

ИССЛЕДОВАНИЕ И РАЗРАБОТКА ЭКСПЕРИМЕНТАЛЬНЫХ СРЕДСТВ ГЕНЕРАЦИИ КТ-ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ ДИФФУЗИОННЫХ НЕЙРОННЫХ СЕТЕЙ

М. С. Жидович

maksim.zhidovich@gmail.com;

Научный руководитель — В. А. Ковалев, кандидат технических наук

Одной из главных проблем при использовании искусственного интеллекта в медицине является недостаточное количество данных для обучения моделей. Современные подходы в машинном обучении требуют сотни тысяч анализов и изображений для достижения высокого качества. В данный момент не представляется возможным подготовить такое количество данных. Альтернативным решением может стать генерация подобных данных с помощью нейронных сетей. В данной работе описана разработка экспериментальных средств генерации КТ-изображений с помощью диффузионных нейронных сетей, а также сравнение с ранее полученными результатами.

Ключевые слова: генерация медицинских изображений; КТ-изображения; диффузионные нейронные сети; DDPM; нейронные сети.

ДИФФУЗИОННЫЕ МОДЕЛИ

Идея моделей данного типа состоит в определении диффузии как обратимого стохастического процесса, в котором структура распределения данных систематически и медленно разрушается посредством итеративного процесса прямого распространения. Затем изучается процесс обратного распространения, который восстанавливает структуру данных. В случае с изображениями прямой итеративный процесс представляет собой добавление шума по известному гауссовскому закону $q(x)$, обратный - удаление шума по некоторому неизвестному обратному гауссовскому закону $p(x)$. Этот закон предлагается аппроксимировать с помощью нейронной сети.

Таким образом, опишем процесс генерации изображений с помощью диффузионной модели: структура исходного изображения постепенно разрушается путем добавления шума, а затем восстанавливается с использованием нейросети. Повторяя описанные действия достаточное количество раз, модель обучается оценивать распределение исходных данных. Затем на вход подаётся случайный шум, который с помощью обратного процесса преобразуется в новое изображение. Отметим тот факт, что итеративность процесса позволяет проводить контролируемое обучение на каждом временном шаге.

Существует несколько основных видов диффузионных моделей. Основное различие между ними заключается в разной организации

итерационного процесса. В данной работе использовалась DDPM (Denoising Diffusion Probabilistic Model) [1]. Описанный ранее процесс здесь представляется в виде цепи Маркова с конечным количеством шагов. Такое решение негативно влияет на время работы, однако позволяет существенно улучшить качество генерируемых образцов. Восстановление изображения сводится к предсказанию добавленного шума на временном шаге t по изображению с шага $t+1$. Размерность шума совпадает с размерностью изображения. Как следствие, в качестве модели использовался UNet, реализованный в соответствии с [2].

ИСХОДНЫЕ ДАННЫЕ

Исходный датасет состоял из 100000 изображений. Было принято решение сконцентрироваться на одном классе – КТ-срезы на уровне сердца. Исходный размер изображений: 512×512 . Затем он был уменьшен с помощью бикубической интерполяции. Данный вид интерполяции часто используется для масштабирования изображений и видео, так как позволяет сохранять детали намного лучше, чем простой билинейный алгоритм. Размеры изображений, используемые в ходе экспериментов: 64×64 , 128×128 .



Рис. 1. Примеры исходных изображений

ОБУЧЕНИЕ МОДЕЛИ. РЕЗУЛЬТАТЫ

Обучение нейросети для генерации изображений в разрешении 128×128 происходило на сервере DeepLab(Tesla V100 16Gb) в течение 40 часов. Количество шагов в цепи было равно 1000. Такой выбор позволяет соблюдать баланс между качеством генерации и её временем. В качестве планировщика дисперсий использовалась линейная сетка. Далее рассмотрим примеры генерации с помощью DDPM:



Рис. 2. Примеры генерации с помощью DDPM

Ранее в задаче генерации КТ-изображений были получены определенные результаты с помощью генеративно-сопоставительных сетей (Рис. 3). Визуальная оценка показывала, что качество изображений нельзя в полной мере назвать удовлетворительным, также существовала проблема малого разнообразия среди генерируемых образцов. Использование диффузионных моделей позволило улучшить качество генерации. Изображения стали более четкими, детализированными, форма грудной клетки стала более правильной с точки зрения анатомии.



Рис. 3. Примеры генерации с помощью DCGAN

Таким образом, в ходе работы были исследованы принципы генерации изображений на основе диффузионных нейронных сетей, а также способы обучения моделей данного типа. Результатом работы стала обученная нейронная сеть, которая может быть использована для генерации КТ-изображений грудной клетки. Визуальная оценка показала, что полученные изображения имеют достаточное качество для использования в других задачах обработки медицинских изображений.

Библиографические ссылки

1. Ho J., Jain A., Abbeel P. Denoising diffusion probabilistic models //Advances in Neural Information Processing Systems. 2020. Т. 33. С. 6840-6851.
2. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation //Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. – Springer International Publishing, 2015. С. 234-241.