

Phân đoạn ảnh mạch máu võng mạc bằng kiến trúc mạng U-Net

Hứa Tấn Sang - 2252 1239
Phạm Đình Anh Duy - 2252 0337

Khoa Khoa học Máy tính
Trường Đại học Công nghệ Thông tin
Đại học Quốc gia Thành phố Hồ Chí Minh

Ngày 02, tháng 06, năm 2025

Mục lục

1. Tổng quan
2. Động lực nghiên cứu
3. Bộ dữ liệu
4. Mô hình U-Net
5. Các độ đo
6. Thực nghiệm và kết quả
7. Tổng kết
8. Tài liệu tham khảo
9. Q&A

Giới thiệu bài toán

- **Phân đoạn mạch máu võng mạc** là một bước quan trọng trong chẩn đoán các bệnh lý về mắt như:
 - Bệnh võng mạc do tiểu đường (*Diabetic Retinopathy*)
 - Tăng huyết áp võng mạc
 - Tắc tĩnh mạch võng mạc
- **Mục tiêu:**
 - Phân tách chính xác vùng mạch máu trong ảnh chụp đáy mắt
 - Hỗ trợ bác sĩ trong quá trình chẩn đoán và theo dõi điều trị
- **Đầu vào / Đầu ra:**
 - **Đầu vào:** Ảnh võng mạc màu (*RGB*)
 - **Đầu ra:** Ảnh nhị phân (1: mạch máu, 0: nền)
- **Thách thức chính:**
 - Mạch máu mảnh, phân nhánh phức tạp
 - Độ tương phản thấp giữa mạch máu và nền
 - Nhiều và biến đổi ánh sáng trong ảnh

Giới thiệu mô hình U-Net

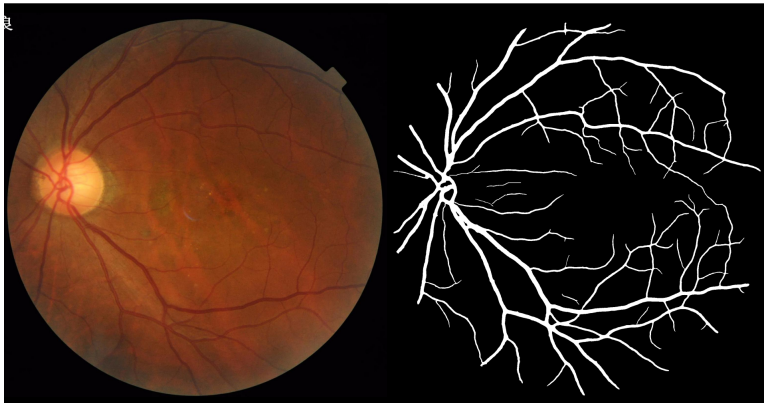
- *U-Net* là một kiến trúc mạng nơ-ron tích chập (CNN) được thiết kế chuyên biệt cho các bài toán phân đoạn ảnh y tế.
- Cấu trúc gồm hai phần chính:
 - **Contracting path (Encoder):** Trích xuất đặc trưng của ảnh thông qua các lớp tích chập và pooling.
 - **Expanding path (Decoder):** Khôi phục độ phân giải ban đầu và tạo ra mặt nạ phân đoạn bằng cách kết hợp đặc trưng từ encoder.
- Đặc điểm nổi bật:
 - Hoạt động tốt với dữ liệu ít, phù hợp với ảnh y tế.
 - Kết nối skip giúp giữ lại thông tin không gian chi tiết.
 - Dễ dàng tùy biến cho nhiều bài toán phân đoạn khác nhau.
- Trong bài toán này, U-Net được sử dụng để phân đoạn mạch máu từ ảnh võng mạc đầu vào.

- **Các bệnh lý vồng mạch** như tiểu đường, tăng huyết áp, và tắc mạch có thể được phát hiện sớm thông qua việc phân tích cấu trúc mạch máu trong ảnh vồng mạch.
- **Phân đoạn mạch máu thủ công** là quá trình tốn thời gian, đòi hỏi chuyên môn cao, và dễ xảy ra sai sót giữa các chuyên gia.
- **Machine Learning và Deep Learning** mang lại khả năng:
 - Tự động hóa quá trình phân đoạn mạch máu với độ chính xác cao
 - Giảm tải cho bác sĩ nhãn khoa trong công tác sàng lọc và chẩn đoán
 - Áp dụng trên quy mô lớn với tốc độ xử lý nhanh

FIVES: A Fundus Image Dataset for Artificial Intelligence based Vessel Segmentation ¹

- Bao gồm 800 cặp ảnh, mỗi cặp bao gồm 1 ảnh RGB ảnh võng mạc và 1 mặt nạ nhị phân của mạch máu.
- Được chia thành 600 ảnh cho tập train và 200 cho tập test.
- Tập train được chia tiếp thành 420 ảnh cho training và 160 cho validate.

¹<https://www.nature.com/articles/s41597-022-01564-3>



Hình 2: Một mẫu trong bộ dữ liệu

U-Net: Là mô hình được phát triển vào năm 2015, nổi bật với hiệu quả cao trong bài toán phân đoạn ảnh y tế. Được giới thiệu lần đầu trong bài báo: "*U-Net: Convolutional Networks for Biomedical Image Segmentation*"².

Giành chiến thắng trong thử thách theo dõi tế bào năm 2015³.

Đặc điểm chính:

- Cấu trúc hình chữ U gồm nhánh mã hóa (encoding) và giải mã (decoding)
- Phù hợp cho các bài toán phân đoạn với dữ liệu y tế hạn chế
- Bảo toàn thông tin không gian nhờ các kết nối skip

²<https://arxiv.org/abs/1505.04597>

³<https://celltrackingchallenge.net/>

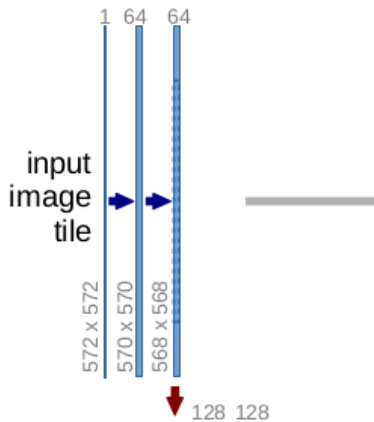
Encoder trong U-Net

Encoder: Là phần bên trái của mô hình U-Net, có nhiệm vụ trích xuất đặc trưng từ ảnh đầu vào.

Thành phần chính:

- Chuỗi các khối **Conv2D** + **ReLU** + **BatchNorm**
- Sau mỗi khối là một lớp **MaxPooling2D** để giảm kích thước không gian
- Số lượng kênh tăng dần để học đặc trưng phức tạp hơn

Kết quả: Mã hóa ảnh đầu vào thành đặc trưng chiều thấp nhưng giàu thông tin.



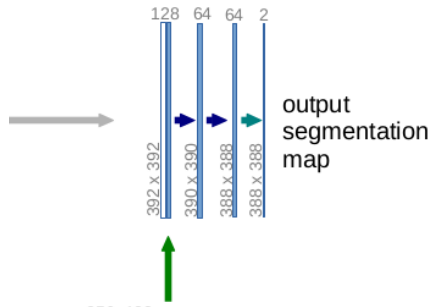
Decoder trong U-Net

Decoder: Là phần bên phải của mô hình U-Net, có nhiệm vụ khôi phục kích thước ảnh và tạo ra ảnh phân đoạn.

Thành phần chính:

- Các lớp **UpSampling2D** hoặc **ConvTranspose2D** để tăng kích thước không gian
- Ghép nối (concatenate) với các đặc trưng tương ứng từ encoder qua các kết nối skip
- Chuỗi các lớp **Conv2D** + **ReLU** giúp tinh chỉnh đặc trưng

Kết quả: Tái tạo lại ảnh đầu ra có cùng kích thước với ảnh gốc, trong đó vùng mạch máu được phân đoạn rõ ràng.



- **IoU (Intersection over Union):**

- Đo mức độ trùng khớp giữa vùng dự đoán và vùng thực.
- $$\text{IoU} = \frac{TP}{TP+FP+FN}$$

- **F1-score:**

- Trung bình điều hòa giữa Precision và Recall.
- $$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

- **Recall (Sensitivity):**

- Tỷ lệ mạch máu thực được phát hiện đúng.
- $$\text{Recall} = \frac{TP}{TP+FN}$$

- **Specificity (SPE):**

- Tỷ lệ vùng nền được dự đoán chính xác.
- $SPE = \frac{TN}{TN+FP}$

- **ROC (Receiver Operating Characteristic):**

- Đường cong thể hiện mối quan hệ giữa **Recall (TPR)** và **FPR** ở nhiều ngưỡng phân loại.
- Càng gần góc trái trên càng tốt.

- **AUC (Area Under the ROC Curve):**

- Diện tích dưới đường cong ROC.
- Giá trị càng gần 1 thì mô hình phân biệt tốt giữa hai lớp (mạch máu / nền).

- Nhóm tiến hành thực nghiệm trên mô hình UNET, ngoài ra còn cài đặt thêm UNET++ để so sánh
- Cấu hình:
 - NVIDIA L40s với 48GB RAM: 4x
 - Batch size: 32
 - Optimzer: Adam Optimizer
 - Learning rate: $3e-4$
 - Weight decay: $1e-5$
 - Epochs: 30
 - Hàm mất mát: ABE Dice Loss

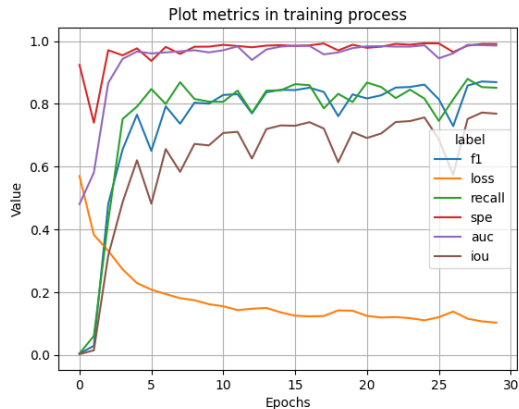
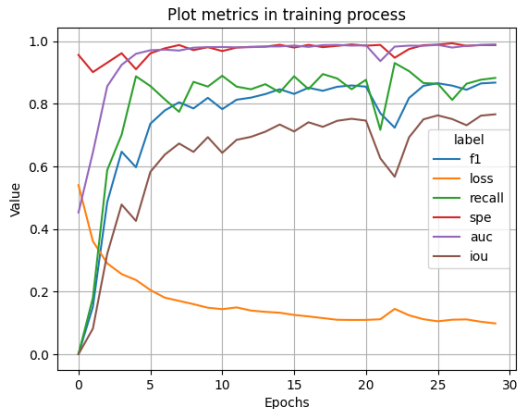
Kết quả

Model	F1	IOU	Recall	SPE	ROC/AUC
UNET	0.8458023	0.7328054	0.8556773	0.9871029	0.986078
UNET++	0.8460372	0.7331581	0.8202765	0.9908601	0.9818744

Bảng: Kết quả của mô hình ⁴

⁴<https://drive.google.com/drive/folders/16Se-MbXDKpSQSo6zQCZRi4OEjSMk2HZR>

Thực nghiệm



Hình 3: So sánh quá trình huấn luyện giữa UNET và UNET++

Tổng kết và hướng phát triển

Tổng kết

- U-Net đã phát triển một kiến trúc đột phá cho bài toán phân đoạn. Giúp cho việc huấn luyện các mô hình phân đoạn nhanh hơn.
- Đặt nền móng cho sự phát triển của hàng loạt các kiến trúc sau này như là Duck-Net, TAOD-CFNET,...

Hướng phát triển:

- Điều chỉnh các lớp bên trong các khối để có thể đạt được độ chính xác cao hơn.
- Ngoài ra cũng có thể cài đặt thêm các phương pháp chú ý hay thay đổi hàm mất mát.



Ronneberger, O., Fischer, P., & Brox, T. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015.
https://doi.org/10.1007/978-3-319-24574-4_28



Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. *UNet++: A Nested U-Net Architecture for Medical Image Segmentation*. In: Deep Learning in Medical Image Analysis, 2018.
https://doi.org/10.1007/978-3-030-00889-5_1

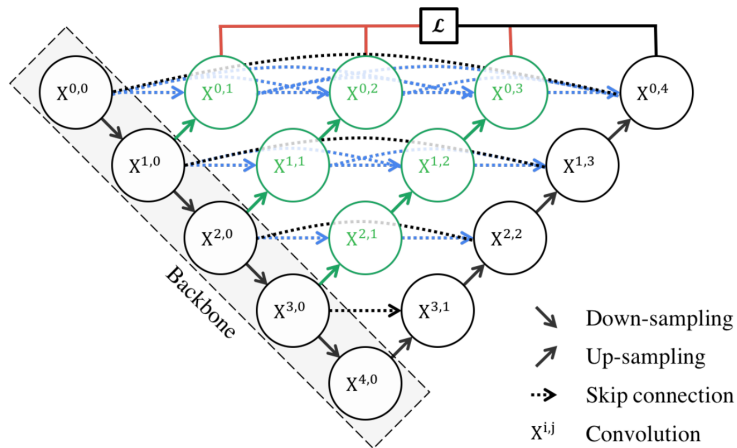


Jin, K., Luo, Y., Zhang, Y., et al. *FIVES: A Fundus Image Dataset for Artificial Intelligence-based Vessel Segmentation*. Scientific Data, 9, 475, 2022. <https://doi.org/10.1038/s41597-022-01564-3>



Zheng, Y., Tian, B., Yu, S., Yang, X., Yu, Q., Zhou, J., Jiang, G., Zheng, Q., Pu, J., & Wang, L. *Adaptive Boundary-Enhanced Dice Loss for Image Segmentation*. Biomedical Signal Processing and Control, vol. 86, 107741, 2025. <https://doi.org/10.1016/j.bspc.2025.107741>





Hình 3: Mô hình UNET++