

Расширение датасета для обучения нейронных сетей в задаче сегментации митохондрий электронной микроскопии мозга

К.Е. Яшин¹, Н.А. Соколов¹, Е.П. Васильев¹, А.А. Гетманская¹

¹ Нижегородский государственный университет им. Н.И. Лобачевского, пр. Гагарина, 23, Нижний Новгород, 603022, Россия

Аннотация

В данной работе представлена адаптация диффузионной нейросети для генерации размеченного синтетического набора данных электронной микроскопии мозга. Была обучена модель, способная генерировать изображения и одновременно разметку для них, что является несомненным преимуществом выбранного подхода.

С помощью обученной диффузионной модели был сгенерирован набор изображений с разметкой для них. Синтетические изображения визуально очень похожи на оригинальные, метрика FID схожести между синтетическим и оригинальным наборами данных равна 27,1. Упрощенная модель U-Net для сегментации, натренированная на смешанном наборе данных (оригинальные данные + синтетические) получила 0,856 по метрике Dice против 0,858 на оригинальном тренировочном наборе. То есть несмотря на хорошее качество синтетических данных их использование при обучении сегментационной сети не улучшило результаты сегментирования.

Ключевые слова

Расширение набора данных, электронная микроскопия, диффузионные нейросети.

Dataset Extension for Neural Networks Training for the Mitochondrial Segmentation Problem of the Brain Electron Microscopy

К.Е. Yashin¹, N.A. Sokolov¹, E.P. Vasiliev¹, A.A. Getmanskaya¹

¹ Lobachevsky State University of Nizhny Novgorod, Gagarina av. 23, N.Novgorod, 603022, Russian Federation

Abstract

This paper presents the adaptation of a diffusion neural network to generate a labeled synthetic dataset of electron microscopy of the brain. A model was trained can generate images and markup for them at the same time, which is an undoubted advantage of the chosen approach.

Using the trained model, a set of labeled images was generated. The synthetic images are visually very similar to the original ones, the FID similarity metric between the synthetic and original datasets is 27.1. A simplified U-Net segmentation model trained on a mixed data set (original data + synthetic data) obtained a Dice score of 0.856 versus 0.858 on the original training set. Despite the good quality of synthetic data, their use in training the segmentation network does not improve the segmentation results.

Keywords

Dataset extension, electron microscopy, diffusion neural networks.

ГрафиКон 2023: 33-я Международная конференция по компьютерной графике и машинному зрению, 19-21 сентября 2023 г., Институт проблем управления им. В.А. Трапезникова Российской академии наук, г. Москва, Россия

EMAIL: nikcsokolof@mail.ru (Н.А. Соколов); evgeny.vasiliev@itmm.unn.ru (Е.П. Васильев); getmanskaya.alexandra@gmail.com (А.А. Гетманская)

ORCID: 0000-0002-9055-2225 (Н.А. Соколов); 0000-0002-7949-1919 (Е.П. Васильев); 0000-0003-3533-1734 (А.А. Гетманская)



© 2023 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

1. Введение

Применение методов искусственного интеллекта для задач обработки данных электронной микроскопии затрудняется тем, что в открытом доступе находится очень малое количество размеченных данных, которые можно использовать для обучения глубокой нейронной сети. Разметка данных электронной микроскопии очень длительная и трудоемкая операция. Разметка одного эксперимента может достигать до полугода ручной работы.

Аугментация [1, 2] — это признанное эффективное решение для расширения обучающего датасета. Однако обычные способы аугментации достаточно простые — поворот, масштабирование, отражение, это значит, что взаимное расположение компарментов не изменяется. С развитием нейросетей, появилась возможность применять модели глубокого обучения для создания синтетических изображений.

Одним из популярных решений была генеративно-состязательная сеть (GAN) [3, 4, 5, 6] для увеличения данных о классах мало представленных в датасете. Используя GAN для создания данных [7] можно получить изображения, практически не отличимые от реальных, однако для обучения этих нейросетей необходим большой объем данных и обучение может быть нестабильно [8].

Использование вариационных автоэнкодеров (VAE) [9] позволяет обучаться на неразмеченных данных, однако качество генерации будет хуже реальных данных.

Нейродиффузионные модели (Neural Diffusion (ND)) являются передовым подходом в области искусственного интеллекта, который моделирует диффузию информации в нейронных сетях, для получения реалистичного результата. Они определяют марковскую цепь шагов диффузии, чтобы медленно добавлять случайный шум к данным, а затем учатся обращать процесс диффузии вспять для создания желаемых выборок данных из шума. Этот метод широко используется в генеративном моделировании, синтезе текстур и восстановлении изображений [10]. Недавно вероятностная модель диффузионного шумоподавления (DDPM) [11] привлекла большое внимание благодаря своей превосходному качеству генерации синтетических изображений [12, 13, 14]. Согласно обзорной статье о применении диффузионных моделей в медицине [15], этот подход за последние несколько лет набирает популярность.

Основными плюсами нейродиффузионных моделей можно назвать:

1. ND позволяет генерировать изображения высокого качества с богатыми деталями, реалистичными текстурами и плавными переходами. Данный подход способен устранять шумы и артефакты, что приводит к генерации чистых и естественных изображений.
2. Подход ND предоставляет контроль над процессом генерации изображений. Через параметры диффузии можно регулировать уровень детализации, размытия или степень сохранения исходной информации. Это позволяет пользователю настраивать генерируемые изображения в соответствии с конкретными требованиями.
3. ND может быть применен к различным типам изображений, включая фотографии, рисунки, текстуры и другие. Это делает его универсальным инструментом для генерации и обработки разнообразных визуальных данных.

Минусом нейродиффузионного подхода является его медлительность, особенно при повышенной сложности модели или при использовании большого числа итераций диффузии.

В данной работе будет подробно рассмотрен именно этот подход к генерации изображений, так как фактически перед нами стоит задача генерации текстуры, для которой диффузионные нейронные сети подходят в качестве возможного решения. Основная мотивация для этой работы связана с недавними достижениями в области генеративных моделей. В данной статье будет рассмотрен пример генерации синтетических данных ЭМ-данных с разметкой митохондрий.

2. Наборы данных и методы генерации

2.1. Наборы данных

В работе мы опирались на набор данных EPFL, доступный по адресу <https://www.epfl.ch/labs/cvlab/data/data-em/>. Данные взяты из области гиппокампа CA1 головного мозга. Разрешение каждого вокселя примерно $5 \times 5 \times 5 \text{ nm}$. Тренировочный набор состоит из 165 фрагментов стека изображений размером 1024×768 пикселей. Данные оригинального набора содержат разметку только для митохондрий. Пример данных приведен на рисунке 1.

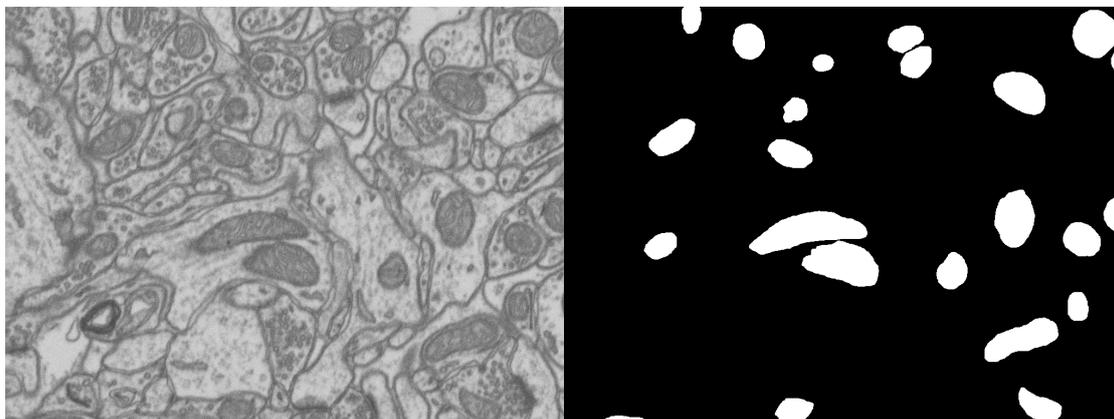


Рисунок 1 – Изображение электронной микроскопии мозга мыши и разметка митохондрий из набора данных EPFL

2.2. Диффузионные модели

Если взять изображение и начать накладывать на него гауссов шум, то после достаточного числа повторений наложения шума, из исходного изображения получится образец чистого шума. Главная идея, лежащая в основе диффузионных моделей – научиться отменять вышеописанный процесс, то есть начиная с шумового изображения постепенно удалять шум и в итоге получить понятное изображение. Исходя из этого можно сделать вывод, что архитектура сети должна удовлетворять следующему требованию – вход и выход должны иметь одинаковые размеры. Таким требованиям удовлетворяет, например, стандартная архитектура U-Net. Такой подход добился успеха в области генерации изображений, а модели, использующие данный метод, начинают конкурировать и даже превосходить другие виды генеративных моделей. Например, такие модели уже превосходят генеративно-сопоставительные сети, известные как GAN, по показателям качества восприятия.

3. Эксперименты

Идея обучения модели, способной генерировать сразу и изображение, и разметку для него, состоит в том, что мы можем создать новую тренировочную выборку, где каждое изображение содержит информацию сразу и об исходном изображении и о разметке этого исходного изображения. В качестве исходных данных был использован датасет EPFL [16]. Новый датасет был сгенерирован таким образом:

1. Каждая пара изображение-разметка из исходной обучающей выборки была помещена в новое изображение, где в каналы R и G попадает информация об интенсивности пикселей исходного изображения, а в канал B попадает информация об интенсивности пикселей разметки исходного изображения. Это сделано чтобы хранить тренировочные данные и получать новые в формате png изображений,
2. Изображение разрезается на части размером 256×256 пикселей. Для того, чтобы во время предварительной обработки датасета перед обучением модели не приходилось

производить масштабирование изображения до заданных параметров входа и выхода сети, которое ведет к потере информации об исходном изображении и об его деталях, а также к размытию изображений.

После всех вышеописанных преобразований была обучена модель U-Net со следующими гиперпараметрами:

1. Размер входного и выходного изображения – 256x256 пикселей
2. Размер батча обучающей выборки – 16
3. Количество эпох – 40
4. Скорость обучения для оптимизатора – 0.0004
5. Количество шагов денойзинга – 700

На рисунках 2 и 3 видно, что сгенерированные изображения имеют хорошую детализацию, похожи на реальный снимок электронной микроскопии и на них явно присутствуют объекты класса митохондрий.

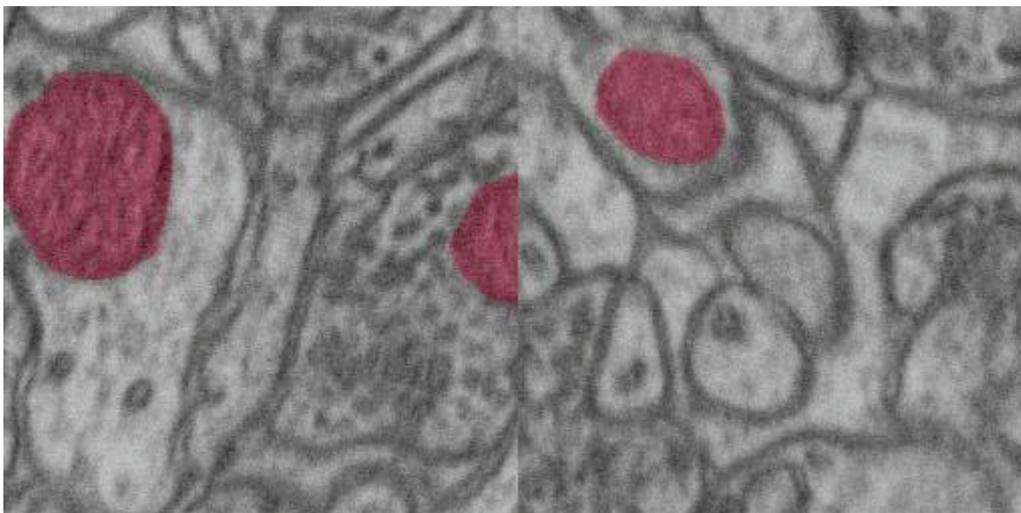


Рисунок 2 – Два сгенерированных изображения. Красным цветом выделены митохондрии

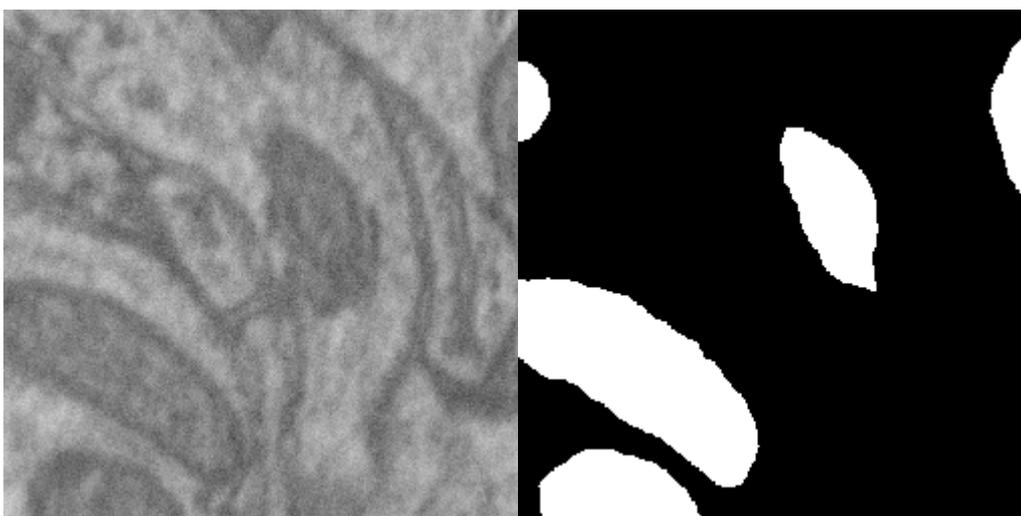


Рисунок 3 – Полученные пары изображение-разметка

В качестве финального эксперимента было решено увеличить количество эпох до 250, чтобы посмотреть на поведение при потенциальном переобучении. Изображения сохранялись каждые 10 эпох, чтобы проследить за динамикой. В итоге, примерно после сотой эпохи, на сгенерированных изображениях практически исчезли целые митохондрии, остались только их куски на границах изображения. Возможно это связано это с тем, что в результате нарезки

исходного датасета, получилось много изображений, где митохондрия присутствует не целиком. Результат генерации после 250 эпох представлен на рисунке 4.

В связи с этим можно сделать вывод, что наблюдается эффект переобучения модели, а оптимальное количество эпох для обучения модели – 40-70.

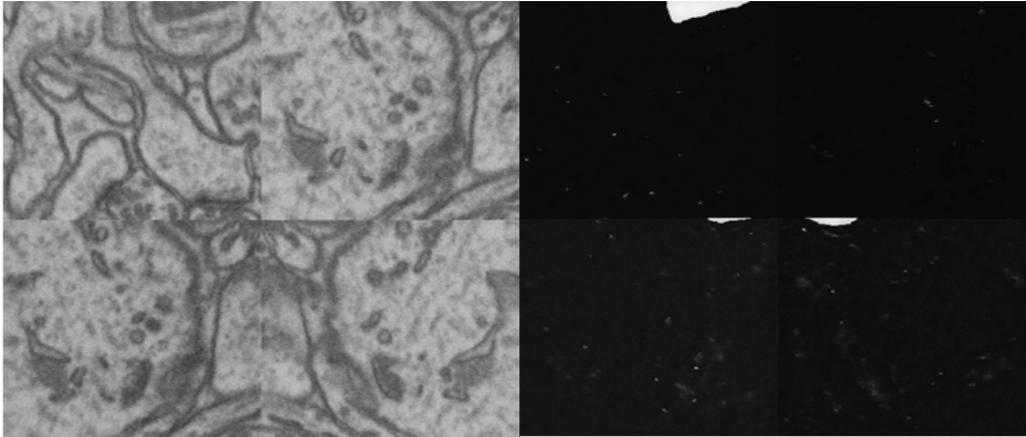


Рисунок 4 – Результат генерации после 250 эпохи

4. Анализ результатов

После этого с использованием модели было сгенерировано 1000 изображений. Анализировать наборы данных изображений предлагается с помощью вычисления метрики FID (Fréchet inception distance) – это мера сходства между двумя наборами данных. В общем случае эта метрика вычисляется по следующей формуле:

$$FID(\mu_r, \Sigma_r, \mu_g, \Sigma_g) = \|\mu_r - \mu_g\|^2 + \text{tr}(\Sigma_r + \Sigma_g - 2\sqrt{\Sigma_r \Sigma_g}) \quad (1)$$

где μ_r и μ_g – средние значения реального и сгенерированного наборов данных, Σ_r и Σ_g – дисперсия реального и сгенерированного набора данных, tr – обозначает операцию суммирования элементов матрицы на главной диагонали

Наши одноканальные изображения можно свести к простому одномерному случаю:

$$FID(\mu_r, \Sigma_r, \mu_g, \Sigma_g) = (\mu_r - \mu_g)^2 + (\sqrt{\Sigma_r} - \sqrt{\Sigma_g})^2 \quad (2)$$

Чем меньше метрики FID, тем более схожи наборы данных. Если метрика FID равна 0, то наборы должны быть идентичны.

Мы вычислили значение метрики для изображений и отдельно только для областей митохондрий.

Таблица 1 – Результаты качества генерации синтетического набора данных

Метрика/Класс	Слой	Митохондрии
FID	27.1	15.41

Для визуализации статистики по наборам данных мы построили гистограммы, представленные на рисунках 5 и 6. Гистограммы реальных и синтетических наборов данных действительно очень схожи, но диапазон часто встречающихся значений пикселей у реального изображения шире. По численным значениям метрик сравнения гистограмм можно сделать вывод о том, что модель генерирует изображения в основном похожие на исходную выборку, но есть и неудачные примеры генерации. Примеры удачной и неудачных генераций приведены на рисунках 7 и 8.

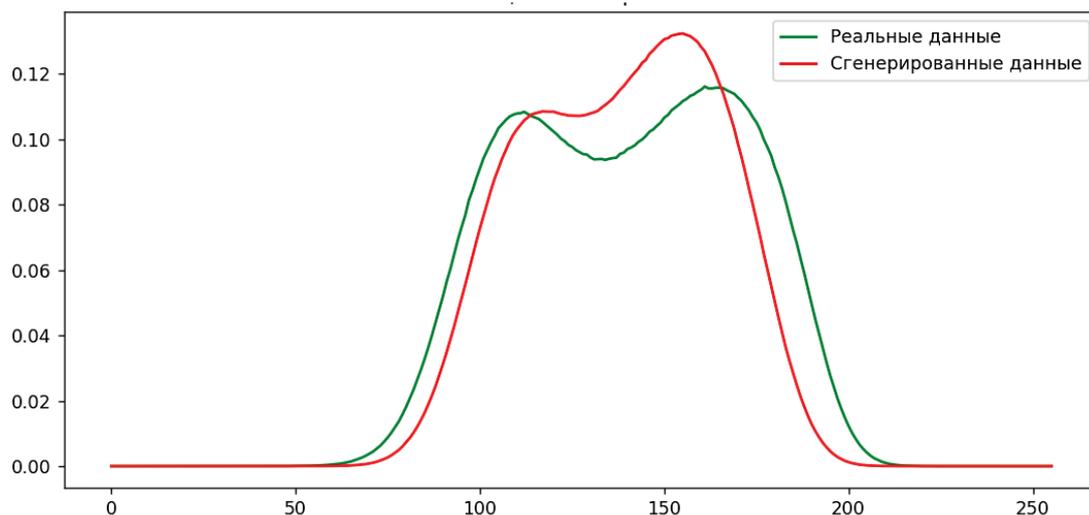


Рисунок 5 – Гистограммы наборов данных. (гистограмма исходного набора данных синим цветом, гистограмма синтетического набора данных красным цветом)

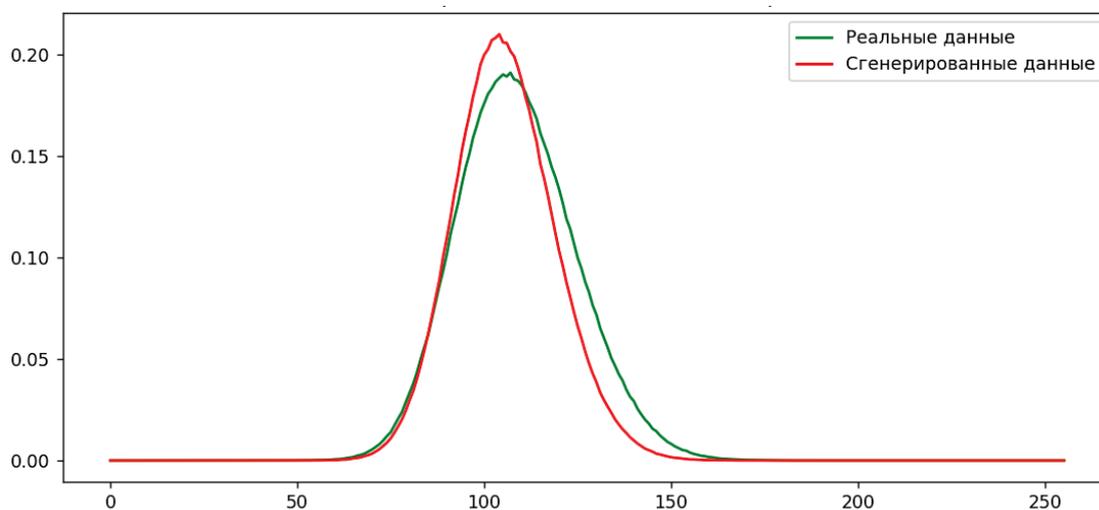


Рисунок 6 – Гистограммы митохондрий (из исходного набора данных – зеленый цвет, из сгенерированного красный)

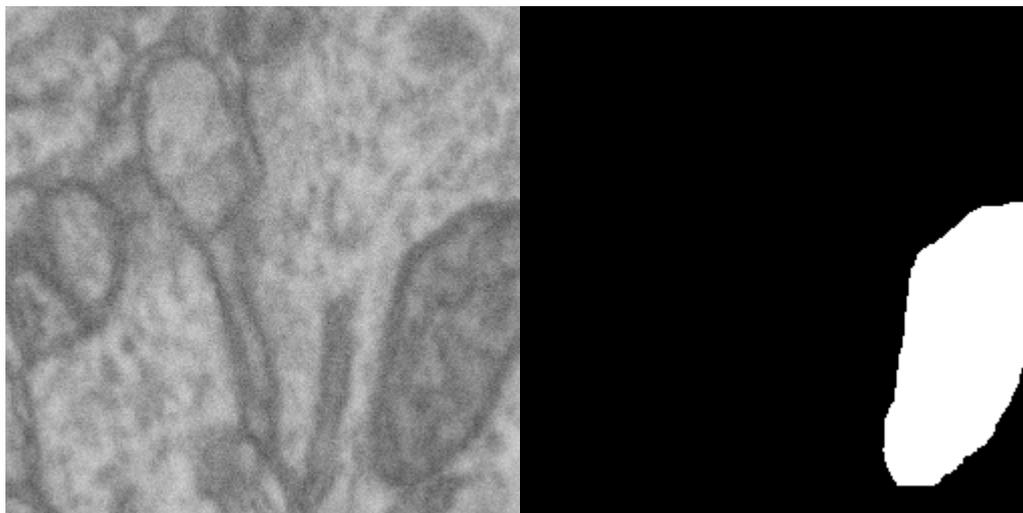


Рисунок 7 – Пример удачной генерации

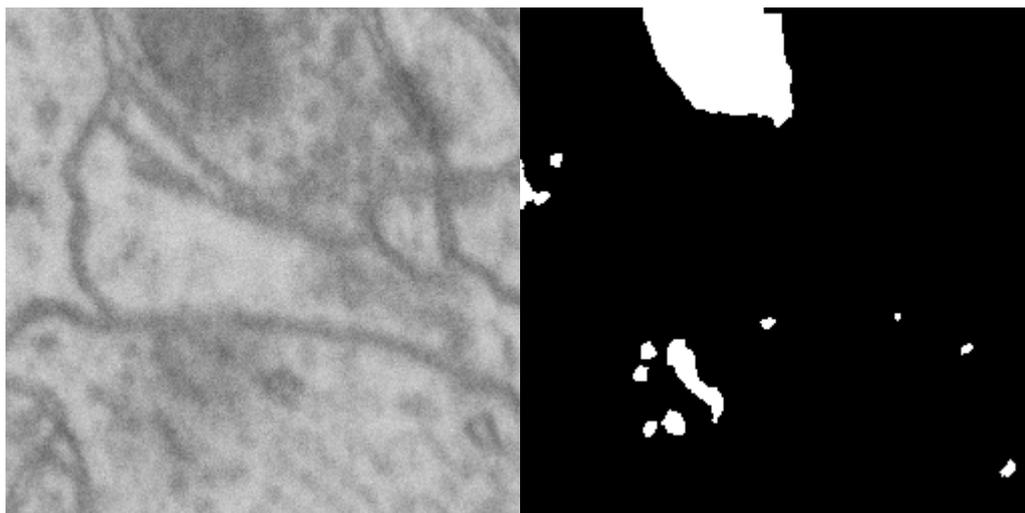


Рисунок 8 – Пример неудачной генерации

5. Обучение сегментационной нейросети на наборе данных с использованием синтетических изображений

В нашей прошлой работе [17] мы использовали упрощенную модель U-Net для сегментации. Для сегментации митохондрий она показала результаты лучше, чем классическая U-Net: 0.938 против 0.93. Облегченная модель содержит в 15.7 раз меньше параметров и занимает 24 МБ оперативной памяти, вместо 364 МБ при использовании оригинальной модели. Архитектура упрощенной модели показана на рисунке 9.

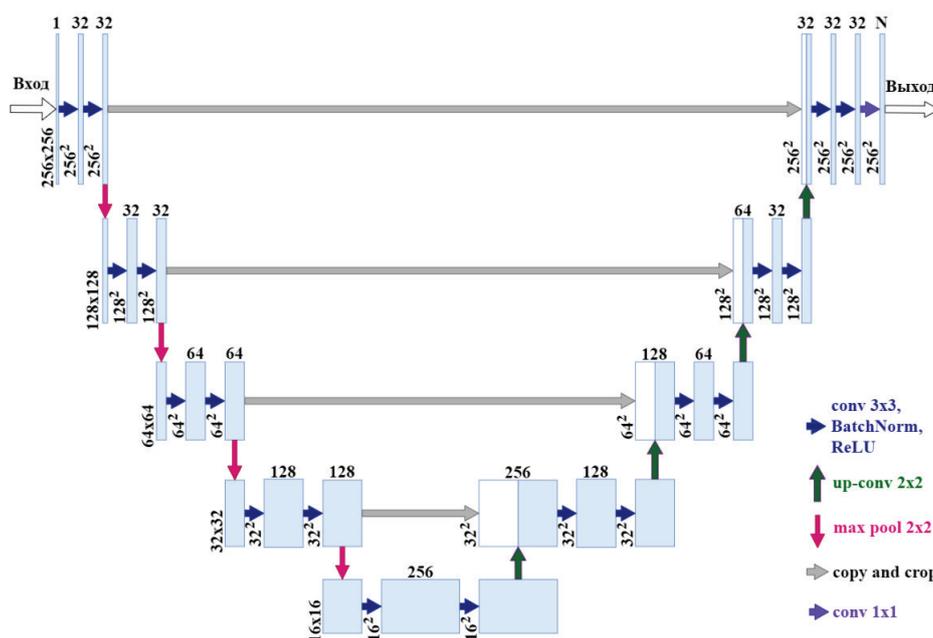


Рисунок 9 – Архитектура облегченной модели U-Net

Модель обучалась со следующими гиперпараметрами:

1. Размер входного и выходного изображения – 256x256 пикселей
2. Размер батча обучающей выборки – 7
3. Количество эпох – 200
4. Оптимизатор Adam, с переменной скоростью обучения от 1×10^{-4} до 1×10^{-6}
5. Количество семплов из оригинального датасета – 1722

6. Количество семплов из синтетического датасета - 1000

Таблица 2 – Результаты обучения для набора данных EPFL по метрике Dice с синтетическим расширением датасета и без

Набор данных	разметка EPFL	разметка Lucchi++
Оригинальный набор данных	0,858	0,943
Оригинальный набор данных + сгенерированные изображения	0,856	0,939

В наших экспериментах использование дополнительного синтетического датасета не привело к улучшению сегментации. Незначительное ухудшение качества может быть вызвано неудачными примерами генерации.

7. Заключение

В данной работе был опробован алгоритм генерации синтетического набора данных, аналогичного набору данных EPFL с разметкой митохондрий. Для генерации изображений был использован подход, основанный на диффузионных нейронных сетях. Была обучена модель, способная генерировать изображения электронной микроскопии одновременно с разметкой для них, что является несомненным преимуществом выбранного подхода.

Сгенерированные данные показывают высокую визуальную схожесть с реальными данными, однако по гистограммам наборов данных заметны небольшие отличия. Несмотря на хорошее качество синтетических данных, их использование при обучении сегментационной сети не улучшило результаты сегментирования. Возможно постобработка и отбор семплов может улучшить качество тренировочного синтетического набора данных.

8. Благодарности

Работа поддержана Минобрнауки, соглашение 075-02-2023-945 от 16.02.2023 г.

9. Список источников

- [1] Chapelle O., Weston J., Bottou L., Vapnik V., Vicinal risk minimization. *Advances in neural information processing systems* 13 (2000).
- [2] Simard P.Y., LeCun Y.A., Denker J.S., Victorri B., Transformation invariance in pattern recognition—tangent distance and tangent propagation. In: *Neural networks: tricks of the trade*, pp. 239–274. Springer (2002).
- [3] Gong X., Chen S., Zhang B., Doermann D., Style consistent image generation for nuclei instance segmentation. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. pp. 3994–4003 (2021).
- [4] Hou L., Agarwal A., Samaras D., Kurc T.M., Gupta R.R., Saltz J.H., Robust histopathology image analysis: To label or to synthesize? In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 8533–8542 (2019).
- [5] Lin Y., Wang Z., Cheng K.T., Chen H., InsMix: Towards realistic generative data augmentation for nuclei instance segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer (2022).
- [6] Wang H., Xian M., Vakanski A., Shareef B., Sian: Style-guided instance-adaptive normalization for multi-organ histopathology image synthesis. *arXiv preprint arXiv:2209.02412* (2022).
- [7] Goodfellow I.J., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., et al., Generative Adversarial Networks. *14062661 2014:2672–80*. URL: <https://arxiv.org/abs/1406.2661>.

- [8] Dhariwal P., Nichol A., Diffusion models beat gans on image synthesis. *Advances in Neural Information Processing Systems* 34, 8780–8794 (2021).
- [9] Kingma D. P., Welling M., Auto-encoding variational bayes, in 2nd International Conference on Learning Representations, ICLR 2014. URL: <https://arxiv.org/abs/1312.6114>.
- [10] Song Y., Sohl-Dickstein J., Kingma D.P., Kumar A., Ermon S., Poole B., Score-based generative modeling through stochastic differential equations, 2020, URL: <https://arxiv.org/abs/2011.13456>.
- [11] Ho J., Jain A., Abbeel P., Denoising diffusion probabilistic models. In: Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., Lin, H. (eds.) *Advances in Neural Information Processing Systems*. vol. 33, pp. 6840–6851. Curran Associates, Inc. (2020), <https://proceedings.neurips.cc/paper/2020/file/4c5bcfec8584af0d967f1ab10179ca4b-Paper.pdf>
- [12] Dhariwal P., Nichol A., Diffusion models beat gans on image synthesis. *Advances in Neural Information Processing Systems* 34, 8780–8794 (2021).
- [13] Nichol A., Dhariwal P., Ramesh A., Shyam P., Mishkin P., McGrew B., Sutskever I., Chen M., Glide: Towards photorealistic image generation and editing with textguided diffusion models. arXiv preprint arXiv:2112.10741 (2021).
- [14] Wolleb J., Bieder F., Sandkühler R., Cattin P.C., Diffusion models for medical anomaly detection. In: *Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part VIII*. pp. 35–45. Springer (2022).
- [15] Kazerouni A., Aghdam E. K., Heidari M., Azad R., Fayyaz M., Hacihaliloglu I. and Merhof D., Diffusion models for medical image analysis: A comprehensive survey, arXiv preprint arXiv:2211.07804, 2022.
- [16] EPFL dataset. URL: <https://www.epfl.ch/labs/cvlab/data/data-em/> (дата обращения 12.01.2023).
- [17] Sokolov N.A., Vasiliev E.P., Getmanskaya A.A., Generation And Study Of The Synthetic Brain Electron Microscopy Dataset For Segmentation Purpose. *Programming and Computer Software* (in printing).