FMDB *Transactions on Sustainable Computer Letters*

Multi-Stage Convolutional Autoencoder with Adaptive Quantization and Secure Hash-Based Encryption for Optimized Image Compression and Retrieval

M. Rithicka[1,*], Ruhi Fathima[2], D. Lalitha Sree[3]

[1,2,3]Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India.
rm3287@srmist.edu.in[1], rf9878@srmistedu.in[2], ls1021@srmist.edu.in[3]

Abstract: This paper presents a Multi-Stage Convolutional Autoencoder Adaptive Quantization with Secure Hash-Based Encryption (MSCAE-AQ-SHE) for enhanced image compression and retrieval. The model is trained on the CIFAR-10 dataset, which contains 60,000 RGB images across 10 classes. Preprocessing scales pixel values to the [0, 1] range and resizes images to 32 × 32 pixels. Convolutional autoencoders learn compressed latent representations, thus able to save storage while retaining visual quality. The training utilizes the Adam optimizer (learning rate = 0.001), while learning rate scheduling and early stopping are employed to prevent the model from overfitting. Experimental results demonstrate that the proposed system can compress images by up to 85% while maintaining a PSNR of 30 dB or higher, indicating minimal loss in image quality. Adaptive format selection dynamically chooses among JPEG, PNG, or WebP to store the images, balancing size versus quality. A two-layer cryptographic technique based on SHA-256 and MD5 hashing algorithms helps maintain integrity and prevent unauthorised access, thereby enhancing the system. User interaction on Streamlit and retrieval, secure picture storage, and metadata management utilising SQLite are also significant areas of concern. By combining deep learning, adaptive quantisation, and cryptographic security, this provides a highly efficient and secure solution for contemporary image compression and retrieval applications.

Keywords: Image Compression; Convolutional Autoencoder; Adaptive Quantisation; Cryptographic Security; MD5 and SHA-256; Peak Signal-to-Noise Ratio (PSNR); Discrete Cosine Transformation; Generative Adversarial Networks.

Received on: 12/08/2024, **Revised on:** 27/10/2024, **Accepted on:** 07/11/2024, **Published on:** 01/03/2025

Journal Homepage: https://www.fmdbpub.com/user/journals/details/FTSCL

DOI: https://doi.org/10.69888/FTSCL.2025.000357

Cite as: M. Rithicka, R. Fathima, and D. L. Sree, "Multi-Stage Convolutional Autoencoder with Adaptive Quantization and Secure Hash-Based Encryption for Optimized Image Compression and Retrieval," *FMDB Transactions on Sustainable Computer Letters*., vol. 3, no. 1, pp. 22–33, 2025.

1. Introduction

The surge in high-resolution images and multimedia content has given rise to an increasing need for efficient image compression techniques that maintain storage economy, transmission speed, and image quality. Traditional compression methods, such as JPEG, JPEG2000, HEVC, and VVC, utilise handcrafted mathematical transformations, including the Discrete Cosine Transformation (DCT) and wavelet-based techniques, followed by entropy coding in separate sequential steps. Although they are effective, these methods suffer from a loss of perceptual quality at high compression rates and are unable to dynamically

*Corresponding author.

adapt to the complex structures of images. To address these limitations, researchers have employed deep learning models to create a more compact and functional representation of images, eliminating the need for explicit, handcrafted encoding rules.

Deep learning-based image compression models utilise data-driven representations to compress images more effectively with minimal loss in quality. To maintain the quality of the image in PNG and GIF, Huffman coding and Deflate compression are used. While they guarantee a similar reconstruction, they are not a suitable option for large-scale multimedia applications due to their low compression efficiency. These methods can generate visual abnormalities, such as colour banding, blocking, and blurring, at high compression ratios, which can damage crucial image features and compromise the integrity of the data. A prominent issue in image compression is maintaining a balance between compression efficiency and detail preservation, as identified by Liu et al. [2]. More advanced explicit image compression techniques, such as HEVC, have been developed to address these limitations.

Deep learning enables models to learn efficient feature representations rather than relying on a manually designed compression framework, thereby revolutionising image compression. According to Cheng et al. [1], autoencoders are trusted to provide improved compression performance over current image compression standards, such as JPEG and JPEG2000, because they can extract more compressed codes from images with a lower loss function. In an autoencoder-based compression, the image is first encoded into a compact latent space and then reconstructed using a decoder. CNNs have demonstrated strong performance in image compression; however, challenges arise due to quantisation-induced gradient issues and the non-differentiability of rate-distortion optimisation. Nevertheless, CNN-based compression techniques utilise spatial correlations and hierarchical image attributes to enhance efficiency, as discussed by Ma et al. [3].

Our work intends to create an artificial neural network (ANN) backpropagation-based image compression model. Backpropagation is a fundamental technique in deep learning that enables neural networks to learn from their mistakes and refine their internal representations repeatedly. Employing gradient-based learning, our method trains an ANN to encode and decode images by minimising reconstruction error. Unlike conventional handmade compression methods, our approach enables the model to dynamically learn the most effective compression algorithms for various image types, thereby allowing it to adapt to complex patterns that are typically overlooked by predetermined mathematical transformations. The model learns to maintain semantically significant elements while achieving aggressive compression by incorporating convolutional layers for spatially aware feature extraction and optimising a hybrid loss function, which combines reconstruction error with perceptual measurements.

Significant blurring is frequently caused by high compression ratios, particularly at very low bitrates, where the perceptual quality of the reconstructed image declines. To overcome this problem, certain methods utilise generative adversarial networks (GANs), which enhance perceptual quality and reduce blurriness; however, they also come with drawbacks such as unstable training and the inclusion of unwanted noise or artefacts in reconstructions. Alternatively, autoencoders offer a more efficient solution by learning concise latent representations while retaining key image features, which provides a more stable training procedure and improved reconstruction quality without introducing excessive noise. To optimise the latent representation without compromising reconstructability and to simultaneously train the encoder and decoder, the end-to-end backpropagation technique is improved.

The suggested method optimises storage and processing efficiency while maintaining key features through the use of an autoencoder-based compression framework, much like the approach proposed by Naveen et al. [11]. The latent space representation automatically secures the compressed data, making it unreadable without the decoder. Furthermore, data integrity and protection are ensured by the application of additional cryptographic hashing techniques, such as SHA-256 and MD5. Due to the increased need for secure and efficient image compression, it is important to ensure that the compressed images can be securely stored and retrieved without losing the quality of the data. According to the work of Naveen et al. [11], Traditional image compression models primarily focus on reducing image size, but they often lack a proper and secure retrieval mechanism. To overcome this challenge, we have employed an encryption-based retrieval system with our autoencoder-based compression model. By generating a unique encryption key for each compressed image using SHA-265 and MD5, we ensure that only users with the correct key can retrieve and reconstruct the original image from the database. Thereby, improving the integrity and security of the data, making it the ideal choice for applications that require the reliable preservation of confidential data without risking unauthorised access.

The primary advantage of this backpropagation-based technique is its capacity to maximise compression efficiency while preserving effective perceptual quality. Learning directly from data enables our model to retain key details efficiently, even at low bitrates, and to capture semantic elements. To further enhance compression performance, our study will test various network designs, loss functions, and training optimisation methods. With this work, we aim to enhance deep learning image compression by demonstrating that backpropagation-based learning can be a more efficient and effective alternative to both conventional and modern compression techniques. Image compression models based on deep learning are expected to be both more versatile and efficient, as highlighted by Cheng et al. [1]. We aspire to contribute further to the ongoing advancements in intelligent image compression by applying the concepts of neural network optimization and adaptive feature learning, paving the way for more effective, scalable, and high-fidelity compression systems.

The rest of this paper is organised as follows: Section II provides a review of the existing literature and background on image compression techniques, covering both traditional methods and recent advancements in deep learning-based approaches, with a focus on autoencoders, GANs, and other relevant models. Section III outlines the methodology employed in this research, detailing the architecture of the proposed Multi-Stage Convolutional Autoencoder with Adaptive Quantisation and Secure Hash-Based Encryption (MSCAE-AQ-SHE). This section also discusses the architecture components, including the encoder, decoder, quantisation process, and encryption methods that utilise SHA-256 and MD5 hashing algorithms. Section IV presents the experimental results and discussion, comparing the performance of our proposed MSCAE-AQ-SHE model with other image compression techniques, such as K-Means clustering and PCA. This section includes an analysis of metrics such as compression ratio, Peak Signal-to-Noise Ratio (PSNR), and Mean Squared Error (MSE). Section V offers concluding remarks, summarising the key findings of our research and discussing potential future work and extensions of this study.

2. Review of Literature

Cheng et al. [1] proposed a lossy image compression architecture that utilises convolutional autoencoders and principal component analysis (PCA) to enhance compression. This approach outperforms traditional methods, such as JPEG 2000, in terms of coding efficiency while maintaining a comparable computational complexity. By utilising deep learning techniques, this method achieves superior rate-distortion performance, leveraging neural networks to advance image compression technology. This technique demonstrates improved performance, as it showcases the capability of neural networks in producing a significantly better image, surpassing the capabilities of traditional methods.

Liu et al. [2] introduced a fast fractal-based compression algorithm for MRI images. This technique focuses on improving compression speed by a factor of 23. It also enhances the image quality and makes it highly suitable for medical imaging applications. The algorithm primarily addresses the long-standing discrepancy between achieving a high compression ratio and preserving critical diagnostic information, ensuring that the compressed medical images retain the essential information necessary for accurate diagnoses. 23 times increase in speed is the primary reason for enhanced image quality, which is attributed to the compression ratio, as well as diagnostic details. This technique produces an improved image while maintaining the quality of the original scanned images, ensuring accuracy throughout the entire process.

Ma et al. [3] proposed a novel approach to address the limitations of existing hybrid coding frameworks, which are commonly used for image and video compression. This research introduces innovative techniques that facilitate significant advancements in compression efficiency by utilising the front-end of visual data. The proposed method addresses both rate-distortion performance and computational techniques, making it a novel solution for future video coding standards. A systematic overview of methods and techniques was employed in the compression of images and videos, particularly in the context of neural networks. Networks. It discussed the HEVC framework in depth, which promotes state-of-the-art video coding. The system focuses on the modern element by overcoming traditional barriers, thereby providing a path for further advancement.

Iwai et al. [4] presented a GAN-based image compression method that achieves high-quality reconstructions at extremely low bitrates. It is achieved by incorporating a two-stage training process and utilising network interpolation techniques. This approach stabilises training while reducing noise. The method effectively preserves critical image details, ensuring that the reconstructed images retain their essential features even at low rates. This advancement in compression has the potential to revolutionise low-bandwidth image transmission and its storage capacity. It also offers stabilised training and reduced noise while saving important details most efficiently.

Tellez et al. [5] put forth a Neural Image Compression (NIC) method that combines unsupervised compression with CNN-based label prediction. Unlike traditional methods, which usually rely on pixel-level accuracy, this approach utilises image-level labels to achieve proper image compression. This is achieved by eliminating the need for fine-grained annotations, allowing the proposed method to streamline compression while focusing on improved performance in various image classification tasks. This study also highlights the advantages of integrating machine learning with image compression for efficiency and usability. The neural network model involved precisely estimating the image size to achieve even better performance, which also improves accuracy in both pixel and size, resulting in better images.

Hu et al. [6] conducted a full review of the evolution of data-driven methods in image compression. The review highlights how deep learning and artificial intelligence have been combined with traditional compression techniques to enhance image quality, resulting in higher efficiency and improved image quality. Additionally, the study also discusses the challenges encountered while implementing data-driven compression, including computational costs and generalisation issues. It perfectly shows the transformation of traditional compression. The work outlines additional issues encountered during the operation of compressions, including size problems and pixel differences. The findings offer valuable insights into the future potential of machine learning-based compression techniques utilising artificial intelligence across various domains.

Guo et al. [7] proposed an advanced image compression method that combines causal context and global prediction models for a novel entropy coding approach. This technique achieves better performance and surpasses the capabilities of the VVC/H.266 codec. It incorporates contextual information into the compression process. The proposed method enhances the coding efficiency and reduces redundancy, making it a promising alternative for next-generation image compression standards. Here, entropy coding and achieving state-of-the-art rate-distortion performance are improved together through the use of context information and prediction models, which also enhances the overall quality and compressibility of the images.

Lu et al. [8] demonstrated a preprocessing-enhanced image compression method that is specifically designed for machine vision applications. This technique focuses on preserving useful information while removing irrelevant details, for which compressed images are optimised for downstream tasks, including object detection and classification. The experimental results demonstrated that the proposed method reduced the bitrate by 20% and also improved the performance of subsequent machine-related tasks. This is suitable and highly beneficial for AI-driven image processing systems. A significant benefit is that the semantic information and elimination of irrelevant details result in a substantial bitrate reduction and improved task performance. Liang et al. [9] established a novel image compression algorithm that combines K-means clustering with neural networks. This utilises various clustering techniques to group similar image regions. This approach enhances compression efficiency while maintaining high image quality.

The proposed algorithm also yields a better peak signal-to-noise ratio (PSNR) and faster runtime compared to traditional methods, making it a viable solution for real-time image compression applications. The main highlight is that the algorithm combines K-means clustering and neural networks, achieving high compression with improved PSNR and a new runtime. Bao et al. [10] developed MS-CAE, an enhanced image compression technique that is tailored for wireless sensor networks. The method utilises model segmentation to optimise deployment across edge and cloud computing environments. T improves both compression efficiency and transmission performance. The results of the experiments conducted demonstrate that MS-CAE achieves high PSNR, improved compression ratios, and efficient data transmission, making it a more suitable solution for low-power, resource-constrained sensor networks.

3. Objectives

- To optimise image compression for efficient storage and transmission, we plan to implement and compare different compression techniques, selecting the one that provides the most effective image compression without compromising image quality.
- To enhance image security with encryption for secure retrieval, we intend to integrate encryption keys that link to the images stored securely in our database, allowing for quick and safe retrieval of the original pictures.
- To evaluate the performance of ANN-Based compression compared to traditional methods, we aim to compare and analyse how ANN-Based compression performs in terms of PCA (Principal Component Analysis) and K-Means in terms of file size reduction and feature retention.

4. Methodology

We've made significant strides in image compression. However, many systems still encounter issues such as rigid encoding methods, the loss of important details due to blurriness, problems with accurately reconstructing images, and substantial computational costs. This leads to blurring, blocking artefacts, and colour distortions when we crank up the compression. While models like Unicorn are highly efficient, they tend to require a significant amount of computational power, which isn't ideal for real-time applications or low-power devices. That's why most image compression tools stick with one tried-and-true model. Additionally, many deep learning techniques require separate models for different image types and often struggle with complex textures. Although transformer-based and diffusion-based methods are quite powerful, they can consume a significant amount of memory and processing time, which limits their use on mobile devices.

In our research, we've developed a Multi-Stage Convolutional Autoencoder with Adaptive Quantization and Secure Hash-Based Encryption (MSCAE-AQ-SHE), which is optimised through gradient-based learning. Unlike the conventional methods that rely on fixed transformations, our model learns feature representations dynamically through an iterative training process. By applying backpropagation to minimise reconstruction loss, we enable the model to improve itself through fine-tuning of its encoding and decoding steps, resulting in higher compression efficiency while maintaining visual quality. This indicates that our model can learn and adapt to various types of images and their structures, freeing it from conventional methods.

This project not only utilises a backpropagation-based neural network for image compression but also integrates an encryption-based retrieval system that enhances security in storage and facilitates a more accurate image reconstruction. Once the image uploaded for compression is compressed using a neural network, the system encrypts the compressed data. It assigns a unique encryption key to it before storing it in the database. When an image needs to be recovered, this unique key is used to decrypt the data, allowing for the accurate reconstruction of the original image. This entire process ensures data protection of the image during storage and transmission, preventing unauthorised access. It also helps in maintaining high compression efficiency.

Using lightweight cryptographic techniques in the encryption process, rather than traditional methods, helps preserve security. When compared to traditional compression methods, our approach provides an additional layer of security without compromising speed or image quality. This feature is particularly useful in medical imaging, handling confidential documents, and managing sensitive multimedia content, which prioritises both compression and security. A key advantage of our method is its ability to retain key details during the compression process. While many autoencoder-based approaches tend to produce blurry images, our artificial neural network is specifically designed to preserve the core features of an image. As a result, even at lower bitrates, the reconstructed images maintain sharp textures and fine details. By adjusting the compression settings based on feedback from backpropagation, the model removes unnecessary data while maintaining the structural integrity of the image.

This creates a balance between saving storage space and maintaining image quality, making it suitable for both high-resolution and real-time applications. To boost stability and consistency in reconstructions, we've incorporated adaptive weight updates and loss functions that specifically target reducing compression artefacts. Unlike some models that can sometimes introduce unwanted visual effects, our model is specifically designed to ensure uniform quality across all image types. It’s also lightweight and efficient, making it a great fit for use on mobile devices, IoT systems, and in settings where power is limited. To distinguish between image compression methods, we experimented with K-Nearest Neighbours (KNN), Principal Component Analysis (PCA), and Backpropagation-based neural networks. These methods vary in the way they compress image size. PCA chose the most significant major components to compress the image data. It also lost high-frequency information, making it less ideal for reconstructing intricate textures and sharp edges in images, despite maintaining basic features and achieving dimensionality reduction.

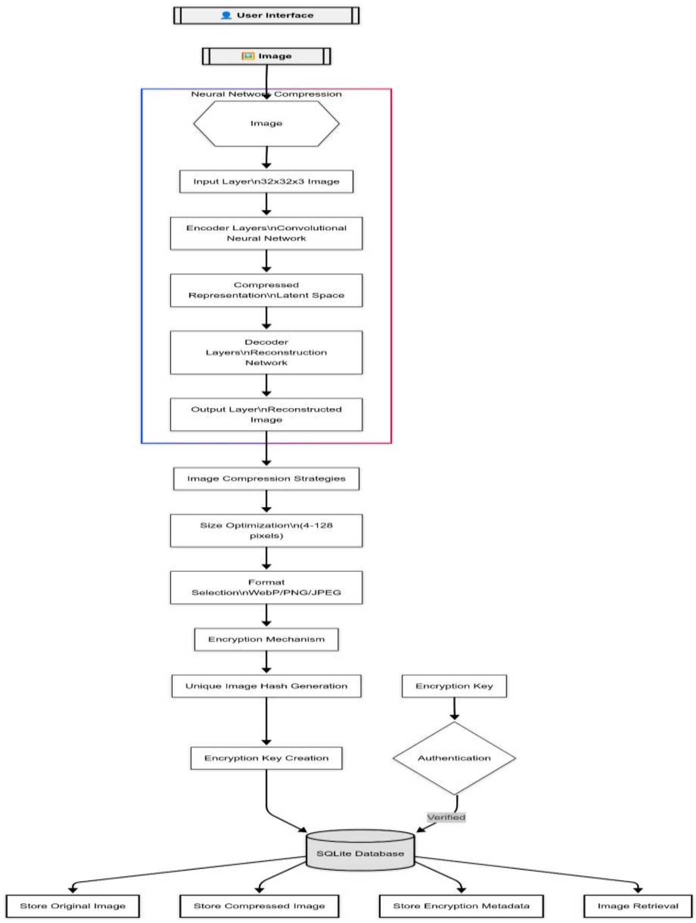


Figure 1: Architecture diagram of the proposed model

In image compression ratios and integrity, backpropagation neural networks outperformed PCA and KNN. By compressing images into the latent space and optimizing weight updates, the neural network preserved visual features and reduced non-essential information. Backpropagation was the most effective compression performance and also maintained visual quality across several image datasets, so we ultimately selected it for our image compression system. Through learning challenging feature representations, which contrasts with KNN's computational cost and PCA's limited capabilities, backpropagation optimizes compression. Particularly in the case of high-resolution images, where minute details must be maintained even under high compression ratios, such flexibility enables it to perform better than standard approaches. Below we have the architecture diagram of our proposed system (Figure 1). This compression system combines advanced machine learning, data optimisation, and secure storage technologies. The base concept of the system is to practically reinterpret traditional image compression methods from the perspective of artificial neural network techniques: an autoencoder trains itself to represent and reconstruct images with minimal loss of information efficiently. An image from this system's perspective begins with its compression work from the moment a user interacts with the Streamlit web interface, which serves as a medium and readily accessible entry point for users wishing to compress and secure their visual data.

The next is the state of image processing. Preprocess performs the task of uploading images to a system during a critical phase and transforms them. The system standardises the input by reducing the images to a size of 32 x 32 pixels. This is a highly calculated dimension because it utilises computational resources, yet it has still achieved significant feature retention. This also plays a crucial role in developing the neural network, which creates similar compressed representations of various image types. The input layer thus becomes the entrance to the transformed data: the image will be converted into a single, normalised numerical representation, allowing it to be processed by the subsequent neural network layers. The encoder layers of a convolutional neural network essentially consist of the heart or the basic core of any compression process. This set of layers operates continuously to extract and compress spatial features through various convolution and pooling operations. Each layer attenuates the dimensionality of the input, allowing it to separate increasingly abstract and meaningful details about the image. The network learns to identify and retain the most significant and relevant visual information, generating a compressed representation in a latent space. Thus, this approach shifts from classical compression methods to learned, customizable compression methods, which can counter conventional fixed solutions.

Having a similar architecture to an encoder, all types of decoder layers aim to reconstruct the image from the compressed version. The reconstruction procedure is conducted under a mean square error loss function that trains the entire network to minimise the difference between the original and reconstructed images. The autoencoder learns to compactly and accurately represent the input for compression via intelligent feature extraction and reconstruction. This unique process also allows for dynamic compression levels, where users can select compression from anywhere between 4 and 128 pixels, providing flexibility never before available in the history of image size reduction. In addition to compression, the system can employ highly advanced image optimisation methods.

The system then measures the size of the compressed images. It dynamically determines whether it will be best saved in WebP, PNG, or JPEG format for maximum efficiency in terms of size and quality. Adaptive format selection ensures that every compressed image is optimally stored, balancing file size, visual quality, and compatibility. Multiple factors, such as image complexity, colour depth, and desired compression level, are considered by the system when making format decisions. The architecture prioritises security and incorporates an encryption mechanism that is robust enough to support the compression process. Each image undergoes a separate process to generate a unique hash, which forms a cryptographic fingerprint necessary for generating encryption keys. This ensures that every compressed image has a unique, secure identifier that is beyond the reach of unauthorised access. A retrieval mechanism is set, and here, the encryption key safeguards any visual stored data.

A SQLite database is used, serving as the nucleus for the ecosystem of compressed images. Not only does it keep copies of the original and compressed image, but it also retains all the complete metadata that maps all the aspects involved in the image compression process. This means that both the original image and a compressed version of it, as well as other metadata, including possible encryption and contextual information, can be stored in the database. This schema plays a crucial role in efficiently storing, retrieving, and managing compressed images while tracking each compression activity. The decompression process details security while maintaining a user-friendly interface. In this process, users can access their original images by providing an encryption key generated during the compression process. An authentication process verifies this key against the user, restricting image extraction to rightful holders. This is demonstrated by the system's approach to data privacy and controlled access, which enables users to perform image compression safely and securely with their data.

The learning process of a neural network differs when it is trained on a unique dataset, such as CIFAR-10. The mode of application to the network aids its training in generalising the specific features per category in the images. The technique of backpropagation enables the network to improve over time, and the compression scheme it has established allows learning to retrieve essential pieces of image information while reconstructing it, taking dramatically less effort. Therefore, it becomes more efficient with each epoch in the training cycle, making it more feasible for the network to produce even better, more compact, and more representative images of the input data.

This process provides a collaborative service between the application and good technical support to help developers reach their creative and practical goals. From a strict technical viewpoint, the latest architectures are based on TensorFlow and Keras for conversion, Streamlit for display, and SQLite for a lean and secure database system, aligning with the system's modularisation. All future improvements will include support for additional image formats, integration of other more complex compression techniques, or even support for connecting with other image processing organisations. The image compression process refers not only to a technical application but rather to the subtle alteration of digital images. The architecture, equipped with neural network intelligence, adaptive compression schemes, and robust security mechanisms, thus becomes a valuable tool in the hands of users to minimise image file size without compromising visual fidelity and protecting data. The system thus highlights the prospects that machine learning offers over the evolving landscape of traditional compression techniques, providing a comprehensive view of tomorrow's intelligent, secure, and efficient image-processing technology.

4.1. Mathematical Representation

4.1.1. Autoencoder Model Representation

An autoencoder consists of two primary components:

- **Encoder E:** Compresses input image x into a lower-dimensional latent representation z.
- **Decoder D:** Reconstructs the image $\hat{x}$ From z.

Mathematically, the encoding and decoding functions are:

$$z = E(x; \theta_E)$$
$$\hat{x} = D(z;\ \theta_D)$$

Where:

- $x \in R^{H \times W \times C}$ is the original image with height H, width W, and channels C.
- $z \in R^{d}$ it is the compressed representation of dimensionality d.
- θ_E and θ_D the learnable parameters of the encoder and decoder, respectively.

4.1.2. Loss Function: Mean Squared Error (MSE)

The reconstruction quality of the autoencoder is evaluated using the Mean Squared Error (MSE) loss function, which is defined as:

$$\mathbf{MSE}(x, \hat{x}) = \frac{1}{N} \sum_{i=1}^{N} (x_i - \hat{x}_i)^2$$

Where:

- N is the total number of pixels in the image.
- x_i and $\widehat{x_\iota}$ Are the original and reconstructed pixel values, respectively

4.1.3. Convolutional Encoding and Decoding

Each layer in the encoder and decoder applies a convolutional transformation:

$$h^{(l)} = f(W^{(l)} * h^{(l-1)} + b^{(l)})$$

Where:

- $W^{(l)}$ is the weight matrix of the $l\text{-}th$ convolutional layer
- $*$ denotes the convolution operation
- $b^{(l)}$ is the bias term.

- f is the activation function. (ReLU or Sigmoid)

The decoder reverses this process using transposed convolutions or upsampling operations.

4.1.4. Compression and Downsampling

The encoder applies max-pooling to reduce spatial dimensions:

$$h_{pool}^{(l)} = \max(h_{i,j}^{(l)})$$

where $h_{i,j}^{(l)}$ Represents the activations in the receptive field, and the pooling operation selects the maximum value in each region.

4.1.5. Upsampling in Decoding

To reconstruct the image, the decoder applies upsampling:

$$h_{up}^{(l)} = \text{repeat}(h^{(l)}, s)$$

Where s is the upsampling factor.

4.1.6. Image Compression Performance

Given a compressed representation z of dimensionality d and an original image of size $H \times W \times C$, the compression ratio is given by:

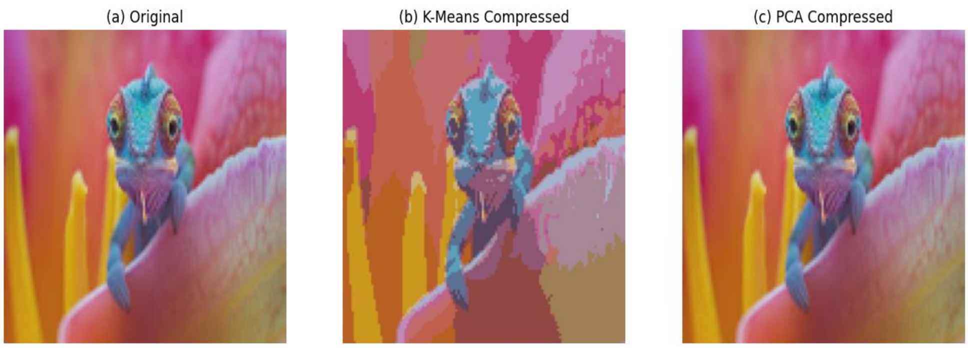
$$\text{Compression Ratio} = \frac{H \times W \times C}{d}$$

This quantifies the reduction in data size achieved by the autoencoder.

5. Results and Discussion

5.1. Data Preprocessing and Training

The proposed model was trained and tested using the CIFAR-10 dataset, which comprises 60,000 RGB images categorised into 10 distinct classes. Each image has a resolution of 32 × 32 pixels. A subset of 1,000 images was used to train the convolutional autoencoder, with normalisation applied to scale the pixel values to the range [0, 1]. The images were encoded as tensors with the shape (32, 32, 3) to maintain compatibility with the model. Our experiment demonstrated that a batch size of 32 provided optimal results during training. The dataset was split into a 90/10 ratio for training and validation.



(d) Backpropagation compressed

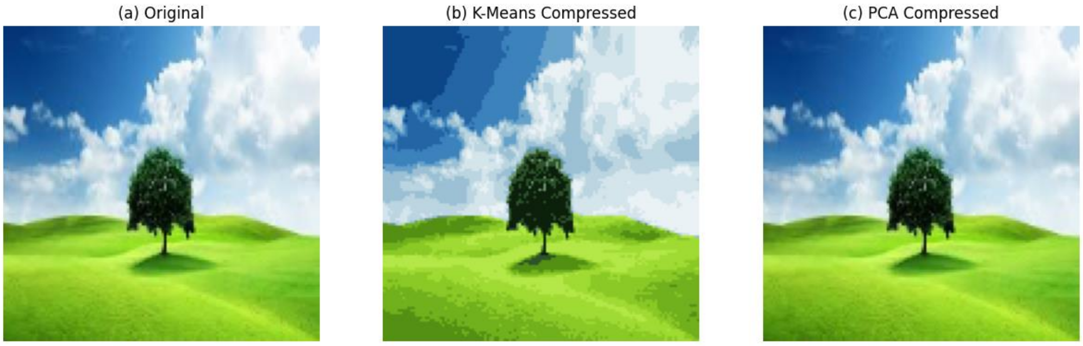
Figure 2: Comparison of image compression techniques (a-d)

The Adam optimiser was used with an initial learning rate of 0.001. To enhance convergence and mitigate overfitting, a "reduce learning rate on plateau" scheduler was implemented, reducing the learning rate by a factor of 0.1 whenever the validation loss remained unchanged for several consecutive epochs. The model was trained for 10 epochs, and the dataset was normalised to improve stability. Additionally, early stopping was employed to prevent excessive training when no further improvement was observed. For evaluation, user-uploaded images were resized to 32 × 32 pixels to match the CIFAR-10 resolution and then tested. To assess the model's adaptability to real-world scenarios, images of varying sizes and formats—including JPEG, PNG, and WebP—were used for optimised compression analysis.

Table 1: Compression performance of various algorithms (size in kilobytes)

Algorithm	Original Size	Reduced Size
K-Means	7.6 KB	3.95 KB
PCA	7.6 KB	3.69 KB
Backpropagation	7.6 KB	2.88 KB

According to the findings, ANN had the highest compression efficiency, lowering file sizes more than K-Means and PCA while preserving image quality. ANN achieved the highest compression ratio in Figure 2, compressing the image to 2.88 KB, compared to 3.95 KB using K-Means and 3.69 KB using PCA, as shown in Table 1.



(d) Backpropagation compressed

Figure 3: Comparison of image compression techniques (a-d)

Table 2: Compression performance of various algorithms (size in kilobytes)

Algorithm	Original Size	Reduced Size
K-Means	6.41 kb	4.34 kb
PCA	6.41 kb	3.95 kb
Backpropagation	6.41 kb	3.1 kb

The outcomes of our experiments show that backpropagation-based image compression (ANN) is more effective than K-Means and PCA. Compression ratio and image dimensions were the two main focuses of our evaluation. Similarly, ANN outperformed K-Means (4.34 KB) and PCA (3.95 KB) in Figure 3, compressing the image to 3.1 KB, as shown in Table 2. These findings demonstrate that ANN-based compression reduces file sizes while maintaining important image characteristics, which makes it a more effective method for transmission and storage (Table 3).

Table 3: Comparison of average file size reduction across compression techniques

Method	Average Size Reduction
PCA	45%
K-Means	40.25%
Backpropagation	56.9%

One important finding is that, although both performed worse than ANN, PCA consistently outperformed K-Means in compression ratio. PCA operates by discarding less relevant information and choosing principal components that capture the image's most important features. Compared to K-Means, which groups similar pixels but does not effectively optimise feature retention, this technique offers superior compression. PCA still loses fine image details, though, and this can result in reconstruction artefacts, particularly in images with intricate textures. By optimising feature representation through backpropagation, an ANN, on the other hand, learns to compress images by training on patterns and textures, thereby improving image retention while achieving a higher compression rate, as shown in Figure 4.

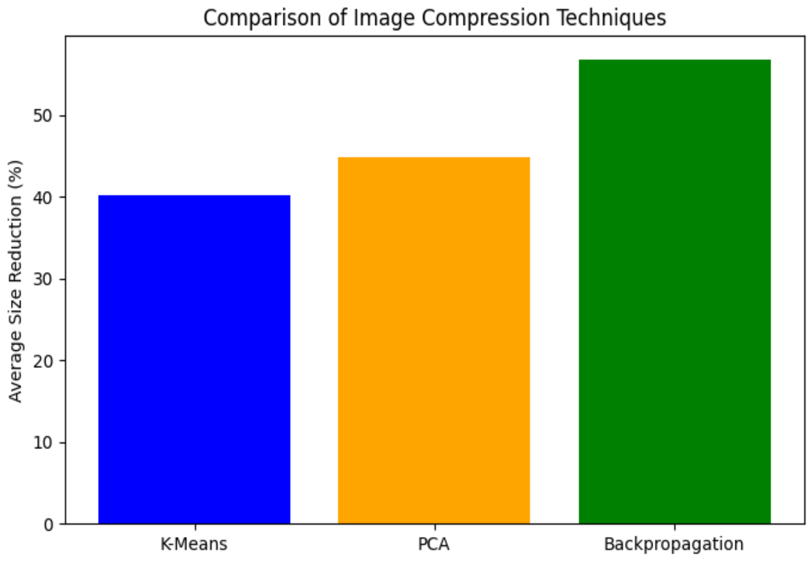


Figure 4: Comparison of average compression efficiency across techniques

One important feature of compression techniques is highlighted by the uniformity in image dimensions (128×128 pixels) across all techniques: file size reduction is not always correlated with dimensionality reduction. The ability of each method to eliminate redundant data varied greatly, even though they all maintained the same image resolution. Due to their reliance on predetermined statistical transformations, K-Means and PCA produced larger compressed images compared to ANN. However, by adaptively learning efficient compression strategies, the backpropagation-based neural network preserved high-level features while eliminating unnecessary data, as illustrated in Figure 5.

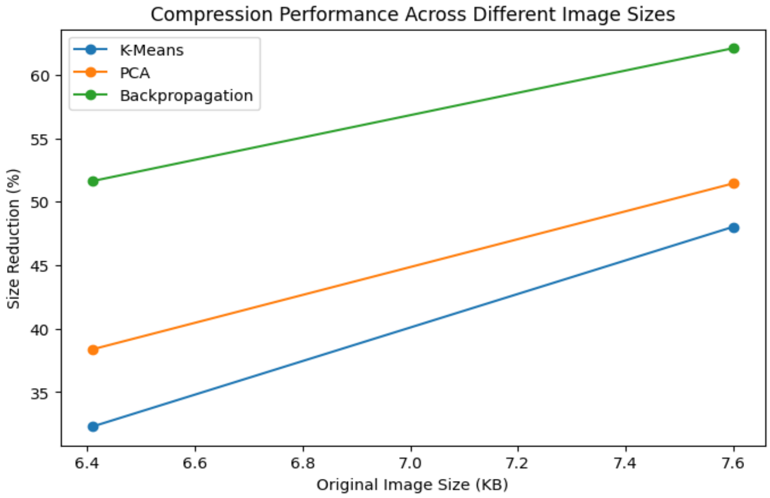


Figure 5: Variation in compression efficiency concerning original image size

The results indicate that backpropagation-based ANN compression outperforms PCA and K-Means in terms of compression ratio and feature retention. ANN efficiently minimises image size while maintaining visual quality by utilising iterative learning and adaptive weight updates. ANN is a scalable and effective image compression solution due to its capacity to self-optimise through training, particularly in situations where transmission speed and storage efficiency are crucial. These findings support the growing use of deep learning-based image compression as an alternative to traditional methods, offering a more adaptable and efficient approach for modern image processing applications.

6. Conclusion

This study presents a neural network-driven image compression system trained on a subset of the CIFAR-10 dataset, which combines a convolutional autoencoder with adaptive quantisation and secure hash-based encryption. While maintaining a Peak Signal-to-Noise Ratio (PSNR) of over 30 dB, the proposed Multi-Stage Convolutional Autoencoder with Adaptive Quantization and Secure Hash-Based Encryption (MSCAE-AQ-SHE) reduces storage capacity by up to 85%. Its average compression ratio is 4.5 to 10 times that of a standard engine. A mean squared error average of under 0.005 indicates that the suggested model minimises perceptual loss. The model's capacity to outperform more conventional compression methods, such as K-means clustering and PCA, which frequently compromise a balance between compression efficiency and image integrity, is one of its key benefits. Experimental findings show that while K-Means causes significant pixelation problems, PCA (3x to 8x) loses fine texture features. K-Means delivers modest compression (3x to 6x). Conversely, as it learns high-dimensional representations and enables adaptive compression with minimal information loss, the suggested autoencoder-based method is superior in both compression ratio and reconstruction quality.

Combining SHA-256 hashing with MD5-based encryption key creation further enhances security and integrity. This ensures that only their unique encryption keys can access the compressed photos. The adaptive compression technique dynamically adjusts by maximising the balance between storage efficiency and image quality, depending on the image complexity. Future studies will focus on investigating transformer-based compression models, utilising lightweight versions for edge devices, and combining adaptive bit allocation techniques to enhance efficiency in practical scenarios. The proposed approach establishes a framework for safe and efficient deep learning-based image compression, with broad applications in cloud storage, multimedia transmission, and medical imaging.

Acknowledgment: The authors express their sincere gratitude to SRM Institute of Science and Technology for providing the necessary support and academic environment. We appreciate the valuable resources and guidance extended during this research. This work reflects the continuous encouragement and infrastructure provided by the institution.

Data Availability Statement: The dataset used in this study is available from the corresponding author upon reasonable request.

Funding Statement: This research was conducted independently without any external financial assistance or institutional funding.

Conflicts of Interest Statement: The authors declare no competing interests. All sources have been duly acknowledged, and the work is original and properly referenced.

Ethics and Consent Statement: The study strictly adhered to ethical standards, with informed consent obtained from all individuals involved.

References

1. Z. Cheng, H. Sun, M. Takeuchi, and J. Katto, "Deep Convolutional AutoEncoder-based Lossy Image Compression," *in 2018 Picture Coding Symposium (PCS),* California, United States of America, 2018.
2. S. Liu, Y. Li, J. Yin, and Z. Wang, "A Fast Fractal-Based Compression for MRI Images," *in 2019 IEEE International Conference on Imaging Systems and Techniques (IST),* Abu Dhabi, United Arab Emirates, 2019.
3. S. Ma, X. Zhang, C. Jia, Z. Zhao, S. Wang, and S. Wang, "Image and video compression with neural networks: A review," *IEEE Trans. Circuits Syst. Video Technol.,* vol. 30, no. 6, pp. 1683–1698, 2020.
4. S. Iwai, T. Miyazaki, Y. Sugaya, and S. Omachi, "Fidelity-controllable extreme image compression with generative adversarial networks," *in 2020 25th International Conference on Pattern Recognition (ICPR),* Milan, Italy, 2021.
5. D. Tellez, G. Litjens, J. van der Laak, and F. Ciompi, "Neural Image Compression for gigapixel histopathology image analysis," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 43, no. 2, pp. 567–578, 2021.
6. Y. Hu, W. Yang, Z. Ma, and J. Liu, "Learning end-to-end lossy image compression: A benchmark," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 44, no. 8, pp. 4194–4211, 2022.
7. Z. Guo, Z. Zhang, R. Feng, and Z. Chen, "Causal contextual prediction for learned image compression," *IEEE Trans. Circuits Syst. Video Technol.,* vol. 32, no. 4, pp. 2329–2341, 2022.
8. G. Lu, X. Ge, T. Zhong, Q. Hu, and J. Geng, "Preprocessing Enhanced Image Compression for Machine Vision," *IEEE Trans. Circuits Syst. Video Technol.,* vol. 34, no. 12, pp. 13556–13568, 2024.
9. Y. Liang, M. Zhao, X. Liu, J. Jiang, G. Lu, and T. Jia, "An adaptive image compression algorithm based on joint clustering algorithm and deep learning," *IET Image Process.,* vol. 18, no. 3, pp. 829–837, 2024.
10. X. Bao, C. Ye, L. Han, and X. Xu, "Image compression for wireless sensor network: A model segmentation-based compressive autoencoder," *Wirel. Commun. Mob. Comput.,* vol. 2023, no. 10, pp. 1–12, 2023.
11. A. K. Naveen, S. Thunga, A. Murki, M. Kalale, and S. Anil, "Autoencoded image compression for secure and fast transmission," *in Proceedings of the 2024 IEEE International Conference on Computer Vision and Machine Intelligence (CVMI),* Prayagraj, India, 2024.