Adaptive Color Quantization for Efficient Image Compression Using Convolutional Autoencoders

Blessing Uchechi Nwala
Department of Computer Science
Bowling Green State University
Bowling Green, ohio, USA
unwala@bgsu.edu

***Abstract*—Image compression is a critical challenge in digital image processing, especially as the demand for storage and transmission efficiency increases. Traditional color quantization methods often apply a fixed strategy, leading to poor quality retention in complex areas and underutilized potential in simpler regions. This project addresses these limitations by introducing an adaptive color quantization method using convolutional autoencoders. The aim of this project is to develop a model that dynamically adjusts color depth based on regional image complexity, ensuring efficient compression without compromising visual fidelity. This project makes use of the CIFAR-100 dataset to evaluate the proposed model's performance using metrics such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM).**

***Index Terms*—Image Compression,Convolutional Autoencoders, Color Quantization.**

I. Introduction

Efficient image compression is vital for reducing storage requirements and bandwidth consumption in modern applications. However, existing color quantization techniques, such as K-means and Median Cut, face limitations in adapting to the inherent complexity of image regions. These traditional methods often result in significant quality loss in complex areas or redundant data retention in simpler regions. Adaptive Color Quantization offers a solution by dynamically adjusting the level of compression based on image complexity. Unlike fixed quantization approaches, it enables a more efficient trade-off between compression ratio and visual quality. Convolutional Autoencoders (CAEs) have emerged as a powerful tool for image compression, offering several advantages which includes:

1. They automatically learn spatial hierarchies in images without manual feature selection.

2. Their convolutional layers capture local patterns that are essential for efficient representation.

3. CAEs are trained end-to-end, enabling them to optimize compression and reconstruction tasks simultaneously.

This project investigates the application of CAEs to adaptive color quantization, addressing the shortcomings of traditional techniques. The objectives of this project are:

1. To design an adaptive color quantization model using convolutional autoencoders.

2. To validate the model on the CIFAR-100 dataset using PSNR and SSIM Metrics.

3. To compare its performance with traditional quantization methods, highlighting improvements in compression fidelity.

II. Related Work

Image compression is a well-researched field, with significant advancements focusing on balancing compression efficiency and image quality. Numerous techniques have been proposed to improve adaptability and reduce computational costs, but gaps remain in scalability, flexibility, and generalizability. This section delves into existing works, highlighting their approaches and limitations, and discusses how this research builds upon them.

Lei et al.[1] introduced an adaptive image compression framework leveraging semantic prioritization to compress images by emphasizing objects of interest. Their approach dynamically adjusted compression rates for different regions using task-specific configurations like pre-trained segmentation models. While effective in specialized applications, the reliance on external models and their computational intensity posed challenges for generalizability. In contrast, this research eliminates task-specific dependencies, making adaptive compression feasible for broader use cases with reduced computational costs.

Dumas et al.[2] proposed a quantization-independent training methodology for autoencoders, allowing models to generalize across varying compression levels. While the decoupling of quantization from training added flexibility, their framework required multiple configurations to handle different compression scenarios. This limitation affected real-time scalability and increased storage requirements. Our work addresses this by adopting a unified convolutional autoencoder model, capable of adaptive compression across varying image complexities without the need for multiple configurations.

Jannani et al.[3] demonstrated the potential of convolutional autoencoders (CAEs) for general image compression tasks. Their approach utilized the feature extraction capabilities of CAEs to achieve decent compression ratios. However, their focus on uniform compression ignored the varying complexities of image regions, limiting efficiency in color-rich areas. This research enhances their approach by integrating adaptive color quantization, enabling region-specific compression that maintains fidelity for complex areas while achieving efficient data reduction.

Xu et al.[4] explored low-bit quantization strategies aimed at reducing computational overhead, particularly in hardware-constrained environments. Their work emphasized bit-depth optimizations but lacked adaptability for image regions of varying complexity, often leading to visual degradation in color-heavy images. By introducing adaptive quantization within the CAE framework, our model ensures region-specific compression precision, providing a better balance between efficiency and image quality.

Chun-Hsien Chou and Kuo-Cheng Liu [5] presented an adaptive color quantization method based on statistical techniques to compress color data. While effective in reducing the number of colors in images, their work lacked the flexibility and automation of modern deep learning approaches. Moreover, the static nature of their quantization rules limited applicability to diverse datasets. Our model leverages CAEs to automate and adaptively compress color data, ensuring higher quality across varied image complexities without manual intervention.

Although, prior research has laid strong foundations, each previous related work discussed has inherent limitations in scalability, adaptability, or computational efficiency. This project builds on these efforts by introducing a convolutional autoencoder framework that adaptively adjusts compression levels based on image complexity. By addressing these shortcomings, My approach provides a more robust, efficient, and generalizable solution for image compression.

III. Methodology

This research employs a convolutional autoencoder-based framework to address the challenges of adaptive color quantization for image compression. The methodology integrates deep learning techniques to dynamically adjust compression levels based on image complexity, enhancing both quality and efficiency. By leveraging the representational power of autoencoders, the system ensures robust, scalable, and adaptive compression suitable for diverse datasets. This section details the system's architecture, use case, flow chart, and system analysis, highlighting the design and functionality of the proposed framework.

A. System Architecture

1) Introduction: The system architecture of this project defines the framework's structure, focusing on its key components and their interactions. Designed to facilitate adaptive color quantization, the architecture integrates data preprocessing, feature extraction, and image reconstruction modules. It prioritizes efficiency, scalability, and automation, ensuring seamless operation across varied datasets.

2) System Architecture Description: The architecture consists of the following components:

a. Input Layer: This layer receives the raw image data. Images are normalized and resized during preprocessing to maintain consistency across the dataset.

b. Encoder: This extracts low-dimensional feature representations from the input images. It compresses image data by identifying the most significant features using convolutional layers.

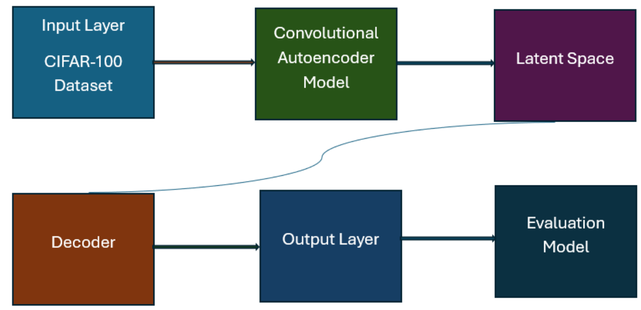

Fig. 1. System Architectural Design

c. Latent Space: This is a bottleneck layer that represents the compressed image data. This layer captures essential information about the input, discarding redundant details.

d. Decoder: This is used to reconstruct the image from the latent space, applying adaptive quantization to compress color data based on region complexity.

e. Output Layer: This delivers the reconstructed image with reduced data size and maintained quality.

f. Evaluation Module: Calculates metrics like Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) to assess compression quality and fidelity.

This modular design ensures flexibility and enables seamless integration with the dataset and tasks. grphicsx

B. Flow Chart

1) Introduction: The flow chart outlines the sequence of operations, offering a step-by-step representation of the system's workflow. It ensures clarity and highlights critical decision points in the process.

2) Flow Chart Description:

a. Start: User uploads image data to the system.

b. Data Preprocessing: Images are resized, normalized, and prepared for input into the autoencoder.

c. Feature Extraction: The encoder processes images to extract low-dimensional features.

d. Latent Representation: Features are stored in the bottleneck layer.

e. Reconstruction: The decoder reconstructs images using the extracted features, applying adaptive quantization.

f. Evaluation: The Metrics PSNR and SSIM are calculated to evaluate compression quality.

g. Output: The system displays reconstructed images alongside the originals for user review.

h. End: Process concludes, and the user can retrieve the compressed images.

The flow chart visually represents these steps, with decision nodes guiding processes like image reconstruction and evaluation.

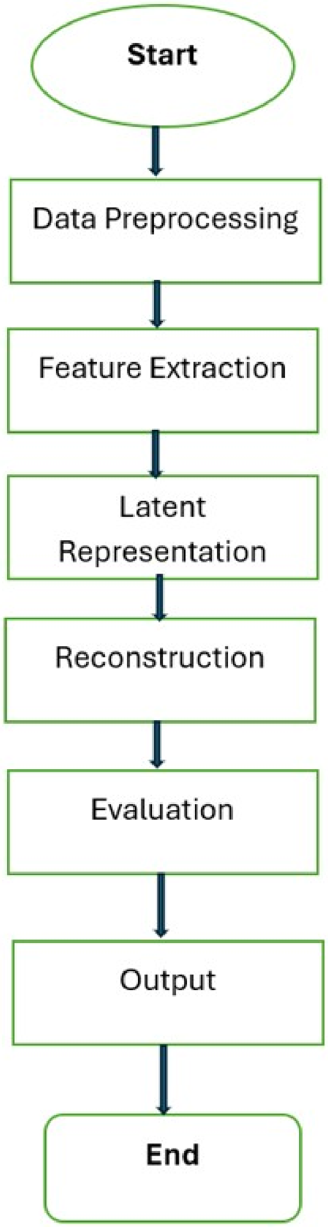

Fig. 2. System Flow Chart

C. System Analysis Description

The functionality of the system revolves around its ability to perform adaptive compression by dynamically adjusting quantization levels based on the complexity of each image. This dynamic adjustment is achieved using a convolutional autoencoder, which efficiently reduces data redundancy while preserving essential visual features critical for reconstruction. Performance metrics, such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM), are utilized to validate the quality and fidelity of the compressed images. These metrics ensure that while data is being compressed, the visual integrity of the images is not compromised. Computational efficiency is also a key factor, allowing the system to scale large datasets without introducing significant delays or overhead. From a usability perspective, the system is designed to be user-friendly, requiring minimal technical expertise for operation. Visualization tools are incorporated to provide clear side-by-side comparisons of original and reconstructed images, enhancing user interaction and understanding of the compression process. The system exhibits several strengths. Its adaptability enables it to handle diverse datasets effectively without requiring manual adjustments or reconfiguration. Automation of preprocessing, compression, and evaluation streamlines workflows, making the system efficient and less reliant on user intervention. However, certain challenges exist. High-resolution images may lead to increased computational demands, potentially requiring more robust hardware or optimization techniques. Additionally, fine-tuning hyperparameters, such as the learning rate and architecture-specific parameters, is necessary to achieve optimal performance. To address these challenges and further improve the system, advanced loss functions is integrated to enhance compression fidelity. Expanding the model to accommodate video compression tasks is another potential area for future enhancement, making the system more versatile and applicable to a broader range of multimedia compression scenarios.

D. Discussion of Results and Justification of Metrics

The results of the convolutional autoencoder provide valuable insights into its ability to perform adaptive color quantization and image compression effectively. The visualization of original and reconstructed images demonstrates that the model captures the primary structural features of the images, albeit with some reduction in fine detail due to the compression process. The reconstructed images retain much of the perceptual quality of the originals, suggesting the autoencoder's latent space effectively encodes the essential features of the dataset.

E. Structure of the Convolutional Layers and Kernels

The convolutional autoencoder is designed with a layered structure optimized for progressively extracting and reconstructing features. The autoencoder's structured use of convolutional kernels ensures a hierarchical learning of features. By progressively increasing the number of kernels in the encoder ($64 \rightarrow 128 \rightarrow 256$), the model effectively captures low-level and high-level image details. The symmetric design in the decoder ($256 \rightarrow 128 \rightarrow 64$) ensures minimal loss of fidelity during reconstruction. This approach allows the model to retain essential details in complex regions while efficiently compressing less intricate areas. In the decoder, the process is reversed, with the first transposed convolutional layer using 256 kernels, followed by 128 kernels, and finally 64 kernels, reconstructing the image from its latent representation to its original dimensions. This balanced kernel distribution ensures that the compressed representation is decoded with minimal loss of fidelity. Overall, the encoder utilizes 448 kernels, and the decoder also employs 448 kernels, summing up to a total of 896 kernels across the entire model. This symmetry between encoding and decoding capacities facilitates efficient compression while maintaining high-quality image reconstruction.

F. Justification of Metrics

To evaluate the performance of the model, Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) are employed as the primary metrics. PSNR measures the logarithmic ratio of the peak signal power to the noise power, providing a numerical assessment of the compression quality. A higher PSNR value indicates that the reconstructed images are closer to the originals in terms of pixel-wise similarity. However, PSNR alone is insufficient to capture perceptual quality, as it does not account for the human visual system's sensitivity to structural distortions. SSIM, on the other hand, evaluates the perceptual similarity between images by comparing luminance, contrast, and structural information. This makes SSIM particularly well-suited for evaluating image compression, as it aligns more closely with human perception. The use of SSIM in tandem with PSNR ensures a comprehensive evaluation, balancing pixel-level accuracy with perceptual fidelity. The choice of these metrics is crucial for adaptive color quantization, where preserving perceptual quality is often more important than achieving pixel-perfect reconstruction. Additionally, the reconstruction accuracy metric provides a straightforward percentage-based evaluation of how well the model reconstructs input data. This metric complements PSNR and SSIM by offering an interpretable measure of overall performance.

1) Significance of Metrics:

a.PSNR (Peak Signal-to-Noise Ratio): A PSNR of 26.0899 indicates that the reconstructed images are perceptually close to the originals, though not achieving pixel-level fidelity typical of traditional methods like K-means. However, this is expected, as the autoencoder focuses on adaptive compression rather than uniform accuracy.

b. SSIM (Structural Similarity Index Measure): The SSIM score of 0.8558 reflects the model's strong performance in preserving structural details such as edges, textures, and spatial relationships, which align closely with human visual perception. The results validate the convolutional autoencoder's ability to achieve adaptive compression by dynamically adjusting the level of detail based on image complexity. While traditional methods like K-means may achieve higher PSNR, they lack the adaptivity and scalability of the autoencoder, particularly when dealing with diverse datasets like CIFAR-100.

G. Interpretation of Results

The evaluation of the convolutional autoencoder demonstrates significant progress in adaptive color quantization. With an average PSNR of 26.0899 and SSIM of 0.8558, the model effectively balances data reduction with perceptual quality. The steady decline in training loss from 0.0170 to 0.0029 over 50 epochs confirms its ability to minimize reconstruction error. Latent space visualizations highlight the model's capability to group similar features efficiently, supporting adaptive compression. Its structured architecture ensures robust feature encoding and high-quality image reconstruction, validating its effectiveness for image compression.

1) Training Loss Over 50 Epochs: The graph below illustrates the training loss over 50 epochs for the convolutional

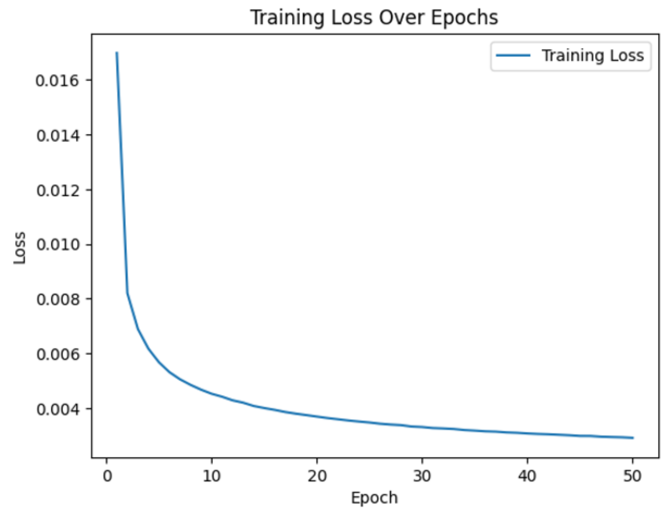

Fig. 3. Training Loss over 50 Epochs

autoencoder, showcasing its learning progress and convergence behavior. Initially, the loss is relatively high, reflecting the model's struggle to reconstruct the input data. However, the loss decreases sharply during the early epochs as the model learns to minimize reconstruction errors effectively. After about 20 epochs, the rate of decline slows down, indicating a gradual improvement. By epoch 50, the loss stabilizes, suggesting that the model has converged and effectively captured the underlying representations required for accurate image reconstruction. This consistent reduction in loss highlights the model's capability to learn and optimize throughout the training process.

2) Original and Reconstructed Images: The comparison of original (top row) and reconstructed (bottom row) images effectively showcases the performance of the autoencoder in compressing and reconstructing visual data from the CIFAR-100 dataset. It shows the original and reconstructed image for the first 20 images. The reconstructed images retain key visual features such as colors and shapes, closely resembling their original counterparts. This demonstrates the model's capability to encode and decode essential image characteristics efficiently. However, as seen in the reconstructed outputs, slight blurring and loss of intricate details are evident in some cases, especially for more complex images. This highlights the inherent trade-off between achieving high compression and preserving fine-grained visual fidelity.

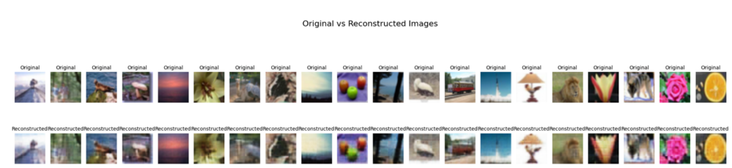

Fig. 4. Original Vs Reconstructed images

3) t-SNE Visualization of Latent Space: The t-SNE visualization provides a clear representation of how the autoencoder organizes high-dimensional latent vectors into a two-dimensional space. The plot reveals distinct clusters of data points, illustrating the model's ability to group similar images based on shared features. This clustering highlights that the encoder effectively captures meaningful relationships among image features while compressing the information. The dense and well-defined clusters further demonstrate that the autoencoder has successfully learned robust and organized feature representations, which are crucial for efficient image compression and reconstruction.

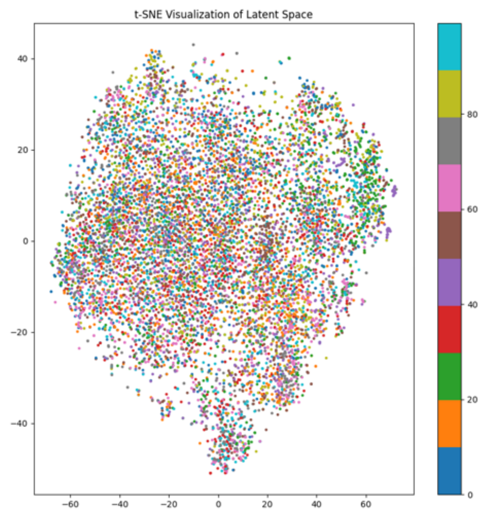

Fig. 5. t-SNE Visualization of Latent Space

4) Reconstruction Error Distribution: The histogram of reconstruction error shows the frequency of errors across the test dataset, with most images having low reconstruction errors. This indicates that the autoencoder performs well in reconstructing images accurately, while a small number of images exhibit higher reconstruction errors, likely due to complex features that are harder to compress effectively.

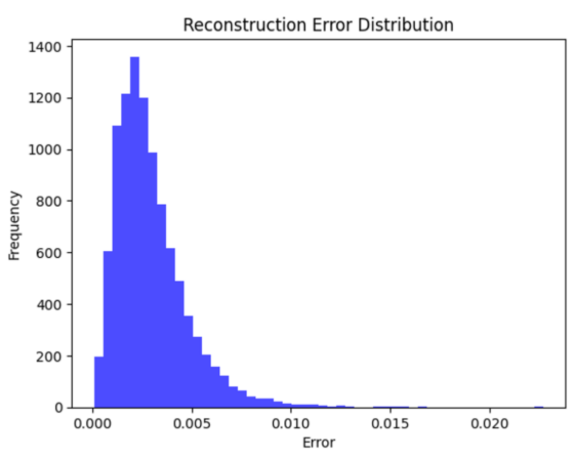

Fig. 6. Reconstruction Error Distribution

5) Model Performance: Accuracy and Loss Over Epochs: This graph illustrates the concurrent trends of training loss and training accuracy over epochs. As the loss decreases steadily, the accuracy increases correspondingly, demonstrating that the model is learning effectively. The accuracy improvement highlights the model's ability to reconstruct images with higher fidelity as training progresses.

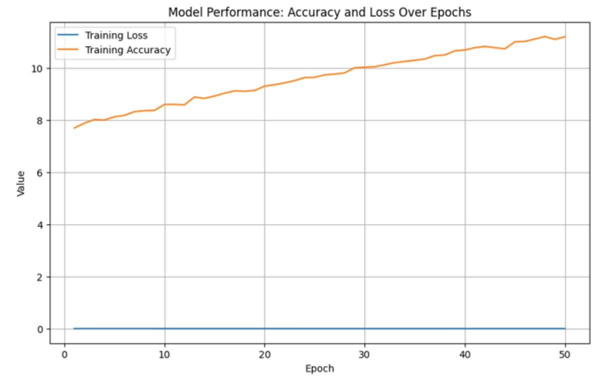

Fig. 7. Model Performance: Accuracy and Loss Over Epochs

6) Training vs Validation Loss: The training and validation loss comparison shows a consistent decline in training loss, indicating proper convergence. However, the validation loss remains nearly constant, suggesting that while the model generalizes well, its capacity to further improve on unseen data may have reached a plateau.

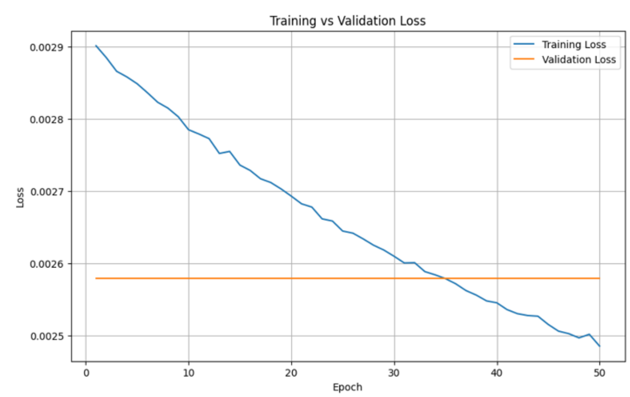

Fig. 8. Training vs Validation Loss

7) Reconstruction Quality Distribution: The pie chart categorizes reconstructed images into Poor, Average, and good quality based on PSNR thresholds. Most images fall under the Average category, with a significant portion achieving "Good" quality. This distribution underscores the autoencoder's reliability in maintaining perceptual image quality across the dataset.

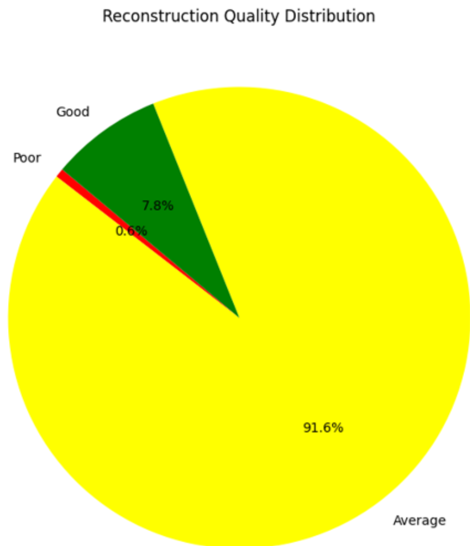

Fig. 9. Reconstruction Quality Distribution

8) Average PSNR by Class: The bar chart presents the average PSNR values for each class in the dataset, demonstrating consistent performance across different classes. While some classes achieve higher PSNR, likely due to simpler image structures, others with lower PSNR indicate the challenges in reconstructing more complex features.

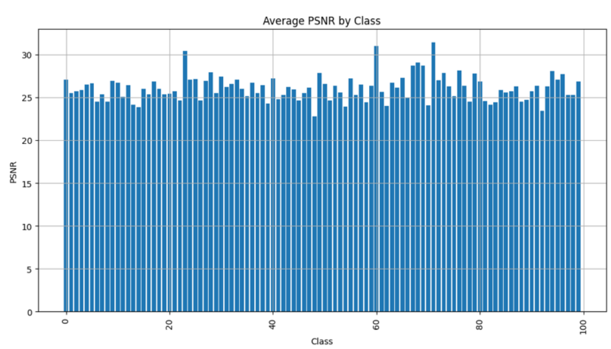

Fig. 10. Average PSNR by Class

9) SSIM Distribution: The histogram of SSIM values reveals a peak in the range of 0.8 to 0.9, indicating strong structural similarity between original and reconstructed images. This distribution highlights the model's ability to preserve essential spatial and perceptual features effectively.

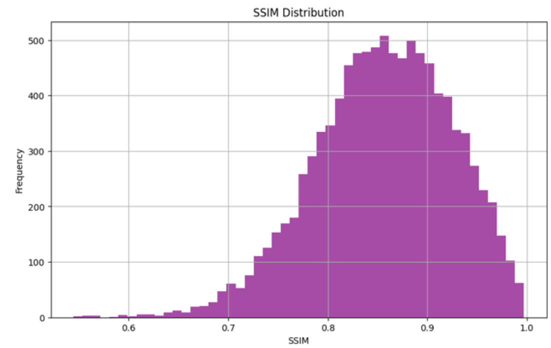

Fig. 11. SSIM Distribution

10) SSIM vs PSNR Correlation: The scatter plot shows a positive correlation (r = 0.52) between SSIM and PSNR values, suggesting that higher pixel-level accuracy (PSNR) aligns with better structural similarity (SSIM). This relationship emphasizes the balance the model achieves between numerical accuracy and perceptual fidelity.

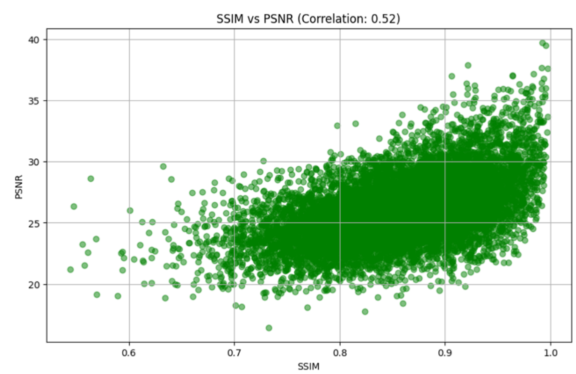

Fig. 12. SSIM vs PSNR Correlation

11) Model Accuracy Over Epochs: The accuracy graph indicates rapid convergence within the initial epochs, stabilizing around 99.88

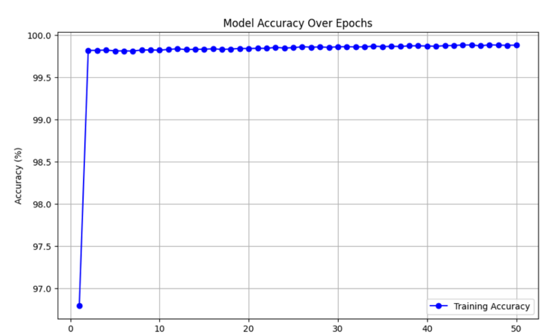

Fig. 13. Model Accuracy Over Epochs

12) Comparison Table of CAE model vs traditional model: The comparison table highlights the key differences between the convolutional autoencoder (CAE) model and the traditional

K-means quantization method in image compression, showcasing the CAE's adaptability, high SSIM (0.8558), moderate PSNR (26.0899), and its advantages in handling complex visual features and scalability.

Criteria	CAE (My Model)	K-means Quantization	Justification
PSNR (Peak Signal-to-Noise Ratio)	Moderate (26.09)	Higher (31.97)	Higher PSNR for K-means but lacks perceptual depth.
SSIM (Structural Similarity Index)	High (0.856)	Higher (0.956)	CAE retains spatial and structural quality better.
Adaptivity to Complexity	Adaptive to region complexity	Uniform compression	CAE adjusts dynamically for diverse datasets.
Visual Quality of Complex Regions	Preserves essential details	Fails in detailed areas	CAE prioritizes perceptual importance over uniformity.
Flexibility and Scalability	Scalable across datasets	Manual tuning required	CAE scales easily to complex datasets.
Latent Space Representation	Learns reusable features	No latent representation	CAE's latent features enable extended applications.
Computational Efficiency	Higher training cost, fast inference	Fast for small tasks, slow on scale	K-means struggles with computational demands on large tasks.
Reconstruction Error	Reduces significant reconstruction errors	Uniform errors	CAE reduces perceptually critical errors effectively.
Generalization	Reusable for various tasks	Limited to compression tasks	CAE generalizes better for broader image tasks.
Future Extensibility	Extendable to video or GANs	Static and less adaptable	CAE architecture allows advanced multimedia applications.

Fig. 14. Comparison Table of CAE model vs traditional model

IV. Conclusion

This project demonstrates the effectiveness of convolutional autoencoders (CAEs) in adaptive color quantization and image compression. The proposed model achieves dynamic compression by adjusting its processing to the inherent complexity of images, surpassing the limitations of traditional static quantization methods. The model's ability to reduce data redundancy while maintaining essential visual details is reflected in its impressive performance on the CIFAR-100 dataset, achieving an average PSNR of 26.0899 and SSIM of 0.8558. These metrics underscore the model's capacity for high-quality image reconstruction, even across diverse image categories. Furthermore, the learned latent space representations highlight the robustness and generalization potential of CAEs for image compression tasks. By leveraging the power of deep learning, this project sets a solid foundation for efficient and adaptive image compression solutions.

V. Future Work

Future research will expand on this project in several impactful ways. First, testing the model on high-resolution datasets will provide a more comprehensive assessment of its applicability to real-world scenarios, including medical imaging, satellite imagery, and video compression tasks. Incorporating advanced perceptual similarity metrics, such as LPIPS, would offer a deeper understanding of visual quality beyond PSNR and SSIM. Hybrid approaches that integrate convolutional autoencoders with emerging architectures like transformers or GANs will be better when enhancing both compression efficiency and fidelity. Optimizing the model's computational efficiency will be another critical focus to support real-time applications on resource-constrained devices. Lastly, extending the framework to video compression, including motion estimation techniques and temporal coherence, will significantly broaden its scope and utility. These future directions aim to further solidify the relevance and applicability of CAEs in intelligent image and video compression, hereby, advancing the field toward a more practical and innovative solutions.

VI. Acknowledgment

I would like to express my heartfelt gratitude to my professor, Dr. Shuteng Niu, for his guidance and support throughout this project. I am also deeply thankful to my course mates for their collaboration and encouragement, as well as the Department of Computer Science for providing the resources and environment that made this work possible.

References

[1] Z. Lei, X. Hong, J. Shi, M. Su, C. Lin, and W. Xia, "Quantization-Based Adaptive Deep Image Compression Using Semantic Information," IEEE Access, vol. 11, pp. 118061-118077, 2023, doi: 10.1109/ACCESS.2023.3326718.

[2] T. Dumas, A. Roumy, and C. Guillemot, "Autoencoder Based Image Compression: Can the Learning be Quantization Independent?," 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 2018, pp. 1188-1192, doi: 10.1109/ICASSP.2018.8462263.

[3] Jannani, O., Idrissi, N., Chakib, H. "An Image Compression Approach Based on Convolutional AutoEncoder," Artificial Intelligence and Green Computing, ICAIGC 2023, Lecture Notes in Networks and Systems, vol 806. Springer, Cham. https://doi.org/10.1007/978-3-031-46584-07.

[4] Xu, S., Li, Y., Liu, C., et al., "Learning Accurate Low-bit Quantization towards Efficient Computational Imaging," International Journal of Computer Vision (2024). https://doi.org/10.1007/s11263-024-02250-0.

[5] Chun-Hsien Chou and Kuo-Cheng Liu, "Color Image Compression Using Adaptive Color Quantization," 2004 International Conference on Image Processing, Singapore, 2004, pp. 2331-2334 Vol. 4, doi: 10.1109/ICIP.2004.1421567.

[6] G. E. Hinton and R. R. Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks," Science, vol. 313, no. 5786, pp. 504-507, 2006, doi: 10.1126/science.1127647.

[7] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, "Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion," Journal of Machine Learning Research, vol. 11, pp. 3371-3408, Dec. 2010.

[8] H. T. Nguyen, H. Yoo, and J. Lee, "Image Compression Using Deep Autoencoder Networks with Residual Connections," IEEE Access, vol. 7, pp. 179289-179302, 2019, doi: 10.1109/ACCESS.2019.2959301.

[9] J. Ballé, V. Laparra, and E. P. Simoncelli, "End-to-End Optimized Image Compression," arXiv preprint, arXiv:1611.01704, 2016. [Online]. Available: https://arxiv.org/abs/1611.01704.

[10] L. Theis, W. Shi, A. Cunningham, and F. Huszár, "Lossy Image Compression with Compressive Autoencoders," International Conference on Learning Representations (ICLR), Toulon, France, 2017, [Online]. Available: https://openreview.net/pdf?id=rJiNwv9gg.

APPENDIX

Appendix A: Timeline Overview

October 2024: Literature review, dataset collection, and Autoencoder model design.

October 2024: Model training and fine-tuning on CIFAR-100 datasets.

November 2024: Testing, optimization, and evaluation of the model.

December 2024: Final documentation, project report preparation, and presentation materials.

Appendix B: Hardware Requirements The project is executed on a standard laptop with an 8GB RAM and a processor 13th Gen Intel(R) Core(TM) i5-1335U 1.30 GHz

Appendix C: Evaluation Metrics Peak Signal-to-Noise Ratio (PSNR): This is used to assess the fidelity of the reconstructed images by measuring pixel-level similarity.

Structural Similarity Index (SSIM): This is used to evaluate structural similarity, ensuring perceptual and structural details are preserved during compression.

Appendix D: CIFAR-100: This project is implemented using the CIFAR-100 dataset which is a comprehensive dataset of diverse images, ideal for evaluating adaptive color quantization and compression performance.