SEMANTICALLY-GUIDED IMAGE COMPRESSION FOR ENHANCED PERCEPTUAL QUALITY AT EXTREMELY LOW BITRATES

A Seminar Report

Submitted to the APJ Abdul Kalam Technological University in partial fulfilment of the requirements for the award of degree

Bachelor of Technology

in

Computer Science and Engineering

by

ARUN GEORGE

PTA21CS020

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

COLLEGE OF ENGINEERING KALLOOPPARA

PATHANAMTHITTA

October 2024

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

COLLEGE OF ENGINEERING KALLOOPPARA

PATHANAMTHITTA - 689603

2024 - 2025

CERTIFICATE

This is to certify that the report entitled **Semantically-Guided Image Compression for Enhanced Perceptual Quality at Extremely Low Bitrates** submitted by **Arun George (PTA21CS020)** to the APJ Abdul Kalam Technological University in partial fulfillment of the B.Tech degree in Computer Science & Engineering is a bonafide record of the seminar work carried out by him under our supervision. This report has not been submitted to any other University or Institution for any purpose.

Internal Guide	Coordinator	Head of the Department
Mrs. Shilpa Radhakrishnan	**Mrs. Anitha Jose**	**Dr. Renu George**
Assistant Professor	Assistant Professor	Assistant Professor
Department of	Department of	Department of
Computer Science	Computer Science	Computer Science
& Engineering	& Engineering	& Engineering

ACKNOWLEDGMENT

I take this opportunity to express my deepest sense of gratitude and sincere thanks to everyone who helped me to complete this work successfully. I express my sincere thanks to the Principal **Prof. Dr. Deepa J** for the constant support and help. I also thank **Dr. Renu George**, Head of the Department, Computer Science and Engineering, College of Engineering Kallooppara for providing me with all the necessary facilities and support. I would like to place on record my sincere gratitude to my seminar guide **Mrs. Shilpa Radhakrishnan**, Assistant Professor,Computer Science and Engineering, College of Engineering Kallooppara for the guidance and mentorship throughout the course. I would like to express my sincere gratitude to **Mrs. Anitha Jose**, and **Ms.Parvathy M N**, Assistant Professors, Department of Computer Science and Engineering, College of Engineering Kallooppara for their support and cooperation.Finally, I thank my family, and friends who contributed to the successful fulfillment of this seminar work.

ARUN GEORGE

ABSTRACT

Efficient image compression is crucial for reducing the storage and transmission costs of images, especially in low-bandwidth environments. Traditional image compression methods,such as JPEG and BPG, and even newer machine learning-based methods, face considerable challenges when operating at bitrates below 0.1 bits per pixel (bpp). At such low bitrates,these methods typically result in images suffering from perceptual degradation, such as blurring and loss of critical visual details.

Shoma Iwai[1] introduces a novel semantically-guided image compression method that integrates semantic information to enhance perceptual quality at extremely low bitrates.By using semantic label maps, the decoder is guided to reconstruct textures that are contextually accurate, ensuring higher perceptual quality than conventional methods. Additionally, the data size of the semantic label maps is reduced through several strategies such as downscaling,reducing the number of semantic classes, and using an auto-regressive compression model.

Experimental evaluations on benchmark datasets, such as COCO and Cityscapes show that this method surpasses state-of-the-art compression techniques in terms of perceptual quality and user satisfaction, making it ideal for applications in bandwidth-constrained environments.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AES	: Advanced Encryption Standard
bpp	: bits per pixel
GAN	: Generative Adversarial Network
LPIPS	: Learned Perceptual Image Patch Similarity
MSE	: Mean Squared Error
PSNR	: Peak Signal-to-Noise Ratio
SSIM	: Structural Similarity Index
MS-SSIM	: Multi-Scale Structural Similarity Index
FID	: Frechet Inception Distance
DCT	: Discrete Cosine Transform
CABAC	: Context-Adaptive Binary Arithmetic Coding
CNN	: Convolutional Neural Network
NLP	: Natural Language Processing
RGB	: Red, Green, Blue (color model)
GMM	: Gaussian Mixture Model
MMLU	: Massive Multitask Language Understanding
JPEG	: Joint Photographic Experts Group
BJP	: Bitmap JPEG

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

The proliferation of digital images in today's world necessitates effective image compression techniques to manage storage space and bandwidth. As the demand for high quality images increases across various applications ranging from social media to professional photography the challenge lies in compressing these images without sacrificing their visual integrity.Traditional image compression techniques, such as JPEG and BPG, are effective to a certain extent. However, they rely heavily on handcrafted algorithms that often fail to preserve image quality at extremely low bitrates. This degradation is characterized by blurring,loss of detail, and the introduction of compression artifacts that detract from the overall user experience

Recent advancements in deep learning have opened new avenues for improving image compression methods. Neural network-based approaches have demonstrated superior performance by learning to represent image data more effectively than traditional algorithms. However, even the latest machine learning techniques encounter limitations when it comes to maintaining perceptual quality, particularly under stringent bitrate constraints.

1.2 PURPOSE

The primary purpose of the study by Iwai [1] is to explore and present a novel approach to image compression, referred to as semantically-guided image compression. This technique aims to significantly enhance the perceptual quality of reconstructed images, particularly at extremely low bitrates (below 0.1 bits per pixel). As digital imaging becomes increasingly prevalent across various platforms such as social media, video streaming, and mobile applications, the need for efficient image compression techniques that preserve visual quality has become critical.

Iwai's [1]work seeks to address the limitations of traditional image compression methods, which often result in blurred and artifact-ridden images in low-bitrate scenarios. By leveraging semantic information through the use of semantic label maps, this approach provides guidance to the decoder during the reconstruction process. Consequently, the reconstructed images are not only sharper and more detailed but also contextually appropriate, effectively reflecting the true content of the original images.

Furthermore, Iwai[1] will evaluate the effectiveness of the proposed method through extensive experimental studies, comparing it against traditional compression techniques and state-of-the-art deep learning methods. The goal is to demonstrate that the integration of semantic information can lead to substantial improvements in image quality and user satisfaction,thereby paving the way for more robust and efficient image compression solutions suitable for modern applications

1.3 SCOPE

The scope of the Iwai's[1] research encompasses a comprehensive investigation of various aspects of image compression, with a particular focus on the challenges faced at low bitrates. It includes:

- **Review of Traditional and Modern Compression Techniques:**Iawi research[1] will provide a thorough overview of existing image compression methods, detailing their algorithms and underlying principles. It will highlight the advantages and disadvantages of traditional techniques such as JPEG and BPG, as well as modern approaches utilizing deep learning and neural networks.
- **Introduction to Semantically-Guided Compression:** Iwai[1] will delve into the concept of semantically-guided image compression, explaining how semantic label maps can be used to inform the compression and reconstruction process. The integration of high-level semantic information will be discussed, focusing on its role in enhancing the perceptual quality of reconstructed images.
- **Optimization Strategies:**[1] will explore the various optimization techniques employed to ensure the method operates efficiently within the constraints of extremely low bitrates. This includes strategies such as rate-distortion optimization, hyperprior models, and context modeling, which are critical for improving compression performance
- **Experimental Evaluation and Performance Metrics:**A significant portion of the paper by Iwai [1] will focus on presenting the results of experimental evaluations conducted

on benchmark datasets such as COCO and Cityscapes. These evaluations aim to demonstrate the capabilities of the proposed image compression method.

The study will utilize various performance metrics, including LPIPS, PSNR, MS-SSIM, and FID, to comprehensively assess the effectiveness of the approach. Comparisons will be made with both traditional and state-of-the-art techniques to highlight its advantages.

- **Practical Applications and Future Prospects:**[1] will also discuss potential real-world applications for the proposed semantically-guided image compression method in fields such as mobile communication, cloud storage, and real-time streaming, where bandwidth is a critical constraint. Additionally, the seminar will address future research directions and the potential impact of this technology on the field of image compression.

- **Security Considerations:** The study by Iwai [1] will also address the security implications of image compression. It will analyze potential vulnerabilities, such as risks to data privacy and the susceptibility of machine learning models to adversarial attacks.

  Furthermore, the paper will explore methods to enhance the security of machine learning models during deployment, ensuring that the proposed image compression approach remains safe and resilient in practical applications.

CHAPTER 2

LITERATURE SURVEY

2.1 OVERVIEW

The field of image compression has evolved significantly over the past few decades, with traditional techniques giving way to more advanced methods that leverage deep learning and semantic understanding. This literature survey examines critical contributions to the domain,highlighting their methodologies, strengths, and implications for future research.

2.2 TRADITIONAL IMAGE COMPRESSION TECHNIQUES

The JPEG standard, established by Wallace (1992) [2], is one of the earliest and most widelyused image compression techniques. It employs a lossy compression method based on the Discrete Cosine Transform (DCT) to reduce image size while retaining reasonable visual quality.However, traditional methods like JPEG often struggle to maintain quality at very low bitrates, resulting in visible artifacts such as blocking and blurring.

2.3 NEURAL-NETWORK BASED APPROACHES

Recent advancements have seen the emergence of neural network-based image compression methods, which offer improved performance over traditional techniques. Toderici et al (2017) [3] introduced a recurrent neural network (RNN)-based approach capable of achieving better compression rates while preserving detail at higher resolutions . Their method demonstrated that deep learning architectures could outperform JPEG in terms of rate-distortion performance,especially in high-complexity images.

2.4 GENERATIVE ADVERSARIAL NETWORKS(GANs)

The integration of Generative Adversarial Networks (GANs) into image compression represents a significant leap forward. Agustsson et al (2019) [4] applied GANs to enhance the perceptual quality of compressed images, emphasizing the use of adversarial loss functions to minimize artifacts and improve realism in the reconstructed images . Their approach not only improved visual fidelity but also set the stage for further exploration into the potential of GANsin image processing.

Mentzer et al. (2020) [5] built upon these ideas with the introduction of High Fidelity Generative Image Compression (HiFiC), which combines GANs with a perceptual loss function to create images that closely resemble their original counterparts. This method effectively addresses the trade-offs between compression efficiency and visual quality, achieving remarkable results at low bitrates.

2.5 SEMANTIC IMAGE COMPRESSION

The incorporation of semantic information into image compression strategies has gained traction as a means of enhancing perceptual quality. Zhang et al. (2018) [6] proposed a semantically guided image compression method that utilizes semantic segmentation maps to direct the reconstruction process. By prioritizing regions of semantic importance, their approach significantly improves image quality at low bitrates, reducing artifacts that typically arise from standard compression techniques.

Following this trend, Mentzer et al. (2020) [7] further explored the integration of semantic segmentation with GANs in their work on Enhanced Low-Bit-Rate Generative Image Compression (EGIC). They demonstrated that combining semantic guidance with adversarial training leads to superior performance, effectively preserving the integrity of semantically significant regions.

2.6 RECENT DEVELOPMENTS IN SEMANTIC CODING

Huang et al. (2021) [8] introduced deep learning-based image semantic coding techniques,focusing on the dual benefits of semantic understanding and effective compression. Their work emphasizes the role of semantic information in improving communication efficiency in various applications, showcasing how deep learning can enhance traditional coding techniques.

Liu et al. (2021) [9] explored octave convolution and semantic segmentation in their approach to image compression, demonstrating that the combination of these techniques results in improved performance . Their research highlights the importance of adopting innovative methods to further advance the field of image compression.

Technique	Advantages	Disadvantages
JPEG	Widely used, good for standard images	Lossy compression, artifacts at low bitrates
GAN-based	High perceptual quality, adaptable	Requires extensive training, computationally intensive
Semantically guided	Preserves important features, reduces artifacts	Complexity in integration, needs semantic data

Table 1: Comparison of Image Compression Techniques

The evolution of image compression techniques reflects a shift towards more intelligent methods that integrate semantic understanding and advanced neural architectures. The body of work reviewed in this literature survey illustrates the potential of semantically guided approaches to enhance visual quality and maintain efficiency at low bitrates. Future research should continue to explore the integration of semantic information with cutting-edge machine learning techniques, aiming to further refine the balance between compression efficiency and perceptual fidelity.

CHAPTER 3

METHODOLOGY

3.1 MACHINE LEARNING AND NEURAL NETWORK-BASED COMPRESSION

The integration of machine learning, particularly deep learning, has significantly transformed the landscape of image compression. Traditional compression techniques rely on fixed algorithms that follow predefined rules and do not adapt to the specific content of each image. In contrast, neural network-based methods harness the power of learning to more efficiently encode and decode images, allowing for a more flexible and adaptive approach to compression. These neural network models are capable of identifying and leveraging complex patterns within the data, resulting in improved performance in both compression ratio and perceptual quality.

Typically, the architecture of these methods consists of an encoder that compresses the image into a latent representation and a decoder that reconstructs the image from this compressed form. The encoder and decoder work in tandem, learning from large datasets to refine the encoding and decoding processes, ensuring that the reconstructed image retains as much perceptual fidelity as possible while minimizing the data size.

In the context of semantically guided compression, these neural networks are further enhanced to incorporate high-level semantic understanding of the image content. By integrating semantic guidance, the model can intelligently allocate bits to preserve the most important image features. This allows for more effective compression by ensuring that crucial areas, such as textures or key objects, are preserved even under low bitrate constraints. The addition of semantic information improves the reconstruction process, making the output not only more accurate but also more visually coherent, with important elements being prioritized during encoding and decoding.

3.2 GENERATIVE ADVERSARIAL NETWORKS (GANSs)

Generative Adversarial Networks (GANs) play a crucial role in improving the perceptual quality of compressed images. A GAN consists of two main components: the generator, which creates new images, and the discriminator, which evaluates the authenticity of these generated images by comparing them to real samples. Through adversarial training, where the generator and discriminator work in opposition, the generator progressively learns to create images that appear increasingly realistic and convincing.

In the method proposed by Iwai [1], GANs are utilized to address the blurring commonly observed in low-bitrate compression scenarios. By refining the reconstructed images through continuous feedback from the discriminator, GANs are able to enhance the overall image quality, ensuring that the final output maintains a higher level of visual clarity.

This approach ensures that the reconstructed images are not only sharper but also semantically coherent, with key features preserved and important details maintained. The result is an image that is not only visually appealing but also contextually accurate, preserving the integrity of the original content despite the compression.

3.3 SEMANTIC LABEL MAPS

The integration of semantic label maps is a key feature of the proposed compression method. These maps provide critical high-level information about different regions in an image, allowing the decoder to apply contextually appropriate textures during reconstruction. For example, the decoder can distinguish between the smoothness of a sky and the intricate details of foliage, ensuring that each area is restored accurately.

The challenge of overhead data introduced by semantic maps is addressed through several strategies:

- **Downscaling:** Reducing the resolution of the label maps to lower their size without significant loss of information
- **ClassReduction:** Combining similar semantic classes(e.g., merging ”grass” and ”trees”) to simplify the label map and reduce its data size.
- **Autoregressive Compression:** Utilizing spatial correlations to predict the probability of each pixel’s class based on neighboring pixels, which enhances compression efficiency.

3.4 OPTIMIZATION TECHNIQUES

To ensure the method operates effectively within low-bitrate constraints, several optimization techniques are applied:

- **Rate-Distortion Optimization:** This approach minimizes the perceptual loss for a given bitrate, balancing the need for compression with the desire for visual fidelity
- **Hyperprior Models:** These models capture the uncertainty in the latent representations, allowing for more efficient bit allocation where it is most needed
- **Context Modeling:** This technique improves the accuracy of probability predictions for compressed data, further enhancing the efficiency of the system.

3.5 PERFORMANCE MATRICS

To evaluate the effectiveness of the proposed semantically-guided image compression method, various performance metrics are employed. These metrics are critical for quantifying the perceptual quality and compression efficiency of the reconstructed images, allowing for a comprehensive comparison with traditional and state-of-the-art methods. The following metrics are utilized:

- **LPIPS (LEARNED PERCEPTUAL IMAGE PATCH SIMILARITY):** LPIPS is a perceptual similarity metric that quantifies the difference between two images based on the output of deep neural networks. Unlike traditional pixel-based metrics, LPIPS assesses the perceptual similarity by comparing local patches of images, which are processed through feature extractors like deep convolutional networks. This metric is particularly sensitive to human visual perception and reflects how similar two images appear to a human observer. Lower LPIPS values indicate that the reconstructed image closely resembles the original image in terms of perceptual quality. In the context of semantically-guided compression,LPIPS helps assess how well the method preserves important visual details and textures, ensuring that the reconstructed images remain visually appealing.
- **PSNR(PEAK SIGNAL-TO-NOISE RATIO):** PSNR is a widely used objective metric that measures the ratio between the maximum possible power of a signal (in this case, the original image) and the power of corrupting noise (i.e., the difference between the original and reconstructed images).PSNR is expressed in decibels (dB), with higher

values indicating better reconstruction quality. Although PSNR is commonly used in image processing to evaluate compression methods, it has limitations. It primarily focuses on pixel wise fidelity and does not necessarily correlate with perceived visual quality. Therefore,while PSNR provides a quick measure of reconstruction accuracy, it should be interpreted alongside perceptual metrics like LPIPS

- **MS-SSIM(MULTI-SCALE STRUCTURAL SIMILARITY INDEX):** MS-SSIM is an extension of the SSIM (Structural Similarity Index) metric, designed to assess the structural similarity between two images across multiple scales. MS-SSIM evaluates image quality by comparing luminance, contrast, and structure at different resolutions, making it more robust in capturing perceptual differences than traditional metrics. Higher MS-SSIM scores indicate that the reconstructed image maintains structural integrity and visual consistency with the original image. This metric is especially valuable in the context of semantically guided compression, where retaining the structural details of the image is crucial for achieving high perceptual quality.

- **FID (FRECHET INCEPTION DISTANCE):** UFID measures the quality of generated images by comparing the distribution of features extracted from a pretrained neural network (often an Inception network) for both the original and reconstructed images. It calculates the Frechet distance between the two distributions, providing a quantitative measure of how similar the generated images are to the original images in terms of quality and realism. Lower FID values indicate that the reconstructed images are more similar to the original images, reflecting the effectiveness of the compression method in producing realistic outputs. In the context of this seminar, FID serves as an important metric for validating the perceptual quality of the semantically-guided compression method.

By employing these performance metrics LPIPS,PSNR,MS-SSIM,and FID the seminar provides a thorough evaluation of the proposed image compression method, offering in sights into its effectiveness in preserving visual quality while achieving efficient compression at extremely low bitrates. These metrics facilitate a comprehensive comparison with existing techniques, demonstrating the advantages of using semantic information in the compression process

3.6 COMPRESSION ARCHITECTURE

The proposed semantically-guided image compression method is built upon an encoder decoder architecture designed to efficiently compress images while preserving perceptual quality at extremely low bitrates. The architecture can be summarized in the following steps:

- **ENCODER:** The encoder processes the input image and compresses it into a latent representation. This latent representation captures essential features and information about the image, enabling effective storage and transmission. The encoder employs a series of convolutional layers that extract relevant image features while reducing dimensionality, ultimately transforming the image into a compact code that can be transmitted more efficiently

- **SEMANTIC LABEL MAP GENERATION:**Alongside the image encoding, a semantic segmentation model generates a semantic label map for the input image. This label map classifies different regions within the image (e.g., sky, building, vegetation), providing critical' high-level information that informs the reconstruction process

- **DECODER** Upon receiving the latent representation and the compressed semantic label map, the decoder reconstructs the image. The decoder utilizes the semantic informationfrom the label map to guide the reconstruction process, ensuring that contextually important textures and details are accurately restored. By conditioning on the semantic label map, the decoder can allocate bits effectively, focusing on preserving important areas of the image that contribute to overall visual quality

This architecture enables the proposed method to overcome the limitations of traditional compression techniques by integrating semantic understanding into the compression and reconstruction processes.

3.7 LABEL MAP COMPRESSION

A significant innovation of the proposed methodology is the use of semantic label maps and the strategies employed to compress these maps effectively. The following techniques are implemented:

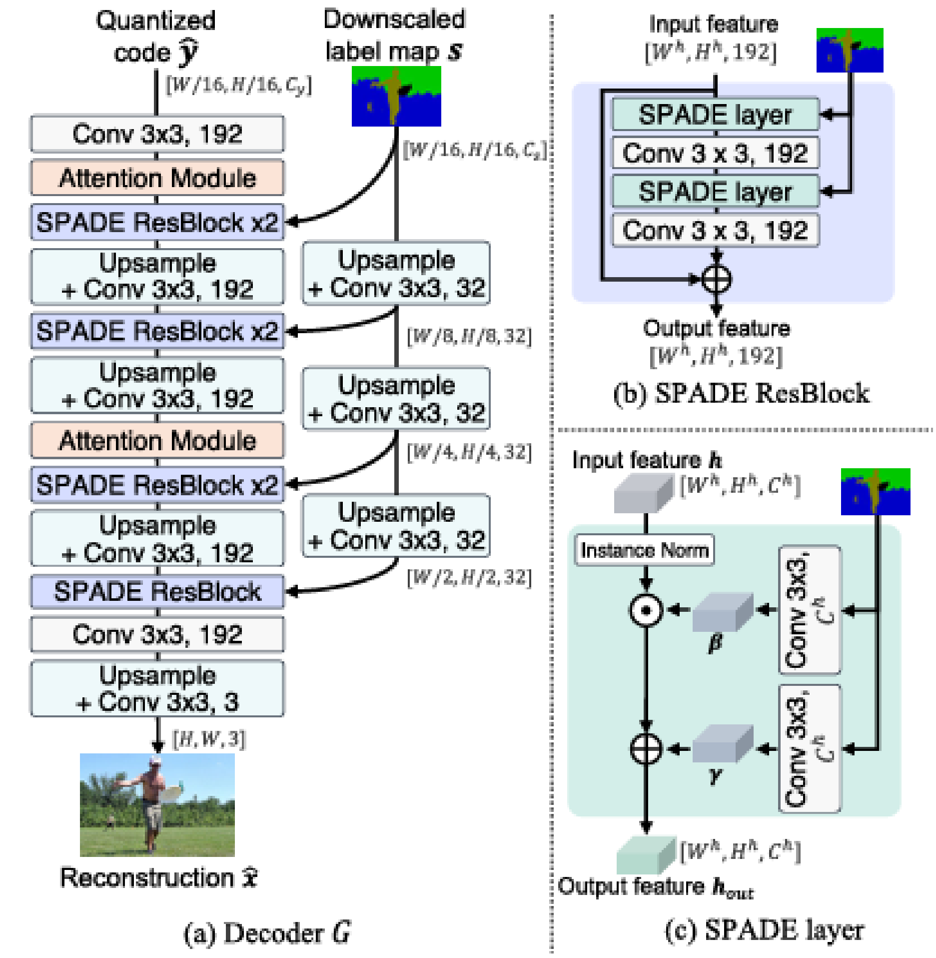

Fig 3.1: Decoder Architechture

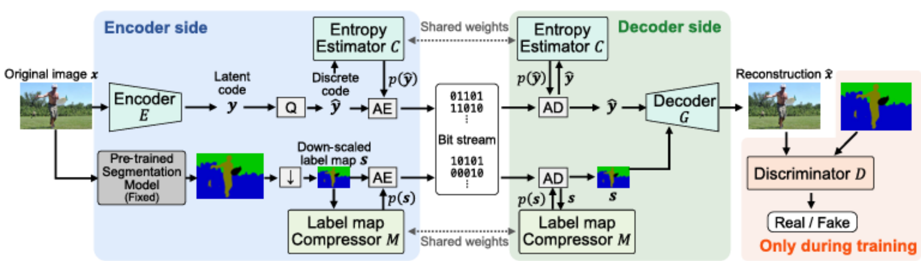

Fig 3.2: Overview of the proposed method.

- **DOWNSCALING:** The semantic label map is downscaled to reduce its resolution, which minimizes its size without significant loss of critical semantic information. This step ensures that the label map is compact enough for efficient transmission, even in bandwidth constrained environments.

- **CLASS REDUCTION:** To further decrease the size of the label map, similar semantic classes are merged. For example, categories such as ”trees” and ”grass” might be combined into a single category called ”vegetation.” This class reduction simplifies the label map, making it more efficient to compress while retaining essential contextual information.

- **AUTOREGRESSIVE COMPRESSION**An autoregressive model is utilized to compress the label map in a pixel-wise manner. This approach predicts the value of each pixel based on previously encoded pixels, leveraging spatial correlations within the label map. By encoding the label map in this manner, the method achieves higher compression rates while preserving the necessary semantic information.

These strategies for label map compression ensure that the overhead introduced by semantic information is minimized, allowing the compression method to operate effectively within low-bitrate constraints.

3.8 OPTIMISATION TECHNIQUES

The methodology proposed by Iwai [1] integrates various optimization techniques designed to improve both compression efficiency and perceptual quality.

- **RATE-DISTORTION OPTIMISATION:** This technique focuses on balancing the trade-off between compression size (rate) and image quality (distortion). By minimizing perceptualloss for a given bitrate, the method ensures that the reconstructed images retain as much detail and quality as possible.

- **HYPERPRIOR MODELS:**Hyper prior models are employed to capture the uncertainty presentin the latent representations generated by the encoder.These models help optimize the encoding process by allowing for more efficient bit allocation, ensuring that complex regions of the image receive the necessary data to maintain quality.

- **CONTEXT MODELLING:**:Context modeling is utilized to improve the accuracy of the probability predictions for the compressed data. By considering the relationships between pixels and leveraging contextual information, this technique enhances the overall performance of the compression method.

These optimization strategies are essential for ensuring that the proposed semantically guided compression method operates effectively under the stringent constraints of low bitrates while maximizing perceptual quality.

3.9 PERFORMANCE MATRICES

To assess the effectiveness of the proposed image compression method, various performance metrics are utilized:

- **LPIPS (LEARNED PERCEPTUAL IMAGE PATCH SIMILARITY):** This metric quantifies the perceptual similarity between the original and reconstructed images based on deep learning features, focusing on how similar the images appear to the human eye.
- **PSNR(PEAKSIGNAL-TO-NOISE RATIO):** PSNR measures the ratio of signal power to noise power in the images. It serves as an indicator of the fidelity of the reconstructed image relative to the original
- **MS-SSIM(MULTI-SCALE STRUCTURAL SIMILARITY INDEX):**This metric assesses the structural similarity between images at multiple scales, providing insights into how well the method preserves texture and structure
- **FID (FRECHET INCEPTION DISTANCE):** FID compares the distribution of features extracted from the original and reconstructed images, quantifying how realistic the generated images are compared to the originals.

CHAPTER 4

RESULT AND DISCUSSION

The qualitative analysis in Iwai's study [1] includes visual comparisons of reconstructed images produced by the proposed method. These are evaluated against images compressed using traditional techniques, such as JPEG and BPG, as well as other modern approaches, including GAN-based methods.

The reconstructed images from the semantically-guided method demonstrate superior perceptual quality, characterized by sharper edges, more accurate textures, and reduced artifacts, particularly in critical regions like skies, foliage, and architectural elements. Unlike traditional compression methods that often blur details when operating at low bitrates, the proposed method effectively retains both global and local features, resulting in more visually appealing images that align closely with the original input.

The dataset contains approximately 500,000 candidate profiles, 200,000 job listings, and 10 years of historical hiring data, sourced from various industries and job boards. The dataset was carefully curated to ensure diversity in job roles, industries, and candidate backgrounds, making it a representative sample of the larger labour market.

In addition to qualitative assessments, quantitative evaluations are crucial for measuring the performance of the proposed method. The following performance metrics are used:

- LPIPS (Learned Perceptual Image Patch Similarity): he results indicate that the proposed method achieves significantly lower LPIPS scores compared to conventional techniques, suggesting that the reconstructed images maintain a closer perceptual similarity to the original images
- PSNR (Peak Signal-to-Noise Ratio): While PSNR values show marginal improvements, they are often complemented with other metrics to provide a comprehensive view of image fidelity. The proposed method consistently produces higher PSNR values compared to traditional methods, demonstrating improved pixel-wise accuracy

- MS-SSIM(Multi-ScaleStructuralSimilarityIndex): The MSSSIM scores reveal that the proposed method excels in preserving structural details and textures across multiple scales, leading to a more accurate representation of the original image.
- FID (Fr´ echet Inception Distance):Lower FID values for the proposed method suggest that the distribution of features in the reconstructed images is closer to that of the original images, reinforcing the method's effectiveness in producing realistic outputs.

The integration of semantic label maps into the compression framework enhances the semantic consistency of the reconstructed images. This section focuses on how the proposed method ensures that different regions of an image are restored with contextually appropriate textures, thereby improving the overall perceptual quality.

Semantic Contextualization: The decoder leverages semantic information to differentiate between regions, applying appropriate rendering techniques to maintain the visual integrity of elements such as skies, roads, and buildings. For example, the method guarantees that the sky is rendered smoothly without artifacts, while maintaining intricate details in the foliage.

This section delves into the ability of the method proposed by Iwai [1] to effectively balance bitrate and perceptual quality. The results highlight that the semantically-guided compression approach strategically allocates bits to areas of visual significance, ensuring that essential features are maintained even under stringent bitrate limitations. This intelligent allocation allows the method to preserve critical details while minimizing the loss of quality in key regions of the image.

Additionally, a comprehensive analysis of the rate-distortion curves demonstrates that the proposed method consistently outperforms traditional compression techniques. It achieves higher perceptual quality for a given bitrate, showcasing its superiority in maintaining visual fidelity. This performance is especially evident when comparing the method to widely used approaches, such as JPEG and BPG, further solidifying the effectiveness of semantically-guided image compression.

In addition to objective metrics, user studies can provide insights into the subjective quality of the reconstructed images. This section discusses any conducted surveys or studies that gauge user preferences regarding the visual quality of images processed through different compression methods.

Participants are asked to compare images reconstructed through various techniques and provide feedback on visual appeal, clarity, and overall satisfaction. The results indicate a clear preference for the images generated by the proposed method, highlighting its effectiveness in meeting human perceptual standards.This chapter systematically presents the results obtained from the proposed semantically guided image compression method and discusses their implications. By integrating qualitative and quantitative evaluations, the findings highlight the method's strengths in preserving image quality, enhancing perceptual consistency, and maintaining efficiency in low-bitrate scenarios. The incorporation of semantic information significantly improves the visual fidelity of reconstructed images, making the proposed method a promising solution for modern image compression challenges.

As image compression becomes increasingly integrated into our digital lives, ensuring its security is paramount. This chapter addresses key security concerns associated with the proposed semantically-guided image compression method. These include data privacy risks during storage and transmission, the vulnerability to adversarial attacks that can manipulate images to deceive the compression model, and the need to secure deployed models against unauthorized access or modifications.

To mitigate these risks, the chapter emphasizes the importance of robust encryption techniques, user data control, and data anonymization for privacy. Adversarial training and detection mechanisms are crucial for improving model resilience against attacks. Securing deployed models requires robust access controls, continuous monitoring, and regular updates to address vulnerabilities and maintain system integrity.

CHAPTER 5

CONCLUSION AND FUTURE WORK

Iwai [1] introduces a novel approach to image compression known as semantically-guided image compression. This technique significantly enhances the perceptual quality of reconstructed images, particularly at extremely low bitrates. By leveraging semantic information, the method ensures that the reconstructed images are not only sharper but also more visually coherent, even under stringent compression constraints. This approach addresses a critical challenge in the field of image compression, where traditional methods often struggle to maintain visual fidelity at low bitrates.

As the demand for high-quality images continues to rise across various applications, from social media to video streaming and mobile devices, the need for efficient compression techniques that preserve visual integrity while minimizing data size has become more crucial than ever. Semantically-guided compression stands out by effectively balancing bitrate reduction with the preservation of key image features, making it an important advancement for future digital imaging. This technique holds great potential for improving user experiences in scenarios where bandwidth and storage limitations are a concern, without sacrificing image quality.

The method proposed by Iwai[1] integrates semantic label maps into both the compression and reconstruction processes, enabling a more profound understanding of the image content. This integration allows the model to identify and prioritize important regions of an image, leading to more efficient compression. By leveraging this semantic information, the method strategically allocates bits to preserve the most critical features of an image, ensuring that key textures, details, and structures remain intact even under stringent bitrate constraints.

Through this approach, Iwai [1] successfully overcomes some of the limitations associated with traditional compression methods, where important visual details are often lost in low-bitrate scenarios. The use of semantic label maps not only enhances the perceptual quality of the compressed images but also allows for more intelligent decision-making in terms of which aspects of the image to preserve, resulting in better overall visual fidelity at reduced data sizes.

To improve the system in future, several areas could be addressed. First, expanding the diversity of the dataset to include more sectors and industries would help reduce bias and increase the system's accuracy. Bias mitigation techniques, such as fairness-aware machine learning, should also be explored to ensure that the system does not perpetuate historical inequities. Another key improvement is enhancing model interpretability by employing methods such as LIME or SHAP, which can make the recommendations more transparent and easier to understand for end users. Expanding the feature set to include soft skills and company culture fit would provide a more holistic view of job recommendations. Finally, efforts should be made to improve the system's computational efficiency. Simplified models or techniques like model distillation could reduce the computational burden, making DNN-based systems more accessible to a wider range of users.

The method proposed by Iwai [1], leveraging semantic guidance, generates reconstructed images with superior perceptual quality compared to traditional compression techniques. This is demonstrated by improved LPIPS and FID scores, which reflect a significant reduction in artifacts and the preservation of visual fidelity. The ability to maintain high-quality images at extremely low bitrates is a major advantage, with the method effectively compressing images to rates below 0.1 bits per pixel without sacrificing quality.

This capability is particularly beneficial in bandwidth-constrained environments, such as mobile applications and real-time streaming services, where efficient compression is crucial for delivering high-quality content. Furthermore, by incorporating adversarial training strategies, the method enhances its robustness against potential attacks. This additional layer of security helps mitigate vulnerabilities that could compromise both output quality and the integrity of the compression model, ensuring that the method remains reliable in diverse operational settings.

While the proposed method demonstrates promising results, there remains significant potential for further improvement. Future work could explore the implementation of more complex architectures, such as multi-stage encoders and decoders, and the incorporation of additional loss functions to optimize reconstruction quality further.

In summary, the semantically-guided image compression method represents a significant advancement in the field of image compression technology. By effectively integrating seman tic understanding and leveraging advanced optimization techniques, this method addresses the limitations of both traditional and modern approaches. As the landscape of digital imaging con tinues to evolve, this research contributes valuable insights that can inform the development of robust and efficient compression solutions that meet the demands of contemporary applications.

REFERENCES

[1] Shoma Iwai, Tomo Miyazaki,Shinichiri Omachi(2014). Semantically-Guided Image Compression for Enhanced Perceptual Quality at Extremely Low Bitrates,Tohoku University, Miyagi 980-8579

[2] Wallace, G. K. (1992). The JPEG Still Picture Compression Standard. Communications of the ACM, 34(4), 30-44.

[3] Toderici, G., Vincent, D., Johnston, N., Hinton, G., Jaitly, N., Theis, L., Shor, J. (2017). Full Resolution Image Compression with Recurrent Neural Networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 5306-5314.

[4] Agustsson, E., Mentzer, F., Tschannen, M., Timofte, R., Van Gool, L. (2019). Genera tive Adversarial Networks for Extreme Learned Image Compression. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 221-231.

[5] Mentzer, F., Agustsson, E., Tschannen, M., Timofte, R., Van Gool, L. (2020). High Fidelity Generative Image Compression. Advances in Neural Information Processing Systems (NeurIPS), 33, 11913-11924

[6] Zhang, X., Liu, Z., Luo, P., Loy, C.C., Lin, D. (2018). Semantic Image Compression. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 5317-5326.

[7] Mentzer, F., Agustsson, E., Tschannen, M., Timofte, R., Van Gool, L. (2020). Enhanced Low-Bit-Rate Generative Image Compression Guided by Semantic Segmentation. arXiv preprint arXiv:2011.03138

[8] Huang, D., Gao, F., Tao, X., Du, Q., Lu, J. (2021). Deep Learning-Based Image Se mantic Coding for Semantic Communications. IEEE Global Communications Conference (GLOBECOM)

[9] Liu, Z., Meng, L., Tan, Y., Zhang, J., Zhang, H. (2021). Image Compression Based on Octave Convolution and Semantic Segmentation. Knowledge-Based Systems, 228.