế bằng mắt một cách chính xác đòi hỏi mức độ chuyên môn cao và tiêu tốn nhiều thời gian. Trước đây, khi mà lượng dữ liệu chưa đủ lớn, các phương pháp machine learning truyền thống đem lại kết quả khá chính xác và hiệu quả. Tuy nhiên, càng về sau có sự bùng nổ về dữ liệu, deep learning đã khẳng định được tầm quan trọng của mình và được đánh giá là hiệu quả tốt hơn. Trong dự án này, chúng tôi đề xuất thực hiện phân đoạn hình ảnh khối u của bệnh nhân ung thư phổi bằng hình ảnh PET/CT sử dụng mô hình deep learning.

2 Cơ sở lý thuyết

2.1 Phân đoạn hình ảnh (image segmentation)

Phân đoạn (segmentation) trong xử lý ảnh là quá trình phân chia hình ảnh thành các vùng hoặc đối tượng có ý nghĩa để hỗ trợ phân tích và hiểu rõ hơn về nội dung của hình ảnh đó. Trong ngữ cảnh này, phân đoạn sẽ cố gắng xác định và phân biệt các cấu trúc cụ thể, chẳng hạn như tế bào, mô hoặc các vùng tổn thương trong cơ thể, để giúp hỗ trợ chẩn đoán và nghiên cứu y khoa. Cụ thể hơn trong bài báo cáo này, ta cần phân đoạn chuẩn xác đến từng pixel (pixel-wise segmentation), hay nói cách khác, mỗi điểm ảnh trong hình ảnh được gán một nhãn lớp, chẳng hạn như "tế bào" hoặc "nền", nhằm tạo ra một bản đồ phân đoạn rõ ràng cho hình ảnh.

2.2 Những thách thức trong nhiệm vụ phân đoạn (segmentation)

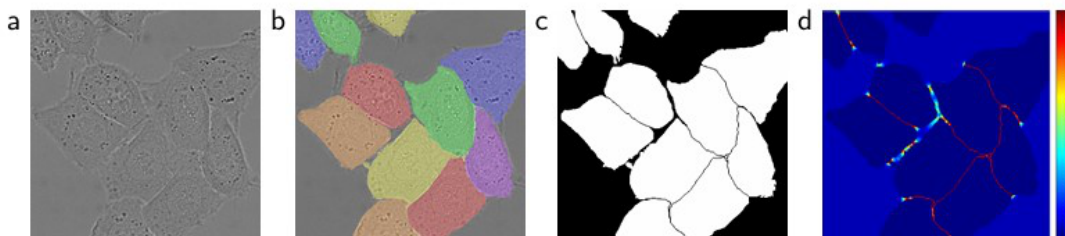
Trước đây, khi lượng dữ liệu chưa đủ lớn thì các phương pháp Machine Learning truyền thống cho kết quả khá chính xác và hiệu quả. Tuy nhiên, theo thời gian, dữ liệu bùng nổ, yêu cầu ngày càng cao hơn

(từ việc xác định đối tượng chỉ bằng “hộp giới hạn” đến cần đường viền tốt hơn để bao phủ đối tượng), mặc dù các phương pháp truyền thống có thể được sử dụng để phân tích dữ liệu, hiệu quả giảm dần, không theo kịp tốc độ tăng trưởng của dữ liệu. Nhiệm vụ phân đoạn hình ảnh phải đối mặt với một số thách thức đáng kể.

- **Độ chính xác:** Một số phương pháp truyền thống gặp khó khăn trong việc phân đoạn hình ảnh một cách chính xác, đặc biệt là những phương pháp có kết cấu, cấu trúc phức tạp hoặc nhiễu khác nhau. Các kỹ thuật như xác định ngưỡng, phân đoạn dựa trên vùng và phát hiện cạnh đã được đề xuất; tuy nhiên, những phương pháp này không mang lại độ chi tiết và độ chính xác cần thiết cho các nhiệm vụ phức tạp như ảnh chụp y tế.

2.3 Quá trình phân đoạn

Một số mô hình học sâu, chẳng hạn như U-Net, Mask R-CNN và Fully Convolutional Network (FCN), đã được sử dụng để phân đoạn hình ảnh. Mặc dù các mô hình này có kiến trúc khác nhau nhưng chúng thường tuân theo một quy trình tương tự để phân đoạn hình ảnh.



Hình 1: Quá trình phân đoạn

- **Chuẩn bị dữ liệu:** Việc có một lượng lớn dữ liệu chú thích là điều cần thiết để đào tạo mô hình deep learning để phân đoạn hình ảnh. Quá trình này liên quan đến việc dán nhãn hình ảnh ở cấp độ pixel và gán nhãn lớp cho từng pixel.
- **Tiền xử lý:** Bước này bao gồm việc thay đổi kích thước hình ảnh và sử dụng các phương pháp tăng cường dữ liệu (data augmentation) để nâng cao độ tin cậy của mô hình. Ngoài ra, nó còn bao gồm việc chuẩn hóa hình ảnh để nâng cao hiệu quả học.
- **Huấn luyện mô hình:** Huấn luyện mô hình trên tập dữ liệu có chú thích, cho phép mô hình học cách phân loại từng pixel của hình ảnh đầu vào vào lớp thích hợp.
- **Hậu xử lý:** Sau khi mô hình dự đoán các lớp pixel, triển khai các bước hậu xử lý, chẳng hạn như loại bỏ các cụm pixel nhỏ bị phân loại sai hoặc sử dụng Conditional Random Fields (CRF) để nâng cao kết quả.
- **Trực quan hóa:** Để trực quan hóa phân đoạn trong bước hậu xử lý, bạn chỉ định một màu cụ thể cho từng giá trị lớp duy nhất. Ánh xạ màu này là tùy ý và được chọn để tối đa hóa độ tương phản giữa các lớp khác nhau, tạo điều kiện phân biệt trực quan dễ dàng hơn.

3 Các công trình liên quan

Phân đoạn khối u sử dụng hình ảnh y tế là rất quan trọng trong thực hành lâm sàng để giảm tỷ lệ tử vong do ung thư. Nhiệm vụ phân đoạn hình ảnh y tế thường sử dụng mạng lưới thần kinh tích chập và một số kiến trúc được sử dụng rộng rãi đã được áp dụng cho vấn đề này.

Một kiến trúc như vậy là U-Net, một mô hình encoder-decoder được phát triển ban đầu để phân đoạn hình ảnh y tế. U-Net thể hiện ưu điểm là tương đối đơn giản và hiệu quả trong khi vẫn đạt được hiệu suất tốt trong các nhiệm vụ phân đoạn hình ảnh y tế khác nhau. Tuy nhiên, nó có thể gặp khó khăn với các hình ảnh đầu vào phức tạp hơn hoặc đa dạng hơn và các phương pháp thay thế có thể phù hợp hơn trong những trường hợp này.

PraNet [3] là một kiến trúc CNN được thiết kế đặc biệt cho việc phân đoạn u tự động trong các hình ảnh y tế. Nó sử dụng một bộ giải mã một phần song song để trích xuất các đặc trưng cấp cao từ hình ảnh và tạo ra một bản đồ toàn cầu làm hướng dẫn ban đầu cho các bước xử lý tiếp theo. Ngoài ra, nó còn sử dụng một mô-đun attention ngược để khai thác các tín hiệu biên, giúp xác lập mối quan hệ giữa các vùng khác nhau trong hình ảnh và biên giới của chúng. PraNet cũng tích hợp cơ chế hợp tác hồi tiếp để điều chỉnh các dự đoán sai lệch và cải thiện độ chính xác của việc phân đoạn. Kết quả đánh giá cho thấy PraNet cải thiện đáng kể độ chính xác phân đoạn và có ưu thế về hiệu quả xử lý thời gian thực, đạt tốc độ khoảng 50 khung hình mỗi giây.

DeepLabV3+ [6] là một phiên bản mở rộng của kiến trúc DeepLabV3 cho phân đoạn hình ảnh ngữ nghĩa. Nó sử dụng các phép biến đổi atrous convolution, cho phép mở rộng tầm nhìn và trích xuất các đặc trưng ở nhiều tỷ lệ khác nhau, từ đó cải thiện khả năng nắm bắt các phụ thuộc ngữ cảnh xa. Cách tiếp cận này giúp phân đoạn chính xác hơn các đối tượng có hình dạng phức tạp hoặc biến thiên quy mô lớn, nhưng đồng thời yêu cầu tính toán nhiều hơn và có thể làm chậm quá trình huấn luyện và suy luận.

MSRF-Net [5] là một kiến trúc CNN được thiết kế đặc biệt cho phân đoạn hình ảnh y tế. Nó sử dụng một khối Dual-Scale Dense Fusion (DSDF) độc đáo để trao đổi các đặc trưng đa tỷ lệ với các trường thu nhận khác nhau, giúp bảo tồn độ phân giải và cải thiện dòng thông tin. Mạng con MSRF sau đó sử dụng một chuỗi các khối DSDF này để thực hiện việc kết hợp đa tỷ lệ, cho phép truyền đạt các đặc trưng cấp cao và cấp thấp để phân đoạn chính xác. Tuy nhiên, một hạn chế của phương pháp này là nó có thể không hoạt động tốt trên các hình ảnh có độ tương phản thấp.

Tổng quan, lĩnh vực này đang là một lĩnh vực nghiên cứu sôi động, với nhiều phương pháp khác nhau được đề xuất và đánh giá. Vì vậy, cần có thêm nghiên cứu để xác định thiết kế tối ưu và các chiến lược huấn luyện cho các mô hình. Việc lựa chọn phương pháp cho một ứng dụng cụ thể cần phải cân nhắc kỹ lưỡng sự đánh đổi giữa độ chính xác, hiệu quả tính toán và các chỉ số hiệu suất khác.

4 Phương pháp thực hiện

4.1 Dataset và tiền xử lý

Chúng tôi sử dụng tập dataset đã được sử dụng trong cuộc thi phân đoạn hình ảnh y tế (Medical Segmentation Decathlon) [1]. Tổng cộng, 2.633 hình ảnh ba chiều đã được thu thập trên nhiều giải phẫu được quan tâm, nhiều phương thức và nhiều nguồn (hoặc tổ chức) đại diện cho các ứng dụng lâm sàng trong thế giới thực. Tập dữ liệu bao gồm ảnh CT (Computed Tomography), MRI (Magnetic Resonance Imaging). Tổng quan về tập dữ liệu được cung cấp ở Hình 2.

ID	Target	Modality	Image Series	N	Labels
Task01_BrainTumour	Brain tumours	MRI	FLAIR, T1w, T1gd, T2w	750	Glioma (necrotic/active tumour), edema
Task02_Heart	Cardiac	MRI	3D	30	Left atrium
Task03_Liver	Liver and tumours	CT	Portal venous phase	201	Liver, tumours
Task04_Hippocampus	Hippocampus	MRI		394	Hippocampus head and body
Task05_Prostate	Prostate	MRI	T2, ADC	48	Prostate central gland, peripheral zone
Task06_Lung	Lung tumours	CT	Non-contrast	96	Lung tumours
Task07_Pancreas	Pancreas	CT	Portal venous phase	420	Pancreas, pancreatic mass (cyst or tumour)
Task08_HepaticVessel	Liver tumours and vessels	CT	Portal venous phase	443	Liver vessels
Task09_Spleen	Spleen	CT	Portal venous phase	61	Spleen
Task10_Colon	Colon tumours	CT	Non-contrast	190	Colon cancer

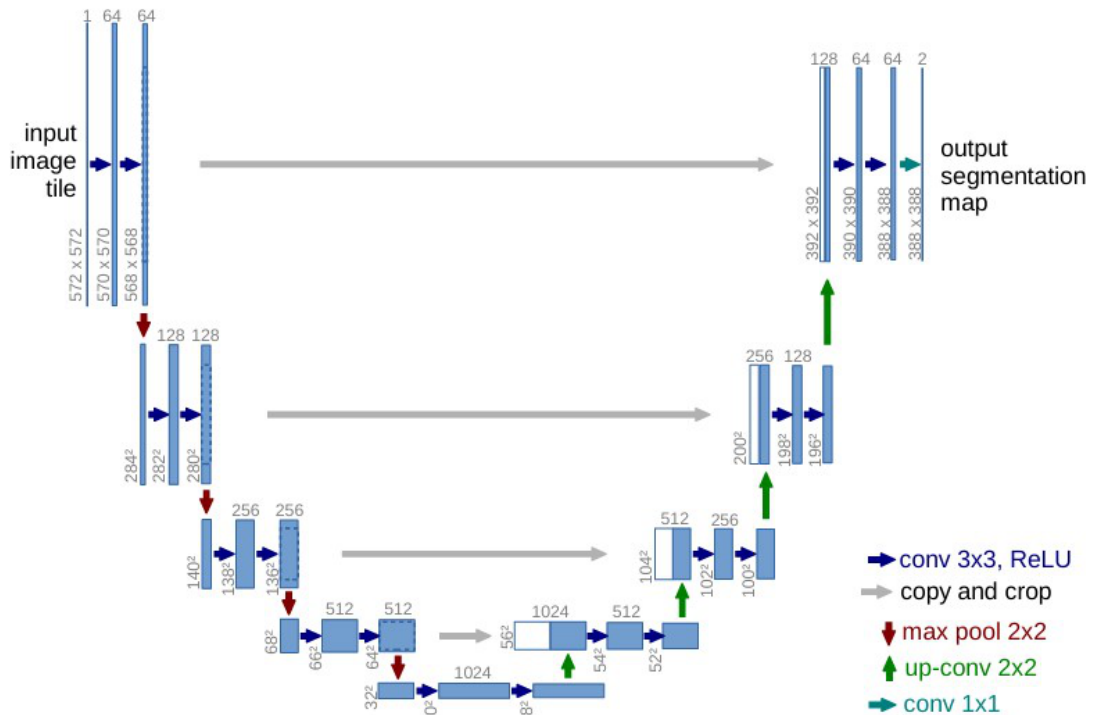
Hình 2: Tổng quan Dataset

Tập dữ liệu có tổng cộng 10 nhiệm vụ phân đoạn. Tất cả hình ảnh đã được định dạng lại từ hình ảnh DICOM tiêu chuẩn sang hình ảnh Neuroimaging Informatics Technology Initiative (NIfTI), một định dạng mở được hỗ trợ bởi National Institutes of Health.

Do trong mỗi file NIfTI số lượng lát cắt của mỗi đối tượng là khác nhau nên cần phải xử lý để các đối tượng có số lượng lát cắt là như nhau. Nhóm sử dụng phần mềm 3D Slicer để tách file NIfTI thành chuỗi các file DICOM, sau đó chia thành các tập nhỏ, tạo thành các đối tượng mới và mỗi đối tượng bao gồm 64 lát cắt. Từ đó chuyển đổi lại thành file NIfTI và xử lý loại bỏ đi các đối tượng không chứa dữ liệu phân đoạn. Ngoài ra các hình ảnh còn được thu nhỏ về kích thước (128, 128, 64) và cường độ được scale về khoảng $[-200, 200]$.

4.2 Kiến trúc của mô hình

Kiến trúc mạng có tên gọi này là do 2 quá trình chính của nó gồm nhánh co (contracting path) và nhánh mở rộng (expanding path) tạo thành hình chữ U đối xứng:



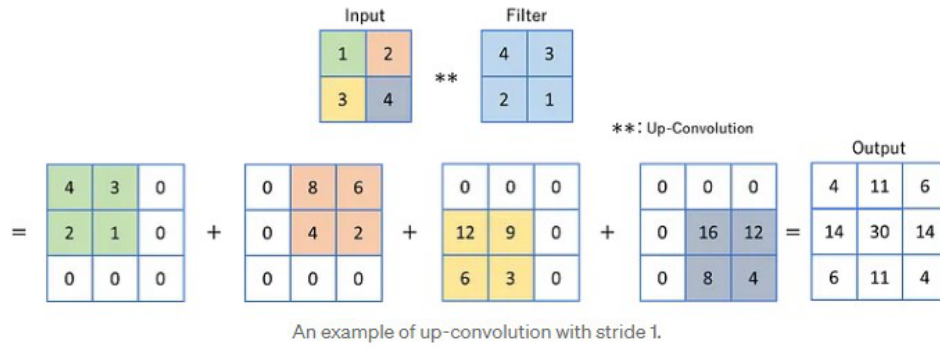
Hình 3: Kiến trúc U-Net

- Nhánh co (Contracting path): là phần quá trình bên trái của mạng U-Net. Giống với mạng neuron tích chập bình thường (Convolutional Neural Network), nhánh này áp dụng các phép tích chập 3x3 liên tục, kèm với các hàm ReLU và phép tính gộp max-pooling 2x2 với bước trượt (stride) là 2. Điều này làm cho ảnh ngày càng bị ép, giảm về chiều không gian nhưng lại tăng về số lượng kênh đặc trưng. Nhiệm vụ chính của nhánh trái là để thu thập được ngữ cảnh của ảnh, và tổng hợp các đặc trưng quan trọng của các vùng lớn bên trong ảnh.
- Nhánh mở rộng (Expansive path): là phần quá trình bên phải của mạng U-Net. Đây là giai đoạn mở rộng (upsample) bản đồ đặc trưng (feature map) để phục hồi các thông tin xung quanh. Mỗi bước là phép upsample kèm với phép tích chập 2x2 (hay "up-convolution") giảm số kênh đặc trưng đi một nửa. Ngoài ra, các bản đồ đặc trưng từ nhánh co còn được ghép với các bản đồ đặc trưng được upsample bên này để tăng khả năng định vị. Nhiệm vụ chính của nhánh phải là để giúp cho mô hình liên kết các chi tiết xung quanh mức thấp với các thông tin ngữ cảnh cấp cao một cách hiệu quả.

Mô hình còn tránh việc sử dụng các lớp kết nối đầy đủ (fully-connected layer), giúp cho nó xử lý tốt trên ảnh của bất kỳ kích thước nào. Hơn thế nữa, đầu ra của mô hình là bản đồ phân đoạn (segmentation map) có đầy đủ nhãn trên từng điểm, phù hợp với yêu cầu y tế.

4.3 Up-convolution và kênh

Tích chập lên (Up-convolution), thường được gọi là giải mã (deconvolution) hoặc chuyển vị tích chập (transpose convolution), là một kỹ thuật được sử dụng để lấy mẫu hình ảnh và khôi phục thông tin không gian.



Hình 4: Up-convolution

Phương pháp hiệu quả nhất để thực hiện tích chập lên bao gồm việc mở rộng và sao chép từng phần tử từ bản đồ tính năng đầu vào để phù hợp với kích thước của bộ lọc. Quá trình này lấy mẫu đầu vào, sau đó bộ lọc được áp dụng cho từng vùng mở rộng này.

Thay đổi số lượng kênh: Trong toàn bộ U-Net, số lượng feature channels liên tục thay đổi. Các phép toán tích chập không ảnh hưởng trực tiếp đến số kênh hiện có. Trên thực tế, nó được xác định bởi số bộ lọc được sử dụng trong lớp tích chập. Nếu 64 bộ lọc được áp dụng trên đầu vào, với mỗi bộ lọc cố gắng trích xuất một tính năng khác nhau, 64 bản đồ tính năng cũng sẽ được tạo ra.

4.4 Skip Connection

Skip connection đóng một vai trò quan trọng trong hiệu quả của U-Net. Bằng cách hợp nhất trực tiếp các bản đồ đặc trưng từ contracting path với các bản đồ từ expanding path, U-Net tích hợp chi tiết cấp thấp với thông tin ngữ cảnh cấp cao trên toàn mạng.

Khôi phục hệ thống phân cấp không gian: Các kết nối này cho phép U-Net kết hợp các tính năng có độ phân giải cao từ contracting path với các đầu ra được lấy mẫu lại từ expanding path. Sự tích hợp này hỗ trợ khôi phục hệ thống phân cấp không gian có thể bị mất trong quá trình hoạt động gộp trong giai đoạn contracting.

4.5 Hàm mất mát

Mô hình sử dụng cross-entropy loss và dice loss làm hàm mất cuối cùng vì nó đã được chứng minh là mạnh mẽ qua các nhiệm vụ phân đoạn hình ảnh y tế khác nhau. Cụ thể, đặt S, G lần lượt biểu thị kết quả phân đoạn và ground truth. Đặt s_i, g_i lần lượt là phân đoạn được dự đoán và ground truth của voxel i . N là số lượng điểm ảnh trong ảnh I . Hàm Binary cross-entropy loss được định nghĩa bởi

$$L_{BCE} = -\frac{1}{N} \sum_{i=1}^N [g_i \log s_i + (1 - g_i) \log(1 - s_i)],$$

và dice loss được định nghĩa bởi

$$L_{Dice} = 1 - \frac{2 \sum_{i=1}^N g_i s_i}{\sum_{i=1}^N (g_i)^2 + \sum_{i=1}^N (s_i)^2}.$$

Hàm mất mát cuối cùng được định nghĩa bởi

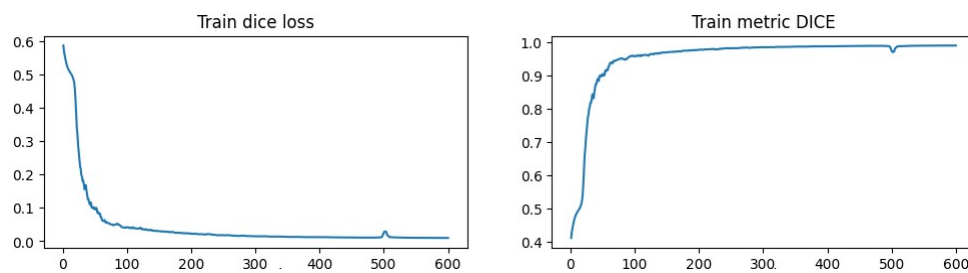
$$L = L_{BCE} + L_{Dice}.$$

5 Đánh giá

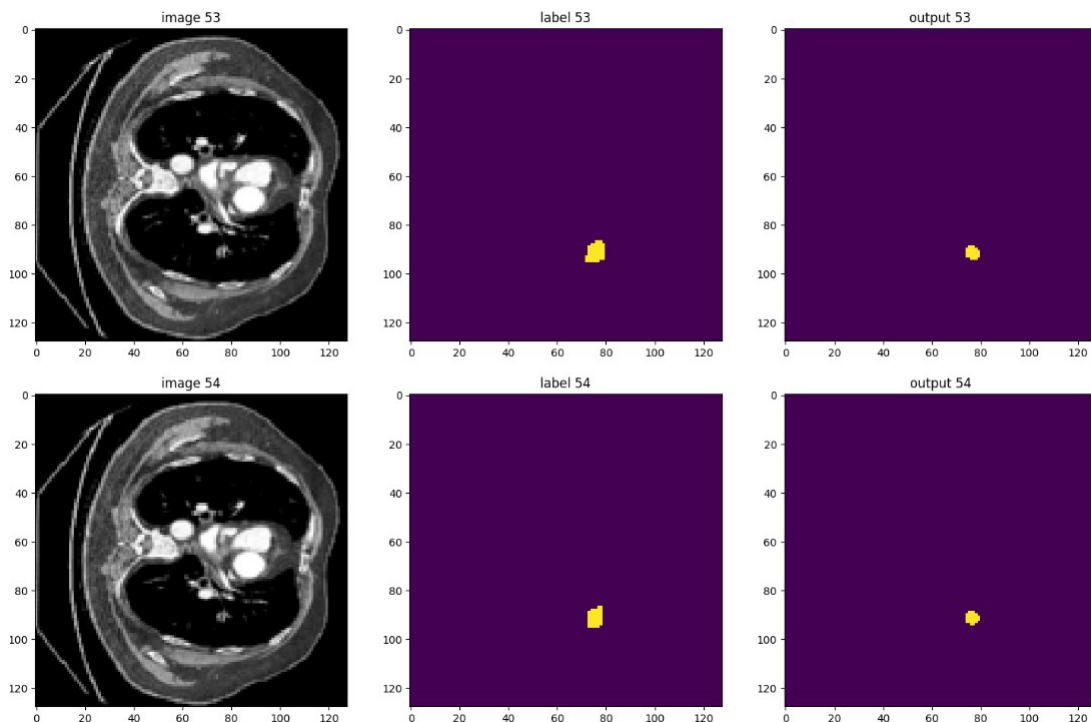
Chúng tôi thực hiện huấn luyện mô hình dựa trên tập dữ liệu của Medical Segmentation Decathlon về Task06_Lung. Tập dữ liệu có tổng cộng 70 đối tượng, chúng tôi chia ra 63 đối tượng để huấn luyện và 7 đối tượng để kiểm tra. Quá trình huấn luyện sử dụng learning rate là $1e^{-4}$, số lượng epoch là 600.

```
Epoch_loss: 0.0098
Epoch_metric: 0.9902
test_loss_epoch: 0.4901
test_dice_epoch: 0.5099
current epoch: 600 current mean dice: 0.5000
best mean dice: 0.5587 at epoch: 35
train completed, best_metric: 0.5587 at epoch: 35
```

Hình 5: Thông số huấn luyện



Hình 6: Train loss



Hình 7: Kết quả phân đoạn

6 Kết luận

Dự án "Medical Image Segmentation" đã sử dụng thành công kiến trúc U-Net để phân đoạn ảnh y tế. Qua quá trình thực hiện và phân tích, mô hình U-Net cho thấy khả năng ổn định trong việc xác định và phân đoạn các vùng mục tiêu trên ảnh y tế.

Kết quả từ mô hình cho thấy rằng U-Net có thể đạt được hiệu suất tốt trên các tập dữ liệu phân đoạn ảnh y tế, giúp hỗ trợ các chuyên gia y tế trong chẩn đoán và điều trị. Tuy nhiên, để cải thiện hơn nữa hiệu suất của mô hình, có thể nghiên cứu và thử nghiệm thêm các biến thể của U-Net cũng như các kỹ thuật tăng cường dữ liệu (data augmentation) và xử lý trước dữ liệu (preprocessing).

Nhìn chung, dự án đã đạt được mục tiêu đề ra trong việc ứng dụng kiến trúc U-Net vào bài toán phân đoạn ảnh y tế, đồng thời mở ra các hướng nghiên cứu mới nhằm tối ưu hóa mô hình và mở rộng phạm vi ứng dụng của nó trong lĩnh vực chăm sóc sức khỏe.

Tài liệu tham khảo

- [1] Michela Antonelli, Annika Reinke, Spyridon Bakas, Keyvan Farahani, Annette Kopp-Schneider, Bennett A. Landman, Geert Litjens, Bjoern Menze, Olaf Ronneberger, Ronald M. Summers, Bram van Ginneken, Michel Bilello, Patrick Bilic, Patrick F. Christ, Richard K. G. Do, Marc J. Gollub, Stephan H. Heckers, Henkjan Huisman, William R. Jarnagin, Maureen K. McHugo, Sandy Napel, Jennifer S. Golia Pernicka, Kawal Rhode, Catalina Tobon-Gomez, Eugene Vorontsov, James A. Meakin, Sebastien Ourselin, Manuel Wiesenfarth, Pablo Arbeláez, Byeonguk Bae, Sihong Chen, Laura Daza, Jianjiang Feng, Baochun He, Fabian Isensee, Yuanfeng Ji, Fucang Jia, Ildoo Kim, Klaus Maier-Hein, Dorit Merhof, Akshay Pai, Beomhee Park, Mathias Perslev, Ramin Rezaifar, Oliver Rippel, Ignacio Sarasua, Wei Shen, Jaemin Son, Christian Wachinger, Liansheng Wang, Yan Wang, Yingda Xia, Daguang Xu, Zhanwei Xu, Yefeng Zheng, Amber L. Simpson, Lena Maier-Hein, and M. Jorge Cardoso. The medical segmentation decathlon. *Nature Communications*, 13(1), July 2022.
- [2] Freddie Bray, Mathieu Laversanne, Hyuna Sung, Jacques Ferlay, Rebecca L Siegel, Isabelle Soerjomataram, and Ahmedin Jemal. Global cancer statistics 2022: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 74(3):229–263, 2024.
- [3] Deng-Ping Fan, Ge-Peng Ji, Tao Zhou, Geng Chen, Huazhu Fu, Jianbing Shen, and Ling Shao. Prant: Parallel reverse attention network for polyp segmentation. In *International conference on medical image computing and computer-assisted intervention*, pages 263–273. Springer, 2020.
- [4] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pages 234–241. Springer, 2015.
- [5] Abhishek Srivastava, Debesh Jha, Sukalpa Chanda, Umapada Pal, Håvard D Johansen, Dag Johansen, Michael A Riegler, Sharib Ali, and Pål Halvorsen. Msrf-net: a multi-scale residual fusion network for biomedical image segmentation. *IEEE Journal of Biomedical and Health Informatics*, 26(5):2252–2263, 2021.
- [6] Salih Can Yurtkulu, Yusuf Hüseyin Şahin, and Gozde Unal. Semantic segmentation with extended deeplabv3 architecture. In *2019 27th Signal Processing and Communications Applications Conference (SIU)*, pages 1–4. IEEE, 2019.