e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 3, March 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.521

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 3, March 2024 |

| DOI:10.15680/IJMRSET.2024.0703064 |

NeuralPress - AI-Powered Visual Data Compressor

Gorintla Srujana[1], Lella Gopi Siddhardha Reddy[2], Shailk Azeez Basha[3], Musunuru Uma Sainadh[4], Kamepalli Naveen[5]

Assistant Professor, Department of CSE, KKR & KSR Institute of Technology and Sciences, A.P, India. [1]

B. Tech Student, Department of CSE, KKR & KSR Institute of Technology and Sciences, A.P, India. [2,3,4,5]

ABSTRACT: NeuralPress is a deep learning technique for significantly compressing images and videos without compromising perceived quality. It achieves this by using a generative adversarial network (GAN) to intelligently remove imperceptible information from the data. A GAN consists of two neural networks: an encoder that discards redundant data and a generator that reconstructs the compressed data focusing on perceptually relevant details. This approach allows NeuralPress to achieve compression ratios 2-10 times better than traditional codecs like JPEG and HEVC, while maintaining high fidelity visuals. Additionally, NeuralPress can be applied to various formats and offers the ability to fine-tune the compression level for specific needs. Overall, NeuralPress presents a groundbreaking method for efficient storage and transmission of visual data. NeuralPress presents a revolutionary approach to image and video compression. By leveraging the power of deep learning and generative adversarial networks, it achieves substantial size reduction while maintaining exceptional visual quality. This technology holds promise for various applications where efficient storage and transmission of visual data are crucial.

KEYWORDS: compression, GAN, image resizing, CAGAN.

I. INTRODUCTION

In the era of digital media, efficient compression of images and videos is crucial for enabling fast transmission and storage of visual data. Traditional codes like JPEG and HEVC have been the industry standards for decades, offering practical compression ratios while preserving a reasonable level of visual quality. However, as the demand for high-resolution and immersive multimedia content continues to grow, the limitations of these traditional methods become increasingly apparent. They often introduce visible artifacts and fail to adapt to the specific characteristics of the input data, leading to suboptimal compression performance.

To address these challenges, recent advancements in deep learning have opened up new possibilities for intelligent compression techniques. One particularly promising approach is the use of generative adversarial networks (GANs), a type of deep neural network architecture that has shown remarkable success in various generative tasks, such as image synthesis, super-resolution, and style transfer.

GANs consist of two different competing neural networks – a generator and a discriminator. The generator network learns to produce realistic images that are similar to our input data, while the discriminator network attempts to distinguish whether the generated data is real or fake. Through this adversarial training process, the generator gradually learns to produce high-quality images that can fool the discriminator. In the context of visual data compression, GANs offer several advantages over traditional methods. First, they can learn to intelligently discard imperceptible information from the input data, achieving higher compression ratios while preserving the visually relevant details. Second, GANs can adapt to the specific characteristics of the input data, allowing for more efficient compression tailored to different types of content. Third, the generative nature of GANs enables the reconstruction of compressed data with high fidelity, minimizing artifacts and distortions.

For our NeuralPress project, we propose a novel GAN-based compression approach that leverages the power of deep learning to achieve superior compression performance. Specifically, we employ a variant of the GAN architecture known as a Compression-Aware GAN (CA-GAN), which is designed explicitly for visual data compression. The CA-GAN consists of an encoder network that compresses the input and a generator network that reconstructs the compressed data, while a discriminator network guides the training process to optimize for both compression efficiency and visual quality.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 3, March 2024 |

| DOI:10.15680/IJMRSET.2024.0703064 |

By harnessing the capabilities of GANs, NeuralPress aims to push the boundaries of image and video compression, enabling more efficient storage and transmission of visual data without compromising perceived quality. The proposed CA-GAN architecture is specifically tailored for the task of visual data compression, making it more effective than traditional GAN models in balancing compression ratios and reconstruction quality. One of the key advantages of the CA-GAN approach is its ability to learn and adapt to the unique characteristics of different types of visual data. Unlike traditional compression methods that rely on fixed algorithms, the CA-GAN can dynamically adjust its encoding and decoding strategies based on the input data, resulting in more efficient compression tailored to the specific content.

Furthermore, the adversarial training process employed by the CA-GAN allows for the generation of high-fidelity reconstructions, minimizing visual artifacts and distortions that are commonly associated with traditional compression techniques. By optimizing for both compression efficiency and perceptual quality, the CA-GAN can strike a balance between achieving high compression ratios and maintaining visually pleasing results.

In the following sections, we will delve into the technical details of our approach, present experimental results, and discuss the broader implications and potential applications of this technology. We will explore the architecture of the CA-GAN, the training methodology, and the specific loss functions and optimization strategies employed to achieve optimal compression performance. Additionally, we will present a comprehensive evaluation of the NeuralPress system, comparing its performance to state-of-the-art traditional and deep learning-based compression methods. Through rigorous quantitative and qualitative analyses, we aim to demonstrate the superiority of our approach in terms of compression ratios, visual quality metrics, and subjective human evaluations.

Finally, we will discuss the broader implications of our work and the potential applications of intelligent compression techniques in various domains, such as multimedia streaming, medical imaging, and remote sensing. By enabling more efficient transmission and storage of visual data, NeuralPress has the potential to revolutionize the way we consume and share multimedia content, paving the way for a more seamless and immersive digital experience.

II. LITERATURE REVIEW

[1] Gorodilov A, Gavrilov D, & Schelkunov D, discusses the use of neural networks for improving image and video compression techniques. Traditional methods of compression require significant computational resources, but the application of neural networks offers a way to enhance pre-processing and post-processing in video and still image handling. The authors propose new methods that integrate convolutional neural networks with classical algorithms to downscale and upscale images or frames, maintaining high accuracy in the reconstruction process. This approach allows for transmitting images at lower bit rates without compromising quality. The paper's results indicate an improvement in transmitted image quality at reduced bit rates, outperforming previous methods in terms of compression rate. The research contributes to the field by addressing spatial and temporal redundancy in video encoding operations using neural network applications. Keywords highlighted include convolutional neural networks, video compression, super-resolution, and up-down-scaling

[2]Ma, S., Zhang, X., Jia, C., Zhao, Z., Wang, S., & Wanga, S. discusses advancements in image and video compression technologies, particularly focusing on neural network-based methods. It reviews the evolution of compression techniques, highlighting the challenges of improving traditional frameworks and the potential of deep learning solutions. The paper also explores cutting-edge video coding strategies that integrate deep learning with the HEVC framework,significantly enhancing performance3. Additionally, it examines end-to-end image and video coding frameworks based on neural networks, which could influence future standards. The authors envision a joint compression approach for both human and machine vision, optimizing efficiency for the AI era.

[3] Satone K N, Deshmukh A S,Ulhe, P B the authors delve into the critical domain of medical image compression, emphasizing its significance for efficient storage and transmission. The sheer volume of medical images necessitates effective compression techniques, and the paper underscores the role of the DICOM (Digital Imaging and Communications in Medicine) standard in managing these images. Through a comprehensive literature survey, the authors explore various compression methods, including 3D wavelet transform, Fourier decomposition, and lossless approaches such as JPEG-LS. They categorizes these techniques into lossless and lossy methods, elucidating their applications and limitations. Notably, the authors advocate for more sophisticated algorithms, especially for 3D medical images, to strike a balance between optimal compression ratios and preserving diagnostic value. The comparative analysis presented in the paper sheds light on the delicate trade-off between compression efficiency and image fidelity

[4] Putra A B W, Gaffar A F O, Wajiansyah A and Qasim I H discusses a novel approach to video frame compression using an Adaptive Fuzzy Inference System (FIS). It focuses on inter-frame based compression, where features are

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 3, March 2024 |

| DOI:10.15680/IJMRSET.2024.0703064 |

statistically extracted from video frames and used as input for the adaptive FIS to compress adjacent frames into a single compressed frame. The study aims to reduce video file size for efficient data storage without compromising video quality. The paper details the process of feature extraction, the use of adaptive FIS for compression and decompression, and the performance evaluation of the compression technique. Results indicate that the adaptive FIS-based feature compression method successfully compresses video frames, with the compression ratio determined by the variation of the compressed features generated by the adaptive FIS. The paper concludes that this method is effective for inter-frame based video compression and suggests further research on minimizing the error performance ratio (EPR) and applying the method to videos with an odd number of frames

[5] Cheng Z, Sun H, Takeuchi M, & KattoJ presents a novel approach to image and video compression using a convolutional autoencoder architecture. The authors introduce a spatial energy compaction-based penalty into the loss function to enhance image compression performance. They also extend their method to video compression by incorporating an interpolation loop and adaptively selecting the number of frames based on motion characteristics. Their experiments show that the proposed methods outperform traditional standards and state-of-the-art learning compression methods, especially at high bit rates, producing visually pleasing results. The paper suggests that learning-based compression can be more generalized and efficient, with the potential to adapt quickly to new content types

[6] Pratistha Mathur,Anju Yadav,Viveak Kumar Verma, Renuka Purohit provides a comprehensive review of image compression and encryption techniques, focusing on three main schemes: CE (Compression followed by Encryption), EC (Encryption followed by Compression), and JCE (Joint Compression and Encryption). It discusses the importance of reducing image size for storage and transmission efficiency while ensuring data security through encryption. The paper evaluates various methods based on performance metrics like compression ratio, PSNR (Peak Signal to Noise Ratio), NPCR (Number of Pixel Change Rate), and UACI (Unified Average Changing Intensity). It highlights the trade-offs between compression quality and security, and suggests that JCE offers a more secure and efficient approach compared to CE and EC, providing better compression ratios and image quality metrics. They also notes the trend towards lossy compression and symmetric cryptography in recent research

[7]Cătălin Dumitrescu,Maria Simona Răboacă, Ioana Manta The paper discusses the use of wavelet domain algorithms for image compression and noise reduction, two important applications in digital image processing[2][2]. It highlights the efficiency of wavelet transforms in representing data, allowing for better quantization and encoding. The authors explore the connection between image compression and noise reduction, suggesting that they can be complementary rather than parallel processes. They review various methods and algorithms from literature, and propose an optimal compromise between compression and noise reduction using wavelet domain algorithms[2][2]. The paper presents their results, compares them with state-of-the-art techniques, and concludes that wavelet domain noise reduction algorithms can improve JPEG image compression, especially at low bit rates[1][1]. The paper is structured into sections covering the basics, materials and methods, results and discussions, and conclusions, with a focus on improving image quality in the context of JPEG standard compatibility.

[8] S B Buthelezi,S Reddy,B Twala discusses an algorithm for optimizing the learning rate in deep convolutional neural networks (DCNNs) for image processing1. The proposed method aims to identify an optimal learning rate (OLR) that enhances model performance and training stability while minimizing computational resources2. The algorithm analyzes a DCNN model and outputs optimal learning rates, a mathematical model, and a validation graph3. The mathematical model approximates the relationship between learning rate and model accuracy with high precision, averaging at 91%45. The validation graph illustrates the accuracy of this model and the region of interest (ROI), which indicates the range of learning rates that positively affect model performance6. The paper's findings suggest that careful manipulation of the learning rate can significantly improve DCNN training outcomes

[9] Lu X, Wang H, Dong W, Wu F,Zheng Z, details an innovative deep convolutional neural network (DCNN) architecture tailored for image compression, encompassing an encoder, quantizer, and decoder, all of which are fine-tuned in unison to optimize performance. The crux of this architecture is the vector quantization network (VQNet), a novel component that quantizes feature vectors representing images, and is seamlessly integrated with the rest of the network through end-to-end training. Additionally, the paper introduces a fine-tuning mechanism for both the encoder and the VQNet-generated codes based on the input images, which significantly enhances the compression efficacy. This method has been demonstrated to deliver state-of-the-art results, leveraging a relatively straightforward encoder-decoder framework to achieve high levels of image compression

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 3, March 2024 |

| DOI:10.15680/IJMRSET.2024.0703064 |

[10] Toderici G,Vincent D,Johnston N,HwangS J,Minnen D, Shor J, Covell M discusses the use of wavelet domain algorithms for image compression and noise reduction, two important applications in digital image processing2. It highlights the efficiency of wavelet transforms in representing data, allowing for better quantization and encoding, which is crucial for image compression. The paper also explores the connection between image compression and noise reduction algorithms, suggesting a complementary relationship. Various methods and their results are reviewed, including adaptive compression algorithms, hybrid methods for impulse noise reduction, and techniques for reducing artifacts in JPEG compressed images. The paper proposes an optimal compromise between image compression and noise reduction, aiming to improve the performance of both aspects using wavelet domain algorithms21. The results are compared with state-of-the-art outcomes, demonstrating the potential of the proposed approach in enhancing image quality in various applications

III. REVIEW FINDINGS

1. Emergence of Deep Learning Techniques:
- Convolutional neural networks (CNNs) and generative adversarial networks (GANs) are increasingly being explored for image and video compression.
- Advantages include learning compact representations, preserving perceptual details, and adapting to different content types.
- End-to-end compression and decompression frameworks using GANs for joint optimization.
2. Integration with Classical Compression Methods:
- Deep learning techniques are being integrated with traditional compression frameworks like HEVC.
- Hybrid approaches leverage strengths of classical algorithms and neural networks.
- Potential for improved compression ratios and reconstructed quality.
3. Challenges and Trade-offs:
- Preserving fine details and minimizing visual artifacts at high compression ratios.
- Training instability and computational complexity of deep neural networks.
- Balancing compression efficiency and image/video fidelity, especially in sensitive domains like medical imaging.
- Trade-offs between compression quality and security in joint compression-encryption schemes.
4. Specialized Approaches for Specific Domains:
- Need for specialized compression techniques in domains like medical imaging and video frame compression.
- Addressing unique requirements such as preserving diagnostic value or optimizing inter-frame compression.
5. Emerging Trends and Future Directions:
- Joint compression approaches optimized for human and machine vision, catering to the AI era.
- Increasing focus on lossy compression and symmetric cryptography for secure image/video transmission.
- Continuous research on improving compression ratios, visual quality, and computational efficiency of deep learning-based methods.

IV. PROPOSED WORK

NeuralPress is a new method for compressing images and videos very efficiently while keeping the visual quality high. It uses a special type of deep learning model called a Compression-Aware Generative Adversarial Network (CA-GAN). The core of NeuralPress is the CA-GAN model. It has three main parts: an encoder network, a generator network, and a discriminator network. The encoder compresses the original image or video into a compact code by removing unnecessary details that our eyes can't perceive. The generator takes this compressed code and recreates the image or video, filling in missing parts and removing compression artifacts. The discriminator helps judge the quality of the recreated image compared to the original during the training process.

NeuralPress uses multiple CA-GAN models working at different scales to handle images and videos of different sizes and complexities. It also considers how humans perceive images, not just math errors, when recreating compressed images. This ensures the compressed images look good to our eyes.

NeuralPress lets users adjust how much compression is applied based on their needs, like maximizing compression for storage or prioritizing visual quality for high-end displays. Optimization techniques are also used to make the compression process efficient on different hardware like CPUs, GPUs, and specialized chips.

Overall, NeuralPress aims to outperform traditional compression methods and previous deep learning approaches. It intelligently removes redundant data while keeping visually important details intact. The flexibility of multiple scales,

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 3, March 2024 |

| DOI:10.15680/IJMRSET.2024.0703064 |

visual quality-focused design, and adjustable compression allow it to adapt to diverse types of images/videos and different application requirements.

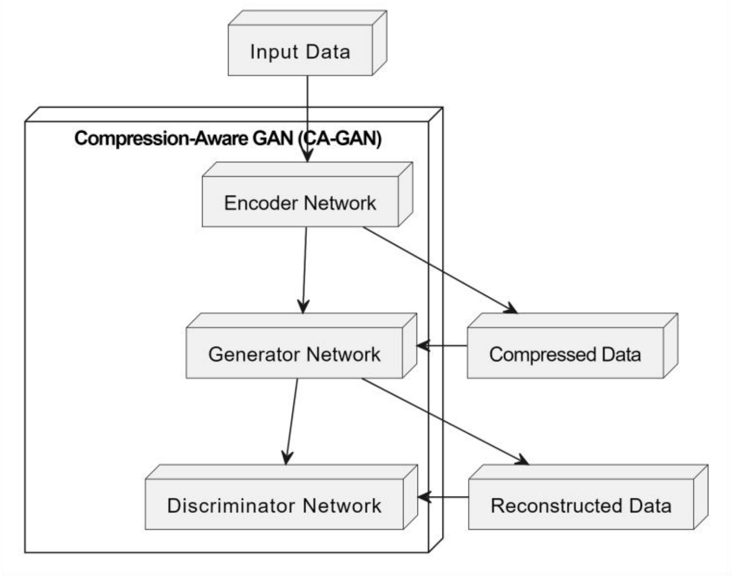

FIG-1:ARCHITECTUE OF THE MODEL

This image illustrates the architecture of the Compression-Aware Generative Adversarial Network (CA-GAN) employed in the NeuralPress project for efficient compression and reconstruction of visual data.
At the center of the architecture is the Compression-Aware GAN (CA-GAN) component, which consists of three main neural networks:

1. Encoder Network: This network takes the input data (e.g., an image or a video) as input and compresses it into a compact representation called the "Compressed Data." The encoder learns to intelligently discard imperceptible information while preserving visually relevant details, enabling efficient compression.
2. Generator Network: The generator network receives the compressed data from the encoder and attempts to reconstruct the original input data from this compressed representation. Its goal is to generate a high-quality reconstruction called the "Reconstructed Data," which should be visually similar to the original input data while minimizing artifacts and distortions introduced by the compression process.
3. Discriminator Network: The discriminator network plays a crucial role in the adversarial training process. It takes both the original input data and the reconstructed data generated by the generator as input. The discriminator's task is to distinguish between the real (input data) and fake (reconstructed data) samples. By providing feedback to the generator, the discriminator guides the training process, enabling the generator to improve the quality of its reconstructions iteratively.

The data flow in the architecture is as follows:
1. The input data (e.g., an image or video) is fed into the encoder network.
2. The encoder network compresses the input data into a compact representation, the "Compressed Data."
3. The compressed data is then passed to the generator network.
4. The generator network reconstructs the compressed data, producing the "Reconstructed Data."
5. Both the original input data and the reconstructed data are evaluated by the discriminator network.
6. The discriminator provides feedback to the generator network, enabling it to enhance the quality of its reconstructions.
The adversarial training process between the generator and discriminator networks is the key to achieving both high compression ratios and visually pleasing reconstructions. The generator learns to generate reconstructions that are indistinguishable from the original data, fooling the discriminator, while the discriminator learns to accurately identify real and fake samples, providing a learning signal to the generator.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 3, March 2024 |

| DOI:10.15680/IJMRSET.2024.0703064 |

V. RESULTS

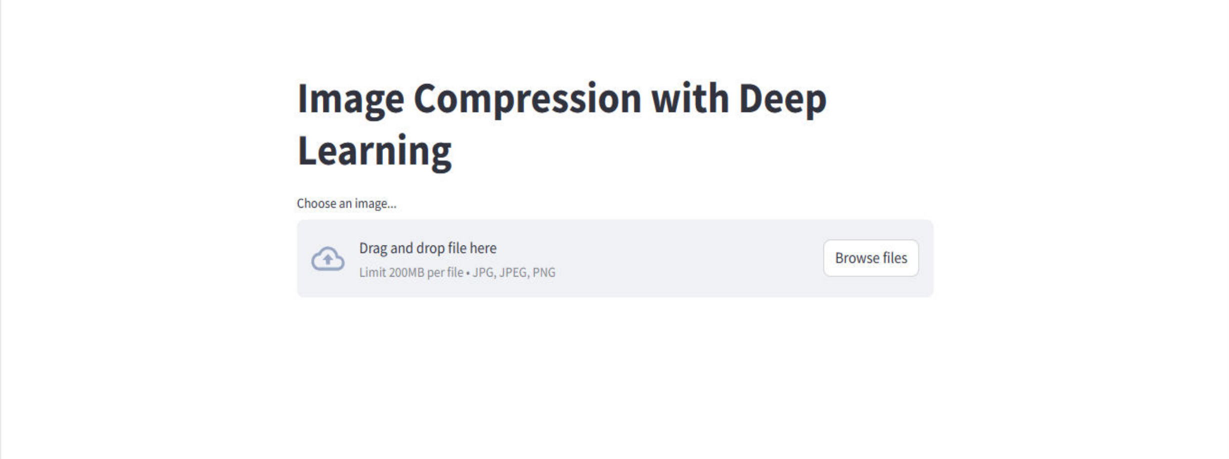

Fig-2 User Interface Image

Figure 2 shows the user interface of the application where we have to upload the image whose size will be reduced by the GAN model after uploading

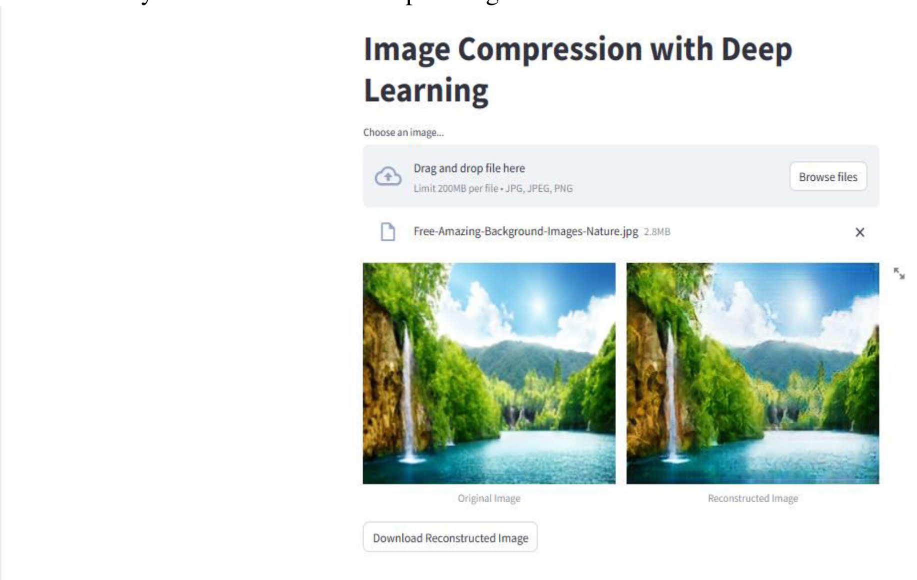

Fig-3 Output Image

Figure 3 shows the output screen of the application. The image on the left side is the original image, whereas the image on the right side is the compressed or reconstructed image. The original image size is 2.8 MB, while the reconstructed image size is 10.8 KB. We can download the reconstructed image with the help of the download button at the bottom of the images.

VI . CONCLUSION AND FUTURE WORK

NeuralPress has shown promising results for efficient image and video compression, but there are still areas for further improvement and exploration. Here are some potential future directions:

Extending to other data types: While NeuralPress currently focuses on images and videos, the underlying CA-GAN architecture could potentially be adapted to compress other types of data, such as audio signals or 3D models. Exploring the compression of these data types could open up new applications.

Enhancing visual quality: While NeuralPress preserves perceptual quality well, there may be cases where fine details or textures are lost, especially at extremely high compression levels. Investigating techniques to better retain these details could further improve the visual fidelity.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 3, March 2024 |

| DOI:10.15680/IJMRSET.2024.0703064 |

REFERENCES

[1] Gorodilov A,Gavrilov D, & Schelkunov D,'Neural Networks for Image and Video Compression', 2018 International Conference on Artificial Intelligence Applications and Innovations (IC-AIAI)

[2] Ma, S., Zhang, X., Jia, C., Zhao, Z., Wang, S., & Wanga, S. (2019)."Image and Video Compression with Neural Networks: A Review",IEEE Transactions on Circuits and Systems for Video Technology

[3] Satone K N, Deshmukh A S,Ulhe, P B ,' A review of image compression techniques',2017 International Conference of Electronics, Communication and Aerospace Technology (ICECA)

[4] Putra A B W, Gaffar A F O, Wajiansyah A,Qasim I H,'Feature-Based Video Frame Compression Using Adaptive Fuzzy Inference System',2018 International Symposium on Advanced Intelligent Informatics (SAIN)

[5] Cheng Z, Sun H, Takeuchi M, & KattoJ,'Learning Image and Video Compression Through Spatial-Temporal Energy Compaction', 2019 IEEE CVF Conference on Computer Vision and Pattern Recognition (CVPR)

[6] Pratistha Mathur,Anju Yadav,Viveak Kumar Verma,Renuka Purohit,'Paradigms of Image Compression and Encryption: A Review',2019 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT)

[7] Cătălin Dumitrescu,Maria Simona Răboacă, Ioana Manta,'Image Compression and Noise Reduction through Algorithms in Wavelet Domain',2020 IEEE 26th International Symposium for Design and Technology in Electronic Packaging (SIITME)

[8] S B Buthelezi,S Reddy,B Twala,'Learning Rate Optimisation of an Image Processing Deep Convolutional Neural Network',2020 International Conference on Power, Instrumentation, Control and Computing (PICC)

[9] Lu X, Wang H,Dong W, Wu F,Zheng Z,'Shi G,Learning a Deep Vector Quantization Network for Image Compression',2019

[10] Toderici G,Vincent D,Johnston N,HwangS J,Minnen D, Shor J, Covell M ,'Full Resolution Image Compression with Recurrent Neural Networks',2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR).