

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И МАШИННОЕ ОБУЧЕНИЕ / ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

DOI: <https://doi.org/10.60797/IRJ.2024.143.158>

ГЕНЕРАЦИЯ БИМЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ ДЛЯ АУГМЕНТАЦИИ ДАННЫХ С ПОМОЩЬЮ ГЕНЕРАТИВНО-СОСТЯЗАТЕЛЬНЫХ СЕТЕЙ

Научная статья

Новоселов И.Э.^{1,*}, Смирнов А.А.²

¹ ORCID : 0009-0005-7078-2663;

² ORCID : 0000-0002-7312-7219;

^{1,2} Уральский Федеральный Университет имени первого Президента России Б. Н. Ельцина, Екатеринбург, Российская Федерация

* Корреспондирующий автор (novivan001[at]gmail.com)

Аннотация

Разработка медицинских приложений является важным направлением в сфере здравоохранения. Однако для создания таких приложений необходимы точные аннотации биомедицинских изображений, которые часто являются редкими и сложными для получения. Использование нейронных сетей для генерации таких изображений может значительно увеличить объем данных.

Одним из наиболее эффективных методов генерации изображений является использование генеративно-сопоставительных сетей (GAN), что может помочь для аугментации данных и улучшения качества сегментации медицинских снимков. Особенно полезен данный метод в случаях, когда доступ к реальным данным ограничен или когда требуется большое количество данных для обучения моделей машинного или глубокого обучения.

Цель исследования заключается в создании метода генерации биомедицинских изображений с помощью GAN для аугментации данных.

Ключевые слова: аугментация, нейронные сети, генеративно-сопоставительные сети, снимки мрт, python.

BIOMEDICAL IMAGES GENERATION FOR DATA AUGMENTATION USING GENERATIVE ADVERSARIAL NETWORKS

Research article

Novoselov I.E.^{1,*}, Smirnov A.A.²

¹ ORCID : 0009-0005-7078-2663;

² ORCID : 0000-0002-7312-7219;

^{1,2} Ural Federal University named after the first President of Russia B. N. Yeltsin, Ekaterinburg, Russian Federation

* Corresponding author (novivan001[at]gmail.com)

Abstract

Medical application development is an important direction in the healthcare. However, this task requires accurate annotations of biomedical images, which are often sparse and difficult to obtain. Neural networks can be used to generate such images to increase the amount of data significantly.

One of the most effective image generation methods is the use of generative adversarial networks (GANs), which can help in data augmentation and improve the quality of medical image segmentation. This technique is especially useful in cases where access to real data is limited or when a large amount of data is required to train machine or deep learning models.

The aim of the research is to create the biomedical image generation method by using GAN for data augmentation.

Keywords: augmentation, neural networks, generative adversarial networks, mri scans, python.

Введение

Точные аннотации биомедицинских изображений необходимы для разработки различных медицинских приложений. Задачи аннотирования требуют высокого уровня медицинских знаний, и, следовательно, можно использовать уже размеченные наборы данных для решения подобных задач. В последние годы наблюдается колоссальный скачок в разработке и применении методов машинного и глубокого обучения для создания новых снимков, используя размеченные изображения из конкретного набора данных. Таким образом, генерация биомедицинских изображений с помощью генеративно-сопоставительной сети (GAN) [1] может быть использована для аугментации данных, что может повысить эффективность сегментации изображений, добиться экономии времени специалистов на ручную аннотацию снимков.

Совместное применение синтетических и реальных биомедицинских данных демонстрирует хорошие результаты в задачах классификации и сегментации изображений. Например, Frid-Adar et al. [2] представили улучшение классификации поражений печени с использованием GAN в качестве метода увеличения данных по сравнению с классическими методами аугментации, Jin et al. [3] продемонстрировали повышение надежности прогрессивной целостной вложенной сети за счет использования GAN в задаче сегментации патологий легкого.

Целью работы является создание метода генерации биомедицинских изображений для аугментации данных с использованием генеративно-сопоставительных сетей.

Методы и принципы исследования

Аугментация – метод увеличения набора данных с помощью модификации существующих данных или создания новых данных из конкретного набора данных.

Методы аугментации показывают хорошие результаты в задачах классификации, сегментации изображений, улучшают обобщающую способность используемой модели. При этом осуществляется модификация лишь тренировочного набора данных, а тестовый и проверочный наборы остаются без изменений. Для небольших наборов данных методы аугментации могут использоваться в качестве регуляризации данных, что препятствует переобучению нейронной сети [4].

Так, для аугментации биомедицинских изображений применяются следующие методики:

- 1) поворот на конкретный угол;
- 2) сдвиг по оси;
- 3) переворот относительно оси;
- 4) изменение яркости;
- 5) изменение размера;
- 6) использование генеративно-сопоставительных сетей.

2.1. Набор данных

Для исследования возможности генерации биомедицинских изображений из ограниченного набора данных с целью аугментации был использован набор снимков МРТ головного мозга с опухолями, который представлялся в рамках соревнования Multimodal Brain Tumor Segmentation Challenge 2020. Датасет содержал 369 снимков в формате NIFTI с разрешением 240 x 240 x 155 вокселей. Каждому изображению сопоставляется набор аннотаций, размеченных вручную экспертами, следуя протоколам, результаты их работы одобрены сертифицированной экспертной комиссией нейрорадиологов [5], [6] (см. рисунок 1). Каждая аннотация представлена масками различных классов опухолей: necrotic/non-enhancing tumor core, peritumoral edema, enhancing tumor core.

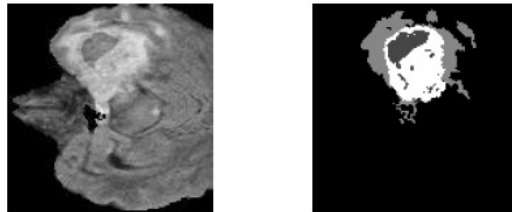


Рисунок 1 - Снимок МРТ головного мозга и соответствующая ему аннотация
DOI: <https://doi.org/10.60797/IRJ.2024.143.158.1>

Из снимков данного датасета был составлен собственный в размере 30 изображений, взятых случайным образом. Затем было изменено разрешение каждого изображения на 256 x 256 пикселей, а также для удобства изменен формат снимков на TIFF.

2.2. Архитектура нейронной сети

Предлагаемая модель состоит из двух сетей: генератора и дискриминатора. Генератор изучает разметку входного изображения, чтобы создать выходное изображение, подобное аннотированному изображению из набора данных, чтобы «обмануть» дискриминатор. В то время как дискриминатор принимает входные данные и синтезированные изображения или аннотированное изображение на вход и представляет вероятность определения реальных и синтезированных изображений. Генератор построен с помощью U-Net [7], а дискриминатор построен с помощью сверточной нейронной сети (CNN) [8]. Архитектура разрабатываемой нейронной сети представлена на рисунке 2.

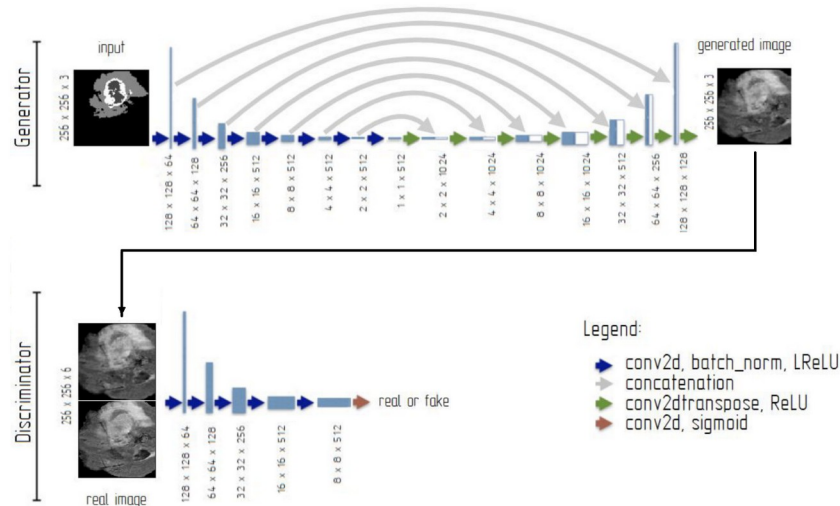


Рисунок 2 - Архитектура нейронной сети
DOI: <https://doi.org/10.60797/IRJ.2024.143.158.2>

Примечание: *Generator* – генератор, *Discriminator* – дискриминатор, *input* – вход, *generated image* – сгенерированное изображение, *real image* – реальное изображение, *real or fake* – реальное или искусственное, *conv2d* (синяя стрелка) – сверточный слой с ядром (4 x 4) и шагом (2 x 2), *batch_norm* – слой пакетной нормализации, *LReLU* – функция активации *leaky ReLU*, *concatenation* – объединение, *conv2dtranspose* – слой транспонированной свертки с ядром (4 x 4) и шагом (2 x 2), *ReLU* – функция активации *ReLU*, *conv2d* (коричневая стрелка) – сверточный слой с ядром (4 x 4) и шагом (1 x 1), *sigmoid* – функция активации *sigmoid*

Из рисунка 2, генератор состоит из:

- входа, на него подается RGB-изображение размером 256 x 256 пикселей;
- семи блоков понижения дискретизации, созданных свертками с ядром (4 x 4), шагом (2 x 2), пакетной нормализацией, функцией активации *Leaky ReLU* и инициализатором ядра *RandomNormal*. Соотношения сверточных фильтров в блоках: 64, 128, 256, 512, 512, 512, 512;
- одного блока «bottleneck», созданный путем свертки с ядром (4 x 4), шагом (2 x 2), функцией активации *ReLU*, инициализатором ядра *RandomNormal* и 512 фильтрами;
- семи блоков повышения дискретизации, созданных транспонированными свертками с ядром (4 x 4), шагом (2 x 2), «dropout» слоем, функцией активации *ReLU* и инициализатором ядра *RandomNormal*. Каждый блок повышения дискретизации объединяется с блоком понижения дискретизации с помощью «пропускных» соединений. Соотношение фильтров транспонированной свертки в блоках: 512, 512, 512, 512, 256, 128, 64;
- выхода, составленного с помощью свертки с ядром (4 x 4), шагом (2 x 2), функцией активации *tanh* для создания изображений в диапазоне [-1; 1] и инициализатор ядра *RandomNormal*. Количество фильтров равно количеству классов изображения.

Дискриминатор состоит из:

- входа, на него поступают изображения RGB (входные и сгенерированные изображения или аннотированные изображения) размером 256 x 256;
- пяти блоков понижения дискретизации, созданных свертками с ядром (4 x 4), шагом (2 x 2), пакетной нормализацией, функцией активации *Leaky ReLU* и инициализатором ядра *RandomNormal*. Соотношения сверточных фильтров в блоках: 64, 128, 256, 512, 512;
- выхода, представленного сверткой с 1 фильтром, ядром (4 x 4), шагом (1 x 1), сигмовидной функцией активации, инициализатором ядра *RandomNormal*.

2.3. Функция потерь

Целевая функция GAN может быть выражена как

$$L_{GAN}(G, D) = E_{x,y} [\log D(x, y)] + E_{x,y} [\log(1 - D(x, G(x, z)))] \quad (1)$$

где *G* – генератор, *D* – дискриминатор, *x* – входное изображение, *y* – выходное изображение, *z* – вектор случайного шума.

Генератор учится минимизировать ошибку против дискриминатора, а дискриминатор учится максимизировать ее. Таким образом, целевая функция GAN:

$$L_G^* = \arg \min_G \max_D L_{GAN}(G, D) \quad (2)$$

Необходимо управлять обучением генератора, чтобы он синтезировал изображения похожие на реально существующие, а не только «обманывал» дискриминатор. Таким образом, полезно комбинировать целевую функцию GAN с традиционной функцией потерь (например, средней абсолютной ошибкой). Следовательно, окончательная функция потерь GAN может быть выражена как:

$$L = L_G^* + \alpha L_1(G), \quad (3)$$

где α – скалярный весовой коэффициент ($\alpha \geq 0$), $L_1(G)$ – средняя абсолютная ошибка генератора.

2.4. Оптимизация и регуляризация

И генератор, и дискриминатор обучаются с помощью оптимизатора Adam [9] с $\text{Learning_rate} = 0,0001$ и $\beta_1 = 0,5$. Для исключения переобучения сети используется регуляризация «dropout» [10] (позволяет не выполнять операции в сверточных слоях с некоторой вероятностью) в блоках повышения дискретизации генератора с вероятностью 0,3.

2.5. Программные средства

Предлагаемая модель была создана с использованием языка программирования Python и его библиотек Keras и TensorFlow для проектирования и обучения нейронной сети. Нейронная сеть была обучена в течение 100 эпох, и ее веса сохранялись каждый раз, когда в конце каждой эпохи достигались более низкие потери на проверочном наборе данных. Модель обучалась на графическом процессоре Nvidia K80 12 GB.

Основные результаты

После обучения нейронной сети, а также сдвигов, поворотов, переворотов изображений по осям удалось увеличить число изображений в наборе данных с 30 до 300 изображений. Примеры сгенерированных снимков МРТ головного мозга представлены на рисунке 3.

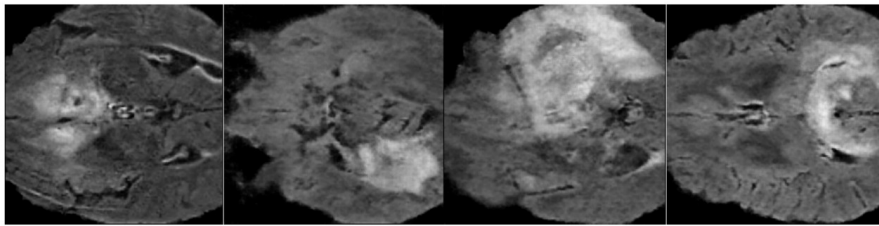


Рисунок 3 - Сгенерированные снимки МРТ головного мозга
DOI: <https://doi.org/10.60797/IRJ.2024.143.158.3>

После аугментации качество модели оценивалось путем сегментации этого набора данных с использованием U-Net. Однако эта архитектура требует аннотированного набора данных для обучения. Эта проблема была решена с помощью онлайн-сервиса APEER для разметки изображений. Затем набор данных был разделен на два набора: обучающий набор (255 изображений) и проверочный набор (45 изображений), которые использовались для обучения U-Net в 50 эпох. Также оценка проведена для тестового набора данных, состоящего из 25 изображений.

По окончании обучения нейронной сети была проведена оценка качества модели с помощью метрик Accuracy, Mean IoU, Dice Score.

Метрику Accuracy можно вычислить по формуле (4).

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

где TP – истинно положительный исход;

TN – истинно отрицательный исход;

FP – ложно положительный исход;

FN – ложно отрицательный исход.

Метрику Mean IOU можно вычислить по формуле (5).

$$\text{Mean IOU} = \frac{1}{N} \cdot \sum_{i=1}^N \frac{TP_i}{TP_i+FP_i+FN_i} \quad (5)$$

где TP_i – истинно положительный исход в i -классе;

FP_i – ложно положительный исход в i -классе;

FN_i – ложно отрицательный исход в i -классе;

N – число классов.

Метрику Dice Score можно вычислить по формуле (6).

$$\text{Dice Score} = \sum_{i=1}^N \frac{2 \cdot T_i \cdot P_i + S}{T_i + P_i + S} \quad (6)$$

где T_i – истинное значение переменной в i -классе;

P_i – предсказанное значение переменной в i -классе;

S – сглаживание;

N – число классов.

В таблице 1 представлены результаты обучения базовой версии U-Net на датасете без аугментации данных и с аугментацией данных.

Таблица 1 - Результаты обучения

DOI: <https://doi.org/10.60797/IRJ.2024.143.158.4>

Метрика	Без аугментации			С аугментацией		
	Train	Val	Test	Train	Val	Test
Accuracy	0,90	0,92	0,87	0,94	0,85	0,90
Mean IoU	0,43	0,37	0,38	0,69	0,58	0,59
Dice Score	0,27	0,25	0,25	0,71	0,60	0,66

Примечание: *Train* – Обучающий набор, *Val* – Проверочный набор, *Test* – Тестовый набор

Из таблицы 1 видно, что показатели качества модели по метрикам Mean IoU и Dice Score значительно улучшились, а по метрике Accuracy остались примерно на том же уровне после аугментации данных.

Заключение

Благодаря использованию генеративно-сопоставительных сетей возможно создание реалистичных биомедицинских изображений с целью увеличения конкретного набора данных.

В этой работе предложен метод создания реалистичных биомедицинских изображений с помощью GAN для увеличения ограниченного набора данных. Модель демонстрирует свою эффективность в задаче сегментации опухоли для случаев с увеличением данных и без нее, достигая значений метрик 0,71 против 0,27 Dice Score, 0,69 против 0,43 Mean IoU на обучающем наборе данных, 0,60 против 0,25 Dice Score, 0,58 против 0,37 Mean IoU на проверочном наборе данных и 0,66 против 0,25 Dice Score, 0,59 против 0,38 Mean IoU на тестовом наборе данных.

Кроме того, в качестве будущей работы предложенную модель можно использовать для создания изображений для других наборов данных и модифицировать путем оптимизации сетей генератора и дискриминатора.

Конфликт интересов

Не указан.

Рецензия

Все статьи проходят рецензирование. Но рецензент или автор статьи предпочли не публиковать рецензию к этой статье в открытом доступе. Рецензия может быть предоставлена компетентным органам по запросу.

Conflict of Interest

None declared.

Review

All articles are peer-reviewed. But the reviewer or the author of the article chose not to publish a review of this article in the public domain. The review can be provided to the competent authorities upon request.

Список литературы на английском языке / References in English

1. Goodfellow I. Generative adversarial nets / I. Goodfellow, J. Pouget-Abadie, M. Mirza [et al.] // Advances in neural information processing systems. — 2014. — P. 2672–2680.
2. Frid-Adar M. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification / M. Frid-Adar, I. Diamant, E. Klang [et al.] // Neurocomputing. — 2018. — Vol. 321. — P. 321–331.
3. Jin D. CT-realistic lung nodule simulation from 3D conditional generative adversarial networks for robust lung segmentation / D. Jin, Z. Xu, Y. Tang [et al.] // International Conference on Medical Image Computing and Computer-Assisted Intervention. — Springer, Cham, 2018. — P. 732–740.
4. Yuheng S. Image segmentation algorithms overview / S. Yuheng, Y. Hao. — 2017.
5. Menze B. H. The multimodal brain tumor image segmentation benchmark (BRATS) / B. H. Menze, A. Jakab, S. Bauer [et al.] // IEEE transactions on medical imaging. — 2014. — Vol. 34. — № 10. — P. 1993–2024.
6. Bakas S. Advancing the Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features / S. Bakas, H. Akbari, A. Sotiras [et al.] // Scientific data. — 2017. — Vol. 4. — № 1. — P. 1–13.
7. Ronneberger O. U-net: Convolutional networks for biomedical image segmentation / O. Ronneberger, P. Fischer, T. Brox // International Conference on Medical image computing and computer-assisted intervention. — Springer, Cham, 2015. — P. 234–241.
8. Li Z. A survey of convolutional neural networks: analysis, applications, and prospects / Z. Li, F. Liu, W. Yang [et al.] // IEEE transactions on neural networks and learning systems. — 2021.
9. Kingma D. P. Adam: A method for stochastic optimization / D. P. Kingma, J. Ba. — 2014.
10. Cai S. Effective and efficient dropout for deep convolutional neural networks / S. Cai, Y. Shu, G. Chen [et al.] — 2019.