A survey of Image Compression Algorithms based on Deep Learning

Lichuan Wang (wanglc1@chinatelecom.cn)
China Telecom (China)
Shuchun Wang
China Telecom (China)

Research Article

Keywords: Image Compression, Deep Learning, Neural Network

Posted Date: April 12th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2794445/v1

Abstract

Images can carry more information than words, but the data space of images format is much larger than the text format when they are containing the same information. Therefore, how to efficiently compress images to improve their storability and transmissibility is one of the key research issues in the field of computer vision. Through consulting the relevant literature, this paper analyzes the development process of the current image compression technology, and introduces traditional compression methods and deep learning compression methods, while focusing on the compression methods based on deep learning. Through comparative experiments, this paper analyzes the performance of various types of neural networks in image compression tasks, and summarizes the advantages and disadvantages of various types of neural networks in compression tasks.

1 Introduction

With the development of image sensor hardware technology, image compression is now essential for applications such as transmission and storage in data bases. Taking the mobile phone which have the widest range as an example, the image pixels taken by mainstream mobile phones on the market are about 20 million, and the size of each image is about 3MB. With the increasing size of image data, higher requirements have been placed on the network bandwidth for transmission and the hardware devices for storage. In addition to improving the external conditions, the compression operation of the image itself is also crucial for storage, transmission and processing. Image compression means that redundant information in an image is eliminated, only valid information is retained, and the image can be reconstructed according to such valid information. Redundant information can be divided into five types: spatial redundancy, temporal redundancy, visual redundancy, information entropy redundancy and knowledge redundancy. Redundant information refers to useless expression information in the image or information that can be derived from other information [1-4]. According to the development of time line, image compression technology can be divided into traditional image compression methods and image compression methods based on deep learning. In the early days, due to the limited computing power of the computer, only the traditional way of image compression coding can be used for compression. The compression process of this kind of compression method is: the input image is preprocessed first, usually including image segmentation, color space conversion, etc; After that, the main function of converting and coding the preprocessed image is to remove the redundant pixels of the image; Secondly, the information representation of the image is quantized to reduce the storage size of the image; Finally, the quantized image is converted into a bit stream by quotient coding. Such traditional compression methods are mainly represented by JPEG [5-6], JPEG2000 [7-8], BPG [9], WebP [10-11] and other algorithms. Most of these algorithms complete data compression by blocking and quantizing. Although they have good quantizing effect, blocking effect, ringing effect and blocking effect will occur under the condition of high compression ratio, which will greatly damage the image quality. At present, with the development of deep learning technology, using the parameter self-learning ability of deep neural network, more complex coding structures can be constructed to achieve higher quality image compression. According to the type of neural network, the methods of image compression using depth learning algorithm can be divided into the following three categories: based on recurrent neural network (RNN), convolution neural network (CNN) and generative adversarial neural network (GAN). Through these neural network, data formats at different levels such as pixels and bit values can be compressed to achieve deeper data compression, making the image compression method based on deep learning better than traditional image compression algorithm in compression ratio and compression image quality [12-15].

Jiang et al[53] presents an extensive survey on the development of neural networks for image compression which covers three categories: direct image compression by neural networks; neural network implementation of existing techniques, and neural network based technology in 1999. Mohammad Haris Baig et al[54] design deep architectures for lossy image compression in the context of multi-stage progressive encoders. Johannes Balle et al [55] finds that the optimized method generally exhibits better rate–distortion performance than the standard JPEG and JPEG 2000 compression methods. Wen bin Yin[56] propose an end-to-end reference resource based image compression scheme to exploit the strong correlations with external similar images. Since the pioneering work of Toderici et al. [20] in 2015 exploited recurrent neural networks for learned image compression, much progress has been made, benefiting from the strong modeling capacity of deep networks. [57-62].

Therefore, this paper will introduce the principles and current achievements of three kinds of neural network compression methods one by one, analyze and summarize their commonalities, advantages and disadvantages.

2 Image Compression Based On Deep Learning

In traditional methods, to improve the compression ratio and image quality, the encoder structure is optimized to make the encoding of pixel information more compact and efficient. However, the framework used in traditional methods contains multiple submodules and a large number of parameters, which makes the optimization space limited and the compression performance does not improve significantly in the later stage. While deep learning relies on the characteristics of the network to automatically extract feature information and update network parameters. In terms of improving compression performance, it can jointly optimize the entire deep learning network framework, which has more exploration space for improving compression performance. Therefore, the compression method based on deep learning has evolved into three types of network structures: convolutional neural network[50], recurrent neural network[51] and generative adversary neural network[44]. Three types of networks complete efficient image compression from different aspects and a brief summary is presented in Table 1.

2.1 Image Compression based on CNN

Convolutional neural network is an artificial neural network architecture based on the connection mode and structure of biological neurons, which is widely used in the field of computer vision because of its excellent ability to process image data. Its architecture has two characteristics: sparse connection and weight sharing, and the network has translation invariance for images. Sparse connection replaces the connection mode of full connection of neurons between adjacent network layers. Sparse connection can effectively control the number of connections between neurons and only connect local neurons; Weight sharing means that neurons in the same layer network use the same weight parameters, effectively reducing the number of parameters. Moreover, the sparse connection method enables the network to effectively extract the local features of the image, and identify the semantic information of the image after stretching, rotation, translation and other operations. The above three characteristics make the convolutional neural network have higher compression ratio and better compression image quality than traditional compression methods in image compression tasks.

Ball é J team applied convolutional neural network to image compression task for the first time. His team proposed an end-to-end image compression network based on convolutional neural network, which is composed of nonlinear encoder, equalizing quantizer and nonlinear decoder. In the construction of nonlinear encoder, Ball é J team adopted GDN (Generalized Divisional Normalization, a nonlinear normalization method of decorrelation) model to reduce the correlation between data and simplify the subsequent entropy coding [16]. Later, his team proposed an image compression algorithm based on the variational self-encoder on the basis of the previous network. This model uses a super prior to capture the spatial dependence of the potential representation of the image. The minimum average coding length of the quantized entropy model is obtained by calculating the minimum value of the compression cross entropy [17]. Ball é J team has laid the foundation for using convolutional neural network to solve the problem of image compression, and in the subsequent research, each team has also proposed a better codec network. For example, Bao Yuting team proposed an improved image compression method based on hybrid domain attention mechanism and post-processing to solve the problem that the inaccurate estimation of entropy estimation model and the reduction of compression efficiency caused by information redundancy in potential representation. It embedded hybrid domain attention module as a nonlinear converter in the codec network and the super prior network, making the network have more compact potential features and super prior knowledge, The potential features are modeled as a parametric Gaussian scale mixture model to obtain more accurate entropy estimation and improve the compression efficiency [18]. In order to improve the image compression picture, Liu Shuai's team proposed a perceptual metric -based learning image compression rate distortion loss function for application in the end-to-end network. By introducing perceptual metrics into rate distortion loss, the compression model can enhance its ability to capture perceptual differences and semantic information in images [19].

To sum up, convolutional neural network is the first network algorithm to use depth learning algorithm to replace traditional algorithm for image compression. This type of network mainly adopts end-to-end network structure. Convolution networks are added at the coding end for image feature extraction, and convolutional networks are added at the decoding end for image restoration. The parameters of the overall framework are jointly optimized through training network models, To improve the image compression efficiency and image quality.

2.2 Image Compression based on RNN

Different from convolutional neural network, recurrent neural network is mainly used to process time data, which can process time series characteristics of different lengths. In addition to the input from other nerves, the neuron in the recurrent neural network can convert its own output at the previous time into its own input at the current time, which enables the recurrent neural network to have memory power and share the parameters in the time dimension; Secondly, the recurrent neural network uses the feedforward gradient descent training method to update the parameters of the network iteratively. Therefore, it can adjust the compression ratio of the image by controlling the compression degree of the data or controlling the code rate to achieve image compression and reconstruction.

Toderici G team first applied the cyclic neural network to the image compression task, and proposed a convolution and deconvolution residual encoder network based on LSTM. The image to be compressed is input into the network to obtain the output residual, and then the residual output is re input into the network as an input, which is iterated until the training is completed [20]. Thanks to the memory and continuity of the recurrent neural network, the compressed image quality will be better and better with the iteration in the training process. However, because the recurrent neural network has a long-term dependence when processing long time series, the network will have the problem of gradient explosion and gradient disappearance, so the maximum image size that the network can compress is 32x32. In the later stage, Toderici G team proposed to add a new network structure integrating GRU (Gated Recurrent Units) and ResNet in the network to solve the problem of limiting the size of compressed pictures due to the failure to take into account the long-term dependence between picture blocks. The network structure includes four network modules: encoder, binary encoder, decoder and entropy encoder [21]. The processing flow of the network for the input image is the same as that in [20]. The main difference is that the GRU module replaces the LSTM module. Compared with the LSTM module, the GRU module reduces the doorway unit and only retains the update gate and reset gate, and adds residual mapping at the input end of the GRU module to speed up the convergence; At the same time, the entropy coding module is added to the network. By analyzing the probability of the occurrence of each pixel value, the entropy coding represents the value with a high probability of occurrence as a character less than one byte, and the value with a low probability of occurrence as a character more than one byte. In this way, the average size of each pixel is less than one byte, so as to achieve the purpose of compressing the image, and the entropy coding process will not cause loss of data information, Lossless compression of pictures can be realized. The recurrent neural network framework designed by Toderici G team has been widely used in the follow-up, and other scholars have proposed different optimization schemes based on the network framework. For example, Johnston team made the network more sensitive to image feature information to extract more feature information by increasing the perception space of the network, and proposed to use space adaptive bit allocation scheme and structural similarity (SSIM) weighted loss function to further improve the compression performance of the network [22].

A summary study of the outputs of each phase has been carried out. Baig et al. [54] proposed and analyzed different types of polymerization schemes. The basic increment scheme adds the outputs of all stages to form the final decoded image. The loss function of an incremental scheme usually includes a term to encourage the output of each stage to approximate the remainder of the previous stage. A different way to combine all the phases is to treat the multilevel structure as a residual network to form a skip connection scheme. For such a scheme, only one of the loss functions requires the sum of all stages to reconstruct the original image. Unlike the incremental structure, there is no clear arrangement of the remaining parts in the skip join structure. The outputs of all phases contribute to eventual reconstruction, and each phase complements the quality of reconstruction relative to the others. In addition to these two schemes, Baig et al. also report that the network achieves optimal performance using the state propagation structure and the corresponding residual-to-image prediction, where each step generates the prediction of the original image instead of the residual signal. In this stateful propagation scheme, it is important to propagate the state of the layer to the next step to construct a finely decoded image.

In conclusion, the recurrent neural network also uses an end-to-end network architecture in image compression tasks, but it can gradually obtain better compressed images in the iterative process by relying on its ability to store and reuse features at different times. Compared with convolutional neural network, it can set accurate compression ratio for image compression.

2.3 Image Compression based on GAN

Unlike convolutional neural network or recurrent neural network, which completes the corresponding task by extracting feature information, the generation of countermeasure network completes the corresponding task by generating false data. The generation of countermeasure network was initially proposed by Goodfellow based on game theory. The overall network structure consists of a generator and a discriminator. The generated false data can be obtained by inputting noise samples to the generator, and then the generated false data is input into the discriminator, which is responsible for measuring the difference between the distribution of data generated by the generator and the distribution of real data [23]. The training process of the network is to constantly optimize the network parameters through the mutual confrontation and learning between the generator and the discriminator, so that the false data generated by the generator can be infinitely close to the real data. Therefore, in the image compression task, the generation countermeasure network can reconstruct the image according to the image coding information to reduce the amount of image data.

With regard to the application of the confrontation generation network in the field of image compression, Rippel O and Bourdev L first proposed a real-time adaptive image compression algorithm based on the confrontation generation network, which proposed an adaptive encoder with a pyramid structure, and added multi-scale confrontation training in the training process to improve the generalization of the model, After testing, the pictures generated by this network under low bit rate conditions are 2.5 times smaller than those generated by JPEG compression method, and the picture definition is also higher than that of traditional algorithms [24]. At the same time, taking advantage of the generatability of the generation countermeasure network, Eirikur Agustsson proposed a full resolution image compression method for low bit rate [25]. In this study, two compression methods based on the generation countermeasure network are introduced: generative compression. During compression, the image is converted into bit stream by generating a structure of different proportions to retain the overall image content, The arithmetic coding is used for coding, and finally the generator generates the compressed picture based on the coding information; selective generative compression: during compression, the semantic tag information is used to map the areas that need to generate images, and the highly detailed areas that need to be reserved in the images can be customized. In this mode, different image blocks are first obtained by semantic/instance segmentation and stored in the form of vectors, and then each semantic tag is encoded to generate compressed images, The size of the compressed image is inversely proportional to the area generated by semantic label mapping. On the other hand, since the structure of the confrontation generation network contains two sub module networks, namely, generator and discriminator, the amount of computation and parameters of the network is relatively large, which is not conducive to the deployment and use of the model. To solve this problem, Ren Y team proposed an online multi granularity distillation network to obtain a lightweight generation countermeasure network, which can generate high fidelity images under low computing requirements [26]. In this distillation network, the student generator only uses the output information of the teacher network to optimize the parameters, and trains in the setting without discriminator. The student generator is no longer tightly bound to the discriminator, making it more flexible to train and obtain further compression. In addition, the network expands the student model into the teacher model based on the two complementary dimensions of depth and width, and improves the image quality of the deep image from more diversified dimensions. On the other hand, the network uses the granularity information of the middle channel as an auxiliary supervision signal to conduct distillation optimization, further simplifying the parameters of the network model.

When the generation countermeasure network is used for image compression task, the encoder first encodes the image to obtain the corresponding compact representation features, and the quantizer quantizes the features to further reduce redundancy. Then the generator in the generation countermeasure network uses the compact feature information to generate the compressed image, and finally the discriminator measures the difference between the generated image and the original image, The difference meets the threshold condition to output the generated compressed picture.

Table 1: Summary of important image compression models in recent years

Model Name	Introduction	Category
Tiled Network [64]	Introduces explicit intraprediction with a tiled structure in the network.	CNN
GDN Transform [16]	Introduce a GDN based on nonlinear transform codes for end-to-end optimization of the rate–distortion performance.	CNN
3D-CNN Entropy Model [65]	3D-CNN is used for learning a conditional probability model for a multiresidual-block-based network.	CNN
Multiscale CNN [67]	Proposes a multiscale model and corresponding contextual entropy estimation to improve compression efficiency.	CNN
Full-Resolution RNN [21]	First practical recurrent model for variable-bit-rate fullresolution image compression.	RNN
Priming RNN [22]	The recurrent compression model is improved with a proposed priming technique and spatial contextual entropy model.	RNN
Variable-Rate RNN [63]	First propose a general framework for variable-rate image compression based on convolutional and deconvolutional LSTM.	RNN
OMGD[26]	First attempt to popularize single-stage online distillation for GAN-oriented compression	GAN
GC[25]	First present a learned image compression system based on GANs, operating at extremely low bitrates.	GAN
Real-Time Adversarial [24]	First try to adopt a multiscale framework with adversarial loss for learned real-time image compression.	GAN

3 Introduction To Image Compression Datasets

There are many kinds of image datasets. This part mainly summarizes the image datasets that have been applied in the image compression methods based on depth learning technology. The use of deep learning network architecture to complete image compression training requires a large number of image data support, and the selection and adoption of correct image data sets is critical to the role of network structure training. The Table 2 mainly introduces the datasets used in the image compression algorithm based on depth learning.

In specific use, it is necessary to select a reasonable data set application according to the researchers' experimental purposes and methods. Specifically, when the feasibility of the experiment needs to be analyzed, you can choose a scenario in cifar-10 and LSUN. This kind of dataset has small data volume, moderate content, and fast training speed, which can meet the feasibility of the experiment design; When the positioning requirements for experimental purposes are high resolution of reconstructed images, data sets such as DIV2K and Flickr can be used; When the network structure of the experimental design is deep, and a large number of data sets with various types are needed, you can use such data sets as

ImageNet and Open Images V4; When some conditional features are required for constraint in the experimental design, such as using image semantic information to constrain the generator when using cGAN for image compression, image training sets with semantic annotation such as Cityscapes, COCO, LSUN can be selected.

The data sets pointed by the reconstructed images for testing mainly include Kodak, PhotoCD [39], CLIC, RAISE-1K, Tecnik [40]. These datasets have high resolution. For example, Kodak and PhotoCD datasets have an image resolution of 762 * 512, and their pixels are about 400000; As a challenge specially launched for image compression, CLIC provides higher resolution images and photos. The resolution of mobile phone pictures is 1913 * 1361, and that of professional cameras is 1803 * 1175. The number of pixels in Tecnick data set is about 1.4 million.

Table 2
Dataset Introduction

Dataset Name	Introduction	Highlight
ImageNet[41]	More than 15 million pictures, more than 20000 kinds of images with labels	ImageNet is usually used for image classification, target detection and other tasks. Because of its large amount of image data, you can choose the type you want in actual experiments
Yahoo Flickr[49]	Image storage and video hosting website, storing more than 600 million photos	Flickr images are uploaded by users with high resolution, various image types and different image sizes
Open Images V4[42]	With object position annotation dataset, including about 9 million pictures	This data set is derived from the image challenge. Its image types are diverse, the image content scene is complex, the image contains many objects, and the location of objects in the image is annotated
DIV2K[43]	1000 high-definition pictures, 800 of which are for training, 100 for validation, and the last 100 for testing	The data set has ultra-high resolution and strong diversity. It is a training data set designated by NTIRE2018 and is mostly used in image super-resolution tasks
CityScapes[44]	5000 images, 500 validation Images and 1525 test Images	It contains images of 50 cities throughout the year, with high-quality image annotation. 5000 images from 27 cities were manually selected for intensive annotation
Place365[45]	Including 1.8 million images of 365 scenes, each scene can contain up to 5000 images	Place365 has two versions: Places 365 standard and Place365 challenge. This data set is marked with the semantic category of the scene, and is mostly used in scene classification tasks
COCO[47]	Including more than 330000 images of 80 categories, including 200000 images with labels	COCO is mostly used in target recognition, object detection, object segmentation and other fields, with a large number of clothing category images and rich categories
CIfar-10[48]	32 * 32 color images include 50000 training samples and 10 categories	Classic image classification and image recognition data set, with low image resolution and few categories
LSUN[46]	Including 10 scene lists, 20 object categories, and about 1 million tag images	It mainly includes 10 kinds of scenes such as bedroom, living room and classroom, covering basic life scenes, which can be used for large-scale scene understanding
ImageNet[41]	More than 15 million pictures, more than 20000 kinds of images with labels	ImageNet is usually used for image classification, target detection and other tasks. Because of its large amount of image data, you can choose the type you want in actual experiments

4 Image Compression Evaluation Index And Model Test

At present, with the optimization and improvement of the algorithm performance, the compressed image quality cannot be seen intuitively by the human eye, so we need to use various objective evaluation indicators to accurately and objectively evaluate the compression performance of the algorithm [27–29]. Common evaluation indicators used in compression image quality evaluation include PSNR, SSIM and MS-SSIM. BPP is mainly used for compression ratio evaluation indicators. The specific meaning and calculation formula of the above indicators are as follows:

PSNR (Peak Signal to Noise Ratio) is used to measure the pixel difference between the original image and the compressed image. PSNR obtains the difference of the enlarged image by comparing the difference between each pixel one by one. Since there are positive and negative differences between the image pixels before and after compression, in order to avoid the offset of the difference caused by superposition, the mean square error is calculated for the pixel difference:

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left[x(i,j)-x'(i,j)\right]^2$$

1

In the above formula, x and x' is the original picture and the compressed picture, m and n is the size of the image, and $(i,j)$is the gray value of the coordinate pixel in the picture. The dynamic range of different image pixel values is also different, so the influence of dynamic range shall be considered when calculating PSNR value. The calculation formula of PSNR is:

$$PSNR = 20 \cdot \log_{10}(\frac{MAX(x)}{\sqrt{MSE}})$$

2

MAX (x) represents the maximum gray value of the pixel in the original image, that is, the right value of the dynamic range. The higher the PSNR value, the higher the similarity between the two images, and the better the compression performance of the network.

SSIM (structural similarity index) is used to measure the correlation of adjacent pixels between the original image and the compressed image. The correlation of adjacent pixels in the image can reflect the structural information of objects in the actual scene. Based on this index, it can judge whether there is structural distortion of objects in the compressed image. The calculation formula is:

$$SSIM(x,x') = \frac{(2\mu_x\mu_{x'}+c_1)(2\sigma_{xx'})}{(\mu_x^2+\mu_{x'}^2+c_1)(\sigma_x^2+\sigma_{x'}^2+c_2)}$$

3

In this formula, $\mu_x and \mu_{x'}$ is the average pixel value of the image, $\sigma_x and \sigma_{x'}$ is the pixel variance of the image, and $\sigma_{xx'}$ is the pixel covariance of the image before and after compression, c_1, c_2 are constant values. The value range of SSIM is [0,1]. The higher the structural similarity between the compressed image and the original image, the closer its value is to 1.

MS-SSIM (Multi scale SSIM index), multi-scale structure similarity. Compared with SSIM index, MS-SSIM index increases the structure similarity judgment of images under different resolution conditions. The calculation process is to conduct multiple down-sampling processing for images, calculate the SSIM values of images on different scales respectively, and finally summarize the SSIM index values on multiple scales to MS-SSIM values. Due to the small range of MS-SSIM changes, it is not easy to reflect the difference of travel, so its unit is converted into decibels:

$$MS-SSIM(dB) = -10\log_{10}(1-MS-SSIM)$$

4

BPP is used to measure the compression rate of image compression. Its value indicates the average number of bits required for each pixel to encode. The calculation formula is:

$$BPP = \frac{8}{m \times n} \times SIZE$$

5

In the above formula, it is the number of bytes of the compressed image. The higher the BPP value, the higher the compression ratio.

In deep learning, image compression is an unsupervised learning task. The data set used for training network does not need to be classified or labeled. Therefore, there are many choices of data sets that can be used for image compression network training. In this experiment, we choose COCO dataset, which contains 80 object categories and 330000 images. According to the above evaluation indexes, the representative network algorithms in convolutional neural network, cyclic neural network and countermeasure generation network in recent years are selected for comparative test. The test results are shown in the following Table 3:

Table 3
Comparison of network compression performance

Model	PSNR/db	SSIM	MS-SSIM	BPP
Sun H et al[30]	31.735	84.1%	95.3%	52.4%
Brummer B et al[31]	33.432	85.7%	96.1%	50.1%
Kamisli F [32]	32.753	83.4%	94.3%	56.7%
Koyuncu A B et al[33]	28.431	80.2%	91.7%	44.8%
Fu H et al[34]	28.982	81.3%	92.1%	46.0%
Islam K et al[35]	26.682	79.5%	90.3%	41.2%
Iwai S et al[36]	22.831	72.7%	88.5%	30.1%
Lin N et al[37]	21.084	71.8%	87.3%	28.6%
Mulkiah et al[38]	23.127	73.9%	87.6%	33.7%

In the Table 3, 1–3 are compression networks based on convolutional neural networks, 4–6 are compression networks based on recurrent neural networks, and 7–9 are compression networks based on generative adversarial network. From the test results in the Table 3, it can be seen that the three types of networks have different emphasis in image compression. Convolutional neural networks are mainly used in encoders and decoders. Their excellent parameter learning ability enables the encoders and decoders to compress better and clearer images; Compared with convolutional neural network, recurrent neural network focuses on the optimization of entropy coding module, so its BPP index is better than that of neural network; The adversary generation network relies on its own generator to generate compressed images, and its advantage is that it can obtain images with high compression ratio.

5 Summary

This paper introduces the development and classification of image compression technology, focuses on the image compression methods based on deep learning, and introduces the ideas and methods of realizing image compression in various types of networks according to the classification of network types, and lists classic algorithm cases in each type of network for illustration. At the same time, through literature review, code reproduction and other ways, according to the network type, we selected the more representative network algorithms in recent years for comparative analysis, and summarized and verified the advantages and disadvantages of each type of network.

Declarations

Author contributions

Lichuan Wang. and Shuchun Wang wrote the main manuscript text. All authors reviewed the manuscript.

Data availability statement:

The Data is available in https://cocodataset.org/#home.

Competing Interest

The author declares that there are no conflicts of interest regarding the publication of this paper.

References

1. Rai Sandeep,Shrivastava Aditi,Nigam Rajesh. Review of Color Image Compression using Discrete Wavelet Transform and Block based Image Coding[J]. International Journal of Computer Applications,2019,178(39).
2. Kaur Suhel,Kaur. Sumeet. Fractal Image Compression - A Review.[J]. International Journal of Advanced Research,2016,4(7).
3. Thaneshwar Kumar,Ramesh Kumar. A Survey On Medical Image Compression Using Hybrid Compression Technique (DWT, DCT and Huffman coding)[J]. International Journal of Management, IT and Engineering,2014,4(10).
4. P. Anandan, R. S. Sabeenian. Curvelet based Image Compression using Support Vector Machine and Core Vector Machine – A Review[J]. International Journal of Advanced Computer Research (IJACR),2014,4(15).
5. Ince, Ibrahim Furkan,Bulut, Faruk,Kilic, Ilker,Yildirim, Mustafa Eren,Ince, Omer Faruk. Low dynamic range discrete cosine transform (LDR-DCT) for high-performance JPEG image compression[J]. The Visual Computer,2022(prepublish).
6. Sekine Asadi Amiri,Hamid Hassanpour. Image compression using JPEG with reduced blocking effects via adaptive down-sampling and self-learning image sparse representation[J]. Multimedia Tools and Applications,2018,77(7).
7. Cai Jin,Wang Shuo. Study of the Discrete Wavelet Transform Based on JPEG2000 Image Compression Algorithm[J]. Applied Mechanics and Materials,2013,464(464-464).
8. David S. Taubman,Michael W. Marcellin. JPEG2000 Image Compression Fundamentals, Standards and Practice[M].Springer, Boston, MA.
9. Mentzer F.,Van Gool L.,Tschannen M.. Learning better lossless compression using lossy compression[J]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition,2020.
10. Lian L , Wei S . Webp:A New Image Compression Format Based on VP8 Encoding[J]. Microcontrollers & Embedded Systems, 2012.
11. Kraichingrith E , Phimoltares S . WebP-based high capacity image steganography. 2018.
12. Gandor Tomasz,Nalepa Jakub. First Gradually, Then Suddenly: Understanding the Impact of Image Compression on Object Detection Using Deep Learning[J]. Sensors,2022,22(3).

13. Tripathi Ramesh Chandra,Guo Y.,Liu Y.,Oerlemans A.,Lao S.,Wu S.,Lew M. S.,Wang H.,Yeung D. Y.,Hao W.,Yeung D. Y.,Memisevic R.,Tyler C. W.,Likova L. T.,Molnar T.,Zhang Y.,Laub J. A.,Zahedi M.,Zhao R. W.,Wu Z.,Li J.,Jiang Y. G.. A review on deep learning for visual understanding[J]. ACADEMICIA: An International Multidisciplinary Research Journal,2021,11(12).
14. Mishra Dipti,Singh Satish Kumar,Singh Rajat Kumar. Deep Architectures for Image Compression: A Critical Review[J]. Signal Processing,2022,191.
15. Li Zewen,Liu Fan,Yang Wenjie,Peng Shouheng,Zhou Jun. A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects.[J]. IEEE transactions on neural networks and learning systems,2021,PP.
16. Ballé J., Laparra V., Simoncelli E. P. End-to-end optimization of nonlinear transform codes for perceptual quality[C]//2016 Picture Coding Symposium (PCS). Nuremberg: IEEE Press, 2016: 1–5.
17. Ballé J., Minnen D, Singh S,et al.Variational image compression with a scale hyperprior[J].arXiv:1802.01436,2018.
18. Bao Yuting and Tao Yuwen and Qian Pengjiang. Image Compression Based on Hybrid Domain Attention and Postprocessing Enhancement[J]. Computational Intelligence and Neuroscience, 2022, 2022 : 4926124-4926124.
19. Liu, Shuai et al. End-to-end image compression method based on perception metric[J]. Signal, Image and Video Processing, 2022, : 1-8.
20. Toderici G O'Malley S M Hwang S J et al.Variablerate image compression with recurrent neural networks[J].arXiv 1511.06085 2015.
21. Toderici G Vincent D Johnston N et al.Full resolution image compression with recurrent neural networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2017 5306-5314.
22. Johnston N Vincent D Minnen D et al.Improved lossy image compression with priming and spatially adaptive bit rates for recurrent networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2018 4385-4393.
23. Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[C]// International Conference on Neural Information Processing Systems. MIT Press, 2014:2672-2680.
24. Rippel O Bourdev L.Real-time adaptive image compression[C]//Proceedings of the 34th International Conference on Machine Learning-Volume 70 2017 2922-2930.
25. Agustsson E Tschannen M Mentzer F et al.Generative adversarial networks for extreme learned image compression[C]//Proceedings of the IEEE International Conference on Computer Vision 2019 221-231.
26. Ren Y , Wu J , Xiao X , et al. Online Multi-Granularity Distillation for GAN Compression[J]. 2021.
27. Atta Rawheyaa E.,Kasem Hossam M.,Attia Mahmoud. A Comparison Study for Image Compression Based on Compressive Sensing[J]. ELEVENTH INTERNATIONAL CONFERENCE ON GRAPHICS AND IMAGE PROCESSING (ICGIP 2019),2020,11373.

28. Image and Video Processing; Findings from University of Girona in the Area of Image and Video Processing Reported (Overview and Evaluation of the Jpeg Xt Hdr Image Compression Standard)[J]. Journal of Technology,2019.
29. Barry Mazor. Image Compression: Standard vs. Proprietary[J]. Advanced imaging,2003,18(4).
30. Sun H , Yu L , Katto J . End-to-End Learned Image Compression with Quantized Weights and Activations[J]. 2021.
31. Brummer B , Vleeschouwer C D . End-to-end optimized image compression with competition of prior distributions[C]// Computer Vision and Pattern Recognition. IEEE, 2021.
32. Kamisli F . End-to-End Learned Block-Based Image Compression with Block-Level Masked Convolutions and Asymptotic Closed Loop Training[J]. 2022.
33. Koyuncu A B , Gao H , Steinbach E . Contextformer: A Transformer with Spatio-Channel Attention for Context Modeling in Learned Image Compression[J]. 2022.
34. Fu H , Liang F , Lei B , et al. An extended context-based entropy hybrid modeling for image compression[J]. Signal Processing. Image Communication: A Publication of the the European Association for Signal Processing, 2021(95-):95.
35. Islam K , Dang L M , Lee S , et al. Image Compression with Recurrent Neural Network and Generalized Divisive Normalization[C]// Computer Vision and Pattern Recognition. IEEE, 2021.
36. Iwai S , Miyazaki T , Sugaya Y , et al. Fidelity-Controllable Extreme Image Compression with Generative Adversarial Networks[C]// International Conference on Pattern Recognition. Institute of Electrical and Electronics Engineers Inc. 2021.
37. Lin N , Zhu Y , Zhang Y , et al. Edge-enhanced Generative Adversarial Network for Reconstruction of Compressed Image. 2021.
38. Mulkiah, Suprapto, Sari A K . Compression-Encryption Model for Digital Images Based on GAN and Logistic Map[C]// 2021 International Seminar on Intelligent Technology and Its Applications (ISITIA). 0.
39. Franzen, Rich. "Kodak lossless true color image suite." source: http://r0k. us/graphics/kodak 4.2 (1999).
40. Asuni N.Tecnick test image library[EB/OL]. 2011 .http //testimages.tecnick.com.
41. Deng, Jia, et al. "Imagenet: A large-scale hierarchical image database." 2009 IEEE conference on computer vision and pattern recognition. Ieee, 2009.
42. Kuznetsova A Rom H Alldrin N et al.The open images dataset v4 unified image classification object detection and visual relationship detection at scale[J].arXiv 1811.00982 2018.
43. Agustsson E Timofte R.Ntire 2017 challenge on singleimage super- resolution dataset and study[C]//Proceedingsof the IEEE Conference on Computer Vision and PatternRecognition Workshops 2017 126-135.
44. Goodfellow, Ian, et al. "Generative adversarial networks." Communications of the ACM 63.11 (2020): 139-144.

45. Zhou, Bolei, et al. "Places: A 10 million image database for scene recognition." IEEE transactions on pattern analysis and machine intelligence 40.6 (2017): 1452-1464.
46. u F Seff A Zhang Yinda et al.LSUN construction of a large- scale image dataset using deep learning with humans in the loopar[J]arXiv 1506.03365 2015.
47. COCO - Common Objects in Context (cocodataset.org)
48. CIFAR-10 and CIFAR-100 datasets (toronto.edu)
49. Flickr Blog
50. Dong, Chao, et al. "Compression artifacts reduction by a deep convolutional network." Proceedings of the IEEE international conference on computer vision. 2015.
51. Mikolov, Tomas, et al. "Recurrent neural network based language model." Interspeech. Vol. 2. No. 3. 2010.
52. Ma, Haichuan, et al. "End-to-end optimized versatile image compression with wavelet-like transform." IEEE Transactions on Pattern Analysis and Machine Intelligence (2020).
53. Jiang, J. "Image compression with neural networks–a survey." Signal processing: image Communication 14.9 (1999): 737-760.
54. Baig, Mohammad Haris, Vladlen Koltun, and Lorenzo Torresani. "Learning to inpaint for image compression." Advances in Neural Information Processing Systems 30 (2017).
55. Ballé, Johannes, Valero Laparra, and Eero P. Simoncelli. "End-to-end optimized image compression." arXiv preprint arXiv:1611.01704 (2016).
56. Yin, Wenbin, et al. "A Reference Resource Based End-to-End Image Compression Scheme." Pacific Rim Conference on Multimedia. Springer, Cham, 2018.
57. Canh, Thuong Nguyen, Motong Xu, and Byeungwoo Jeon. "Rate-distortion optimized quantization: A deep learning approach." Proc. IEEE High Perform. Extreme Comput. Conf.. 2018.
58. Li, Qing, and Yang Chen. "Learning to compress using deep AutoEncoder." 2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton). IEEE, 2019.
59. Wang, Yefei, et al. "Ensemble learning-based rate-distortion optimization for end-to-end image compression." IEEE Transactions on Circuits and Systems for Video Technology 31.3 (2020): 1193-1207.
60. Yilmaz, M. Akin, and A. Murat Tekalp. "End-to-end rate-distortion optimization for bi-directional learned video compression." 2020 IEEE International Conference on Image Processing (ICIP). IEEE, 2020.
61. Eusébio, Paulo, João Ascenso, and Fernando Pereira. "Optimizing an Image Coding Framework with Deep Learning-based Pre-and Post-Processing." 2020 28th European Signal Processing Conference (EUSIPCO). IEEE, 2021.
62. Hu, Yueyu, et al. "Learning end-to-end lossy image compression: A benchmark." IEEE Transactions on Pattern Analysis and Machine Intelligence (2021).

63. Toderici, George, et al. "Variable rate image compression with recurrent neural networks." arXiv preprint arXiv:1511.06085 (2015).
64. Minnen, David, et al. "Spatially adaptive image compression using a tiled deep network." 2017 IEEE International Conference on Image Processing (ICIP). IEEE, 2017.
65. Mentzer, Fabian, et al. "Conditional probability models for deep image compression." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.
66. Tschannen, Michael, Eirikur Agustsson, and Mario Lucic. "Deep generative models for distribution-preserving lossy compression." Advances in neural information processing systems 31 (2018).
67. Klopp, Jan, et al. "Learning a Code-Space Predictor by Exploiting Intra-Image-Dependencies." BMVC. 2018.