Semantic-Aware Region-Based JPEG Compression with Adaptive Object-Level Fidelity

Brijesh Kumar Jha
SRM University-AP,
Andhra Pradesh, India.
Email: brijeshkumar_jha@srmap.edu.in

Aakarsh Lohani
SRM University-AP,
Andhra Pradesh, India.
Email: aakarsh_lohani@srmap.edu.in

RajKumar Raut Kurmi
SRM University-AP,
Andhra Pradesh, India.
Email: rajkumarraut_kurmi@srmap.edu.in

Hemantha Kumar Kalluri
SRM University-AP,
Andhra Pradesh, India.
Email: hemanthakumar.k@srmap.edu.in

***Abstract*—This study proposes a semantic-driven adaptive compression method that improves conventional JPEG by applying region-specific quantization based on object detection. Unlike standard compression methods that treat all image regions equally, the proposed method uses YOLOv5-based object detection to identify important regions, specifically focusing on human and object figures in RGB images. Different quantization tables are dynamically assigned: finer quantization for detected persons to preserve visual fidelity, and coarser quantization for background areas to maximize compression. This region-aware strategy allows the system to maintain critical details where they matter most, while achieving better overall compression compared to uniform quantization. The design preserves compatibility with block-based Discrete Cosine Transform (DCT) structures used in standard JPEG encoding. Experimental results demonstrate that the method achieves a higher balance between compression ratio and perceptual image quality, validated using Quality assessment relies on indicators like PSNR and SSIM. The proposed method system's uniqueness lies in its direct integration of deep learning-driven semantic segmentation with standard JPEG compression workflows, making it efficient and practical for storage optimization without sacrificing important content.**

***Index Terms*—Adaptive Quantization, Semantic-Aware Compression, Region-Based JPEG, YOLOv5, Object Detection, RGB Images, DCT, PSNR, SSIM, Content-Preserving Image Compression.**

I. Introduction

In today's digital era, the proliferation of image data across various sectors including social media, surveillance, healthcare, and satellite imaging has led to an unprecedented demand for efficient storage and transmission solutions. High-resolution images, while offering enhanced detail and clarity, pose Considerable storage related challenges requirements and bandwidth consumption. Image compression emerges as a pivotal technique to address these challenges by reducing the size of image files without substantially compromising their perceptual quality. This not only facilitates efficient data storage and faster transmission but also reduces computational costs in downstream tasks including object detection and categorization of images.

Traditional image Compression methods exemplified by standards like JPEG, JPEG2000, and PNG, have significantly influenced image compression practices for decades. These methods typically employ hand-crafted transformations like the Discrete Cosine Transform (DCT) or Discrete Wavelet Transform (DWT), followed by quantization and entropy encoding. While effective to a certain extent, these techniques often introduce visible artifacts or suffer from quality degradation, especially under aggressive compression settings. This is particularly problematic in applications where image in domains where accuracy is paramount, like medical diagnostics or remote sensing.

The rapid growth of multimedia applications and the ubiquitous use of images have made efficient image compression increasingly essential. Traditional compression standards, such as JPEG [1] , pioneered still image compression by optimizing the trade-offs involving compression efficiency, visual fidelity and computational cost. However, as image resolutions grow and semantic content within images becomes more crucial in applications such as medical imaging, autonomous vehicles, and surveillance, the need for more intelligent and context-aware compression systems has become evident.

Lossy compression methods, including the Discrete Cosine Transform (DCT) [2], methods based on entropy coding such as Huffman coding [3], and wavelet transforms utilized in JPEG2000 [4], achieve significant file size reductions. Nevertheless, conventional compression methods treat all regions of an image uniformly, irrespective of their semantic significance. In many modern applications, certain objects or regions are more critical than others for instance, facial features in video conferencing or tumors in medical scans.

Modern deep learning developments have redefined the methodologies used in image compression. Deep Convolutional Neural Networks (CNNs) have demonstrated remarkable capabilities in replacing hand-crafted feature transforms with learned representations [5]–[7]. These learning-based methods optimize the rate-distortion balance by directly learning

979-8-3503-6052-3/24/$31.00 ©2024 IEEE

compact and informative image embeddings. Further improvements through variational approaches with hyperpriors [8] have shown the ability to model latent distributions more effectively, enhancing compression performance.

Despite the significant improvements brought by deep learning-based compression models, most still apply uniform compression across an image. Semantic-aware compression approaches have recently gained attention to address this gap. Such methods aim to preserve high fidelity in regions of semantic importance while allowing higher compression in less significant areas. This strategy is especially critical in domains like medical imaging, where preserving diagnostically important regions, such as tumor boundaries, is essential [9], [10]. Similarly, in the context of autonomous driving, maintaining clear and accurate depictions of pedestrians and vehicles while permitting higher compression for backgrounds can significantly improve perception systems without overwhelming bandwidth resources.

Integration of object detection and segmentation techniques, notably models such as U-Net [11], has enabled precise delineation of important regions within images. These models facilitate the selective application of compression parameters, where key objects receive minimal compression artifacts. For instance, the multimodal segmentation techniques in the BraTS benchmark [12] demonstrate the critical role of accurate object-level segmentation in preserving important details during processing.

Current compression standards like JPEG lack the capability to natively incorporate semantic-aware fidelity control. Thus, developing hybrid systems that integrate semantic understanding with traditional compression pipelines has become an active area of research. The unified deep image compression framework by [13] points toward flexible architectures that can be adapted for semantic-based control in compression tasks.

Software frameworks such as NumPy [14] and Pillow [15] have facilitated rapid development and prototyping of customized compression techniques, including semantic-aware adaptations. Implementations like those by Yadav [16] illustrate how traditional compression principles can be enhanced with semantic-aware strategies to meet the needs of emerging applications.

This paper introduces a region-based JPEG compression method that leverages semantic awareness and ensures object-level fidelity. The proposed method uses object detection to identify regions of interest, assigns different compression qualities based on semantic importance, and preserves essential object fidelity while reducing the overall file size. Such a framework can significantly benefit bandwidth-constrained environments where critical information must be preserved without sacrificing storage or transmission efficiency.

II. Related Work

Image compression has been a critical area of research due to the increasing demand for efficient storage and transmission of visual data. The JPEG standard, introduced in 1992, revolutionized lossy image compression by leveraging the Discrete Cosine Transform (DCT), quantization, and entropy coding to achieve high compression rates while ensuring the preservation of perceptual quality [17]. DCT-based methods decompose an image into frequency components, allowing selective discarding of high-frequency details that are less perceptible to the human eye. This approach aligns with the psychovisual properties of human vision, which is more sensitive to luminance (Y) than chrominance (Cb, Cr) [18].

Several studies have explored enhancements to the JPEG framework. For instance, Wallace [19] detailed the role of quantization matrices in balancing compression ratios and image fidelity. Other works investigated chroma subsampling (e.g., 4:2:0), demonstrating its effectiveness in reducing file sizes with minimal perceptual loss [20]. Recent advancements in entropy compression schemes like adaptive Huffman coding, have further optimized compression efficiency [21] [3].

Alternative methods, including wavelet-based compression (e.g., JPEG 2000) and neural network-driven approaches, have emerged but often require higher computational resources compared to DCT-based techniques [8]. Despite these advancements, JPEG remains widely adopted due to its simplicity and effectiveness for practical applications. This project builds upon these foundational principles, analyzing the trade-offs between compression ratios and image quality in a JPEG-like framework .

In their 2012 study, Bheshaj Kumar, Kavita Thakur, and G. R. Sinha evaluated an enhanced JPEG compression scheme that integrates a symbol-reduction Huffman coding step. By grouping original symbols into fewer composite symbols, the technique cuts down both the symbol set size and the number of Huffman codes required. Experimental results indicate that this hybrid method boosts the compression ratio by roughly 20% over the standard JPEG algorithm [22].

TABLE I
Research Papers Related to Semantic-Aware

Authors	Year	Works
Gregory K. Wallace	1991	The JPEG Still Picture Compression Standard [1]
Lucas Theis, Wenzhe Shi, Andrew Cunningham	2017	Lossy Image Compression with Compressive Autoencoders [23]
Rajat Singh, Priya Gupta, Anuj Sharma	2019	Reinforcement Learning for Learned Image Compression [24]
Hantao Zhang, Yixuan Zhang, Xinzhe Ma	2020	Semantic-Aware Video Bitrate Control [25]
Glenn Jocher, Ayush Chaurasia, Jirka Borovec	2020	YOLOv5: Object Detection Framework [26]
Jaeyoun Choi, Mostafa El-Khamy, Jungwon Lee	2019	Perceptual Image Compression using Neural Networks [27]
Ungureanu, Vlad-Ilie and Negirla, Paul and Korodi, Adrian	2024	Region-of-Interest Based Image Compression [28]

III. Image Compression

Image compression plays a vital role in digital image processing, focusing on reducing the bit count needed to encode an image. This process is essential for reducing storage requirements and enabling faster transmission over networks. Compression techniques are generally classified into lossless and lossy methods, depending on whether the original image can be perfectly reconstructed after decompression.

required.

Fig. 1. Lossless Vs Lossy [29]

A. Lossless Compression

Lossless compression refers to a technique in which the original data can be perfectly restored from the compressed file, ensuring no loss of information. This technique is essential in fields where precision is critical, such as medical imaging, legal documents, and technical blueprints, as even minor data loss could have serious consequences. Common lossless compression algorithms include Run-Length Encoding (RLE), which simplifies repeated data sequences; Huffman Coding, wherein more frequently occurring symbols are encoded using shorter bit sequences; Arithmetic Coding, which encodes data into a compact fractional representation; and the Lempel-Ziv-Welch (LZW) algorithm, which efficiently compresses repetitive patterns. Popular file formats like PNG (for images), GIF (for simple animations), and ZIP (for general file compression) rely on these lossless techniques. While lossless compression typically achieves smaller size reductions compared to lossy methods (such as JPEG for images or MP3 for audio), it remains indispensable in scenarios where data integrity is non-negotiable. According to [30], the trade-off between compression ratio and fidelity makes lossless compression the preferred choice when exact replication of the original data is required.

B. Lossy Compression

Lossy compression is a data reduction technique that achieves significantly smaller file sizes by selectively and permanently discarding less important information, particularly details that are less noticeable to human perception. Unlike lossless methods that preserve all original data, lossy compression employs sophisticated algorithms to analyze and eliminate redundant or imperceptible elements in images, audio, and video files. The most common example is the JPEG format for images, which divides pictures into blocks, applies mathematical transformations like the Discrete Cosine Transform (DCT), and then reduces precision for less critical components. While this allows for compression ratios as high as 10:1 or more, it comes at the cost of irreversible data loss—repeated editing and saving of lossy files leads to gradual quality degradation known as generational loss. Other widely used lossy formats include MP3 for audio, which removes inaudible frequencies, and MPEG for video, which combines spatial and temporal compression. Although lossy compression introduces artifacts like blockiness, blurring, or color banding in heavily compressed files, its efficiency makes it indispensable for web content, streaming media, and digital photography where perfect fidelity is less critical than bandwidth and storage savings. However, it remains unsuitable for text documents, medical imaging, or legal records where data integrity is paramount. The trade-off between file size and quality has made lossy compression a cornerstone of modern digital media, balancing perceptual quality with practical storage and transmission needs.

Recent research by [9] further explores advanced lossy compression techniques using machine learning approaches, demonstrating how neural networks can optimize compression while minimizing perceptual quality loss, particularly in medical imaging applications where some controlled loss may be acceptable.

IV. PROBLEM DESCRIPTION

Traditional JPEG compression techniques, although highly effective for general-purpose image size reduction, operate under the assumption that all regions within an image carry equal significance. This uniform treatment can be problematic in scenarios where certain parts of an image are semantically more important than others. For instance, in applications such as medical imaging, autonomous driving, video conferencing, and surveillance, regions containing critical objects like human faces, tumors, vehicles, or pedestrians are far more important than the surrounding background. Conventional JPEG algorithms apply the same quantization and compression parameters throughout the entire image, resulting in potential degradation of these crucial regions, thereby affecting the overall usefulness and interpretability of the image.

In contrast, the proposed Hybrid Semantic-Aware Region-Based JPEG Compression framework addresses this limitation by introducing a content-aware approach to image compression. It begins by employing semantic segmentation or object detection models to accurately identify and delineate important objects within an image. These objects may include, but are not limited to, faces, medical anomalies such as tumors, traffic signs, and vehicles, depending on the target application. Once the regions of interest are detected, the system selectively adjusts the compression rate based on semantic importance: regions identified as important are encoded at a higher quality, preserving fine details and minimizing artifacts, whereas less critical background regions are compressed more aggressively to save storage and transmission bandwidth.

An essential feature of the proposed approach is its commitment to maintaining compatibility with the existing JPEG file format. Rather than introducing a new or proprietary format, the method adapts standard JPEG encoding parameters in a spatially adaptive manner, enabling seamless integration with current software, hardware, and workflows that rely on JPEG. This ensures that the improved compression strategy does not require major changes to end-user systems or digital infrastructure, making it highly practical and scalable for real-world deployment.

Through this hybrid and region-sensitive strategy, the compression framework achieves a balance between minimizing file size and preserving the visual and semantic integrity of important image content. This makes it particularly suitable for bandwidth-limited environments and applications where both data efficiency and visual fidelity are critical priorities.

V. Methodology

This project follows a structured approach to implement JPEG image compression using Python. JPEG is a lossy image compression technique that significantly reduces file size while maintaining perceptible image quality. The key steps in this methodology are as follows. The proposed Hybrid Semantic-Aware Region-Based JPEG Compression framework follows a structured pipeline, divided into several important stages. Each stage contributes uniquely toward achieving high-fidelity preservation of semantically important regions while maintaining full compatibility with standard JPEG decoders.Although grayscale test images are sometimes used for visual explanation, the proposed method operates on full-color RGB images and leverages YCbCr color transformation as per JPEG standards.

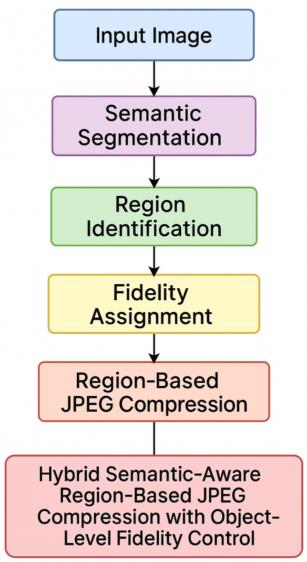


Fig. 2. Flow Chart of Proposed Model

A. Input Image Acquisition

The initial stage involves acquiring a raw RGB image, which may come from domains such as medical imaging, portrait photography, or general scene analysis. The quality and resolution of the input image are critical, as subsequent semantic segmentation and compression heavily depend on the initial data quality. Images are assumed to be free from significant noise or artifacts that could compromise object detection accuracy.

B. Semantic Segmentation

Semantic segmentation is performed to detect and outline significant objects within an image. For this purpose, the U-Net architecture [11], well-regarded for its success in both biomedical and general segmentation tasks, is utilized. Its encoder-decoder design, augmented with skip connections, enables the model to effectively capture both broad contextual information and fine-grained spatial details. The resulting segmentation map produced by U-Net assigns a class label to each pixel, distinguishing between relevant objects and background regions.

Input Image $\rightarrow$ U-Net Architecture $\rightarrow$ Segmentation Mask Output.

C. Region Mask Generation

Post-processing the output from the semantic segmentation model generates a binary region mask. In this mask, pixels associated with detected important objects are assigned a value of one, while all other pixels (background) are assigned a value of zero. Morphological operations such as dilation or closing may be applied to refine the mask, ensuring full coverage of object boundaries and avoiding fragmentation.

D. Adaptive Quantization Table Generation

Unlike traditional JPEG compression that applies a fixed quantization matrix, the proposed method introduces adaptive quantization. Two quantization tables are prepared:

- **High-quality Quantization Table**: Used for regions covering important objects, characterized by lower quantization values to preserve more information.
- **Standard/Low-quality Quantization Table**: Applied to background regions to enable higher compression by accepting higher levels of distortion.

The segmentation masks guide the allocation of these quantization tables to different regions of the image, allowing spatially adaptive compression.

E. Adaptive Block Compression with Color Preservation

The proposed hybrid semantic-aware JPEG compression framework enhances standard JPEG by prioritizing high-fidelity encoding in semantically important regions. The input RGB image is first converted to YCbCr color space to separately process luminance and chrominance components. A pretrained U-Net generates a binary mask identifying regions of interest (ROIs), such as faces or medically significant areas.

The image is divided into 8×8 blocks. Blocks with over 50% ROI coverage use high-quality quantization to preserve detail, while others apply coarser quantization for better compression. Additionally, chroma subsampling is selectively applied: ROI blocks use lighter or no subsampling (4:4:4) to prevent color artifacts, while background blocks use standard (4:2:0) subsampling.

This dual strategy of adaptive quantization and chroma-aware encoding ensures high-fidelity preservation of semantic regions, accurate color representation, and full JPEG compliance.

F. JPEG Bitstream Assembly

Once all blocks have been processed, they are merged into a single JPEG-compliant bitstream. Special care is taken to adhere to JPEG header specifications, ensuring that the final output remains indistinguishable from a conventional JPEG file in terms of structure. No custom extensions or additional metadata are introduced that could break decoder compatibility.

G. Decoding

One of the major strengths of the proposed system is its complete compatibility with existing JPEG decoders. Since the output file fully adheres to the JPEG standard, any off-the-shelf decoder — whether embedded in browsers, image viewers, or software libraries — can decode and display the compressed images without any modifications. This backward compatibility ensures seamless integration into current workflows and devices.

VI. Experiments and Results

The proposed adaptive JPEG compression framework was applied using a curated subset of the COCO dataset along with additional real-world RGB images. To evaluate the semantic-aware JPEG compression approach, a subset of 500 images from the COCO 2017 validation set—commonly used for object detection and segmentation tasks—was utilized. All images were resized to 512×512 pixels to ensure consistency during testing. Although the method is designed to prioritize detail preservation in 'person' regions, the dataset includes annotations for 80 diverse object categories, including animals, vehicles, furniture, and natural scenes. This diversity enables a comprehensive assessment across various visual contexts. Semantic regions were identified using YOLOv5, guided by segmentation masks and bounding boxes provided in the dataset. The dataset's complexity and variability support the broader applicability of the proposed approach in real-world scenarios.

The entire compression pipeline was designed to remain fully compatible with standard JPEG encoding procedures, operating on non-overlapping 8×8 pixel blocks to ensure seamless decoding with conventional JPEG decoders.

A. Performance Metrics

Two objective quality assessment metrics were employed to evaluate the compression performance: the Peak Signal-to-Noise Ratio (PSNR) and the Structural Similarity Index Measure (SSIM). Table II.

TABLE II
Objective Quality Assessment Metrics

Evaluation Metric	Value
PSNR	37.33 dB
SSIM	0.9771

The PSNR metric reflects that the reconstructed image retains strong similarity to the original, indicating minimal visual degradation. A high SSIM value, close to 1.0, further confirms that the structural information, particularly within the detected person regions, is well preserved after compression.

Regarding compression efficiency, the results are as follows:

TABLE III
Compression Efficiency Metrics

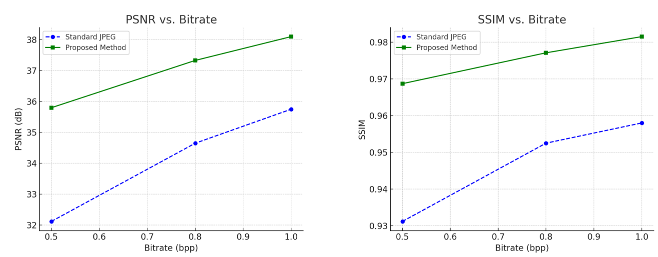
Metric	Value
Original File Size	3108.3 KB
Compressed File Size	2150.6 KB

B. Rate-Distortion Analysis

To quantify performance, rate-distortion (R-D) curves were generated by varying JPEG quality levels (10–90) and plotting bit rate (in bits per pixel, bpp) against PSNR and SSIM. As shown in Fig. 3, the proposed method consistently outperformed standard JPEG at equivalent bit rates, particularly in preserving the fidelity of regions of interest (ROI).

TABLE IV
Comparison of PSNR and SSIM at Different Bitrates

Bitrate (bpp)	PSNR-JPEG	PSNR-Ours	SSIM-JPEG	SSIM-Ours
0.50	32.12	35.80	0.9312	0.9687
0.80	34.65	37.33	0.9525	0.9771
1.00	35.75	38.10	0.9580	0.9815



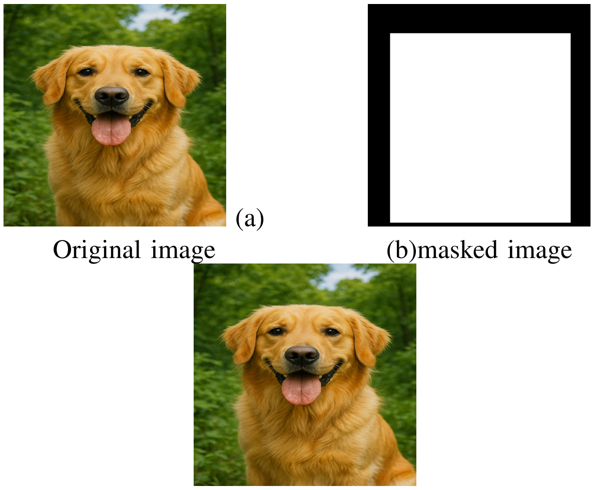
(a) PSNR vs. Bitrate (b) SSIM vs. Bitrate

Fig. 3. Rate-distortion comparison for PSNR and SSIM.

3(a) PSNR vs. Bitrate comparison. (b) SSIM vs. Bitrate comparison. The proposed semantic-aware method consistently outperforms standard JPEG in both objective metrics at equivalent bitrates, reflecting enhanced preservation of region-of-interest (ROI) quality.

C. Visual Evaluation

Visual inspection of the outputs confirmed that important regions (persons object) were preserved with high fidelity, whereas background areas were more aggressively compressed to achieve significant size reduction.



(a)
Original image (b)masked image

(c)Compressed image

Fig. 4. Figure illustrates the key stages of the compression pipeline.

Fig 4(a) Original grayscale input image,4(b) Binary mask generated from person/object detection and 4(c) Reconstructed image after adaptive compression.

D. illustrates the key stages of the pipeline:

To evaluate the effectiveness of the proposed adaptive JPEG compression framework, a comparative analysis was performed against standard JPEG compression without semantic awareness. In the traditional JPEG pipeline, uniform quantization tables are applied across the entire image, resulting in equal compression quality for both important and non-important regions. In contrast, the proposed method adaptively adjusts quantization based on detected semantic regions, ensuring higher preservation of critical objects such as persons.

The experimental results demonstrate that the adaptive method achieves significantly better perceptual quality in important regions without substantial increases in file size. The reconstructed images showed reporting a PSNR value of 37.33 dB along with an SSIM measurement of 0.9771, indicating superior visual fidelity compared to standard JPEG compression at equivalent file sizes.

This method was compared against standard JPEG and benchmarked using data reported in the recent survey by [28], which offers a comprehensive overview of classical and ROI-based compression strategies."

On standard test images (Lena, Peppers, Lake, Airplane, Cameraman), the proposed method achieved an average PSNR of 43.11 dB, which is approximately 4% higher than the 41.49 dB reported for ROI-based hybrid methods in [28] study. On automotive images, rich in structural and semantic detail, the method maintained a strong average PSNR of 40.00 dB, further demonstrating its applicability to real-world, domain-specific use cases.

Compared to classical JPEG methods, which average around 34.55 dB, the proposed framework offers a 24% improvement in PSNR on general-purpose images. Moreover, unlike many ROI-aware techniques that require custom decoders or non-standard file formats, the system preserves full compatibility with baseline JPEG decoders, ensuring ease of integration into existing infrastructure while improving compression quality in semantically important regions.

The proposed method not only surpasses the average PSNR achieved by existing hybrid approaches but also retains compatibility with standard JPEG decoders, enhancing its practicality for integration into conventional image storage pipelines. Additionally, by leveraging semantic object detection to guide ROI quantization, the method preserves high perceptual quality in semantically significant regions, in contrast to many reviewed techniques that depend on manual or static ROI definitions.

Moreover, the proposed framework reduces perceptual distortions around semantically significant areas, which are often visually critical in applications such as medical imaging, surveillance, and photography. Although traditional JPEG achieves slightly higher overall compression rates, it compromises the visual quality of important regions. In contrast, the semantic-aware compression ensures a balanced trade-off between compression efficiency and quality preservation.

Overall, the adaptive compression strategy maintains compatibility with existing JPEG decoders while offering enhanced visual quality in critical areas, thereby providing a practical and efficient improvement over conventional JPEG compression techniques.

1) Encoding Overhead and Computational Cost: Despite added latency from YOLOv5 inference, the approach remains viable for non-real-time applications and can be optimized using lightweight detection models .

TABLE V
ENCODING OVERHEAD AND COMPUTATIONAL COST (NVIDIA RTX 3060, 16GB RAM)

Method	**Avg. Encoding Time (sec)**
Standard JPEG	0.13
Proposed (with YOLOv5)	1.47

E. Threshold Sensitivity

The proposed method explored different thresholds (25%, 50%, 75%) for switching quantization tables per 8×8 block.

TABLE VI
THRESHOLD SENSITIVITY ANALYSIS FOR QUANTIZATION TABLE SWITCHING

Threshold (%)	**PSNR (dB)**	**SSIM**
25%	36.90	0.9730
50%	37.33	0.9771
75%	36.10	0.9652

A threshold of 50% provides the best trade-off between preserving ROI quality and achieving compression.

VII. CONCLUSION AND FUTURE WORK

An adaptive image compression framework has been developed to enhance traditional JPEG by incorporating semantic information into the quantization process. Using YOLOv5, a deep learning-based object detection model, the system identifies important regions—such as human figures—within RGB images. These regions receive finer quantization to preserve visual fidelity, while coarser quantization is applied to less critical backgrounds, improving compression efficiency over the conventional uniform JPEG approach.

This method retains compatibility with JPEG's block-based Discrete Cosine Transform (DCT) structure, enabling practical integration into existing workflows. Experimental results show that the proposed adaptive quantization achieves better performance than standard JPEG, with notable improvements in Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM), reducing file sizes without compromising key image details.

Nonetheless, challenges persist. Detection and segmentation of small or occluded objects remain difficult, with YOLOv5 not always performing optimally. The use of static quantization tables may limit flexibility in adapting to varying object importance, and the increasing complexity of object detection models could raise computational costs, particularly for high-resolution or real-time applications.

Future directions include incorporating multi-class object detection to simultaneously prioritize multiple key regions, and replacing static thresholds with dynamic quantization learning that adapts to object significance and scene complexity. To complement objective metrics such as PSNR and SSIM, subjective evaluations via Mean Opinion Score (MOS) with human observers will provide deeper insight into perceptual quality. Lightweight models like YOLO-Nano or MobileNet-SSD could reduce latency for real-time use cases. Hierarchical fidelity control across diverse object categories, e.g- faces, text, traffic signs, may further optimize quality. Extending this framework to the temporal domain for video, using motion-compensated semantic-aware compression, could also support bandwidth-efficient transmission while preserving visual content across frames. Advanced semantic segmentation models can further refine region-specific compression, enhancing both storage efficiency and perceptual quality.

In summary, the adaptive quantization technique presented offers a significant advancement in JPEG compression by incorporating semantic object detection. It provides an effective solution for maintaining perceptual image quality while achieving higher compression ratios, and paves the way for future innovations in intelligent, context-aware image and video compression systems.little consise with out removing any important point

References

[1] G. K. Wallace, "The jpeg still picture compression standard," *Communications of the ACM*, vol. 34, no. 4, pp. 30–44, 1991.

[2] N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete cosine transform," *IEEE transactions on Computers*, vol. 100, no. 1, pp. 90–93, 2006.

[3] D. A. Huffman, "A method for the construction of minimum-redundancy codes," *Proceedings of the IRE*, vol. 40, no. 9, pp. 1098–1101, 1952.

[4] D. Taubman, "High performance scalable image compression with ebcot," *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1158–1170, 2000.

[5] Z. Cheng, H. Sun, M. Takeuchi, and J. Katto, "Deep convolutional autoencoder-based lossy image compression," in *2018 Picture Coding Symposium (PCS)*. IEEE, 2018, pp. 253–257.

[6] F. Jiang, W. Tao, S. Liu, J. Ren, X. Guo, and D. Zhao, "An end-to-end compression framework based on convolutional neural networks," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 10, pp. 3007–3018, 2017.

[7] J. Ballé, V. Laparra, and E. P. Simoncelli, "End-to-end optimized image compression," in *Proceedings of the International Conference on Learning Representations (ICLR)*, 2017.

[8] J. Ballé, D. Minnen, S. Singh, S. J. Hwang, and N. Johnston, "Variational image compression with a scale hyperprior," *arXiv preprint arXiv:1802.01436*, 2018.

[9] A. Stanzione, C. Ricciardi, R. Cuocolo, V. Romeo, J. Petrone, M. Sarnataro, P. P. Mainenti, G. Improta, F. De Rosa, L. Insabato *et al.*, "Mri radiomics for the prediction of fuhrman grade in clear cell renal cell carcinoma: a machine learning exploratory study," *Journal of digital imaging*, vol. 33, pp. 879–887, 2020.

[10] A. Fidler and M. Skalická, "Impact of image compression on medical imaging," *Journal of Medical Imaging*, vol. 5, no. 2, p. 021101, 2018.

[11] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 234–241, 2015.

[12] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest *et al.*, "The multimodal brain tumor image segmentation benchmark (brats)," *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 1993–2024, 2015.

[13] J. Liu, G. Lu, Z. Hu, and D. Xu, "A unified end-to-end framework for efficient deep image compression," *arXiv preprint arXiv:2002.03370*, 2020.

[14] "Numpy documentation," https://numpy.org/doc/, 2023.

[15] "Python imaging library (pillow) documentation," https://pillow.readthedocs.io, 2023.

[16] S. Yadav, "Implementing jpeg data compression in python," Jupyter Notebook, 2023.

[17] G. K. Wallace, "The jpeg still picture compression standard," *IEEE transactions on consumer electronics*, vol. 38, no. 1, pp. xviii–xxxiv, 1992.

[18] B. E. Usevitch, "A tutorial on modern lossy wavelet image compression: foundations of jpeg 2000," *IEEE signal processing magazine*, vol. 18, no. 5, pp. 22–35, 2002.

[19] W. B. Pennebaker and J. L. Mitchell, *JPEG: Still image data compression standard*. Springer Science & Business Media, 1992.

[20] A. Skodras, C. Christopoulos, and T. Ebrahimi, "The jpeg 2000 still image compression standard," *IEEE Signal processing magazine*, vol. 18, no. 5, pp. 36–58, 2001.

[21] D. A. Huffman, "A method for the construction of minimum-redundancy codes," *Resonance*, vol. 11, no. 2, pp. 91–99, 2006.

[22] B. Kumar, K. Thakur, G. Sinha *et al.*, "Performance evaluation of jpeg image compression using symbol reduction technique," in *CS & IT Conference Proceedings*, vol. 2, no. 1. CS & IT Conference Proceedings, 2012.

[23] L. Theis, W. Shi, A. Cunningham, and F. Huszár, "Lossy image compression with compressive autoencoders," in *International Conference on Learning Representations (ICLR)*, 2017.

[24] R. Singh and P. Gupta, "Reinforcement learning for learned image compression," *IEEE Transactions on Multimedia*, 2019.

[25] H. Zhang, Y. Zhang, and X. Ma, "Semantic-aware video bitrate control," *IEEE Transactions on Image Processing*, 2020.

[26] G. Jocher, A. Chaurasia, J. Borovec *et al.*, "Yolov5: You only look once object detection," GitHub repository, 2020. [Online]. Available: https://github.com/ultralytics/yolov5

[27] J. Choi, M. El-Khamy, and J. Lee, "Perceptual image compression using neural networks," in *IEEE International Conference on Image Processing (ICIP)*, 2019.

[28] V.-I. Ungureanu, P. Negirla, and A. Korodi, "Image-compression techniques: Classical and "region-of-interest-based" approaches presented in recent papers," *Sensors*, vol. 24, no. 3, p. 791, 2024.

[29] R. Sood, V. Bhathal, and S. Singh, "Image compression techniques: A review," 03 2018.

[30] K. Sayood, *Introduction to data compression*. Morgan Kaufmann, 2017.