

АЛГОРИТМИ ТА МЕТОДИ СТИСНЕННЯ ІНФОРМАЦІЇ

Керівник курсової роботи
Кандидат технічних наук, доцент
Савченко Т.В.

Виконала
студентка 3 року навчання
Свіріна Д.В.

Актуальність: У цифрову епоху обсяг даних зростає експоненційно (175 ZB до 2025 року), що створює навантаження на інфраструктуру. Стиснення даних дозволяє зменшити обсяг збереження і передачі, покращити продуктивність, знизити витрати та енергоспоживання, що критично важливо для мобільних пристроїв, хмарних сервісів і штучного інтелекту.

Мета роботи: Аналіз існуючих алгоритмів стиснення даних та розробка програмного забезпечення для їх реалізації — зокрема, Huffman і LZW (без втрат) для тексту та JPEG (з втратами) для зображень. Реалізація базується на поетапному схематичному поданні роботи кожного з алгоритмів.

Практичне значення:

Розроблене ПЗ забезпечує ефективне стиснення тексту й зображень (Huffman, LZW, JPEG) та може використовуватись для оптимізації зберігання і передачі даних в ІТ-системах.

Класифікація методів стиснення

За ступенем збереження інформації

Тип	Приклади	Характеристика
Без втрат	Huffman, LZW, RLE, BWT	Повне відновлення даних. Використовується для текстів, коду, фінансів
З втратами	JPEG, MP3, МВБ	Часткові втрати, непомітні для людини. Ефективне стискання медіа
Гібридні	PDF, HEIF	Поєднання обох підходів: текст без втрат, зображення з втратами

Таблиця 1- Класифікація методів стиснення

Алгоритм Хаффмана (Huffman coding)

Алгоритм Хаффмана (1952) є статистичним ентропійним методом безвратного стиснення, що ґрунтується на побудові префіксного бінарного дерева за частотами символів.

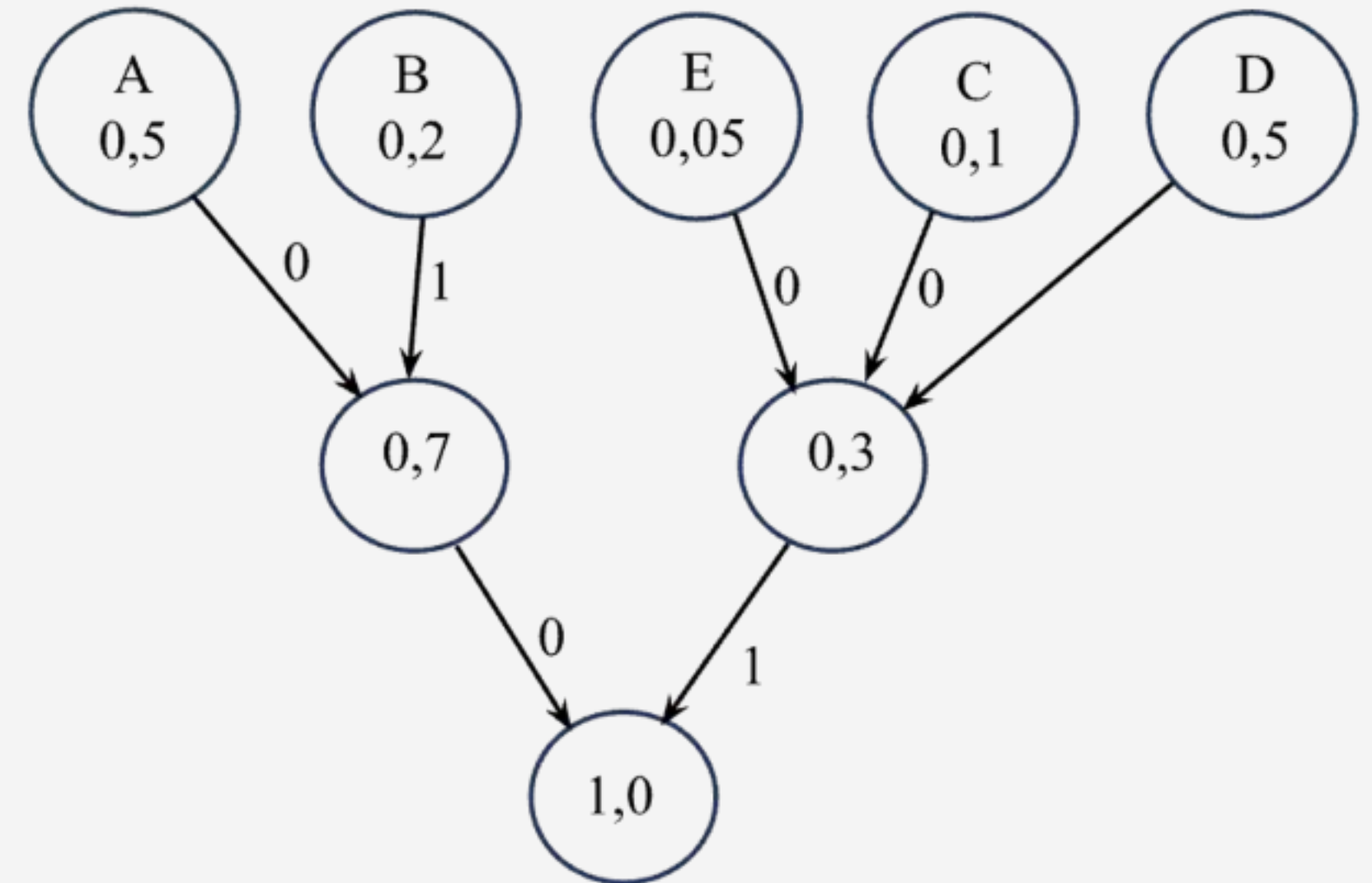
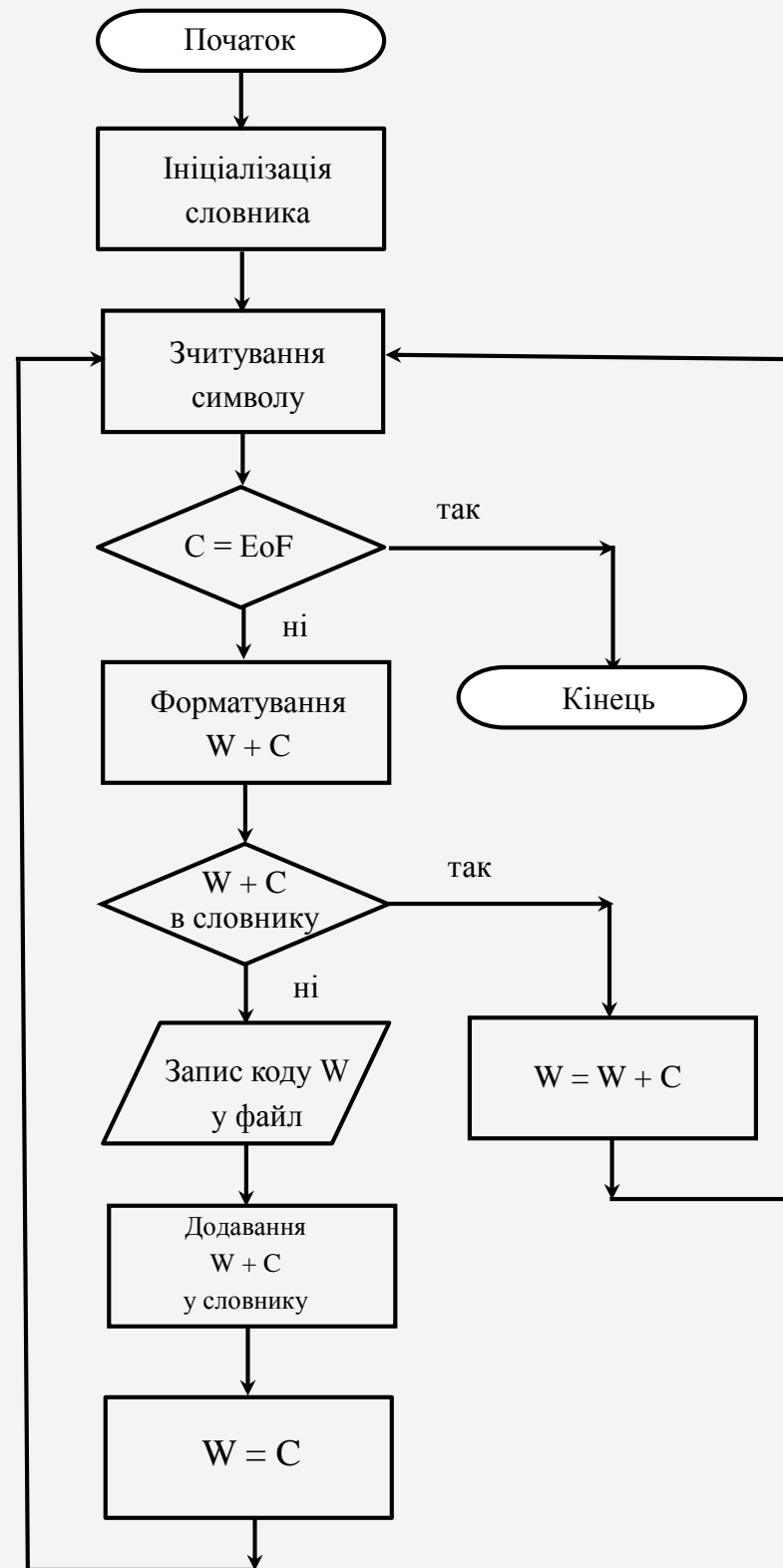


Рисунок 1- Дерево кодування за алгоритмом Хаффмана для прикладу з 5 символами

Алгоритм Лемпеля–Зіва–Велча (LZW coding)



Алгоритм LZW (1984) формує динамічний словник повторюваних послідовностей під час обробки вхідного потоку. Це дозволяє стиснути дані без попереднього аналізу частот, зберігаючи повну відновлюваність.

Рисунок 2 - Схема роботи алгоритму LZW

Алгоритм JPEG (стиснення з втратами)

JPEG — класичний алгоритм стиснення з втратами, що базується на блоковій обробці зображення, частотному аналізі та видаленні малопомітних даних.

Ключові етапи стиснення:

1. Перетворення кольору RGB \rightarrow YCbCr
2. Даунсемплінг кольорових компонент (4:2:0)
3. Блокування зображення на фрагменти 8×8
4. DCT + квантування: перехід у частотну область і округлення коефіцієнтів
5. Зигзаг-сканування: згортання у лінійну послідовність
6. Ентропійне кодування (RLE + Хаффман) \rightarrow бітовий потік.

Порівняльна характеристика алгоритмів

Алгоритм	Тип стиснення	Принцип	Втрати	Застосування
Huffman	Без втрат	Статистичне префіксне дерево	Ні	Тексти, коди, дані з повтореннями
LZW	Без втрат	Словникове динамічне кодування	Ні	Текстові файли, GIF, TIFF
JPEG	З втратами	Трансформаційне DCT+ квантування	Так	Цифрові зображення, фото

Таблиця 2 – Порівняльна характеристика алгоритмів

Реалізація програмного забезпечення Архітектура ПЗ

Основна функціональність програми:

- Вибір файлу або введення тексту
- Вибір алгоритму стиснення (Huffman / LZW / JPEG)
- Стиснення та розпакування
- Попередній перегляд результату
- Збереження (опційно)
- Обмеження на розмір файлів (1 МБ / 5 МБ)
- Підтримка форматів
- Виведення логів та повідомлень

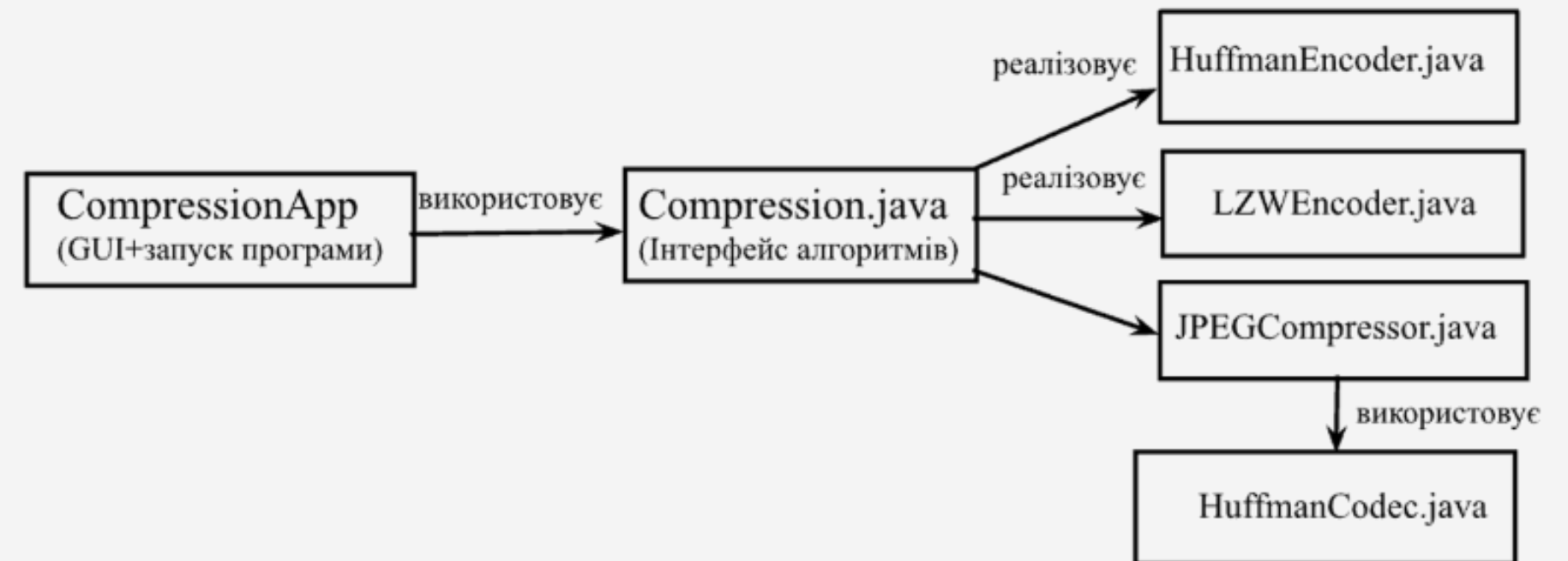


Рисунок 3 – Архітектура програмного забезпечення

Інтерфейс програми та його демонстрація

Реалізовано за допомогою Java Swing — міжплатформенний графічний інтерфейс, що забезпечує зручність для користувача.

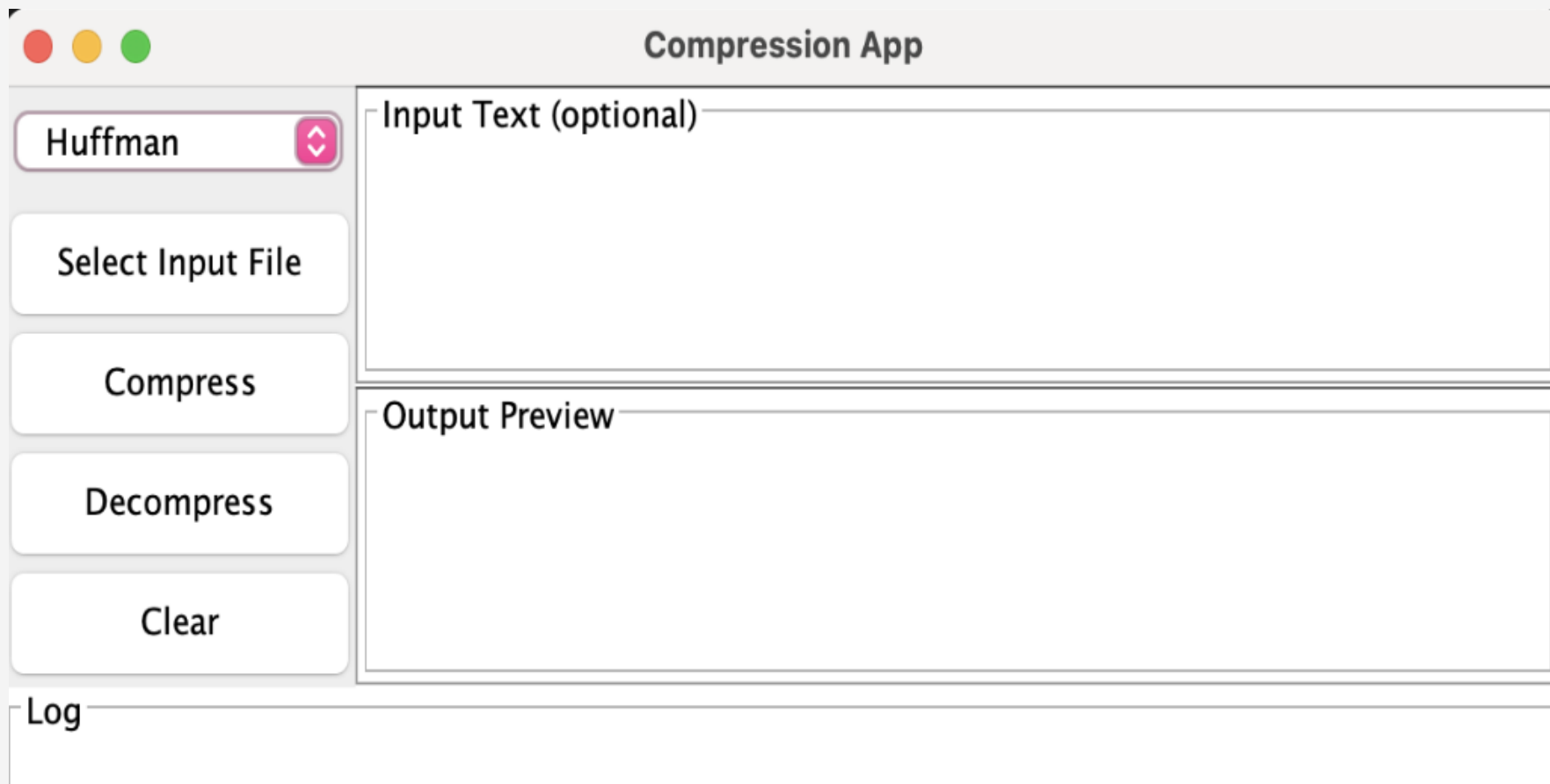


Рисунок 4 – Інтерфейс програми

Панель керування містить кнопки для основних дій:

- **Select Input File** — вибір файлу (текст/зображення)
- **Compress** — стиснення
- **Decompress** — розпакування
- **Clear** очищення полів і стану

Усі дії відображаються в журналі подій або спливаючими повідомленнями.

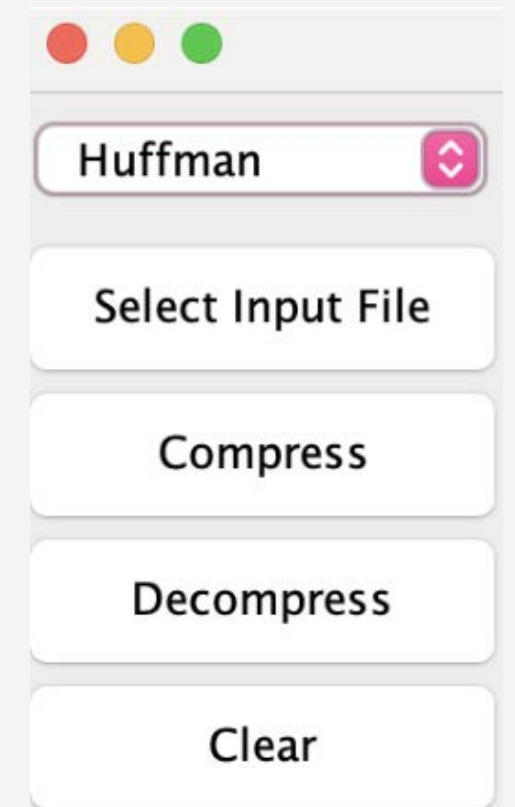


Рисунок 5
Панель керування



Рисунок 6 – Випадаючий список вибору алгоритму

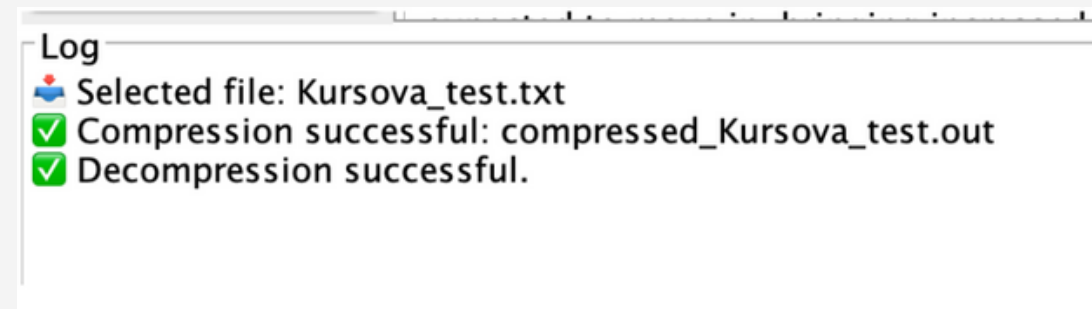


Рисунок 7 – Відображення журналу подій

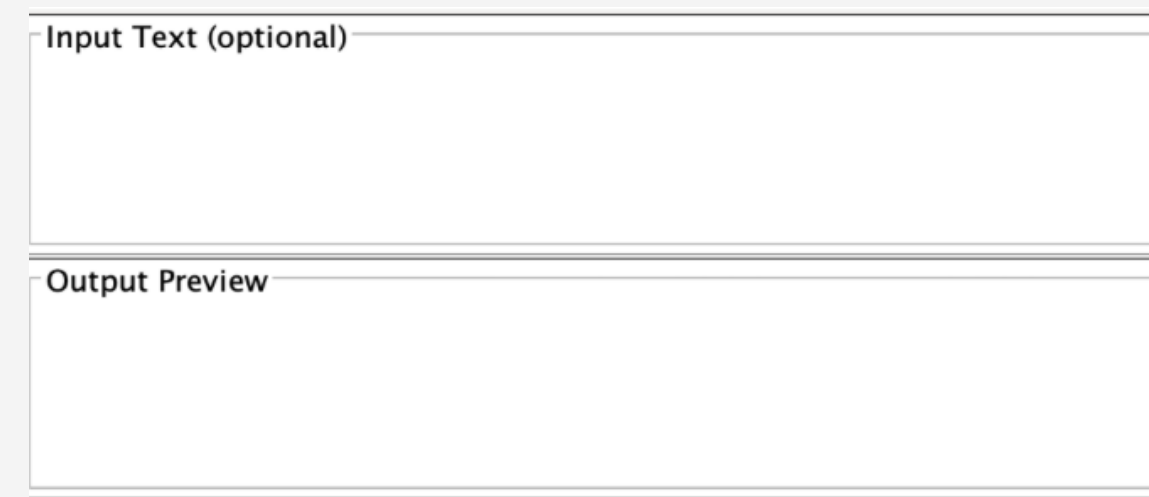
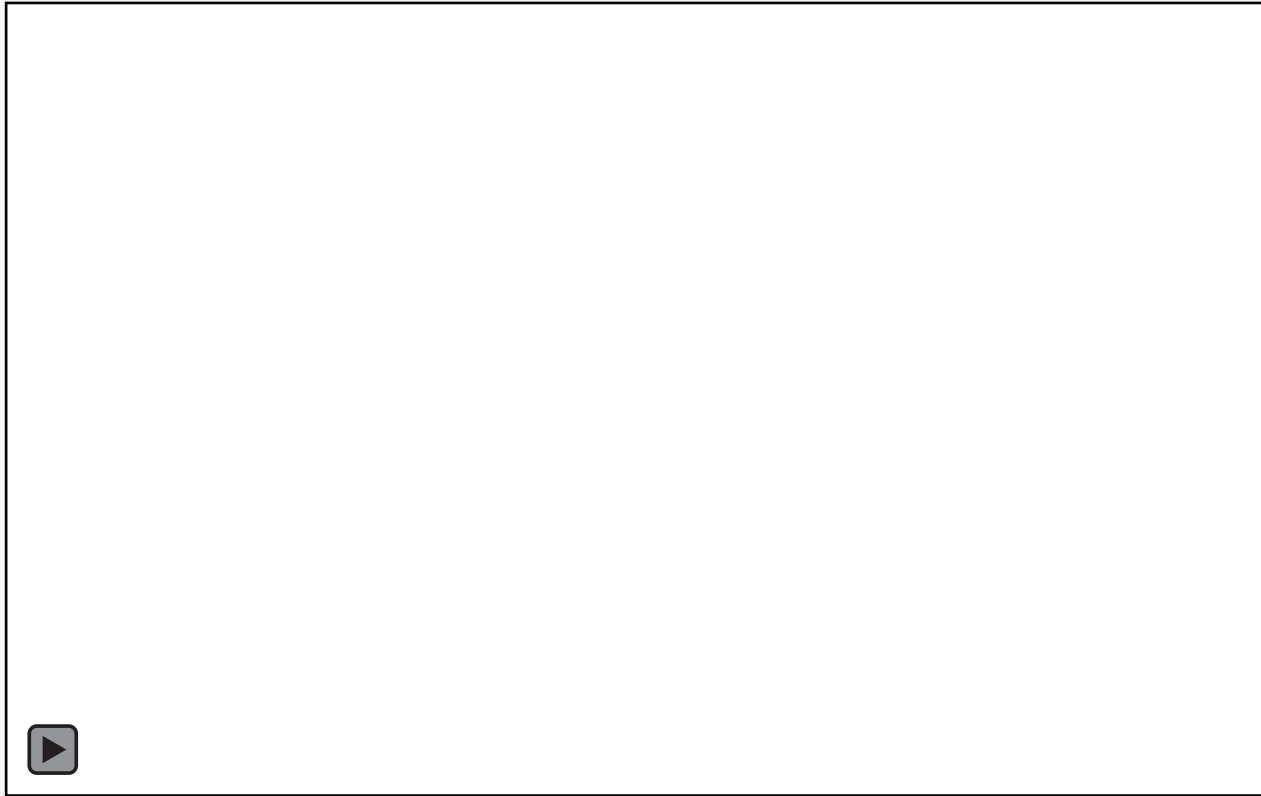
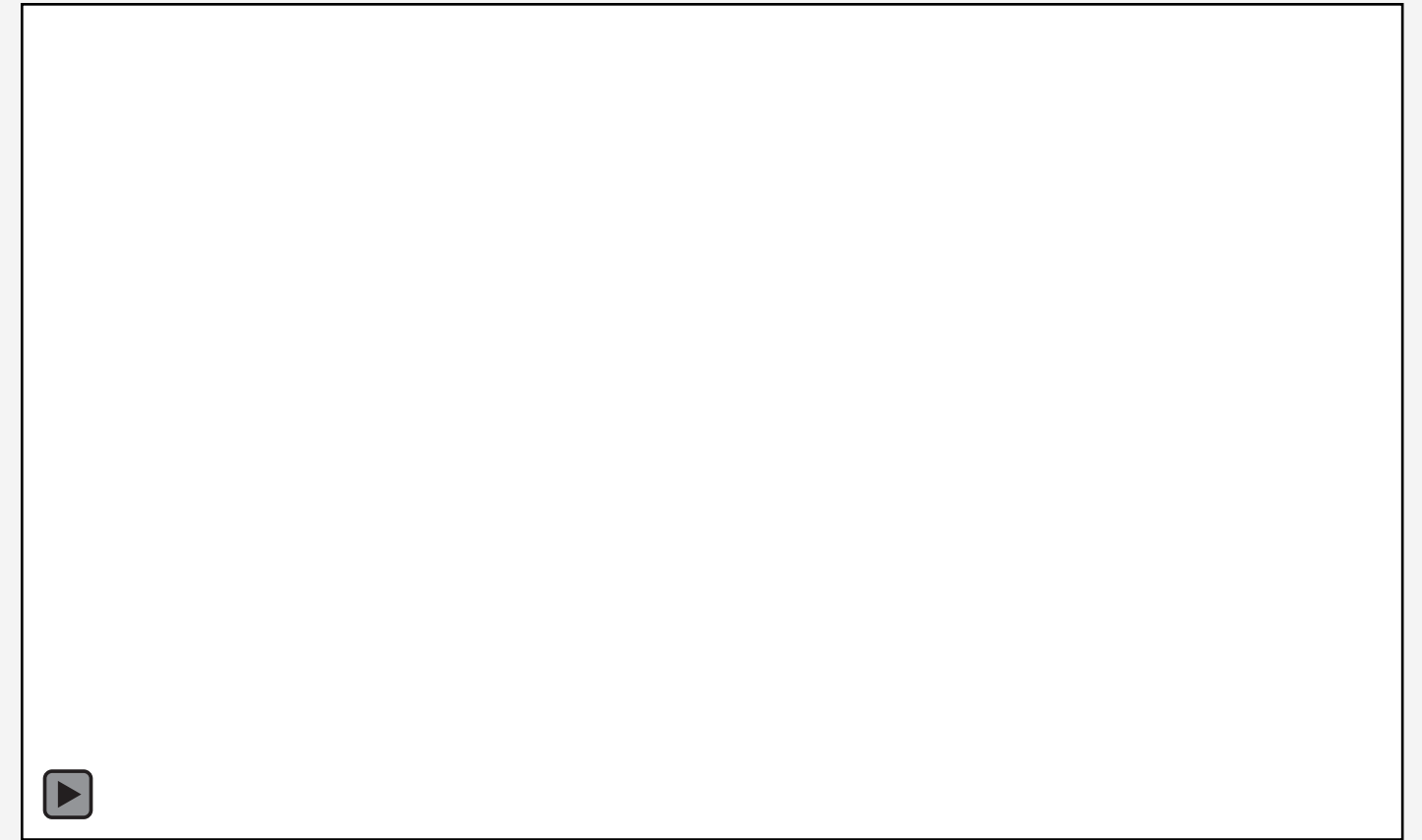


Рисунок 8 – Зона введення та виводу тексту

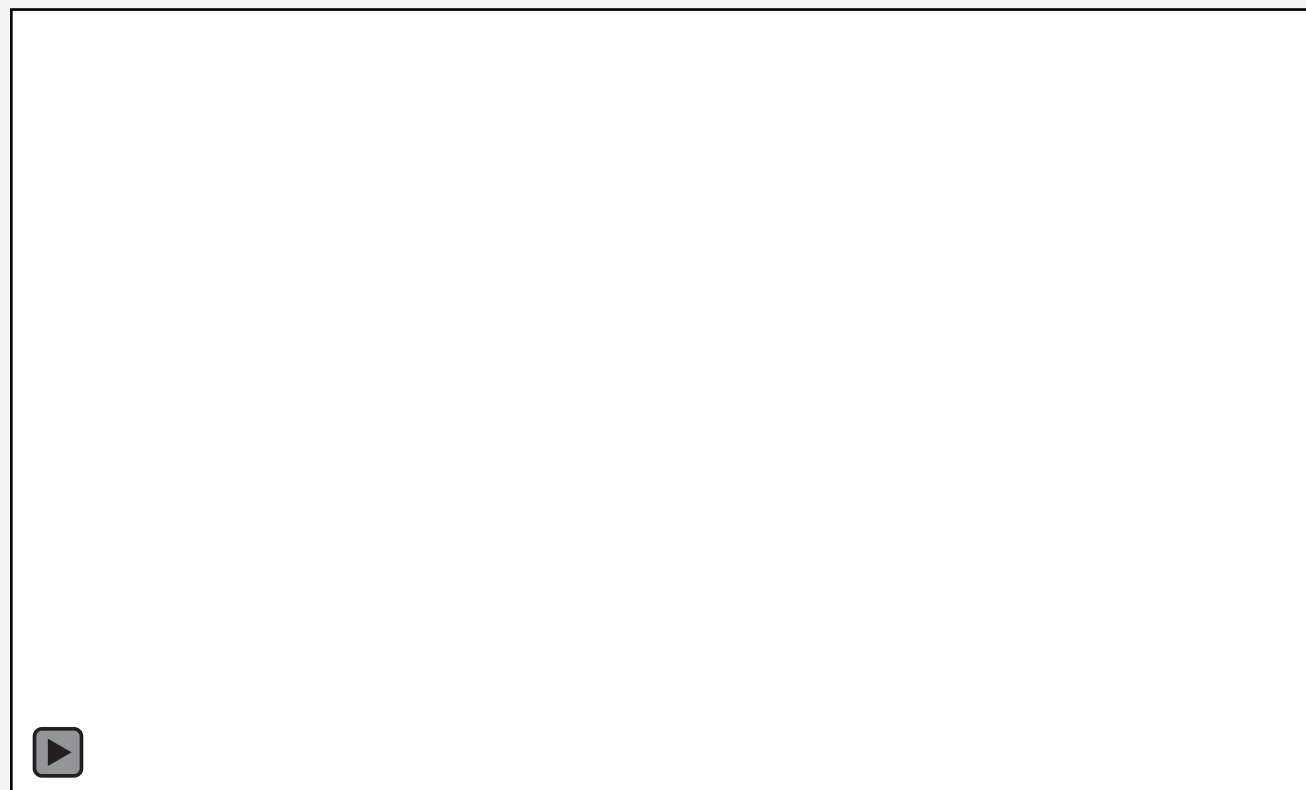
Демонстрація сценаріїв використання



Сценарій 1 – Вибір вхідного файлу



*Сценарій 3 – Стиснення та розпакування
текстового файлу*



Сценарій 2 – Вибір алгоритму стиснення

Демонстрація сценаріїв використання

Сценарій 4 - Стиснення та розпакування зображення у форматі JPEG

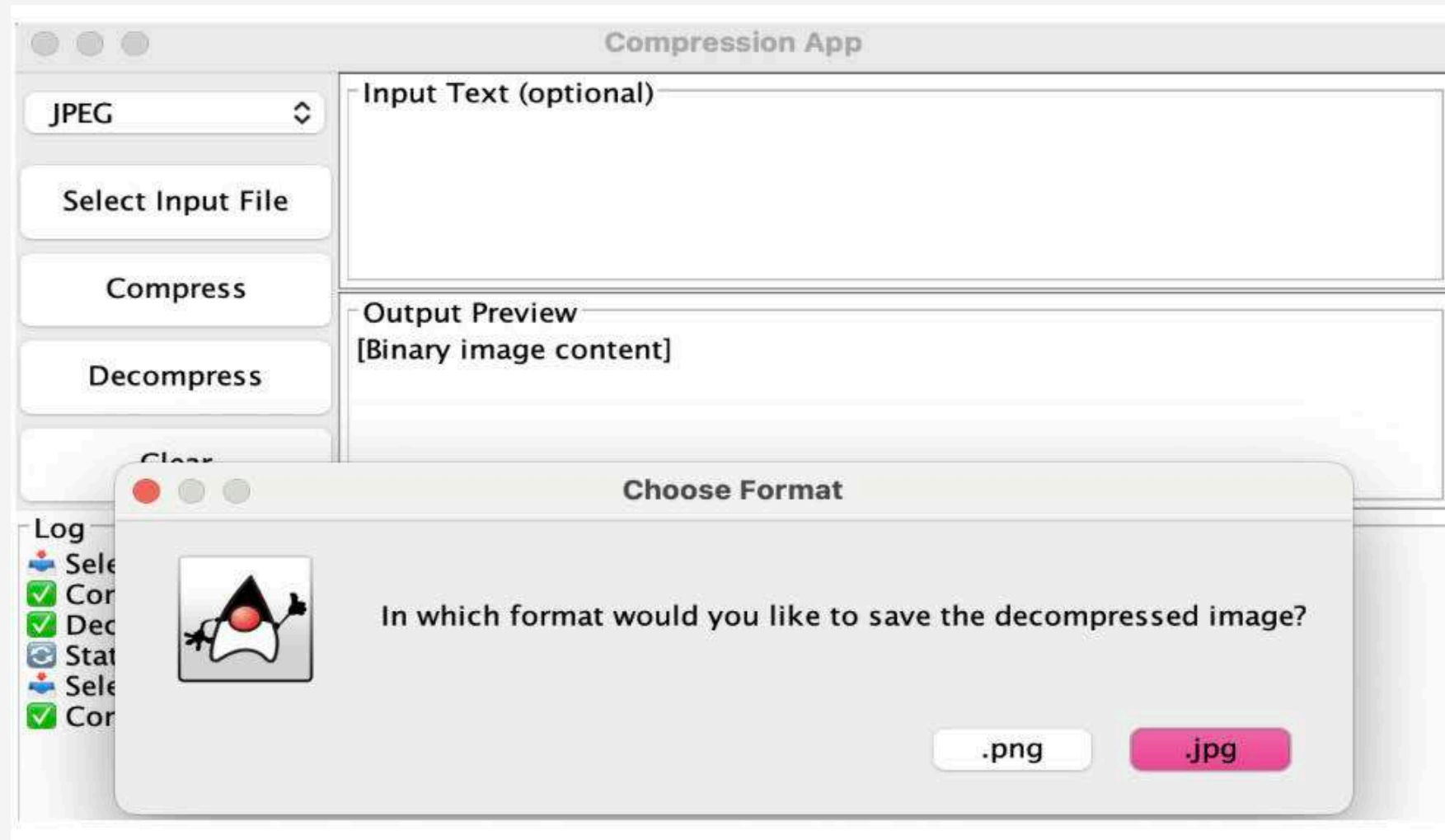


Рисунок 9 – Формат збереження результату

compressed_Kursova_demo_jpegcomp	Сьогодні, 20:11	761 КБ	Документ
decompressed_compressed_Kursova_demo_.jpg	Сьогодні, 20:12	1,1 МБ	JPEG
Kursova_demo_.JPG	16 лют. 2025 р., 10:45	3,6 МБ	JPEG

Рисунок 10 – Демонстрація результатів стиснення та розпакування зображення у форматі jpg

compressed_Test_image_jpegcomp	Сьогодні, 22:44	907 КБ	Документ
decompressed_compressed_Test_image_.jpg	Сьогодні, 22:40	1,2 МБ	JPEG
decompressed_compressed_Test_image_.png	Сьогодні, 22:45	1,2 МБ	PNG
Test_image_.JPG	16 лют. 2025 р., 08:54	3,9 МБ	JPEG

Рисунок 11- Демонстрація результатів стиснення та розпакування зображення у форматі png

Порівняння оригінального та відновленого зображення

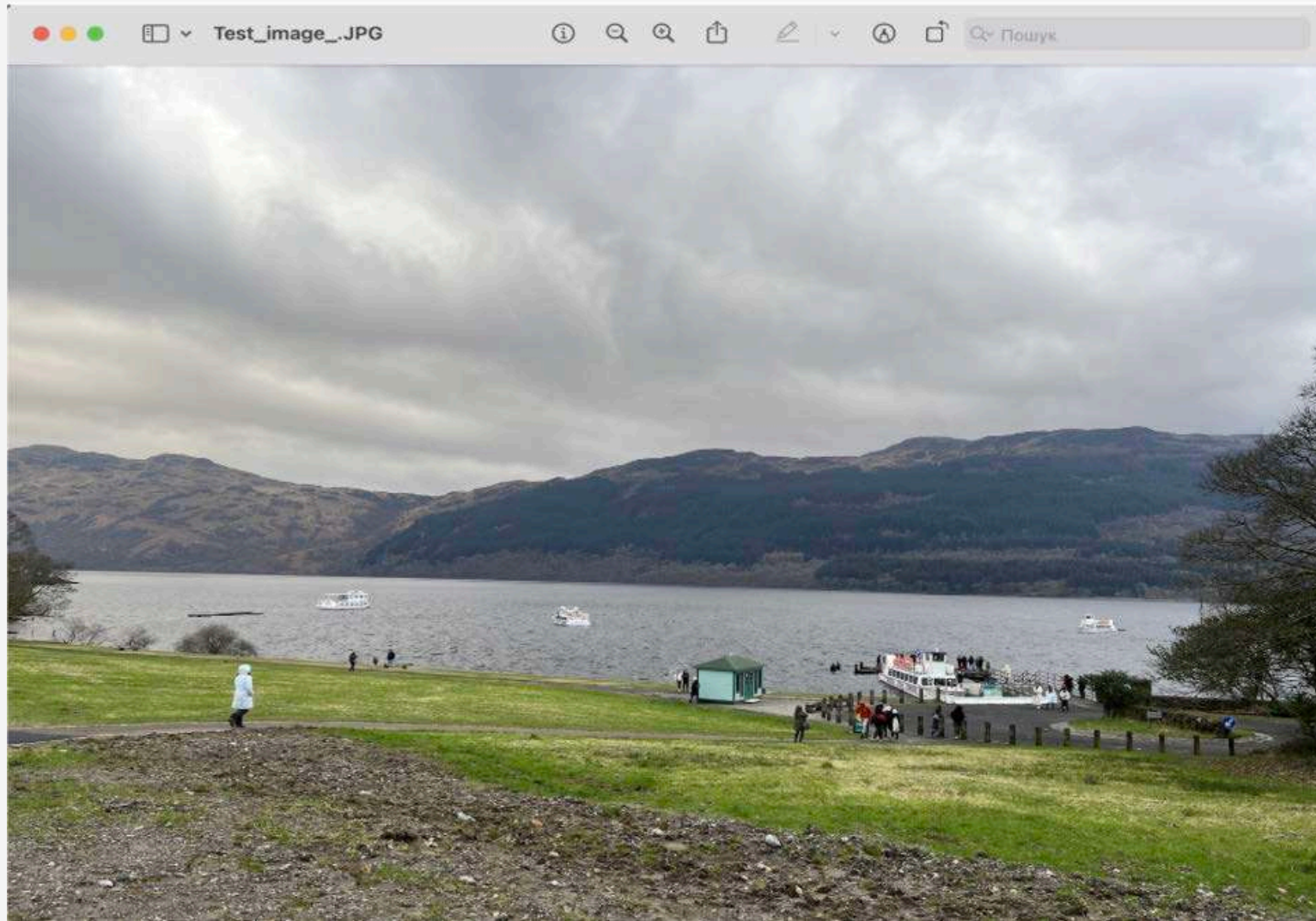


Рисунок 12 – Оригінальне зображення



Рисунок 13 – Зображення після стиснення

Висновки

Аналіз алгоритмів: досліджено методи стиснення без втрат (Huffman, LZW) та з втратами (JPEG), їхні переваги, недоліки і сфери застосування.

Практична реалізація: створено програмне забезпечення на Java з графічним інтерфейсом, яке підтримує стискання і розпакування даних за допомогою трьох алгоритмів. Реалізація не використовує сторонні бібліотеки для стиснення.

Тестування: програма протестована на текстових та графічних даних з обмеженнями розміру файлів і попереднім переглядом результату. Алгоритм JPEG показав високі результати стиснення зберігаючи якість.

Можливості для розвитку:

- Оптимізація продуктивності
- Адаптивне кодування
- Інтеграція з веб- і мобільними платформами

Основні використані джерела

1. Nelson M. The Data Compression Book. – IDG Books Worldwide, 1992. – 624 p.
2. Salomon D. Data Compression: The Complete Reference. – 4th ed. – Springer, 2007. – 1092 p. – Режим доступу: <https://drive.google.com/file/d/1-3FOIXj46HKMwXlSxmruWuLMHMP30wDR/view>.

**Дякую за
увагу!**