

УНИВЕРСАЛЬНЫЙ АЛГОРИТМ УЛУЧШЕНИЯ ИЗОБРАЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ

© 2022 Н. И. Бережнов✉, А. А. Сирота

*Воронежский государственный университет
Университетская пл., 1, 394018 Воронеж, Российская Федерация*

Аннотация. Рассматривается задача повышения качества изображений при воздействии различных видов шумов и искажений. Проводится сравнительный анализ стандартных алгоритмов и нейросетевых алгоритмов улучшения качества изображений, основанных на использовании глубоких нейронных сетей. Последние позиционируются как универсальное средство решения задачи улучшения качества изображений. Исследуется влияние гиперпараметров глубоких нейронных сетей на качество восстанавливаемых изображений. В первой части рассматривается теоретическая часть проблемы восстановления изображения, как решение сложной неустойчивой оптимизационной задачи. Исследуются различные техники аугментации, основанные на применении методов принудительно-го зашумления изображений и искусственной генерации дефектов, а также на применении нового подхода к аугментации данных путем частичной стилизации изображений. Помимо этого, используются различные метрики оценки качества восстановленных изображений, определяются их недостатки и границы применимости. Во второй части проводится сравнение известных классических алгоритмов с предложенными нейросетевыми, исследуется зависимость качества восстановления изображений от уровня шума. В ходе экспериментов по оценке эффективности предложенных подходов для улучшения изображений рассматриваются аддитивные и импульсные типы шумов, встречающиеся на изображениях медицинских снимков, а также на изображениях, полученных при помощи оптической когерентной томографии. Исследуется возможность работы алгоритмов в режиме реального времени. Определяется способ сравнения качества изображений через задачу сегментации с помощью сети U-net. Было показано, что нейронные сети не уступают классическим алгоритмам при улучшении качества изображений, а в некоторых случаях даже превосходят их.

Ключевые слова: обработка изображений, улучшение изображений, нейронные сети, аугментация данных, качество изображений, аддитивные и импульсные шумы.

ВВЕДЕНИЕ

В настоящее время в системах компьютерного зрения используются разнообразные методы обработки и улучшения изображений. Однако всегда возникает проблема их применимости к конкретным изображениям.

К сожалению, на сегодняшний день не существует универсального алгоритма их обработки. Так, например, в работе [1] авторы приводят эффективный алгоритм восстановления текстурных изображений малого масштаба через метод градиентных гистограмм, но при этом показывают, что его применимость существенно ограничена. Кроме того, современные исследователи фокусируются на отдельных типах изображений и их помехах.

✉ Бережнов Никита Игоревич
e-mail: beregnovnikita@yandex.ru



Контент доступен под лицензией Creative Commons Attribution 4.0 License.

The content is available under Creative Commons Attribution 4.0 License.

Например, медианный фильтр хорошо борется с шумом «соль и перец», но делает изображение размытым. Наоборот, алгоритмы повышения резкости страдают от присутствия помех на изображениях, которые также усиливаются с увеличением резкости.

Существует большое количество стандартных классических алгоритмов обработки изображений с целью повышения их качества, использующих

пространственную и частотную фильтрацию, адаптивные методы коррекции изображений, а также разнообразные нелинейные фильтры [2, 3]. Но они не являются универсальными и, в большинстве случаев, необходимо человеческое вмешательство для определения перечня применяемых алгоритмов и их параметров, весьма субъективное из отсутствия истинных, не зашумлённых изображений. В этом плане следует отметить работу [4], где авторы проводят восстановление изображений в микроволновом диапазоне длин волн и указывают на сложность проверки качества результата из-за отсутствия эталонов и тонкой настройки их параметров.

В связи с этим на сегодняшний день всё большую популярность набирает нейросетевой подход в обработке изображений. Особенно эффективны методы глубокого обучения и, в частности, свёрточные нейронные сети, так как они могут автоматически извлекать релевантные признаки из изображений. Это позволяет им быть инвариантными к множеству различных преобразований над данными, обращая внимание только на самые важные части изображения.

Существует огромное множество видов шумов и искажений, воздействующих на изображения. В данной работе основное внимание среди них было уделено аддитивным и импульсным помехам. В тоже время есть области, где использование традиционных моделей шумов неэффективно. В частности, в статье [5] авторы изучают шумы на изображениях, полученных при помощи электронного микроскопа. Они приходят к выводу, что данный шум нельзя аппроксимировать ни гауссовским, ни пуассоновским распределением. Он имеет более сложную природу, поэтому

необходимо исследовать другие, более сложные модели и техники генерации дефектов.

Одной из самых популярных архитектур для восстановления изображений является архитектура «encoder-decoder». Основная цель ее состоит в отображении исходного изображения в скрытое пространство признаков с последующим его восстановлением [6]. В 2017 году авторами работы [7] была предложена модель REDNet с остаточными связями, благодаря которым теряется меньше информации об исходном изображении, что увеличивает качество его последующего восстановления. В настоящей работе авторы как раз ориентировались на такую архитектуру и предложили свою модификацию этой архитектуры.

Проблемы восстановления изображения также исследуются на основе технологий обучения без учителя. В работе [8] авторы описывают U-net-подобную архитектуру, обучающуюся только на зашумлённых изображениях с использованием техники слепых пятен.

При решении указанных задач исследователи часто сталкиваются с проблемой переобучения, когда нейронная сеть становится чувствительна к обучающему набору данных и теряет свою обобщающую способность. Во избежание этого используются различные техники и приёмы аугментации данных. Простых техник аугментации данных не всегда бывает достаточно, иногда даже они приводят к искажению исходного распределения данных. Использование генеративно-состязательных сетей [9] может решить эту проблему, но исследователи сталкиваются с неустойчивостью и сложностью их обучения. Исходя из этого, в общем контексте решаемой задачи в данной работе был предложен легковесный метод стилизации изображений в реальном времени по аналогии с [10].

Целью работы является сравнительный анализ классических методов и алгоритмов улучшения изображений и предложенных нейросетевых алгоритмов, базирующихся на использовании технологий машинного обучения и аугментации данных. При этом исследовался новый способ аугментации для зашумления изображений, основанный на их частичной стилизации [10].

1. МЕТОДЫ И МАТЕРИАЛЫ

Классические алгоритмы устранения шумов и искажений. Как уже отмечалось, существует много разнообразных источников шума. Зашумление может возникать по различным причинам, например, из-за плохих условий съёмки, неидеального оборудования для захвата изображения и др. Эти шумы можно классифицировать по виду искажающего воздействия: аддитивного, мультипликативного и импульсного (аппликативного).

Для проведения экспериментов потребовалось использовать стандартные модели и алгоритмы внесения шумовых воздействий. В том числе: алгоритмы генерации импульсных и аддитивных помех, добавления градиентной засветки, нерезкого маскирование и алгоритмы размытия на основе виде гауссовского и медианного фильтров.

В общем случае искомое изображение может быть найдено путём решения интегрального уравнения Фредгольма второго рода:

$$z(x, y) = \int_{a_1}^{b_1} \int_{a_2}^{b_2} K(x, s_1, y, s_2) f(s_1, s_2) ds_1 ds_2 + n(x, y) \quad (1)$$

$$a_1 \leq x \leq b_1, \quad a_2 \leq y \leq b_2,$$

где $f(s_1, s_2)$ — исходное неискаженное изображение, $K(x, s_1, y, s_2)$ — функция ядра, $n(x, y)$ — аддитивный шум, $z(x, y)$ — результат искажения.

Решение уравнения (1) является обратной некорректной задачей по Адамару из-за её неустойчивости [11]. Поэтому используют специальные методы регуляризации для нахождения устойчивого решения. Среди них можно выделить методы итеративной, статистической, локальной регуляризации, регуляризацию Тихонова, а также методы оптимальной фильтрации Калмана и Винера. Последние являются более точными, однако они требуют гораздо больше дополнительной информации об изображении.

Так как большая часть помех имеет искусственное происхождение, то можно в ряде случаев, довольно точно определить отношение сигнал-шум и использовать винеровскую фильтрацию. Она работает эффективно, если

функцию ядра K можно представить в виде свёртки $h(x, y)$. Тогда уравнение (1) можно преобразовать в следующий вид [11]:

$$Z(u, v) = H(u, v)F(u, v) + N(u, v), \quad (2)$$

где $Z(u, v)$, $H(u, v)$, $F(u, v)$, $N(u, v)$ — фурье-образы функций $z(x, y)$, $h(x, y)$, $f(x, y)$, $n(x, y)$ соответственно.

Восстановленное изображение можно получить следующим образом:

$$\hat{F}(u, v) = \frac{1}{H(u, v)} \frac{|H(u, v)|^2 G(u, v)}{|H(u, v)|^2 + S_n(u, v) / S_f(u, v)}. \quad (3)$$

Функцией S здесь обозначены энергетические спектры шума и исходного изображения соответственно. На практике дробь $S_n(u, v) / S_f(u, v)$ заменяется на некоторую константу K , которую можно охарактеризовать как отношение сигнал-шум.

При восстановлении изображений фильтром Винера из-за неполной информации о сигнале могут возникать краевые эффекты; также он малопригоден для восстановления изображения от импульсных помех.

Поэтому для устранения такого рода помех часто используются разного рода эвристические и морфологические методы восстановления изображений, а также методы адаптивной фильтрации.

Существует множество вариантов адаптаций [12]. В работе решено было использовать адаптивную медианную фильтрацию и адаптивную линейную усредняющую фильтрацию.

Постановка задачи оптимизации и её решение на основе сверточных нейронных сетей. Сформулируем задачу оптимизации. Для этого представим уравнение (1) через линейный интегральный оператор A :

$$z = Af + n, \quad (4)$$

где f — тензор исходного неискаженного изображения, A — линейный интегральный оператор, n — аддитивный шум, z — результат искажения.

Задачу оптимизации тогда можно записать в следующем виде [13]:

$$\hat{z} = \inf_f \rho(Af, z), \quad (5)$$

где ρ — метрическая функция, \hat{z} — восстановленное изображение

Как показала практика, данная задача имеет не единственное решение. Даже при малых искажениях изображения можно найти несколько решений. Поэтому необходимо добавить некоторую априорную информацию об исходном изображении:

$$\hat{z} = \inf_f (\rho(Af, z) + \lambda\Phi(f)), \quad (6)$$

Функция $\Phi(f)$ должна гарантировать устойчивость и единственность решения. Её вид определяется типом задачи. Она часто задаётся в виде L_2 нормы. Параметр регуляризации λ позволяет контролировать визуальное качество восстанавливаемого изображения и обычно также определяется на практике.

Данное уравнение можно переписать в рамках статистического подхода, используя максимальную апостериорную оценку (MAP):

$$\hat{x} = \arg \max_x \log p(y|x) + \log p(x), \quad (7)$$

где $p(y|x)$ — условная плотность вероятности, а $p(x)$ представляет собой априорную информацию об истинном изображении и не зависит шума.

Подход к обучению нейронной сети основан на минимизации заданной функции потерь для нахождения оптимального параметра Θ , характеризующего весовые коэффициенты нейронной сети. Большинство функций потерь могут быть получены на основе MAP. В рамках нейросетевого подхода уравнение (6) можно скорректировать следующим образом:

$$\hat{z} = \arg \min_f \rho(Af, z) + \lambda\Phi(f; \Theta), \quad (8)$$

где Θ — набор весовых коэффициентов нейронной сети

На практике нейронные сети могут иметь более высокую производительность и использоваться в реальном времени, в отличие от численных методов решения обратных задач, которые страдают от недостатка априорной информации об искажении и не всегда бывают устойчивы.

Архитектура сверточной нейронной сети. В ходе исследований было решено использовать сверточные нейронные сети. В частности, рассматривалась архитектура

REDNet. Нейронная сеть состояла из сверточной части, извлекающей необходимые признаки из изображения, и развёртывающей части (deconvolutional part), которая восстанавливала изображение. Для того, чтобы нейронная сеть могла запомнить большее количество искажений, решено было сделать её глубже, чем предполагалось в REDNet, и также увеличить количество ядер в сверточных слоях и затем подавать на её вход изображения с различными типами помех.

Авторы [7] в своём решении не используют слои понижающей дискретизации — слои пулинга. Они считают, что это приводит к ухудшению качества восстанавливаемого изображения, так как слои пулинга выделяют абстрактные признаки на изображении, но при этом теряют его мелкие детали.

В последнее время при восстановлении изображений набирают популярность остаточные связи, когда часть признаков, полученных на ранних слоях, объединяется или суммируется с признаками, полученными на более глубоких. Это позволяет сохранить больше значимой информации об изображении и эффективно его восстанавливать. Также это приводит к быстрой сходимости при обучении и решает проблему исчезающего градиента для глубоких нейронных сетей. В данной архитектуре также используются остаточные связи, но при этом полученные после суммирования отрицательные значения обнуляются.

Изложенные выше соображения решено было использовать при формировании окончательной архитектуры сети. Также на основе модификации данной архитектуры сети был предложен новый способ искажения изображений, основанный на их частичной стилизации в реальном времени. В его основе лежит использование слоёв AdaIN (adaptive instance normalization). Слои AdaIN похожи на обычную нормализацию экземпляров, но при этом не имеют обучаемых параметров, а вес и смещение вычисляются на основе изображения стиля. Формула нормализации представлена ниже:

$$AdaIN(x, y) = \sigma(y) \left(\frac{x - \mu(x)}{\sigma(x)} \right) + \mu(y), \quad (9)$$

где $\mu(x)$, $\sigma(x)$ — математическое ожидание и стандартное отклонение исходного изображения, а $\mu(y)$, $\sigma(y)$ — математическое ожидание и стандартное отклонение изображения стиля.

Все преобразования выполняются не над самими изображениями, а над их признаками. Авторы [10] используют для этих целей предобученную VGG сеть, однако, в ходе исследования было обнаружено, что данная сеть не позволяет эффективно извлекать признаки для чёрно-белых изображений медицинских снимков.

Исходя из этого, решено было использовать собственный кодировщик, который обучался извлекать признаки непосредственно во время обучения стилизации изображений. В качестве функции потерь была выбрана среднеквадратичная ошибка (MSE). При этом возникала проблема устойчивости при обучении данной сети, так как она не имела глобального ориентира и могла восстанавливать нереальные текстурные изображения. Её удалось решить с использованием дополнительной функции потерь — identity loss. Декодер помимо декодирования стилизованного изображения, также декодировал и исходное изображение.

Помимо этого, данная архитектура позволяет использовать параметр alpha, регулирующий степень стилизации изображений. Предложенная архитектура сети представлена на рис. 1. Она состоит из трёх частей — кодировщика, декодировщика и слоёв AdaIN. Стрелки на схеме указывают, что AdaIn применяется для каждой остаточной связи исходной нейронной сети.

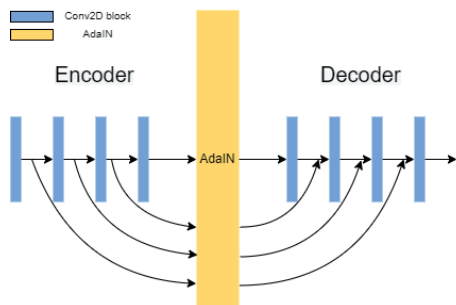


Рис. 1. Предложенная архитектура нейронной сети для стилизации изображений
 [Fig. 1. Proposed neural network architecture for image stylization]

Генерация дефектов. Помимо частичной стилизации изображений было рассмотрено несколько способов генерации дефектов.

1. *Генерация дефекта на основе гауссовского распределения.* Случайным образом задавался пик распределения, случайно определялись линейные размеры помехи и формировалось усечённое нормальное распределение.

2. *Генерация дефекта при помощи наращивания.* Выбирались центры дефектов согласно пуассоновскому случайному процессу. Затем происходило их наращивание. Данный тип генерации был подробно описан в [14]. Авторы проводят исследование формы дефекта путём изменения определённых масок, меняющих топологию его наращивания. Изменяя вероятности искажения соседних пикселей относительно текущего заданием нормального и экспоненциальных распределений для наращивания области, можно изменять прозрачность генерируемой помехи, что было показано в работе [15].

В данный подход была добавлена постобработка помех: применялся фильтр Гаусса с размером окна 3×3 для их размывания или, наоборот, использовался приём нерезкого маскирования для повышения их резкости и контрастности.

При использовании указанных выше двух способов не учитывалось наложение помех друг на друга. Рассматривались три случая искажения ими исходного изображения:

- помеха умножалась на коэффициент и складывалась с исходным изображением;
- помеха перемножалась с исходным изображением, происходило моделирование локальных мультипликативных помех.
- помеха замещала исходные пиксели изображения, происходило моделирование импульсных и аппликативных помех.

Примеры дефектов, сгенерированных способами 1 и 2 представлены на рис. 2 в увеличенном размере и на чёрном фоне.

3. *Генерация дефектов в виде засветки изображения по краям.* Часто возникают помехи при неравномерном освещении исследуемого объекта. Предположим, что источник света находится сбоку от изображённых объектов. Согласно физическому закону обратных ква-

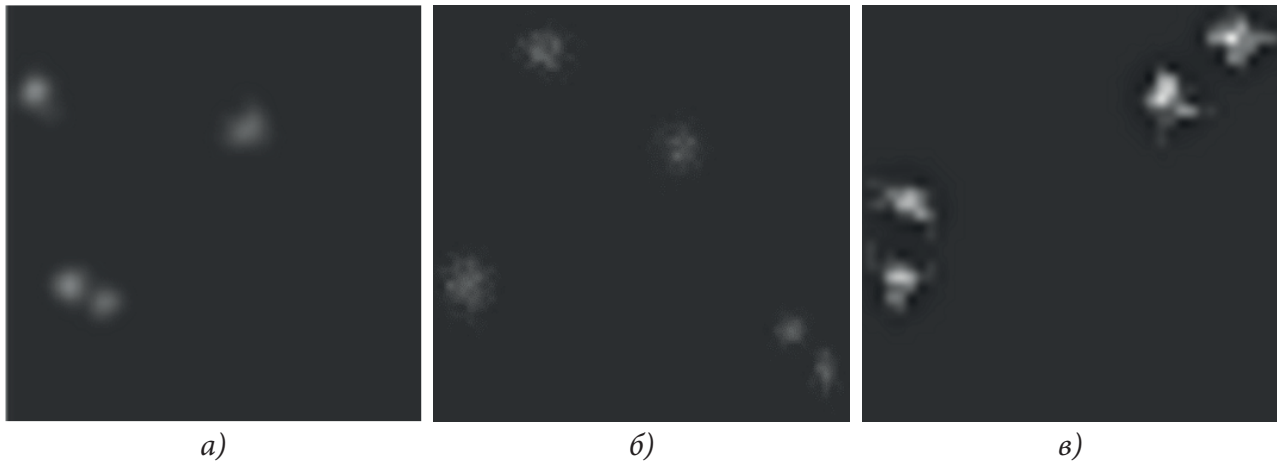


Рис. 2. Сгенерированные дефекты на чёрном фоне: а) — генерация дефектов первым способом; б) и в) — генерация дефектов вторым способом с применением гауссовского фильтра и нерезкого маскирования соответственно

[Fig. 2. Generated defects on a black background: a) — generation of defects by the first method; b) and c) — generation of defects by the second method using a Gaussian filter and non-sharp masking respectively]

дратов, интенсивность света обратно пропорциональна квадрату расстояния от его источника. Для простоты будем считать, что световые волны, распространяющиеся от источника, имеют плоский фронт.

Засветку представим, как аддитивную помеху и будем добавлять её только с одной стороны изображения согласно следующей формуле:

$$\hat{I}(x, y) = I(x, y) + h / (d + x)^2, \quad (10)$$

где d — предполагаемое расстояние до источника освещения, h — высота изображения, x, y — координаты пикселя исходного изображения I .

Указанная выше формула используется для генерации засветки от источника, находящегося справа от изображения.

4. *Генерация двоящихся границ объекта.* Данный тип помех очень часто возникает при расфокусировки датчика камеры при съёмке. Для его моделирования исходное изображение было размыто, чтобы не усиливать присутствующие на нём шумы и другие нежелательные артефакты. Затем был дважды применён приём нерезкого маскирования. После этого производилось смещение пикселей изображения на шаг d . На практике оно не задавалось больше 10. Далее полученное изображение накладывалось на исходное изображение.

Используемые датасеты. При проведении исследований были использованы изображения из датасета ImageNet, а также чёрно-белый сильно зашумленный набор медицинских снимков изображений сетчатки глаза, полученный при помощи оптической когерентной томографии. Все изображения были приведены к размеру 256×256 пикселей для возможности сравнения алгоритмов между собой. Затем изображения были переведены в оттенки серого и исследовалась только их яркостная составляющая, так как предполагалось, что искажающая функция действует на каждый цветовой канал одинаково.

Далее происходила генерация различных видов дефектов с целью зашумления исходных изображений и последующего их восстановления, чтобы иметь возможность сравнить эффективность классических и нейросетевых алгоритмов. Медицинские снимки изначально были достаточно сильно зашумлены, поэтому возникла необходимость оценки параметров шума, уже имеющегося на них.

Оценка качества изображений PSNR и SSIM. Для оценки качества изображений в первую очередь использовался визуальный подход. Сравнивались два восстановленных изображения с исходным. Одно из которых было получено с помощью сверточной ней-

ронной сети, второе при помощи классических алгоритмов, рассмотренных выше.

Для количественной оценки использовалось среднеквадратичное расстояние между пикселями истинного изображения и восстановленного. Помимо этого, вычислялись известные метрики: отношение сигнал-шум PSNR и метрика структурного подобия SSIM для проверки качества изображений. Эти метрики являются общепринятыми в данной области, однако требуют наличия изображения эталона. Исходя из этого, данные метрики можно было применять только для датасета ImageNet.

Основным недостатком этих метрик является их не полное соответствие человеческому восприятию. Это было экспериментально продемонстрировано в работе [16] Использование универсальной метрики качества, не требующей эталонов, до сих пор остаётся актуальной и нерешённой задачей.

В данной работе также был использован подход для определения стандартного отклонения аддитивного гауссовского шума [17]. Он основан на определении уровня шума на изображении на основе усреднения дисперсии в его блоках. Это позволило оценить исходный уровень шума для датасета медицинских изображений сетчатки глаза.

Помимо этого, решено было сравнить качество восстановления изображений через задачу сегментации. Нейронная сеть Unet обучалась на зашумлённых и восстановленных изображениях. Далее сравнивалась точность работы нейронной сети, чтобы доказать факт того, что на восстановленных точность больше.

2. РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ И ИХ ОБСУЖДЕНИЕ

Классические методы аугментации и зашумления непосредственно использовались через ImageDataGenerator в Keras — высокоуровневым API Tensorflow. Генераторы в Keras имеют возможность осуществлять аугментацию данных в режиме реального времени и позволяют случайно применять к изображению перечисленные выше искажения в начале каждой эпохи обучения. Параметры алгоритмов задаются в виде допустимых диапазонов, из которых случайным образом выбирается значение. Это позволяет не только случайно применять к изображению различные методы аугментаций, но и контролировать их силу. Помимо этого, была добавлена частичная стилизация исходных изображений через AdaIN.

На рис. 3 представлены изображения, полученные при помощи частичной стилизации. В качестве стиля решено было выбрать однотонные и текстурные изображения с различной степенью яркости и зашумлённости для того, чтобы явно видеть эффект стилизации. На рис. 3 стилизованное изображение получилось размытым и более ярким по сравнению с оригиналом. Такого же эффекта можно добиться с помощью эвристических алгоритмов, однако, нейронная сеть делает это автоматически.

В табл. 1 приведены результаты сравнения классических алгоритмов и нейросетевого алгоритма восстановления изображений. В таблице использовались следующие

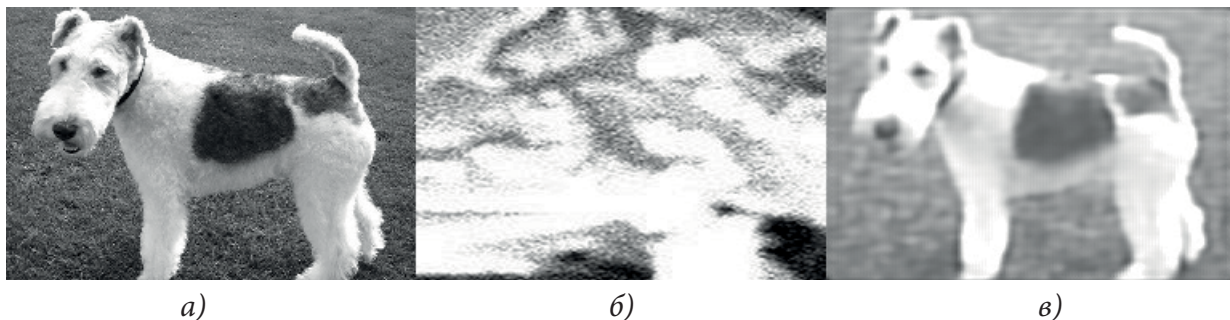


Рис. 3. Частичная стилизация изображения. а) — исходное изображение, б) — накладываемый стиль, в) — стилизованное изображение.

[Fig. 3. Partial image stylization. a) — original image, b) — style, c) — stylized image]

сокращения: МФМО — медианный фильтр с последующей морфологической операцией открытие, АУФ — адаптивный усредняющий фильтр. Из полученных результатов видно, что нейронная сеть не уступает по метрикам классическим алгоритмам. За основу последнего был взят адаптивный усредняющий фильтр [2], а дисперсия шума находилась при помощи метода усреднения дисперсии в блоках [17]. Контрастность изображения определялась как среднее квадратичное отклонение всех пикселей изображения.

Все эксперименты проводились с 500 случайными изображениями из датасета ImageNet. Использовались различные способы искажений, описанные выше.

Исходное и восстановленное нейросетевым алгоритмом изображение представлено на рис. 4. В качестве искажений использовались импульсные помехи, сгенерированные вторым способом. В некоторых случаях можно разглядеть низко контрастные белые пятна на восстановленном изображении, однако удалось при этом сохранить его хорошее качество и мелкие детали. При обучении нейронной сети для восстановления изображения

от импульсных помех решено было добавить дополнительную функцию потерь: $\max D$ — максимальное абсолютное отклонение между пикселями зашумленного и восстановленного изображениями.

Отдельно было проведено исследование зависимости MSE от среднеквадратического отклонения, используемого для добавления аддитивного белого гауссовского шума. Использовались алгоритмы, не требующие априорной информации об источнике шума.

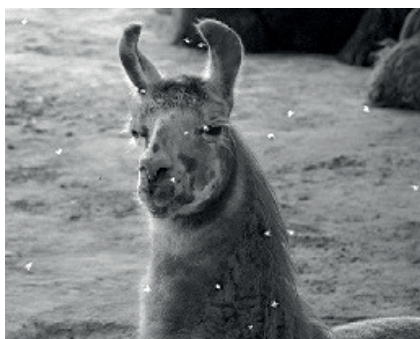
При зашумлении использовалось 500 различных изображений из датасета ImageNet. MSE рассчитывался для каждого изображения отдельно, а потом производилось усреднение.

На рис. 5 показано, что нейронная сеть лучше справляется с восстановлением исходных изображений, чем алгоритм АУФ. Однако MSE всё равно уменьшается с ростом среднеквадратического отклонения, так как уменьшается соотношение сигнал-шум и затрудняется восстановление исходного изображения.

В табл. 2 представлено время работы алгоритмов в секундах. Измерения проводились

Таблица 1. Сравнение метрик качества восстановления изображений
[Table 1. Comparison of image recovery quality metrics]

Алгоритм	PSNR	SSIM	Контрастность
REDNet	21.0	0.81	0.20
МФМО	27.1	0.73	0.216
АУФ	25.83	0.63	0.21
Фильтр Винера	20.17	0.71	0.236



а)



б)

Рис. 4. Восстановление изображения от импульсных помех.

а) — зашумленное изображение, б) — восстановленное изображение
[Fig. 4. Image recovery from pulse noise. a) — noisy image, b) — restored image]

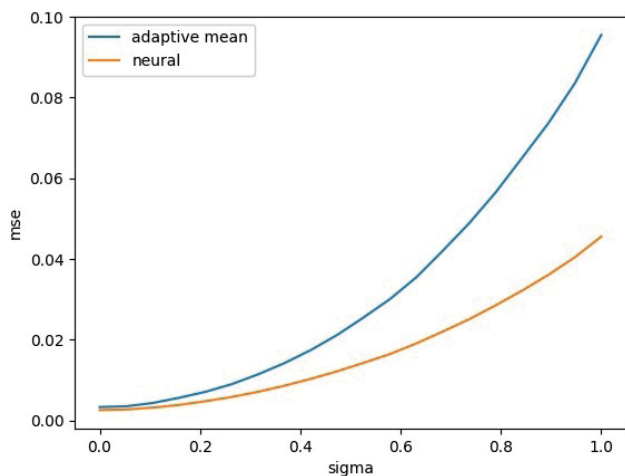


Рис. 5 График зависимости среднеквадратичной ошибки восстановления от стандартного отклонения [Fig. 5. Graph of the dependence of the MSE on the standard deviation]

на процессоре Intel Core i7 10 поколения с 16 гигабайтами оперативной памяти. Каждый из алгоритмов запускался несколько сотен раз и вычислялось среднее время его работы. Все измерения проводились с помощью библиотеки timeit. Стоит отметить, что в алгоритмах МФМО и фильтра Винера все тяжеловесные операции проведены с помощью библиотеки OpenCV, включающей в себя C++ код, что дало значительное увеличение производительности.

Таблица. 2 Время работы алгоритмов в секундах

[Table 2. The running time of the algorithms in seconds]

Алгоритм	Время работы (с)
REDNet	2.29
МФМО	0.13
АУФ	2.75
Фильтр Винера	0.8

Из табл. 2 видно, что применение нейросетевого алгоритма имеет ограничения для применения в режиме реального времени и с трудом подходит для восстановления изображений в видеопотоке.

На заключительном этапе было проведено обучение сети Unet для объективной проверки качества восстановления путем сегмента-

ции изображений. Каждый раз веса нейронной сети инициализировались одинаковым образом, чтобы можно было оценить эффективность её работы. В качестве датасета был выбран изначально зашумлённый набор медицинских снимков сетчатки глаза. В качестве искомой метрики была использована F1-мера, так как присутствовал дисбаланс классов. В результате удалось добиться F1-меры 0.81 на исходных изображениях и 0.85 на восстановленных. Таким образом, нейросетевой алгоритм, основанный на REDNet, сделал изображения более качественными, по крайней мере с точки зрения нейронной сети Unet.

ЗАКЛЮЧЕНИЕ

В работе был проведен сравнительный анализ существующих классических и нейросетевых алгоритмов. Исследованы различные способы зашумления изображений и эффективные методы их восстановления, в том числе генерация импульсных помех, засветки, двоящихся границ.

Использованы различные метрики для сравнения качества изображения, а также применена архитектура Unet в задаче сегментации для оценки качества восстановленных изображений. Результаты исследования показали, что на восстановленных изображениях сеть смогла добиться лучшей F1-меры.

В тоже время было обнаружено, что нейронная сеть ведет себя недостаточно точно при наличии больших аппликативных помех. Целью дальнейшего исследования будет восстановление изображения от этого вида помех.

Другим недостатком предложенной архитектуры нейронной сети является её неприемлемость для потоковой обработки видео в реальном режиме времени. В дальнейшем также планируется усовершенствовать и оптимизировать текущую архитектуру нейронной сети.

КОНФЛИКТ ИНТЕРЕСОВ

Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

СПИСОК ЛИТЕРАТУРЫ

1. *Zuo, W.* Texture Enhanced Image Denoising via Gradient Histogram Preservation / W. Zuo, L. Zhang, C. Song, D. Zhang // CVPR. – 2013. – DOI: 10.1109/CVPR.2013.159.
2. *Гонсалес, Р.* Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. – М. : Техносфера, 2012. – 1104 с.
3. *Тихонов, В. И.* Статический анализ и синтез радиотехнических устройств и систем связи / В. И. Тихонов, В. Н. Харисов. – М. : Радио и связь, 2004. – 608 с.
4. *Клочко, В. К.* Методы восстановления изображений и оценивания аппаратной функции по прореженной матрице наблюдений / В. К. Клочко, В. П. Кузнецов // Автометрия. – 2016. – Т. 52. – №6. – С. 12–20. – DOI: 10.15372/AUT20160602.
5. *Zhang, J.* Quantile analysis of image sensor noise distribution / J. Zhang, K. Hirakawa, X. Jin // ICASSP. – 2015. – DOI: 10.1109/ICASSP.2015.7178240.
6. *Ваняшкин, Ю. Ю.* Применение автокодировщиков для устранения шумов с изображений / Ю. Ю. Ваняшкин, Д. А. Макаров // Научно-образовательный журнал для студентов и преподавателей «StudNet». – 2020. – № 10. – 8 с.
7. *Xiao-Jiao, M.* Image Restoration Using Very Deep Convolutional Encoder-Decoder Networks with Symmetric Skip Connections / M. Xiao-Jiao, S. Chunhua, Y. Yubin // NIPS. – 2016.
8. *Deya, B.* SEM image denoising with Unsupervised Machine Learning for better defect inspection and metrology / B. Deya, S. Haldera, K. Khalil // SPIE Advanced Lithography. – 2021. – DOI: 10.1117/12.2584803.
9. *Wang, Z.* Defect simulation in SEM images using generative adversarial networks / Z. Wang, Y. Liangjiang, P. Lingling // SPIE Advanced Lithography. – 2021. – DOI: 10.1117/12.2581881.
10. *Huang, X.* Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization / X. Huang, S. Belongie // arXiv: 1703.06868. – 2017.
11. *Сизиков, В. С.* Устойчивые методы обработки результатов измерений / В. С. Сизиков. – Санкт-Петербург : СпецЛит, 1999. – 240 с.
12. *Бондина, Н. Н.* Адаптивные алгоритмы фильтрации и изменения контраста изображения / Н. Н. Бондина, Р. Ю. Мураров // Вестник НТУ. – 2014. – № 35. – 8 с.
13. *Milukova, O.* Image Restoration Spectral Techniques/ O. Milukova, V. Kober, I. A. Ovsseevich // PRIP. – 2009. – 4 p.
14. *Алгазинов, Э. К.* Анализ и компьютерное моделирование информационных процессов и систем / Э. К. Алгазинов, А. А. Сирота. – М. : Диалог-МИФИ, 2009. – 416 с.
15. *Сирота, А. А.* Анализ потенциальных и реальных характеристик оценивания случайных полей (изображений) в условиях аддитивных и импульсных помех / А. А. Сирота, П. В. Калинин // Вестник ВГУ. Серия: Системный анализ и информационные технологии. – 2011. – № 1. – С. 41–50.
16. *Huynh-Thu, Q.* Scope of validity of psnr in image/video quality assessment / Q. Huynh-Thu and M. Ghanbari // Electronics letters. – 2013. – № 44 – P. 800–801. – DOI:10.1049/el:20080522.
17. *Калинкина, Д. А.* Определение уровня шума на изображении на основе усреднения дисперсии в блоках / Д. А. Калинкина // Международная конференция студентов и аспирантов по фундаментальным наукам «Ломоносов 2005».

Сирота Александр Анатольевич — д-р техн. наук, проф., заведующий кафедрой технологий обработки и защиты информации, факультет компьютерных наук, Воронежский государственный университет. E-mail: sir@cs.vsu.ru
ORCID iD: <https://orcid.org/0000-0002-5785-8513>

Бережнов Никита Игоревич — студент 2 курса магистратуры кафедры технологий обработки и защиты информации, факультет компьютерных наук, Воронежский государственный университет. E-mail: beregnochnikita@yandex.ru
ORCID iD: <https://orcid.org/0000-0002-3532-1002>

UNIVERSAL IMAGE ENHANCEMENT ALGORITHM USING DEEP NEURAL NETWORKS

© 2022 N. I. Berezhnov✉, A. A. Sirota

Voronezh State University

1, Universitetskaya Square, 394018 Voronezh, Russian Federation

Annotation. The article considers the problem of image quality improvement under the influence of various types of noise and distortion. A comparative analysis of standard algorithms and neural network algorithms for improving image quality based on deep neural networks has been carried out. The neural network algorithms are considered to be a universal mean for improving image quality. This article also offers the investigation of the influence of hyperparameters of deep neural networks on the quality of reconstructed images. The first part deals with theoretical issues of the image restoration problem which is regarded as a solution to a complex, unstable optimization task. Various augmentation techniques based on the application of methods of forced image noise and artificial defect generation are also considered. Finally, we introduce a new approach for data augmentation by partial image stylization. In addition, we use various metrics for assessing the quality of restored images, determine their disadvantages and limits of their applicability. In the second part of the article, the well-known classical algorithms are compared with the proposed neural network algorithms, and the dependence of image recovery quality from noise level is investigated. We evaluate the effectiveness of the proposed approaches for image improvement, consider additive and pulse types of noise, which can be found on medical images, as well as on images obtained using optical coherence tomography. The possibility of algorithms working in real time is also investigated. We propose the method of comparing the quality of images through the segmentation task by using the Unet network. The research procedures have enabled us to conclude that neural networks are not inferior to classical algorithms in improving image quality, and in some cases even outperform them.

Keywords: image processing, image enhancement, neural networks, data augmentation, image quality, additive and pulse noise.

CONFLICT OF INTEREST

The authors declare the absence of obvious and potential conflicts of interest related to the publication of this article.

REFERENCES

1. Zuo W., Zhang L., Song C. and Zhang D. D. (2013) Texture Enhanced Image Denoising via Gradient Histogram Preservation. 2013 *IEEE Conference on Computer Vision and Pattern Recognition*. P. 1203–1210. Available from: doi: 10.1109/CVPR.2013.159.
2. González R. C. and Woods R. E. (2008) Digital image processing, 3rd Edition.
3. Tihonov V. I. and Harisov V. N. (2004) Staticheskiy analiz i sintez radio-tehnicheskikh ustroystv i sistem svyazi. *Moscow, Radio i svyaz*. (in Russian)
4. Klochko V. K. and Kuznecov V. P. (2016) Metody vosstanovleniya izobrazhenij i ocele-nivaniya apparatnoj funkicii po prorezhen-noj matriche nabljudenij. *Avtometrija*. 52, (6). P. 12–20. Available from: doi: 10.15372/AUT20160602. (in Russian)
5. Zhang J., Hirakawa K. and Jin X. (2015) Quantile analysis of image sensor noise distribution. 2015 *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. P. 1598–1602. Available from: doi: 10.1109/ICASSP.2015.7178240.
6. Vanjashkin Ju. Ju. and Makarov D. A. (2020) Primenenie avtokodirovshhikov dlja ustraneniya

✉ Berezhnov Nikita I.
e-mail: berezhnovnikita@yandex.ru

shumov s izobrazhenij. *Nauchno-obrazovatel'nyj zhurnal dlja studentov i prepodavatelej «StudNet»*. (10). (in Russian)

7. Mao X., Shen C. and Yang Y. (2016) Image Restoration Using Very Deep Convolutional Encoder-Decoder Networks with Symmetric Skip Connections. *NIPS*.

8. Dey B., Halder S., Khalil K., Lorusso G. F., Severi J., Leray P. and Bayoumi M. A. (2021) SEM image denoising with unsupervised machine learning for better defect inspection and metrology. *Advanced Lithography*. Available from: doi: 10.1117/12.2584803.

9. Wang Z., Yu L. and Pu L. (2021) Defect simulation in SEM images using generative adversarial networks. *Advanced Lithography*. Available from: doi: 10.1117/12.2581881.

10. Huang X. and Belongie S. J. (2017) Arbitrary Style Transfer in Real-Time with Adaptive Instance Normalization. *2017 IEEE International Conference on Computer Vision (ICCV)*. P. 1510–1519.

11. Sizikov V. S. (1999) Ustojchivye metody obrabotki rezul'tatov izmerenij. *Saint-Petersburg, SpecLit*. (in Russian)

12. Bondina N. N. and Murarov R. Ju. (2014) Adaptivnye algoritmy fil'tracii i izmenenija kontrasta izobrazhenija. *Vestnik NTU*. (35). (in Russian)

13. Milukova O., Kober V. and Ovseevich I. A. (2009) Image Restoration Spectral Techniques. *PRIP*.

14. Algazinov Je. K. and Sirota A. A. (2009) Analiz i komp'juternoe modelirovanie informacionnyh processov i sistem. *Moscow, Dialog-MIFI*. (in Russian)

15. Sirota A. A. and Kalinin P. V. (2011) Analiz potencial'nyh i real'nyh harakteristik ocenivaniya sluchajnyh polej (izobrazhenij) v uslovijah additivnyh i im-pul'snyh pomeh. *Vestnik VGU Serija: Si-stemnyj analiz i informacionnye tehnologii*. (1). P. 41–50. (in Russian)

16. Huynh-Thu Q. and Ghanbari M. (2008) Scope of validity of PSNR in image/video quality assessment. *Electronics Letters*. 44. P. 800–801. Available from: doi: 10.1049/el:20080522.

17. Kalinkina D. A. Opredelenie urovnja shuma na izobrazhenii na osnove usrednenija dispersii v blokah. Mezhduna-rodnaja konferencija studentov i aspirantov po fundamental'nym naukam «Lomonosov 2005». (in Russian)

Sirota Alexander A. — DSc in Technical Sciences, Head of the Department of Information Security and Processing Technologies, Faculty of Computer Sciences, Voronezh State University.

E-mail: sir@cs.vsu.ru

ORCID iD: <https://orcid.org/0000-0002-5785-8513>

Berezhnov Nikita I. — 2nd year master's student, Department of Information Security and Processing Technologies, Faculty of Computer Sciences, Voronezh State University.

E-mail: bereznovnikita@yandex.ru

ORCID iD: <https://orcid.org/0000-0002-3532-1002>