

Министерство образования Республики Беларусь
Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

УДК 519.688

Никитин
Глеб Юрьевич

Квантователь изображений на основе нейронных сетей

АВТОРЕФЕРАТ

на соискание степени магистра технических наук
по специальности 1-40 80 01 «Элементы и устройства вычислительной
техники и систем управления»

Научный руководитель
Петровский Н.А.
канд. техн. наук

Минск 2020

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования.

В настоящее время широко применяются различные алгоритмы компрессии изображений практически во всех существующих системах связи, поэтому актуальность данной темы состоит в поиске новых или улучшении существующих алгоритмов компрессии изображений.

Цель и задачи исследования.

Целью данной работы является разработка модели квантования изображений как комбинации нейронных сетей и вейвлет преобразования. Соответственно в работе сформулированы и решены следующие задачи:

1. изучить существующие модели компрессии изображений;
2. изучить теорию вейвлет преобразования и существующие вейвлеты и произвести выбор вейвлета для исследования;
3. разработать модель квантования изображений на основе вейвлет преобразования и нейронных сетей;
4. провести экспериментальное исследование разработанной модели.

Объектом исследования являются методы и способы квантования изображений.

Предметом исследования является применение вейвлет преобразования в алгоритмах квантования изображения на основе нейронных сетей.

Область исследования и содержание диссертационной работы соответствуют образовательному стандарту высшего образования второй ступени (магистратуры) специальности 1-40 80 01 «Элементы и устройства вычислительной техники и систем управления».

Научная новизна диссертационной работы заключается в разработке и экспериментальном исследовании модели квантования изображений на основе нейронных сетей и вейвлет преобразования.

Положения, выносимые на защиту.

1. разработанная модель квантования изображений на основе нейронных сетей и вейвлет преобразования.
2. экспериментальные исследования эффективности полученной модели на базе языка Python и библиотек tensorflow и pywavelets.

Апробация результатов диссертации.

Основные положения и результаты диссертационной работы докладывались и обсуждались на 55-я научной конференции аспирантов, магистрантов и студентов БГУИР (Минск, 2019).

Опубликованность результатов исследования.

По результатам исследований, представленных в диссертации, опубликован тезис в сборнике и материале научной конференции.

Структура и объем диссертации.

Структура диссертационной работы обусловлена целью, задачами и логикой исследования. Работа состоит из введения, четырёх глав, заключения, библиографического списка и приложений. Общий объем диссертации - 88 страниц. Работа содержит 12 таблиц, 30 рисунков. Библиографический список включает 71 наименование, графический материал включает 15 слайдов презентации (Приложение Б).

Библиотека БГУИР

ВВЕДЕНИЕ

Под цифровыми изображениями понимают двумерный набор пикселей, несущих в себе некоторую информацию и расположенных в пространстве понятным человеческому глазу образом. Каждое цифровое изображение имеет формат, который представляет собой способ его кодирования, позволяющий с той или иной степенью эффективности сохранять изображения в памяти компьютера. На текущий момент существует множество вариантов такого кодирования, как с потерями, так и без. Каждый из таких форматов имеет свои преимущества и недостатки, например:

1. у форматов кодирования с потерями в качестве преимущества можно выделить малый размер кодированного представления изображений, однако качество восстановленных изображений ниже по сравнению с методами сжатия без потерь;
2. у форматов кодирования без потерь же всё наоборот: такие форматы не оставляют заметных артефактов на изображении, однако изображения требуют значительно больше места в памяти.

Помимо стандартизированных алгоритмов компрессии изображений постоянно появляются и разрабатываются новые. Одним из новых направлений в этой проблеме являются алгоритмы компрессии на нейронных сетях. Нейросетевые алгоритмы показывают сравнимые результаты относительно классических алгоритмов (например, JPEG), однако часто такие алгоритмы оперируют изображениями, имеющими высокую информационную избыточность, которая может быть устранена при помощи различных преобразования цифровой обработки сигнала.

Поэтому основной целью данной работы является улучшение результатов работы нейросетевого квантователя при помощи вейвлет преобразования, одного из известных и популярных инструментов не только компрессии, но и анализа сигнала. Для достижения поставленной цели сформулированы следующие задачи:

1. изучить существующие способы снижения энтропии изображений;
2. изучить теорию вейвлет преобразования и существующие вейвлеты и произвести выбор вейвлета для исследования;
3. разработать модель квантования изображений на основе вейвлет преобразования и нейронных сетей;
4. провести экспериментальное исследование разработанной модели.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

1 Анализ методов снижения энтропии в изображении

В первой главе диссертации проведён обзор и анализ литературы с точки зрения методов снижения энтропии в изображении для последующего сжатия. Так, например, были выделены следующие области и направление исследуемой тематики:

1. Стандартизированные методы компрессии изображений.
Здесь были рассмотрены классические методы компрессии и квантования изображения, например: рассмотрен один из самых популярных алгоритмов компрессии изображений JPEG2000, а также стандарт для кодирования видео HEVC.
2. Многослойный перцептрон.
Показаны теоретические основы, скрывающиеся за словом перцептрон, а также рассмотрена сеть Сонехараета и алгоритм, предложенный Сикуранзаэтом.
3. Случайные нейронные сети.
В этом разделе рассмотрены методы, разработанные Геленбе и Крамелом, основанные на случайных нейронных сетях. Сигналы в случайной нейронной сети передаются в виде пиков единичной амплитуды. Связь между нейронами в таких сетях исходит из процесса Пуассона, где положительные сигналы представляют собой возбуждающие сигналы, а отрицательные сигналы – тормозящие.
4. Методы на основе свёрточных нейронных сетей.
В этом разделе был сделан упор на работы американского учёного Балле с соавторами, представившем нейросетевого квантователя, которая легла в основу этой работы.
5. Рекуррентные нейронные сети.
Методы компрессии с использованием рекуррентных нейронных сетей базируются на особом виде памяти, называемом Long short-term memory, а также более продвинутой её версии, Gated Recurrent Unit. Эти структуры легли в основу рассмотренных методов, предложенных Хохрейтером и Тодеричи.
6. Методы на основе генеративных состязательных нейронных сетей.
Один из новейших способов снижения энтропии в изображении. По своей структуре генеративные состязательные сети представляют собой две сети: генератор и дискриминатор. Дискриминатор пытается

различить, генерируются ли образцы из генератора, а генератор - обмануть дискриминатор и производить образцы, которые проходят проверку. Здесь были рассмотрены алгоритмы, предложенные Бурдевым и Грегоретом.

2 Применение вейвлет анализа для декомпозиции изображения на субполосы

В этой главе рассмотрена теория вейвлет преобразования и обозначены основные способы его применения. Также было показано приложение вейвлетов для двумерных сигналов и, в частности, вейвлетов. Показан алгоритм, по которому происходит декомпозиция изображения в его вейвлет области. В конце главы был проведён обзор следующих существующих вейвлетов:

1. Вейвлет Хаара.
2. Вейвлеты Добеши.
3. Биортогональные вейвлеты.
4. Коифлеты.
5. Симлеты.
6. Вейвлеты Морле.
7. Вейвлет мексиканская шляпа.
8. Вейвлет Мейера.
9. Другие вейвлеты, определённые в действительной и комплексной областях.

Как результат было постулировано, что семейств вейвлетов достаточно большое количество и правильный выбор вейвлета для анализа сигнала – крайне нетривиальная задача.

3 Архитектура системы

В этой главе было проведено обоснование выбора уровня вейвлет преобразования и варианта вейвлета, используемых для исследований. В общем случае выбор вейвлета зависит лишь от исследуемого сигнала, однако существует набор общих критериев, которые помогают осуществить такой выбор, который и был описан. После чего было обосновано, что уровень вейвлета был принят равным единице в силу ограниченности ресурсов системы для обучения, а в качестве вейвлета был использован третий вейвлет

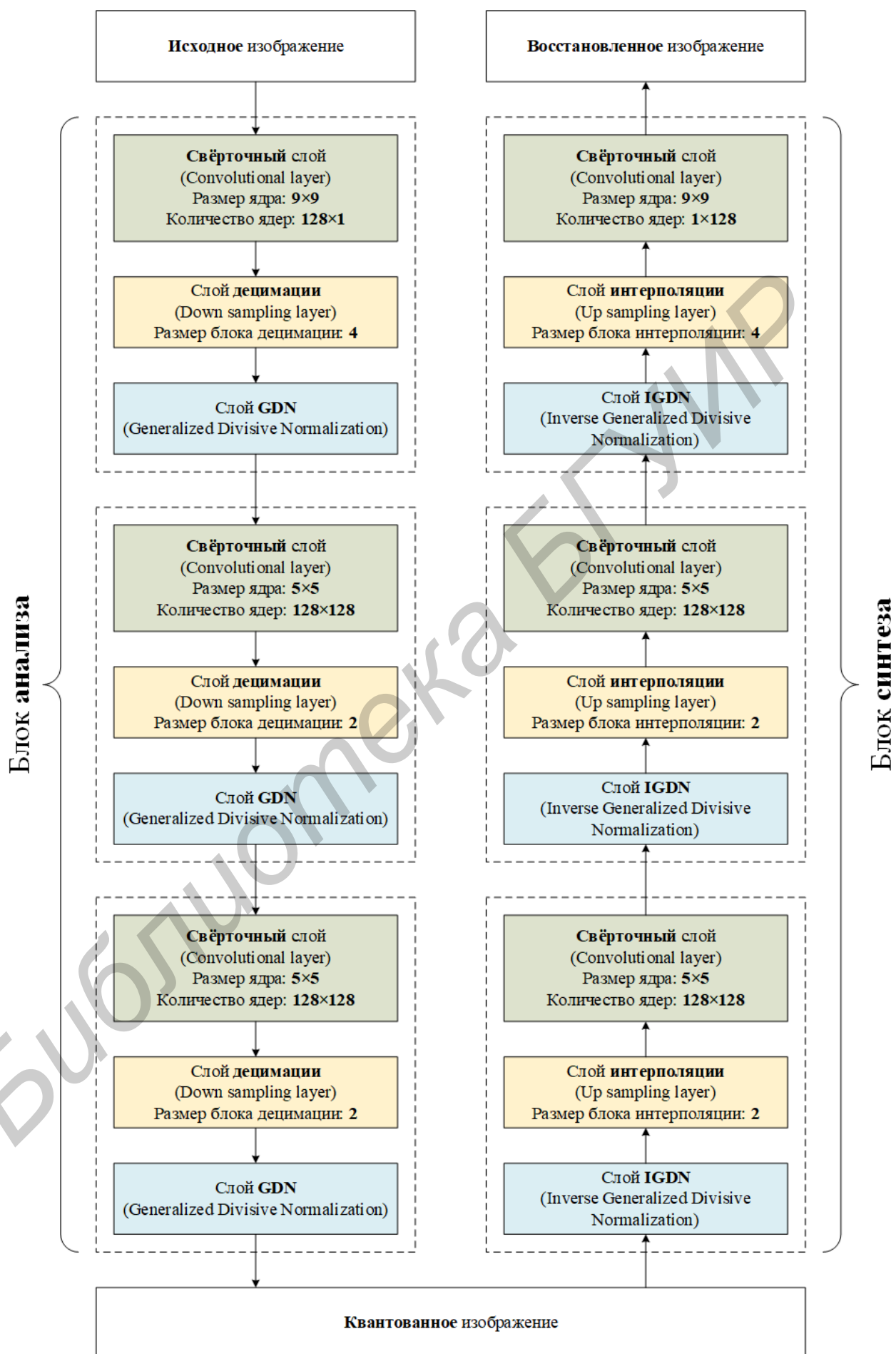


Рисунок 1. – Архитектура автоэнкодера

семейства Добеши (db3). Затем была описана архитектура используемого нейросетевого квантователя для исследований, рисунок 1.

В преобразовании анализа слева на рисунке 1 имеется три условных блока преобразований, каждый из которых состоит из операции аффинной свертки, операции децимации и специальной нормализации. Аффинная свертка определяется следующей формулой:

$$v_i^{(k)}(m, n) = \sum (h_{k,ij} * u_j^{(k)})(m, n) + c_{k,i} \quad (1)$$

где $u_j^{(k)}$ – входной канал k-го этапа в пространственном местоположении (m, n);

* – операция двумерной свертки;

$h_{k,ij}$ – параметр свертки;

$c_{k,i}$ – параметр смещения нейронной сети свертки.

Затем над результатом свёртки выполняется операция децимации, позволяющая уменьшить размерность данных:

$$\omega_i^{(k)}(m, n) = v_i^{(k)}(s_k m, s_k n) \quad (2)$$

где s_k – фактор понижающей дискретизации.

Результат операции децимации подвергается специальной нормализации, называемой *Generalised Divisive Normalisation* (GDN):

$$u_i^{(k+1)}(m, n) = \frac{\omega_i^k(m, n)}{\sqrt{\beta_{k,i} + \sum_j \gamma_{k,ij} (\omega_j^k)^2}} \quad (3)$$

где $\gamma_{k,ij}$ – параметр масштаба для операции нормализации;

$\beta_{k,i}$ – параметр смещения для операции нормализации.

Этот блок чем-то напоминает функцию активации, используемую в классических нейронных сетях, однако таковой не является, поскольку имеет тренируемые параметры.

Блок синтеза состоит из инверсных операций блока анализа для всех трёх его этапов и начинается с инверсной нормализации *Inverse Generalised Divisive Normalisation* (IGDN):

$$\hat{\omega}_i^{(k)} = \hat{u}_i^{(k)}(m, n) \cdot \sqrt{(\hat{\beta}_{k,i} + \sum_j \hat{\gamma}_{k,ij} (\hat{u}_j^{(k)}(m, n))^2)} \quad (4)$$

За которой следует операция интерполяции, возвращающей слоям исходную размерность данных:

$$\hat{v}_i^{(k)} = \begin{cases} \hat{\omega}_i^{(k)}(m/\hat{s}_k, n/\hat{s}_k) & \text{if } m/\hat{s}_k \text{ and } n/\hat{s}_k \text{ are round,} \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

И завершается блок операцией аффинной свёртки:

$$\hat{u}_i^{(k+1)} = \sum_j (\hat{h}_{k,ij} * \hat{v}_j^{(k)})(m, n) + \hat{c}_{k,i} \quad (6)$$

Затем была пояснена замена вейвлета db3 на первый вейвлет семейства Добеши (db1) из-за особенностей архитектуры нейросетевого квантователя (требование на размер входного изображения в $256 * n \times 256 * n$ пикселей) и синтезирована общая архитектура системы, показанная на рисунке 2.

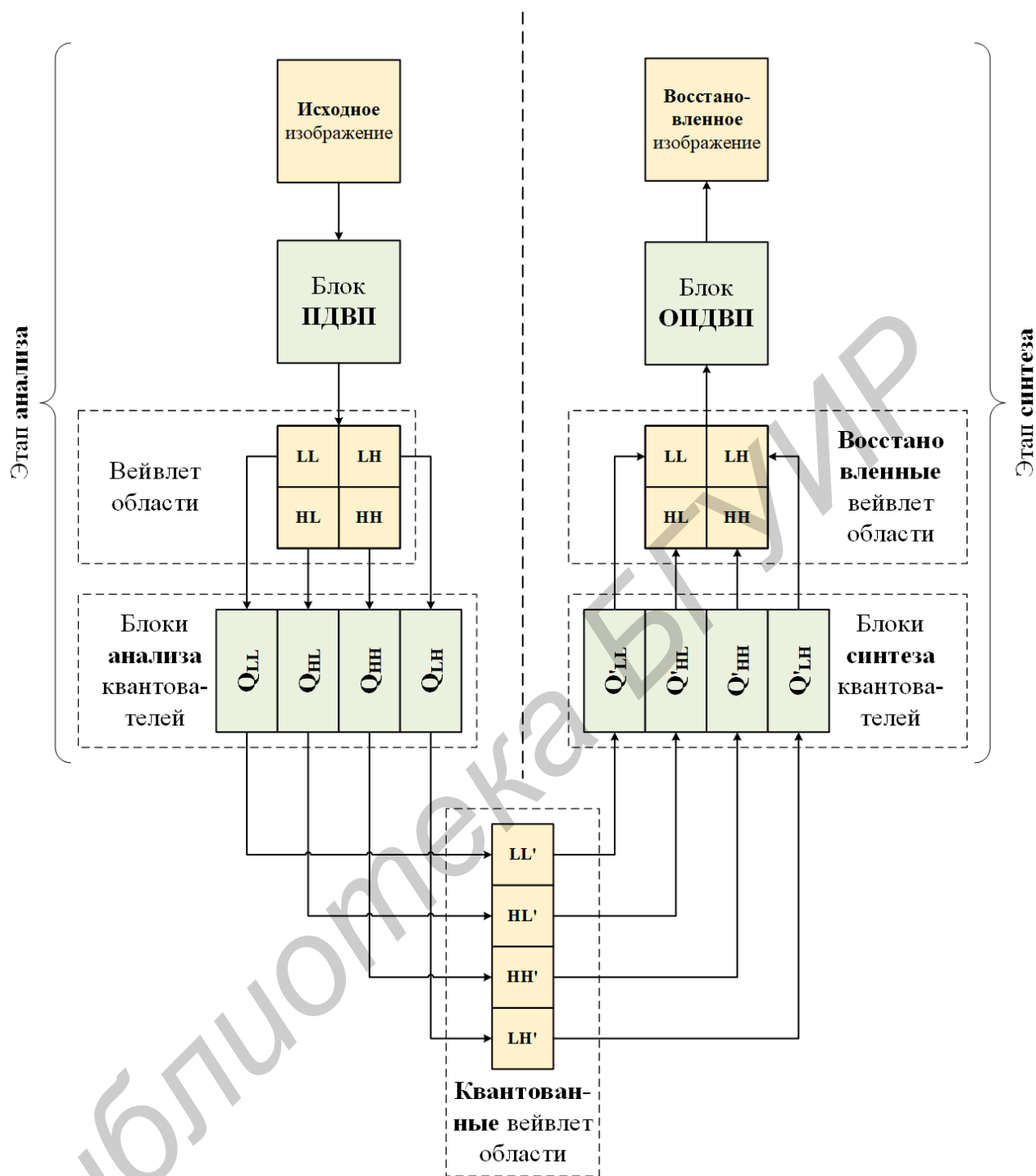
4 Экспериментальные результаты

В начале главы описываются объективные характеристики изображения (метрики), по которым производилось сравнение между оригинальными и восстановленными изображениями. В качестве метрик были использованы: MSE, PSNR, MS-SSIM (индекс структурного сходства), MS-SSIM-DB (его логарифмическая форма). Также проведён обзор доступных ресурсов для обучения и описана тестовая выборка, использованная при обучении нейросетевых квантователей. Основу выборки составила база «The USC-SIPI Image Database», из которой в последствии был подготовлен набор изображений, состоящий из 91 монохромного изображения с 8-ю бит на пиксель с размерами 512×512 пикселей для проведения экспериментов. Затем был описан алгоритм проведения экспериментов и серия из 4 экспериментов по обучению нейросетевых квантователей.

4.1 Модель model_11

Отличительная особенность модели: модель создана и обучена без использования вейвлет преобразования.

В качестве первого эксперимента была обучена нейронная сеть без модификаций. Проведение данного эксперимента позволило использовать его в качестве эталонного для сравнения между оригинальным методом и его модификациями. Общие характеристики эксперимента показаны в таблице 1.



Q_{XX} – блок анализа нейросетевого квантователя для соответствующей вейвлет области,
 Q'_{XX} – блок синтеза нейросетевого квантователя для соответствующей вейвлет области.

Рисунок 2. – Общая архитектура системы

4.2 Модель model_12

Отличительная особенность модели: модель создана и обучена в соответствии с модифицированным алгоритмом.

Этот эксперимент проведён в соответствии с предложенным алгоритмом. В качестве вейвлета использован первый вейвлет семейства Добеши

Таблица 1. – Таблица характеристик модели model_11

Файл модели	model_11.py
Вейвлет	–
Уровень вейвлета	0
Итераций обучения на сеть	2500
Среднее время обучения	5 min
Суммарное время обучения	5 min

(db1) для 1 уровня преобразования. Характеристики модели показаны в таблице 2.

Таблица 2. – Таблица характеристик модели model_12

Файл модели	model_12.py
Вейвлет	db1
Уровень вейвлета	1
Итераций обучения на сеть	2500
Среднее время обучения	5 min
Суммарное время обучения	20 min

4.3 Модель model_13e

Отличительная особенность модели: обучена только одна нейронная сеть для набора изображений низкочастотной вейвлет области. Недостающие вейвлет области взяты из блока ПДВП, минуя нейросетевые квантователи.

В этом эксперименте был модифицирован предложенный алгоритм работы: была обучена лишь 1 сеть для низкочастотной области вейвлет преобразования 1 уровня для вейвлета db1. В качестве недостающих 3 вейвлет областей используются вейвлет области, полученные непосредственно из блока ПДВП. Общие характеристики такой модели показаны в таблице 3.

4.4 Модель model_14

Отличительная особенность модели: длительное обучение нейронных сетей (10000 итераций).

В этом эксперименте модель для обучения имела гораздо больше времени для обучения, что и послужило основным отличием от предыдущих

Таблица 3. – Таблица характеристик модели model_13e

Файл модели	model_13.py
Вейвлет	db1
Уровень вейвлета	1
Итераций обучения на сеть	2500
Среднее время обучения	5 min
Суммарное время обучения	5 min

экспериментов. В таблице 4 показаны общие характеристики такой модели.

Таблица 4. – Таблица характеристик модели model_14

Файл модели	model_14.py
Вейвлет	db1
Уровень вейвлета	1
Итераций обучения на сеть	10000
Среднее время обучения	20 min
Суммарное время обучения	80 min

4.5 Сравнение полученных результатов

Все эксперименты были оценены при помощи метрик MSE, PNSR, MS-SSIM и MS-SSIM-DB. Сводные таблицы метрик PSNR и MS-SSIM всех экспериментов показаны на таблицах 5 и 6 соответственно. Также по полученным результатам построены гистограммы, показанные на рисунках 3 и 4 соответственно.

Рассмотрев гистограммы метрик ПОСШ и MS-SSIM можно увидеть, что по ходу проведения экспериментов были получены результаты, сравнимые с оригинальным методом, но не превосходящие его. Из чего следует вывод о том, что для достижения системой результатов, превосходящих оригинальный метод, предложенных изменений недостаточно.

Таблица 5. – Сводная таблица метрик PSNR для проведённых экспериментов

Имя	model_11	model_12	model_13e	model_14
4.2.05.png	28.84	26.36	27.84	25.76
4.2.06.png	28.05	24.87	26.3	25.71
4.2.07.png	29.3	27.02	28.67	26.27
5.2.10.png	26.05	21.81	23.72	22.62
7.1.03.png	31.02	29.91	31.73	30.35
7.1.09.png	29.71	28.23	29.99	28.7
barbara.png	24.95	23.96	26.69	24.58
boat.512.png	29.11	26.04	27.52	26.05
car.png	27.13	24.96	26.19	25.54
lena.png	31.25	28	29.14	28.43

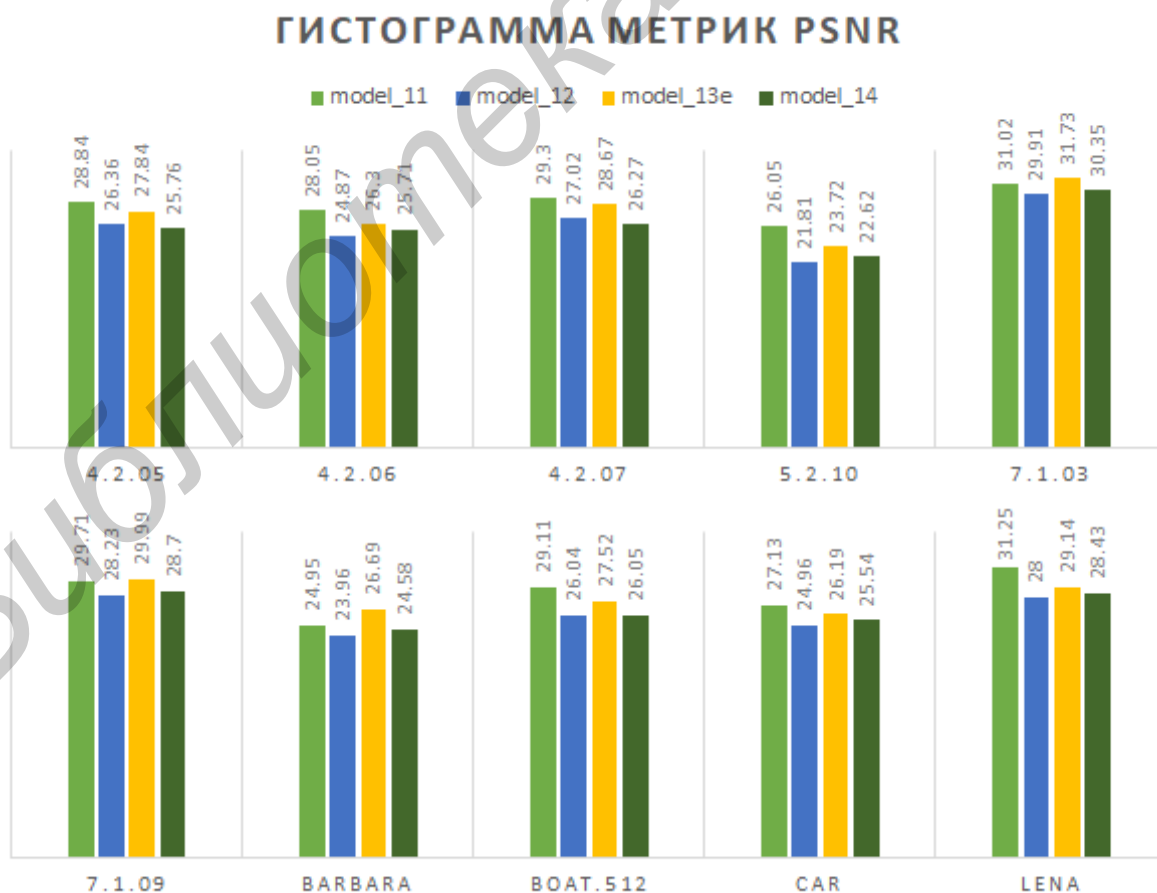


Рисунок 3. – Гистограмма метрик PSNR для проведённых экспериментов

Таблица 6. – Сводная таблица метрик MS-SSIM для проведённых экспериментов

Имя	model_11	model_12	model_13e	model_14
4.2.05.png	0.93	0.88	0.89	0.88
4.2.06.png	0.95	0.88	0.89	0.9
4.2.07.png	0.95	0.9	0.91	0.9
5.2.10.png	0.97	0.87	0.89	0.9
7.1.03.png	0.94	0.91	0.92	0.92
7.1.09.png	0.95	0.9	0.91	0.92
barbara.png	0.93	0.88	0.89	0.9
boat.512.png	0.95	0.89	0.9	0.9
car.png	0.95	0.89	0.89	0.9
lena.png	0.95	0.91	0.91	0.92

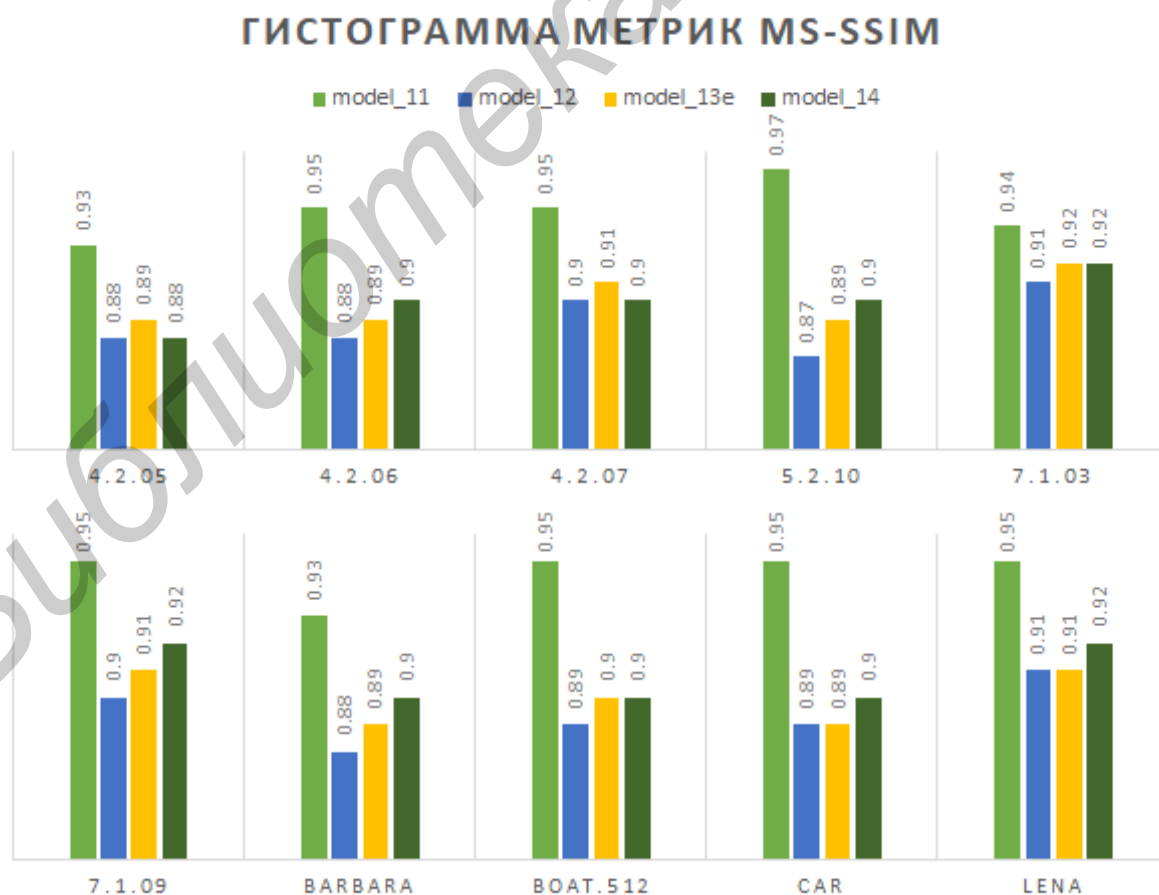


Рисунок 4. – Гистограмма метрик MS-SSIM для проведённых экспериментов

Заключение

Таким образом было проведено исследование нейросетевого квантователя изображений при помощи вейвлет преобразования. Были получены результаты, не превосходящие оригинальный метод, но сравнимые с ним. Из чего следует вывод о том, что для достижения системой результатов, превосходящих оригинальный метод, предложенных изменений недостаточно. Это объясняется следующим набором причин:

1. ограниченность ресурсов для обучения (малый объём выборки, ограниченная мощность системы для обучения);
2. ограниченность выбора типа вейвлета (db1) и уровня вейвлет преобразования (1);
3. некогерентность работы нейросетевых квантователей, т.е. отсутствие информации у нейронных сетей о том, как протекает обучение для соседних подобластей вейвлет преобразования.

В качестве нейросетевого квантователя была использована сеть из работ Балле и его соавторов. В качестве вейвлета был выбран третий вейвлет семейства Добеши (db3) для 1 уровня преобразования, как вейвлет, имеющий наименьшую ошибку декомпозиции-восстановления из всех исследованных вейвлетов. Однако показанная архитектура нейронной сети позволяет работать лишь с изображениями, кратными 256×256 пикселям, что послужило причиной замены вейвлета db3 на первый вейвлет семейства Добеши (db1). Выбор уровня вейвлета был обусловлен ограничениями ресурсов для обучения, поскольку с увеличением уровня вейвлет преобразования растёт и количество нейронных сетей, требующих обучения.

В рамках данного исследования были полностью достигнуты задачи, сформулированные во введении:

1. изучены существующие способы снижения энтропии изображений;
2. изучена теория вейвлет преобразования и существующие вейвлеты, а также произведён и обоснован выбор вейвлета для исследований;
3. разработана модель квантователя изображений на основе вейвлет преобразования и нейронных сетей;
4. произведены экспериментальные исследования разработанной модели.

Таким образом, после анализа полученных результатов можно сделать вывод о достижении нейросетевыми квантователями результатов, сравнимых с оригинальным методом при сочетании с вейвлетами в предложенном виде, однако для достижения результатов, превосходящих оригинальный метод, предложенных изменений недостаточно.

Список публикаций соискателя

1–А. Никитин, Г. Ю. Сжатие изображений на основе нейронных сетей / Г. Ю. Никитин // Компьютерные системы и сети: 55-я юбилейная научная конференция аспирантов, магистрантов и студентов, Минск, 22-26 апреля 2019 г. / Белорусский государственный университет информатики и радиоэлектроники. – Минск, 2019. – С. 275 – 277. (<https://libeldoc.bsuir.by/handle/123456789/35272>)

Библиотека БГУИР