

# Аугментация наборов изображений для обучения нейронных сетей при решении задач семантической сегментации

И.А. Ложкин, М.Е. Дунаев, К.С. Зайцев, А.А. Гармаш

**Аннотация.** Целью настоящей работы является исследование эффективности методов аугментации (дополнения) наборов изображений при их нехватке в обучающей выборке нейронных сетей для решения задач семантической сегментации. Для этого были рассмотрены основные группы методов аугментации и исследована их эффективность при решении задач семантической сегментации медицинских изображений. Для обучения, тестирования и валидации использовались две глубокие архитектуры DeeplabV3+ с энкодером EfficientNetB6. Целевыми метриками, по которым проводились сравнения качества семантической сегментации изображений, были выбраны Intersection over Union и коэффициент Дайса, что позволило определить модели с наилучшими показателями предсказания. Полученные результаты подтвердили эффективность предложенного набора методов аугментации. Итогом работы стало создание эффективного подхода к аугментации наборов медицинских изображений для решения задачи семантической сегментации.

**Ключевые слова** — глубокое обучение, аугментация, семантическая сегментация.

## 1. ВВЕДЕНИЕ

Методы машинного обучения активно проникают в самые разные области человеческой деятельности для решения широкого спектра задач. Медицинская сфера не является исключением. С помощью компьютерного зрения снимков рентгенологических и ультразвуковых исследований (УЗИ), магнитно-резонансной томографии (МРТ) и др. решаются задачи классификации и сегментации медицинских изображений, локализация образований [1-3].

Узлы щитовидной железы являются одним из наиболее часто диагностируемых новообразований у людей. По данным эпидемиологов, их распространенность в популяции достигает 50% [4]. Снимки УЗИ позволяют локализовать узлы и определить класс заболевания по одной из шкал

Статья получена 28 октября 2022.

Ложкин Илья Александрович, Национальный Исследовательский Ядерный Университет МИФИ, магистрант, Lozkin.ilya@gmail.com

Дунаев Максим Евгеньевич, Национальный Исследовательский Ядерный Университет МИФИ, аспирант, ст. инженер ИИКС, преподаватель кафедры №132, MEDunaev@mephi.ru

Зайцев Константин Сергеевич, Национальный Исследовательский Ядерный Университет МИФИ, профессор, KSZajtsev@mephi.ru

Гармаш Александр Александрович, Национальный Исследовательский Ядерный Университет МИФИ, директор инженерно-физического института биомедицины, AAGarmash@mephi.ru

классификации, например, EU-TIRADS [5], т.е. дать количественную оценку новообразования. Для решения указанных задач сегодня преимущественно используются нейросетевые подходы, в которых на вход модели подается исходное изображение или кинопетля (последовательность изображений) продольного или поперечного медицинского исследования, а на выходе получается класс заболевания, а также сегментированное или локализованное изображение узла [6, 7].

Для успешного построения хорошо обобщающих глубоких нейросетевых моделей необходимо большое количество достоверных данных, чтобы избежать переобучения и “запоминания” обучающих наборов [8, 9]. Подготовка таких данных медицинскими учреждениями еще более усложняется значительными затратами времени на разметку границ снимков обучающего набора. Кроме того, наборы изображений зачастую являются несбалансированными по классам. Проблема отсутствия необходимого объема сбалансированных данных для обучения нейронных сетей в медицине присутствует во многих клиниках [10].

Одним из способов решения описанной проблемы является применение методов дополнения (аугментации) обучающего набора данных новыми синтезированными изображениями. В настоящей статье как раз и исследуется эффективность таких методов применительно к медицинским изображениям.

## 2. ПОСТАНОВКА ЗАДАЧИ

Формально задача аугментации набора изображений выглядит так. Имеется исходный размеченный набор из  $N$  медицинских изображений  $X = \{x_1, x_2, \dots, x_N\}$  с масками  $Y = \{y_1, y_2, \dots, y_N\}$ . Этот набор делится на обучающую и тестовую выборки, например, в соотношении 80% и 20% соответственно.

Рассматривается обучение некоторой нейросетевой модели  $M$ , используемой для решения задачи семантической сегментации изображений [11], с параметрами обучения  $P$  в течение  $N_e$  эпох на наборе из изображений  $X$  с масками  $Y$ . Качество сегментации обученной модели оценивается некоторой метрикой качества  $M_t$  на тестовой выборке изображений  $X_{test}$  с масками  $Y_{test}$ . Обозначим наилучшее качество сегментации модели на тесте для обученной сети как:

$$Q(M(P, N_e), X, Y) = \min(Mt_i),$$

где  $Mt_i$  — значение метрики качества сегментации модели на тесте на  $i$ -ой эпохе ( $i$  - натуральное число).

Необходимо задать такое множество методов аугментации  $F$ , чтобы

$$Q(M(P, N_e), F[X], Y) > Q(M(P, N_e), X, Y),$$

т.е. необходимо сформировать такое множество методов аугментации наборов медицинских изображений, которое улучшило бы значение выбранной метрики качества решения задачи семантической сегментации моделью.

### 3. АНАЛИЗ ПОДХОДОВ К АУГМЕНТАЦИИ ИЗОБРАЖЕНИЙ

Методы аугментации изображений можно классифицировать по нескольким признакам. Исходя из анализа [8-9, 12], по типу вносимых изменений в наборы данных выделяют методы геометрических (среди которых часто упоминаются аффинные) преобразований, методы преобразований на уровне пикселей, методы создания искусственных данных с помощью генеративно-сопоставительных нейросетей.

Множество преобразований, применяемых к обучающей и тестовой выборкам, зачастую различаются. Для обобщения прогнозов к тестовым данным применяют также преобразования *test-time augmentation* (ТТА), суть которых в выполнении нескольких различных модификаций для каждого изображения [9, 13-14].

По частоте использования преобразования можно разделить на постоянные, т.е. такие, которые применяются ко всем изображениям заданного набора, и непостоянные — которые применяются с некоторой вероятностью или случайно из заданного множества преобразований [15].

Проанализируем особенности методов аугментации изображений по типу вносимых изменений.

#### 3.1. Методы геометрических преобразований

Геометрические преобразования являются наиболее часто используемыми подходами к аугментации наборов данных. Среди них выделяется аффинные преобразования, к которым относятся поворот, зеркальное отражение, перенос, сдвиг и масштабирование. Есть еще и другие не аффинные преобразования, так или иначе связанные с геометрическими изменениями в изображениях. Рассмотрим методы геометрических преобразований подробнее.

*Поворот (rotation)*. Это преобразование осуществляет поворот изображения на заданный угол или случайный угол из заданного диапазона. За этой операцией зачастую следует операция интерполяции, чтобы размер полученного изображения соответствовал размеру исходного.

*Зеркальное отражение (flip)*. Это преобразование выполняется вдоль одной или нескольких

указанных осей. В отличие от естественных изображений, где в зависимости от типа изображенного объекта вертикальные отражения могут не представлять интереса (в силу нереальности), для медицинских изображений отражение как по горизонтали, так и по вертикали позволяет повысить репрезентативность обучающих наборов данных [16].

*Перенос (translate, shift)*. Это преобразование выполняют перенос всего изображения на заданное количество пикселей или на заданный процент от длины или ширины изображения вдоль одной или нескольких указанных осей. За этой операцией может следовать операция интерполяции, чтобы размер полученного изображения соответствовал размеру исходного. Перенос, подобно повороту и зеркальному отображению, направлен на то, чтобы модель не фокусировалась на объектах, присутствующих в одной конкретной пространственной области, а видела пространственно-инвариантные объекты.

*Сдвиг (shear)*. Это преобразование смещает части изображения по указанной оси в разных направлениях. Например, при горизонтальном сдвиге в противоположных направлениях смещаются верхняя и нижняя части изображения, а при вертикальном сдвиге — левая и правая части. Ввиду того что изображение искажается, сдвиг осуществляется на малые расстояния.

*Масштабирование (scale, zoom)*. Эти преобразования масштабируют (увеличивают/уменьшают) изображение до заданного процента. В зависимости от используемой библиотеки, реализующей преобразование, итоговое изображение может сохранять новые размеры, а может по умолчанию приводиться к исходным значениям. В первом случае для приведения к требуемым размерам понадобится преобразование обрезаки. Во втором случае по умолчанию остается часть изображения, соответствующая исходным размерам изображения, а затем при необходимости следует операция интерполяции.

*Обрезка (crop, crop with padding)*. Это преобразование обрезает изображение в соответствии с указанными пикселями или процентами для обрезки. Для восстановления исходных размеров в зависимости от заданных параметров оставшееся изображение масштабируется или выполняется “заливка” области вокруг него (для соответствующих пикселей устанавливается значение “0” или рассчитанные посредством интерполяции значения).

Примеры геометрических методов аугментации изображений представлены на рис. 1.

#### 3.2. Методы преобразований на уровне пикселей

Особенностью методов преобразований на уровне пикселей является то, что такие методы не совершают геометрических преобразований над изображениями, а влияют на значения пикселей

либо локально в определенных областях, либо по обнуление некоторых пикселей изображений в

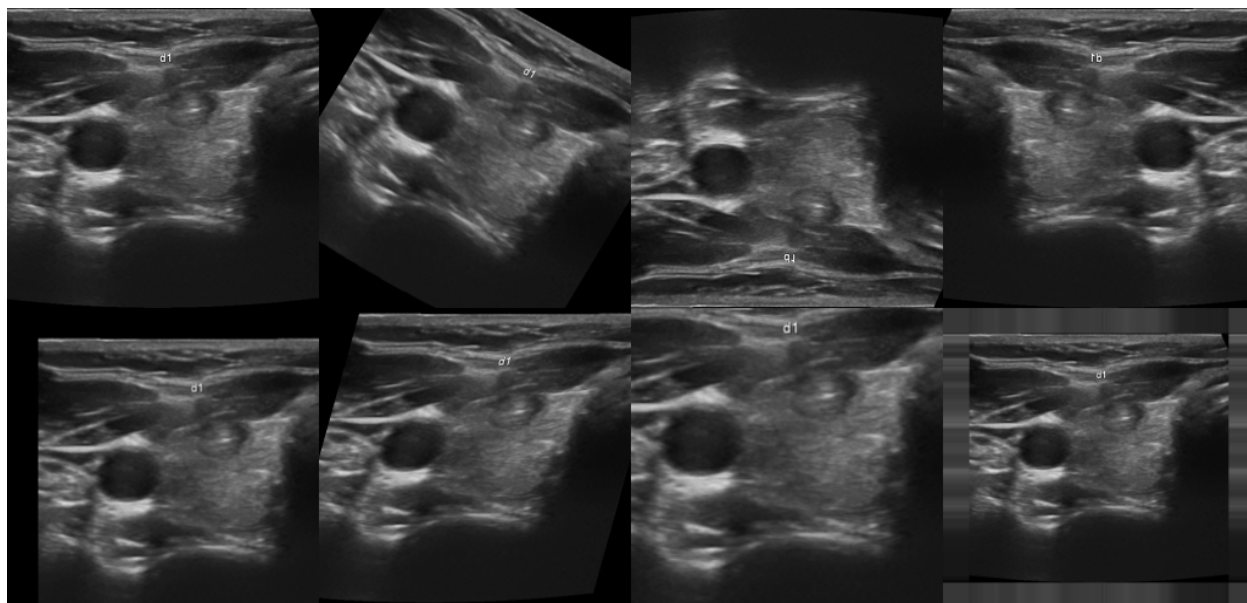


Рис. 1 — Геометрические методы аугментации изображений.

Слева-направо построчно: исходное изображение, поворот, зеркальное отражение вдоль вертикальной оси, зеркальное отражение вдоль горизонтальной оси, перенос, сдвиг, масштабирование, обрезка

всему изображению.

Такие преобразования способны быть полезными при анализе медицинских изображений потому, что наборы изображений для обучения получают с разного медицинского оборудования (установок, аппаратов, сканеров и т.п.), и, поэтому, такие изображения могут быть неоднородными по яркости, контрастности, резкости и иным параметрам.

Рассмотрим подробнее методы преобразований на уровне пикселей.

*Упругие преобразования (elastic transformations).* С помощью преобразований осуществляется перемещение пикселей изображения, как правило, в соответствии с заданным полем смещения. В [9] данный тип аугментации выделяют в отдельную категорию. Такие преобразования могут вносить значительные изменения в исходные данные, создавая эффект нахождения изображения в воде. Имеются работы, как подтверждающие положительный результат использования упругих преобразований (работа по сегментации магнитно-резонансной томографии сердца), так и работы, в которых такие преобразования ухудшили показатели модели (работа по определению опухоли головного мозга).

*Зашумление* представляет собой такое изменение значений пикселей изображения, которое ухудшает качество исходного изображения. Зашумление может быть разных типов, например: случайный шум (*random noise*), Гауссовский шум (*Gaussian noise*), Пуассоновский шум (*Poisson noise*), импульсный шум (*impulse noise, salt-and-pepper noise*). Также к данной категории можно отнести

соответствии с заданными условиями (*dropout*).

*Изменение яркости, насыщенности, контрастности.* Такие преобразования отвечают за регулирование яркости (*brightness*), насыщенности (*saturate*), контрастности (*contrast*) изображения. Как правило, устанавливается некоторый порог, в зависимости от которого меняются значения пикселей изображения. Так, контрастность увеличивается, когда темные пиксели изображения затемняются, а светлые — освещаются: контрастное изображение содержит большее количество черного и белого. Примеры специальных преобразований контрастности: линейный контраст (*linear contrast*), гамма-контраст (*gamma contrast*).

*Применение фильтров* используется для того, чтобы сделать изображение более четким, размыть или сгладить. Фильтрация выполняется с помощью свертки, сверточное ядро перемещается по изображению для изменения значения каждого пикселя на основе значений окружающих пикселей. Некоторые ядра способны обнаруживать и усиливать края объектов, найденных на изображении. Примеры фильтров: размытие по Гауссу (*Gaussian blur*), среднее размытие (*average blur*), медианный фильтр (*median blur*), двусторонний фильтр (*bilateral filter*), фильтр *sharpen* для изменения резкости.

Примеры методов аугментации изображений посредством преобразований на уровне пикселей представлены на рис. 2.

### 3.3. Методы создания искусственных данных

Наиболее часто упоминаемыми методами являются генеративно-состязательные нейросети (*GAN*), с помощью которых синтезируются

искусственные данные, подобные исходным [17, 18]. В таком подходе одна сеть генерирует правдоподобные изображения, а другая старается отличить сгенерированные изображения от реальных. Описанным образом может быть синтезировано большое количество новых данных, однако стоит учитывать, что генерация изображений сама по себе является достаточно сложной

файлами в формате tif. Подбор пациентов и снимков узловых образований проводились в рамках реализации проекта №22-15-00135 гранта Российского научного фонда.

Предобработка исходных данных включала:

- преобразование файлов УЗИ щитовидной железы и масок из формата tif в изображения формата PNG;

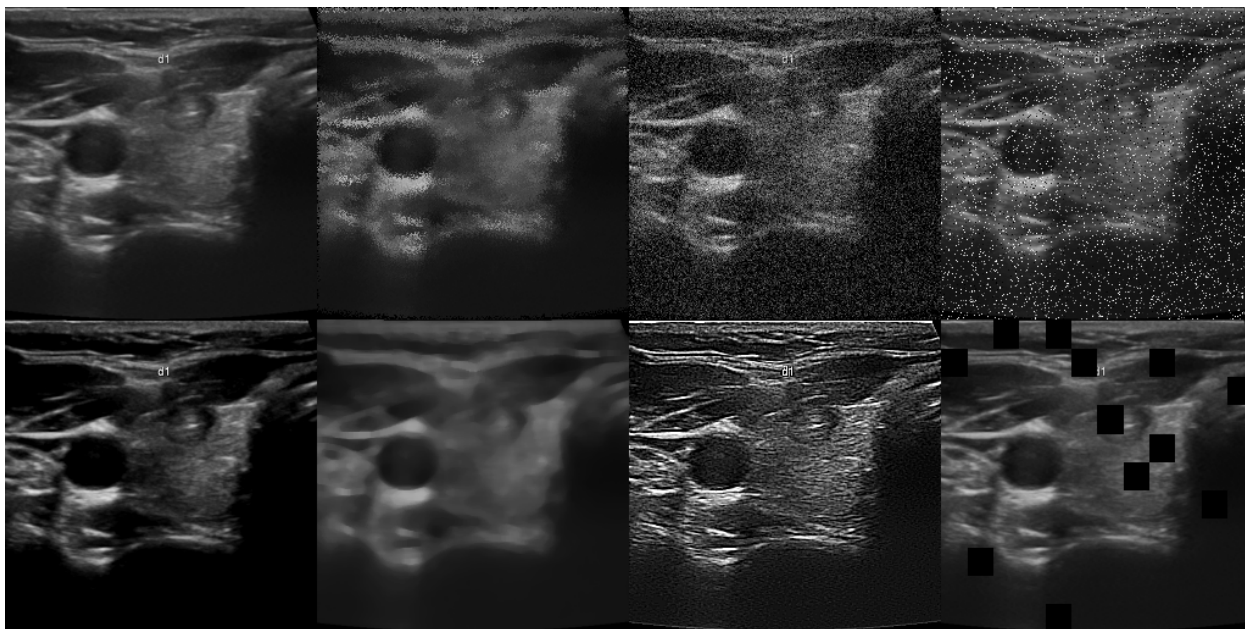


Рис. 2 — Методы аугментации изображений посредством преобразований на уровне пикселей. Слева-направо построчно: исходное изображение, упругое преобразование, Гауссовский шум, шум salt-and-pepper, линейный контраст, медианный фильтр, фильтр sharpen, преобразование dropout

вычислительной задачей.

#### 4. ИССЛЕДОВАНИЯ ПО ОТБОРУ МЕТОДОВ АУГМЕНТАЦИИ

##### 4.1. Предложенный подход к семантической сегментации изображений

Для решения задачи семантической сегментации изображений использовались две сети с одной и той же структурой энкодер-декодер, подробное описание подхода последовательного применения которых представлен в [7]. Обе сети имеют архитектуру DeeplabV3+ [19] с энкодером EfficientNetB6 [20].

На вход сетям подавались изображения в оттенках серого. Первая сеть сегментации обеспечивала “грубую” локализацию узловых образований щитовидной железы на изображениях размером 256x256 пикселей. На вход второй сети сегментации поступала только та область, в которой был “грубо” локализован узел, увеличенная до размеров 512x512 пикселей (region of interest, ROI). Назначением второй сети являлась точная локализация узла.

##### 4.2. Подготовка набора данных для экспериментов

Исходный набор данных состоял из снимков УЗИ щитовидной железы 80 пациентов в продольном (long) и поперечном (cross) срезах и размеченных масок. Снимки УЗИ и маски были представлены

- удаление текстовой информации;
- удаление черных нерелевантных областей;
- нормализация изображений;
- приведение изображений к оттенкам серого;
- изменение размера изображений и масок до 256x256 пикселей для подачи на вход первой сети сегментации;
- создание изображений и масок размером 512x512 пикселей, содержащих увеличенные области с узловыми образованиями щитовидной железы (ROI) для подачи на вход второй сети сегментации.

При преобразованиях из tif в PNG составляющие tif изображения брались с шагом 5 ввиду высокой схожести соседних изображений.

Итоговый набор данных — изображения в формате PNG: 730 изображений и масок продольного среза щитовидной железы, 894 изображения и маски поперечного среза щитовидной железы.

##### 4.3. Описание эксперимента

Проводилось обучение описанных в подразделе 4.1 сетей: отдельно сети 1 и сети 2, на продольных, поперечных и всех снимках одновременно, без аугментации, с простой аугментацией и со сложной аугментацией. Общее количество обученных моделей — 18.

Простая аугментация наборов изображений включала в себя основные методы геометрических преобразований: поворот, зеркальное отражение по горизонтальной и/или вертикальной оси, перенос, сдвиг, масштабирование, обрезка.

Сложная аугментация наборов изображений предусматривала преобразования не только обучающего набора, но и тестового. Аугментация обучающего набора включала в себя геометрические преобразования всего изображения (поворот, зеркальное отражение по горизонтальной и/или вертикальной оси, перенос, сдвиг, масштабирование, обрезка), локальное аффинное преобразование (масштабирование), преобразования на уровне пикселей (упругие преобразования, Гауссовский шум, линейный контраст, изменение резкости, размытие по Гауссу, среднее размытие, медианный фильтр, обнуление пикселей). Описанные преобразования применялись не все одновременно, а использовались с заданной

вероятностью, в случайном порядке, на выбор по одному или нескольким методам для однотипных. Аугментация тестового набора изображений представляла собой ТТА, включающую поворот, зеркальные отражения по горизонтальной и вертикальной осям.

#### 4.4. Метрики качества

Для оценки качества сегментации использовались метрики IoU (Intersection over Union) и коэффициент Дайса (Dice coefficient, DC), формулы вычисления которых для предсказанного изображения представлены ниже.

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (1)$$

$$DC = \frac{2|A \cap B|}{|A| + |B|}, \text{ где} \quad (2)$$

A — множество пикселей узла на маске, B — множество предсказанных пикселей узла. Для нескольких изображений рассчитывались

Таблица 1 — Графики значений функции потерь на обучении.

| Датасет           | Сеть 1 | Сеть 2 |
|-------------------|--------|--------|
| Продольные снимки |        |        |
| Поперечные снимки |        |        |
| Все снимки        |        |        |

средние значения соответствующих метрик.

#### 4.5. Результаты эксперимента

Графики значений функции потерь на обучении представлены в таблице 1.

На графиках в таблице 1 видно, что усложнение аугментации обучающего набора изображений при равных параметрах обучения приводит к увеличению значений функции потерь на этапе обучения. Это свидетельствует о том, что методы преобразований вносят разнообразие в исходные наборы данных.

Графики значений метрики качества сегментации IoU на тесте представлены в таблице 2.

Из таблицы 2 видно, что для сети 1, на вход которой поступали полные изображения, усложнение аугментации обучающего и тестового наборов улучшает показатели IoU, что свидетельствует об увеличении обобщающих способностей моделей. Однако для сети 2, на вход

которой поступали ROI, в которых узел занимает большую часть изображения, усложнение аугментации оказывает негативный эффект.

Графики значений метрики качества сегментации DC на тесте представлены в таблице 3.

Выводы по данной таблице 3 аналогичны ранее описанным выводам по анализу результатов IoU на тесте.

В таблице 4 представлены значения метрик качества сегментации, рассчитанные по результатам последовательного применения двух сетей ( $IoU_{ср.}$ ,  $DC_{ср.}$ ), в сравнении со значениями метрик качества сегментации только первой сети ( $IoU_{сети 1}$ ,  $DC_{сети 1}$ ).

Из анализа таблицы 4 можно сделать вывод, что последовательное применение двух сетей на качество сегментации влияет незначительно: либо не изменяет показателей метрик качества, либо улучшает их лишь на 1%. Это можно связать с тем, что первая сеть сегментации обучалась 100 эпох, в то время, как обучение второй сети проводилось

Таблица 2 — Графики значений IoU на тесте

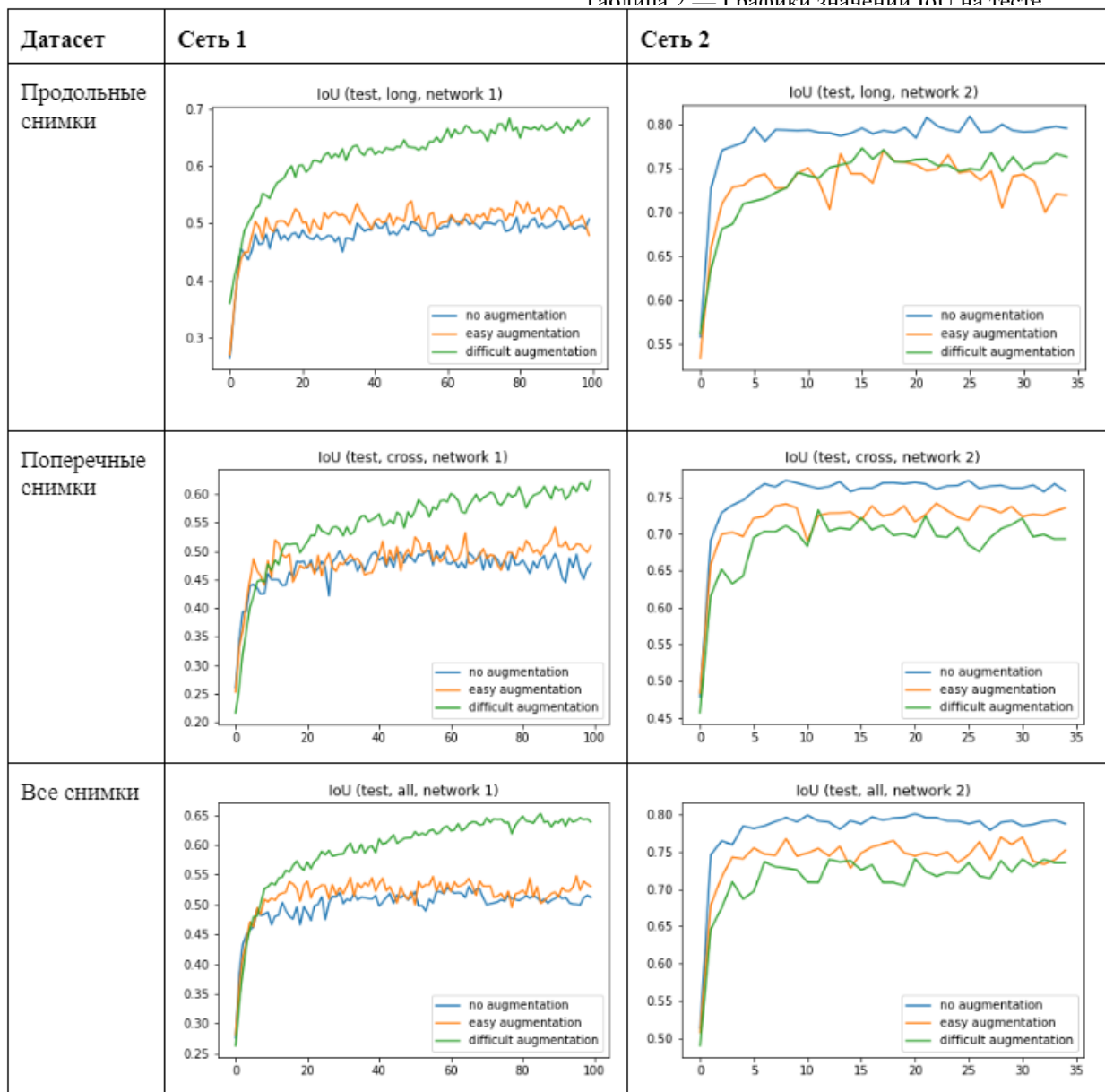


Таблица 3 — Графики значений DC на тесте

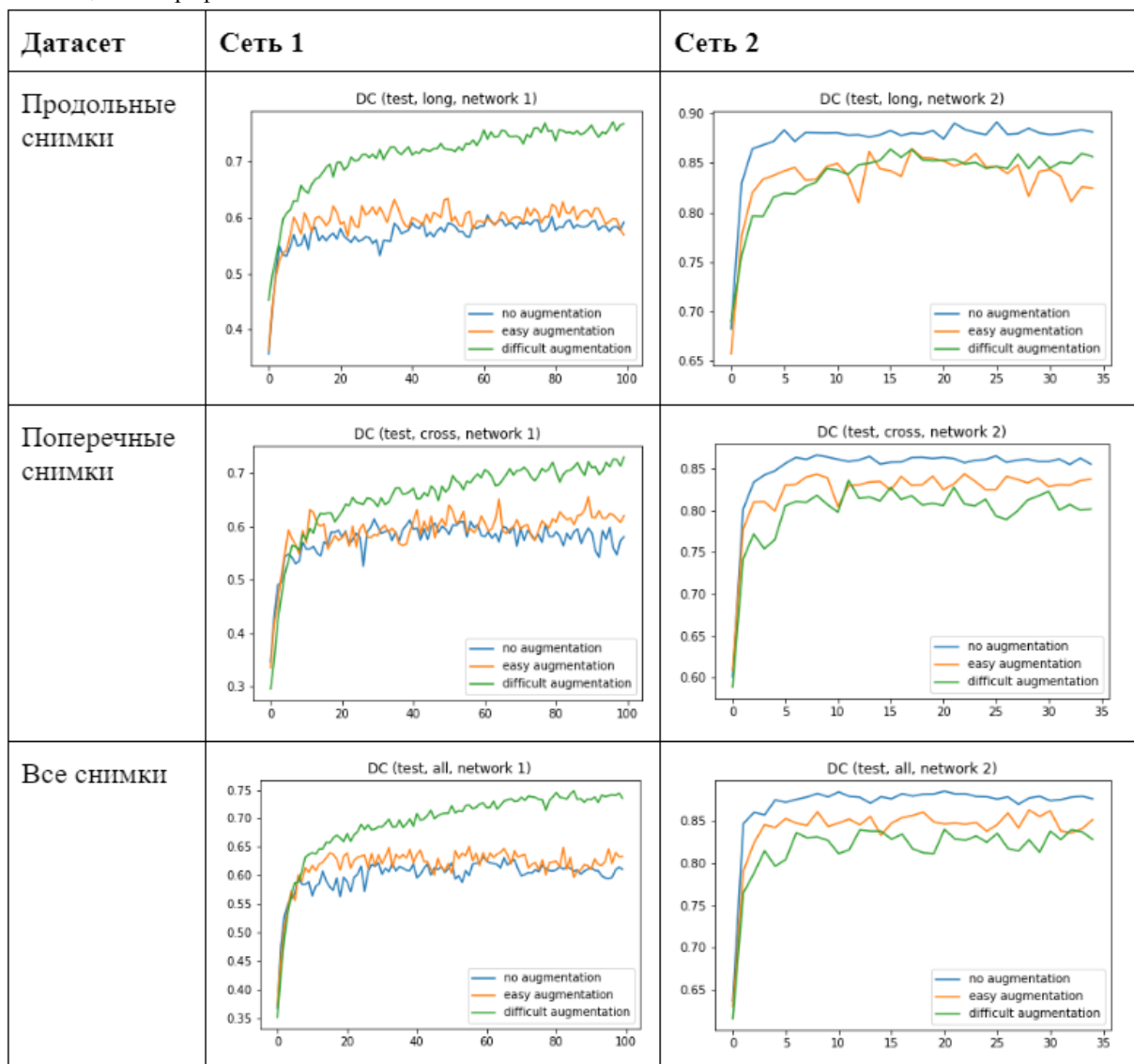


Таблица 4 — Значения метрик качества сегментации первой сети и последовательно двух сетей

| Сеть 1    | Сеть 2 | IoU <sub>сети 1</sub> |       |      | DC <sub>сети 1</sub> |       |      | IoU <sub>ср.</sub> |       |      | DC <sub>ср.</sub> |       |      |
|-----------|--------|-----------------------|-------|------|----------------------|-------|------|--------------------|-------|------|-------------------|-------|------|
|           |        | long                  | cross | all  | long                 | cross | all  | long               | cross | all  | long              | cross | all  |
| no        |        | 0,51                  | 0,50  | 0,53 | 0,60                 | 0,61  | 0,63 | 0,52               | 0,51  | 0,53 | 0,61              | 0,62  | 0,63 |
| easy      |        | 0,54                  | 0,54  | 0,55 | 0,63                 | 0,66  | 0,65 | 0,54               | 0,55  | 0,55 | 0,63              | 0,67  | 0,65 |
| difficult |        | 0,69                  | 0,62  | 0,65 | 0,77                 | 0,73  | 0,75 | 0,69               | 0,62  | 0,65 | 0,77              | 0,73  | 0,75 |
| dif.      | easy   | 0,69                  | 0,62  | 0,65 | 0,77                 | 0,73  | 0,75 | 0,70               | 0,63  | 0,66 | 0,78              | 0,74  | 0,76 |

лишь 35 эпох.



## 5. ЗАКЛЮЧЕНИЕ

В работе проанализирована эффективность применения различных методов аугментации наборов медицинских изображений в обучающей выборке нейронных сетей при решении задач семантической сегментации. Для это были классифицированы существующие подходы к расширению наборов данных.

Проведен эксперимент по обучению сети семантической сегментации DeeplabV3+ с применением разных методов аугментации, на нескольких наборах медицинских изображений УЗИ в областях с новообразованиями. Были обучены и исследованы 18 нейросетевых моделей.

Анализ полученных результатов позволяет сделать вывод о том, что предложенный для использования набор методов аугментации внес разнообразие в исходные данные, улучшив показатели метрик качества сегментации и повысив обобщающие способности моделей, на вход которым поступали снимки УЗИ целиком. Однако усложнение процесса аугментации наборов изображений, содержащих увеличенные области с новообразованиями, сказывается негативно на показателях качества сегментации моделей, обучающихся на таких данных.

## БЛАГОДАРНОСТИ

Авторы выражают благодарность Высшей инженеринговой школе НИЯУ МИФИ за помощь в возможности опубликовать результаты выполненной работы.

## БИБЛИОГРАФИЯ

1. Deo R.C. Machine Learning in Medicine // *Circulation*. Vol. 132. 20. 2015. P. 1920-1930.
2. Acs B., Rantalainen M., Hartman J. Artificial intelligence as the next step towards precision pathology // *Journal of Internal Medicine*. 2020. 288. P. 62-81.
3. Lipkova J., Chen R.J., Chen B., Lu M.Y., Barbieri M., Shao D., Vaidya A.J., Chen C., Zhuang L., Williamson D.F.K., Shaban M., Chen T.Y., Mahmood F. Artificial intelligence for multimodal data integration in oncology. 2022. Vol. 40. P. 1095-1110.
4. Ванушко В.Э. Узлы щитовидной железы — не всегда патология. *ИнфоМедФарм Диалог*, 2022 [Электронный ресурс]. — Режим доступа: <https://imfd.ru/2022/03/15/yzlishitzelez/>, свободный — (27.10.2022).
5. Botz B. European Thyroid Association TIRADS, 2021 [Электронный ресурс]. — Режим доступа:

- <https://radiopaedia.org/articles/european-thyroid-association-tirads>, свободный — (27.10.2022).
6. Kang Q., Lao Q., Li Y., Jiang Z., Qiu Y., Zhang S., Li K. Thyroid nodule segmentation and classification in ultrasound images through intra- and inter-task consistent learning // *Medical Image Analysis*. 2022. Vol. 79.
  7. Wang M., Yuan C., Wu D., Zeng Y., Zhong S., Qiu W.. Automatic Segmentation and Classification of Thyroid Nodules in Ultrasound Images with Convolutional Neural Networks // *MICCAI 2020: Segmentation, Classification, and Registration of Multi-modality Medical Imaging Data*. 2020. P. 109-115.
  8. Maharana K., Mondal S., Nemade B. A review: Data pre-processing and data augmentation techniques // *Global Transitions Proceedings*. 2022. Vol. 3. P. 91-99.
  9. Nalepa J., Marcinkiewicz M., Kawulok M. Data augmentation for brain-tumor segmentation: A review // *Frontiers in Computational Neuroscience*. 2019. Vol. 13.
  10. Lee J., Liu C., Kim J., Chen Z., Sun Y., Rogers J.R., Chung W.K., Weng C. Deep Learning for Rare Disease: A Scoping Review // *Journal of Biomedical Informatics*. 2022. Vol. 135.
  11. Zhou S., Nie D., Adeli E., Wei Q., Ren X., Liu X., Zhu E., Yin J., Wang Q., Shen D. Medical image segmentation using deep semantic-based methods: A review of techniques, applications and emerging trends // *Medical Image Analysis*. 2022. Vol. 82.
  12. Chlap P., Min H., Vandenberg N., Dowling J., Holloway L., Haworth A. A review of medical image data augmentation techniques for deep learning applications // *Journal of Medical Imaging and Radiation Oncology*. 2021. Vol. 65, Issue 5. P. 545-563.
  13. Hoar D., Lee P.Q., Guida A., Patterson S., Bowen C.V., Merrimen J., Wang C., Rendon R., Beyea S.D., Clarke S.E. Combined Transfer Learning and Test-Time Augmentation Improves Convolutional Neural Network-Based Semantic Segmentation of Prostate Cancer from Multi-Parametric MR Images. 2021. V. 210.
  14. Image Test Time Augmentation with PyTorch. TTACH [Электронный ресурс]. — Режим доступа: <https://github.com/qubvel/ttach>, свободный — (27.10.2022).
  15. Документация библиотеки imgaug для аугментации изображений [Электронный ресурс]. — Режим доступа: <https://imgaug.readthedocs.io/en/latest/>, свободный — (27.10.2022).
  16. Hussain Z., Gimenez F., Yi D., Rubin D. Differential Data Augmentation Techniques for Medical Imaging Classification Tasks // *AMIA Symposium*. 2017. P. 979-984.
  17. Chen Y., Yang X. H., Wei Z., Heidari A. A., Zheng N., Li Z., Chen H., Hu H., Zhou Q., Guan Q. Generative Adversarial Networks in Medical Image augmentation: A review // *Computers in Biology and Medicine*. 2022. Vol. 144.
  18. Shi G., Wang J., Qiang Y., Yang X., Zhao J., Hao R., Yang W., Du Q., Kazihise N. G. Knowledge-guided synthetic medical image adversarial augmentation for ultrasonography thyroid nodule classification // *Computer Methods and Programs in Biomedicine*. 2020. Vol. 196.
  19. Chen L., Zhu Y., Papandreou G., Schroff F., Adam H. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation // *ECCV 2018. Computer Science, Computer Vision and Pattern Recognition*. 2018.
  20. Tan M., Le Q. V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks // *ICML 2019. Machine Learning, Computer Vision and Pattern Recognition*. 2019.



# Augmentation of Image Sets for Training Neural Networks in Solving Semantic Segmentation Problems

I.A. Lozhkin, M.E. Dunaev, K.S. Zaytsev, A.A. Garmash

**Annotation.** The purpose of this work is to study the effectiveness of augmentation methods of image sets when they are insufficient in training sample of neural networks for solving semantic segmentation problems. For this purpose, the main groups of augmentation methods were considered and their effectiveness in solving problems of semantic segmentation of medical images was investigated. Two deep architectures DeepLabV3+ with the EfficientNetB6 encoder were used for training, testing and validation. The Intersection over Union and Dice coefficient were chosen as the target metrics for comparing the quality of semantic segmentation of images, which made it possible to determine the models with the best predictions. The obtained results confirmed the effectiveness of the proposed set of augmentation methods. The result of the work was the creation of an effective approach to augmentation of medical image sets to solve the problem of semantic segmentation.

**Keywords** — deep learning, augmentation, semantic segmentation

## REFERENCES

1. Deo R.C. Machine Learning in Medicine // Circulation. Vol. 132. 20. 2015. P. 1920-1930.
2. Acs B., Rantalainen M., Hartman J. Artificial intelligence as the next step towards precision pathology // Journal of Internal Medicine. 2020. 288. P. 62-81.
3. Lipkova J., Chen R.J., Chen B., Lu M.Y., Barbieri M., Shao D., Vaidya A.J., Chen C., Zhuang L., Williamson D.F.K., Shaban M., Chen T.Y., Mahmood F. Artificial intelligence for multimodal data integration in oncology. 2022. Vol. 40. P. 1095-1110.
4. Vanushko V.E. Thyroid nodules are not always pathology. InfoMedFarm Dialogue, 2022 [Internet source]. — Access mode: <https://imfd.ru/2022/03/15/yzlishitzelez/>, free — (27.10.2022).
5. Botz B. European Thyroid Association TIRADS, 2021 [Internet source]. — Access mode: <https://radiopaedia.org/articles/european-thyroid-association-tirads>, free — (27.10.2022).
6. Kang Q., Lao Q., Li Y., Jiang Z., Qiu Y., Zhang S., Li K. Thyroid nodule segmentation and classification in ultrasound images through intra- and inter-task consistent learning // Medical Image Analysis. 2022. Vol. 79.
7. Wang M., Yuan C., Wu D., Zeng Y., Zhong S., Qiu W.. Automatic Segmentation and Classification of Thyroid Nodules in Ultrasound Images with Convolutional Neural Networks // MICCAI 2020: Segmentation, Classification, and Registration of Multi-modality Medical Imaging Data. 2020. P. 109-115.
8. Maharana K., Mondal S., Nemade B. A review: Data pre-processing and data augmentation techniques // Global Transitions Proceedings. 2022. Vol. 3. P. 91-99.
9. Nalepa J., Marcinkiewicz M., Kawulok M. Data augmentation for brain-tumor segmentation: A review // Frontiers in Computational Neuroscience. 2019. Vol. 13.
10. Lee J., Liu C., Kim J., Chen Z., Sun Y., Rogers J.R., Chung W.K., Weng C. Deep Learning for Rare Disease: A Scoping Review // Journal of Biomedical Informatics. 2022. Vol. 135.
11. Zhou S., Nie D., Adeli E., Wei Q., Ren X., Liu X., Zhu E., Yin J., Wang Q., Shen D. Medical image segmentation using deep semantic-based methods: A review of techniques, applications and emerging trends // Medical Image Analysis. 2022. Vol. 82.
12. Chlap P., Min H., Vandenberg N., Dowling J., Holloway L., Haworth A. A review of medical image data augmentation techniques for deep learning applications // Journal of Medical Imaging and Radiation Oncology. 2021. Vol. 65, Issue 5. P. 545-563.
13. Hoar D., Lee P.Q., Guida A., Patterson S., Bowen C.V., Merrimen J., Wang C., Rendon R., Beyea S.D., Clarke S.E. Combined Transfer Learning and Test-Time Augmentation Improves Convolutional Neural Network-Based Semantic Segmentation of Prostate Cancer from Multi-Parametric MR Images. 2021. V. 210.
14. Image Test Time Augmentation with PyTorch. TTach [Internet source]. — Access mode: <https://github.com/qubvel/ttach>, free — (27.10.2022).
15. Library imgaug documentation for image augmentation [Internet source]. — Access mode: <https://imgaug.readthedocs.io/en/latest/>, free — (27.10.2022).
16. Hussain Z., Gimenez F., Yi D., Rubin D. Differential Data Augmentation Techniques for Medical Imaging Classification Tasks // AMIA Symposium. 2017. P. 979-984.
17. Chen Y., Yang X. H., Wei Z., Heidari A. A., Zheng N., Li Z., Chen H., Hu H., Zhou Q., Guan Q. Generative Adversarial Networks in Medical Image augmentation: A review // Computers in Biology and Medicine. 2022. Vol. 144.
18. Shi G., Wang J., Qiang Y., Yang X., Zhao J., Hao R., Yang W., Du Q., Kazihise N. G. Knowledge-guided synthetic medical image adversarial augmentation for ultrasonography thyroid nodule classification // Computer Methods and Programs in Biomedicine. 2020. Vol. 196.
19. Chen L., Zhu Y., Papandreou G., Schroff F., Adam H. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation // ECCV 2018. Computer Science, Computer Vision and Pattern Recognition. 2018.
20. Tan M., Le Q. V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks // ICML 2019. Machine Learning, Computer Vision and Pattern Recognition. 2019.