

1 Аналитический раздел

В этом разделе будет произведен анализ предметной области: описаны основные понятия в области удаления шумов из цветных изображений. Будут рассмотрены существующие методы удаления импульсных шумов из цветных изображений, приведены результаты сравнительного анализа. Будет сформирована цель работы, постановка задачи формализована в виде IDEF0-диаграммы.

1.1 Анализ предметной области

Шум – дефект изображения, в основе которого лежит эффект появления на фотографии пикселей случайного цвета на изображении [1].

Причины возникновения такого эффекта делятся на два типа: естественные и искусственные. Основным источником естественных помех на изображениях является фотосенсор [1]. Существует несколько физических объяснений появления шума на изображении [2]:

1. При дефектах потенциального барьера происходит утечка заряда. В этом случае шум на изображениях проявляется в виде темных точек на светлом фоне.
2. При подаче потенциала на электрод может возникнуть темновой ток, который отображается на картинке в виде светлых точек на темном фоне. Основная причина возникновения темнового тока — это примеси в кремниевой пластине или повреждение кристаллической решетки кремния.

Также шум на изображениях может быть вызван умышленным вмешательством человека или состязательной атакой [3]. Состязательная атака – это манипуляция обучающими данными, архитектурой модели или манипулирование тестовыми данными таким образом, что это приведет к неправильному выходу из модели машинного обучения [3].

Существует несколько основных типов естественных шумов, возникающих на фотографиях [4]. От точного определения характеристики шума зависит то, какой метод требуется выбрать для автоматического определения дефектных

пикселей на изображении и последующего его устранения. Для выявления причины требуется понять, на каком устройстве была сделана фотография, и через какие этапы обработки она прошла. При этом для реальных цветных фотографий результирующий шум является комбинацией всех возможных типов дефектов, соответственно, нельзя точно сказать о происхождении каждого дефекта на рассматриваемой фотографии [5].

1.1.1 Гауссов шум

Так как квантовым процессам свойственна случайность, то такие процессы можно отнести к Гауссовым, следовательно, они обладают следующим свойством: распределение суммы независимых случайных величин сходится к нормальному, вне зависимости от характера распределения слагаемых [6].

Пусть I – интенсивность изначального пикселя, а ν – интенсивность шума, распределенная по нормальному распределению. Тогда интенсивность загрязненного пикселя можно представить по формуле (1) [4]:

$$I_f = I + \nu, \nu \sim N(0, \sigma^2). \quad (1)$$

Вид гауссова шума представлен на рисунке 1:

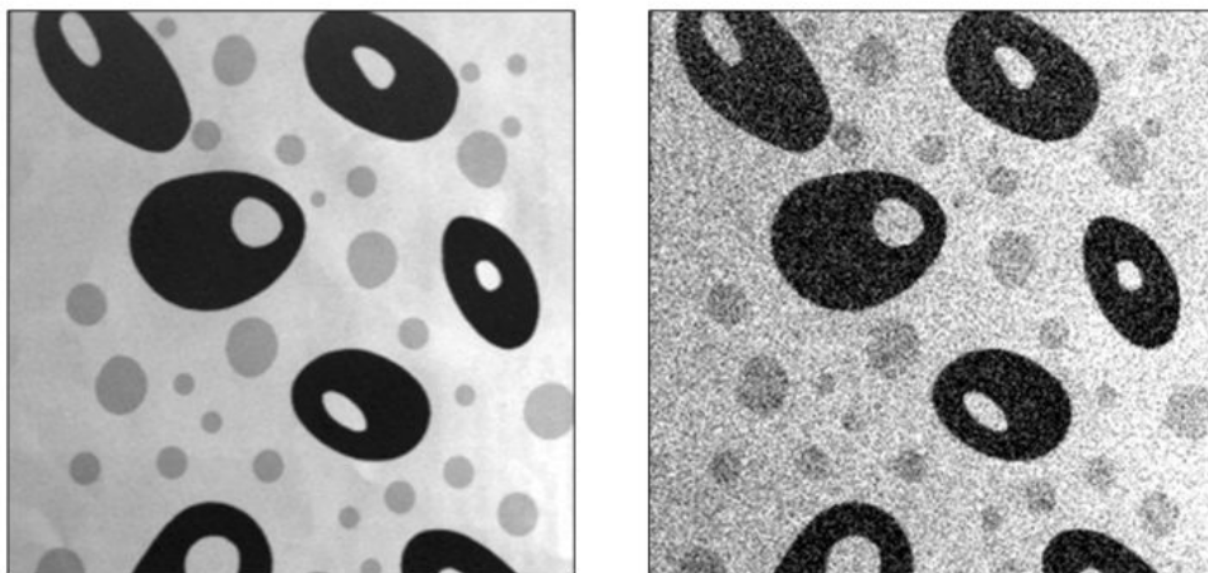


Рисунок 1: Вид гауссова шума

Причиной возникновения таких шумов как правило является некорректная

работа фотосенсора, что часто происходят в условиях ночного освещения или высокой температуры. Именно этот вид шумов на практике встречается чаще всего [4].

1.1.2 Импульсный шум

Импульсный шум проявляется в том, что на изображениях в случайных местах появляются черные и белые пиксели [7]. Основной причиной их возникновения является темновой ток и утечка заряда в фотосенсоре, а также наличие пикселей с дефектами [1].

Когда пиксель имеет белый цвет на темном фоне, то говорят, что здесь присутствует шум соли. Шум перца – обратная ситуация: на светлом изображении заметны темные точки.

Описать появление шума можно по формуле (2) [7]:

$$P(S_{i,j} = 1) = p, \quad (2)$$

где S – исходное изображение, i, j – координаты пикселя, p – вероятность появления шума.

Вероятность появления шума зависит от характеристик фотоаппарата, а также от внешних условий, поэтому точно рассчитать значение невозможно.

Вид шума соли и перца представлен на рисунке 2:

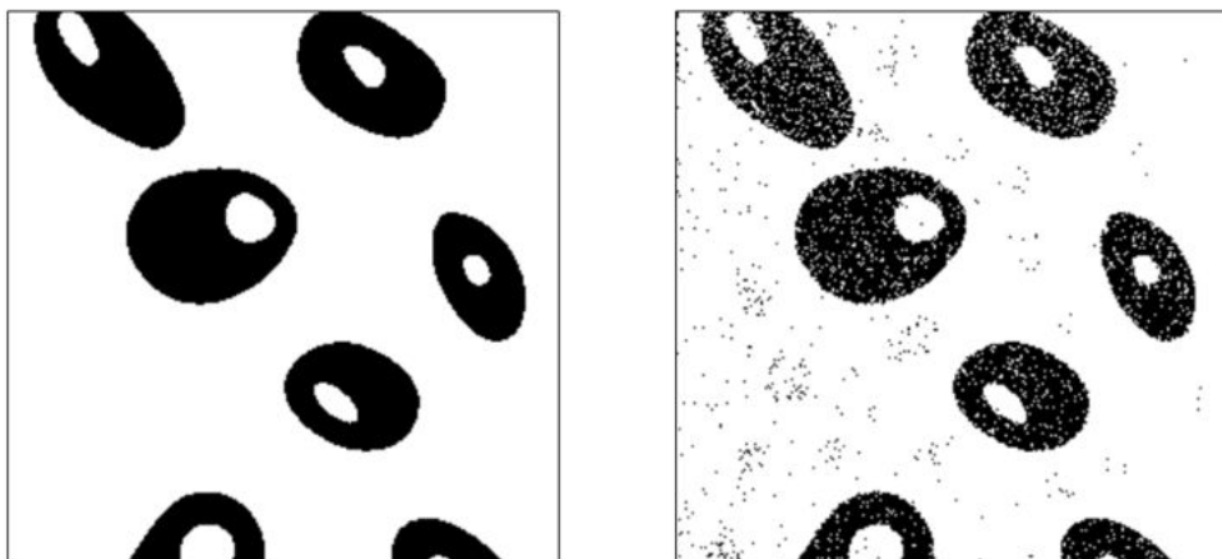


Рисунок 2: Шум соли и перца

Основным методом борьбы с таким видом шумов является медианный фильтр и сверточные нейронные сети.

1.2 Нейронные сети

Нейронная сеть – математическая модель, основанная на принципах функционирования биологических нейронных сетей – например, нервных клеток человека.

Требуется ввести следующие основные определения в математическом контексте нейронной сети:

1. **Нейрон** – вычислительная единица, получающая на вход информацию, обрабатывает ее, и возвращает результат на выходе.
2. **Синапс** – связь между двумя нейронами, построенная по следующему принципу: выход одного нейрона является входом для другого нейрона. У каждого синапса есть один параметр – вес, который влияет на результат вычислений в нейроне.
3. **Слой** – множество нейронов, которые получают на вход информацию от всех нейронов из предыдущего слоя.

Все нейроны можно разделить на три больших типа:

1. **Входной** – он получает на вход исходный вектор, которую требуется обработать в нейронной сети, и без предварительной обработки передаёт на вход следующему слою. Графически принято обозначать входной нейрон синим цветом.
2. **Выходной** – значение, полученное на выходе нейрона, считается итоговым результатом всех вычислений нейронной сети. На схемах для обозначения выходного нейрона используется зеленый цвет.
3. **Скрытый** – нейрон, который не является ни входным, ни выходным, соответственно, ему на вход поступает информация от других нейронов, и его выход также переходит другим нейронам. Скрытые нейроны являются основным вычислительным центром всей сети. Выделяются красным цветом.

Математическая модель одного нейрона представляется формулой (3):

$$Y = f\left(\sum_{i=1}^N w_i * x_i + w_0 * x_0\right), \quad (3)$$

где Y – выход нейрона, x_i – выход i -го нейрона предыдущего слоя, w_i – вес синапса, связывающего x_i нейрон и текущий, x_0 и w_0 – дополнительный вход и его вес, которые используются для инициализации нейронной сети, f – передаточная функция.

Пусть x – значение из входного вектора, y – результат вычислений нейрона. На эти значения накладываются ограничения, указанные в формуле (4):

$$x \in [0, 1]; y \in [0, 1]. \quad (4)$$

Передаточная функция является основным параметром нейронной сети. Как правило, она обладает следующими ключевыми свойствами:

- передаточная функция f монотонно возрастает:
 $\forall(x_i \leq x_j) : f(x_i) \leq f(x_j);$
- область значений функции f лежит в интервале $[0, 1]$: $D(f) = [0, 1];$

Передаточная функция позволяет добавлять нелинейность нейронной сети. Существует несколько основных типов передаточных функций, используемых в нейронных сетях:

1. Линейная передаточная функция

В этом случае информация на выходе линейно зависит от результата суммирования входных сигналов:

$f(x) = k * x$, где k – коэффициент. В основном она используется для входного слоя нейронной сети, причём часто $k = 1$.

2. Шаговая передаточная функция

 Алгоритм работы этой функции пред-

ставлен на формуле (5):

$$f(x) = \begin{cases} 1, & \text{если } x \geq 1 \\ 0, & \text{если } x \leq 0 \\ x, & \text{иначе.} \end{cases} \quad (5)$$

Самая распространенная вариация функции – исправленная линейная передаточная функция или **ReLU** [8]. Математически она описывается формулой (6) [8]:

$$f(x) = \begin{cases} x, & \text{если } x \geq 0 \\ 0, & \text{иначе.} \end{cases} \quad (6)$$

Она получила широкое распространение в сверточных нейронных сетях из-за простоты использования [9].

3. Сигмоидальная передаточная функция

Эта передаточная функция позволяет усиливать слабые сигналы значительно заметнее, чем сильные, поэтому область сильных сигналов представлена пологими участками [8].

На формуле (7) представлен вид сигмоидальной передаточной функции [9]:

$$f(x) = \frac{1}{1 + \exp(-tx)}, \quad (7)$$

где t – параметр функции, регулирующий крутизну функции.

4. Гиперболическая передаточная функция

Эта передаточная функция использует операцию гиперболического тангенса для расчёта итогового значения [8]. По свойствам она похожа на сигмоиду, но слабые сигналы усиливаются еще сильнее.

Формула гиперболической передаточной функции (8):

$$f(x) = \frac{2}{1 + \exp(-2x)} + 1. \quad (8)$$

Перед тем, как нейросеть начнет вычислять результат, требуется провести ее обучение.

Для этого используется метод обратного распространения ошибки.

Предположим, для заданного сигнала заранее известен вектор результатов Y_0 , который ожидается на выходе нейронной сети. Нейронная сеть после подачи на вход сигнала выдала вектор Y_1 .

Функция ошибки для отдельного синапса, полученная методом наименьших квадратов, имеет следующий вид (9):

$$Ew = \frac{1}{2} * \sum N_i = 0(Y_{0i} - Y_{1i})^2, \quad (9)$$

Для проведения обучения нейронной сети используется стохастический градиентный спуск. Требуется изменять веса после вычислений результат для каждого входящего нейрона.

Добавленное значение к весу представлено на формуле (10):

$$\Delta w_{ij} = -\eta * \frac{\partial E}{\partial w_{ij}}, \quad (10)$$

где w_{ij} – вес синапса между нейронами i и j , η – коэффициент, влияющий на скорость обучения.

Суть обучения нейронной сети состоит в минимизации функции ошибки для отдельного входящего сигнала. С помощью корректировки коэффициентов синапсов сеть пытается достигнуть точки сходимости или локального минимума. Именно для этого используется градиентный спуск.

1.2.1 Сверточная нейронная сеть

Сверточная нейронная сеть – один из подвидов нейронных сетей, получивший широкое распространение в области обработки изображений [10].

В контексте СНС нужно ввести следующие определения:

- **Ядро свертки** – квадратная матрица весов, используемая для операции свертки [11]. Ее размер, как правило, определяется программистом при инициализации сети.
- **Свертка** – математическая операция, в результате которой уменьшается

размерность вектора исходного сигнала [12]. Ее можно описать следующей формулой (11):

$$x_i^l = f(\sum_j x_i^{l-1} * k_j^l + b_j^l), \quad (11)$$

где x_i^l – выходной сигнал слоя l , f – передаточная функция, x_i^{l-1} – входной сигнал для входа l , k_j^l – ядро свертки слоя l , b_j^l – коэффициент сдвига для слоя l .

- **Карта признаков** – результат, полученный после операции свертки, который отражает наличие какого-либо признака во входном сигнале [12].

Процесс обучения сверточной нейронной сети осуществляется методом обратного распространения ошибки [12]. В сверточной нейронной сети несколько ядер свертки, которые отвечают за отдельные признаки, но такое разбиение возникает в результате обучения – программист не предусматривает это при проектировании архитектуры сети. Таким образом, каждый набор весов формирует собственную карту признаков, в результате возникает явление многоканальности в нейронной сети: на одном слое появляются независимые карты признаков [13]. Во время выполнения операции свертки ядро матрицы передвигают не на полный шаг, а на маленькое расстояние – несколько пикселей, в результате искомый признак отражается в нескольких весах выходного вектора.

Так как в результате свертки итоговая размерность векторов уменьшается, то величина сигнала стремится к размерам ядра свертки. Последовательность таких действий называется *субдискретизацией* [13].

В качестве передаточной функции чаще всего используют ReLu и сигмоида [13].

Архитектура сверточной нейронной сети состоит из нескольких компонентов [10]:

- **Начальный слой** – принимает на вход исходное изображение;
- **Сверточные слои** – состоят из ядер свертки, каждый слой выполняют субдискретизацию полученных карт признаков;
- После прохождения субдискретизации карты признаков вырождаются до

векторов или скаляров. В дальнейшем они попадают на нейроны из полносвязной нейронной сети, которая рассчитывает итоговый результат.

При прохождении через сверточные слои размерность каждого слоя уменьшается, в то время как количество каналов – растет. Это позволяет в дальнейшем использовать полносвязную нейронную сеть для определения результата [12].

Схематично архитектура сверточной нейронной сети представлена на рисунке 3:

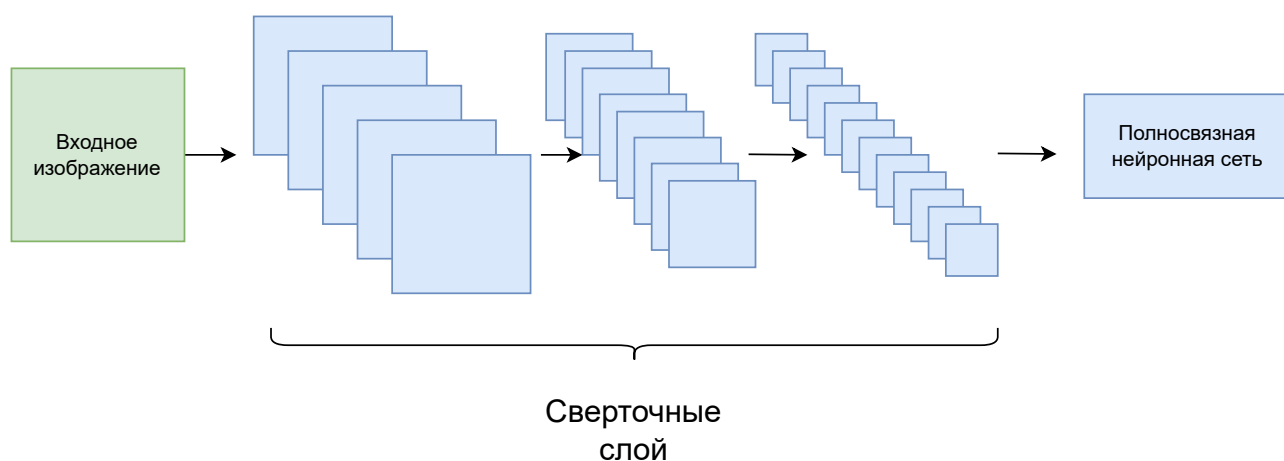


Рисунок 3: Архитектура сверточной нейронной сети

Во время обучения принято на некоторых итерациях убирать из сети одиночные нейроны, что позволяет точнее настроить другие нейроны для более точного обучения [14].

Признакам, сформированные в результате работы нейронной сети, невозможно дать объяснение, поэтому в случае, если нейронная сеть перестает выполнять поставленную задачу, разработчики изменяют ее конфигурацию или выбирают другой этап применения в общем алгоритме [11].

На текущий момент не существует архитектур нейросетей, которые бы в точности позволяли бороться с шумами в изображениях, поэтому подходящие настройки сети подбираются экспериментально [15].

Нейронные сети активно применяются в области удаления шумов из изображений, так как они позволяют распознать и удалить шумы, с которыми не справляются стандартные методы. Для трехмерных изображений ядра матрицы составляются для каждого из составляющих цветов пикселя, итоговые резуль-

таты по цветам суммируется [14]. Это позволяет менять вариативность нейронной сети, например, настраивать на более приоритетный цвет.

1.3 Обзор существующих методов

Шумы искажают исходную картинку и портят ее качество так, что это способен распознать человеческий глаз. Однако могут возникнуть трудности в обнаружении помех, поскольку они трудно различимы при совпадении цвета фона и цвета пикселя, например, светлые точки будут плохо заметны на ярком фоне.

Было разработано несколько алгоритмов, которые производят бинарную классификацию пикселей и определяют, какие из них можно идентифицировать как шумы и затем их устранить.

В качестве классификации алгоритмы можно разделить на два типа:

1. **Изотропная фильтрация** – такие методы устраняют помехи, но не учитывают детали пикселя и увеличивают размытость.
2. **Анизотропная фильтрация** – алгоритмы устраняют эффекты сглаживания, уменьшают размытость и сохраняют детали пикселя, устраняя при этом непосредственно шум из изображения.

1.4 Общий алгоритм работы фильтров

Алгоритмы, анализирующие наличие шумов в изображениях, имеют дело с различными характеристиками одного пикселя. Например, цвет пикселя можно разбить на три составляющие – синюю, красную и зеленую [16].

В таком случае метод работает с каждой из составляющих пикселя, вычисляя новое значение для каждой характеристики. Результат работы в этом случае является объединением подсчетов по всем характеристикам.

Общая схема работы алгоритмов представлена на рисунке 4:

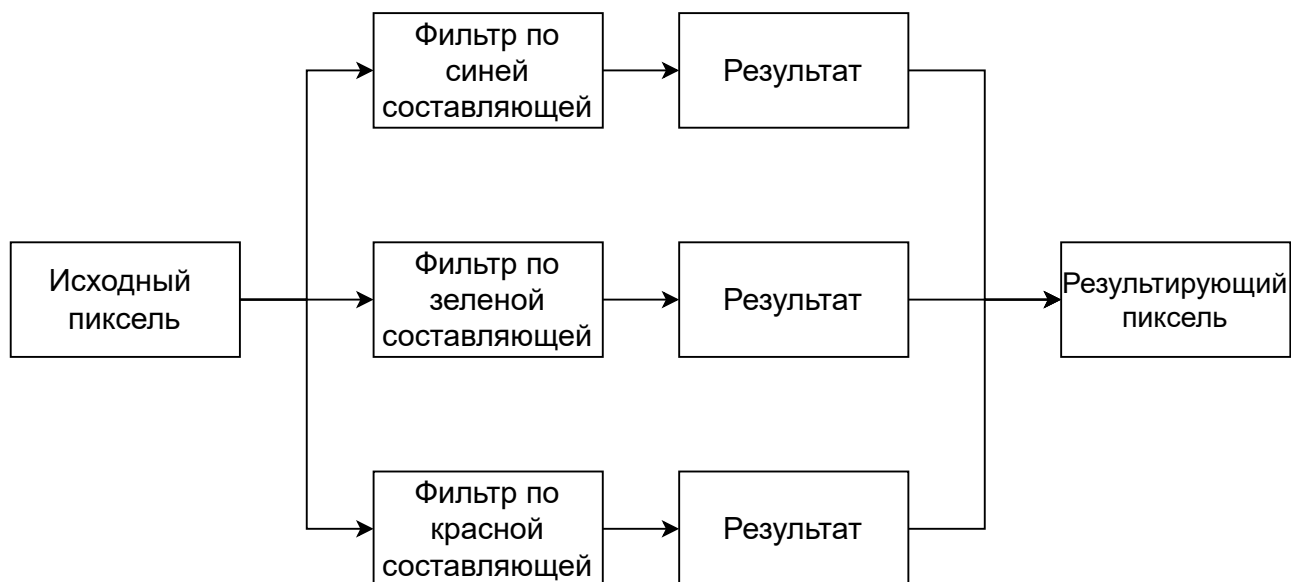


Рисунок 4: Общая схема работы всех алгоритмов

Каждый из фильтров, перечисленный ниже, работает с каждым из параметров пикселя одинаково, поэтому для корректной работы алгоритмов требуется вычислить значение каждого свойства для результирующего пикселя [4].

1.4.1 Медианный фильтр

Под медианным фильтром понимается семейство одностипных алгоритмов, относящихся к классу нелинейных фильтров [17].

Метод работает в цикле с каждым пикселем изображения [4]. В окрестности каждого пикселя находится восемь соседних, каждый обладает собственными свойствами.

На рисунке 5 изображена сетка, с которой работает алгоритм:



Рисунок 5: Рассматриваемая сетка пикселей при работе алгоритма

Пусть $C_{i,j}$ – один из параметров рассматриваемого пикселя, а Ω – все пиксели сетки. Алгоритм подсчитывает медиану от такого же параметра соседних клеток и заменяет параметр пикселя на значение этой медианы [17]. Итоговое значение можно посчитать по формуле (12):

$$C_{i,j} = median(\Omega_i). \quad (12)$$

Схема работы алгоритма изображена на рисунке 6:

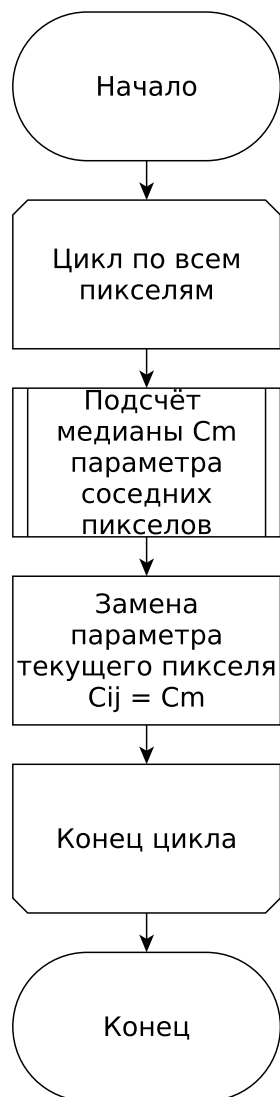


Рисунок 6: Схема работы алгоритма медианного фильтра

К недостаткам данного метода можно отнести то, что алгоритм может убрать значительные детали из изображения, посчитав их за шум [18].

1.4.2 Гауссовский фильтр

Работа алгоритма гауссовского фильтра также зависит от значений цветовых свойств пикселей в сетке, рассмотренной на рисунке 5.

В этом случае для каждого соседнего рассчитывается вес, с которым он влияет на новое значение рассматриваемого пикселя. Пусть d – расстояние до центрального пикселя сетки, σ – стандартное отклонение, подсчитанное для всех значений определенного параметра текущей сетки. Тогда вес w пикселя рассчитывается по формуле (13) [4]:

$$w_{ij} = \exp \left(\frac{-d^2}{2\sigma^2} \right). \quad (13)$$

Подсчитав вес для каждого пикселя в сетки, можно рассчитать новое значение свойства рассматриваемого пикселя по формуле 14:

$$p_i = \frac{1}{\sum_{j \in \Omega} w_{ij}} * \sum_{j \in \Omega} w_{ij} * p_j. \quad (14)$$

Схема алгоритма гауссовского фильтра представлена на рисунке 7:



Рисунок 7: Схема работы алгоритма гауссовского фильтра

1.4.3 Билатеральный фильтр

Алгоритм билатеральной фильтрации является улучшением метода Гауссовского фильтра [19].

Для каждого пикселя сетки соседних пикселей используется сразу два веса: один аналогичный параметру из исходного алгоритма, а второй отвечает за анизотропную составляющую.

Расчет веса w_s , отвечающего за изотропную составляющую, происходит по формуле (13). Коэффициент, регулирующий анизотропные свойства фильтрации, рассчитывается по формуле (15) [19]:

$$w_r = \exp\left(\frac{-|p_i - p_j|}{2\sigma^2}\right). \quad (15)$$

В таком случае результат работы некоторого пикселя можно посчитать по формуле (16):

$$p_i = \frac{1}{\sum_{j \in \Omega} w_s w_r} * \sum_{j \in \Omega} w_s w_r p_j. \quad (16)$$

Схема работы алгоритма представлена на рисунке 8:



Рисунок 8: Схема работы алгоритма билатеральной фильтрации

1.4.4 Алгоритм Цзяньвэй

Этот алгоритм был описан в 2014 году индонезийским ученым Ван Цзяньвэй [20]. Алгоритм позволяет эффективно убирать шумы соли и перца даже в случае сильной загрязненности изображения.

Процедура заключается в обходе всех пикселей фотографии в заданном порядке и определении того, соответствуют ли значения пикселей функции плотности вероятности импульсного шума или нет. Если пиксель на первом классифицируется как шум, то подсчитывается количество импульсного шума в маске определенной формы. Результатом операции маски является замена значения пикселя. В противном случае это не рассматривается как шум, значение пикселя остается неизменным.

Схема алгоритма Цзяньвэй представлена на рисунке 9:

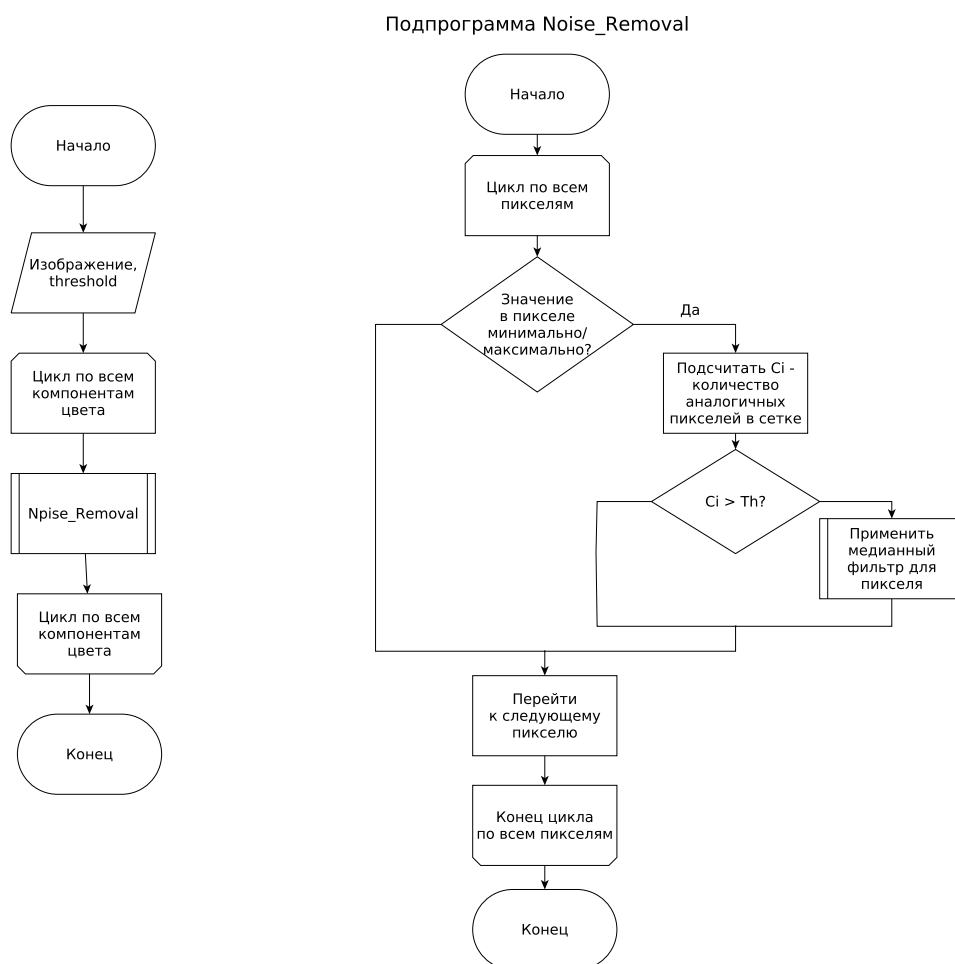


Рисунок 9: Схема работы алгоритма Цзяньвэй

1.4.5 DnCNN

Алгоритмы, основанные на сверточных нейронных сетях, получили широкое распространение в области удаления шумов из изображений. Они применяются в ситуациях, требующих восстановления исходного изображения по имеющемуся загрязненному фото, например, при анализе снимков из телескопов [15]. DnCNN относится к этому классу методов.

Перед запуском алгоритма требуется задать суммарную глубину слоев D , а также количество цветовых составляющих c – для цветных изображений $c = 3$, для серых – $c = 1$ [21]. Всего в конфигурации нейронной сети используется три типа слоев [22]:

1. Первый слой состоит из 64 фильтров размером $3 \times 3 \times c$, в качестве функции активации используется ReLU.
2. Следующие $2 * (D - 1)$ слоев состоят из фильтров размером $3 * 3 * 64$, между каждым слоем применяется пакетная нормализация.
3. Последний слой состоит из c фильтров размером $3 * 3 * 64$ каждый.

Во всех слоях, кроме последнего, используется ReLU в качестве функции активации нейрона [21].

Пусть x – входное значение нейрона.

Тогда выход считается по формуле (17):

$$f(x) = \max(0, x). \quad (17)$$

Конфигурация нейронной сети представлена на рисунке 10:

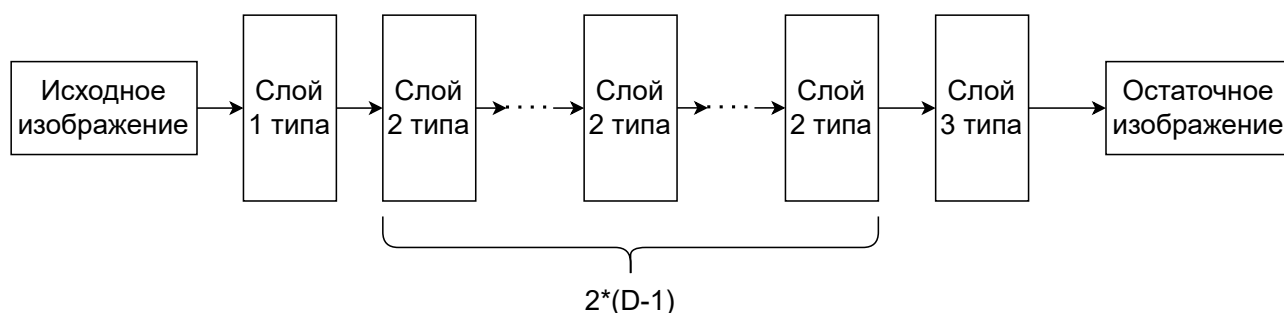


Рисунок 10: Конфигурация нейронной сети в алгоритме DNCNN

На выходе алгоритма получается остаточное изображение R . Пусть Y – исходное изображение. Тогда очищенный снимок X можно посчитать по формуле (18):

$$X = Y - R. \quad (18)$$

1.4.6 RIDNet

Алгоритм RIDNet также использует сверточные нейронные сети, однако в его основе лежит другой подход [23]. Идея состоит в том, что не во всех ситуациях требуется, чтобы модель одинаково оценивала каждую цветовую составляющую пикселя, поэтому у каждой из составляющих будет свой вес [14].

Конфигурация сети состоит из трех основных модулей [14]:

1. **Модуль извлечения признаков** – он состоит из одного слоя, который производит свертку изображения и инициализирует начальные признаки f_0 .
2. **Модуль извлечения остаточных признаков из остаточного изображения** – выполняет основную работу в сети, состоит из четырех дополнительных модулей – EAM (Enhancing attention module).
3. **Модуль реконструкции исходного изображения** – использует один слой, восстанавливает исходное изображение. ReLU используется в качестве передаточной функции.

При этом модули в алгоритме связаны между собой, например, после прохождения первого модуля полученный результат попадает на вход как второго модуля, так и третьего.

Общая конфигурация нейронной сети представлена на рисунке 11:

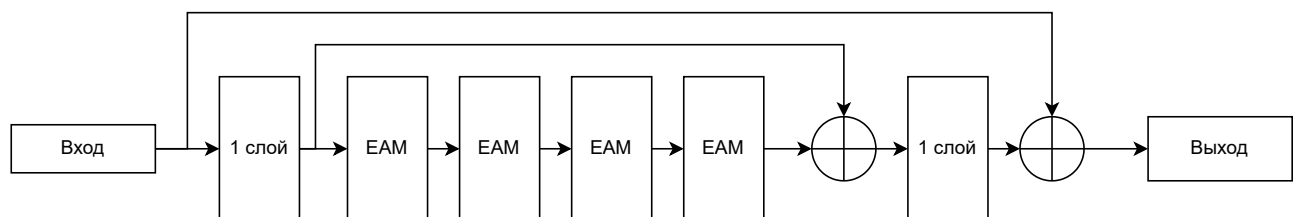


Рисунок 11: Общая конфигурация нейронной сети в алгоритме RIDNet

ЕАМ внутри себя состоит из множества различных слоев, которые комбинируются между собой [14]. Каждый модуль внутри себя выполняет три действия: выделяет основные признаки из изображения, сжимает изображение для большей производительности и увеличивает веса по наиболее значимым признакам.

Внутри модуля используются сетки размером 3×3 . В качестве функции активации на выходе последнего слоя ЕАМ используется сигмоида. Пусть x – вход нейрона, тогда выход можно посчитать по формуле (19):

$$f(x) = \frac{1}{1 + e^x}. \quad (19)$$

При этом авторы алгоритма предложили использовать больше, чем четыре ЕАМ-модуля, однако это не увеличило результаты работы алгоритма [23]. Также в алгоритме предусмотрено использование регуляризации для борьбы с переобучением системы.

1.5 Критерии сравнения алгоритмов

Для верификации работы алгоритмов используются открытые наборы реальных изображений. Алгоритмам предоставляют на вход загрязненные изображения, в результате получается очищенное изображение. У начальных изображений зафиксирован одинаковый размер, так как это позволяет зафиксировать размер входа для нейронной сети. Для полученного результата подсчитываются метрики, и на их основе устанавливается эффективность для метода.

Для сравнения результатов используется две основные метрики: PSNR и SSIM [24].

1.5.1 PSNR

PSNR – метрика, обозначающее пиковое отношение сигнала к шуму [25]. Она универсальна и используется не только для удаления шумов, а также для измерения уровня искажения при сжатии изображений. PSNR рассчитывается по логарифмической шкале.

Пусть есть два изображения I – исходное, и K – полученное в результате обработки. Размер каждого составляет $M \times N$ пикселей.

Разница между ними рассчитывается по формуле (20) [25]:

$$\text{MSE} = \frac{1}{MN} \sum_0^{M-1} \sum_0^{N-1} |I(i, j) - K(i, j)|. \quad (20)$$

Максимальное значение любого пикселя обозначим за MX . Чаще всего оно равно 255 [25]. Теперь искомую метрику PSNR можно рассчитать по формуле (21):

$$\text{PSNR} = 20 \log_{10} \left(\frac{\text{MX}}{\sqrt{\text{MSE}}} \right). \quad (21)$$

Метрика изначально была создана для монохромных изображений, поэтому для цветных изображений результат усредняется по каждой из цветовых составляющих пикселя.

1.5.2 SSIM

SSIM – метрика, обозначающая индекс структурного сходства изображений [26]. В отличие от PSNR, эта характеристика позволяет учесть не только разницу в фактических величинах между пикселями, но и взаимосвязь между ними, то есть структурное сходство.

Пусть есть два изображения I – исходное, и K – полученное в результате обработки. Размер каждого составляет $M \times N$ пикселей. Для каждого изображения были рассчитаны характеристики: μ_i – среднее по цветовой составляющей для изображения i , σ_i^2 – дисперсия цветовой составляющей для фотографии i , σ_{ij} – ковариация цветовой составляющей для изображений i и j .

Также вводится константа L , которая обозначает динамический диапазон. Как правило, она равняется 255 [26].

Исходя из этого, метрику SSIM можно рассчитать по формуле (22) [26]:

$$\text{SSIM} = \frac{(2\sigma_I^2\sigma_K^2 + 0.01L)(2\sigma_{IK}^2 + 0.03L)}{(\mu_I^2 + \mu_K^2 + 0.01L)(\sigma_I^2 + \sigma_K^2 + 0.03L)}. \quad (22)$$

1.6 Классификация существующих решений

Классификацию существующих алгоритмов можно привести к таблице. В качестве основных параметров классификации были использованы следующие признаки:

- учитывается ли анизотропная составляющая при очистке изображения от шумов;
- учитывают ли алгоритмы взаимосвязь пикселей между собой;
- используются ли алгоритмом нейронные сети;
- происходит ли сглаживание в ходе работы алгоритма;
- какой размер изображения требуется для его работы;
- для какого вида шумов используется алгоритм.

Классификация существующих алгоритмов представлена на таблице 1. Под заголовком Gauss приведены критерии для алгоритма гауссовского фильтра, под заголовком Median – для алгоритма медианного фильтра, Bilat – для била-терального фильтра, Jianwei – для алгоритма Цзяньвэй.

Таблица 1: Таблица сравнения существующих алгоритмов

	Median	Gauss	Bilat	Jianwei	DnCNN	RIDNet
Учитывается ли анизотропия?	нет	нет	да	да	да	да
Учитывается ли взаимосвязь пикселей?	нет	да	да	да	нет	да
Происходит ли сглаживание?	нет	да	да	да	нет	нет
С каким шумом работает метод?	им-пульс-ный	гауссов	гауссов	им-пульс-ный	любой	гауссов
Используются нейронные сети?	нет	нет	нет	нет	да	да

Полученные результаты можно свести к следующим тезисам:

- шум на изображениях возникает вследствие несовершенства технических средств. На практике применяются алгоритмические средства борьбы;
- существующие алгоритмы, не использующие нейронные сети, справляются лишь с определенным видом шумов, имея проблемы с устранением остальных типов;
- в алгоритмах DnCNN и RIDNet детали реализации сети могут быть изменены: функция активации, количество слоев;
- новые решения в области удаления шумов являются лишь улучшением или комбинированием уже существующих алгоритмов.

1.7 Формулировка цели

В ходе выполнения работы требуется разработать метод, позволяющий удалить импульсные шумы из цветных изображений. Метод должен отвечать следующим требованиям:

1. Должны удаляться импульсные шумы для цветных изображений любого размера.
2. Максимальное количество процентов шума – до 40% зашумленных пикселей.
3. Для выполнения цели должны быть использованы нейронные сети.

1.8 Формализация задачи. IDEF0-диаграмма

Выводы

Был проведен анализ предметной области, введены основные понятия. Описаны физические причины появления шумов на изображениях. Разобраны основные алгоритмы, которые используются в области удаления шумов, представлены схема их работы. Результаты сравнительного анализа представлены в табличном виде.

Была сформирована цель работы. Задача была формализована в виде IDEF0-диаграммы.

2 Конструкторский раздел

В этом разделе выделяются требования к разрабатываемому методу. Кратко описывается спроектированный метод и его особенности. Описываются конфигурации нейронных сетей и датасет для их обучения. Составляется структура программного комплекса, описывается взаимодействия отдельных частей системы.

2.1 Требования к разрабатываемому методу

Метод очищения цветного изображения от импульсных шумов должен соответствовать следующим требованиям:

1. Принимать на вход изображение в форматах JPG или JPEG, а также вид шума – шум соли или перца. Гарантируется, что процент шума на изображении не превышает 40%.
2. Очистить изображение от шумов.
3. Вывести на экран результат работы алгоритма.
4. Использовать в своей работе нейронные сети.

2.2 Требования к разрабатываемому программному комплексу

Программный комплекс, реализующий интерфейс к разработанному методу, должен позволять выполнять следующие действия:

- возможность загрузки изображений с помощью графического интерфейса;
- возможность выбора типа искомого шума: шум соли или шум перца;
- возможность просмотра результата работы метода и сравнения с исходным изображением.

2.3 Проектирование метода удаления шумов

Разработанный метод должен учитывать следующие особенности ранее известных алгоритмов:

1. Метод медианной фильтрации позволяет очистить изображения от большинства проблемных пикселей, но искажает цветовую составляющую исходного изображения.
2. Нейронная сеть RIDNET плохо справляется с импульсными шумами. В случае обучения на картинках с импульсными шумами она подстраивается на определенный их процент, а обучать на все возможные проценты различные модели является нецелесообразным.
3. Конфигурация нейронной сети RIDNET позволяет обучить ее для того, чтобы корректировать цвета пикселей, полученные в результате применения медианного фильтра.

2.3.1 Основной алгоритм работы

Учитывая особенности известных методов, перечисленные ранее, был разработан метод, позволяющий бороться с шумами на изображениях.

В первую очередь, картинку требуется разбить на изображения небольшого размера, называемые патчами. Это следует сделать для того, чтобы корректнее обучить нейронные сети и получить вменяемый результат.

Затем нужно оценить количество шума на патче. В случае, если оно больше, чем некое пороговое значение, то требуется предварительно применить медианный фильтр. Результаты работы медианного фильтра нужно обработать в нейронной сети, способной к коррекции исходных цветов изображений. По итогу предварительный результат попадает на нейронную сеть, предназначенную для очистки изображения от шумов.

Для получения конечного результата требуется склеить изображение из патчей в нужной последовательности.

Схема алгоритма представлена на рисунке 12:

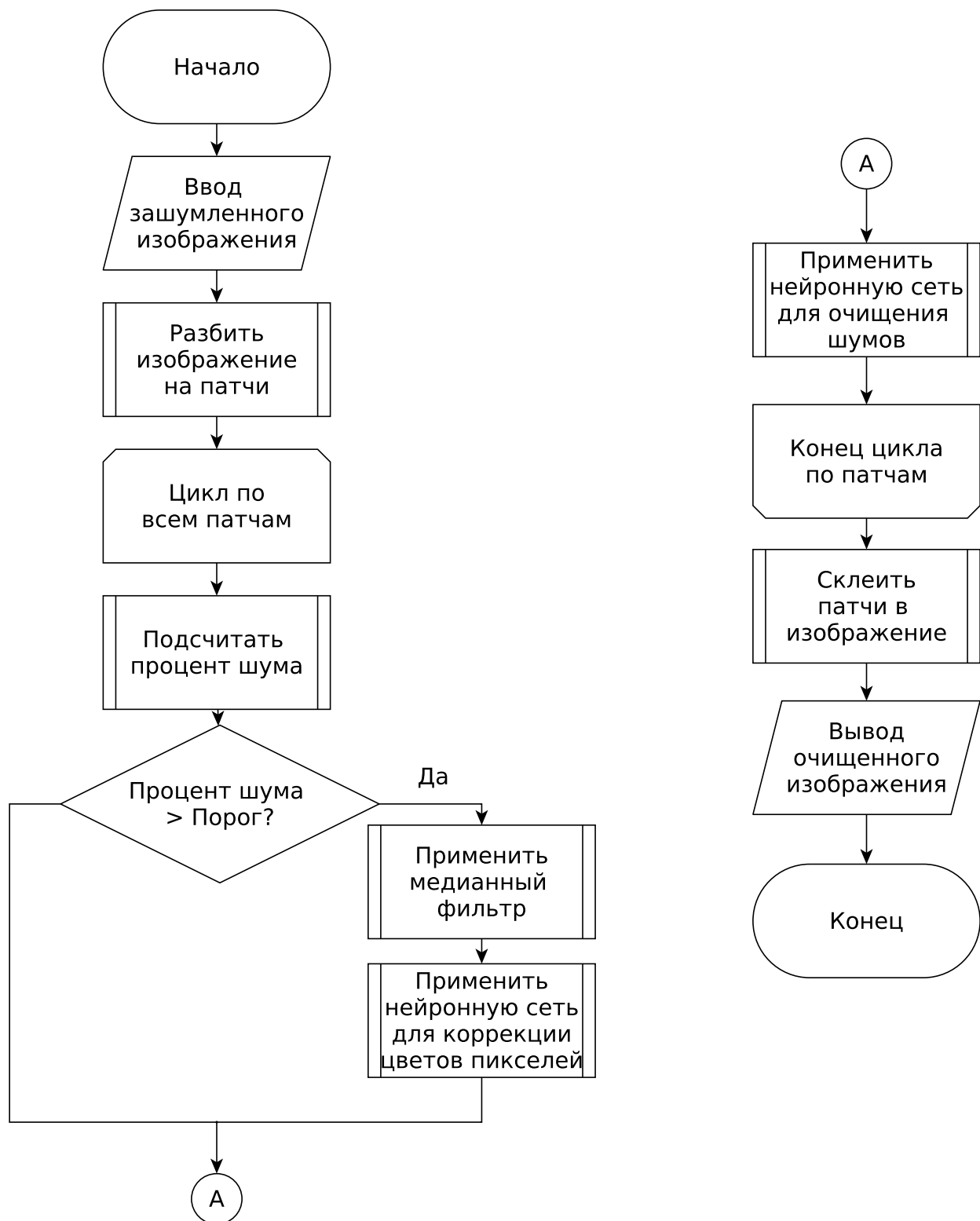


Рисунок 12: Верхнеуровневая схема алгоритма

2.3.2 Подсчет количества шумов в изображении

Для принятия решения об использовании медианного фильтра требуется подсчитать, превышает ли количество импульсных шумов необходимый порог.

Все цвета пикселей делятся на следующие три категории:

1. Цвета, принимающиеся за шум соли.
2. Цвета, принимающиеся за шум перца.
3. Цвета, не принимающиеся за шум.

Однако главная сложность состоит в том, что пиксель может быть ошибочно принят за шум, если он находится на фоне, соответствующим шуму. Например, если рассматривать яркое небо, то почти все пиксели в этом случае можно интерпретировать как шум соли. Соответственно, нужно учитывать контекст для принятия решения.

Для этого требуется оценить цвет рассматриваемого пикселя по сравнению с соседями. Нужно подсчитать среднюю разницу цветов с соседними пикселями, в соответствии с сеткой, используемой в медианном фильтре. Если она получается больше, чем заданная разница, то пиксель определяется как шум, иначе – как исходный пиксель. Заданная разница рассчитывается экспертной оценкой.

В результате требуется посчитать количество пикселей, принятых за шум. Так как размер патча заранее известен, достаточно найти отношения импульсного шума к общему числу пикселей.

Схема алгоритма подсчета количества шумов представлена на рисунке 13:

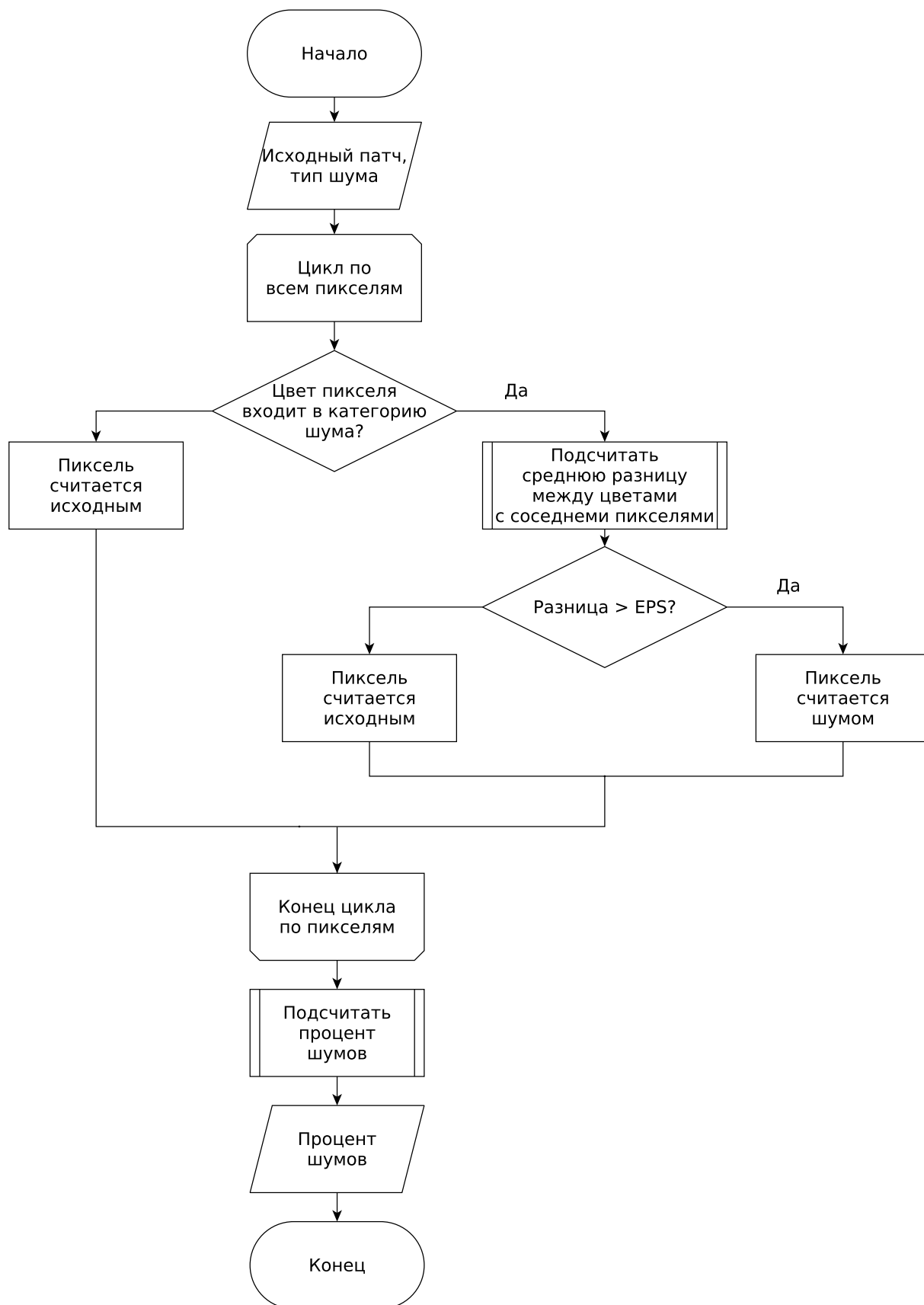


Рисунок 13: Схема алгоритма подсчета количества шумов

2.3.3 Конфигурация нейронных сетей

В качестве основы для итоговой конфигурации нейронной сети была выбрана модель RIDNET.

В нее были добавлены соответствующие изменения:

1. Она состоит из трех ЕАМ-модулей, а не из четырех, как в исходном алгоритме.
2. В качестве функции активации выбрана сигмоида на всех элементах ЕАМ-модуля, а не ReLu.
3. Функция потерь – среднеквадратичное отклонение от результата, а не абсолютная ошибка, как в RIDNET.

Конфигурация нейронной сети для очищения изображения от шумов также основана на RIDNET, но в качестве функции активации внутри ЕАМ-модуля использовалась сигмоида для частей, склеивающих результат с нескольких каналов в один выход.

В качестве функции оптимизации был использован Adam (адаптивная оценка момента). Вместо того чтобы адаптировать скорость обучения параметров на основе среднего первого момента (среднего значения), Adam также использует среднее значение вторых моментов градиентов.

2.3.4 Данные для обучения модели

В качестве данных, используемых для обучения моделей, был использован датасет BSDS500, использующий для верификации алгоритмов, специализирующихся на очищении гауссовых шумов из изображений. Выбор обусловлен тем, что он уже использовался для оценки качества работы других нейронных сетей, перечисленных в аналитическом разделе, соответственно, не придется дополнительно обучать другие модели и применять фильтры для сравнения результата. Также множество изображений было составлено таким образом, что цветовая гамма изображений в них отличается, что позволит проверить метод в различных ситуациях.

Однако для моделирования импульсных шумов требуется разработать собственные функции.

2.4 Структура разрабатываемого программного комплекса

Разрабатываемый программный комплекс состоит из трех модулей:

1. Модуль разработки и обучения нейронных сетей, в котором также происходит разработка датасета.
2. Модуль обработки изображений.
3. Модуль взаимодействия с пользователем.

На рисунке 14 представлена схема модуля разработки и обучения нейронных сетей:

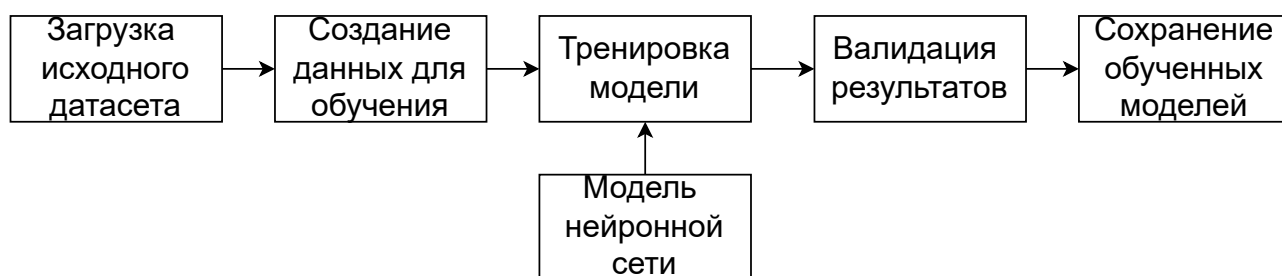


Рисунок 14: Схема модуля разработки и обучения нейронных сетей

Валидация результата должна происходить по следующим правилам:

1. Значение PSNR для очищенного изображения превышает показатель для загрязненного изображения.
2. Значение PSNR больше, чем для реализации исходных алгоритмов: RIDNET и медианный фильтр.

Схема модуля взаимодействия с пользователем представлена на рисунке 15:

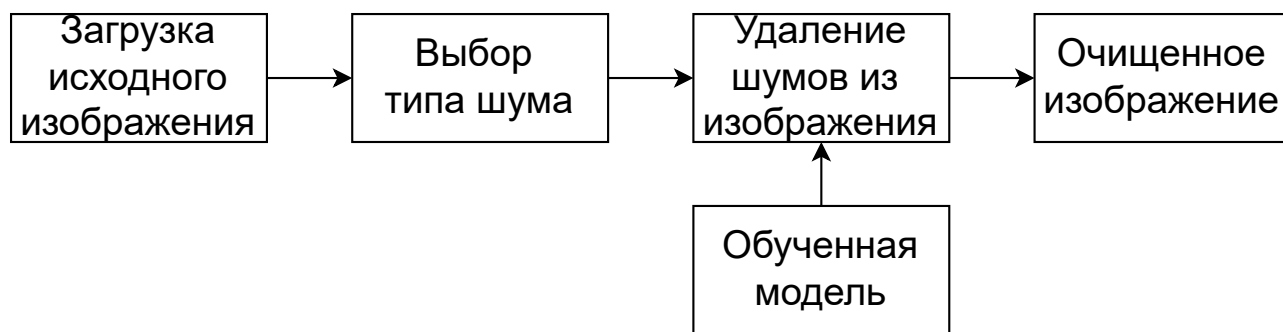


Рисунок 15: Схема модуля взаимодействия с пользователем

Выводы

Были составлены требования к разрабатываемому методу. Был спроектирован метод с учетом особенностей уже существующих способов, описаны его особенности, составлены схемы основных алгоритмов. Была разработана конфигурация нейронных сетей, описан путь составления датасета для его обучения. Была составлена структура программного комплекса, описаны взаимодействия отдельных частей системы.

3 Технологический раздел

3.1 Средства реализации программного комплекса

3.1.1 Выбор языка программирования

Для написания программного обеспечения был использован язык Python 3.10. Это обусловлено следующими причинами:

- имеется опыт разработки на данном языке;
- для Python было реализовано множество библиотек, позволяющих работать со сверточными нейронными сетями.

3.1.2 Выбор библиотек для разработки ПО

Для работы с нейронными сетями была выбрана библиотека tensorflow версии 2.12.

Это обусловлено тем, что на текущий момент tensorflow показывает более высокие показатели производительности, чем другие аналоги, например, pytorch. Также у автора имеется опыт использования данной библиотеки для разработки сверточных нейронных сетей.

В качестве библиотеки для работы с изображениями был выбран OpenCV, так как в ней присутствуют модули, позволяющие взаимодействовать с функциями из tensorflow. Для вывода изображений использовалась библиотека matplotlib.

3.2 Реализация программного комплекса

3.2.1 Нейронная сеть

Конфигурация всех нейронных сетей основана на методе RIDNET, но количество модулей EAM было уменьшено до трех в целях уменьшения времени, потраченного на обучения моделей. При этом в качестве функции активации используется сигмоида, что позволило улучшить точность результатов, однако увеличило время обучения сети.

Исходную базу изображений, требуемых для обучения модели, потребовалось разбить на небольшие изображения размером 40×40 пикселей, так как это предусмотрено в модели. Для увеличения набора тестовых данных некоторые изображения подвергались предварительным преобразованиям, таким как поворот и обрезка.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Thanh Hien Nguyen. Image Noise Removal Method Based on Thresholding and Regularization Techniques // IEEE Access. 2022. С. 71584–71597.
2. Зотов П.В. Цифровой шум изображения и его прикладное значение в криминалистике // Вестник СГЮА. 2015.
3. Голдобин И.А. Климова Е.И. Влияние шумов на алгоритмы цифровой обработки изображений // Актуальные вопросы развития современной цифровой среды. 2021. С. 396–402.
4. Hambal Abdalla Mohamed. Image Noise Reduction and Filtering Techniques // International Journal of Science and Research (IJSR). 2017. С. 2033–2037.
5. H.W. Xiangming. An improved nonlocal means-based correction strategy for mixed noise removal // IET Image Processing. 2022.
6. Čeranić Š.T. Noise reduction quality test for two-photon laser scanning microscopic images // IEEE Access. 2022.
7. G.M. Ganesan. Noise Detection in Images using Moments // Research Journal of Applied Sciences, Engineering and Technology. 2015. С. 307–314.
8. Aylwin Rubén, Henríquez Fernando, Schwab Christoph. ReLU Neural Network Galerkin BEM // Journal of Scientific Computing. 2023. 03. Т. 95.
9. Image Classification Based on Light Convolutional Neural Network Using Pulse Couple Neural Network / Maminaiaina Rafidison, Hajasoa Ramafiarisona, Paul Randriamitantsoa [и др.] // Computational Intelligence and Neuroscience. 2023. 03. Т. 2023. С. 1–17.
10. Lin Xiyu. Research of Convolutional Neural Network on Image Classification // Highlights in Science, Engineering and Technology. 2023. 04. Т. 39. С. 855–862.
11. Karrach Ladislav, Pivarciova Elena. Using a Convolutional Neural Network for Machine Written Character Recognition. 2023. 03.

12. Sle-CNN: a novel convolutional neural network for sleep stage classification / Zhenman Zhang, Yu Xue, Adam Slowik [и др.] // Neural Computing and Applications. 2023. 04. C. 1–16.
13. Liu Ruishen. Face Recognition Based on Convolutional Neural Networks // Highlights in Science, Engineering and Technology. 2022. 11. T. 16. C. 32–39.
14. X. Li. RIDNet: Recursive Information Distillation Network for Color Image Denoising. 2019. C. 3896–3903.
15. M. Fujimoto. Noise Removal for Degraded Images Using Neural Networks // IEEJ Transactions on Electronics, Information and Systems. 2002. C. 1301–1308.
16. A. Bandyopadhyay. High Density Impulse Noise Removal from Color Images by K-means Clustering based Detection and Least Manhattan Distance-oriented Removal Approach // International Journal of Advanced Computer Science and Applications. 2021. C. 608–614.
17. S. Khan. An adaptive dynamically weighted median filter for impulse noise removal // EURASIP Journal on Advances in Signal Processing. 2017.
18. Kunsoth R. Biswas M. Modified decision based median filter for impulse noise removal. 2016. C. 1316–1319.
19. M. Nabahat. Optimization of bilateral filter parameters using a whale optimization algorithm // Research in Mathematics. 2022.
20. Wang Jianwei. A Noise Removal Algorithm of Color Image // TELKOMNIKA Indonesian Journal of Electrical Engineering. 2014.
21. V. Murali. Image Denoising Using DnCNN: An Exploration Study // IEEE Access. 2020. C. 847–859.
22. K. Zhang. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising // IEEE Transactions on Image Processing.
23. N. Loubere. RiDNet Practical Fieldwork Notes. 2012.

24. D.M. Setiadi. PSNR vs SSIM: imperceptibility quality assessment for image steganography // Multimedia Tools and Applications. 2021.
25. M. Nouman. Evaluation of Image Super Resolution Deep Learning Technique based on PSNR Value. 2022. C. 93–122.
26. V. Subramanian. SSIM Compliant Modeling Framework With Denoising and Deblurring Applications // IEEE Transactions on Image Processing. 2021.