

Научная статья
УДК 004.932
ББК 004.818.1
К 56
DOI: 10.53598/2410-3225-2024-3-346-51-62

**Архитектура U-Net для подавления шума на изображениях
с механизмами частотного и частотно-временного внимания**
(Рецензирована)

Алексей Сергеевич Коваленко

*Институт математики, механики и компьютерных наук имени И. И. Воровича,
Южный федеральный университет, Ростов-на-Дону, Россия, akov@sfedu.ru*

Аннотация. Задача подавления шума является одной из наиболее распространенных в области цифровой обработки изображений. Современные методы подавления шума основаны на применении U-Net-подобных глубоких сверточных сетей. Для улучшения качества подавления шума в эти архитектуры добавляют слои визуального внимания. Они позволяют выявлять пространственные области на изображении, имеющие важность для предсказания. С другой стороны, существуют подходы реализации слоев с частотным вниманием, основанным на выборе значащих частот из спектра Фурье. В работе предлагается подход для построения блока внимания, объединяющего методы визуального пространственного внимания и частотного внимания.

Ключевые слова: обработка цифровых изображений, подавление шума, глубокие нейронные сети, механизмы визуального внимания, вейлвет-преобразования

Для цитирования: Коваленко А. С. Архитектура U-Net для подавления шума на изображениях с механизмами частотного и частотно-временного внимания // Вестник Адыгейского государственного университета. Сер.: Естественно-математические и технические науки. 2024. Вып. 3 (346). С. 51–62. DOI: 10.53598/2410-3225-2024-3-346-51-62

Original Research Paper

**U-Net neural network architecture for image noise denoising
with mechanisms of frequency and time-frequency attention**

Aleksey S. Kovalenko

*Institute of Mathematics, Mechanics and Computer Science named after I. I. Vorovich,
Southern Federal University, Rostov-on-Don, Russia, akov@sfedu.ru*

Abstract. The noise reduction problem is one of the most common in the digital image processing field. State-of-the-art noise reduction techniques are based on the use of U-Net-like deep convolutional networks. To improve the quality of noise reduction, layers of visual attention are added to models. It provides possible to identify spatial regions in the image that are important for prediction. On the other hand, there are approaches to implementing frequency-aware layers based on selecting significant frequencies from the Fourier spectrum. This paper proposes an approach for constructing an attention block that combines the methods of visual spatial attention and frequency attention.

Keywords: digital image processing, denoising, deep neural networks, visual attention mechanisms, wavelet-transforms

For citation: Kovalenko A. S. U-Net neural network architecture for image noise denoising with mechanisms of frequency and time-frequency attention // The Bulletin of the Adyghe State University. Ser.: Natural-Mathematical and Technical Sciences. 2024. Iss. 3 (346). P. 51–62. DOI: 10.53598/2410-3225-2024-3-346-51-62

Введение

Подходы, основанные на использовании глубоких сверточных нейронных сетей, показывают наилучшие результаты в задаче подавления шума на цифровых изображениях. Обычно такие сети имеют архитектуру, подобную модели U-Net [1], где сеть работает как автокодировщик со сквозной передачей сигнала между слоями кодировщика и декодировщика. Общая структура модели U-Net показана на рисунке 1. В работе Comparing U-Net Based Models for Denoising Color Images [2] исследуются различные варианты построения архитектуры U-Net для удаления шума на входных изображениях и методы обучения таких моделей.

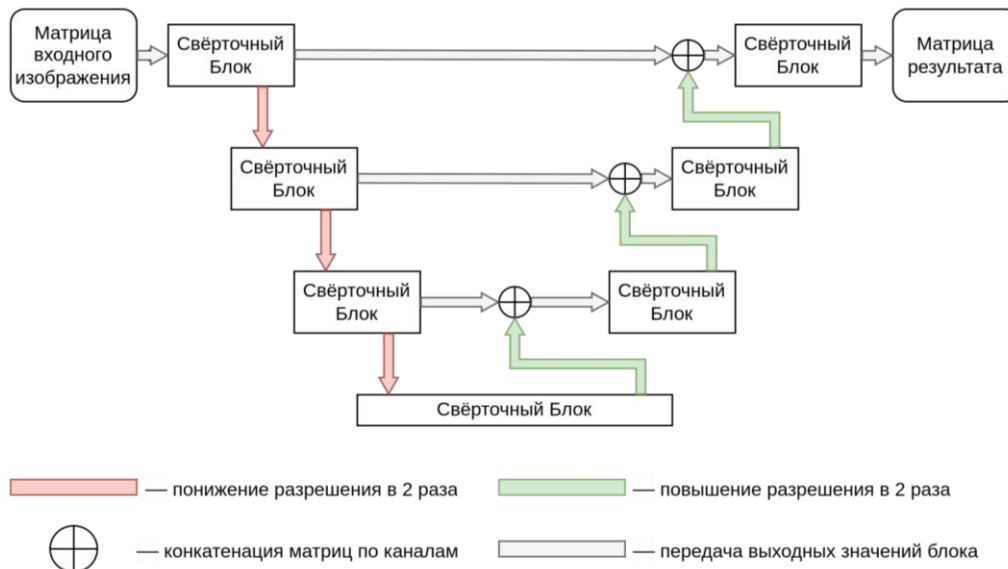


Рис. 1. Общая схема архитектуры U-Net

Fig. 1. General U-Net architecture diagram

Сверточные слои используют локальную контекстную информацию для построения последующих карт признаков. Но одни пространственные области или отдельные каналы в картах признаков могут иметь большее значение, чем другие. Для выделения этих значимых признаков применяются механизмы межканального и пространственного внимания [3]. Широко распространенным является метод Convolutional Block Attention Module (CBAM) [4], объединяющий эти механизмы. Сверточные архитектуры, использующие блоки CBAM, демонстрируют прирост качества в задачах подавления шума на изображениях [3, 5].

Для решения задач из области компьютерного зрения с помощью нейронных сетей широкое применение получили архитектуры, использующие механизм самовнимания (self-attention) [6, 7]. Эти нейросетевые архитектуры называются трансформерами и изначально были разработаны для эффективной обработки естественных языков. При анализе изображений впервые данную архитектуру применили к задачам классификации и детекции объектов на изображениях [8, 9]. Механизм самовнимания также применяется и для построения архитектур, подавляющих шум на изображениях [10], демонстрируя высокое качество восстановления изображений.

Помимо применения механизмов внимания, к пространственным признакам также используются и другие представления сигнала для выделения значимых компонент. Так авторы работы [11] применяют блок «сжатия и возбуждения» [12] для предсказания значимости каналов входной матрицы признаков по ее спектру Фурье. Преобразование Фурье представляет сигнал входного изображения в пространстве частот. Применение механизма частотного межканального внимания позволило повысить ка-

чество семантической сегментации медицинских изображений. Но применение механизма межканального внимания к частотным спектрам Фурье также обобщается и на задачу подавления шума на цифровых изображениях [13].

Для построения шумоподавляющих архитектур применяют и вейвлет-преобразования (ВП) [14]. ВП позволяет получать частотно-временную характеристику сигнала. Использование ВП совместно с механизмами внимания реализует метод частотно-временного внимания [15].

В данной работе предлагается блок внимания с концепцией, схожей с реализацией модуля внимания СВМ [4], но объединяющий механизмы частотного и частотно-временного внимания для выделения наиболее значимых признаков в частотных и пространственных областях.

Предлагаемый блок внимания

Блок внимания СВМ [4] состоит из двух частей, последовательно применяемых к входной матрице признаков: блок межканального внимания и блок пространственного внимания.

Реализацию части с межканальным вниманием предлагается строить на применении механизма внимания для выделения наиболее значимых каналов входной матрицы в ее частотном представлении. Схожий механизм применяется в работе FRUNet [11], его схема приведена на рисунке 2.

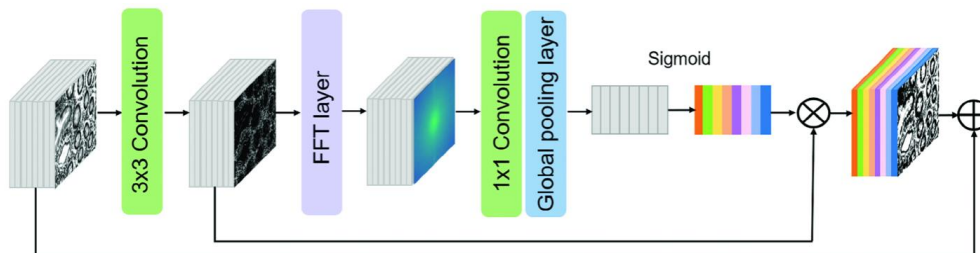


Рис. 2. Схема механизма внимания архитектуры FRUNet

Fig. 2. FRUNet attention mechanism diagram

В реализации метода блока внимания FRUNet для выбора канала для предсказания вектора со значениями значимости каналов [12] используется одно значение из каждого канала. Данное значение получается с помощью применения операции глобального объединения значений в канале (Global Pooling Layer). При этом выбирается значение с максимальной амплитудой. При такой реализации для описания частотной характеристики каждого канала используется мало информации.

В данной работе предлагается для каждого канала предсказывать комплекснозначную матрицу признаков следующего размера:

$$H \in \mathbb{C}^{c \times 8 \times 8},$$

где c – количество каналов матрицы признаков.

Эту матрицу предлагается получать с помощью применения блоков сжатия с остаточным соединением из архитектуры ResNet [16]. Используемые сверточные слои в предлагаемой реализации содержат комплекснозначные ядра сверток. Каждый такой блок сокращает размер признаков в два раза по каждой из размерностей. Для сжатия входной двумерной матрицы размера $N \times N$ потребуется применить блок сжатия $\log \frac{N}{8}$ раз.

Для понижения размерности признаков в архитектуре ResNet применяется операция объединения по максимальному значению (MaxPool). В случае комплексных матриц в качестве критерия выбора максимального элемента можно использовать зна-

чения амплитуд комплексных чисел:

$$|F(k, q)| = \sqrt{\text{Re}^2(k, q) + \text{Im}^2(k, q)}.$$

Схема операции объединения по максимальному значению для комплексных матриц приведена на рисунке 3.

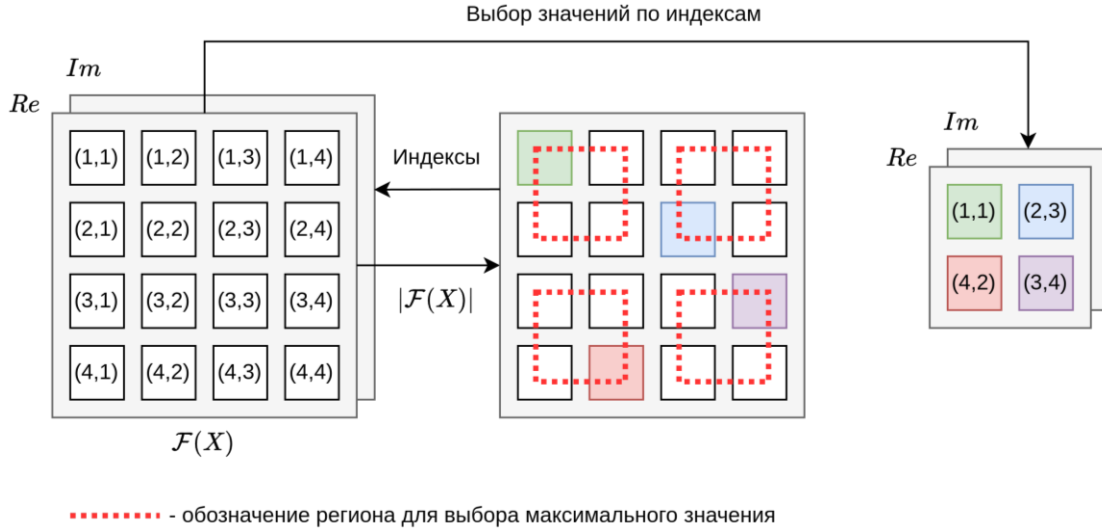


Рис. 3. Схема операции MaxPool для комплекснозначной матрицы

Fig. 3. Diagram of MaxPool for complex matrix

Вектор со значениями значимости каналов предсказывается с помощью многослойного перцептрона, имеющего веса в виде матрицы вещественных чисел. В этот перцептрон передается матрица H , которая содержит амплитуды значений комплекснозначной матрицы, получаемой из блоков обработки спектра. В результате предсказывается вектор следующего размера:

$$f \in R^{c \times 1 \times 1}.$$

Далее к полученному вектору применяется сигмоидальная функция активации. Функция сигмоиды имеет следующий вид:

$$\sigma(x) = \frac{1}{1 + e^{-x}}.$$

На этот вектор умножается входная матрица пространственных признаков. Общая схема предлагаемого блока частотного межканального внимания приведена на рисунке 4.

В отличие от метода межканального частотного внимания из FRUNet [11] предлагаемая реализация строит вектор внимания не по одному значению, обобщающему канал, а предсказывает набор частот, по которым определяется степень важности канала.

При построении второго блока межканального внимания используется двумерное быстрое вейвлет-преобразование Хаара (БВПХ). ВПХ по входному изображению строит 4 матрицы коэффициентов: A – коэффициенты аппроксимации, H , V , D – коэффициенты горизонтальных, вертикальных и диагональных разностей. Размерности матриц коэффициентов БВПХ (DWT) имеют следующий вид:

$$A, H, V, D = DWT(X), X \in R^{c \times H \times W}, A, H, V, D \in R^{\frac{c \times H}{2} \times \frac{W}{2}}.$$

Матрицы коэффициентов БВПХ H , V и D содержат высокочастотную информацию сигнала. Применение к ним блоков пространственного координатного внимания [17] позволяет выделять наиболее значимые коэффициенты матриц БВПХ. Результатом

такого объединения БВПХ и блока координатного внимания является механизм частотно-временного внимания, поскольку значимые признаки выделяются в частотно-временном представлении.

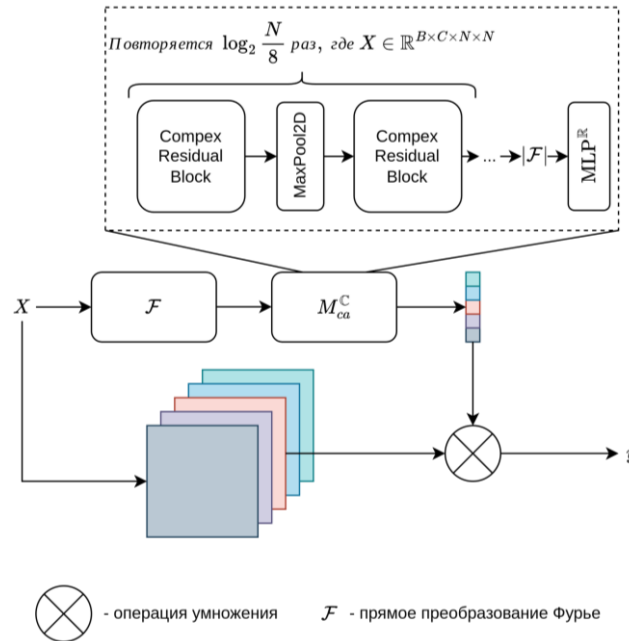


Рис. 4. Схема предлагаемого блока частотного межканального внимания

Fig. 4. Diagram of proposed frequency channel attention mechanism

Для предсказания карт координатного внимания для каждой матрицы коэффициентов сначала извлекаются общие признаки. Это производится путем применения сверточного слоя к матрице, полученной объединением всех матриц БВПХ по каналам. Затем по этим общим признакам строятся признаки для каждого вида коэффициентов отдельно. Это реализовано с помощью применения отдельных сверточных слоев. После этого к результатам слоев применяются блоки координатного внимания, которые предсказывают вектора внимания для столбцов и строк матриц БВПХ. В итоге данные вектора перемножаются с коэффициентами БВПХ входного сигнала, и в конце блока применяется обратное БВПХ. Схема описанного блока приведена на рисунке 5.

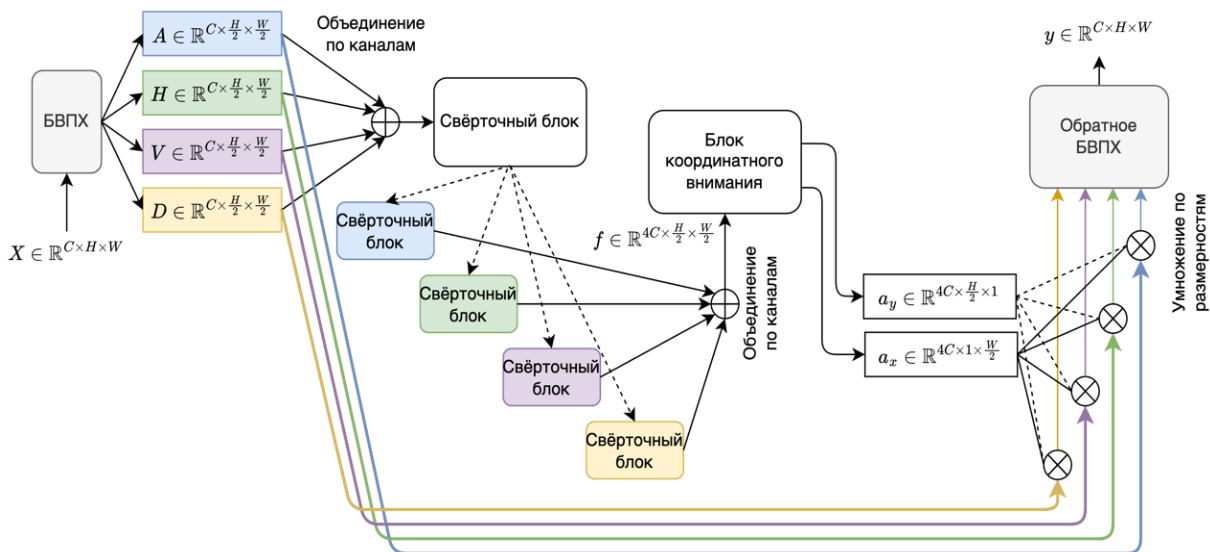


Рис. 5. Схема предлагаемого блока частотно-временного внимания

Fig. 5. Diagram of proposed time-frequency attention mechanism

Последовательное применение предлагаемых блоков частотного межканального внимания и частотно-временного внимания реализуют механизм внимания, учитывающий частотные и пространственные характеристики обрабатываемого сигнала.

Предлагаемый блок внимания основан на остаточной связи, построенной на входном сигнале, и применении механизма внимания к признакам этого сигнала [11]. Для извлечения признаков из сигнала используются сверточные блоки, состоящие из последовательного применения сверточного слоя с двумерным ядром размера 3×3 , слоя пакетной нормализации и функции активации LeakyReLU [18]. Такая реализация блока позволяет усилить значимые признаки в сигнале и сократить эффект затухания градиента при обучении [16].

Построение архитектуры U-Net с блоками внимания

Архитектура U-Net состоит из двух частей: кодировщика и декодера. В отличие от классической реализации автоэнкодера [19] у U-Net передается информация от кодирующих слоев в декодирующие слои. Это позволяет объединять информацию о декодировании сигнала с представлением его признаков после сжатия. Для проведения экспериментов блоки внимания применяются к признакам, передаваемым из блоков энкодера в блоки соответствующего уровня декодера. Это позволяет выделять наиболее полезные признаки для восстановления сигнала чистого изображения [5].

Для экспериментов по обучению моделей шумоподавления в качестве блоков внимания используются: блоки СВМ [4], FCA [11] и механизм объединения частотно-го и частотно-временного внимания, представленный в данной работе.

В реализации оригинальной модели [1], как правило, используются два варианта реализации блоков повышения размерности [20]. В первом варианте применяются операторы обратной свертки (деконволюция), во втором – последовательное применение метода билинейной интерполяции и сверточного слоя. Для понижения размерности в модели U-Net используется операция объединения по максимальному значению (Max-Pool) [1].

В проводимых экспериментах используется широко распространенная модификация архитектуры [20] на основе метода билинейной интерполяции, поскольку применение обратных сверточных слоев накладывает на предсказываемое изображение артефакты «шахматной доски» [21].

Полученная архитектура с примерами предсказанных карт внимания приведена на рисунке 6.

Дополнительные матрицы Vis строятся для визуализации выделяемых признаков после применения предлагаемого блока частотного межканального внимания Fa . Эти матрицы вычисляются как модуль разности между входными признаками в блок внимания и выходными из него. Затем эти разности усредняются по каналам:

$$Vis(X) = \frac{1}{c} \sum_{k=1}^c |Fa(X)^i - X^i|.$$

В приведенных экспериментах предлагаемая модифицированная архитектура U-Net обозначается как Time-Frequency-Attention-U-Net (TFAUnet).

Обучающий набор данных

Для обучения используется объединение нескольких наборов данных, предназначенных как для метода обучения с учителем (supervised learning), так и для обучения с частичным привлечением учителя (self-supervised learning) [22].

Набор для обучения с учителем представлен в виде пар изображений, состоящих из чистого примера x^i и изображения y^i с шумом η^i :

$$X^m = \{(x^i, y^i)\}_{i=1}^m, \quad y^i = x^i + \eta^i, \quad (y^i, x^i, \eta^i) \in R^{3 \times H \times W}.$$

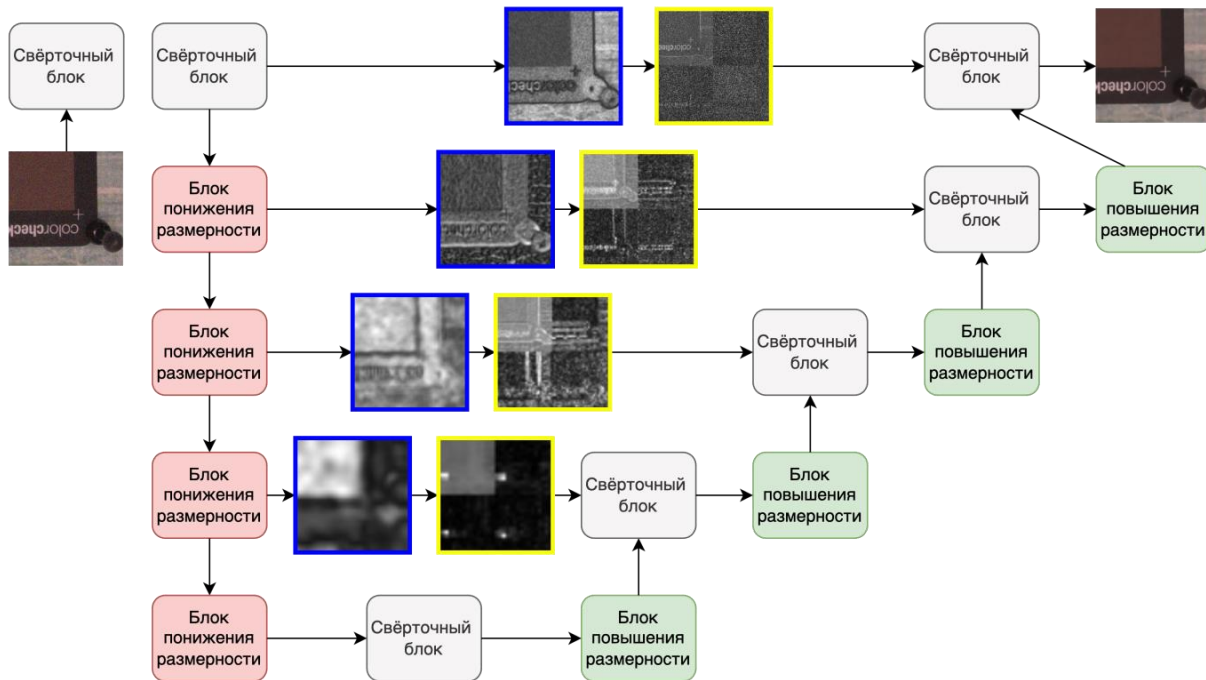


Рис. 6. Схема архитектуры U-Net с предлагаемыми блоками внимания: в синих прямоугольниках отображены признаки межканального внимания, в желтых прямоугольниках изображены карты пространственного внимания для четырех матриц вейвлет-коэффициентов

Fig. 6. Diagram of the U-Net architecture with proposed attention blocks: blue box figures show channel attention maps, yellow boxes show spatial attention maps for four wavelet coefficient matrices

В качестве такой обучающей выборки используется открытый набор данных Smartphone Image Denoising Dataset (SIDD) [23]. Набор SIDD предоставляет реальные зашумленные изображения и соответствующие им чистые изображения. Обучающая часть набора содержит 320 изображений высокого разрешения, а проверочная часть содержит 1280 пар изображений, имеющих размер 256 на 256 точек. Съемка производится авторами на 5 мобильных устройств с КМОП-сенсорами. В проводимых экспериментах эта часть используется для тестирования обученных моделей.

При обучении модели методом частичного привлечения учителя используются изображения, не содержащие шумовую компоненту. При формировании обучающей выборки на эти изображения накладывается синтетический шум [22]. В экспериментах для обучения используются изображения из наборов DIV2K [24], Sun-Hays 80, Set5, Set14 [25].

К чистым изображениям из этих наборов добавляется добавочный гауссовский шум с параметром математического ожидания, равным нулю, и изменяемым значением среднеквадратичного отклонения. При каждом использовании изображения параметр среднеквадратичного отклонения выбирается случайным образом из равномерного распределения $R(\sigma|a,b)$ с диапазоном $a=0$, $b=90$. Шум для каждого пикселя изображения семплируется независимо от остальных пикселей.

Семплирование изображения приведено в следующей формуле:

$$\tilde{y}_{i,j} = x_{i,j} + \eta_{i,j}, \quad \eta_{i,j} \sim N(0, \sigma | R(0,90)), \quad i \in [1, H], \quad j \in [1, W].$$

Изображения из наборов Kodak24 [26] и BSD68 [27] используются для этапа тестирования с фиксированными параметрами шума:

$$\sigma_{train} \in [5, 10, 15, 25, 50].$$

Результаты экспериментов и сравнение результатов

При обучении нейронной сети f решается задача поиска набора весов слоев сети ω , при котором будет достигнут минимум функции ошибки Шарбонье [28] на обучающем наборе данных:

$$\frac{1}{MHW} \sum_{k=1}^M \sum_{i=1}^H \sum_{j=1}^W \sqrt{(f(x_{i,j}^k, \omega) - y_{i,j}^k)^2 + 1} - 1 \xrightarrow{\omega} \min.$$

Для обучения выбрана функция потерь Шарбонье, поскольку она менее чувствительна к выбросам по сравнению с функциями L_1 , L_2 , а также обеспечивает более стабильное обучение [28].

В рамках проведенных экспериментов обучена предлагаемая модель TFAUnet. Также для сравнения обучались две версии архитектуры U-Net: оригинальная и с внедренным в нее блоком внимания CBAM.

В качестве валидационных метрик применялись: индекс структурного сходства изображений (SSIM) и пиковое отношение сигнала к шуму (PSNR). В качестве метода оптимизации параметров модели использовался AdamW [29]. Параметр скорости обучения (learning rate) был задан 0,001, а также применялась регуляризация параметров с параметром снижения (weights decay) 0,0001. Обучение модели останавливалось при выходе графика результата метрики PSNR на плато.

Для запуска обучения и тестирования использовалась версия 2.2.1 фреймворка PyTorch и операционная система Ubuntu 22.04 с CUDA версии 11.8.

Результаты работы обученных моделей сравнивались с подходами, которые используют механизм внимания, основанный на вейвлет-преобразовании [15]. Дополнительно для сравнения использовались результаты подходов SUNet [30] и Restformer [31], использующих U-Net подобную архитектуру с механизмом самовнимания. Вместе с перечисленными подходами результаты работы также сравнивались с моделью FEUSNet [32]. Архитектура FEUSNet построена на применении преобразования Фурье и обработки частотной характеристики сигнала.

Сравнение результатов для наборов с цветными изображениями производилось на Y канале цветового пространства YCrCb [30, 31]. Результаты сравнений на наборе BSD68 [27] с применением добавочного гауссова шума приведены в таблице 1. В случае неуказания авторами результатов работы архитектуры на сравниваемом уровне шума в таблицах с результатами ставится прочерк.

Таблица 1

Сравнение качества предлагаемой архитектуры с другими моделями на наборе данных BSD68

Table 1. Quality comparison of the proposed architecture with other models on the BSD68 dataset

Название модели	$\sigma = 10$		$\sigma = 15$		$\sigma = 25$		$\sigma = 50$	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Restformer [31]	-	-	34,39	-	31,78	-	28,59	-
FEUSNet [32]	35,88	0,969	-	-	-	-	27,87	0,848
SUNet [30]	35,23	0,967	-	-	-	-	26,90	0,844
U-Net [1]	30,06	0,922	29,96	0,901	29,05	0,871	27,75	0,818
U-Net + FCA [11]	35,73	0,951	34,56	0,938	33,25	0,915	31,76	0,877
U-Net + CBAM [4]	35,11	0,951	34,01	0,937	32,76	0,918	31,38	0,882
TFAUNet	35,94	0,957	34,83	0,944	33,58	0,925	32,14	0,892

Также в таблице 2 приведены сравнения на тестовом наборе Kodak24 с добавочным синтетическим шумом [26].

Таблица 2

Сравнение качества предлагаемой архитектуры с другими моделями на тестовом наборе данных Kodak24

Table 2. Quality comparison of the proposed architecture with other models on the test part of Kodak24 dataset

Название модели	$\sigma = 5$		$\sigma = 10$		$\sigma = 15$		$\sigma = 50$	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
MFEWA [16]	36,92	0,953	33,78	0,913	31,89	0,878	-	-
Restformer [34]	-	-	-	-	35,44	-	30,01	-
FEUSNet [32]	-	-	37,11	0,971	-	-	29,51	0,867
SUNet [30]	-	-	36,79	0,953	-	-	29,54	0,810
U-Net [1]	33,43	0,959	32,38	0,937	31,51	0,916	28,94	0,825
U-Net + FCA [11]	39,95	0,970	38,02	0,955	36,65	0,940	33,24	0,878
U-Net + CBAM [4]	36,25	0,969	36,90	0,955	35,09	0,942	32,44	0,884
TFAUNet	39,92	0,972	38,12	0,961	36,88	0,948	33,73	0,896

Для сравнения на примерах с шумом, полученных с помощью реальных сенсоров камер, использовалась валидационная часть набора SIDD [23].

Результаты работы обученных моделей на SIDD приведены в таблице 3.

Таблица 3

Сравнение качества предлагаемой архитектуры с другими моделями на тестовом наборе данных SIDD

Table 3. Quality comparison of the proposed architecture with other models on the SIDD test dataset

Название модели	Restformer [31]	U-Net [1]	U-Net + FCA [11]	U-Net + CBAM [4]	TFAUNet
PSNR	40,02	33,30	40,16	39,63	40,56
SIDD	0,960	0,793	0,923	0,926	0,931

Из проведенных экспериментов можно выделить, что применение механизмов частотного и частотно-пространственного внимания позволяет повысить качество подавления как добавочного шума к изображениям в градациях серого, так и реального шума, получаемого при съемке на КМОП-сенсоры камер. Пример работы обученных моделей на изображении из набора Kodak24 [26] изображен на рисунке 7.

Заключение

Реализованный механизм внимания позволяет выделять сильные признаки входного сигнала на основе его частотной характеристики. Также объединение методов межканального частотного и пространственно-частотного внимания при выделении признаков позволяет учитывать их пространственное расположение.

Обученные модели демонстрируют высокое качество восстановления чистого изображения как при обработке входных изображений с добавочным гауссовым шумом, так и при обработке изображений с сенсоров камер.

Исходный код на языке Python с реализацией подхода и запуска обучения предлагаемой модели содержится в GitFlic-репозитории по следующей URL-ссылке: <https://gitflic.ru/project/alexeykov/tfaunet/>



Рис. 7. Пример работы обученных моделей на изображении с добавлением шума разного уровня

Fig. 7. Example of trained models working on an image with different levels of noise added

Примечания

1. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation // Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Munich : Springer International Publishing, 2015. P. 234–241.
2. Komatsu R., Gonsalves T. Comparing U-Net Based Models for Denoising Color Images // AI. 2020. Vol. 1, No. 4. P. 465–486.
3. SCA-CNN: Spatial and Channel-Wise Attention in Convolutional Networks for Image Captioning / L. Chen [et al.] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Los Alamitos : IEEE Computer Society, 2017. P. 6298–6306. DOI: 10.1109/CVPR.2017.667
4. CBAM: Convolutional Block Attention Module / S. Woo [et al.] // ArXiv. 2018. Vol. abs/1807.06521.
5. Autoencoder Combined with CBAM Improves Denoising of MR Images / T. Lu [et al.] // 2021 11th International Conference on Information Technology in Medicine and Education (ITME). 2021. P. 209–213. DOI: 10.1109/ITME53901.2021.00050
6. Attention is all you need / A. Vaswani [et al.] // Advances in neural information processing systems. 2017. P. 5998–6008.
7. Андриянов Н. А., Дементьев В. Е., Ташлинский А. Г. Обнаружение объектов на изображении: от критериев Байеса и Неймана-Пирсона к детекторам на базе нейронных сетей EfficientDet // Компьютерная оптика. 2022. Т. 46, № 1. С. 139–159. DOI: 10.18287/2412-6179-CO-922
8. Stand-alone self-attention in vision models / P. Ramachandran [et al.] // Proceedings of the 33rd International Conference on Neural Information Processing Systems. N. Y. : Curran Associates Inc., 2019. P. 68–80.
9. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale / A. Dosovitskiy [et al.] // ArXiv. 2020. Vol. abs/2010.11929.
10. Uformer: A General U-Shaped Transformer for Image Restoration / Z. Wang [et al.] // 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022. P. 17662–17672. DOI: 10.1109/CVPR52688.2022.01716
11. Fourier Channel Attention Powered Lightweight Network for Image Segmentation / F. Zou [et al.] // IEEE Journal Translational Engineering Health Med. United States, 2023. Vol. 11. P. 252–260.
12. Hu J., Shen L., Sun G. Squeeze-and-Excitation Networks // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018. P. 7132–7141. DOI: 10.1109/CVPR.2018.00745
13. Evaluation and development of deep neural networks for image super-resolution in optical microscopy / C. Qiao [et al.] // Nature Methods. 2021. Vol. 18. P. 1–9. DOI: 10.1038/s41592-020-01048-5
14. Multi-level Wavelet-CNN for Image Restoration / P. Liu [et al.] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2018. P. 886–895. DOI: 10.1109/CVPRW.2018.00121

15. Kim M. C., Park J. H., Sunwoo M. H. Multilevel Feature Extraction Using Wavelet Attention for Deep Joint Demosaicking and Denoising // IEEE Access. 2022. Vol. 10. P. 77099–77109. DOI: 10.1109/ACCESS.2022.3192451
16. Deep Residual Learning for Image Recognition / K. He [et al.] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016. P. 770–778. DOI: 10.1109/CVPR.2016.9
17. Hou Q., Zhou D., Feng J. Coordinate Attention for Efficient Mobile Network Design // 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2021. P. 13708–13717.
18. Reluplex made more practical: Leaky ReLU / J. Xu [et al.] // 2020 IEEE Symposium on Computers and Communications (ISCC). 2020. P. 1–7. DOI: 10.1109/ISCC50000.2020.9219587
19. Фильтрация цифровых изображений на основе автоэнкодера / А. А. Ипатов [и др.] // Цифровая обработка сигналов. 2016. № 3. С. 79–84.
20. Defect segmentation for multi-illumination quality control systems / D. Honzatko [et al.] // Machine Vision and Applications. 2021. Vol. 32, Iss. 6. 16 p. DOI: 10.1007/s00138-021-01244-z
21. Odena A., Dumoulin V., Olah C. Deconvolution and Checkerboard Artifacts // Distill. 2016. No. 1 (10). DOI: 10.23915/distill.00003
22. Unleashing the Power of Self-Supervised Image Denoising: A Comprehensive Review / D. Zhang [et al.] // ArXiv. 2023. Т. abs/2308.00247.
23. Abdelhamed A., Lin S., Brown M. S. A High-Quality Denoising Dataset for Smartphone Cameras // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2018. P. 1692–1700. DOI: 10.1109/CVPR.2018.0018.
24. Agustsson E., Timofte R. NTIRE 2017 Challenge on Single Image Super-Resolution: Dataset and Study // The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. 2017. P. 1110–1121. DOI: 10.1109/CVPRW.2017.14.
25. Huang J.-B., Singh A., Ahuja N. Single Image Super-Resolution From Transformed Self Exemplars // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015. Vol. 07-12. P. 5197–5206.
26. Franzen R. Kodak Lossless True Color Image Suite. 1999. URL: <https://r0k.us/graphics/kodak> (дата обращения: 07.05.2024).
27. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics / D. Martin [et al.] // Proceedings Eighth IEEE International Conference on Computer Vision // ICCV. 2001. Vol. 2. P. 416–423. DOI: 10.1109/ICCV.2001.937655
28. Barron J. T. A General and Adaptive Robust Loss Function // IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019. P. 4326–4334. DOI: 10.1109/CVPR.2019.00446
29. Loshchilov I., Hutter F. Fixing Weight Decay Regularization in Adam // CoRR. 2017. Vol. abs/1711.05101. arXiv: 1711.05101.
30. Fan C.-M., Liu T.-J., Liu K.-H. SUNet: swin transformer UNet for image denoising // IEEE International Symposium on Circuits and Systems (ISCAS). 2022. February. P. 2333–2337.
31. Restormer: Efficient Transformer for High-Resolution Image Restoration / S. W. Zamir [et al.] // IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022. Iss. 39. P. 5718–5729. DOI: 10.1109/CVPR52688.2022.00564
32. FEUSNet: Fourier Embedded U-Shaped Network for Image Denoising / X. Li [et al.] // Entropy. 2023. Vol. 25, No. 10. P. 1418. DOI: 10.3390/e25101418/

References

1. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation // Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Munich : Springer International Publishing, 2015. P. 234–241.
2. Komatsu R., Gonsalves T. Comparing U-Net Based Models for Denoising Color Images // AI. 2020. Vol. 1, No. 4. P. 465–486.
3. SCA-CNN: Spatial and Channel-Wise Attention in Convolutional Networks for Image Captioning / L. Chen [et al.] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Los Alamitos : IEEE Computer Society, 2017. P. 6298–6306. DOI: 10.1109/CVPR.2017.667
4. CBAM: Convolutional Block Attention Module / S. Woo [et al.] // ArXiv. 2018. Vol. abs/1807.06521.
5. Autoencoder Combined with CBAM Improves Denoising of MR Images / T. Lu [et al.] // 2021 11th International Conference on Information Technology in Medicine and Education (ITME). 2021. P. 209–213. DOI: 10.1109/ITME53901.2021.00050
6. Attention is all you need / A. Vaswani [et al.] // Advances in neural information processing systems. 2017. P. 5998–6008.
7. Andriyanov N. A., Dementyev V. E., Tashlinsky A. G. Detection of objects in the images: from Bayesian and Neyman-Pearson criteria to EfficientDet neural network-based detectors // Computer Optics. 2022. Vol. 46, No. 1. P. 139–159. DOI: 10.18287/2412-6179-CO-922
8. Stand-alone self-attention in vision models / P. Ramachandran [et al.] // Proceedings of the 33rd International Conference on Neural Information Processing Systems. N. Y. : Curran Associates Inc., 2019. P. 68–80.
9. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale / A. Dosovitskiy

[et al.] // ArXiv. 2020. Vol. abs/2010.11929.

10. Uformer: A General U-Shaped Transformer for Image Restoration / Z. Wang [et al.] // 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022. P. 17662–17672. DOI: 10.1109/CVPR52688.2022.01716

11. Fourier Channel Attention Powered Lightweight Network for Image Segmentation / F. Zou [et al.] // IEEE Journal Translational Engineering Health Med. United States, 2023. Vol. 11. P. 252–260.

12. Hu J., Shen L., Sun G. Squeeze-and-Excitation Networks // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018. P. 7132–7141. DOI: 10.1109/CVPR.2018.00745

13. Evaluation and development of deep neural networks for image super-resolution in optical microscopy / C. Qiao [et al.] // Nature Methods. 2021. Vol. 18. P. 1–9. DOI: 10.1038/s41592-020-01048-5

14. Multi-level Wavelet-CNN for Image Restoration / P. Liu [et al.] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2018. P. 886–895. DOI: 10.1109/CVPRW.2018.00121

15. Kim M. C., Park J. H., Sunwoo M. H. Multilevel Feature Extraction Using Wavelet Attention for Deep Joint Demosaicking and Denoising // IEEE Access. 2022. Vol. 10. P. 77099–77109. DOI: 10.1109/ACCESS.2022.3192451

16. Deep Residual Learning for Image Recognition / K. He [et al.] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016. P. 770–778. DOI: 10.1109/CVPR.2016.9

17. Hou Q., Zhou D., Feng J. Coordinate Attention for Efficient Mobile Network Design // 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2021. P. 13708–13717.

18. Reluplex made more practical: Leaky ReLU / J. Xu [et al.] // 2020 IEEE Symposium on Computers and Communications (ISCC). 2020. P. 1–7. DOI: 10.1109/ISCC50000.2020.9219587

19. Filtration of digital images based on autoencoder / A. A. Ipatov [et al.] // Digital Signal Processing. 2016. No. 3. P. 79–84.

20. Defect segmentation for multi-illumination quality control systems / D. Honzatko [et al.] // Machine Vision and Applications. 2021. Vol. 32, Iss. 6. 16 p. DOI: 10.1007/s00138-021-01244-z

21. Odena A., Dumoulin V., Olah C. Deconvolution and Checkerboard Artifacts // Distill. 2016. No. 1 (10). DOI: 10.23915/distill.00003

22. Unleashing the Power of Self-Supervised Image Denoising: A Comprehensive Review / D. Zhang [et al.] // ArXiv. 2023. T. abs/2308.00247.

23. Abdelhamed A., Lin S., Brown M. S. A High-Quality Denoising Dataset for Smartphone Cameras // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2018. P. 1692–1700. DOI: 10.1109/CVPR.2018.0018.

24. Agustsson E., Timofte R. NTIRE 2017 Challenge on Single Image Super-Resolution: Dataset and Study // The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. 2017. P. 1110–1121. DOI: 10.1109/CVPRW.2017.14.

25. Huang J.-B., Singh A., Ahuja N. Single Image Super-Resolution From Transformed Self Exemplars // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015. Vol. 07-12. P. 5197–5206.

26. Franzen R. Kodak Lossless True Color Image Suite. 1999. URL: <https://r0k.us/graphics/kodak> (дата обращения: 07.05.2024).

27. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics / D. Martin [et al.] // Proceedings Eighth IEEE International Conference on Computer Vision // ICCV. 2001. Vol. 2. P. 416–423. DOI: 10.1109/ICCV.2001.937655

28. Barron J. T. A General and Adaptive Robust Loss Function // IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019. P. 4326–4334. DOI: 10.1109/CVPR.2019.00446

29. Loshchilov I., Hutter F. Fixing Weight Decay Regularization in Adam // CoRR. 2017. Vol. abs/1711.05101. arXiv: 1711.05101.

30. Fan C.-M., Liu T.-J., Liu K.-H. SUNet: swin transformer UNet for image denoising // IEEE International Symposium on Circuits and Systems (ISCAS). 2022. February. P. 2333–2337.

31. Restormer: Efficient Transformer for High-Resolution Image Restoration / S. W. Zamir [et al.] // IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022. Iss. 39. P. 5718–5729. DOI: 10.1109/CVPR52688.2022.00564

32. FEUSNet: Fourier Embedded U-Shaped Network for Image Denoising / X. Li [et al.] // Entropy. 2023. Vol. 25, No. 10. P. 1418. DOI: 10.3390/e25101418/

Статья поступила в редакцию 10.09.2024; одобрена после рецензирования 15.09.2024; принята к публикации 16.09.2024.

The article was submitted 10.09.2024; approved after reviewing 15.09.2024; accepted for publication 16.09.2024.