**SAR**

Радиолокационные системы с синтезированной апертурой (Synthetic Aperture Radar, SAR) используют электромагнитные волны для получения изображений земной поверхности. Эти устройства состоят из передатчика, который излучает сигналы, и приемника, который принимает отраженные радиоволны от объектов в окружающей среде. Спутник с данной аппаратурой движется в пространстве и в определённые моменты времени с фиксированным периодом зондирует поверхность. Этот процесс сбора данных создаёт синтетическую апертуру большой длины, что в свою очередь позволяет получить радиолокационные изображения с заданными пространственным разрешением и детализацией.

Описанный подход к визуализации поверхности имеет ряд преимуществ по сравнению со стандартным оптическим методом:

1. Невосприимчивость к погодным условиям. Сигналы, отправляемые передатчиком, соответствуют частотам, которые способны проникать сквозь облака, туман, снегопад, дождь и прочие метеорологические помехи, тем самым достигая поверхности Земли и беспрепятственно возвращаться.
2. Независимость от времени суток. Так как для зондирования используются только отправляемые сигналы и их отражённые копии, следовательно отсутствует необходимость в освещённости исследуемой поверхности.
3. Создание объёмной карты территории. Технология SAR позволяет вычислять время, которое понадобилось сигналу на преодоления расстояния от спутника до поверхности и обратно. По полученным данным о разности затраченного времени разными сигналами появляется возможность выстроить трёхмерную модель поверхности зондируемой области.
4. Исследование геологических образований. Благодаря способности излучаемых сигналов проникать сквозь некоторые структуры, открывается возможность изучать поверхности, которые невозможно увидеть со спутника при помощи оптических изображений. Например, почву в лесу, где кроны деревьев перекрывают обзор.

Изображения SAR находят широкое применение в различных областях, включая геодезию, картографию, экологию, геологию, сельское хозяйство и безопасность. Изображения SAR также используются для изучения климата и погоды, контроля за площадями лесных пожаров, мониторинга изменений ледников и ледяного покрова, а также для поиска и спасения людей в экстремальных условиях.

**Speckle noise**

Спекл-шум представляет собой мультипликативный паттерн случайных ярких и темных точек на изображении, которые могут искажать искомый объект и делать его неразличимым.

В отличие от оптических изображений, где световые волны не меняются при отражении от неоднородных объектов, радиолокационные волны, излучаемые когерентными источниками, могут испытывать случайные изменения фазы и амплитуды при отражении от объектов в окружающей среде. В результате пересечения этих волн их интерференция приводит к возникновению спекл-шума на изображении.

Для некоторых типов задач анализ распределения спекл-шума и его изменение во времени позволяет делать сложные выводы о структуре объекта, таким образом, даже данный тип помех способен нести в себе полезную информацию, например для изучения двойных звёзд, их массы и расстояния до них. Но в случае зондирования поверхности земли, спекл-шум затрудняет визуальное анализирование радиолокационных изображений и понижает качество автоматической обработки, например сегментацию объектов, классификацию, детектирование и другие. Поэтому, чтобы получить более точную информацию об исследуемой поверхности, необходимо проводить предобработку данных: фильтрацию спекл-шума.

**Related work**

На сегодняшний день существуют десятки различных фильтров радиолокационных изображений., которые можно условно разделить на несколько категорий.

Первые подходы к фильтрации радиолокационных изображений были представлены ещё с момента появления самих радиолокационных изображений. Изначально разработанные методы опирались на использование локальных статистических данных(spatial domain) в исследуемой квадратной области для извлечения и удаления спекл-шума. Наиболее известными и эффективными оказались Lee filter [1], Frost filter [2], Kuan filter [3]

Другой подход к фильтрации радиолокационных изображений заключается в применении вейвлет-преобразований[4, 5, 6] (wavelet-domain methods). Если стандартное представление сигнала во временной области не даёт информации о частотной составляющей, а использование Фурье-преобразования, наоборот, сохраняет только информацию о частоте сигнала, не оставляя информации о времени, то вейвлет-преобразование является обобщением спектрального анализа и заключается в извлечении частотных признаков из сигналов, но при этом с сохранением временных параметров. Способы фильтрации радиолокационных изображений, основанные на использовании вейвлет-преобразовании показывают более хорошие результаты, чем подходы, связанные с использованием локальных статистических данных в одной области(без частотной). Применение вейвлет-преобразований для обработки изображений стало широко распространенным в обработке сигналов и изображений, включая такие области, как компьютерное зрение, медицинская диагностика, видеообработка и другие.

Отдельной группой являются подходы, использующие для фильтрации и генерации нового denoised пикселя информацию о всём изображении – non-local methods [7, 8]. Один из первых таких фильтров: Non-local means filter [9], в основе которого лежит вычисление евклидового или другого расстояния от каждого пикселя до каждого и введения весовых коэффициентов, которые зависят от полученного значения расстояния. Таким образом, наиболее отдалённые пиксели будут оказывать меньший вклад, чем находящиеся рядом.

С ростом производительности и возможностью параллельных вычислений начали активно развиваться подходы к обработке изображений, использующие глубокие нейронные сети(DNN). Сложность применения DNN для фильтрации радиолокационных изображений состоит в том, что для создания модели, способной удалить шум с изображения, необходимо эту модель обучить: подавать на вход зашумлённые изображения, и изображения без шума. Так как на радиолокационных изображениях всегда присутствует спекл-шум, то не существует исходных примеров без шума. Это ограничение можно обойти, искусственно накладывая спекл-шум на чистые оптические изображения.

Как показала практика, применение архитектуры на основе свёрточных нейронных сетей (Convolutional Neural Networks, CNN) в задачах обработки изображений даёт хорошие результаты за счёт использования информации о пространственной структуре. Во время обучения нейронная сеть настраивается на извлечение признаков из SAR-изображений и на удаление шума, чтобы получить нешумную версию изображения. В статье [10] описывается механизм избавления от аддитивного белого Гауссовского шума с фиксированным уровнем, основанный на вычитании извлечённого при помощи CNN шума из изображения. Данный подход был развит и в [11] авторы статьи обобщили его на любой уровень шума. Конкретно задача удаления спекл-шума хорошо решается при помощи архитектуры нейронной сети, основанной на свёрточных слоях и остаточных соединениях[12].

Архитектура на основе автоэнкодера позволяет обучить нейронную сеть сжимать исходное изображение до более компактного представления с меньшей размерностью и оставлять наиболее значимые компоненты, , после чего расширить полученное представление обратно в изображение исходного размера, но без шума. На основе данной идеи авторы статьи[13] предложили одновременное использование нейронной сети на базе свёрточной архитектуры для увеличения receptive field с применением механизма автокодировщика для извлечения важных признаков.

Более сложным является применение архитектуры на основе GAN, идея которой базируется на двух нейронных сетях: Генератор получает на вход зашумлённое изображение и старается его отфильтровать. Дискриминатор получает на вход либо изображение с выхода Генератора, либо изображение без шума, после чего определяет, является ли это изображение изначально без шума, или получено после фильтрования генератором. Таким образом две нейронные сети состязаются и обучаются вместе. Главная проблема заключается в сложности выбора функции потерь для обучения и длительности процесса. На основеё такого подхода разработана система[14], задача которой заключается в повышении разрешения радиолокационных изображений.

Отдельно можно выделить обучение нейронной сети на основе Трансформеров. Впервые такой тип архитектуры был представлен в 2017 году, её отличительной особенностью стало появление механизма внимания[15] для извлечения зависимостей в последовательностях. Это позволило проектировать модели, способные справляться с решением задач обработки естественного языка и распознавания звука[16] намного лучше альтернативных методов, в частности, рекуррентных нейронных сетей. Адаптацией данного метода для обработки изображений стал проект Visual Transformers (ViT)[17]. Использование механизмов внимания в комбинации с другими методами применяется во различных областях. На основе данного подхода были решены многие задачи, в частности и фильтрация радиолокационных изображений. Например, применение вейвлет-преобразования и глубокой нейронной сети на базе Трансформеров[18]. Ключевая идея заключается в извлечении высокочастотной и низкочастотной информации с последующим нахождением зависимостей между ними при помощи механизма внимания.

**Описание модели**

Для фильтрации радиолокационных изображений были разработаны и обучены различные модели глубоких нейронных сетей, направленные на решение задач регрессии или классификации. Процесс получения готовой модели можно разделить на несколько этапов, часть из которых посвящены предобработке данных.

1. Подготовка данных для обучения

В первую задаётся фиксированный размер скользящего окна , от которого зависит количество нейронов во входном слое модели. К исходному оптическому изображению добавляются по краям зеркальные копии границ, где глубина границы составляет половину ширины выбранного окна. После на изображение накладывается мультипликативный спекл-шум, имеющий, согласно научным исследованиям, распределение Рэлея с параметром масштаба 0.27, по формуле (1)

1. Создание датасета

Получение набора данных заключается в генерации пар: данные на входе и ожидаемые данные на выходе. В качестве входных данных выбирается квадратное окно заданной ширины из зашумлённого изображения, ожидаемое значение на выходе – это пиксель в центре окна на исходном изображении без шума. Таким образом, окно скользит вдоль всего изображения с заданным шагом. На рисунке 1 графически показан данный процесс на изображении размера , к которому добавили границы, равные размеру окна где ширина окна составляет 3 и шаг окна равняется 1

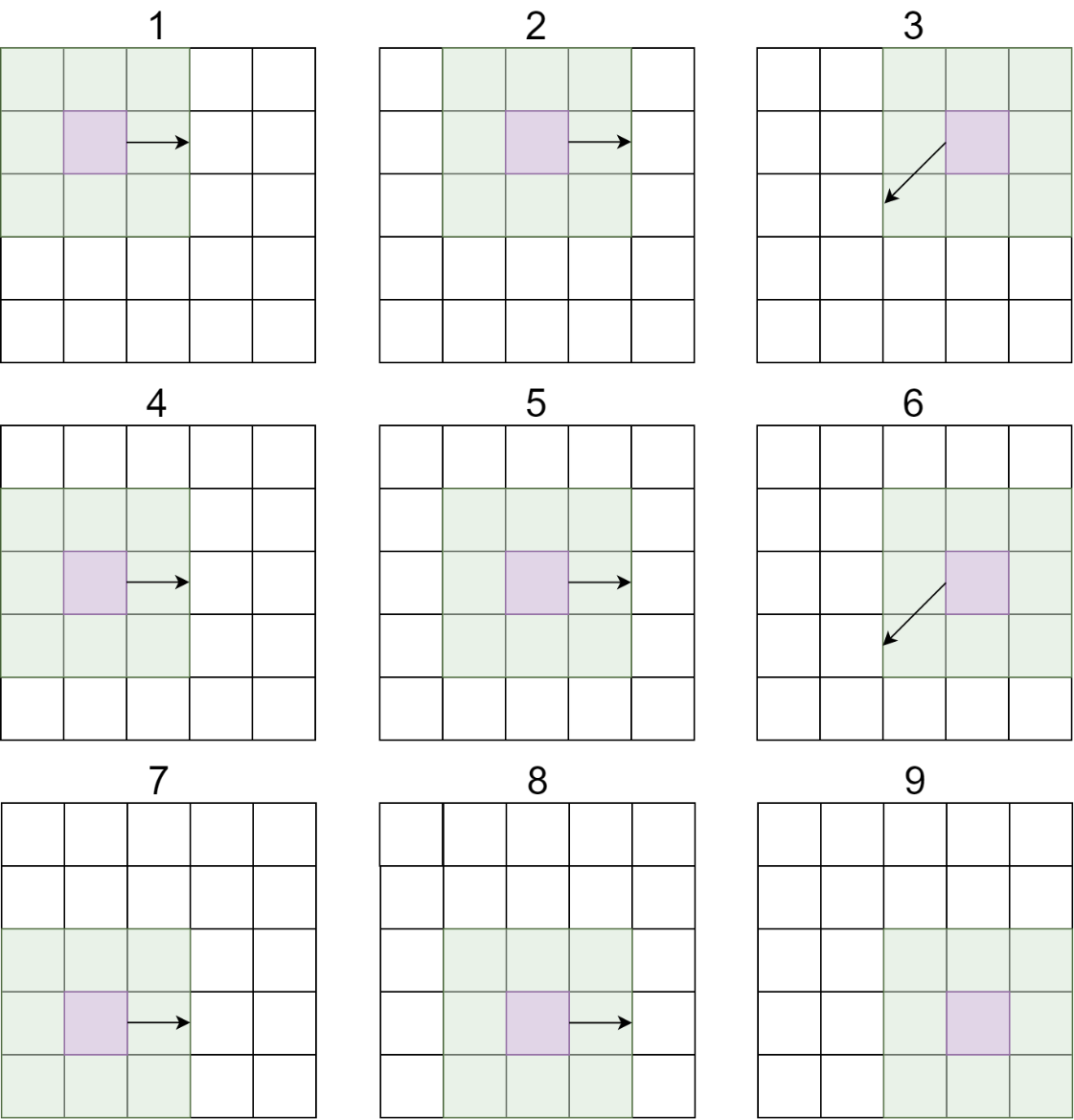
**

Рисунок 1 Процесс прохода окна по изображению с шириной окна N = 3.

Полученные данные преобразуются в одномерный массив и записываются в файл, либо переменную для дальнейшего использования. Рисунок 2 иллюстрирует, каким образом двумерные данные конвертируются в одномерный вектор.

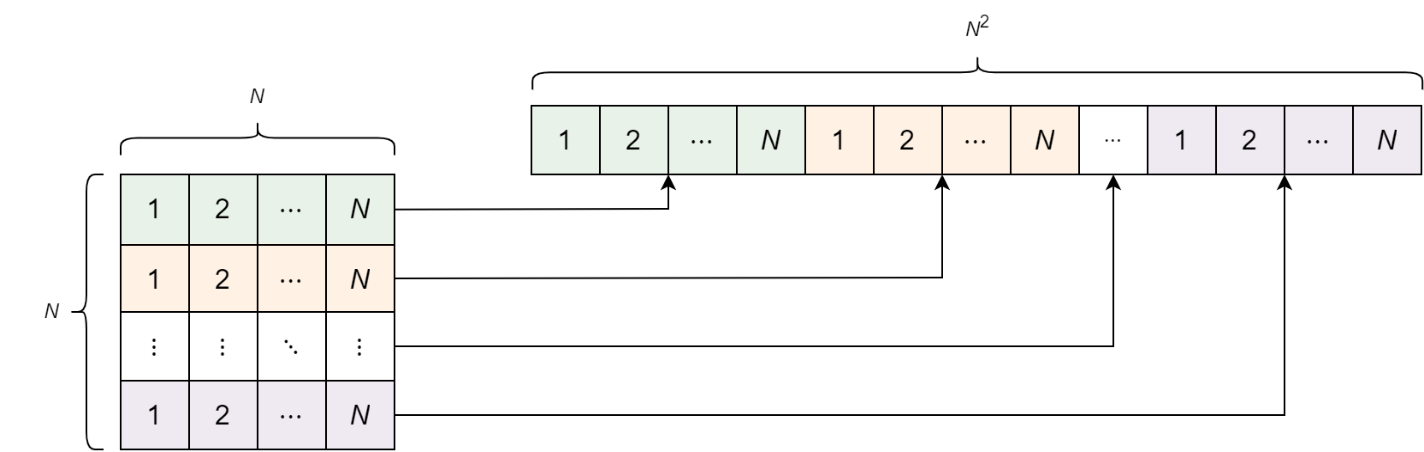


Рисунок 2. Процесс преобразование матрицы в вектор. N - количество элементов в строке и количество строк

Изложенный алгоритм действий применяется последовательно ко всем изображениям.

1. Описание архитектуры

Архитектура нейронной сети состоит из полносвязных слоёв: входной слой, выходной и *L* скрытых, количество которых можно изменять. Для повышения качества работы модели используется нормировка данных между слоями[19]. В качестве функции активации выбрана ReLU[19]. ~~Количество нейронов входного слоя равняется квадрату ширины окна (). С выходным слоем сложнее: в задаче регрессии на выходе только один нейрон, если же задача классификации, в таком случае на выходе должен быть вектор длинной 255.~~ Общая структура нейронной сети для задачи регрессии и классификации приведены на рисунках 1 и 2 соответственно. Количество нейронов в слое также указано на изображении.

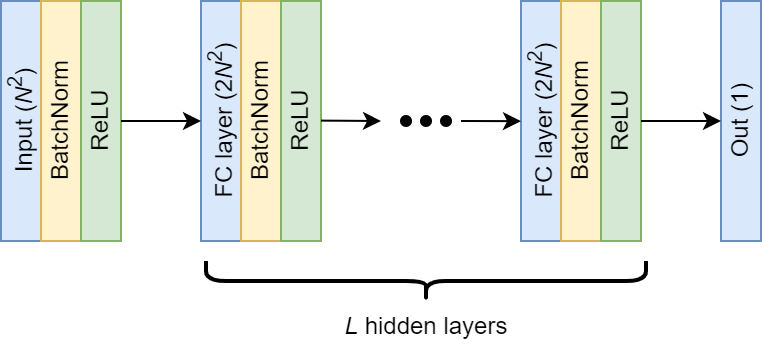


Рисунок 3. Архитектура полносвязной нейронной сети для решения задачи регрессии.

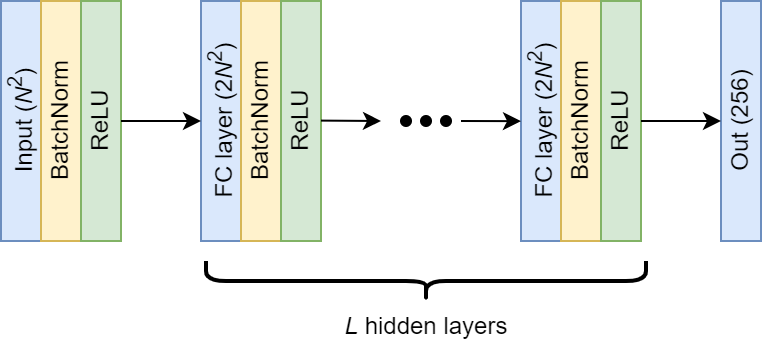


Рисунок 4. Архитектура полносвязной нейронной сети для решения задачи классификации.

1. Обучение модели

Обучение модели происходит на синтезированных данных. На вход подаётся случайная часть изображения – окно.

Для задачи регрессии используется среднеквадратичное отклонение в качестве функции потерь, которое сравнивает полученное одно числовое значение, характеризующее яркость пикселя с реальным значением яркости пикселя из изображения без шума.

Так как задаче классификации модель возвращает не одно значение, а вектор, то необходимо реальное значение пикселя предварительно закодировать также в вектор длины 256, состоящий из нулей и одной единицы, индекс которой в массиве характеризует яркость пикселя. Для обучения данной архитектуры используется перекрёстная энтропия.

Оптимизатором выбран алгоритм Adam[20]. Таким образом, происходит минимизация функции потерь и поиск минимума функции.

1. Фильтрация изображения

Чтобы отфильтровать зашумлённое изображение с предварительно добавленными границами, о которых говорится в пункте 1, на вход модели подаётся окно, в котором модель предсказывает пиксель в центре. Таким образом, итеративным процессом удаётся восстановить исходное изображение без шума.

1. Оценка модели

Для оценивания эффективности модели необходимо задать метрику, которая будет сравнивать отфильтрованное изображение с исходным. В данной работе оценивание производилось при помощи метрик SSIM[21] и GMSD[22]. Результаты представлены в таблице 1. Оценка модели проводилась на 4778 оптических изображениях, на которые накладывался шум с распределением Рэлея.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Approach | Window size | Hidden layers | Neurons in hidden layers | Is normalized | SSIM | GMSD |
| Regression (Our) | 7x7 | 5 | 49, 49, 98, 147, 147, 1 | No | 0.875 | 0.056 |
| Regression (Our) | 7x7 | 5 | 49, 49, 98, 147, 147, 1 | Yes | 0.846 | 0.064 |
| Classification (Our) | 7x7 | 5 | 49, 49, 98, 147, 147, 256 | No |  |  |
| Classification (Our) | 7x7 | 5 | 49, 49, 98, 147, 147, 256 | Yes |  |  |
| CNN |  |  |  |  |  |  |
| Anis diff |  |  |  |  |  |  |

Таблица 1. Результаты работы нейронной сети с различными параметрами

**Анализ**

Использование нейросетевых подходов для фильтрации радиолокационных изображений показало ожидаемо хорошие результаты по сравнению с классическими методами…

1. J.-S. Lee, “Speckle analysis and smoothing of synthetic aperture radar images,” Computer Graphics and Image Processing, vol. 17, no. 1, pp. 24 – 32, 1981. [Online]. Available: http://www.sciencedirect.com/ science/article/pii/S0146664X81800056
2. V. S. Frost, J. A. Stiles, K. S. Shanmugan, and J. C. Holtzman, “A model for radar images and its application to adaptive digital filtering of multiplicative noise,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-4, no. 2, pp. 157–166, March 1982.
3. D. Kuan, A. Sawchuk, T. Strand, and P. Chavel, “Adaptive restoration of images with speckle,” IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 35, no. 3, pp. 373–383, March 1987.
4. F. Argenti and A. Alparone, “Speckle removal from SAR images in the undecimated wavelet domain,” IEEE Transactions on Geoscience and Remote Sensing , vol. 40, pp. 2363–2374, Nov. 2002.
5. F. Argenti, T. Bianchi, and A. Alparone, “Multiresolution MAP despeckling of SAR images based on locally adaptive generalized gaussian pdf modeling,” IEEE Transactions on Image Processing, vol. 15, no. 11, pp. 3385–3399, Nov. 2006.
6. M. Dai, C. Peng, A. K. Chan, and D. Loguinov, “Bayesian wavelet shrinkage with edge detection for SAR image despeckling,” IEEE Transactions on Geoscience and Remote Sensing, vol. 42, no. 8, pp. 1642– 1648, Aug. 2004.
7. C. Deledalle, L. Denis, and F. Tupin, “Iterative weighted maximum likelihood denoising with probabilistic patchbased weights,” IEEE Transactions on Image Processing, vol. 18, no. 12, pp. 2661–2672, Dec. 2009.
8. H. Zhong, J. Xu, and L. Jiao, “Classification based nonlocal means despeckling for SAR image,” proc. of SPIE, vol. 7495, Oct. 2009.
9. P. Coupe, P. Hellier, C. Kervrann, and C. Baril- ´ lot, “Bayesian non local means-based speckle filtering,” Proceedings of IEEE International Symposium on Biomedical Imaging, pp. 1291–1294, May 2008.
10. B. Coll and J.-M. Morel, “A review of image denoising algorithms, with a new one,” SIAM Journal on Multiscale Modeling and Simulation, vol. 4, 01 2005.
11. G. Chierchia, D. Cozzolino, G. Poggi, and L. Verdoliva, “SAR image despeckling through convolutional neural networks,” in 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), July 2017, pp. 5438–5441.
12. P. Wang, H. Zhang, and V. M. Patel, “SAR Image Despeckling Using a Convolutional Neural Network,” IEEE Signal Processing Letters, vol. 24, no. 12, pp. 1763–1767, Dec 2017.
13. Qianqian Zhang, Ruizhi Sun “SAR Image Despeckling Based on Convolutional Denoising Autoencoder” <https://arxiv.org/abs/2011.14627>
14. Dongyang Ao, Corneliu Octavian Dumitru, Gottfried Schwarz, Mihai Datcu. “Dialectical GAN for SAR Image Translation: From Sentinel-1 to TerraSAR-X” <https://arxiv.org/abs/1807.07778>
15. Vaswani A. et al. Attention is all you need //Advances in neural information processing systems. – 2017. – Т. 30. <https://arxiv.org/abs/1706.03762>
16. Radford A. et al. Robust speech recognition via large-scale weak supervision //arXiv preprint arXiv:2212.04356. – 2022. https://arxiv.org/abs/2212.04356
17. Dosovitskiy A. et al. An image is worth 16x16 words: Transformers for image recognition at scale //arXiv preprint arXiv:2010.11929. – 2020. https://arxiv.org/abs/2010.11929
18. Li H. et al. DnSwin: Toward real-world denoising via a continuous Wavelet Sliding Transformer //Knowledge-Based Systems. – 2022. – Т. 255. – С. 109815. https://arxiv.org/abs/2207.13861
19. Ian Goodfellow, Yoshua Bengio, & Aaron Courville (2016). Deep Learning. MIT Press.
20. Kingma, D. P. & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980. <https://arxiv.org/abs/1412.6980>
21. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process. 2004 Apr;13(4):600-12. doi: 10.1109/tip.2003.819861. PMID: 15376593.
22. B. Zhang, P. V. Sander and A. Bermak, "Gradient magnitude similarity deviation on multiple scales for color image quality assessment," 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 2017, pp. 1253-1257, doi: 10.1109/ICASSP.2017.7952357.