

TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI

Viện Công nghệ Thông tin và Truyền thông

ĐỒ ÁN MÔN HỌC

Học phần: Project 3 (IT3940)

Colon polyp segmentation and Neoplasm detection

Giảng viên hướng dẫn:

TS. Đinh Viết Sang

Mã lớp

709165

HÀ NỘI, 1/2022

Lời cảm ơn

Thưa thầy, em rất vui vì được thầy hướng dẫn môn học project 3. Em cảm ơn thầy đã đưa em một vài gợi ý cũng như định hướng, giúp em hoàn thành đồ án môn học này. Em hi vọng trong tương lai có thể được thầy dẫn dắt trong một số môn khác.

Em xin chân thành cảm ơn.

Tóm tắt nội dung đồ án

Đồ án môn học tìm hiểu một số công nghệ về phân đoạn ảnh hiện tại, cụ thể là các mạng cơ bản là Unet, Unet++, model đề xuất mới như Neo Unet, áp dụng cho cuộc thi BKAI-IGH NeoPolyp. Bên cạnh đó đồ án trình bày một số thay đổi trên model gốc về Loss function, pretrain backbone và khảo sát ảnh hưởng của chúng.

MỤC LỤC

CHƯƠNG 1. TỔNG QUAN ĐỀ TÀI	1
1.1 Tổng quan các bài toán về computer vision	1
1.2 Giới thiệu bài toán nhận dạng ảnh ý tế	3
CHƯƠNG 2. CÁC NGHIÊN CỨU LIÊN QUAN	5
2.1 Kiến thức cơ bản	5
2.2 Các backbone CNN	5
2.3 HarDNet	10
2.4 Loss Function	11
2.5 Attention Unet	13
CHƯƠNG 3. KIẾN TRÚC ĐỀ XUẤT	15
3.1 Unet	15
3.2 Unet với backbone MobileNet	16
3.3 Unet ++.....	16
3.4 Neo Unet	17
CHƯƠNG 4. KẾT QUẢ THỰC NGHIỆM	19
4.1 Unet	19
4.2 Unet với backbone MobileNet	19
4.3 Unet ++.....	21
4.4 Neo Unet	21
CHƯƠNG 5. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN	23
5.1 Kết luận	23
5.2 Hướng phát triển của đề án trong tương lai	23
TÀI LIỆU THAM KHẢO	24

DANH MỤC HÌNH VẼ

Hình 1.1 Các bài toán liên quan trong computer vision	3
Hình 1.2 Bài toán về nhận diện và phân loại ảnh y tế	4
Hình 2.1 Kiến trúc CNN cơ bản	5
Hình 2.2 Các Backbone thường dùng hiện nay	5
Hình 2.3 Kiến trúc Le-Net 5	6
Hình 2.4 Kiến trúc AlexNet	6
Hình 2.5 Kiến trúc VGG16	7
Hình 2.6 Kiến trúc Inception v2	7
Hình 2.7 ResNet Block	8
Hình 2.8 Kiến trúc DenseNet	8
Hình 2.9 Pyramid Networks	9
Hình 2.10 Depthwise Convolution	9
Hình 2.11 Kiến trúc EfficientNet	10
Hình 2.12 Kiến trúc HarDNet	10
Hình 2.13 Attention Gate	14
Hình 3.1 Kiến trúc Unet	16
Hình 3.2 Kiến trúc Unet ++	17
Hình 3.3 Kiến trúc Neo Unet	17

DANH MỤC BẢNG

Bảng 2.1 Lịch sử cải tiến của các backbone CNN	10
Bảng 2.2 Các hàm Loss cho bài segmentation	11
Bảng 3.1 Các công nghệ hiện tại cho nhận diện ảnh y tế	15
Bảng 4.1 Kết quả thực nghiệm với Unet	19
Bảng 4.2 Kết quả thực nghiệm Unet với backbone MobileNet	21
Bảng 4.3 Kết quả thực nghiệm với Unet ++	21
Bảng 4.4 Kết quả thực nghiệm với Neo Unet	22

CHƯƠNG 1. TỔNG QUAN ĐỀ TÀI

1.1 Tổng quan các bài toán về computer vision

Object Recognition là bài toán liên quan nhận diện vật thể trong ảnh, mục đích chính là có thể hiểu được nội dung tấm ảnh.



Các phương pháp thường được sử dụng gồm có:

- HOG (Histogram of oriented Gradients) feature Extractor and SVM (Support Vector Machine) model
- Bag of features model
- Viola-Jones algorithm

Sau đó là sự ra đời của các phương pháp deep learning (mạng CNN...)

Object Recognition bao gồm 2 nhiệm vụ con là **Image Classification** và **Object Localization**.

Phân loại hình ảnh (image classification)	Định vị vật thể (object localization)
liên quan đến việc gán nhãn cho hình ảnh.	liên quan đến việc vẽ một hộp giới hạn (bounding box) xung quanh một hoặc nhiều đối tượng trong hình ảnh nhằm khoanh vùng đối tượng.
An image of two golden retriever puppies sitting in a field of yellow flowers. The text "Output: 'DOG'" is overlaid in yellow.	The same image of two golden retriever puppies, but with yellow bounding boxes drawn around each puppy to indicate object localization.

Kết hợp lại ta có bài toán **Object Detection**, chú ý là trong bài toán này thì không đơn giản là một loại object nữa mà có thể là nhiều object

CHƯƠNG 5. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

5.1 Kết luận

Đồ án đã trình bày một số phương pháp nhận diện điểm bất thường trong ảnh y tế và phân loại nó làm neoplastic hoặc non-neoplastic. Đồ án có tính ứng dụng trong y tế trong việc chẩn đoán bệnh.

Qua đồ án môn học, em đã tìm hiểu được thêm kiến thức về các backbone cnn, các kiến trúc mạng, hàm loss cho image segmentation. Em học thêm được các kỹ năng về thư viện pytorch, tensorflow... cho bài toán xử lý ảnh

5.2 Hướng phát triển của đồ án trong tương lai

Nếu có thêm thời gian tìm hiểu, đồ án tập trung cải tiến thêm một số phần như giao diện chương trình, tìm hiểu và khảo sát thêm thêm một số mô hình khác như Pranet, Caranet...

Trong phạm vi của môn học, một vài mô hình cũng như phương pháp đã không được em chỉ rõ ràng công thức, cách hoạt động cụ thể mà chỉ gọi sơ về ý tưởng. Khi mở rộng có thể đi sâu vào trình bày những phần này.

Trong quá trình làm đồ án môn học đôi khi có một số phần chưa được hoàn thiện và cần được sửa đổi, bổ sung. Rất mong được thầy góp ý, nhận xét để bài báo cáo của em trở lên hoàn thiện hơn.

TÀI LIỆU THAM KHẢO

- [1] Chollet, François. "Xception: Deep learning with depthwise separable convolutions." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
- [2] Li, Zewen, et al. "A survey of convolutional neural networks: analysis, applications, and prospects." *IEEE Transactions on Neural Networks and Learning Systems* (2021).
- [3] Howard, Andrew G., et al. "Mobilenets: Efficient convolutional neural networks for mobile vision applications." *arXiv preprint arXiv:1704.04861* (2017).
- [4] Lei, Tao, et al. "Medical image segmentation using deep learning: a survey." *arXiv preprint arXiv:2009.13120* (2020).
- [5] Lin, Tsung-Yi, et al. "Feature pyramid networks for object detection." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
- [6] Lan, Phan Ngoc, et al. "NeoUNet: Towards accurate colon polyp segmentation and neoplasm detection." *arXiv preprint arXiv:2107.05023* (2021).
- [7] Unet-zoo <https://github.com/Andy-zhujunwen/UNET-ZOO>
- [8] Neo-unet <https://github.com/lanPN85/neounet>
- [9] Oktay, Ozan, et al. "Attention u-net: Learning where to look for the pancreas." *arXiv preprint arXiv:1804.03999* (2018).
- [10] Chao, Ping, et al. "Hardnet: A low memory traffic network." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019.
- [11] Các hàm loss <https://www.kaggle.com/bigironsphere/loss-function-library-keras-pytorch#Lovasz-Hinge-Loss>
- [12] Focal - tversky loss <https://github.com/nabsabraham/focal-tversky-unet/blob/347d39117c24540400dfe80d106d2fb06d2b99e1/losses.py#L6>
- [13] Tổng hợp một số model segmentation https://github.com/qubvel/segmentation_models
- [14] Độ đo <https://towardsdatascience.com/metrics-to-evaluate-your-semantic-segmentation-model-6bcb99639aa2>
- [15] Focal Tversky Loss: <https://towardsdatascience.com/dealing-with-class-imbalanced-image-datasets-1cbd17de76b5>