

ỨNG DỤNG CNN TRONG VIỆC XÁC ĐỊNH MÔ TẾ BÀO UNG THƯ GAN TỪ ẢNH CHỤP CT

APPLYING CNN IN DETERMINING CANCER CELLS FROM CT SCANS

SVTH: Châu Trường Long, Võ Minh Trí

Lớp 18TCLC_DT2, Khoa Công Nghệ Thông Tin, Trường Đại Học Bách Khoa Đà Nẵng; Email: truonglongchau@gmail.com, trithemoonlight3@gmail.com

GVHD: TS. Phạm Công Thắng

Khoa Công Nghệ Thông Tin, Trường Đại Học Bách Khoa Đà Nẵng; Email: pacotha@gmail.com, pcthang@dut.udn.vn

Tóm tắt – Ung thư gan là một trong những căn bệnh nguy hiểm và là nguyên nhân hàng đầu gây tử vong. Việc ứng dụng khoa học công nghệ trong việc chẩn đoán và xác định các mô tế bào ung thư của lá gan đóng vai trò rất quan trọng. Điều này góp phần hỗ trợ các bác sĩ trong việc lên kế hoạch điều trị và chữa trị cho bệnh nhân. Trong bài báo này, chúng tôi nghiên cứu việc ứng dụng của mạng nơ-ron tích chập (convolutional neural network – CNN) trong việc xác định các tế bào ung thư của lá gan người. Việc huấn luyện (training) được thực hiện trên tập dữ liệu các bức ảnh CT 3d của đoạn cơ thể chứa lá gan. Sau đó, chúng tôi thực hiện xuất kết quả thành mô hình huấn luyện. Dựa vào các model này để thực hiện thực nghiệm với các mẫu test khác nhau.

Từ khóa – mạng nơ-ron tích chập; phân đoạn ảnh; ung thư gan; ảnh CT.

Abstract - Liver cancer is one of the most dangerous diseases and is one of causes leading of death. The application of science and technology in the diagnosis and identification of cancerous tissues of the liver plays a very important role. This assists the doctor in planning and treating the patient. In this paper, we study the application of convolutional neural networks (CNN) in the determination of cancerous tissues of the human liver. The training are performed on a 3D CT image dataset of the body segment containing the liver. We then run the results into a train model, on which experiments are performed with different test samples.

Key words – convolutional neural network (CNN); segmentation; liver cancer ; CT image.

1. Đặt vấn đề

Ung thư gan là một trong những căn bệnh ung thư phổ biến nhất trên thế giới nói chung, tại Việt Nam nói riêng và nó gây ra cái chết hàng loạt mỗi năm. Ung thư gan là ung thư xuất phát từ các loại tế bào của gan. Gan có kích thước cỡ trái banh bầu dục nằm ở vùng bụng bên phải, nằm dưới cơ hoành và phía trên của dạ dày (hình 1). Có nhiều dạng ung thư xuất phát từ gan. Thường gặp nhất là ung thư tế bào gan nguyên phát, xuất phát từ chính tế bào gan. Các loại khác ít gặp hơn như ung thư đường mật trong gan, u nguyên bào gan.

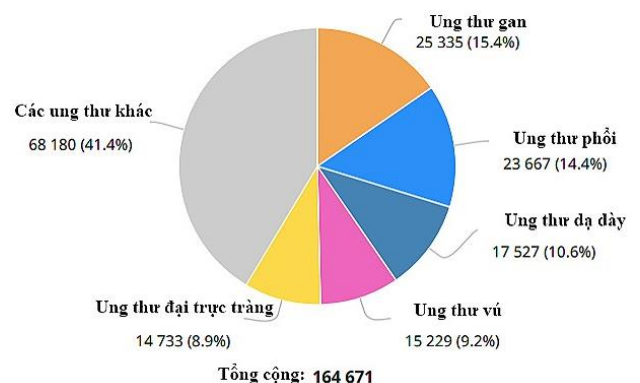


Hình 1: Vị trí của lá gan.

Không phải tất cả các khối u ở gan đều được xem như là ung thư gan. Ung thư xuất phát từ các vị trí khác trong cơ thể, như đại tràng, phổi, vú,... có thể lan đến gan gọi là ung thư di căn gan. Loại ung thư này được gọi tên theo cơ quan mà ung thư xuất phát. Ví dụ, ung thư đại tràng di căn gan được miêu tả là các tổn thương ung thư xuất phát từ đại tràng và lan đến gan. Các ung thư di căn gan thường gặp hơn các ung thư từ các loại tế bào của gan [4,5].

Theo những con số thống kê từ bộ y tế về ung thư gan trong năm 2018 [6,7], người ta đã phát hiện thêm 840.000 ca mắc ung thư gan mới trên toàn thế giới, trong đó số ca

tử vong chiếm khoảng 700.000 ca. Tỷ lệ tử vong do ung thư gan rất cao, đa số là vì bệnh nhân phát hiện bệnh quá muộn, các biện pháp trị liệu không thể tác động hiệu quả. Nguyên nhân của hầu hết các trường hợp mắc ung thư gan là do nhiễm virus viêm gan B (HBV) và virus viêm gan C (HVC), chiếm khoảng 78%. Tỷ lệ ung thư gan cũng có sự thay đổi đáng kể tùy theo giới tính, nam giới mắc ung thư gan cao hơn nữ giới. Riêng Việt Nam có hơn 10.000 trường hợp mới mắc/năm. Tỷ lệ này cao nhất thế giới, có hơn 7,8 triệu bệnh nhân mắc viêm gan B và khoảng 991.000 bệnh nhân mắc viêm gan C. Trong số các trường hợp viêm gan B, có hơn 51.000 trường hợp tiến triển thành xơ gan và hơn 14.000 ca khác tiến triển thành ung thư. Trong số các bệnh nhân viêm gan C, có hơn 13.000 ca chuyển sang xơ gan và 6.000 ca chuyển sang ung thư gan (hình 2).



Hình 2: Tỷ lệ các loại ung thư tại Việt Nam năm 2018 trong đó ung thư gan chiếm 15,4%.

Như vậy, dựa vào tất cả các thông tin ở trên có thể thấy căn bệnh ung thư gan là một căn bệnh rất nguy hiểm.

Nguy cơ chữa trị là rất thấp nếu như không được phát hiện kịp thời. Do đó để góp phần làm giảm thiểu tỷ lệ tử vong cũng như có thể giúp bác sĩ lên kế hoạch chữa trị kịp thời cho bệnh nhân, thì việc giúp các bác sĩ có thể phát hiện được vị trí của các mô tế bào ung thư nằm trong lá gan là rất quan trọng. Hiện nay việc ứng dụng các mô hình học máy trong y học rất được chú trọng đến vì tính hữu dụng và sự chính xác trong việc chẩn đoán mà nó đem lại. Đặc biệt việc vận dụng kiến thức về mô hình mạng nơ-ron tích chập FCN để giải quyết bài toán phân đoạn ảnh của lá gan từ ảnh chụp CT.

2. Cơ sở nghiên cứu

2.1. Các công cụ hỗ trợ cho việc nghiên cứu

- **Python 3:** ngôn ngữ lập trình hỗ trợ cho việc phát triển và training mô hình.
- **Anaconda:** hỗ trợ cho việc cài đặt môi trường ảo python và để cài một số thư viện khó cài như: tensorflow-gpu.
- **Google colab:** là một nền tảng của google để hỗ trợ cho việc huấn luyện(training) model cho mô hình học máy.
- **Theia [3] và ImageJ:** công cụ để xem được hình ảnh CT có định dạng NifTi-1 dưới dạng 3d và 2d.

2.2. Image segmentation

Với deep learning, máy tính ngày càng có khả năng xử lý được những hình ảnh phức tạp ở nhiều tác vụ và mức độ khác nhau. Để hiểu rõ về unet segmentation ta cần hiểu về quá trình xây dựng và phát triển của bài toán xác định các vật thể của một bức ảnh [9].

2.2.1. Image classification

Phân lớp hình ảnh là tác vụ cơ bản nhất của bài toán. Có thể hiểu đơn giản như sau: có một tấm ảnh và nhiều cái hộp, mỗi hộp chứa các bức ảnh cùng loại với nhau. Chúng ta cần đặt bức ảnh đó vào đúng cái hộp có chứa ảnh cùng loại.

Ví dụ như: hình ảnh của con mèo sẽ được bỏ vào cái hộp chứa các ảnh mèo. Hình ảnh của con chó sẽ được bỏ vào hộp chứa các ảnh chó, ...

2.2.2. Image detection

Nâng cao hơn một chút là nếu trong một bức ảnh không đơn giản chỉ chứa một vật thể như con mèo, mà còn chứa nhiều vật thể khác như: chó, gà, vịt, con người,... Từ đó image detection hay object detection ra đời để giải quyết bài toán này. Hiểu đơn giản là detect các vật thể trong ảnh thành các "bounding box" là một dạng hình chữ nhật bên trong là vật thể đó.

Những kiến trúc mạng được thiết kế riêng cho tác vụ này như: RCNN, FAST CNN, FASTER CNN, YOLO,... Nó được ứng dụng rộng rãi vào các ứng dụng như: Nhận diện khuôn mặt để điểm danh, phát hiện biển số xe trong hệ thống quản lý xe, phát hiện các phương tiện vi phạm luật giao thông, theo dõi lưu lượng khách hàng ra vào siêu thị, ...

2.2.3. Image segmentation

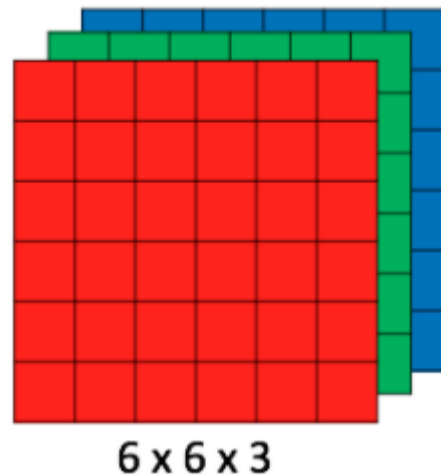
Image segmentation là một nâng cấp của image detection. Chúng không xác định vật thể thành các "bounding box" nữa. Mà chúng cần một lớp đường viền (baseline) tốt hơn để đi bao quanh vật thể. Do đó

nó sẽ không chứa các hình ảnh thừa xung quanh của vật thể như "bounding box". Hay nói cách khác chúng image segmentation có thể xác định từng pixel ảnh thuộc về lớp nào.

2.3. Convolutional neural network (CNN)

Hiện nay trong cộng đồng deep learning đã phát triển rất nhiều mô hình, kiến trúc mạng khác nhau. Phải nói là rất nhiều. Do đó kiến trúc mạng được nghiên cứu cho tác vụ image segmentation cũng không hề nhỏ. Một vài kiến trúc mạng về segmentation như: CNN, FCN, Mask RCNN, ... Tuy nhiên trong nghiên cứu này sẽ chỉ đề cập đến hai kiến trúc mạng CNN và Unet [1].

Trong mạng neural, mô hình mạng neural tích chập (CNN) là một trong những mô hình phân loại hình ảnh được sử dụng rộng rãi. Một trong những ứng dụng phổ biến của nó là nhận diện khuôn mặt. CNN phân loại hình ảnh bằng cách lấy 1 hình ảnh đầu vào, xử lý và phân loại nó theo các hạng mục nhất định (Ví dụ: Chó, Mèo, Hổ, ...). Máy tính coi hình ảnh đầu vào là 1 mảng pixel và nó phụ thuộc vào độ phân giải của hình ảnh. Dựa trên độ phân giải hình ảnh, máy tính sẽ thấy $H \times W \times D$ (H: Chiều cao, W: Chiều rộng, D: Độ dày). Hình 3 biểu diễn mảng ma trận RGB $6 \times 6 \times 3$ [11].

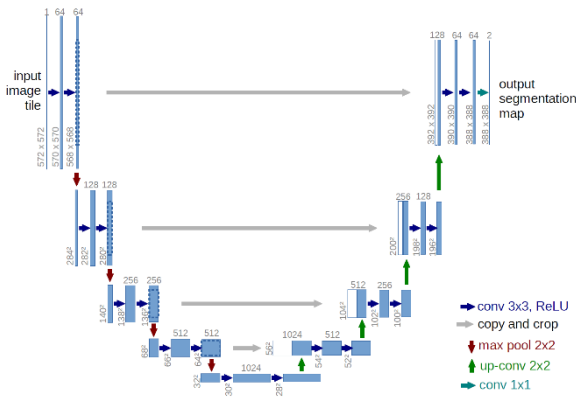


Hình 3: Mảng ma trận RGB $6 \times 6 \times 3$.

Về kỹ thuật, mô hình CNN để training và kiểm tra, mỗi hình ảnh đầu vào sẽ chuyển nó qua 1 loạt các lớp tích chập với các bộ lọc (Kernels), tổng hợp lại các lớp được kết nối đầy đủ (Full Connected) và áp dụng hàm Softmax để phân loại đối tượng có giá trị xác suất giữa 0 và 1. Hình dưới đây là toàn bộ luồng CNN để xử lý hình ảnh đầu vào và phân loại các đối tượng dựa trên giá trị [11].

2.4. Unet

Kiến trúc mạng Unet là một loại của convolutional neural network (CNN), có nghĩa là nó cũng dựa trên kiến trúc mạng CNN cụ thể ở đây là dựa vào kiến trúc FCN(cũng là một loại khác dựa vào CNN có tên đầy đủ là fully convolutional network) được phát triển cho việc nghiên cứu và phân đoạn ảnh trong y sinh (hình 4) [1, 9].



Hình 4: Kiến trúc mạng Unet.

Điểm đặc biệt của Unet so với các kiến trúc khác đó là:

- Mặc dù nó dựa trên kiến trúc mạng FCN nhưng toàn bộ kiến trúc không hề sử dụng một lớp fully connected như các mô hình end-to-end khác. Đó là nhờ nữa “chữ U” thứ hai đảm nhận việc đó, điều này giúp chấp nhận input với kích thước bất kì.
- Unet sử dụng phương pháp padding, giúp kiến trúc có thể phân đoạn hình ảnh hoàn toàn.

3. Kết quả nghiên cứu và khảo sát

3.1. Dataset và tiền xử lý (preprocessing)

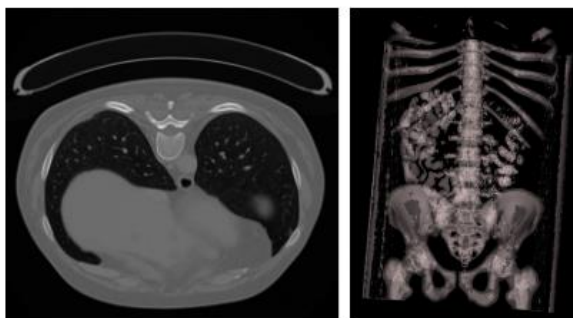
Trước khi bắt đầu quá trình train model dựa vào tập dữ liệu LITS Challenge [2] bao gồm:

- Train model: 131 bức ảnh chụp cắt lớp CT 3d và 131 bức ảnh của lá gan sau khi phân đoạn.
- Test: 70 mẫu test chụp cắt lớp CT 3d.

Ở giai đoạn tiền xử lý các bức ảnh trong tập dữ liệu trên. Tôi đã làm giảm đi cường độ của các bức ảnh và làm sao các bức ảnh chỉ có giá trị thuộc $[-200, 250]$ HU (Hounsfield Scale) (hình 5), mục đích là để loại bỏ các chi tiết thừa không cần thiết trong việc phân đoạn hình ảnh và các mô tế bào của lá gan.

3.2. Kết quả training model

Công việc training model được thực hiện trên google colab với Dataset LITS [2] được lưu trữ trên google driver và được mount đến google colab.



Hình 5: Một dữ liệu training ở dạng 2d và 3d.

Kết quả sau khi traing ta được một file dữ liệu ở định dạng HDF5 (.h5) [12] là file model của sau quá

trình training. Sau này muốn chạy test chỉ cần load file model vào, chứ không cần thực hiện lại quá trình training.

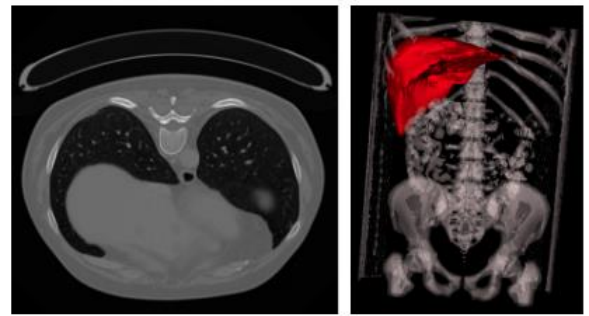
Để training model, hàm chéo trọng số sẽ được xem như là hàm loss để kiểm tra độ mất mát của kết quả

$$L(y, \hat{y}) = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^3 w_i^c y_i^c \log \hat{y}_i^c$$

3.3. Kết quả chạy thử nghiệm test

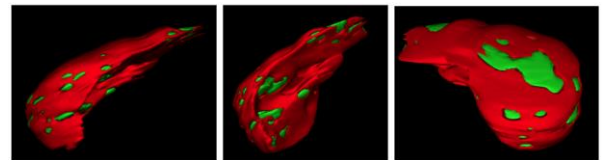
Sau quá trình training, mô hình đã tự học được cách tự phân đoạn lá gan và các mô tế bào ung thư. Lúc này để chạy test ta load file model có định dạng HDF5 (.h5) [12] vào và chạy thử các file test ở dataset LITS Chllenge [2].

Kết quả sau khi test là một file ảnh CT 3d chứa hình ảnh là gan (lá gan) và các mô tế bào ung thư (màu xanh lá) (hình 6-7). Sử dụng phần mềm Theia [3] để xem kết quả ở dạng 3d.



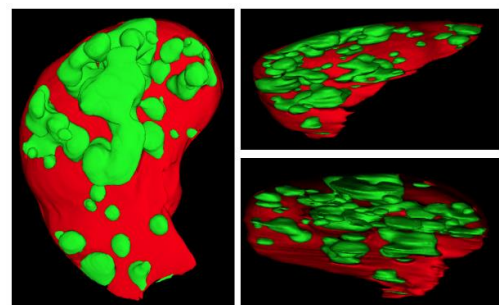
Hình 6: Hình ảnh sau khi test.

Phóng to lá gan và xem nó ở nhiều góc độ khác nhau.



Hình 7: Hình ảnh một lá gan ở nhiều góc độ khác nhau với các mô tế bào ung thư là các hình màu xanh. Có thể thấy lá gan này đã bị hư hại khá nhiều.

Tuy nhiên hình ảnh này chỉ minh họa được các tế bào ung thư nằm trên bề mặt của lá gan. Để xem được bên trong lá gan như thế nào, chúng ta sẽ tăng độ trong suốt của nó lên (hình 8).



Hình 8: Có thể thấy các tế bào ung thư đã ăn sâu vào lá gan rất nhiều.

4. Bàn luận

4.1. Tài nguyên phần cứng cho việc training model

Trong công việc nghiên cứu, một trong những vấn đề khó khăn nhất để có thể hoàn thành nghiên cứu là việc training model. Tập dữ liệu LITS [2] mà chúng tôi thực hiện chứa các file ảnh 3d có định dạng NifTi-1 (.nii), trung bình dung lượng một file từ 150-300 MB. Có thể thấy được nguồn dữ liệu là rất lớn, do đó mà training model cũng rất yêu cầu cấu hình phần cứng rất cao đặc biệt là GPU và RAM.

Thời gian cho việc training khoảng hơn 1 ngày nếu dữ liệu nguyên tập dữ liệu gốc, nhưng trong khi google colab chỉ giới hạn 12 tiếng training. Do đó giải pháp là bỏ bớt đi dữ liệu sao cho thời gian training chỉ khoảng 12 tiếng trở xuống hoặc tự đầu tư phần cứng.

4.2. Sử dụng mô hình kiến trúc mạng khác

Trong nghiên cứu này chỉ sử dụng mô hình kiến trúc mạng Unet để thực hiện. Trong thời gian tới chúng tôi sẽ cố gắng thử nghiệm với các mô hình phân đoạn ảnh khác để so sánh kết quả có gì khác nhau không?.

5. Kết luận

Phân đoạn gan và khối u đóng vai trò quan trọng trong việc chẩn đoán lâm sàng. Nó cung cấp chính xác vị trí của lá gan và khối u nằm trên lá gan đó.

Mô hình hệ thống đào tạo end-to-end Unet cho phân đoạn gan và khối u từ ảnh chụp cắt lớp CT của

một đoạn cơ thể người từ phần sườn tới hông, là một mô hình đã cho thấy được sự hiệu quả và cũng như kết quả mà nó đem lại rất đáng được tin cậy. Trong tương lai nó có thể sẽ là một đóng góp rất to lớn cho nền y học và có ý nghĩa hi vọng mà nó đem lại cho các bệnh nhân và bác sĩ.

Tài liệu tham khảo

- [1] Xiaomeng Li, Hao Chen, Xiaojuan Qi, Qi Dou, Chi-Wing Fu, Pheng Ann Heng. Paper: [“H-DenseUNet: Hybrid Densely Connected UNet for Liver and Tumor Segmentation from CT Volumes”](#), tham khảo source code, github: [“xmengl999/H-DenseUnet”](#).
- [2] Patrick Christ, data set: for training [“LITS-Challenge”](#), for test [“LITS-Challenge-Test-Data”](#).
- [3] Adam Wolf, tool to show 3d volume, github [“adamkwolf/3d-nii-visualizer”](#).
- [4] Bệnh viện Vinmec, [“Sự nguy hiểm của ung thư gan”](#).
- [5] Ths. BS. Nguyễn Hải Nam, [“Tổng quan về ung thư gan”](#).
- [6] Bệnh viện quân y 103, [“Thực trạng và cập nhật ung thư tại việt nam”](#).
- [7] Phương Quỳnh, tham vấn y khoa: BS. Nguyễn Thường Hanh, [“Các con số thống kê về bệnh ung thư gan có thể khiến bạn giật mình”](#).
- [8] Vũ Hữu Tiệp, sách [“Machine learning cơ bản”](#).
- [9] Bùi Tiến Tùng, viblo asia: [“Unet: Kiến trúc mạnh mẽ cho segmentation”](#).
- [10] Zhixuhao, github: [“zhixuhao/unet”](#).
- [11] Phạm Văn Chung, [“\[Deep Learning\] Tìm hiểu về mạng tích chập \(CNN\)”](#).
- [12] Wikipedia, Định dạng HDF5: [“Hierarchical Data Format”](#).