МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

федеральное государственное автономное образовательное учреждение высшего образования
«Самарский национальный исследовательский университет имени академика С.П. Королева»
(Самарский университет)

Институт информатики и кибернетики
Кафедра геоинформатики и информационной безопасности

ОТЧЕТ ПО ПРАКТИКЕ

Вид практики: производственная
(учебная, производственная)

Тип практики: Проектно-технологическая практика

Сроки прохождения практики: с 18.08.2025 г. по 31.08.2025 г.
по направлению подготовки 10.05.03 Информационная безопасность автоматизированных систем
(уровень академического специалитета)
направленность (профиль) «Безопасность открытых информационных систем»

Обучающийся группы № 6413-100503D Ле Лок Тхо

Руководитель практики от университета профессор, д.т.н. Сергеев В.В.

Дата сдачи 31.08.2025 г.
Дата защиты 31.08.2025 г.

Оценка _____

Самара 2025

СОДЕРЖАНИЕ

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

федеральное государственное автономное образовательное учреждение высшего образования
«Самарский национальный исследовательский университет имени академика С.П. Королева»
(Самарский университет)

Институт информатики и кибернетики
Кафедра геоинформатики и информационной безопасности

Индивидуальное задание на практику

Студенту группы № 6413-100503D Ле Лок Тхо
Направление на практику оформлено приказом по университету от _378 от 18.07.25 г. на кафедру геоинформатики и информационной безопасности Самарского университета
(наименование профильной организации или структурного подразделения университета)

Планируемые результаты освоения образовательной программы (компетенции)	**Выполнение определенных видов работ, связанных с будущей профессиональной деятельностью (сбор и анализ данных и материалов, проведение исследований)**	**Результаты практики**
ОПК-1	Способен оценивать роль информации, информационных технологий и информационной безопасности в современном обществе, их значение для обеспечения объективных потребностей личности, общества и государства	Выполнен обзор стеганографии и протокола WAVES; показана роль ЦВЗ для защиты ИС и прослеживаемости ИИ-контента. Сформулированы цель, задачи и актуальность работы; отражены правовые/этические аспекты использования ЦВЗ.
ПК-1	Способен выполнять интеграцию программных модулей и компонент, проводить верификацию программных продуктов	Реализованы модули встраивания/извлечения DWT_DCT, DWT_DCT_SVD, LSB. Настроены воспроизводимые эксперименты WAVES, автоматизированы запуск и логирование; подготовлены графики PSNR и радар-диаграммы.
ПК-2	Способен разрабатывать требования и проектировать программное обеспечение	Составлены требования к бенчмарку: датасеты, сценарии атак, метрики (PSNR, TPR/FPR). Спроектирован пайплайн: загрузка → встраивание → атаки → детектирование/оценка
ПК-4	Способен разрабатывать системы защиты информации автоматизированных систем	Оценена робастность алгоритмов к JPEG/Noise/Blur и геометрии; выявлены уязвимости.
ПК-6	Способен оценивать уровень безопасности компьютерных систем и сетей	Проанализированы показатели обнаружения (**TPR/FPR**, при необходимости **BER**); выбран рабочий порог детектора.

Срок предоставления на кафедру отчета о практике: 31.08.2025 г.

Руководитель практики от университета, д.т.н., профессор _____ В.В.Сергеев
(подпись)

Руководитель практики от профильной организации _____ А.И. Максимов
(подпись)

Задание принял к исполнению обучающийся группы № 6413-100503D _____ Ле Лок Тхо
(подпись)

О Т Ч Е Т

о выполнении индивидуального задания

по производственном практике

ВВЕДЕНИЕ

Цель: исследование протокола WAVES, освоение запуска экспериментов; проведение бенчмаркинга алгоритмов цифровых водяных знаков (ЦВЗ) с акцентом на метод НЗБ-встраивания и анализ не менее трёх алгоритмов.

При прохождении практики, руководителем были поставлены следующие задачи:

- Изучить статью по протоколу WAVES и воспроизводимость экспериментов в системе WAVES.
- Спланировать и выполнить бенчмаркинг нескольких алгоритмов ЦВЗ (не менее трёх), начав с НЗБ-встраивания и далее исследуя альтернативные подходы.
- Систематизировать и визуализировать результаты (таблицы, диаграммы), выполнить сравнение и анализ.

Задания необходимо было выполнять последовательно в течение всего времени практики, предоставляя руководителю промежуточные отчеты.

ВЫПОЛНЕНИЕ ЗАДАНИЯ

A. Теоретическая справка по WAVES

WAVES (Watermark Analysis via Enhanced Stress-testing) представляет собой новый протокол для стандартизированной оценки алгоритмов цифрового водяного знака. В отличие от традиционных подходов, WAVES обеспечивает комплексное и объективное тестирование, учитывающее как качество изображения, так и устойчивость встроенного водяного знака к различным видам атак.

Основная цель WAVES:

- Разнообразные стресс-тесты. Протокол моделирует множество типов «атак» (geometric/photometric искажения, сжатие, шум, adversarial- и diffusion-атаки), а также сложные сценарии вроде многократной регенерации (rinsing).
- WAVES одновременно анализирует качество изображения и робастность водяного знака.

WAVES выявляет скрытые слабости современных методов и задаёт практичный стандарт для дальнейшей разработки более устойчивых алгоритмов.

1. Введение

Зачем создан WAVES? Это комплекс для продвинутого стресс-тестирования водяных знаков, созданный на смену разрозненным частным проверкам.

Ключевые особенности WAVES:

- Полный набор проверок: классические трансформации (сжатие JPEG/WebP, масштабирование, поворот, обрезка), фотометрические изменения, добавление шума, а также adversarial-атаки и атаки регенерации (diffusion-/VAE-модели).
- Новые «жёсткие» сценарии: многократная регенерация (rinsing) и многоразовое встраивание (multi-embedding), нацеленные на разрушение знака.

Конечная цель: оценка устойчивости при фиксированном качестве и метриках обнаружения (TPR/FPR), с упором на низкие ложные срабатывания.

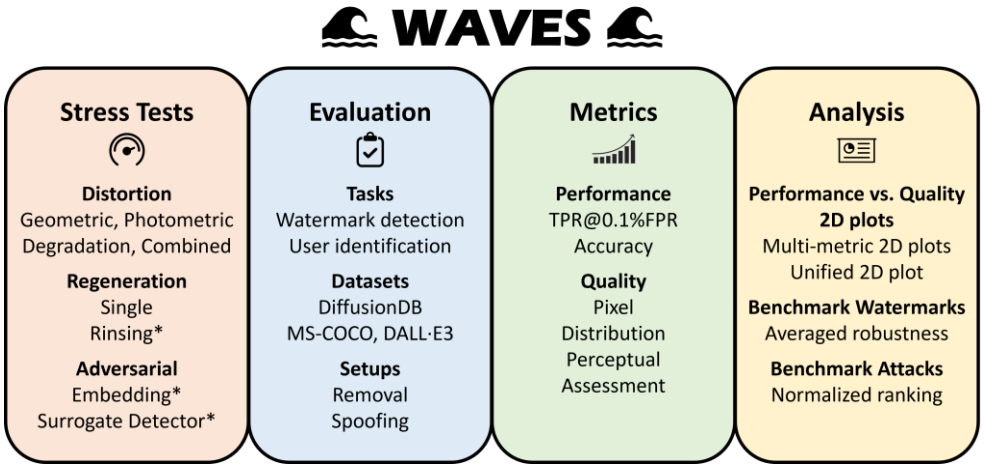

Рис. 1. Стандартизированная структура оценки WAVES и набор стресс-тестов, включая новые усиленные атаки

1.1. Типы атак (Stress-testing)

Протокол WAVES включает три основных класса атак:

Distortion Attacks – геометрические и фотометрические искажения (поворот, кадрирование, сжатие JPEG, добавление шума, изменение яркости/контрастности). Эти атаки моделируют наиболее распространённые сценарии обработки изображений.

Regeneration Attacks – повторная генерация изображений с использованием моделей машинного обучения (Diffusion models, VAE). Данный тип атак способен полностью уничтожить или значительно ослабить водяной знак, особенно в случае многократной регенерации (rinsing).

Adversarial Attacks – целенаправленные малозаметные возмущения, которые приводят к ошибкам в системах обнаружения. Эти атаки являются наиболее сложными и опасными, поскольку эксплуатируют уязвимости моделей распознавания.

1.2. Оценивание

После атак система применяет обнаружение или декодирование, чтобы извлечь/подтвердить знак.

- Watermark Detection — проверка наличия знака.
- Decoding — извлечение встроенной информации.
- Model Prediction — сопоставление извлечённого знака с исходным шаблоном/пользователем.

1.3. Метрики

Для объективного сравнения алгоритмов WAVES использует следующие группы показателей:

- Качество изображения: PSNR, SSIM, LPIPS, DISTS, CLIP-FID, Aesthetic Score. Все значения объединяются в интегральный индекс Normalized Quality Degradation (Q).
- Производительность обнаружения: TPR (True Positive Rate), FPR (False Positive Rate). Введён новый показатель TPR@0.1%FPR, позволяющий оценивать точность при минимальном числе ложных срабатываний.

Идентификация пользователей: способность правильно соотносить водяной знак с конкретным пользователем при большом числе возможных меток.

1.4. Аналитика

a. Графики «Производительность vs. Качество»

Multi-metric 2D plots. Двумерные графики показывают связку Performance ↔ Quality вместе с метриками обнаружения (Detection Rate/Accuracy) или качества (PSNR/LPIPS), что позволяет визуально сравнивать методы.

Unified 2D plot. Сводный график объединяет несколько метрик в одно наглядное поле, избегая необходимости рассматривать десятки отдельных диаграмм.

b. Benchmark Watermarks

Averaged Robustness. Усреднение устойчивости метода по всем видам атак даёт интегральный показатель «средней прочности», помогающий быстро выявлять наиболее стойкие решения

c. Benchmark Attacks

Normalized Ranking. Помимо алгоритмов, WAVES ранжирует атаки по эффективности разрушения. Это помогает исследователям понимать, какие угрозы наиболее опасны и на что фокусировать защиту.

2. Сильные стороны WAVES

WAVES обладает рядом преимуществ, которые выделяют его среди существующих протоколов оценки:

- Широкий охват атак — включены 26 сценариев, охватывающих как базовые преобразования, так и современные adversarial-атаки.
- Использование крупных и реалистичных датасетов (5000 изображений), что снижает вероятность переобучения и обеспечивает репрезентативность.
- Полный набор метрик — одновременный учет качества изображения и устойчивости водяного знака, включая новый показатель TPR@0.1%FPR.
- Эффективная визуализация — графики Performance vs. Quality позволяют интуитивно сопоставлять разные алгоритмы.
- Оценка не только алгоритмов, но и атак — WAVES ранжирует сами методы разрушения водяного знака, что помогает выделить наиболее опасные угрозы.
- Выявление новых уязвимостей — протокол впервые показал слабые стороны популярных методов (Tree-Ring, Stable Signature, StegaStamp).
- Формирование нового стандарта — WAVES задаёт единую основу для дальнейших исследований и разработки более надёжных алгоритмов.

3. Стандартизированная оценка с помощью WAVES

3.1. Процедура и метрики

WAVES задаёт единый протокол: 3 датасета, 26 атак в 3 классах и 9 метрик качества. Цель — прозрачное и воспроизводимое сравнение устойчивости водяных знаков с учётом компромисса между незаметностью и прочностью. Баланс показывается на 2D-графиках Performance vs. Quality. Производительность оценивается по задачам AI Detection и User Identification с использованием TPR/FPR и точности классификации «со знаком/без знака». Качество изображения измеряется набором метрик (например, PSNR/SSIM, LPIPS/DISTS, CLIP-FID, Aesthetic), которые нормируются и агрегируются в Normalized Quality Degradation.

3.2. Стресс-тесты

Distortion — типовые преобразования без регенерации (поворот, обрезка, resize, шум, сжатие JPEG/WebP, изменение яркости/контраста); базовый практический минимум.

Regeneration — перегенерация генеративными моделями (diffusion, VAE), включая режим rinsing (многократная регенерация), что постепенно «вымывает» знак.

Adversarial — малозаметные возмущения, вводящие в заблуждение детектор (классификация «со знаком» как «без знака» и наоборот).

WAVES охватывает все ключевые сценарии — от повседневных искажений до ИИ-регенерации и целенаправленных атак — и даёт целостную, сопоставимую оценку методов.

4. Результаты анализа

Сравнение трёх популярных алгоритмов (Tree-Ring, Stable Signature, StegaStamp) показало, что:

- StegaStamp демонстрирует наибольшую общую устойчивость, но снижает эстетическое качество изображения.
- Tree-Ring сохраняет надёжность при простых искажениях, однако уязвим к adversarial-атакам и многократной регенерации.
- Stable Signature относительно устойчив к искажениям, но полностью разрушается под воздействием атак регенерации.

Анализ уязвимостей позволил выявить природу слабых мест:

- Уязвимость Tree-Ring объясняется зависимостью от латентного пространства VAE, где целенаправленные возмущения легко разрушают сигнал водяного знака.
- Stable Signature оказывается неэффективным при смене декодера в процессе регенерации.
- Высокая устойчивость StegaStamp обусловлена обучением с использованием множества физических искажений, хотя это приводит к артефактам в изображении.

WAVES формирует новый стандарт оценки цифровых водяных знаков, объединяющий разнообразные атаки, крупные датасеты и интегральные метрики качества. Эксперименты показали, что даже самые современные методы обладают критическими уязвимостями. Таким образом, WAVES служит как инструмент

выявления слабых мест, так и основа для разработки более надёжных алгоритмов водяного знака в будущем.

В. Алгоритм цифрового водяного знака на основе DWT–DCT

Цифровой водяной знак (ЦВЗ) является важным инструментом защиты авторских прав и аутентификации мультимедийных данных. При разработке алгоритмов водяного знака необходимо учитывать два ключевых требования:

- Незаметность (imperceptibility): встроенная метка не должна ухудшать визуальное качество изображения.
- Устойчивость (robustness): метка должна сохраняться после распространённых атак — сжатия, шума, масштабирования и других искажений.

С этой целью широкое распространение получили гибридные методы, в частности алгоритм на основе комбинации дискретного вейвлет-преобразования (DWT) и дискретного косинусного преобразования (DCT). Использование DWT обеспечивает многомасштабное разложение изображения, а DCT позволяет выбрать частоты, оптимальные для внедрения метки.

1. Теоретические основы

1.1. Дискретное вейвлет-преобразование (DWT)

DWT разбивает изображение на четыре поддиапазона:

- LL – аппроксимация (низкие частоты, глобальная структура).
- HL, LH, HH – детали (высокочастотные компоненты по горизонтали, вертикали и диагонали).

Для встраивания метки в коде используется двухуровневое DWT:

- Первое разложение формирует LL, LH, HL, HH.
- Второе разложение применяется к компоненте HL, после чего в поддиапазоне HL2 осуществляется внедрение.

Такой выбор обеспечивает баланс между незаметностью (не искажается LL) и устойчивостью (HL содержит достаточно энергии для надёжного встраивания).

1.2. Дискретное косинусное преобразование (DCT)

Каждый блок изображения размером n × n подвергается DCT. Полученные коэффициенты упорядочиваются по схеме зигзага.

- Низкочастотные коэффициенты (в начале зигзага) отвечают за общее восприятие изображения.
- Высокочастотные коэффициенты легко теряются при сжатии.
- Поэтому встраивание выполняется в среднечастотной зоне (mid-band).

В коде диапазон mid-band выбирается по долям: low = 0.3, high = 0.7 (т.е. отбираются коэффициенты с 30% по 70% длины зигзага).

Псевдослучайные последовательности (PN-sequences)

Для повышения надёжности используется распространённый спектр (spread spectrum):

- Для каждого бита водяного знака генерируются две PN-последовательности (−1,+1).
- Если встраивается бит 0, добавляется PN0, если 1 — PN1.
- Коэффициенты DCT модифицируются по правилу:

$$C' = C + \alpha \cdot PN$$

где α — коэффициент силы встраивания.

2. Процесс встраивания (Encode)

- Применяется двухуровневое DWT к исходному изображению.
- Из поддиапазона HL2 извлекаются блоки 4×4.
- Для каждого блока выполняется DCT.
- В mid-band коэффициенты встраиваются биты водяного знака с использованием PN-последовательностей и коэффициента α.
- Выполняются обратные преобразования: IDCT → IDWT → восстановленное изображение с внедрённым водяным знаком.

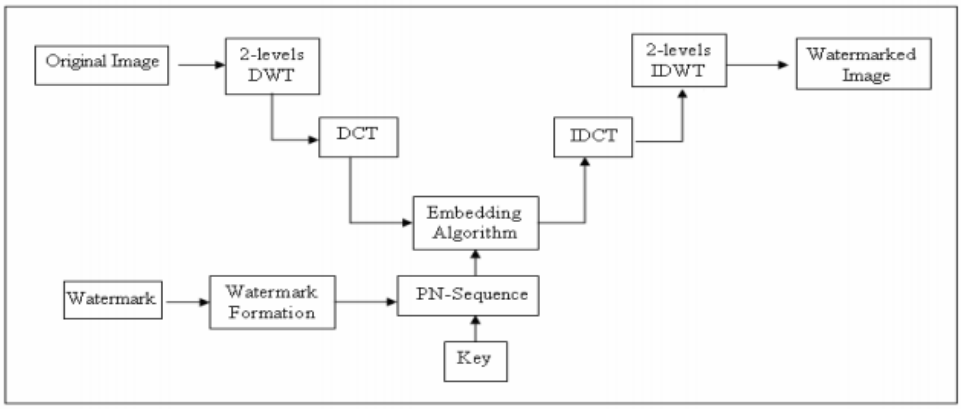

Рис.2: Комбинированная процедура встраивания водяного знака с использованием DWT-DCT.

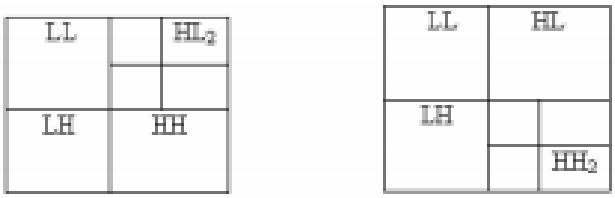

Рис.3: Многомасштабные поддиапазоны DWT исходного изображения.

3. Процесс извлечения (Decode)

- К изображению с водяным знаком применяется то же двухуровневое DWT.
- В поддиапазоне HL2 выполняется блочное DCT.
- Для каждого блока вычисляется корреляция mid-band коэффициентов с PN0 и PN1.
- Побеждает последовательность с большей корреляцией → восстанавливается бит (0 или 1).
- Итоговый водяной знак формируется усреднением результатов по всем блокам.

4. Метрики оценки

Для количественной оценки применяются:

- PSNR (Peak Signal-to-Noise Ratio) — измеряет искажение изображения. Значение выше 30 дБ считается приемлемым.
- Bit Accuracy — доля правильно восстановленных битов:

$$Acc = \frac{\text{число совпавших битов}}{\text{длина водяного знака}} \times 100\%$$

5. Преимущества и ограничения

Преимущества:

- Высокая устойчивость к сжатию JPEG и добавлению шума.
- Незаметность благодаря использованию среднечастотной зоны.
- Возможность регулирования баланса «качество/устойчивость».

Ограничения:

- Относительно высокая вычислительная сложность по сравнению с методами только DWT или DCT.
- Уязвимость к сильным геометрическим атакам (крупное кадрирование, масштабирование).

Алгоритм DWT–DCT является одним из наиболее эффективных гибридных методов цифрового водяного знака. Использование двухуровневого вейвлет-разложения и выбор среднечастотных коэффициентов в DCT позволяет достичь баланса между незаметностью и устойчивостью. Такой подход делает метод перспективным для защиты авторских прав и аутентификации цифровых изображений в реальных условиях эксплуатации

C. Алгоритм цифрового водяного знака на основе DWT–DCT–SVD

Защита авторских прав и аутентификация источника цифрового контента требуют встраивания метки, сохраняющейся после типичных преобразований (сжатие, шум, размытие, умеренные геометрические искажения) и при этом незаметной для зрителя. Для достижения указанного компромисса рассматривается гибридная схема DWT–DCT–SVD.

Её идея заключается в том, чтобы: (i) выбрать устойчивую по энергии область изображения с помощью 2D-вейвлет-разложения; (ii) перейти к блочному частотному представлению через DCT и адресно работать в среднечастотной зоне (mid-band); (iii) кодировать биты посредством квантования индексов (QIM) на наибольшем сингулярном значении SVD, вычисленном для mid-band-части каждого блока.

В дальнейшем излагаются теоретические предпосылки, процедурные шаги, метрики оценки и ключевые проектные решения.

1. Теоретические основы и выбор архитектуры

1.1. Одноуровневое 2D-DWT (вейвлет Хаара)

Пусть входное изображение в градациях серого обозначено II. Одноуровневое разложение даёт:

$$I \xrightarrow{\text{DWT}} (LL_1, LH_1, HL_1, HH_1).$$

Компонента LL_1 аккумулирует основную энергию и глобальные структуры, что делает её сравнительно стабильной к множеству искажений. В предлагаемой конфигурации метка встраивается именно в LL_1; потенциальный риск визуальной заметности компенсируется маской (при наличии), ограничивающей модификации текстурными областями.

1.2. Блочное DCT и mid-band

Поддиапазон LL_1 разбивается на блоки 4 × 4. Для блока BB вычисляется C = DCT(B). Коэффициенты упорядочиваются «змейкой» (zigzag), а mid-band выбирается как долевой интервал $[0.30, 0.70]$ длины зигзага. Такой выбор исключает DC-компоненту (вносит наибольшую заметность) и крайние высокие частоты (наиболее чувствительны к компрессии), обеспечивая тем самым баланс между незаметностью и робастностью.

1.3. SVD для mid-band и QIM-кодирование

Из матрицы C формируется «маскированная» матрица C_{mb}, нулевая вне mid-band. Выполняется разложение

$$C_{\text{mb}} = U\,\Sigma\,V^{\top}, \quad \Sigma = \text{diag}(s_0, s_1, s_2, s_3),\ s_0 \geq s_1 \geq \dots$$

и модифицируется только наибольшее сингулярное значение s_0 с помощью QIM. Пусть шаг квантования $\Delta > 0$. Для бита $b \in \{0,1\}$ задаётся правило:

$$s_0' = \left(\left\lfloor \frac{s_0}{\Delta} \right\rfloor + 0.25 + 0.5\,b\right)\Delta,$$

то есть для b = 0 остаток после деления на Δ тяготеет к 0.25Δ, а для b = 1 — к 0.75Δ. При декодировании бит восстанавливается тестом порога по остатку:

$$\hat{b} = \mathbf{1}\left[(s_0 \bmod \Delta) > 0.5\Delta\right].$$

Широкая разделительная полоса 0.5Δ обеспечивает помехоустойчивое бинарное решение.

2. Процедура метода

2.1. Предобработка и маска

Изображение подрезается до размеров, кратных 4, чтобы корректно разложить LL_1 на блоки 4×4. Бинарная маска (если задана) масштабируется к размеру LL_1 методом ближайшего соседа и использует порог по среднему значению блока (> 0.5) для включения/исключения блока из обработки. Это позволяет локализовать модификации в менее заметных (текстурных) областях.

2.2. Встраивание и извлечение: идеи агрегирования

Каждый блок LL_1 , прошедший маску, несёт один локальный вклад в восстановление очередного бита. Нумерованные блоки циклически «закрепляются» за битами сообщения по индексу n mod L (где LL — длина метки). На этапе извлечения локальные решения агрегируются (усреднение/большинство) по группам, что повышает итоговую надёжность.

3. Метрики оценки и интерпретация

Для количественной проверки применяются:

- PSNR (дБ) — измеряет незаметность посредством отношения пиковой мощности сигнала к мощности ошибки между исходным изображением и с меткой; значения ≥ 30 обычно соответствуют хорошему визуальному качеству.
- Bit Accuracy (%) — доля корректно восстановленных битов по сравнению с исходной меткой; близость к 100% свидетельствует о высокой робастности в рассматриваемых условиях.

4. Анализ проектных решений

Почему LL_1 . Эта компонента энергетически стабильна, что положительно сказывается на стойкости к компрессии и сглаживанию. Потенциальное влияние на видимое качество компенсируется маской, исключающей однородные области.

Роль mid-band и блока 4 × 4. Работа в средних частотах уменьшает как визуальные искажения (не затрагиваем низкие частоты), так и риск стирания при JPEG (не уходим в самый высокий диапазон). Малый размер блока снижает вычислительную сложность и обеспечивает стабильность SVD.

SVD + QIM. Наибольшее сингулярное значение аккумулирует существенную часть энергии и более стабильно к возмущениям; двудольное размещение остатков (окна 0.25Δ и 0.75Δ) с разделителем 0.5Δ повышает надёжность двоичного решения при шуме.

Агрегирование по блокам. Циклическое сопоставление блоков битам сообщения создаёт множество независимых «голосов» для каждого бита; последующая агрегация существенно снижает вероятность ошибки.

5. Параметры и практические рекомендации

- Шаг квантования Δ: увеличение Δ улучшает декодирование в шуме (robustness), снижая при этом PSNR; уменьшение — наоборот. На практике целесообразно подбирать Δ в диапазоне, обеспечивающем PSNR ≥ 30 при требуемой точности.
- Mid-band $[0.30, 0.70]$: разумный выбор для блока 4×4; при других размерах блока границы можно адаптировать, чтобы учесть особенности целевых искажений.
- Маска: полезно строить по картам текстур/краёв (например, Sobel/LoG или saliency), чтобы ограничить модификации перцептуально «дешёвыми» зонами.
- Длина метки LL: при фиксированном числе валидных блоков меньшее LL даёт больше наблюдений на бит и, как следствие, лучшую итоговую точность.

Представленная схема DWT–DCT–SVD реализует устойчивое и перцептуально щадящее встраивание: DWT выбирает стабильную область LL_1 , блочное DCT с mid-band минимизирует видимые артефакты и потери при сжатии, а связка SVD+QIM предоставляет чёткую бинарную решающую процедуру с хорошей помехоустойчивостью. В совокупности с маскированием и агрегированием по блокам метод демонстрирует перспективность для задач защиты прав и аутентификации в реалистичных сценариях обработки изображений.

D. Алгоритм LSB (Least Significant Bit)

Алгоритмы LSB (Least Significant Bit) внедряют двоечную метку в младшие разряды значений пикселей (или коэффициентов в частотной области). Цель — достичь незаметности для наблюдателя при высокой ёмкости (payload). Ограничение подхода: низкая робастность к сильным преобразованиям (сжатие, фильтрация, геометрия). Поэтому LSB целесообразен для задач скрытия данных/лёгкого водяного знака в контролируемой среде, либо как подсистема в гибридных схемах.

1. Модель и обозначения

- Изображение в оттенках серого: $X \in \{0,\ldots,255\}^{H \times W}$. Для цветных изображений рассматриваются каналы R,G,B.
- Сообщение (метка): $m = (m1,\ldots,mL)$, $mi \in \{ 0, 1 \}$.
- Скорость внедрения r (бит/пиксель или бит/канал/пиксель).
- Битовые плоскости: от 0 (LSB) до 7 (MSB) для 8-битного пикселя.
- Канал искажений $T(\cdot)$: сжатие JPEG, шум, размытие, ресэмплинг и т.д.
- Метрики: PSNR/SSIM (качество/незаметность), BER/Accuracy (точность извлечения).

2. Принцип LSB

2.1. LSB Substitution (прямая подстановка)

Идея: присвоить младшему биту пикселя требуемое значение. Если LSB уже равен нужному — пиксель не меняется; иначе изменяется на ± 1.

Плюсы: крайне простая и быстрая реализация, высокая ёмкость (до ~1 bpp для серого и ~3 bpp для RGB при 1 LSB/канал).

Минусы: характерные статистические следы (нарушение парности 2k, 2k+1) — уязвимость к χ^2-тестам, RS-анализу.

2.2. LSB Matching (встраивание ± 1)

Если LSB не совпадает с битом сообщения, пиксель случайно увеличивается или уменьшается на 1.

Плюсы: лучше сохраняет распределение пар (2k, 2k+1), сложнее обнаружить, чем Substitution.

Минусы: робастность не растёт существенно — по-прежнему чувствителен к сжатию/фильтрации.

2.3. Matrix Encoding и адаптивные схемы

- Matrix Encoding (например, Хэмминг (7,4)) уменьшает среднее число изменений на полезный бит ⇒ выше PSNR при той же полезной нагрузке.
- Content-adaptive (по текстуре/краям): встраивание преимущественно в текстурные области, что повышает незаметность и устойчивость.

3. Ёмкость, искажения и незаметность

3.1. Ёмкость

- Серое изображение, 1 LSB: максимум ≈ H × W.
- RGB, 1 LSB/канал: максимум ≈ 3HW бит.
- Использование k > 1 младших битовых плоскостей повышает ёмкость до k bpp/канал, но резко ухудшает качество и устойчивость.

3.2. Искажения и PSNR (оценка)

Для 1-LSB-Substitution при скорости r bpp вероятность изменения пикселя ≈ r / 2. Тогда:

$$\mathrm{MSE} \approx \frac{r}{2}, \qquad \mathrm{PSNR} \approx 10 \log_{10}\left(\frac{255^2}{r/2}\right).$$

Пример: r = 1 bpp ⇒ PSNR ≈ 51 дБ, визуально практически незаметно. При k > 1 (несколько LSB) возможная амплитуда изменения до 2k - 1, MSE растёт, PSNR падает значительно.

4. Устойчивость к преобразованиям

- JPEG/размытие: LSB в пространственной области очень чувствителен — агрессивное сжатие легко разрушает метку.
- Добавочный шум/лёгкая фотокоррекция: переносится ограниченно; устойчивость повышают повторение битов и голосование.
- Геометрия (кроп, поворот, масштаб): возникает рассинхронизация позиций, извлечение затруднено.

- Выбор канала и плоскости: в RGB канал B обычно менее заметен, чем R/G; плоскость 0-я (LSB) надёжна по незаметности, но слабее по устойчивости; 1-я плоскость может немного повысить устойчивость ценой заметности.

LSB — это не «жёсткий» робаст-водяной знак в строгом смысле; он подходит для высокой ёмкости при мягких преобразованиях. Для реальной устойчивости нужен переход к частотным или гибридным методам и/или ECC.

5. Детектирование (стегоанализ) и меры снижения риска

5.1. Инструменты стегоанализа

χ^2-тест, RS-анализ, Sample Pairs, признаки в остаточной области, марковские модели и др. особенно эффективны при большой нагрузке или последовательном (неадаптивном) встраивании.

5.2. Как снизить детектируемость

- Уменьшать скорость внедрения rr.
- Случайно распределять позиции с секретным ключом (PRNG).
- Подготавливать сообщение: сжимать и шифровать (равномернее распределение бит).
- Применять Matrix Encoding для уменьшения числа изменений.
- Использовать адаптивные маски (края/текстуры).
- Избегать насыщенных пикселей (0/255) в LSB Matching.

6. Проектирование системы

- Подготовка сообщения: сжатие + шифрование + при необходимости ECC (Хэмминг/BCH/LDPC).
- Выбор области внедрения: пространство (LSB) — для высокой ёмкости и незаметности; при требованиях устойчивости — частотная область или гибрид.
- Выбор каналов/бит-плоскостей и адаптивной маски (по краям / текстурам / салиентности).
- План позиций по ключу (PRNG) для безопасности.
- Встраивание: Substitution / Matching / Matrix Encoding при контроле нагрузки rr.
- Извлечение: синхронизация позиций, чтение битов, олосование / декодирование ECC, проверка целостности.

- Оценка: PSNR/SSIM (качество), BER/Accuracy (надёжность), базовые тесты стегоанализа (детектируемость).

7. Сильные и слабые стороны

Преимущества

- Крайняя простота и скорость, высокая ёмкость (особенно в RGB).
- Высокая незаметность при умеренной нагрузке (часто PSNR > 50 дБ).

Ограничения

- Низкая устойчивость к JPEG/размытию/геометрии.
- Детектируемость при больших нагрузках или неадаптивном размещении.
- При увеличении числа LSB качество быстро ухудшается.

LSB — базовый и практичный инструмент для скрытия данных и «лёгкого» водяного знака: он обеспечивает высокую ёмкость и отличную незаметность при низкой вычислительной цене. Однако из-за ограниченной устойчивости и детектируемости при повышенной нагрузке LSB требует осмотрительной настройки (малая скорость, случайные позиции по ключу, адаптивные маски, Matrix Encoding/ECC). Для задач, где критична реальная робастность, LSB следует рассматривать как подсистему в составе гибридной архитектуры, а не как самостоятельное решение.

E. Результаты бенчмарков и выводы

В этой части я, опираясь на две диаграммы (радар робастности и столбики PSNR), кратко фиксирую, что именно показывает каждый график, а затем формулирую практические выводы.

1. Результаты бенчмарков

1.1. Незаметность (PSNR)

Это средний PSNR (дБ) по всему набору изображений после встраивания знака. Чем выше PSNR, тем менее заметны артефакты и тем лучше сохраняется качество картинки. Диаграмма показывает чистоту встраивания каждого алгоритма.

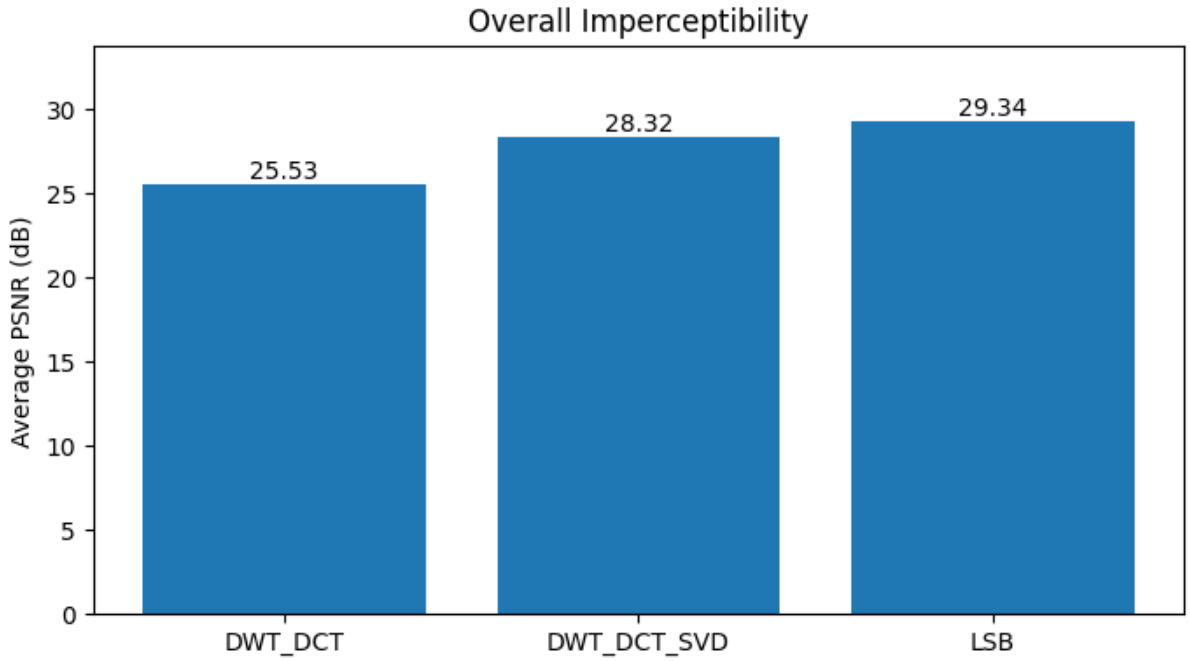

Рис.2: Незаметность после встраивания: средний PSNR (дБ).

- LSB — 29.34 дБ: качество изображения почти не меняется, артефактов минимум.
- DWT_DCT_SVD — 28.32 дБ: лучше, чем DWT_DCT, но чуть хуже LSB.
- DWT_DCT — 25.53 дБ: заметнее всего портит картинку среди трёх.

Итог по качеству:

Порядок ожидаемый — *LSB* > *DWT_DCT_SVD* > *DWT_DCT*.

1.2. Устойчивость к атакам

Каждая ось — отдельный **тип атаки** (JPEG, шум, blur, поворот, ресайз/кроп и т.д.). Радиус — нормированная метрика обнаружения (чем ближе к 1, тем устойчивее алгоритм к этой атаке). Такой график нужен, чтобы видеть сильные и слабые стороны метода по классам искажений, а не только «среднюю температуру».

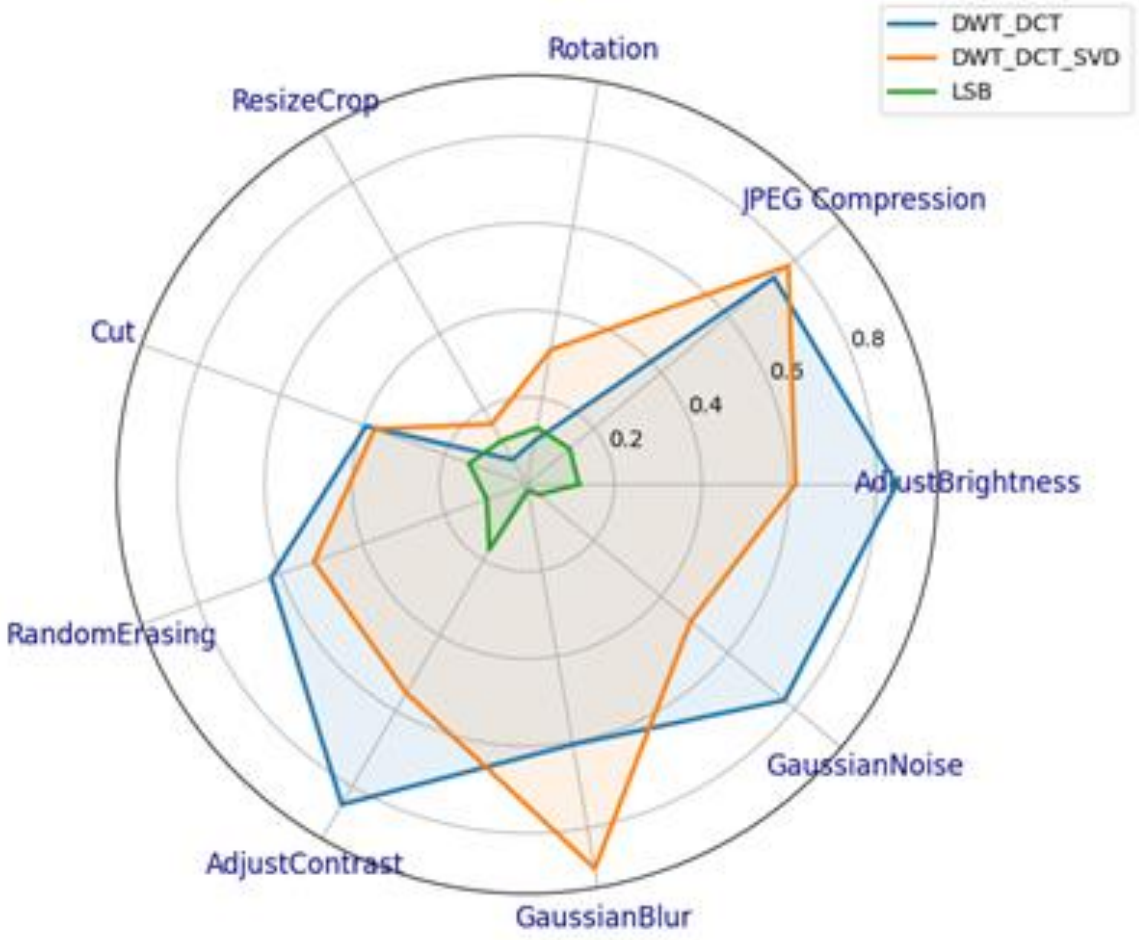

Рис.5: Устойчивость к атакам по видам искажений.

- **DWT_DCT (синий):**

  сильный на JPEG, GaussianNoise, AdjustBrightness/AdjustContrast;

  средний на GaussianBlur, RandomErasing, Cut;

  слабый на Rotation, ResizeCrop.
- **DWT_DCT_SVD (оранжевый):**

  лучший на Blur, хорош на JPEG;

  средний на Rotation, RandomErasing, Cut, GaussianNoise;

  слабее на AdjustBrightness/AdjustContrast и ResizeCrop.
- **LSB (зелёный):**

  значения низкие почти везде → метод хрупкий к большинству искажений.

2. Выводы

Компромисс «качество ↔ робастность» виден чётко:

- LSB — максимум качества, минимум устойчивости;

- DWT_DCT — наоборот: хорошая робастность к сжатию/шуму/фотометрии, но проблемы с геометрией;
- DWT_DCT_SVD — вариант «посередине», особенно хорош при *blur*.

Это классический trade-off: чем незаметнее встраивание, тем ниже робастность.

Где что применять:

- контент для соцсетей/мессенджеров (часто JPEG/шум) → DWT_DCT + желательно геометрическое выравнивание при детекте;
- пайплайны с размытием/ретушью → DWT_DCT_SVD;
- если важна максимальная сохранность вида и пост-обработка минимальна → LSB.

LSB — «чисто, но хрупко»; DWT_DCT — «крепко, но проседает на геометрии»; DWT_DCT_SVD — «компромисс с плюсом на blur».

ЗАКЛЮЧЕНИЕ

В течение практики по получению профессиональных умений и опыта профессиональной деятельности были успешно выполнены поставленные задачи: были изучены протокол WAVES, развернуты воспроизводимые эксперименты и проведён бенчмаркинг трёх алгоритмов ЦВЗ (DWT_DCT, DWT_DCT_SVD, LSB) по метрикам PSNR и устойчивости к типовым искажениям (JPEG, шум, blur, яркость/контраст, поворот, ресайз/кроп, стирание фрагментов).

По качеству изображения порядок ожидаемый — LSB > DWT_DCT_SVD > DWT_DCT. По робастности DWT_DCT силён против сжатия/шума/фотометрии и слаб на геометрии; DWT_DCT_SVD лучше переносит blur, но уступает на яркости/контрасте и геометрии; LSB наиболее хрупок. Тем самым подтверждён компромисс «незаметность ↔ устойчивость».

Практические рекомендации: для сред с компрессией/шумом — DWT_DCT (желательно с геометрическим выравниванием при детектировании); при ожидаемом размытии/ретуши — DWT_DCT_SVD; когда критична визуальная «чистота» и пост-обработка минимальна — LSB. Направления доработки: добавить adversarial/diffusion-атаки, фиксировать TPR@0.1%FPR, включить SSIM/LPIPS и оценить BER/производительность.

За время прохождения практики освоены необходимые компетенции, в частности овладела основами государственной политики РФ в сфере информационной безопасности в условиях современного информационного общества.

ОТЗЫВ
О ПРОХОЖДЕНИИ ПРАКТИКИ

Вид практики производственная
(учебная, производственная, преддипломная)

Тип практики Проектно-технологическая практика
(учебная, производственная, преддипломная)

Сроки прохождения практики: с 18.08.2025 г. по 31.08.2025 г.
по направлению подготовки 10.05.03 Информационная безопасность автоматизированных систем (уровень академического специалитета)
направленность (профиль) «Безопасность открытых информационных систем»
студентом группы № 6413-100503D Ле Лок Тхо

№ п/п	Критерии оценки	Оценка (по 5-балльной шкале)
1.	Общая систематичность и ответственность работы в ходе практики	
2.	Степень личного участия и самостоятельности практиканта в представляемой работе	
3.	Выполнение поставленных целей и задач	
4.	Корректность в сборе, анализе и интерпретации представляемых данных	
5.	Качество оформления отчетной документации	
ИТОГОВАЯ ОЦЕНКА*		

Руководитель практики
от профильной организации _____ А.И. Максимов
(подпись)

* Итоговая оценка выставляется как средняя арифметическая оценок по пяти критериям оценки