

СЕКЦИЯ 4

МОДЕЛИ И АЛГОРИТМЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Ю. ВАН, С. В. АКСЕНОВ

Томский политехнический университет, г. Томск

УДК 004.912

СОВЕРШЕНСТВОВАНИЕ U-NET ДЛЯ ДОСТИЖЕНИЯ РЕГИОНАЛЬНОЙ СЕГМЕНТАЦИИ ЛЕГОЧНЫХ ИНФЕКЦИЙ

Аннотация. Для решения проблемы, связанной с тем, что традиционная U-сеть влияет на эффект сегментации изображений легочной инфекции с переменной формой и разрозненными областями, была предложена новая сеть сегментации изображений, основанная на улучшенной U-Net.

Ключевые слова: сегментация медицинских изображений, U-Net, свертка полостей, область заражения легких.

Введение. Компьютерная диагностика (computer-aided diagnosis, CAD) является одним из важных средств помощи врачам в постановке диагноза, а анализ медицинских изображений включает в себя анализ определенной информации о признаках изображения для его классификации и последующего решения клинических проблем [1]. Сегментация медицинских изображений является одной из самых сложных задач в системах автоматизированного проектирования и играет решающую роль при изучении поражений мозга [2], клеток [3] и поражений кожи [4].

Методы глубокого обучения (Deep Learning, DL), основанные на сверточных нейронных сетях (convolutional neural network, CNN), обладают хорошей способностью к распознаванию признаков. Классические сверточные нейронные сети, такие как VGG [5], R-CNN [6] и ResNet [7], хорошо работают в задачах компьютерного зрения, а производительность сегментации может быть значительно улучшена путем наложения признаков, полученных из нескольких сверточных слоев. Однако CNN в основном решают задачи классификации, а сегментация медицинских изображений требует сегментации изображений на уровне пикселей. Чтобы решить проблему потери традиционными CNN важной контекстной информации в медицинских задачах, в 2015 году была предложена структурная полностью сверточная нейронная сеть (FCN) [8] для кодирования и декодирования, которая хорошо сочетает высокоуровневые семантические признаки и неглубокие детальные текстовые признаки с помощью прыжковых связей. Кроме того, важнейшей особенностью FCN является внедрение апсемплинга и возможность принимать на вход изображения любого размера, а в итоге выдавать результат сегментации с тем же разрешением, что и исходное изображение. Это закладывает основу для последующей семантической сегментации. Однако FCN не в полной мере учитывает взаимосвязь между пикселями при семантической сегментации изображений, что приводит к недостаточной пространственной согласованности, нечувствительности к деталям, нечетким результатам апсемплинга и недостаточно точным результатам сегментации. Ronneberger и др. [9] предложили улучшенную полностью сверточную нейронную сеть U-Net,

U-Net заменяет полностью связанный слой CNN сверточным слоем, вводит концепцию апсемплинга и механизм прыжкового соединения, а также предлагает классическую U-образную полностью сверточную модель сети, основанную на структуре кодера, которая больше подходит для классификации на уровне пикселей.

Проблема исследования. В клинической практике эти медицинские изображения должны быть тщательно изучены опытными радиологами для идентификации заболевания и определения его причины, но этот процесс занимает много времени и требует от рецензента высокой профессиональной компетентности и глубоких знаний в области патологии [10]. Для решения проблемы ручной сегментации и ее неэффективности в литературе [12] используются автоматизированные методы машинного обучения (machine learning, ML) для сегментации медицинских изображений. U-Net отлично справляется с задачами компьютерного зрения, а функции, полученные путем наложения нескольких сверточных слоев, могут значительно улучшить качество сегментации. Однако изображения легочной инфекции по-прежнему сталкиваются с такими проблемами, как разный размер признаков, незаметные области поражения и легкая потеря краевой информации. Чтобы еще больше повысить точность сегментации медицинских изображений, мы предлагаем усовершенствовать на основе сети U-Net.

Цель нашего исследования — автоматизированная сегментация областей поражения на изображениях легочной инфекции с улучшенной точностью сегментации с помощью усовершенствованной модели U-Net. Для достижения цели были поставлены следующие задачи:

- 1) добавление гибридного механизма внимания на слиянии прыжковых соединений каждого слоя модели U-Net, извлечение поверхностных и глубоких признаков соответственно, полное использование контекстной информации признаков разного масштаба и повышение эффективности модели сегментации;
- 2) транспонирующая сверточная билинейная интерполяция в модели U-Net для сохранения одинакового масштаба изображения после дискретизации вверх и вниз, чтобы избежать потери важных признаков после дискретизации вверх.

Материалы и методы. Улучшенная структура U-Net показана на рис. 1. U-Net — это сквозная сеть сегментации медицинских изображений, построенная на основе полной сверточной нейронной сети. Это типичная модель кодирования и декодирования с лево-правой симметричной структурой. Сеть состоит из трех частей: сокращающего пути (кодера) с левой стороны, расширяющего пути (декодера) с правой стороны и прыжковых соединений. Контур сужения состоит из сверточной нейронной сети, которая используется для извлечения признаков изображения. Каждый этап свертки состоит из двух сверток 3×3 и функции активации ReLU. Кроме того, путь сокращения уменьшает выборку карты признаков с помощью слоя максимального объединения 2×2 и удваивает количество каналов признаков после каждого уменьшения выборки. Контур расширения увеличивает выборку карты признаков, транспонируя свертку, восстанавливая разрешение и вдвое уменьшая количество каналов карты признаков. Затем входящие карты признаков, полученные при сокращении, обрезаются, чтобы иметь ту же размерность, что и карты признаков, полученные при повышении дискретизации, а затем они сращиваются и поступают в следующий сверточный слой. Усовершенствованная архитектура U-Net добавляет модуль механизма внимания PSA к оригинальному кодеру U-Net, чтобы учесть различные масштабы пространственных признаков, и разрабатывает многомасштабный гибридный блок нулевой свертки с использованием нескольких слоев различных

сверточных ядер, чтобы избежать потери информативных признаков в процессе понижающей дискретизации. Выходное изображение оригинальной сети немного меньше входного изображения из-за использования свертки без подстановки и операции обрезки при объединении карт признаков. В нашей сети U-Net используется свертка с размахом 1, поэтому размер выходного изображения и размер входного изображения могут быть одинаковыми. В последнем слое сети используется свертка 1×1 для сопоставления вектора признаков с количеством классов, необходимых для сегментации, и вывода сегментированного изображения.

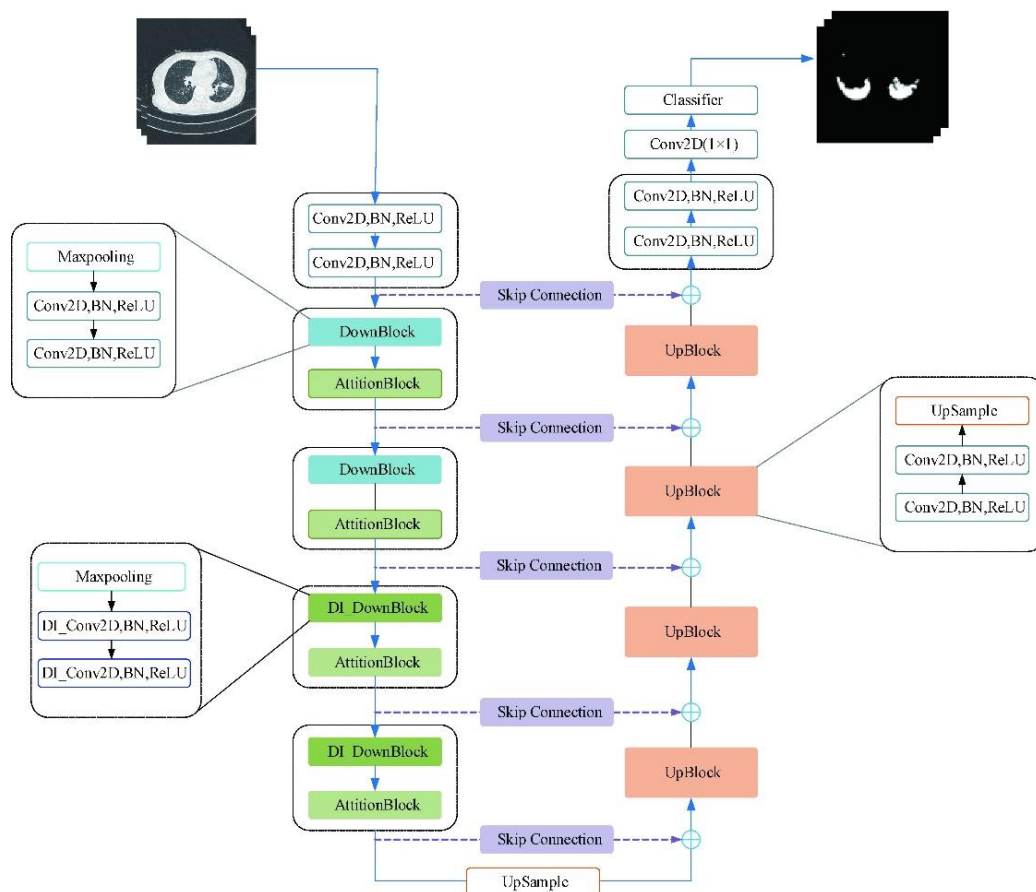


Рис. 1. Улучшенная модель U-Net

Механизм внимания в CNN может улавливать зависимости на больших расстояниях, но этот подход более сложен и чувствителен к шуму. Пирамидальное разделение внимания (pyramid split attention, PSA) — это модуль внимания, который захватывает пространственную информацию в разных масштабах для обогащения пространства признаков. Благодаря объединению контекстной информации в разных масштабах и установлению зависимостей на больших расстояниях для улучшения внимания на уровне пикселей, PSA является простой, легкой, простой и эффективной, и может эффективно улучшить производительность модели. Путем группировки и свертки входного тензора извлекается информация разных масштабов, а весовые коэффициенты внимания для карт признаков разных масштабов получает модуль внимания каналов. Наконец, PSA объединяет и нормализует каждую группу весов внимания каналов, а затем взвешивает их с выходом модуля SPC, чтобы получить окончательный ре-

зультат. Модуль SPC — это модуль свертки, который способен извлекать информацию в разных масштабах. Его основная операция заключается в разделении входного тензора из канала на S групп, и каждая группа свертывается с ядрами свертки разных размеров для получения различных масштабов рецептивных полей. Наконец, модуль SPC объединяет карты признаков после свертки разных размеров на канале в качестве выходного сигнала модуля. Таким образом, модуль SPC может эффективно извлекать информацию разных масштабов и предоставлять более богатое пространство признаков для последующей обработки. Структура модуля PSA показана на рис. 2.

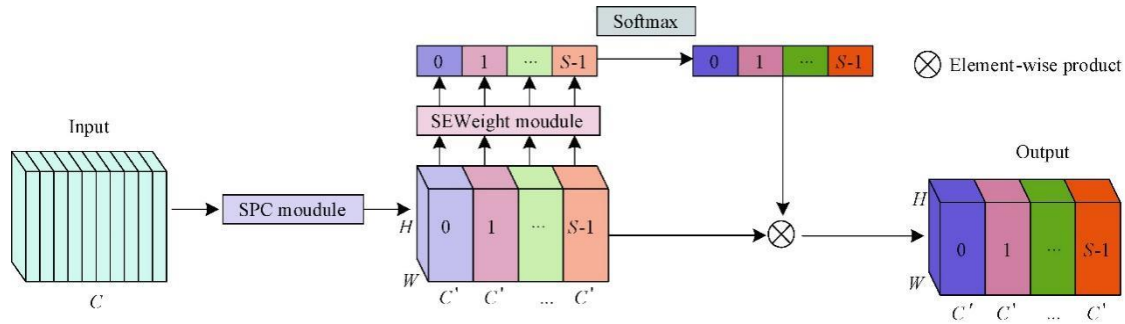


Рис. 2. Улучшенная модель U-Net

Эксперименты проводились на Windows 10 с использованием среды Anaconda, фреймворка глубокого обучения Pytorch и аппаратного устройства NVIDIA RTX3060. В качестве параметров оптимизации использовался алгоритм адаптивной оценки момента (adaptive moment estimation, Adam) со 100 итерациями, размером партии 16, начальной скоростью обучения 0,001 и дополнением параметров обучения с помощью данных оптимизатора Adam. Используется набор данных COVID-19CT scans, в котором соотношение обучающего, тестового и валидационного наборов составляет 7:2:1.

Для количественной оценки точности результатов сегментации сетевой модели, задействованной в данном исследовательском эксперименте, в качестве оценочных индексов в данной работе выбраны коэффициент Dice, точность пикселей (ассигасу, Acc) и средний коэффициент пересечения и слияния (MIoU). Все вышеперечисленные оценочные индексы являются комплексными показателями, которые могут измерить степень правильной сегментации, недостаточной сегментации и неправильной сегментации в результатах сегментации.

$$Dice = \frac{2TP + FP + FN}{2TP} \quad (1)$$

$$Acc = \frac{TP + TN + FT + FN}{TP + TN} \quad (2)$$

$$MIoU = \frac{TP + FP + FN}{TP} \quad (3)$$

где TP — истинно положительный результат для правильно сегментированной области поражения; FP — ложноположительный результат для неправильно классифицированной области поражения; TN — истинно отрицательный результат для правильно сегментированной области без поражения; FN — ложноотрицательный результат для неправильно классифицированной области без поражения.

Результаты. Сравнительные эксперименты по сегментации поражения легких проводились с использованием нескольких моделей глубокого обучения, и результаты сравнительных экспериментов представлены на рис. 3. Как видно из рис. 3, сегментированная область поражения методом данной статьи ближе к реальному образцу.

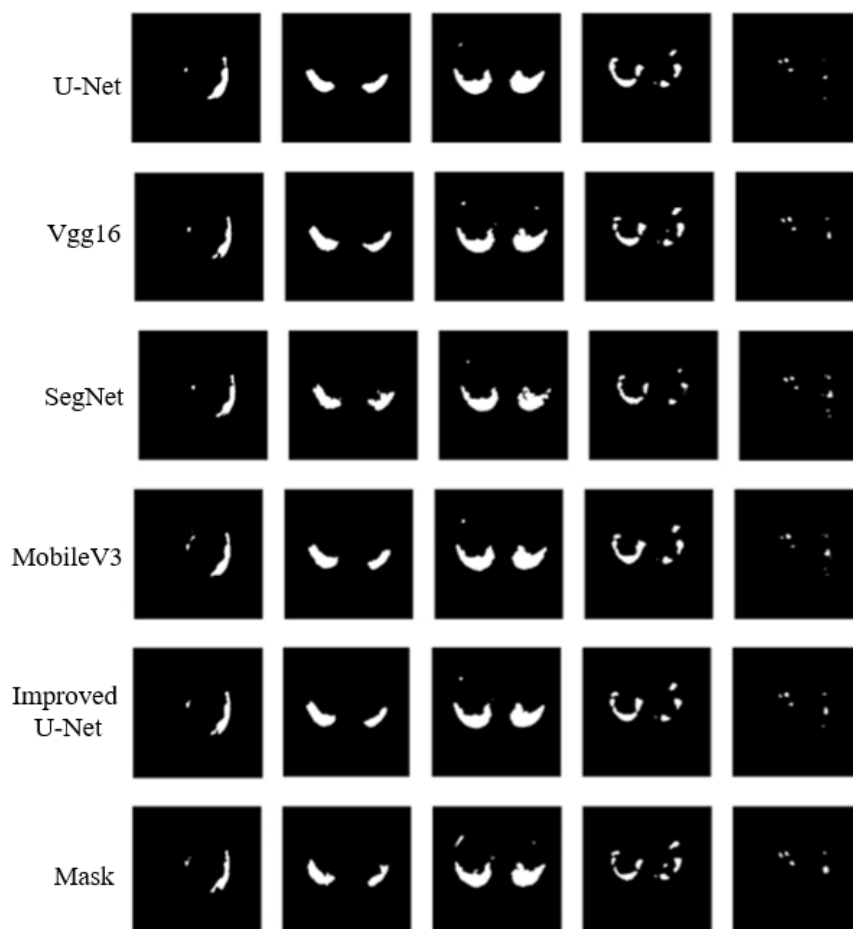


Рис. 3. Влияние различных моделей сегментации

Дальнейший анализ показывает, что точность предложенного нами модельного подхода также выше, чем у других основных методов (табл. 1). Из табл. 1 видно, что с точки зрения точности и скорости сегментации изображений на КТ-изображениях используемый нами метод улучшается на 0.009, 1.9%, 1.47% и 0.0087 с по четырем аспектам, а именно Dice, Acc, MIoU, и времени вывода (Benchmark Time), по сравнению с U-Net.

Таблица 1

Таблица точности сегментации поражений на СТ-изображениях

<i>Model</i>	<i>Dice</i>	<i>Acc/%</i>	<i>MIoU/%</i>	<i>Benchmark Time/s</i>
U-Net[9]	0.847	97.89	87.83	0.0179
VGG16U-Net[12]	0.737	97.49	85.82	0.0390
SegNet[13]	0.882	97.92	88.24	0.0229
MobileV3U-Net[14]	0.879	97.67	86.86	0.0058
Improved U-Net	0.856	99.79	89.3	0.0092

Заключение. СТ-изображения легких были сегментированы с помощью улучшенной сети U-Net, и были получены хорошие результаты. Экспериментальные результаты показывают, что улучшенная сеть U-Net может эффективно сегментировать медицинские изображения с высокой точностью и надежностью, а MIOU сегментации поражений в СТ достигает 89.3%, что явно лучше, чем у существующей сетевой модели, что обеспечивает новое решение в области сегментации медицинских изображений.

СПИСОК ЛИТЕРАТУРЫ

1. DUG, CAO X., LIANG J. et al. Medical image segmentation based on U-Net:A review[J] // Journal of Imaging Science and Technology. — 2020. — 64 (2). — P. 1-12.
2. VALVERDE J.M., SHATILLO A., DE FEO R. et al. Automatic cerebral hemisphere segmentation in rat MRI with ischemic lesions via attention-based convolutional neural networks[J] // Neuroinformatics. — 2023. — 21 (1). — P. 57-70.
3. MENZE B.H., JAKAB A., BAUER S. et al. The multimodal brain tumor image segmentation benchmark (BRATS) [J].IEEE // Transactions on Medical Imaging. — 2015. — 34 (10). — P. 1993-2024.
4. BAID U., GHODASARA S., MOHAN S. et al. The RSNA-ASNR-MICCAIBraTS 2021 benchmark on brain tumor segmentation and radiogenomic classification[J] // Computer Science. — 2021. — DOI:10.48550/arXiv.2107.02314.
5. SIMONYAN K., ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J] // Computer Science. — 2014. — DOI:10.48550/arXiv.1409.1556.
6. GIRSHICK R., DONAHUE J., DARRELL T. et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. 2014 IEEE // Conference on Computer Vision and Pattern Recognition. — 2014.
7. HE K., ZHANG X., REN S. et al. Deep residual learning for image recognition [C]. 2016 IEEE // Conference on Computer Vision and Patter Recognition (CVPR). — 2016.
8. SHELHAMER E., LONG J., DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE // Transactions on Pattern Analysis and Machine Intelligence. — 2017. — 39 (4). — P. 640-651.
9. RONNEBERGERO, FISCHER P., BROX T. U-Net: Convolutional networks for biomedical image segmentation [J] // Computer Science. 2015. DOI: 10.48550/arXiv.1505.04597.
10. UMMADI V. U-Net and its variants for medical image segmentation: A short review[J] // Electrical Engineering and Systems Science. 2022. DOI:10.48550/arXiv.2204.08470.
11. GALISOT G., RAMEL J., BROUARD T. et al. Visual and structural feature recombination in an interactive machine learning system for medical image segmentation[J] // Machine Learning with Applications. 2022. 8. P. 100294.
12. GUDMUNDSSON E., STRAUSC M., ARMATO S.G. Pre-trained deep convolutional neural networks for the segmentation of malignant pleural mesothelioma tumor on CT scans[C] // SPIE Conference on Computer-Aided Diagnosis. — 2019.
13. VIJAY B., ALEX K., ROBERTO C. Seg Net: A deep convolutional encoder-decoder architecture for image segmentation [J] // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2017. — 39 (12). — P. 2481-2495.
14. HOWARD A., SANDLER M., CHU G. et al. Searching for MobileNetV3 [C]. 2019 IEEE/CVF // International Conference on Computer Vision (ICCV). — 2019.