W-OP8: Enhancing JPEG XL's Lossless Compression through Genetic Algorithm-Optimized Predictor Weights

Final Report – Undergraduate Thesis

Xavier Hill Roy

CS4490Z

Department of Computer Science

Western University

Date: April 3rd, 2025

Project Supervisor: Mahmoud El-Sakka, Department of Computer Science

Course Instructor: Nazim Madhavji, Department of Computer Science

Glossary

1. **W-OP8**: Weighted average Optimization with 8 Predictors - An enhancement to JPEG XL's compression that expands the original Weighted Average Predictor 4 sub -predictor set to 8 sub- predictors with genetically optimized weights for improved lossless compression efficiency.
2. **JPEG XL**: JPEG XL a new, next-generation image encoding standard developed by the JPEG group
3. **Weighted Average Predictor**: The specific JPEG XL predictor function that is the focus of your research
4. **Modular mode**: JPEG XL's mode specifically designed for lossless compression
5. **Lossless compression**: A data compression technique that allows the original data to be perfectly reconstructed from the compressed data
6. **GA**: Genetic Algorithm
7. **Compression Ratio**: The ratio between the size of the original (uncompressed) data and the size of the compressed data. A higher compression ratio indicates more efficient compression. For example, a compression ratio of 10:1 means the original data is ten times larger than the compressed data.
8. **MED**: Median Edge Detector
9. **GAP**: Gradient-Adjusted Predictor
10. **MA Tree**: Meta-Adaptive Tree
11. **Pixel prediction**: The process of estimating pixel values based on neighboring pixels in lossless compression
12. **CALIC**: Context-based, Adaptive, Lossless Image Codec
13. **JPEG LS**: JPEG Lossless Standard
14. **PNG**: Portable Network Graphics
15. **Codec**: Coder-decoder, a device or program that encodes/decodes digital data
16. **MAE**: Mean Absolute Error
17. **Entropy**: A measure of randomness or unpredictability in image data
18. **Spatial redundancy**: The statistical correlation between neighboring pixels in an image
19. **FLIF**: Free Lossless Image Format
20. **GED**: Gradient Edge Detector
21. **DIV2K**: Dataset of 2K resolution images
22. **CLIC**: Challenge on Learned Image Compression
23. **IAM**: Intelligent Archive Management (handwritten forms dataset)
24. **JPEG**: Joint Photographic Experts Group
25. **AVIF**: AV1 Image File Format
26. **WebP**: Web Picture format

Abstract

Context and Motivation

Lossless image compression preserves all original data, essential for medical imaging, archival storage, and professional photography. JPEG XL, a modern codec, offers significant improvements over legacy formats. Its Weighted Average Predictor function in modular mode uses fixed initial weights that serve as scaling factors, presenting an opportunity for optimization not previously explored.

Research Question

Can JPEG XL's lossless compression techniques be improved to achieve higher compression ratios, and if not, can compression be optimized for specific image types or validated as optimal?

Principal Ideas

This research expands JPEG XL's Weighted Average Predictor from four to eight sub-predictors and employs genetic algorithms to optimize weights for specific image types, leveraging complementary strengths rather than replacing individual components.

Research Methodology

We conducted an empirical study modifying JPEG XL's reference code to incorporate four additional predictors and developed a genetic algorithm to optimize weights across diverse datasets including benchmarking collections, medical ultrasounds, and document images.

Results of Research

The optimized implementation (W-OP8) achieved compression improvements from 0.64% to 2.98% across all datasets, with ultrasound images (2.44%) and document images (2.56%) showing the highest gains, while photographic content demonstrated moderate improvements (1.20%).

Novelty

This research introduces the first ensemble optimization of JPEG XL's Weighted Average Predictor through genetic algorithms and provides the first empirical evaluation of predictor optimization for specialized imagery like medical ultrasounds.

Impact of Results

The compression gains translate to substantial resource savings in large-scale storage systems, with even a 0.99% improvement yielding a 3.5MB reduction across 1.16GB of data—savings that multiply significantly in real-world image repositories.

Table of Contents

Glossary *2*

Abstract *3*

Context and Motivation **3**

Research Question **3**

Principal Ideas **3**

Research Methodology **3**

Results of Research **3**

Novelty **3**

Impact of Results **3**

1 Introduction *7*

2 Background and Related Work *8*

2.1 Overview of Image Compression **8**

2.2 JPEG XL Predictor Functions and Weight Optimization **8**

2.3 Lossless Compression Efficiency: JPEG XL vs Other Formats **9**

2.4 Predictor Functions in Lossless Compression **9**

2.5 Genetic Algorithms in Image Compression Optimization **9**

2.6 Prior Efforts to Enhance Lossless JPEG XL **9**

2.7 Research Gap: Evaluation of JPEG XL Predictor Optimization **10**

3 Research Objectives *10*

4 Methodology *11*

4.1 Datasets and Experimental Setup **11**

4.1.1 Image Datasets 11

4.1.2 Baseline configuration and comparison 11

4.2 Weighted Average Predictor: Sub-Predictor Design & Enhancements **11**

4.3 Genetic Algorithm Optimization and Weight Selection **12**

4.4 Evaluation Methodology **12**

4.4.1 Compression Evaluation and Comparative Analysis 12

4.4.2 Validation Strategy to Prevent Overfitting 12

4.4.3 Lossless Validation 12

4.5 Code Availability and Reproducibility **12**

5 Results *13*

5.1 Baseline Compression Performance **13**

5.2 Individual Predictor Performance: **13**

5.3 Genetic Algorithm Optimization Results **14**

5.4 W-OP8 Performance Evaluation **16**

5.5 Novelty of Results **17**

6 Discussion ***19***

6.1 Interpretation of Key Results **19**

6.1.1 Photography Category 19

6.1.2 Ultrasound Category 19

6.1.3 Document Category 19

6.2. Threats and Limitations **20**

6.3 Implications of Research Results **20**

6.4 Generalizability of Results **21**

7 Conclusion ***21***

8 Future Works and Lessons Learnt ***22***

8.1 Future Work **22**

8.2 Lessons Learnt **22**

9 Acknowledgements ***23***

10 References ***24***

11 Appendix ***27***

Appendix A: Image Datasets **27**

A.1 Standard Benchmarking Datasets 27

A.2 Medical Imaging Datasets 27

A.3 Document Images 27

Appendix B: Details and Technical implementation of 4 new predictors **28**

B.1 Adaptive MED Predictor (prediction[4]) 28

B.2 Adaptive Median Predictor (prediction[5]) 28

B.3 Paeth Predictor (prediction[6]) 28

B.4 GAP-based Predictor (prediction[7]) 28

B.5 Rationale for Selecting These Prediction Strategies 30

B.6 Implementation Details in the JPEG XL Codebase 30

Appendix C: Genetic Algorithm Implementation and Parameters **31**

C.1 Algorithm Design 31

C.2 Chromosome Representation and Initialization 31

C.3 Fitness Function Definition 31

C.4 Selection, Crossover, and Mutation 31

C.5 Elitism and Generational Progression 31

C.6 Parameter Selection 31

C.7 Final Weight Selection 32

C.8 Pseudocode 32

Appendix D Hardware and software **34**

D.1 Hardware: 34

D.2 Software: 34

Appendix E Predictor Weights Distributions **35**

E.1 Predictor Weight Distribution Table 35

E.2 Predictor Weight Distribution Bar Chart 35

E.3 Predictor Weight Distribution Heat Map 36

Appendix F Image Characteristic Analysis **37**

F.1 Image Characteristics and Correlation with Compression Improvements 37
F.2 Correlation Matrix of Image Characteristics and Compression Improvement 37
F.3 Image Characteristics Summary 38

1 Introduction

In the digital era, the efficient storage and transmission of image data remain significant challenges. Lossless image compression, which ensures perfect reconstruction of original files, plays a vital role in applications where data integrity is essential – such as medical imaging, professional photography and document archiving. The challenge of lossless compression methods is to preserve all information while reducing redundancy (Ungureanu, Negirla, & Korodi, 2024).

JPEG XL, a modern image codec standardized in 2023, represents the state of the art in image compression, offering substantial improvements over legacy formats (Alakuijala, Sneyers, Versari, & Wassenberg, 2023). Its modular mode for lossless compression employs sophisticated prediction techniques to exploit spatial redundancies in image data (Alakuijala,et al., 2023). A core function in this system is the Weighted Average Predictor, which dynamically adjusts its weights based on recent prediction errors, making it self-correcting (Rhatushnyak et al., 2019a). Despite JPEG XL's advanced design, the initial weights of this predictor – which serve as scaling factors through the compression process – come from a fixed set of weights in the standard implementation (Rhatushnyak et al., 2019b). This presents an opportunity for optimization that has not been previously explored in the literature.

This research focuses on improving the Weighted Average Predictor function of JPEG XL modular mode using an experimental methodology centred on empirical evolution, comparative benchmarking, and algorithmic optimizations through genetic algorithms. This study addresses a gap in JPEG XL optimization research by expanding the predictor set from four to eight sub-predictors and employing genetic algorithms to optimize their weights for specific image types. These collective improvements are called W-OP8 for **op**timizing the **8** predictor **w**eights.
The key research question guiding this investigation is: Can JPEG XL's lossless compression techniques be improved to achieve higher compression ratios, and if not, can compression be optimized for specific image types or validated as optimal? This question is particularly relevant given that prior attempts to modify JPEG XL predictors have not led to aggregate improvements across diverse datasets (Mamedov, 2024).

Our experimental evaluation across seven diverse datasets – including standard benchmarking collections (Kodak, 1999; Agustsson & Timofte, 2017; CLIC, 2018), specialized medical images (Al-Dhabyani, Gomaa, Khaled, & Fahmy, 2020; Burgos-Artizzu et al., 2020), and document images (Marti & Bunke, 2002) – yielded compression ratio improvements ranging from 0.64% to 2.98%. The most substantial gains were observed for ultrasound images (2.44% average improvement) and document images (2.56% average improvement), while photographic content showed moderate but meaningful improvements (1.20% average improvement). These results confirm that JPEG XL's compression efficiency can indeed be enhanced through targeted predictor optimization.

The novelty of this research lies in three key contributions. First, unlike previous approaches that evaluated individual predictor replacement, our W-OP8 expands the Weighted Average Predictor function to leverage the complementary strengths of multiple predictors simultaneously. Second, we introduce genetic algorithm optimization to JPEG XL's predictor weights, demonstrating intelligent predictor selection across varied datasets. Third, we provide the first empirical

evaluation of JPEG XL predictor optimization for specialized images such as medical ultrasounds, establishing dataset-specific improvement that was potentially unexplored. The significance of these improvements extends beyond academic interest. Even modest compression gains translate to substantial savings in large-scale storage systems or bandwidth-constrained applications. For instance, the 0.99% improvements achieved for the CLIC M dataset resulted in a 3.5 MB reduction across approximately 1.16GB of original data – savings that would multiply significantly in real-world image repositories containing terabytes of data.

This thesis is structured as follows: Section 2 provides background on image compression and related works, establishing the research gap. Section 3 outlines the research objectives. Section 4 details our methodology, including dataset selection, predictor design, genetic algorithm implementation and evaluation approach. Section 5 presents comprehensive results across diverse datasets. Section 6 discusses key findings, implications, threats and limitations and generalisability of our findings. Section 7 concludes with a summary of contributions, while section 8 explores future research directions and lessons learnt. Technical implementation details and supplementary analyses are included in the appendices.

2 Background and Related Work

2.1 Overview of Image Compression

Image compression is categorized into lossy and lossless methods. Lossy compression achieves higher compression ratios by discarding specific image details that are considered less critical (Khobragade & Thakare, 2014). Meanwhile, lossless compression preserves all the image data, ensuring perfect construction of the original file (Khobragade & Thakare, 2014). JPEG XL's modular mode offers an advanced lossless compression approach with dynamic predictor choices based on context and error estimations (Alakuijala et al., 2023).

2.2 JPEG XL Predictor Functions and Weight Optimization

JPEG XL's Predictors: JPEG XL's lossless mode (modular mode) utilizes multiple predictors and a Meta-Adaptive (MA) tree to decide which predictor to use based on its context (Alakuijala et al., 2023). A particularly interesting predictor is the Weighted Average Predictor (Error-correcting predictor), which adjusts its weights based on recent prediction errors, making the overall predictor self-correcting (Rhatushnyak et al., 2019a). The weights are updated and computed using a fixed heuristic function of local prediction errors. In this process, the initial predictor weights get passed into the max weight parameter. This parameter serves as a scaling factor in the error weight calculation. This implementation approach is documented in the JPEG XL codebase by Rhatushnyak et al. (2019b), specifically in the "context_predict.h" file within the "ErrorWeight" function. These weight adjustment formulas (and the four initial weight settings) are built-in and fixed in the standard implementation (Rhatushnyak et al., 2019b, file: context_predict.h, function: ErrorWeight, Predictor_Mode).

Prior Work on Predictor Weights: Despite JPEG XL's advanced prediction system, we found no previous research specifically devoted to optimizing its internal weight parameters and modifying the Weighted Average Predictor. The literature on JPEG XL has focused on performance as is or on introducing new Predictors (Alakuijala et al., 2020; Mamedov, 2024). In

short, while new Predictors have been suggested (Mamedov, 2024), modifying the Weighted Average Predictor and fine-tuning predictor weights remain an open problem. This gap motivates our work to experimentally optimize JPEG XL's Weighted Average Predictor, sub-predictor set and weights to see if compression gains can be made in lossless compressions.

2.3 Lossless Compression Efficiency: JPEG XL vs Other Formats

JPEG XL outperforms modern Codecs like AVIF and WebP (Barina, 2021). However, Barina (2024) observed that FLIF marginally outperformed JPEG-XL in some lossless compression tests. Hence, it suggests that further improvement (even a few percent) in its lossless performance would be noteworthy.

2.4 Predictor Functions in Lossless Compression

Predictors play a crucial role in lossless image compression by estimating a pixel's value based on its context and encoding only the residual (error) (Khobragade & Thakare, 2014). This process takes advantage of the spatial correlation between pixels, reducing redundancy and making entropy coding more efficient (Khobragade & Thakare, 2014). Simple lossless codecs use fixed predictors such as the Median Edge Detector (MED) in JEPG-LS and the Paeth predictor in PNG (Weinberger, Seroussi, & Sapiro, 2000; World Wide Web Consortium, 2003). More advanced approaches, such as the Gradient-Adjusted predictor (GAP) used in CALIC, adapt predictions based on local gradient ranges (Wu & Memon, 1996). While JPEG XL selects between multiple predictors (Alakuijala et al., 2023).

2.5 Genetic Algorithms in Image Compression Optimization

Genetic Algorithms, originally introduced by Holland (1975) and later popularized in engineering by Goldberg (1989), have since been applied in numerous optimization contexts, including image compression. Mitra, Murthy, and Kundu (1998) proposed a GA-based technique to improve fractal image compression by optimizing parameters to achieve better convergence and image quality. Kumar and Karpagam (2015) compared Differential Evolution and GAs in optimizing the quantization table of the JPEG baseline algorithm, demonstrating that GAs can be effectively used to enhance compression performance. These studies highlight the potential of GAs in scenarios where compression quality depends on tuning multiple interdependent parameters. Given that JPEG XL's Weighted Average Function relies on multiple sub-predictors with fixed initial weights, this suggests that a GA could be a promising experimental approach for discovering alternative weight configurations that yield improved compression performance.

2.6 Prior Efforts to Enhance Lossless JPEG XL

Some research has attempted to introduce different predictors in modular mode, such as incorporating the Gradient-Adjusted Predictor (GAP) from CALIC, a modified Median Edge detector (MED) and a gradient Edge Detector (GED) (Mamedov, 2024). These modifications did not lead to aggregate improvement for the datasets used over the baseline JPEG XL (Mamedov, 2024). Our work tests whether adding new sub-predictors to the Weighted Average Predictor and refining weight assignments enhances JPEG XL's lossless compression across different image types.

2.7 Analysis and Research Gap

Existing literature has established that JPEG XL is highly efficient in lossless compression, and attempts to modify the predictor set have not led to aggregate improvements thus far (Mamedov, 2024). However, a significant research gap exists regarding the optimization potential of the Weighted Average Predictor specifically. While traditional lossless codecs have demonstrated compression gains through predictor enhancements (Alonso, 2021), the possibility of dataset-specific improvements through modifications to JPEG XL's Weighted Average Predictor weights and sub-predictor configurations remains unexplored. This study experimentally evaluates whether alternate weight and sub-predictor configurations can improve compression performance or validate the default implementation as already optimal across different image categories.

3 Research Objectives

This research aims to systematically investigate potential improvements to JPEG XL's lossless compression through predictor optimization. The following specific objectives guide this work:

- **O1**: Expand JPEG XL's predictor function for the Weighted Average Predictor by implementing four additional sub-predictors that complement the existing set.
- **O2**: Develop and implement a genetic algorithm optimization framework to determine optimal predictor weight configurations for different image types.
- **O3**: Evaluate the compression performance of the optimized predictor weights across diverse image datasets, including standard photographic content, medical images and document images.
- **O4**: Analyse the relationship between predictor weights, image characteristics, and compression efficiency to identify patterns that explain performance variation across image types.
- **O5**: Quantify improvement potential in JPEG XL's lossless compression through predictor optimization while ensuring perfect reconstruction of original images.
- **O6**: Validate whether dataset-specific optimization yields meaningful compression improvements compared to the standard JPEG XL implementation.

These research objectives correctly address whether JPEG XL's lossless compression can be improved universally through targeted predictor optimization or for specific image types.

4 Methodology

Our methodology combined empirical evaluation with algorithmic optimization to improve JPEG XL's Weighted Average Predictor function. We expanded the predictor set from four to eight and modified the weighting mechanism for determining optimal configurations. Unlike JPEG XL's fixed initial weights, our approach used a GA to evaluate compression performance across multiple generations of candidate weight sets, each assessed using a fitness function based on unaltered JPEG XL compression size.

4.1 Datasets and Experimental Setup

4.1.1 Image Datasets

We used diverse datasets (standard benchmarks, medical, document images) to evaluate predictor performance across varied content. The data was partitioned by selecting the first 10% of the images sorted by name for training and the remaining 90% for testing. Our final analysis focused on the following datasets: Standard Benchmarking Datasets (Kodak Lossless True Color Suite, DIV2K Dataset, CLIC 2018 Compression Challenge Test Datasets P and M); Medical Imaging Datasets (Breast Ultrasound Images, Fetal Ultrasound Images); and Document Images (IAM Handwritten Forms Dataset) (Kodak, 1999; Agustsson & Timofte, 2017; CLIC, 2018; Al-Dhabyani, Gomaa, Khaled, & Fahmy, 2020; Burgos-Artizzu et al., 2020; Marti & Bunke, 2002). Detailed information about each dataset, including their characteristics, sizes, and sources, can be found in Appendix A.

4.1.2 Baseline configuration and comparison

Our testing framework established a baseline using the default JPEG XL library with lossless settings (effort 7, modular_predictor 6 - ensuring only the weighted predictor is used). The experimental pipeline involved: computing baseline compression sizes; training the GA on 10% of each dataset; testing on the remaining 90%; and evaluating performance differences (MAE, total/average size differences) against the baseline.

4.2 Weighted Average Predictor: Sub-Predictor Design & Enhancements

This section goes through the way we addressed objective O1.
The four default predictors are the simple gradient, the adaptive North, the adaptive West and the adaptive multicontext (Rhatushnyak et al., 2019b). Three of the four default sub-predictors use an error correction mechanism (which we refer to as Adaptive) (Rhatushnyak et al., 2019a). The prediction algorithm keeps track of previous errors and feedback them into new predictions (Rhatushnyak et al., 2019b). This information causes the predictors to correct past mistakes to some extent (Rhatushnyak et al., 2019a). In our proposed sub-predictors, we have deliberately maintained this adaptive error correction mechanism in two instances to maintain the nuance of the Weighted Average Predictor.

We implemented four additional predictors to enhance the JPEG XL modular compression system: Adaptive MED Predictor, Adaptive Median Predictor, Paeth Predictor, and GAP-based Predictor. These predictors were selected to complement the existing set. Detailed descriptions and technical implementations can be found in Appendix B.

4.3 Genetic Algorithm Optimization and Weight Selection

This section describes how we addressed objective O2. We implemented a genetic algorithm to optimize predictor weights, representing each candidate solution as a vector of 8 integer weights (0-15). Our fitness function calculated the improvement in compression size relative to the baseline JPEG XL implementation. The optimization process employed tournament selection, uniform crossover, mutation, and elitism strategies across 24 generations. The candidate with the highest fitness from the final generation was selected as the optimal weight configuration. Complete implementation details, including parameter selection rationale and pseudocode, are provided in Appendix C.

4.4 Evaluation Methodology

4.4.1 Compression Evaluation and Comparative Analysis

We assessed optimization effectiveness by calculating compression ratios for both our modified JPEG XL encoder (W-OP8) and the baseline implementation. These ratios were determined by comparing the compressed file sizes against the original uncompressed image sizes. Both encoders used identical encoding parameters, and improvements were quantified as percentage improvement in compression ratio relative to the baseline (Addressing O5).

4.4.2 Validation Strategy to Prevent Overfitting

To ensure generalization, we utilized only 10% of available data for training. Coordinate descent optimization was deliberately omitted after initial experiments showed minimal gains but increased overfitting risk. Our approach balanced specialization and generalization through the genetic algorithm with appropriate population diversity and mutation rates.

4.4.3 Lossless Validation

To verify our W-OP8 implementation preserved lossless compression (addressing objective O5), we conducted pixel-by-pixel comparisons between original images and their decompressed versions. Images were compressed using both baseline and W-OP8 implementations (--distance=0, --modular_predictor=6, --effort=7), then decompressed for comparison. The Mean Absolute Error (MAE) was calculated between original and reconstructed images, as described by Ungureanu et al. (2024), with true lossless compression requiring an MAE of exactly zero, confirming perfect reconstruction.

4.5 Code Availability and Reproducibility

To ensure reproducibility and facilitate further research, all code used in this study, including the modified JPEG XL implementation with the W-OP8 enhancement, genetic algorithm optimization scripts, and evaluation frameworks, is available in our open-source repository at: https://github.com/xavierhillroy/libjxl-lossless-optimization

5 Results

5.1 Baseline Compression Performance

Table 1 presents the baseline compression results, highlighting the significant variation in compression efficiency across different image categories. Note that photographic content: natural images (Kodak, DIV2K, CLIC P, CLIC M) had smaller compression ratios ranging from 2.70:1 to 3.24:1, reflecting their complex textures, broad colour spectra and varied detail levels. Meanwhile, the Ultrasound and Document Images categories showed intermediate compression ratios (8.69:1 to 10.31:1), balancing their characteristics between highly compressible medical images and the more complex photographic content.

Table 1: Baseline Compression Performance

Dataset	Original Size (MB)	Baseline Size (MB)	Compression Ratio
Kodak	24.75	9.17	2.70:1
DIV2K	731.05	258.41	2.83:1
CLIC P	669.54	218.72	3.06:1
CLIC M	1163.13	359.52	3.24:1
Breast Ultrasound	122.41	14.09	8.69:1
Fetal Ultrasound	77.97	7.57	10.31:1
IAM	1331.44	151.44	8.79:1

5.2 Individual Predictor Performance:

Table 2 summarizes the percentage improvement (or degradation) in compression achieved by each predictor compared to the baseline. Figure 1 presents a heatmap of predictor performance across all datasets to visualize better these performance patterns, where red cells indicate improved compression and blue cells indicate degraded compression relative to the baseline. The rightmost column shows the performance of W-OP8 for comparison.

The table and heatmap visualization reveal that various predictors were more suited for specific image types. **Adaptive MED** Performed exceptionally well on photographic and document datasets (Kodak: +2.26%, IAM +2.69%). Meanwhile, **Adaptive West** was effective for ultrasound images (Breast Ultrasound: +2.36%, Fetal Ultrasounds: +0.44%), standing out as the only predictor with positive values in this row. **Paeth** and **GAP-based** showed moderate improvements in photographic images and document datasets, particularly for Kodak and IAM images.

Table 2 Individual Predictor Performance

Category	Dataset	Simple Gradient	Adaptive North	Adaptive West	Adaptive Multicontext	Adaptive MED	Adaptive Median	Paeth	GAP-Based	W-OP8
Photography	Kodak	-2.16%	-4.95%	-2.37%	-1.28%	2.26%	-0.96%	1.53%	1.22%	2.37%
	DIV2K	-4.22%	-5.10%	-2.56%	-3.42%	-0.82%	-2.93%	-1.83%	-1.37%	0.64%
	CLIC P	-3.30%	-4.72%	-2.38%	-2.98%	-0.04%	-1.97%	-0.85%	-0.71%	0.79%
	CLIC M	-3.24%	-5.00%	-2.57%	-2.53%	0.17%	-1.67%	-0.64%	-0.50%	0.99%
Ultrasounds	Breast US	-5.33%	-12.29%	2.36%	-14.38%	-1.43%	-11.69%	-4.48%	-5.66%	2.98%
	Fetal US	-9.17%	-9.70%	0.44%	-14.24%	-3.08%	-16.34%	-8.86%	-7.95%	1.89%
Document	IAM	-2.72%	-5.21%	-1.26%	-3.65%	2.69%	-0.60%	1.32%	1.12%	2.56%

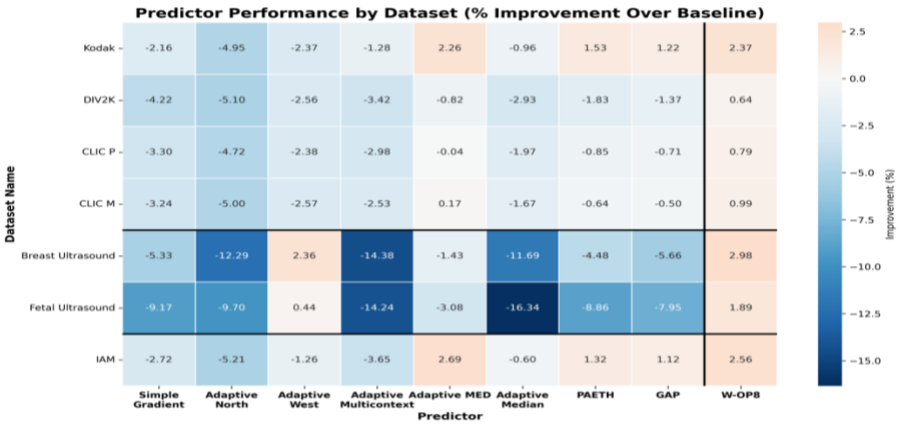

Figure 1 Heatmap of Individual Predictors Across Datasets

Simple Gradient and **Adaptive North** showed negative improvements across all datasets. Meanwhile, **Adaptive Multicontext** and **Adaptive Median** performed poorly on ultrasound images, with degradation exceeding -10% in several cases. Meanwhile, the GA-optimized W-OP8 consistently outperforms individual predictors across all datasets.

The contrast between predictor performance patterns underscores the importance of predictor selection based on dataset analysis, demonstrating the potential value of an optimized ensemble approach that leverages multiple predictors' strengths while mitigating their weaknesses.

5.3 Genetic Algorithm Optimization Results

The genetic algorithm was employed to identify optimal predictor weight configurations for each dataset. Figure 2 illustrates the convergence pattern observed during optimization for the Kodak dataset, showing the progression of fitness scores across generations. As shown in Figure 2, the optimization process demonstrates the rapid initial improvement in the first 3-5 generations, then more gradual improvement in the middle generations (5-15), and finally relative stabilization in later generations (15-23) with smaller incremental gains.

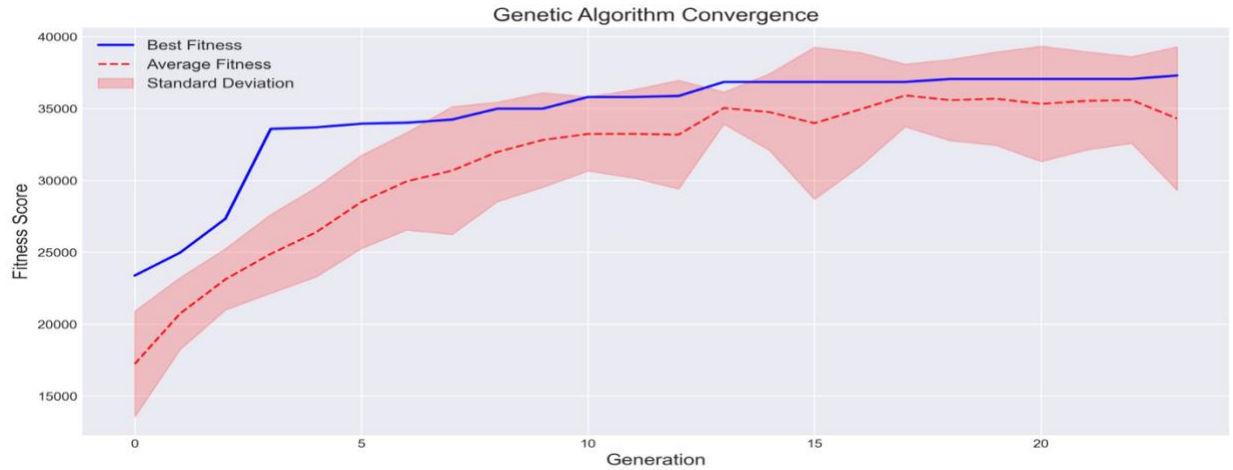

Figure 2 Genetic Algorithm Convergence for Kodak

Table 3 provides specific numerical data points for this optimization process, highlighting key generations and their corresponding weight configurations. Table 3 and Figure 2 information reveal that the steepest improvements occurred in early generations, with fitness gains of over 45% to the initial random weights. The rate of improvement diminished in later generations (17-23), with incremental gains of less than 2% despite continued optimization efforts.

Table 3 Genetic Algorithm Key Generations

Generation	Best Fitness	Average Fitness	Best Weights	Adaptive MED Weight	Improvement (%)	Observation
0	23385	17237.5	[3, 4, 4, 6, 11, 10, 8, 13]	11	0.00	Initial random population
3	33585	24890.9	[1, 0, 1, 8, 14, 2, 12, 12]	14	43.61	Refinement of predictor weights
6	34012	29938.92857	[1, 0, 1, 8, 15, 2, 12, 12]	15	45.44	Fine-tuning near-optimal solution
9	34989	32811.93333	[1, 0, 1, 14, 10, 0, 15, 12]	10	49.62	Fine-tuning near-optimal solution
12	35872	33186.7	[1, 0, 1, 8, 15, 0, 15, 13]	15	53.40	Fine-tuning near-optimal solution
14	36849	34756.16667	[1, 0, 1, 8, 15, 0, 15, 3]	15	57.58	Fine-tuning near-optimal solution
17	36849	35910.83333	[1, 0, 1, 8, 15, 0, 15, 3]	15	57.58	Fine-tuning near-optimal solution
20	37058	35322.3	[1, 1, 1, 8, 15, 0, 15, 3]	15	58.47	Fine-tuning near-optimal solution
22	37058	35591.06667	[1, 1, 1, 8, 15, 0, 15, 3]	15	58.47	Fine-tuning near-optimal solution
23	37297	34305.43333	[1, 1, 1, 8, 15, 0, 11, 3]	15	59.49	Final optimized solution

Similar optimization patterns were observed in other datasets, with convergence occurring within 20-24 generations—the final optimized weights varied by dataset category, reflecting the distinct compression characteristics of each image type. See the final weights selected in Appendix E.

Table 4 shows the Pearson correlation coefficient between individual performance and the Genetic Algorithm weights. This table reveals many critical insights into the genetic algorithm's weight optimization strategy: There is a substantial variation in how closely the genetic algorithm's weight assignments correlate with individual predictor performance across datasets. Strong positive correlations (>0.70) were observed for Kodak, Breast Ultrasound, Fetal Ultrasound, and IAM datasets, indicating that the algorithm effectively identified and prioritized well-performing predictors. Meanwhile, negative correlations were observed for two datasets (DIV2K, CLIC M).

Table 4 Individual Predictor vs GA weight Assignment

Dataset	Category	Best Individual Predictor	Performance (%)	Highest Weighted Predictor	Weight	Match?	Correlation
Kodak	Photography	Adaptive MED	2.26	Adaptive MED	15	Yes	0.72
DIV2K	Photography	Adaptive MED	-0.82	Adaptive North / West	12	No	-0.39
CLIC P	Photography	Adaptive MED	-0.04	Adaptive MED	14	Yes	0.21
CLIC M	Photography	Adaptive MED	0.17	Adaptive North	14	No	-0.35
Breast US	Medical	Adaptive West	2.36	Adaptive West	14	Yes	0.84
Fetal US	Medical	Adaptive West	0.44	Adaptive West	15	Yes	0.78
IAM	Document	Adaptive MED	2.69	Adaptive MED	15	Yes	0.89

As shown in Table 4, the genetic algorithm successfully identified the best-performing predictor in 5 out of 7 datasets and assigned it the highest weight. For example, in Kodak and IAM, Adaptive MED (best performer at 2.26% and 2.69%, respectively) received maximum weight (15). Meanwhile, in Breast Ultrasound, Adaptive West (best performer at 2.36%) received the highest weight (14). At the same time, in Fetal Ultrasound, Adaptive West (best performer at

0.44%) received maximum weight (15). In accordance with objective O2, our genetic algorithm successfully identified optimal predictor weight configurations across diverse datasets.

For datasets where tended predictors performed poorly individually (DIV2K, CLIC M), the genetic algorithm employed a more balanced weight distribution, identifying synergistic combinations that achieved positive overall improvement despite negative individual performances. Surprisingly, the CLIC P dataset had a positive correlation despite no individual predictor performing positively.

These findings demonstrate that the genetic algorithm effectively learned dataset-specific predictor patterns, with stronger performance-weight correlations in datasets where clear individual predictor superiority existed. The algorithm showed intelligence in overcoming scenarios where no individual predictor performed well, discovering combinations that leveraged complementary strengths to achieve positive compression improvements.

5.4 W-OP8 Performance Evaluation

The final W-OP8 implementation, integrating all eight predictors with genetically optimized weights, was evaluated against the baseline JPEG XL implementation. Table 5 summarizes the overall compression performance across datasets, addressing O3. Figure 3 illustrates the differences in improvement visually. We categorized results by image type in Table 6 to better understand performance patterns.

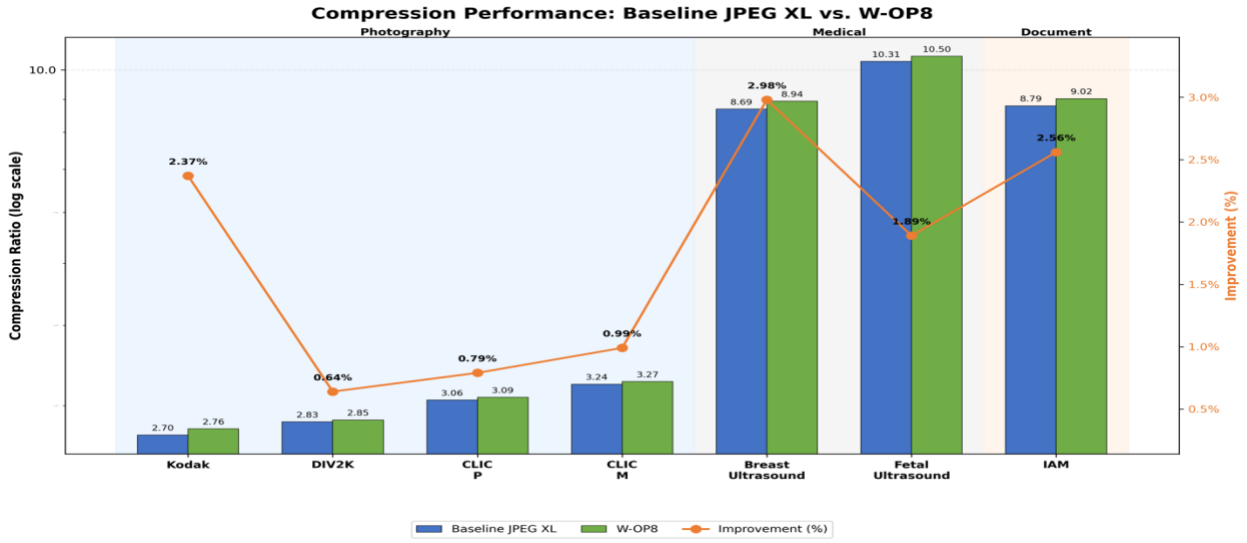

Figure 3 Compression Performance JXL vs W-OP8

Table 5 Baseline VS W-OP8 Compression Ratios

Dataset	Baseline Size (MB)	Baseline Compression Ratio	W-OP8 Size (MB)	W-OP8 Compression Ratio	Improvement (%)	Size Reduction (MB)
Kodak	9.175	2.70	8.963	2.76	2.37	0.212
DIV2K	258.407	2.83	256.765	2.85	0.64	1.642
CLIC P	218.724	3.06	217.001	3.09	0.79	1.723
CLIC M	359.524	3.24	356.001	3.27	0.99	3.523
Breast US	14.093	8.69	13.686	8.94	2.98	0.408
Fetal US	7.566	10.31	7.425	10.50	1.89	0.141
IAM	151.443	8.79	147.664	9.02	2.56	3.779

Table 6 Compression Performance Categorized

Image Category	Datasets	Average Improvement	Improvement Range
Photography	Kodak, DIV2K, CLIC P, CLIC M	1.20%	0.64% - 2.37%
Ultrasounds	Breast Ultrasound, Fetal Ultrasound	2.44%	1.89% - 2.98%
Documents	IAM	2.56%	2.56% - 2.56%

W-OP8 demonstrated apparent performance variation across image categories. For example, Document images showed the highest average improvement (2.56%). Ultrasound images achieved substantial average improvements (2.44%). Photographic content showed moderate improvement (1.20%) (O6).

W-OP8 consistently outperformed individual predictors across all datasets but IAM, demonstrating the effectiveness of the ensemble approach. Even in challenging datasets like DIV2K, W-OP8 achieved positive improvements where individual predictors typically failed.

W-OP8 successfully improved compression across all tested datasets, with optimization effectiveness varying across image types. The most significant benefits were observed for ultrasound and document images.

5.5 Novelty of Results

This research presents several novel contributions to JPEG XL's lossless compression capabilities. First, unlike Mamedov's approach (2024), which evaluated individual predictors (MED, GAP, and GED) as replacements, our W-OP8 implementation expands the Weighted Average Predictor function from four to eight sub-predictors. This ensemble approach leverages complementary strengths of multiple predictors simultaneously rather than substituting individual components.

Second, we introduce genetic algorithm optimization to JPEG XL's predictor weights. Our approach automatically discovers optimal weight configurations for specific image categories, with correlation analysis (ranging from -0.39 to 0.89 across datasets) demonstrating intelligent adaptation to varied image characteristics.

Third, our research expands Mademov's dataset coverage by incorporating specialized medical imaging (breast and fetal ultrasound) and document images (IAM), with ultrasound images

achieving substantial average improvements of 2.44%. This represents the first empirical evaluation of JPEG XL predictor optimization for medical imagery.

Key differences emerge when compared with Mamedov's findings. While Mamedov reported that the GAP predictor performed best on selected images with well-defined edges and flat areas, improvements were inconsistent and generally underperformed the baseline in aggregate. In contrast, W-OP8 achieved positive improvements across all tested datasets (0.64% to 2.98%), demonstrating consistent effectiveness. Furthermore, unlike Mamedov's approach, which requires the manual identification of suitable images, our implementation automatically optimizes for dataset characteristics.

The consistently positive results across diverse image types, coupled with our novel weighted ensemble approach and genetic algorithm optimization, establish this research as a significant advancement in JPEG XL lossless compression optimization, providing a framework for dataset-specific optimization previously unexplored.

6 Discussion

6.1 Interpretation of Key Results

W-OP8 consistently improved compression efficiency across all datasets (0.64%-2.98%) (addressing O6). Though seemingly modest, these gains are significant for lossless compression, particularly for large-scale storage systems and bandwidth-constrained applications.

6.1.1 Photography Category

The photographic datasets revealed variable optimization patterns that resist simple categorization, highlighting the heterogeneous nature of photographic content (see Appendix E for complete predictor weight distributions). The Kodak dataset preferred Adaptive MED (weight 15), achieving a 2.37% improvement, while DIV2K distributed weights more evenly, emphasizing directional predictors (Adaptive North and West, both 12), yielding the lowest improvement of 0.64%.

This variability appears tied to specific image characteristics. Analysis revealed that different predictors performed optimally depending on image content features. For example, as shown in Figure 4 from CLIC Mobile (CLIC 2018), images with repeated horizontal patterns (notice the floor) showed strong performance with the Adaptive West predictor (addressing O4).

Figure 4 Image from CLIC M with strong adaptive West Performance

CLIC P showed preference for Adaptive MED (weight 14) while also giving significant weight to Adaptive West (11). Meanwhile, CLIC M distributed weights more evenly across multiple predictors with Adaptive North and West both at 14, and Adaptive MED at 12. These datasets showed moderate improvements (0.79% and 0.99% respectively).

The performance differences across photographic datasets demonstrate that predictor optimization must be tailored to specific image collections, as even within the photography category, optimal predictor weights vary significantly based on dataset characteristics.

6.1.2 Ultrasound Category

Medical ultrasound images showed the highest compression improvements (~2.44%), with both breast and fetal datasets favoring Adaptive West predictor. This effectively exploits the strong row-wise correlation in ultrasound data (addressing O4) (Hoskins, Martin, & Thrush, 2010). The substantial improvements (2.98% and 1.89%) demonstrate specialized image categories with consistent acquisition parameters benefit most from our optimization approach.

6.1.3 Document Category

Document images (IAM) achieved 2.56% improvement, favoring Adaptive MED (weight 15), with significant contributions from GAP-based and Paeth predictors (both weight 12). Adaptive MED and the GAP-based predictors likely handle the high-contrast edges in handwritten letters,

while and Paeth predictor may have identified structural regularities, suggesting value for document archiving.

6.2. Threats and Limitations

This research faces potential internal validity concerns due to the fixed parameters used in the genetic algorithm (population size, crossover rate, mutation probability), which may limit exploration of better configurations. Additionally, compression results depend on a modified JPEG XL library, which could introduce undocumented behavior. Using only the first 10% of dataset images for training may have introduced selection bias if the images were not randomly distributed, potentially leading to overfitting to specific image characteristics. Computational constraints further limited the number of generations and population size, potentially preventing convergence to the most optimal solutions. A further limitation is the exclusive focus on compression ratio, which overlooks other practical considerations such as encoding speed, memory usage, and computational cost.

To mitigate these issues, we applied consistent benchmarking procedures, validated lossless performance through pixel-by-pixel comparisons, and tested across a wide range of datasets to ensure robustness. Nonetheless, generalizability may be constrained due to the limited diversity of image types and our assumption that datasets are internally homogeneous. Future work should incorporate multiple dataset splits, expanded GA search space (larger populations, more generations), and additional measurement metrics. These adjustments could reduce overfitting, better reflect real-world variability, and enhance the applicability of results across broader image domains.

6.3 Implications of Research Results

Impact on Related Research: Our research provides empirical evidence that JPEG XL's Weighted Average Predictor function can be enhanced through sub-predictor expansion and weight optimization. While JPEG XL already employs adaptive predictor selection during compression (Alakuijala et al., 2023), our findings demonstrate that the initial weights—which serve as scaling factors throughout the compression process (Rhatushnyak et al., 2019b) — should be dynamically determined rather than fixed as in the current implementation. The consistent improvements across diverse datasets (0.64% to 2.98%) suggest substantial value in developing mechanisms that analyze image characteristics and automatically select appropriate initial weights for specific content types. This insight advances understanding of predictor optimization by revealing that optimal initial weight configurations vary significantly by image type, with strong correlations between predictor performance and assigned weights for some datasets (correlation coefficients ranging from 0.72 to 0.89 for Kodak, ultrasound, and IAM datasets). Future codec development could incorporate intelligent initial weight selection based on detected image characteristics rather than relying on universally fixed values.

Impact on Practice: The practical benefits of our work are most evident at scale. Consider Ontario's Hospital Diagnostic Imaging Repository Services, which managed 1.6 petabytes of data as of 2016 (Ontario Ministry of Health, 2016). Applying our W-OP8 optimization's 2.38% average absolute size reduction on ultrasound datasets would yield approximately 38 terabytes of storage savings—a substantial reduction without any compromise in image quality or additional

hardware investment. Similar benefits would apply to document archives and large-scale photography repositories, with these efficiency gains multiplying as data volumes continue to grow.

6.4 Generalizability of Results

Our GA-based weight optimization methodology demonstrates potential generalizability to other compression codecs that employ weighted average prediction functions, extending beyond JPEG XL. Our approach is most suitable for scenarios where compression efficiency takes precedence over encoding speed and computational resources, such as large-scale archival systems and bandwidth-constrained applications. These findings establish both the potential applications and practical limitations of predictor weight optimization across different compression contexts.

7 Conclusion

This research addressed optimizing JPEG XL's lossless compression capabilities, specifically focusing on improving the Weighted Average Predictor. The study investigated whether JPEG XL's compression could be enhanced through sub-predictor weight optimization and sub-predictor expansion (Section 4.2 and 4.3), addressing the research questions of whether higher compression could be improved for specific image types.

Our experimental approach expanded the Weighted Average Predictor sub-predictor set from four to eight sub-predictors and employed genetic algorithm optimization to identify optimal weight configurations directly addressing objectives O1 and O2 (Section 4.3). Testing across diverse image datasets, including photography (Kodak, DIV2K, CLIC M, CLIC P), medical ultrasound images, and document images (section 4.1) demonstrated consistent compression improvements ranging from 0.64% to 2.98% (Table 5, Figure 3) (O3 and O6).

The key findings reveal several novel contributions to lossless image compression:
First, our W-OP8 implementation successfully achieved compression gains across all tested datasets (Table 5), with ultrasound images showing the highest average improvement at 2.44% and document images at 2.56% (Table 6). Unlike previous attempts to modify JPEG XL through direct substitution (Mamedov, 2024), our ensemble approach simultaneously leverages the complementary strengths of multiple predictors.

Second, GA effectively identified dataset-specific optimal weight configurations, with correlation analysis revealing strong relationships between predictor performance and assigned weights (Table 4). A high correlation ranged from 0.72 to 0.89 for datasets with clear predictor preferences. The algorithm showed intelligent adaption, choosing a more balanced strategy with datasets that did not have clear predictor preferences.

We conclude that JPEG XL's lossless compression can indeed be improved through predictor optimization, particularly for specialized image types such as ultrasound and document images. While highly effective, the current JPEG XL implementation leaves room for enhancement through targeted optimization. This research provides a methodological framework for dataset-specific compression improvements that could be particularly valuable for large-scale storage systems or bandwidth-constrained applications handling specialized image types.

8 Future Works and Lessons Learnt

8.1 Future Work

Building on this research, several promising directions for future investigation emerge:

1. **Performance Analysis Across Effort Level**: W-OP8 was tested in an isolated environment (--Modular_predictor=6 & --effort=7), evaluating W-OP8's compression gains across JPEG XL's various effort levels (7, 8 and 9) to determine whether the optimized predictor weights maintain their effectiveness when more complex compression techniques are used.
2. **W-OP4 Comparative Analysis**: Investigate optimizing weights for only the original predictors (W-OP4) to determine whether compression improvements derive primarily from weight optimization or the expanded predictor set, helping establish the minimum intervention needed for meaningful gains. Or even W-OPN where the GA optimizes the weight distribution of N predictors.
3. **Cross-Dataset Generalizability Testing:** Systematically evaluate how predictor weights optimized for one dataset perform when applied to different datasets within the same category (e.g., various types of ultrasound imagery or document collections) to establish the transferability of optimized weights.
4. **Image Characteristic-to-Predictor Heuristic Function:** Develop a heuristic function that analyzes statistical properties of an image (entropy, edge density, etc) and automatically determines appropriate predictor weights without requiring GA optimization, enabling real-time initial weight setting. Could have a built-in dictionary of weights selected for similar properties.
5. **GA Parameter Optimization:** Using different dataset splits, explore larger population sizes (50-200), extended generations (50-100), and varying crossover and mutation rates to potentially discover superior weight configurations, particularly for datasets that showed weaker improvements.

8.2 Lessons Learnt

This research yielded several novel insights that contribute to the understanding of lossless image compression optimization:

1. **Compression Headroom in Modern Codecs**: Our research quantifies previously undocumented optimization potential in JPEG XL, demonstrating improvements of up to 2.98% through targeted predictor optimization alone (Table 5). This challenges the assumption about theoretical compression limits in state-of-the-art codecs.
2. **Ultrasound-Specific Compressive Behaviour:** We identified a pattern where ultrasound images uniquely benefit from the Adaptive West Predictor (Table 2), while most other predictors performed poorly for this modality. This specific insight into medical ultrasound compression characteristics has not been previously documented in JPEG XL research.
3. **Complex GA Optimization Dynamics:** Our analysis revealed that the genetic algorithm can discover non-obvious predictor combinations, shown by the negative correlation between individual predictor performance and final weight assignments for specific

datasets (Table 4 and Appendix E), demonstrating the GA's ability to identify synergistic effects, not apparent from individual predictor analysis.

9 Acknowledgements

I would like to express my sincere gratitude to my supervisor, Professor Mahmoud El-Sakka, for his invaluable guidance and unwavering support throughout this research project. His dedication to meeting regularly, providing timely feedback, and sharing his expertise in image compression was instrumental to the success of this work. His mentorship has not only shaped this thesis but has significantly contributed to my growth as a researcher.

10 References

Agustsson, E., & Timofte, R. (2017). NTIRE 2017 Challenge on single image super-resolution: Dataset and study. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (pp. 1122–1131). https://doi.org/10.1109/CVPRW.2017.150

Alakuijala, J., Boukortt, S., Ebrahimi, T., Kliuchnikov, E., Sneyers, J., Upenik, E., Vandevenne, L., Versari, L., & Wassenberg, J. (2020). Benchmarking JPEG XL image compression. In *Optics, Photonics and Digital Technologies for Imaging Applications VI* (Vol. 11353, 113530X). SPIE. https://doi.org/10.1117/12.2556264

Alakuijala, J., Sneyers, J., Versari, L., & Wassenberg, J. (Eds.). (2023, January). *JPEG white paper: JPEG XL image coding system (Version 2.0)*. ISO/IEC JTC 1/SC 29/WG1. https://jpeg.org/jpegxl/

Al-Dhabyani, W., Gomaa, M., Khaled, H., & Fahmy, A. (2020). Dataset of breast ultrasound images. *Data in Brief, 28*, 104863. https://doi.org/10.1016/j.dib.2019.104863

Alonso, T., Sutter, G., & López de Vergara, J. E. (2021). LOCO-ANS: An optimization of JPEG-LS using an efficient and low-complexity coder based on ANS. *IEEE Access, 9*, 106606–106625. https://doi.org/10.1109/ACCESS.2021.3100747

Alfio, V. S., Costantino, D., & Pepe, M. (2020). Influence of image TIFF format and JPEG compression level in the accuracy of the 3D model and quality of the orthophoto in UAV photogrammetry. *Journal of Imaging, 6*(5), 30. https://doi.org/10.3390/jimaging6050030

Baluja, S. (1995). *Removing the genetics from the standard genetic algorithm*. Carnegie Mellon University.

Barina, D. (2021). Comparison of lossless image formats. *arXiv*. https://doi.org/10.48550/arXiv.2108.02557

Burgos-Artizzu, X. P., Coronado-Gutierrez, D., Valenzuela-Alcaraz, B., Bonet-Carne, E., Eixarch, E., Crispi, F., & Gratacós, E. (2020). FETAL_PLANES_DB: Common maternal-fetal ultrasound images (Version 1.0) [Data set]. *Nature Scientific Reports, 10*, 10200. Zenodo. https://doi.org/10.5281/zenodo.3904280

Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley.

Holland, J. H. (1975). *Adaptation in natural and artificial systems*. University of Michigan Press.

Hoskins, P. R., Martin, K., & Thrush, A. (Eds.). (2010). *Diagnostic ultrasound: Physics and equipment* (2nd ed.). Cambridge University Press.

Jancovic, M. (2017). Lossless compression and the future of memory. *Interactions: Studies in Communication & Culture, 8*(1), 45–61. https://doi.org/10.1386/iscc.8.1.45_1

Khobragade, P. B., & Thakare, S. S. (2014). Image compression techniques—A review. *International Journal of Computer Science and Information Technologies, 5*(1), 272–275.

Kodak. (1999). *Kodak lossless true color image suite*. http://r0k.us/graphics/kodak/

Mamedov, R. (2024). Analysis and enhancement of lossless image compression in JPEG XL [Preprint]. *arXiv*. https://arxiv.org/abs/2404.19755

Marti, U. V., & Bunke, H. (2002). The IAM-database: An English sentence database for off-line handwriting recognition. *International Journal on Document Analysis and Recognition, 5*(1), 39–46. https://doi.org/10.1007/s100320200071

Ontario Ministry of Health. (2016). *Hospital diagnostic imaging repository services (HDIRS)* [PDF]. https://files.ontario.ca/22._hospital_diagnostic_imaging_repository_services.pdf

Rhatushnyak, A., Wassenberg, J., Sneyers, J., Alakuijala, J., Vandevenne, L., Versari, L., Obryk, R., Szabadka, Z., Kliuchnikov, E., Comșa, I.-M., Potempa, K., Bruse, M., Firsching, M., Khasanova, R., van Asseldonk, R., Boukortt, S., Gomez, S., & Fischbacher, T. (2019, August 5). *JPEG XL image coding system* [Committee Draft ISO/IEC 18181]. ISO/IEC JTC 1/SC 29/WG 1. https://jpeg.org/jpegxl/

Rhatushnyak, A., Wassenberg, J., Sneyers, J., Alakuijala, J., Vandevenne, L., Versari, L., Obryk, R., Szabadka, Z., Kliuchnikov, E., Comșa, I.-M., Potempa, K., Bruse, M., Firsching, M., Khasanova, R., van Asseldonk, R., Boukortt, S., Gomez, S., & Fischbacher, T. (2019). *JPEG XL reference implementation* [Computer software]. http://gitlab.com/wg1/jpeg-xl

Scholand, A. J. (n.d.). *GA parameter guidelines*. EIS Lab, Georgia Institute of Technology. Retrieved April 3, 2025, from https://www.eislab.gatech.edu/people/scholand/gapara.htm

Sundararajan, V., & Ayswarya, S. (2019). Genetic algorithm-based feature selection for effective data classification—A survey. *Information, 10*(12), 390. https://doi.org/10.3390/info10120390

Ukrit, M. F., Umamageswari, A., & Suresh, G. R. (2011). A survey on lossless compression for medical images. *International Journal of Computer Applications, 31*(8), 47–50. https://doi.org/10.5120/3850-5353

Ungureanu, I., Negirla, P., & Korodi, A. (2024). Image-compression techniques: Classical and "region-of-interest-based" approaches presented in recent papers. *Sensors, 24*(3), 791. https://doi.org/10.3390/s24030791

Weinberger, M. J., Seroussi, G., & Sapiro, G. (2000). The LOCO-I lossless image compression algorithm: Principles and standardization into JPEG-LS. *IEEE Transactions on Image Processing, 9*(8), 1309–1324. https://doi.org/10.1109/83.855427

Workshop and Challenge on Learned Image Compression (CLIC). (2018). *CLIC 2018 datasets*. https://archive.compression.cc/2018/challenge/

World Wide Web Consortium. (2003). *Portable network graphics (PNG) specification (second edition)*. W3C. https://www.w3.org/TR/2003/REC-PNG-20031110/

Wu, X., & Memon, N. (1996). CALIC—A context-based adaptive lossless image codec. In *Proceedings of the 1996 IEEE International Conference on Acoustics, Speech, and Signal Processing* (Vol. 4, pp. 1890–1893). IEEE. https://doi.org/10.1109/ICASSP.1996.544819

11 Appendix

Appendix A: Image Datasets

This appendix provides detailed information about the datasets used in our study.

A.1 Standard Benchmarking Datasets

A.1.1 Kodak Lossless True Color Suite

A widely-used benchmark consisting of 24 natural photographic images with diverse scenes, textures, and color distributions (Kodak, 1999). This dataset serves as a standard reference point in image compression research and contains images of size 768×512 pixels in 24-bit color.

A.1.2 DIV2K Dataset

High-resolution images with 2K resolution, from which we selected the 100 images in the training directory (Agustsson & Timofte, 2017). These images feature rich details and complex textures that challenge compression algorithms, with resolutions of approximately 2040×1356 pixels.

A.1.3 CLIC 2018 Compression Challenge Test Dataset P

Contains 118 professional photographs with varied content, resolutions, and dynamic ranges, representing high-quality photographic content (CLIC, 2018). These images were captured with professional-grade equipment under controlled conditions.

A.1.4 CLIC 2018 Compression Challenge Test Dataset M

Consists of 168 mobile-captured images with typical smartphone photography characteristics, including varied lighting conditions and occasional motion blur (CLIC, 2018). This dataset represents everyday casual photography.

A.2 Medical Imaging Datasets

A.2.1 Breast Ultrasound Images

133 normal ultrasound images (after removing mask images from directory), characterized by speckle patterns and tissue-specific acoustic properties (Al-Dhabyani et al., 2020). These grayscale images present unique compression challenges due to their noise-like texture patterns.

A.2.2 Fetal Ultrasound Images

Selected first 124 images from a larger dataset (FETAL_PLANES_DB) containing 12,400 common maternal-fetal ultrasound images (Burgos-Artizzu et al., 2020). These images feature varying levels of contrast and speckle patterns.

A.3 Document Images

A.3.1 IAM Handwritten Forms Datasets

59 images from the first inner file (000 directory) containing handwritten text forms (Marti & Bunke, 2002). These high-contrast binary-like images represent document content with distinctive stroke patterns.

Appendix B: Details and Technical implementation of 4 new predictors

Some implementation details were determined through direct examination of the JPEG XL reference code, as certain weight adaptation mechanisms are not explicitly described in published specifications.

B.1 Adaptive MED Predictor (prediction[4])

A modified implementation of the Median Edge Detector (MED) predictor used in the JPEG-LS standard (Weinberger, Seroussi, & Sapiro, 2000). This implementation incorporates error feedback to enhance prediction accuracy. The modular mode of JPEG XL already utilizes a clamped gradient (Rhatushnyak et al., 2019a), which is a variation of the MED, but we aimed to investigate whether this simple predictor could contribute effectively to the Weighted Average Predictor.

B.2 Adaptive Median Predictor (prediction[5])

A median-based predictor incorporating error feedback to improve prediction accuracy. This predictor calculates the median value of neighboring pixels and applies a correction based on recent prediction errors to adapt to local image characteristics.

B.3 Paeth Predictor (prediction[6])

An adaptation of the Paeth predictor from PNG (World Wide Web Consortium, 2003), selecting the neighboring pixel that most closely approximates the local gradient. This predictor is particularly effective for images with sharp edges and rectangular shapes.

B.4 GAP-based Predictor (prediction[7])

This predictor is based on the Gradient-Adjusted Predictor (GAP) used in CALIC (Wu & Memon, 1996). However, our implementation simplifies GAP by focusing solely on edge detection without the complete context-based adaptive coding of CALIC. This predictor adjusts pixel estimates based on local horizontal and vertical gradients, allowing for better adaptation to image structures with sharp edges and smooth regions.

Adaptive MED Predictor

```
int error_const = 4;
pixel_type_w loco_pred;
if (NW >= std::max(N, W)) {
   MED_pred= std::min(N, W) - (((std::min(teN, teW)) * error_const) >> 5);
} else if (NW <= std::min(N, W)) {
   MED_pred= std::max(N, W) - (((std::max(teN, teW)) * error_const) >> 5);
} else {
   MED_pred= (N+W-NW);
}
```

The adaptive MED predictor uses a context-based approach:

- Detects vertical or horizontal edges and applies min/max neighbour values with error correction.

- Defaults to gradient-based prediction in smooth regions.
- Selected error constant and shift amount through manual testing on the Kodak dataset.

Adaptive Median Predictor

```
int avg_error = (teN + teW + teNW) / 3;
int median_pred = (N > W) ?
            ((W > NW) ? W : (N > NW) ? NW : N) :
            ((N > NW) ? N : (W > NW) ? NW : W);
error_const = 5;
int shift = 6;
int corrected_median = median_pred - ((avg_error * error_const) >> shift);
```

The adaptive Median predictor:

- Determines the median of three neighbouring pixels.
- Applies an error correction based on averaged past errors.
- Selected error constant and shift amount through manual testing on the Kodak dataset.

Paeth Predictor

```
int p = N + W - NW;
int pa = abs(p - N);
int pb = abs(p - W);
int pc = abs(p - NW);
prediction[6] = (pa <= pb && pa <= pc) ? N : (pb <= pc) ? W : NW;
```

The Paeth predictor:

- Estimates based on gradient continuity.
- Choose the neighbour with the smallest absolute gradient difference.

GAP-based Predictor

```
pixel_type_w d_h = std::abs(N-NW) + std::abs(NE-N);
pixel_type_w d_v = std::abs(W-NW) + std::abs(N-NN);
pixel_type_w base_pred = (W + N)/2 + (NE - NW)/4;

if (d_v - d_h > 80) {
   GAP_pred= W;
} else if (d_h - d_v > 80) {
   GAP_pred= N;
} else {
   GAP_pred= base_pred;
   if (d_v - d_h > 32) {
      GAP_pred= (GAP_pred+ W) / 2;
   } else if (d_h - d_v > 32) {
      GAP_pred= (GAP_pred+ N) / 2;
   } else {
      if (d_v - d_h > 8) {
         GAP_pred= (3 * GAP_pred+ W) / 4;
      } else if (d_h - d_v > 8) {
         GAP_pred= (3 * GAP_pred+ N) / 4;
      }
   }
```

}

The GAP-Based predictor:

- Detects horizontal and vertical edges with gradient magnitude.
- Adjusts prediction based on the detected edge strength (sharp, medium, weak).

B.5 Rationale for Selecting These Prediction Strategies

These predictors were selected based on their established performance and characteristics in the literature:

- **Adaptive MED**: Implemented as the core predictor in the JPEG-LS standard due to its balance of complexity and effectiveness. Weinberger et al. (2000) demonstrated that MED provides competitive compression performance while maintaining low computational complexity through its simple edge detection mechanism.
- **Adaptive Median**: Implements a computationally efficient approach that incorporates error feedback mechanisms. This predictor builds upon the median filter's robustness while adding adaptive error correction.
- **Paeth**: Chosen for its established role in the PNG standard (World Wide Web Consortium, 2003), demonstrating its reliability in production environments. This predictor offers a unique computation approach compared to the others in our selection, potentially providing complementary strengths to our prediction set.
- **GAP-based**: Incorporated due to its documented success in the CALIC algorithm (Wu & Memon, 1996), where it demonstrated superior prediction accuracy compared to simpler predictors.

These predictors collectively cover a broad range of image content and conditions.

B.6 Implementation Details in the JPEG XL Codebase

Modifications made include:

- Expanded kNumPredictors from 4 to 8.
- New predictors were added to the Predict() method in the Weighted namespace.
- Updated initialization in PredictorMode().
- Modified the Header structure for additional weights.
- Enhanced error tracking and adaptive weighting mechanisms.

These integrations maintain backward compatibility while significantly enhancing compression across diverse image datasets The complete implementation of these modifications can be found in the context_predict.h file in the following repository: https://github.com/xavierhillroy/libjxl-lossless-optimization

Appendix C: Genetic Algorithm Implementation and Parameters

C.1 Algorithm Design

Our genetic algorithm implementation was a modification of the standard design proposed by Goldberg (1989), with parameters based on but modified from literature recommendations (Sundararajan & Ayswarya, 2019; Baluja, 1995).

C.2 Chromosome Representation and Initialization

Each candidate solution was represented as a vector of 8 integer weights, with each weight ranging from 0 to 15. These weights directly correspond to the predictor weights in the Weighted Average Predictor. The genetic algorithm maintained a population of 30 candidates, initialized randomly to ensure genetic diversity.

C.3 Fitness Function Definition

The fitness function quantified compression improvement relative to a baseline, defined as:

$$Fitness = Baseline\ Compression - Candidate\ Compression$$

- **Baseline Compression**: Total compressed size using the factory JPEG XL.
- **Candidate Compression**: Total compressed size using our candidate predictor weights.

C.4 Selection, Crossover, and Mutation

- **Selection**: Tournament selection was utilized, involving three randomly chosen candidates per tournament, with the highest fitness candidate selected as a parent.
- **Crossover**: Uniform crossover was applied at a rate of 0.7, producing offspring by randomly combining weights from parent pairs.
- **Mutation**: Each weight had a 5% chance of mutation, randomly reassigned from 0 to 15, to prevent premature convergence and explore diverse solutions.

C.5 Elitism and Generational Progression

To preserve solution quality across generations, an elitism strategy was employed, retaining the two best-performing individuals from each generation. This approach helps ensure that high-quality solutions are not lost due to the randomness introduced by crossover or mutation. As emphasized by Baluja (1995), elitist selection allows the best solution in a population to persist, contributing to steady generational improvement. This strategy was particularly important in this study due to the limited hardware resources, which limited opportunities for extended convergence.

C.6 Parameter Selection

The parameters for the genetic algorithm—population size (30), crossover rate (0.7), mutation probability (0.05 per weight), and elitism (2 candidates)—were guided by common recommendations and best practices from existing genetic algorithm literature but modified to run within our resource limitations (Goldberg, 1989; Sundararajan & Ayswarya, 2019; Scholand, n.d.; Baluja, 1995). Explicit parameter tuning experiments were not performed due to

computational constraints; instead, parameters were selected to balance exploration effectiveness and practical computational limitations. With additional computational resources, we would have explored larger population sizes and parameters as suggested by the literature (Sundararajan & Ayswarya, 2019; Scholand, n.d.).

C.7 Final Weight Selection

The optimal weight configuration was chosen as the candidate with the highest fitness from the final generation. Due to elitism, this candidate reliably represented the most effective predictor weight set discovered within the given computational constraints.

C.8 Pseudocode

Genetic Algorithm Pseudocode

INPUT:

- NUM_PREDICTORS // Number of genes per candidate
- POPULATION_SIZE // Number of candidates in the population
- GENERATIONS // Total number of generations
- MUTATION_RATE // Mutation probability per gene
- CROSSOVER_RATE // Probability of performing crossover
- ELITISM_COUNT // Number of top candidates to carry over
- MIN_WEIGHT, MAX_WEIGHT // Allowed gene values
- BASELINE // Constant used in fitness evaluation

PROCEDURE GeneticAlgorithm:

```
population ← InitializePopulation(POPULATION_SIZE)
best_candidate ← null
best_fitness ← -∞

FOR generation = 0 TO GENERATIONS - 1 DO:
   Fitnesses ← an empty list

   // Evaluate each candidate
   FOR each candidate IN population DO:
      fitness ← EvaluateCandidate(candidate)
      Append fitness to fitnesses
      LogResult(generation, candidate, fitness)
      IF fitness > best_fitness THEN:
         best_fitness ← fitness
         best_candidate ← copy(candidate)
      ENDIF
   ENDFOR

   Print "Generation", generation, "Best fitness =", best_fitness, "Candidate =", best_candidate

   // Elitism: preserve top candidates
   sorted_population ← population sorted in descending order by fitness
   elites ← first ELITISM_COUNT candidates from sorted_population

   new_population ← copy(elites)

   // Generate offspring until the population is full
```

```
    WHILE size(new_population) < POPULATION_SIZE DO:
        parent1 ← TournamentSelection(population, fitnesses, tournament_size = 3)
        parent2 ← TournamentSelection(population, fitnesses, tournament_size = 3)

        IF random() < CROSSOVER_RATE THEN:
            (child1, child2) ← Crossover(parent1, parent2)
        ELSE:
            child1 ← copy(parent1)
            child2 ← copy(parent2)
        ENDIF

        child1 ← Mutate(child1, MUTATION_RATE, MIN_WEIGHT, MAX_WEIGHT)
        child2 ← Mutate(child2, MUTATION_RATE, MIN_WEIGHT, MAX_WEIGHT)

        Append child1 and child2 to new_population
    ENDWHILE

    population ← first POPULATION_SIZE candidates from new_population

ENDFOR

Print "Best overall candidate:", best_candidate, "with fitness:", best_fitness
```

Appendix D Hardware and software

D.1 Hardware:

- Computing platform: MacBook Pro
- CPU: 2.3 GHz 8-Core Intel Core i9
- RAM: 16 GB
- Storage: 1.0 TB

D.2 Software:

- Operating system: macOS 15.3.1
- Programming Language: Python 3.13.1
- JPEG XL Implementation: v0.11.1 (Rhatushnyak et al., 2019b).
- Key Libraries:
    - **PIL (Pillow)**: Version 11.0.0
    - **Matplotlib**: Version 3.10.1
    - **Numpy**: Version 2.2.0
    - **Pandas**: Version 2.2.3
- Genetic algorithm framework: Custom implementation using Python
- Testing Framework: custom implementation using Python

Appendix E Predictor Weights Distributions

This appendix presents the predictor weight distributions across the seven datasets used in this study, visualized in both bar and heatmap formats. These weights correspond to the eight sub-predictors employed in the W-OP8 Weighted Average Predictor, as determined through the genetic algorithm optimization process described in Appendix C. The figures highlight key patterns in predictor selection across image categories, supporting the analysis in Section 6.1.

E.1 Predictor Weight Distribution Table

Table 7 Predictor Weight Distribution

Dataset	Simple gradient Predictor	Adaptive North Predictor	Adaptive West Predictor	Adaptive Multicontext Predictor	Adaptive MED Predictor	Adaptive Median Predictor	Paeth Predictor	GAP Predictor
Kodak	1	1	1	8	15	0	11	3
DIV2k	5	12	12	0	4	8	0	5
CLIC P	5	7	11	4	14	5	3	4
CLIC M	6	14	14	9	12	7	6	8
Breast Ultrasound	6	1	14	0	8	1	1	1
Fetal Ultrasound	3	5	15	0	5	1	1	0
IAM	2	1	5	0	15	1	12	12

E.2 Predictor Weight Distribution Bar Chart

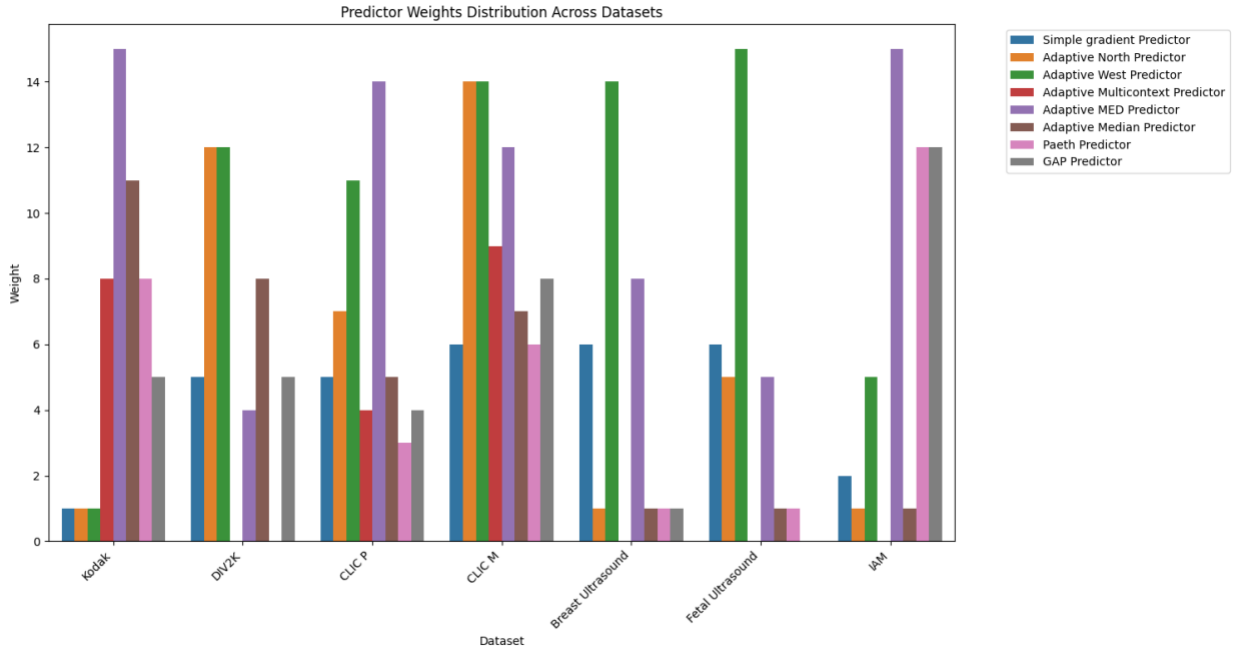

Figure 5 Predictor Weight Distribution Bar Chart

E.3 Predictor Weight Distribution Heat Map

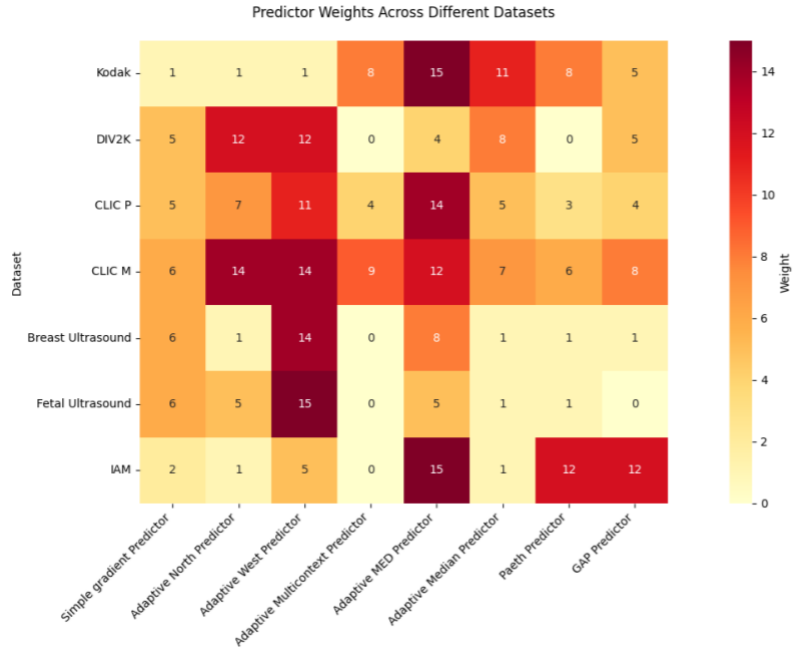

Figure 6 Predictor Weight Distribution Heat Map

Appendix F Image Characteristic Analysis

To better understand the relationship between image characteristics and compression performance, we analyzed key statistical properties across all datasets. This analysis supports the findings presented in Chapter 6 by quantifying how specific image attributes correlate with observed compression gains.

F.1 Image Characteristics and Correlation with Compression Improvements

Table 8 Image Characteristic and Correlation to Improvements

Category	Dataset	Entropy (bits/pixel)	Edge Density	Spatial Autocorrelation	W-OP8 Improvement (%)	Best Individual Predictor	Predictor Improvement (%)
Photography	Kodak	7.1	0.101	0.937	2.37	Adaptive LOCO-I	2.26
Photography	DIV2K	7.24	0.094	0.962	0.64	Adaptive LOCO-I	-0.82
Photography	CLIC P	7.17	0.077	0.965	0.79	Adaptive LOCO-I	-0.04
Photography	CLIC M	7.38	0.08	0.971	0.99	Adaptive LOCO-I	0.17
Medical	Breast Ultrasound	7.28	0.032	0.984	2.98	Adaptive West	2.36
Medical	Fetal Ultrasound	5.96	0.055	0.951	1.89	Adaptive West	0.44
Document	IAM	3.82	0.017	0.945	2.56	Adaptive LOCO-I	2.69

F.2 Correlation Matrix of Image Characteristics and Compression Improvement

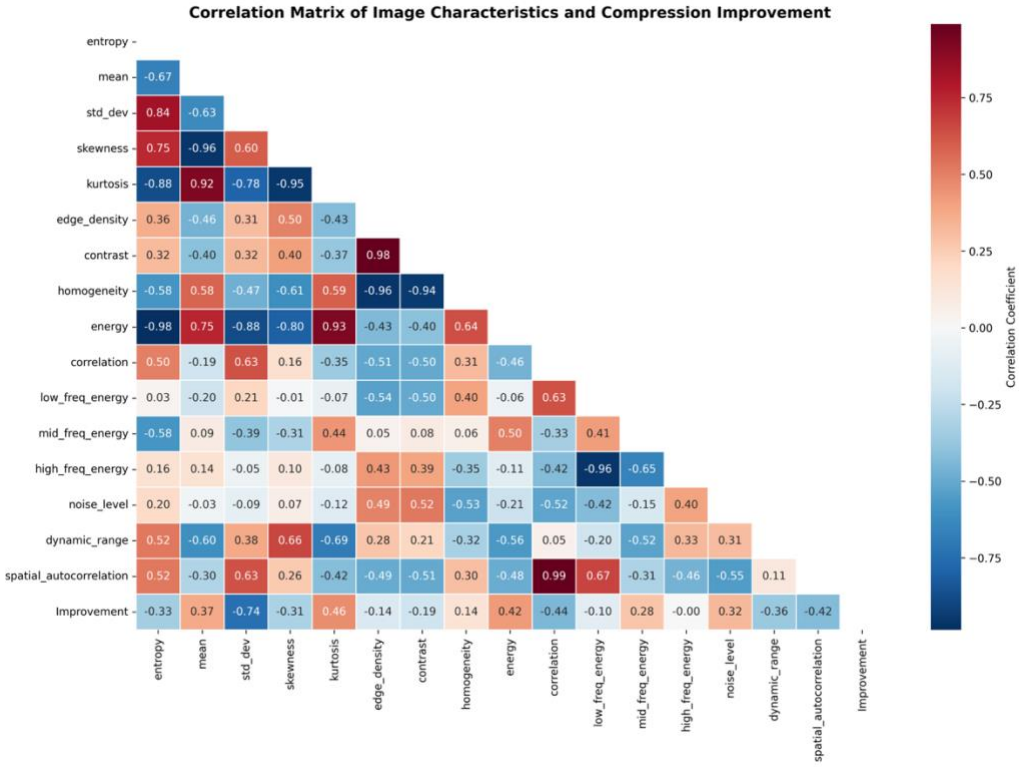

Figure 7 Correlation Matrix of Image Characteristics and Compression Improvement

Pearson correlation coefficients for statistical features across all datasets. Features include entropy, edge density, spatial autocorrelation, and their relationship with compression improvement.

F.3 Image Characteristics Summary

The following observations illustrate the statistical diversity across image categories and their relevance to compression behavior:

Content Entropy

- Low entropy (~3.82 bits/pixel) was observed in document-style images (IAM dataset).
- Moderate entropy (~5.96 bits/pixel) characterized the fetal ultrasound dataset.
- High entropy (~7.10–7.38 bits/pixel) was typical of photographic datasets such as Kodak, CLIC M, and DIV2K.

Edge Density

- Sparse edge density was observed in the IAM dataset (~0.017).
- Moderate edge density was found in ultrasound datasets (0.032–0.055).
- High edge density appeared in photographic datasets, especially Kodak (~0.101) and DIV2K (~0.094).

Spatial Autocorrelation

- Ultrasound images displayed the highest spatial autocorrelation (~0.984), indicative of smooth, coherent pixel regions.
- Lower autocorrelation values were observed in Kodak (~0.937) and IAM (~0.945), reflecting more texture variability and sharper transitions.