Extended Essay

Computer Science

Topic: Image compression

Research Question: What is the comparative effectiveness of traditional image compression techniques versus AI-driven methods in terms of preserving image quality and reducing file size?

Word count: 3898

Contents

1 Introduction

In the time of ever-evolving technologies, information that is transferred and stored rises exponentially. Surge in 4.0 technologies industry emerged an increasing demand for robust data transfer, which can be facilitated through robust data compression techniques (*What Is Industry 4.0 and How Does It Work?* | *IBM*, n.d.). In the raising popularity of messengers, social media, and webpages, online images became an indispensable part of the network (*Number of Worldwide Social Network Users 2027* | *Statista*, 2023).

Even though the traditional image compression techniques like JPG or PNG perform decently, the raising popularity of the artificial intelligence (AI) created a new application of it in image compression field, where machine learning (ML) architectures are widely used (Minnen et al., 2018).

This research focuses on the comparison of traditional techniques with AI-based methods to analyze the possible improvements in the field of image compression. These findings, will help to answer the important question on the need of transition or use of novel image compression method in fields with high image transfer. This may include medical imaging, astronomy, GIS, media, and more.

For the comparative analysis a set of traditional and AI-based techniques will be chosen to evaluate both methods extensively. This set confines to JPG, HEIF for traditional methods and AI-based compression models such as cheng2020-anchor,

mbt2018-mean, and bmshj2018-hyperprior (Zhou et al., 2023). The choice was made considering the popularity, and effectiveness of image compressions methods. The comparison will be mainly based on quantitative measures of full-reference methods, which involve comparing compressed image to undistorted original image. These methods include mean square error (MSE), peak signal-to-noise ratio (PSNR), and structural similarity index (SSIM) to analyze the image similarity (Kunwar, 2023). Apart from this, the space saving percentage will be calculated between original and compressed image to measure the efficiency of compression techniques.

2 Background Knowledge

A typical image is represented as a matrix, where values of this matrix correspond to pixel intensity values: larger numbers mean brighter pixels and smaller numbers represent darker pixels. A colored picture has different channels for red, green, and blue color components. Theoretically, an image of 512x512 pixels with 3 color channels will have a size of $512^2 \times 3$ (786 432) bytes, which is too big (*Image Compression*). Because of this, to ensure efficient data storage image compression is used.

Image compression is the size reduction process of digital images without decreasing the image quality below an acceptable threshold (Sheldon, 2022). There are two compression methods: lossless, and lossy. Formal one removes statistical redundancy–non-essential–information from the image and stores it in

index file, while latter one permanently removes some of the information with insignificant difference for observer. Because of this, lossy method is much more efficient in terms of memory space savings; however, this also creates a tradeoff between image quality and size (*Lossy Vs Lossless Compression: Differences & Advantages | Adobe*, n.d.).

Traditional standards of compression use are divided into few main steps

- Color space transformation

The first step is typically a color space conversion. Instead of representing an image with its original color space (e.g., RGB), they are converted into a color space where one channel represents the light intensities, while other two represent the colors. This separates the (Y) luminance – light intensity and (Cb, Cr) chrominance – color components (*How JPEG Works (Christopher G. Jennings)*, n.d.). This partition then allows to manipulate chrominance and luminance to save more image size.

- Block-Based processing

In this step, images are subdivided into block, where each block is then compressed. In JPEG blocks are divided into 8x8 pixel grid, while in HEIF each picture is partitioned into tiles and then into Coding Tree Units (CTUs), which is are then divided into Coding Units of up 32x32 sizes (*HEVC: An Introduction to High Efficiency Coding*, n.d.).

- Transform coding

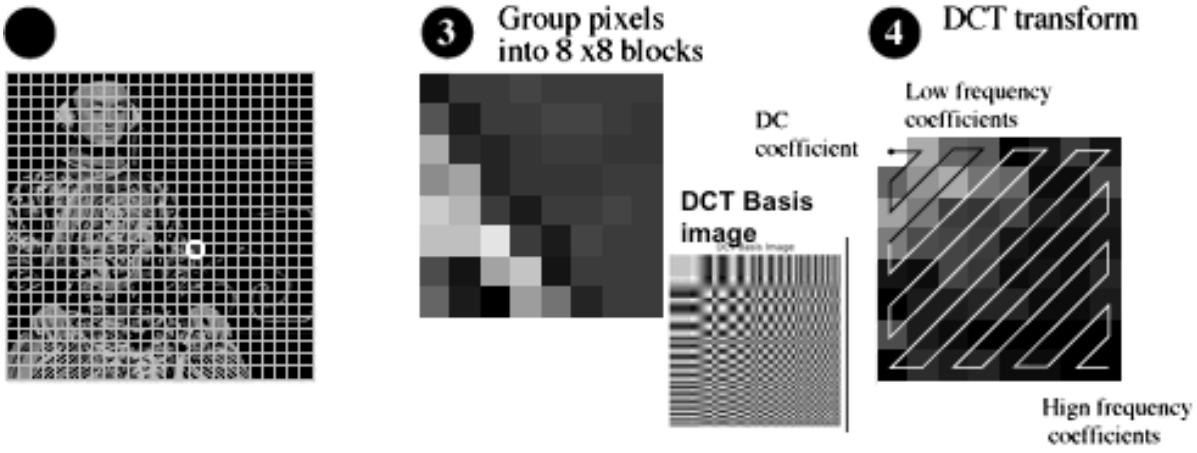

Figure 1 Discrete Cosine Transform with basis image (*JPEG: Joint Photographic Experts Group*, n.d.)

Here JPEG and HEIF implemented algorithms vary. Discrete Cosine Transform (DCT) is used in JPEG. The main idea behind this is to transform 8x8 blocks that represent luminance into a product of constant set of base images defined by the JPEG.

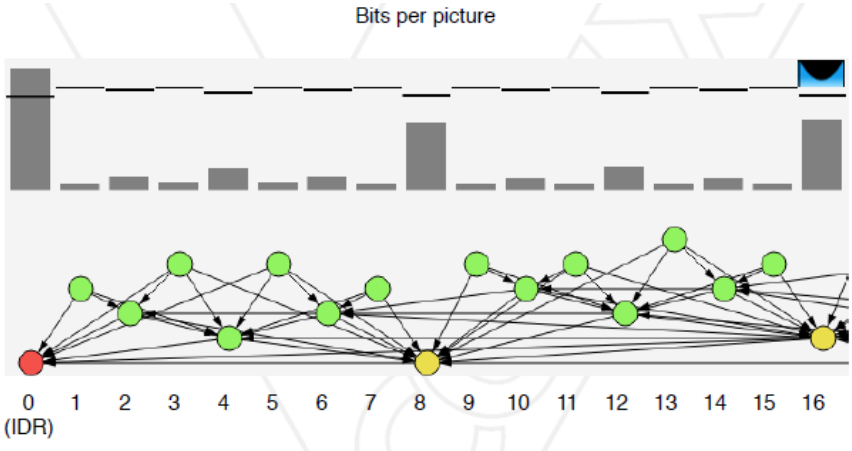

Figure 6: Sequence of coded pictures (source: Parabola Research)

Figure 2 Coded image prediction in HEVC (*HEVC: An Introduction to High Efficiency Coding*, n.d.)

In HEIF, frames are computer using intra or inter prediction (*HEVC: An Introduction to High Efficiency Coding*, n.d.). The main idea is to predict each part of coding unit using neighboring image data in the same picture.

- Quantization and Encoding

In this step the selected data is removed from an image to save more space and thus compress the image. In the case of JPEG, quantization tables are utilized, which are a pre-defined constant 8 x 8 tables for luminance and chrominance. The upper-left corner of each of 8 x 8 (Y) luminance image block will be divided by the number in the upper-left corner of 8 x 8 block in the (Cb, Cr) – chrominance table. Finally, block data will have a smaller size. Then, the similar values in the blocks are condensed using entropy encoding. HEIC, however, uses context-adaptive binary arithmetic coding (CABAC) for entropy coding, which is the main algorithm that decreases the file size.

AI-based image compression

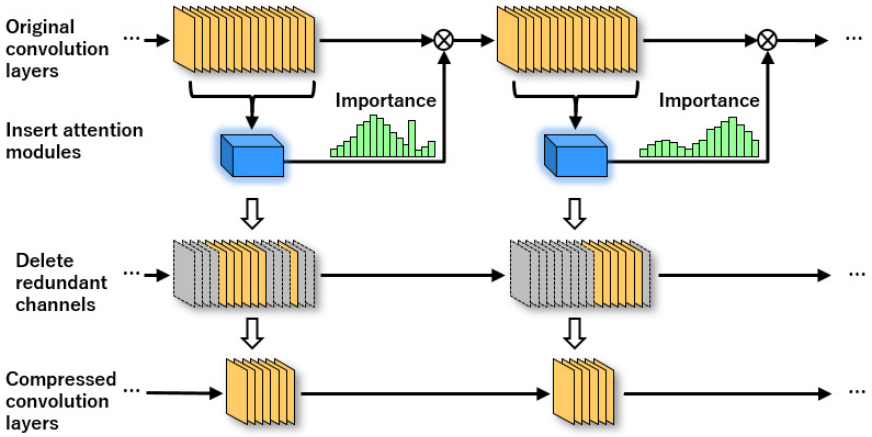

Figure 3 Neural Network based image compression (*Improvement of Power Consumption Efficiency for Advanced AI Processing Using Deep Learning Compression Technology* | *Challenge Zero*, n.d.)

Neural networks are utilized to understand the inherent structure and features of an image. This process is typically divided into two main stages: training and compression During the training stage, the neural network is exposed to a vast dataset of diverse images, where it tries to find an optimal tradeoff in the distortion for the image quality (Zafari et al., 2023).

After training, when presented with an image, the network encodes image into a compact representation by maintaining the most critical features while removing redundant and unimportant data. This then creates a compressed image with much lower file size.

3 Rationale for Selected Compressions Methods

Due to the better performance of lossy algorithms in comparison to lossless, comparison will be based only implementing the lossy method.

Traditional

JPEG (Joint Photographic Experts Group) format can be rationalized based on its popularity and decent performance. Despite its early development in 1990, 77.8% of all websites implements JPEG and it is compatible with most of the technologies used nowadays (*Usage Statistics of Image File Formats for Websites, January 2024*, n.d.). Even though more innovative algorithms were developed (JPEG2000) or (JPEG XR), they are not used universally (*"Jpeg" | Can I Use. . . Support Tables for HTML5, CSS3, Etc*, n.d.). Thus, JPEG method choice was chosen to evaluate the universally applied image compression technique with new ones.

HEIF (High Efficiency Image Format) was chosen for its novelty and high efficiency. It was first applied by the Apple company in 2017 (*Wikiwand - High Efficiency Image File Format*, n.d.). However, some android devices recently started to support it. Due to the application of more innovative algorithms, it compresses almost 2 times more efficient than JPEG, while conserving the better quality. This standard was decided to be compared for more extensive understanding of the non-AI based techniques, as this method outperforms most of the traditional techniques and is used by many popular systems.

ML-based

cheng2020-anchor is a model which utilizes discretized Gaussian Mixture Likelihoods and attention modules to achieve accurate and flexible entropy mode (Cheng et al., 2020)l. mbt2018-mean model employs convolutional neural network (CNN) to perform lossy image compression (Minnen et al., 2018). Finally, bmshj2018-hyperprior applies artificial neural networks (ANN) to build its own model (Ballé et al., 2018). These network architectures perform well in most of the comparisons across related works. This makes provided models to be the most interesting for this research (Jamil et al., 2022).

All these models are pre-trained and provided by the “CompressAI” toolbox on Python programming language (Bégaint et al., 2020).

4 Methodology

Images used

The images of similar dimensions (768 x 512) were used to analyze the performance of different compression methods. It is because AI-based models were trained using these dimensions and change may lead to dimensions-related bias which can hinder the assessment. Images can be separated to 7 categories: cityhall, coffee, hike, gym, racoon, nba, and sea with 3 images for each category. This gives a total of 21 images. Tests were performed on all 21 images for each of the described compression method (JPEG, HEIF, cheng2020-anchor, bmshj2018-

hyperprior, mbt2018-mean). Images were chosen to showcase different levels of detailing for all-round analysis. Appendix 1 can be referred for image samples for each category.

Qualitative Analysis

Because images transmit information through vision, for the qualitative metric visual inspection was used. Compressed images will be cropped to (468x212) dimensions to show how AI-based models handle details on the image.

Quantitative Experimentation

Image compressions field widely uses several metrics for the image similarity analysis which were applied in this analysis. However, they only evaluate two images based on their pixel-wise comparison that is why the qualitative analysis is also crucial. All metrics calculations are applied using scikit-image library, specifically the skimage.measure module.

Space Saving

Space saving formula is given by

$$\text{Space Saving } (\%) = (1 - \frac{\text{Compressed Size}}{\text{Originall Size}}) * 100\%$$

With higher percentage of this metric, the higher the compression efficiency, that is, an image decreases in sizes. Conversely, with lower parameters of space saving, image compression is less efficient (*Wikiwand - Data Compression Ratio*, n.d.).

Mean Squared Error (MSE)

MSE between two images is defined as

$$\mathrm{MSE} = \frac{1}{\mathrm{MN}} \sum_{n=0}^{M} \sum_{m=1}^{N} [\mathrm{I}(n, m) - \hat{\mathrm{I}}(n, m)]^2$$

where $\mathrm{I}(n, m)$ is the original image and $\hat{\mathrm{I}}(n, m)$ is the compressed image and N, M is number of rows. It is possible to see that this formula represents the absolute error between two images. Basically, a lower value of MSE indicates better similarity, as it shows that the absolute error is small, therefore the difference is insignificant (Kunwar, 2023). Provided skimage.metrics.normalized_root_mse function was used to calculate MSE of two images

Peak Signal-to-Noise Ratio (PSNR)

PSNR between two images is stated as

$$\mathrm{PSNR} = 10 \log_{10} \left(\frac{\mathrm{MAX}^2}{\mathrm{MSE}} \right)$$

where MAX is the maximum value in the image data. For an 8-bit unsigned integer, the MAX is 255, which is white. PSNR shows how well a compressed image retains the quality with higher values suggesting less perceptual difference (Kunwar, 2023). Since PSNR is a logarithmic function with base 10, every incremental increase is in fact increase by factor of 10. The

skimage.metrics.peak_signal_noise_ratio function was applied from the scikit-image library.

Multi-Scale Structure Similarity Index Method (MS-SSIM)

SSIM of two images is expressed as

$$\mathrm{SSIM} = \frac{(2\mu_X\mu_Y + c_1)(2\sigma_{XY} + c_2)}{(\mu_X + \mu_Y + c_1)(\sigma_X^2 + \sigma_Y^2 + c_2)}$$

Where μ – mean, σ – standard deviation of original image X and the compressed image Y. σ_{XY} is the covariance of X and Y (Kunwar, 2023). This metrics evaluates the luminance and contrast information of two images. Luminance is the intensity or brightness of pixels, while contrast is the comparison of saturation of pixels. MS–SSIM extends SSIM metrics, calculating it on each scale which gives better results. (*Wikiwand - Structural Similarity*, n.d.)

5 Experimental Results

Qualitative analysis

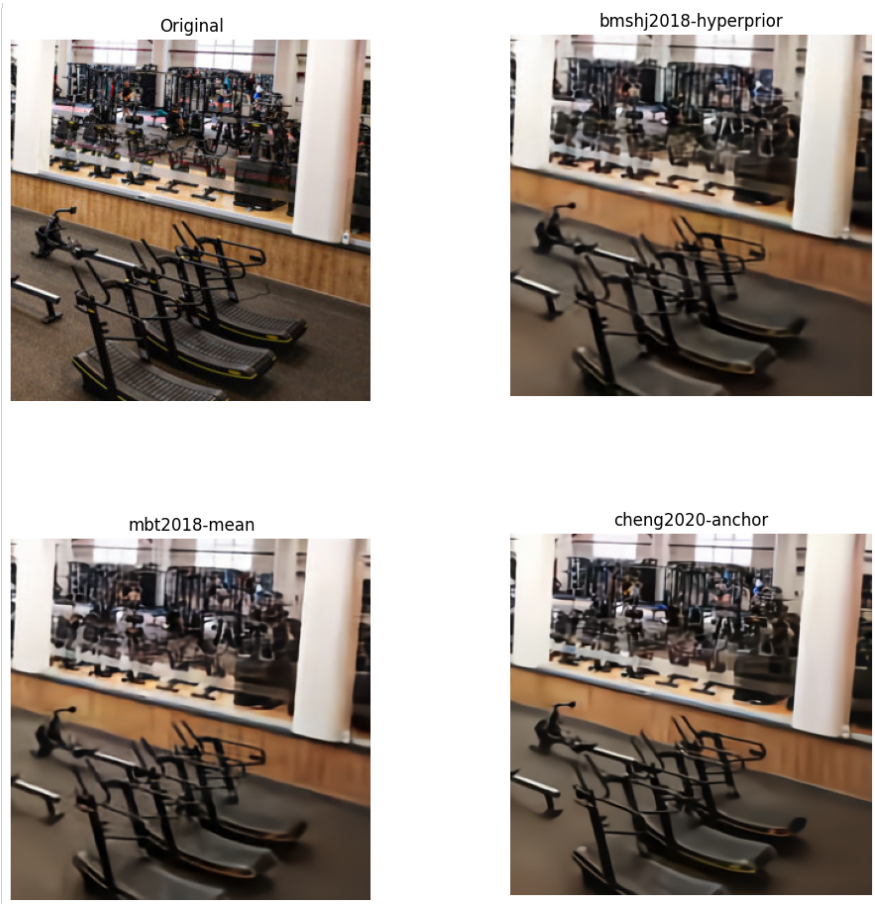

Figure 4 AI-based compressions of cropped “gym” image

Here is the cropped “gym” image with 468x212 dimensions. This is a very detailed image which consists of objects on the mirror and objects in the front. It is possible to see how machine learning models remove details from the images. In most of the cases, these small details such as objects on the back, and surface of the treadmills are blurred because of which quality degrades. No evident differences can be mentioned on the better performance of one model over another.

Figure 5 AI-based compressions of cropped “coffee” image

In contrast to the well detailed “gym” image, the “coffee” image was analyzed to indicate the effectiveness of AI-based compression with lower detailed type images. The overall file size of the original image is 52kB which is small. This also contributes to fewer details presented on the picture. After compression, even though the details on the image such as surface structure of the spoon or the plate are blurred, the difference is small. The visual aspects of images after compression are similar.

Figure 6 Original “nba” image

Figure 7 Cheng2020-anchor compression of the “nba” image

Figure 8 mbt2018-mean compression of the “nba” image

Figure 9 bmshj2018-hyperprior compression of the “nba” image

Figure 10 JPEG compression of the “nba” image

Figure 11 HEIF compression of the “nba” image

This is a comparison of “nba” images with dimensions of 768x512. The original file is located on the top left corner. It is possible to see that the background, having

little to no details, retains its original quality and there is almost no difference between the original image and three other AI-based compressions. However, the human skin on the original image captures more details, as it is possible to see sweat and freckles. After compression these details are blurred, and sharpness decreases.

In contrast, the traditional compressions methods (JPEG, HEIF) retain the original image quality of details. There is no visible difference between HEIF and JPEG images, which suggests that the compression quality will be mainly based on the quantitative metrics for the comparison of these two traditional methods.

Metrics of Machine Learning based compressions

Photo Name	**Space Saving (%)**			**Original Photo Size (kB)**
	cheng2020-anchor	mbt2018-mean	bmshj2018-hyperprior	
cityhall	50,4	53,4	52,6	133
coffee	28,8	26,9	26,9	52
hike	42,6	48,2	48,2	141
gym	59,3	61,4	61,4	147
racoon	47,5	48,8	50,0	80
nba	41,0	42,6	41,0	61
sea	44,1	42,6	42,8	98
Average	44,8	46,3	46,1	102

Table 1 Space Savings of the test images using AI.

According to space saving metrics, the most efficient model is mbt2018-mean which shows the best result pf 46.3% space saving, meaning that it compresses an image by almost half of its original size. However, the difference between these models is not significant, as the range between average results is only 1.5% which is negligible. One of the evident outliers is the “coffee” image, where space saving drops to an average of 27.5% (1). This might be due the extremely small photo size of the original image file. Because when there is low data on the original image, the amount of data with insignificant value to the observer decreases, therefore compression efficiency also decreases.

$$\frac{28.8\% + 26.9\% + 26.9\%}{3} = 27.5\% \ (1$$

Photo Name	MSE			Original Photo Size (kB)
	cheng2020-anchor	mbt2018-mean	bmshj2018-hyperprior	
cityhall	131	164	165	133
coffee	27,8	32,5	33,6	52
hike	266	292	295	141
gym	101	132	133	147
racoon	78,9	83,2	84,3	80
nba	22,8	29,3	31,0	61
sea	88,6	109	109	98
Average	**102**	120	122	**102**

Table 2 Mean Squared Error of the test images using AI.

In spite of the worst results in space saving, cheng2020-anchor model performed the best in MSE metrics, which demonstrates that there is lower pixel difference of compressed image and original image in comparison to other AI-based models. Outliers can be found on the “hike”, “nba”, and “coffee” images, where MSE value is significantly higher or lower.

The “nba” and “coffee” images have significantly lower original photo sizes (61kB) and (52kB) respectively. This is because in smaller file size images, the total resolution, as well as, the details are fewer, therefore there is smaller range of pixel values, which decreases the span of errors.

In contrast, the “hike” image has higher photo size = (141kB) and MSE = (284) (2) is correspondingly higher. This is because this photo captures more details which is difficult for AI to analyze.

$$\frac{266 + 292 + 295}{3} = 284\ (2)$$

Photo Name	PSNR (dB)			Original Photo Size (kB)
	cheng2020-anchor	mbt2018-mean	bmshj2018-hyperprior	
cityhall	26,8	25,8	25,8	133
coffee	32,6	31,9	31,7	52
hike	23,5	23,0	23,0	141
gym	27,6	26,3	26,3	147
racoon	28,9	28,7	28,6	80
nba	33,4	32,2	31,9	61
sea	27,7	26,9	26,8	98
Average	28,6	27,8	27,7	102

Table 3 Peak Signal-to-Noise ratio of the test images using AI.

The cheng2020-anchor model outperforms two other models in PSNR metrics with average value of 28.6, compared to 27.8 and 27.7. This indicates that cheng2020-anchor compressed images were more similar to the original image in quality.

It is seen that “coffee”, “nba” and “hike” images deviate from the average value. This is due to the reasons described in MSE table because PSNR and MSE values inversely relate to each other. However, PSNR is a logarithmic measure with base point 10, meaning that an increase in 1 is equal to the increase by factor 10. This means that cheng2020-anchor introduces higher image quality relative to the noise.

Photo Name	**MS-SSIM (dB)**			**Original Photo Size (kB)**
	cheng2020-anchor	mbt2018-mean	bmshj2018-hyperprior	
cityhall	0,967	0,961	0,962	133
coffee	0,966	0,964	0,964	52
hike	0,911	0,900	0,901	141
gym	0,954	0,947	0,947	147
racoon	0,920	0,918	0,917	80
nba	0,981	0,979	0,978	61
sea	0,943	0,935	0,935	98
Average	0,949	0,943	0,943	102

Table 4 Multi-Scale Structure Similarity Index Method metrics of the test images using AI.

In the multi-scale metrics cheng2020-anchor has the best result of (0.949) compared to (0.943) in both mbt2018-mean and bmshj2018-hyperprior. This shows that cheng2020-anchor model retains the best quality in accordance to luminosity and contrast of pixels.

Metrics of Traditional compression methods

Photo Name	Space Saving (%)		Original Photo Size (kB)
	JPEG	HEIF	
cityhall	38,0	61,4	133
coffee	26,9	30,5	52
hike	22,4	33,3	141
gym	46,0	59,0	147
racoon	35,0	50,6	80
nba	33,0	47,4	61
sea	51,0	60,1	98
Average	36,0	48,9	102

Table 5 Space Savings of the test images using traditional compression.

It is seen that HEIF method provide more efficient space saving option for the image compression, performing on average 26% better. It compresses both smaller and larger images better.

Photo Name	MSE		Original Photo Size (kB)
	JPEG	HEIF	
cityhall	15,8	8,20	133
coffee	4,98	3,70	52
hike	32,8	25,9	141
gym	12,2	8,80	147
racoon	15,8	10,1	80
nba	3,57	2,87	61
sea	2,09	2,01	98
Average	12,5	8,80	102

Table 6 Mean Squared Error of the test images using traditional compression.

From the perspective of MSE metrics, both standards perform decently, however, more novel HEIF method has better results across all images.

Photo Name	PSNR (dB)		Original Photo Size (kB)
	JPEG	HEIF	
cityhall	36,1	36,8	133
coffee	41,2	41,7	52
hike	33,0	34,5	141
gym	37,3	37,6	147
racoon	36,1	36,9	80
nba	42,6	43,9	61
sea	44,9	45,0	98
Average	38,7	39,5	102

Table 7 Peak Signal-to-Noise ratio of the test images using traditional compression.

Since MSE and PSNR are mathematically related it is not a surprise that HEIF will provide more identical image that maintains more quality even after a higher compression.

Photo Name	MS-SSIM (dB)		Original Photo Size (kB)
	JPEG	HEIF	
cityhall	0,986	0,989	133
coffee	0,981	0,987	52
hike	0,965	0,973	141
gym	0,984	0,986	147
racoon	0,951	0,957	80
nba	0,992	0,995	61
sea	0,986	0,991	98
Average	0,978	0,983	102

Table 8 Multi-Scale Structure Similarity Index Method metrics of the test images using traditional compression.

MS-SSIM metrics also shows positive and better results for HEIF. Even though the results are decent for both methods, HEIF performs better.

6 Data analysis

AI-based methods

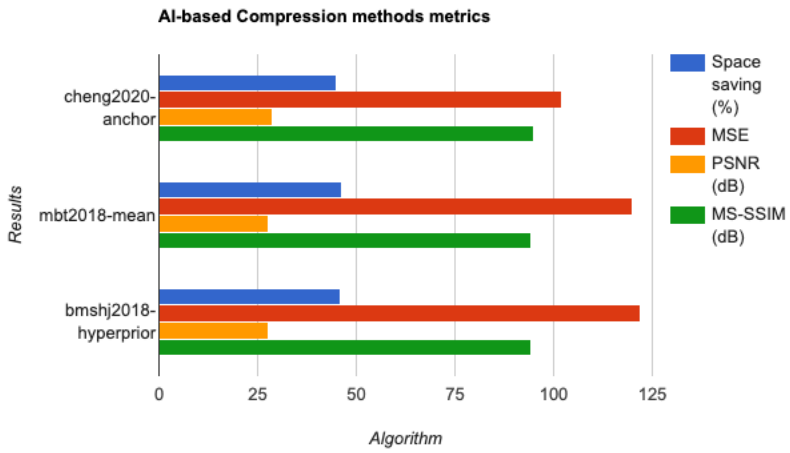

Figure 12 AI compression metrics results

Summarizing the results gathered using the quantitative experimentation, it is possible to state that the cheng2020-anchor model has worse space saving results, yet the difference of this metrics between models is insignificant. However, according to other 3 metrics that evaluates the similarity of compressed image and original image, this model performs the best in all three metrics. This indicates that cheng2020-anchor architect retains the quality of the original image most efficiently with a negligible tradeoff of space saving efficiency.

Apart from that, despite different techniques used to develop the mbt2018-mean and bmshj2018-hyperprior models, they performed equally in most all cases

with slight differences in some metrics. Thus, these models present an equal tradeoff between image compression and quality maintaining.

Some images like “hike”, “nba”, and “coffee” demonstrated an interesting correlation between the impact of original file size on image similarity. According to findings in the table 1, with higher original image size and details, the MSE between compressed image and original increases, which suggests that AI-based models have limitations when processing a high detailed picture. Conversely, the “coffee” and “nba” images, having smaller sizes had also lower MSE between original and compressed images (Figure 2). This indicates that with lower file size, AI is less likely to make an error, which is why the absolute error is smaller. Similarly, SM-SSIM metrics shows that compressed image quality is higher in low detailed pictures than in high detailed pictures.

Based on qualitative analysis, it is possible to emphasize that in some cases AI developed compressions distorted image and blurred (Figure 2) the picture because of which some objects or details worsened in quality. This is especially seen on the details presented in images or background objects that are heavily smoothed out. Despite this, for the objects with less details (Figure 3), the distortion was minimal and quality acceptable.

Traditional methods

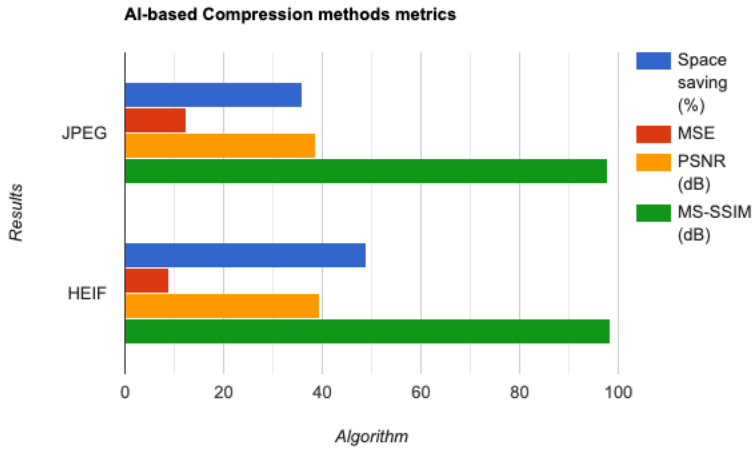

Figure 13 Traditional compression methods metrics results

After close analysis of the data, it was revealed that HEIF format outperforms JPEG standard in all qualitative metrics that was provided. Not only it compresses data more efficiently, but it also retains better quality of the compressed image. This is mainly because HEIF is a relatively new standard which was only applied 26 years later after the JPEG. However, one aspect where JPEG wins is in popularity and compatibility on all devices. Unlike HEIF format that was developed specially for the IOS and Safari systems, JPEG is much more accessible and used world widely.

Apart from that, the qualitative analysis showed that observers saw little to no difference between HEIF and JPEG compressions, indicating that for a person both images for compression look alike. This in fact, favors the HEIF method which has higher space saving efficiency despite similar visionary appeal.

AI-based and Traditional methods

Combining both qualitative and quantitative analysis of different compression methods, it is possible to state that for now the traditional methods perform generally better than the recently emerged AI-based models.

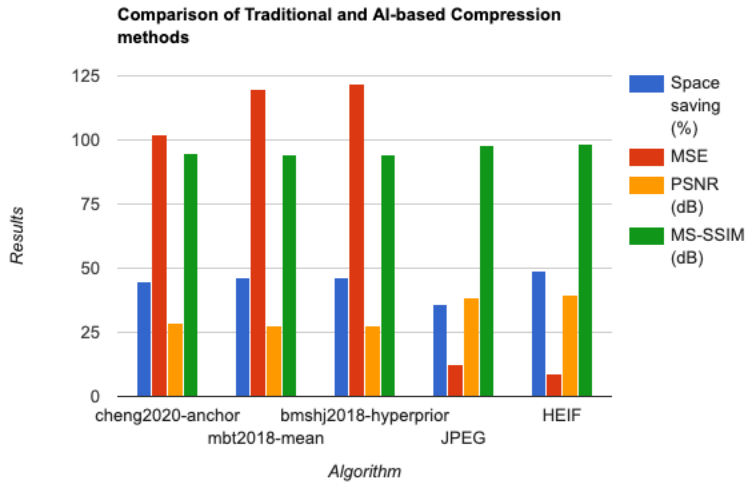

Figure 14 Representation of obtained results from different metrics

The mean squared error of ML compressions methods was significantly higher than the results obtained for traditional methods. The best result for AI-based was 102 performed by cheng2020-anchor model, while the best number was 8.80 for traditional HEIF technique. Since PSNR and MSE are related, this metrics was higher for JPEG and HEIF, indicating more efficient image quality conservation compared to cheng2020-anchor, mbt2018-mean, and mbshj2018-hyperprior. The MS-SSIM metrics also has shown the more efficient quality retaining for traditional methods. According to these metrics, traditional methods conserve the image

quality much better than the three presented models do. It shows that after compression with traditional methods, new images will look more like the original image, while the AI-based models tend to change images to a higher degree. This is also testified by the comparison of the “nba” images, as AI models tended to blur the details, while JPEG and HEIF maintained the quality of details.

On the other hand, most efficient cheng-2020anchor machine learning architecture outperforms the JPEG compressions standard in the space saving dimension (44.8% and 36%). This suggests that AI compression method gives higher space saving options for the cost of lower image quality. In fact, all models perform better than JPEG if space saving is compared. Thus, JPEG presents lower space saving, but higher compressed image quality in comparison to novel compression methods.

Despite slight excellence of AI model compared to JPEG, HEIF method provides option which provides a good space saving and quality retaining. Apple adopted method has higher results in both qualitative and quantitative metrics, indicating is dominance in visual quality and space saving efficiency.

However, the fact that AI models save more space that JPEG and almost as much as HEIF suggest that this novel approach has promising future. In this research these models performed not as good as in other researches because cheng2020-anchor, mbt2018-mean, and mbshj2018-hyperprior are pretrained models that

were not trained on specific datasets, while the images that were evaluated were diverse. In fact, in the work of Zhou et al. 2023 “0.922 in SSIM and 0.949 in correlation, regarding the compression ratio of 97.74 %” was achieved using the cheng2020-anchor method. Similarly, in the other work of Zafari et al. 2023, the AI model performed better than the traditional methods. It is mainly because in these works, the models were specifically trained for their task – compressing microscopy or astronomical images. This research, however, did not trained on a specific data set, but instead used pre-trained models for evaluation. According to these findings, AI has a decent potential when trained for a specific data set, however it is not as efficient for the general use. It will take a lot of time and data to train a model that will be able to perform decently on all kinds of photos, that is why the traditional methods still have better results than the new way of compressing data.

7 Limitations

In this study the reference metrics, were mainly used. Reference metrics refer to the comparison of compressed image to original image. However, as it was seen in the comparison of HEIF and JPEG compression methods, the visual difference between images is negligible, yet reference metrics (MSE, MS-SSIM, PSNR) suggest that HEIF performs better. Thus, the problem arises where full-reference metrics do not always demonstrate the visual similarity of images as it is seen for a human eye or “the effectiveness is not clear” (Ma et al., 2016). Future research should include the new non-reference evaluation metrics such as Universal Image

Quality Index (UIQI), which presents an efficient way to assess the quality of images based on human perception (Zhou Wang & Bovik, 2002).

8 Conclusion

In this research, the popular traditional compressions methods as well as efficient AI-based compression methods are evaluated and discussed using quantitative and qualitative analysis metrics. It was found that AI-based compression methods can outperform the classical compression methods if trained correctly and applied for specific needs. However, traditional methods have in general cases better performance compared to new models.

Hopefully, this paper will shed light on the possible application of novel compression method for better data storage and transmission, leading to more innovative approaches and development of AI image compression field as a new field of study.

9 Bibliography

Ballé, J., Minnen, D., Singh, S., Hwang, S. J., & Johnston, N. (2018). *Variational image compression with a scale hyperprior* (arXiv:1802.01436). arXiv. http://arxiv.org/abs/1802.01436

Bégaint, J., Racapé, F., Feltman, S., & Pushparaja, A. (2020). *CompressAI: A PyTorch library and evaluation platform for end-to-end compression research* (arXiv:2011.03029). arXiv. http://arxiv.org/abs/2011.03029

Cheng, Z., Sun, H., Takeuchi, M., & Katto, J. (2020). *Learned Image Compression with Discretized Gaussian Mixture Likelihoods and Attention Modules* (arXiv:2001.01568). arXiv. http://arxiv.org/abs/2001.01568

HEVC: An introduction to high efficiency coding. (n.d.). Vcodex. https://www.vcodex.com/hevc-an-introduction-to-high-efficiency coding/#:~:text=HEVC%20supports%20highly%20flexible%20partitioning,is%20transformed%20and%20entropy%20encoded.

How JPEG works (Christopher G. Jennings). (n.d.). https://cgjennings.ca/articles/jpeg-compression/

Improvement of power consumption efficiency for advanced AI processing using deep learning compression technology | Challenge Zero. (n.d.). Improvement of Power Consumption Efficiency for Advanced AI Processing Using Deep Learning Compression Technology | Challenge Zero. https://www.challenge-zero.jp/en/casestudy/350

Jamil, S., Piran, M. J., & MuhibUrRahman. (2022). *Learning-Driven Lossy Image Compression; A Comprehensive Survey* (arXiv:2201.09240). arXiv. http://arxiv.org/abs/2201.09240

"jpeg" | Can I use. . . Support tables for HTML5, CSS3, etc. (n.d.). https://caniuse.com/?search=jpeg

JPEG: Joint Photographic Experts Group. (n.d.). https://www.cs.auckland.ac.nz/compsci708s1c/lectures/jpeg_mpeg/jpeg.html

skimage.metrics — skimage 0.22.0 documentation. (n.d.). https://scikit-image.org/docs/stable/api/skimage.metrics.html#skimage.metrics.mean_squared_error

Kunwar, S. (2023). *Comprehensive Image Quality Assessment (IQA) of JPEG, WebP, HEIF and AVIF Formats*. https://doi.org/10.13140/RG.2.2.31276.03202

Lossy vs Lossless Compression: Differences & Advantages | Adobe. (n.d.). https://www.adobe.com/uk/creativecloud/photography/discover/lossy-vs-lossless.html

Ma, C., Yang, C.-Y., Yang, X., & Yang, M.-H. (2016). *Learning a No-Reference Quality Metric for Single-Image Super-Resolution* (arXiv:1612.05890). arXiv. http://arxiv.org/abs/1612.05890

Minnen, D., Ballé, J., & Toderici, G. (2018). *Joint Autoregressive and Hierarchical Priors for Learned Image Compression* (arXiv:1809.02736). arXiv. http://arxiv.org/abs/1809.02736

Number of worldwide social network users 2027 | Statista. (2023, August 29). Statista. https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/

Sara, U., Akter, M., & Uddin, M. S. (2019). Image Quality Assessment through FSIM, SSIM, MSE and PSNR—A Comparative Study. *Journal of Computer and Communications*, *07*(03), 8–18. https://doi.org/10.4236/jcc.2019.73002

Sheldon, R. (2022, July 12). *image compression*. WhatIs. https://www.techtarget.com/whatis/definition/image-compression

Søgaard, J., Krasula, L., Shahid, M., Temel, D., Brunnström, K., & Razaak, M. (2016). Applicability of Existing Objective Metrics of Perceptual Quality for Adaptive Video Streaming. *Electronic Imaging*, *28*(13), 1–7. https://doi.org/10.2352/ISSN.2470-1173.2016.13.IQSP-206

Usage statistics of image file formats for Websites, January 2024. (n.d.). https://w3techs.com/technologies/overview/image_format

What is Industry 4.0 and how does it work? | IBM. (n.d.). https://www.ibm.com/topics/industry-4-0

Wikiwand - Data compression ratio. (n.d.). Wikiwand. https://www.wikiwand.com/en/Data_compression_ratio

Wikiwand - High Efficiency Image File Format. (n.d.). Wikiwand. https://www.wikiwand.com/en/HEIF

Zafari, A., Khoshkhahtinat, A., Grajeda, J. A., Mehta, P. M., Nasrabadi, N. M., Boucheron, L. E., Thompson, B. J., Kirk, M. S. F., & da Silva, D. (2023). Neural-based Compression Scheme for Solar Image Data. *IEEE Transactions on*

Aerospace and Electronic Systems, 1–18. https://doi.org/10.1109/TAES.2023.3332056

Zhou Wang, & Bovik, A. C. (2002). A universal image quality index. *IEEE Signal Processing Letters*, *9*(3), 81–84. https://doi.org/10.1109/97.995823

Zhou, Y., Sollman, J., & Chen, J. (2023). *Deep learning based Image Compression for Microscopy Images: An Empirical Study* (arXiv:2311.01352). arXiv. http://arxiv.org/abs/2311.01352

Wikiwand - Structural similarity. (n.d.). Wikiwand. https://www.wikiwand.com/en/SSIM

Appendix

Category Name	Sample image	Link
cityhall		https://savannahswaterfront,com/wp-content/uploads/2022/02/WaterfrontAssoc-CaseyJones-1-City-hall-1-768x512,jpg
coffee		https://volshebnaya-eda,ru/wp-content/uploads/2020/10/%D0%BB%D0%B0%D1%82%D1%82%D0%B5-768x512-1,jpg
hike		https://www,starvedrocklodge,com/wp-content/uploads/2021/04/guided_hikes_2400x1600-768x512,jpg
gym		https://oxyfitness,kz/wp-content/uploads/2023/08/fr-768x512,jpg

racoon		https://uhnovgrad,ru/wp-content/uploads/2022/12/1234-768x512-1,jpg
nba		https://www,nbamaniacs,com/wp-content/uploads/2023/08/GettyImages-1469015642-768x512,jpg
sea		https://exploresouthwales,com/explore/oxwich-beach/

Appendix 1 Sample images for each of the category