

# СЕГМЕНТАЦІЯ БАГАТОВИМІРНИХ МЕДИЧНИХ ЗОБРАЖЕНЬ

Скорік В. А.

Науковий керівник – к.т.н., ст. викл. Дейнеко А.О.

Харківський національний університет радіоелектроніки  
(61166, Харків, пр. Науки, 14, каф. ШІ, тел. (057) 702-13-37)

e-mail: [viktoriiia.skorik@nure.ua](mailto:viktoriiia.skorik@nure.ua)

An ever-growing amount of medical data and the development of intelligent technologies open up a wide range of research opportunities. Thus, this work is dedicated to researching the modern methods for medical segmentation.

Методи сегментації медичних зображень знаходять межі об'єкта інтересу в зображенні і призначають семантичні ярлики пікселям або вокселям. Ефективні та дієві методи сегментації медичних зображень дають безліч важливих клінічних переваг:

1. Персоналізація зображення цільового регіону конкретного пацієнта таким чином, щоб не опромінювати пацієнта більше, ніж потрібно, але і не втратити цінну інформацію. Тому ефективна детекція органів в попередніх зображеннях дозволяє персоналізувати сканування, і при цьому зменшує час, вартість і одержувану дозу опромінення сканування.

2. Читання медичних зображень для діагностики, терапії та планування операцій. Під час читання медичних зображень з метою виявлення уражень (хвороби) в конкретних органах і частинах тіла, рентгенологу необхідно відмітити потрібний регіон, де була виявлена хвороба, і скласти звіт. Сегментація медичних зображень дозволяє автоматизувати процес виявлення і визначення уражених регіонів. Що прискорює час визначення діагнозу і зменшує вартість діагностики.

3. Вимірювання обсягу органу є важливим для кількісної діагностики хвороби. Але визначення цільового органу і кількісні виміри забирають багато часу лікаря, особливо коли зображення багатовимірні. Сегментація зображень дозволяє автоматизувати кількісні вимірювання і надає матеріал для моделювання.

Завдяки прогресу і повсюдного поширення технологій візуалізацій медичних даних ростуть потреби в нових і більш досконалих методах і алгоритмах аналізу даних, таких, наприклад, як методи глибокого навчання.

U-Net вже став класикою в сегментації зображень, в тому числі і медичних. А так же став прародичем безлічі його модифікацій таких, наприклад, як: 3D U-Net, UNet++, R2U-Net, Graph U-Net та інших.

На високому рівні абстракції U-Net можна уявити як архітектуру, яка бере карту ознак (англ. feature map) отриману під час конвертації вхідного зображення в вектор в цілях конвертації вектора в сегментоване зображення.

Даня архітектура складається з трьох частин: стиснення, вузького місця і розширення. Вхідне зображення подається в енкодер, який складається з послідовних шарів стиснення зображення до вектора. Кожен блок приймає вхідні дані, застосовує два згортальних шару  $3 \times 3$ , за якими слід максимізаційне агрегування (англ. max pooling) розміром  $2 \times 2$ . Кількість ядер після кожного блоку подвоюється, тому архітектура здатна вивчати складні структури. Самий нижній шар є посередником між блоками стиснення і розширення. Він використовує два згортальних шари  $3 \times 3$ , за яким слідує шар розгортки (англ. convolution up)  $2 \times 2$ .

Декодер подібен енкодеру тільки з шарами розгортки (convolution up). Кожен шар передає вхідні дані двом згортальним шарам  $3 \times 3$ , за якими слід upsampling шар  $2 \times 2$ . До кожного блоку розширення на вхід надсилається карта ознак з блоку стиснення. Це гарантує, що функції, які були вивчені під час стиснення зображення, будуть використані для його реконструкції. Кількість блоків розширення дорівнює кількості блоків стиснення. Після цього результуючий вектор проходить через інший шар CNN  $3 \times 3$  з кількістю ознак, що дорівнює кількості бажаних сегментів.

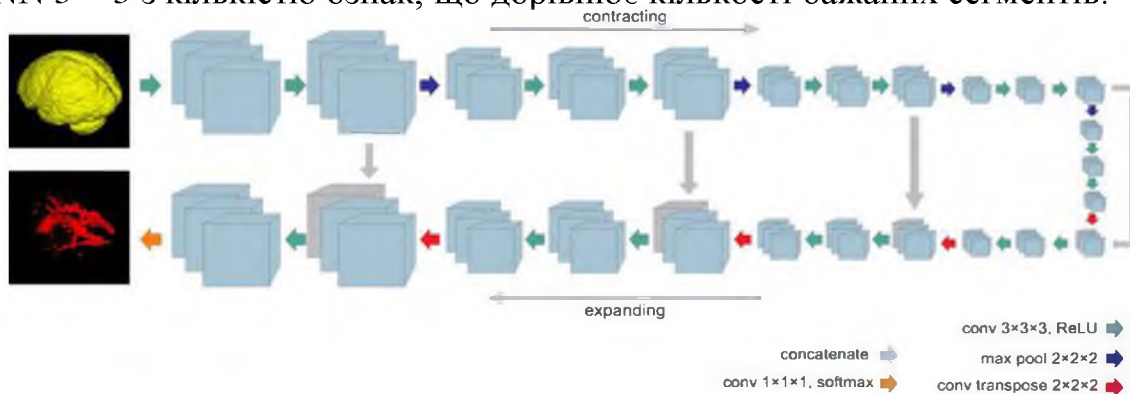


Рисунок 1 – архітектура на основі 3D U-Net

Архітектура 3D U-Net натхненна архітектурою U-Net. Отличее тривимірної архітектури полягає в тому, що енкодері по два згортальних шари з розміром ядра  $3 \times 3 \times 3$ , за якими слід ReLU, а потім max pooling  $2 \times 2 \times 2$  з кроком два в кожному вимірі. У блоці розширення шари розгортання з розміром ядра  $2 \times 2 \times 2$  з кроками по два в кожному вимірі, за якими слідує по два згортальних шари з розміром ядра  $3 \times 3 \times 3$ , за якими слід ReLU. У вузькому місці шар згортки  $1 \times 1 \times 1$  зменшує кількість вихідних каналів до кількості ярликів, що дорівнює 3.

#### Список джерел

1. U-Net: Convolutional Networks for Biomedical Image Segmentation [Електронний ресурс] – 2015. – Режим доступу до ресурсу: <https://arxiv.org/abs/1505.04597>.
2. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation [Електронний ресурс] – 2016. – Режим доступу до ресурсу: <https://arxiv.org/abs/1606.06650>.