JPEG2000-Based Semantic Image Compression using CNN

Original Scientific Paper

Anish Nagarsenker
Dr Vishwanath Karad MIT World Peace University,
School of Electronics and Communication Engineering
Kothrud, Pune, India
anishnagarsenker@gmail.com

Prasad Khandekar
Dr Vishwanath Karad MIT World Peace University,
School of Electronics and Communication Engineering
Kothrud, Pune, India
prasad.khandekar@mitwpu.edu.in

Minal Deshmukh
Vishwakarma Institute of Information and Technology,
Department of Electronics and Telecommunication Engineering
Pune, India
minal.deshmukh@viit.ac.in

Abstract – *Some of the computer vision applications such as understanding, recognition as well as image processing are some areas where AI techniques like convolutional neural network (CNN) have attained great success. AI techniques are not very frequently used in applications like image compression which are a part of low-level vision applications. Intensifying the visual quality of the lossy video/image compression has been a huge obstacle for a very long time. Image processing tasks and image recognition can be addressed with the application of deep learning CNNs as a result of the availability of large training datasets and the recent advances in computing power. This paper consists of a CNN-based novel compression framework comprising of Compact CNN (ComCNN) and Reconstruction CNN (RecCNN) where they are trained concurrently and ideally consolidated into a compression framework, along with MS-ROI (Multi Structure-Region of Interest) mapping which highlights the semiotically notable portions of the image. The framework attains a mean PSNR value of 32.9dB, achieving a gain of 3.52dB and attains mean SSIM value of 0.9262, achieving a gain of 0.0723dB over the other methods when compared using the 6 main test images. Experimental results in the proposed study validate that the architecture substantially surpasses image compression frameworks, that utilized deblocking or denoising post-processing techniques, classified utilizing Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index Measures (SSIM) with a mean PSNR, SSIM and Compression Ratio of 38.45, 0.9602 and 1.75x respectively for the 50 test images, thus obtaining state-of-art performance for Quality Factor (QF)=5.*

Keywords: *Computer Vision, Neural Networks, CNN, MS-ROI, Compression, JPEG2000*

1. INTRODUCTION

Image Compression is a form of data compression procedure that is used on digital images to reduce their size as well as cost and make them more suitable for storage or communication. Properties of Image Compression include Scalability (Component Progressive, Quality Progressive, and Resolution Progressive), Meta Information, Interest Coding, and Processing Power [1]. Lossy image compression standards attain a greater compression ratio in contrast to lossless image compression standards, but there is a loss of image quality. Lossless image compression is mainly used in medical imaging, clip arts, technical drawings, or comics whereas Lossy image compression maybe used for natural images such as photographs [2, 3]. A subset of Machine Learning, Deep Learning algorithms aim to extract identical inferences as humans would, by continuously analyzing data with the specified analytical structure, and hence, the multiple-layered structures of algorithms are employed called neural networks [4].

In Deep learning, a class of ANN that is usually applied to image processing and image recognition to analyze visual imagery is called as Convolutional Neural Network (CNN). Some advantages of using neural networks include (a) producing output irrespective of the input provided to them because they learn by themselves (b) performing multiple tasks concurrently without affecting system performance (c) instead of a database, the input is stored in its own network and hence loss of data does not affect its working (d) it works through real-time events because it learns from events and acts accordingly when a similar event arises (e) ability of the network to detect the fault by themselves and produce the output accordingly[5, 6]. Numerous layers exist in CNN like fully connected layers, integration layers, intermittent layers, and convolutional layers whereby intermittent layers and integration layers have no variables associated with them whereas fully-connected layers and convolutional layers have variables associated with them [7].

ComCNN gains an excellent compact characterization from the image at input hence conserves the structural details of the image [8]. RecCNN has been utilized to rebuild the image with excellent quality, thus producing an output image of high quality and maintaining all the structural details of the input image [9]. The process of portioning every pixel in an image in its region that has semantic value with a specific label is called as semantic segmentation. MS-ROI mapping is a form of region-based semantic segmentation that can be used to successfully prime on localization functions irrespective of the amount of classes. MS-ROI mapping is used to represent numerous prominent areas of a picture in one pass and hence can be used to boost the visual attributes of only the semantic objects, thus making the image visually attractive [10].

PSNR and SSIM are two parameters that are widely used for the assessment of picture quality. The ratio of the highest plausible power from the signal provided to the highest viable power of the distorted noise which impacts the accuracy of characterization is called PSNR. In lossy compression, PSNR is widely used to check the restoration standard of videos and images, where PSNR is inversely proportional to Mean Squared Error (MSE). A perceptual metric that measures image caliber variation generated from rectification like data transmission or data compression losses is known as Structural Similarity Index(SSIM). SSIM values range between 0 and 1 whereas PSNR values range between 0 and ∞ [11].

In [9], a framework is proposed comprising of ComCNN and RecCNN, that are ideally consolidated into a compression framework to achieve image compression of superlative quality at comparatively lower bit rates using JPEG, giving high values of PSNR and SSIM at QF=5 and QF=10. In [10], using MSROI heat map, semantically salient regions of the image are highlighted using CNN and these are then encrypted at a higher caliber than the background portions of the image using JPEG giving significantly high values of PSNR and SSIM. The framework utilised in this paper, is a combination of the methods used in these 2 papers giving a novel framework standard for Image Compression.

In this paper, a framework is proposed comprising 2 CNNs that are ideally consolidated into a compression framework along with MS-ROI (Multi Structure-Region of Interest) to achieve image compression of the highest quality at comparatively lower bit rates using JPEG2000 giving high values of PSNR and SSIM. The structural information is maintained by the Compact CNN (1st CNN) which grasps excellent compact characterization from the given image followed by which an MS-ROI map is generated where semantically salient regions are highlighted and passed through the JPEG2000 encoder. The decrypted image is reconstructed with high standards using Reconstruction CNN (2nd CNN) to obtain the final reconstructed image after compression. Concurrently learning ComCNN and RecCNN eases the precise reformation of the decrypted image using RecCNN [9]. When tested using 50 test images, this framework is able to achieve a mean PSNR, SSIM and Compression Ratio of 38.45, 0.9602 and 1.75x respectively, thus outperforming other image enhancement methods for QF=5. The following sections give a brief overview of the methods employed, methodology performed, results obtained, and conclusions procured from the results.

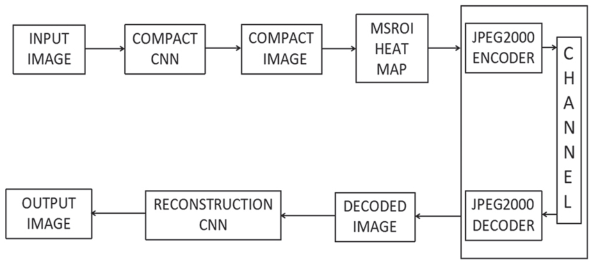

Fig. 1. Proposed Block Diagram

2. COMPACT CNN(COM-CNN) AND RECONSTRUCTION CNN(REC-CNN)

CNNs, which are composite feed-forward neural structures, because of their immense accuracy, are used in image recognition and classification, Segmentation and Object Detection [12]. The CNN follows a progressive model which deals with building an organization and finally provides a fully-associated layer where the neurons are associated with one another and the output is extracted. CNNs consist of the Convolutional layer in which filters are utilized for the input image or attribute maps. The integration layer is utilized to lower the scale of the attribute map and the Fully-Connected layer which are positioned prior to classification yield and is utilized to normalize the outcomes before classification [11].

ComCNN gains an excellent compact characterization from the image at input hence conserves the structural details of the image. ComCNN normally consists of 3 layers compared to RecCNN which contains a higher number of layers (Around 20 layers) out of which, most layers undergo the process of convolution plus batch normalization plus ReLU. As such, RecCNN has been utilized to rebuild the image with excellent quality, thus producing an output image of high quality and maintaining all the structural details of the input image. ComCNN and RecCNN attempt to build the reconstructed image identical to the input image. Training both the CNNs concurrently yields better results as compared to training the CNNs individually [9].

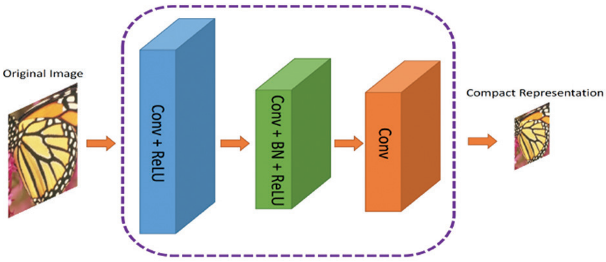

Fig. 2. ComCNN Architecture

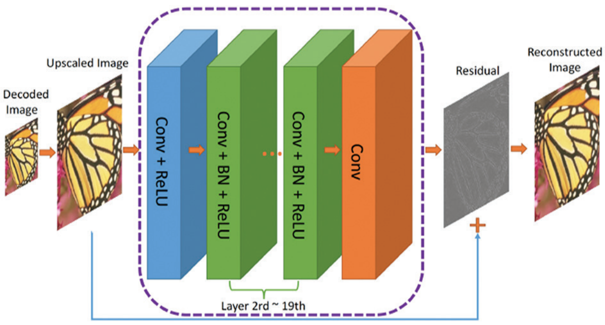

Fig. 3. RecCNN Architecture

3. MULTI-STRUCTURE REGION OF INTEREST (MS-ROI)

The process of portioning every pixel in an image in its region that has a semantic value with a specific label is called semantic segmentation. MS-ROI heat map is a method that stabilizes the activation for numerous entities and hence doesn't agonize from the concern of global average pooling. Unconstrained by the amount of classes, this method permits to productively train on localization tasks. MS-ROI is able to spot and roughly detect all the objects in the given compact image. In this framework, a group of 3D attribute plots is preferred in which each attribute plot is grasped for a discrete domain and is grasped autonomously for the plots for auxiliary domains. In MS-ROI, the number of classes is reduced by combining homogenous types of classes into a broader class as the aim is to associate all the objects present in the image. The plot created will be practically indistinguishable till the time the entities of these merged domains have homogenous shapes and are inside the identical universal type. By utilizing the MS-ROI heat map technique, similarity merit for each element in the domain [0,1] is procured in which 1 specifies maximal similarity [3, 10]. This technique is able to highlight the semantically notable areas of the image which are then encrypted at a superior grade in contrast to the other less notable areas. Therefore, MS-ROI mapping is used to represent numerous prominent areas of a picture in one pass and hence can be used to increase the visual aspect of only the semantic objects in the picture, thus making the image visually attractive [10].

Fig. 4. MSROI Heat Map Examples

4. JPEG2000

JPEG 2000, invented by the Joint Photographic Experts Group, is an international-grade compression technique established on the concept of wavelet transform and provides an exceptionally superior level of robustness and accessibility. The prime advantages of utilizing this approach comprise intensified transmission and locale of interest coding. It is a strong and suitable mechanism that is specifically used in RS picture processing and capacity. It is an image coding system that provides lossy image compression for storage but can maintain an image quality similar to that of the original image [13].

In the JPEG-2000 method, the input image is given to a group of discrete wavelet filters which converts the pixel details into wavelet coefficients, which are then assembled into various bands. Every band in the image carries the wavelet coefficients that portray a discrete level and sharp structural frequency domain of the complete input image. These bands are then given for quantization where the bit streams, accommodated in each code block, are pruned (re-quantized). These bit streams are impeccably pruned utilizing a method called as PCRC (post-compression rate control). The code blocks are assembled conforming to their significance where the information related to the magnitude of bits is determined by a process called context modeling [13, 14].

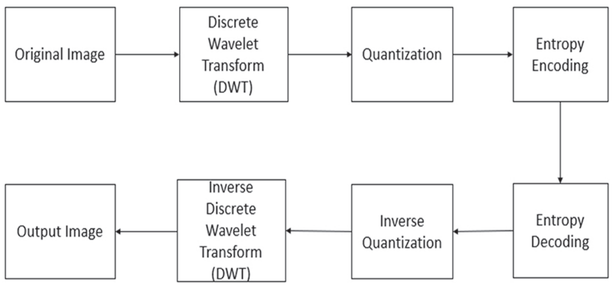

Fig. 5. JPEG2000 Encoder and Decoder

In the decoder section, the bit streams are given for entropy decoding followed by inverse quantization, where the bits are decoded according to their significance. These are then driven through a group of inverse discrete wavelet filters where the wavelet coefficients are transformed back to pixels. The decoded image is obtained which contains pixels according to their significance where the most important contents are observed earlier than the other contents. The final image obtained is a compressed variant of the actual image with very little loss in image quality [15].

5. METHODOLOGY

Image Acquisition

In this phase, an input image is given to the proposed model. The pre-processed and normalized image should be either RGB or Grayscale. The complete dataset is split into 2 sets (training and testing) with the 70-30% data splitting or 10-fold cross-validation method.

Compact CNN

This compact solution minimises the number and sophistication of lower-level kernels in a CNN by separating the colored information of the input picture. It is presented a compact construction of fully convolutional neural networks including worldwide color histogram features. The design proposed decreases the number of processors required to extract high-level given grayscale input. The compact structure has 40% fewer parameters to tune and performs comparably to the standard CNN formed on ROB pictures. In the first layer, two CNNs with variable amounts of kernels are trained: an original title and a compact counterpart. The 'original' network is a carbon copy of CNN. When providing the final results for both the regular Network and the compressed CNN, however, we only use single-view testing. Both the traditional and compact CNNs include four convolution layers. The condensed CNN's first stratum has just 32 kernels, while the traditional CNN's first layer contains 64 kernels. This cuts the number of variables in layer 3 (the intermediate convolution layer) by half, resulting in a compacted CNN with 40% lower dimensionality to tweak than the previous form. The concatenation is denoted as:

$$y^{j(r)} = \mathrm{ReLU}(b^{j(r)} + \textstyle\sum_i k^{ij(r)} * x^{i(r)}) \quad (1)$$

MSROI

The objective of ROI (Region of Interest) involves isolating a section of an image that we want to refine or manipulate. To start with, a binary mask is generated which is basically a binary image, the same size as the image we want to examine, with all other pixels set to 0 and only the ones forming the ROI set to 1. Preprocessing (Edge detection), Polygonal approximation, and Shape signature extraction are all factors to consider while extracting MS-ROI from a compact picture. The steps pursued in the implementation of MS-ROI include defining a group of boundary pixels analogous to the ROI, converting every pixel in the set to a parametric curve in the parameter space, supplementing the pixels in the accumulator matrix 'A' as guided by the parameter curve, evaluating the provisional feature vector, vector mean, and feature vector normalization.

JPEG Encoder and Decoder

The JPEG encoder and decoder provide an inbuilt library to achieve the functionality of compression. The steps followed in JPEG encoding and decoding include taking the picture to encode, getting the resulting image encoded and decoded by the program, choosing the quantization coefficient and quantization matrix for luminance, viewing the details of the block (square of 64 pixels) and size of the image and then, passing on the resulting image to the RecCNN.

Reconstruction CNN

Using a revised DAS algorithm (back projection-based) D: Rm Rn, the technique used first computes a low-graded appraisal of x, indicated by $\bar{x}$ using measurement recorded with a single graphical representation. We define DWH, where W:RnRn is the ith transmitter's "pixel-wise" re-evaluation operator (diagonal matrix).

$$w_{i(r)} = \left(\overline{h_i^{tx}}(r) \sum_{j=1}^{n_r} \overline{h_j^{rx}}(r) \right)^{-1} \quad (2)$$

As post-DAS picture equalisation, it compensates for the amplitude-related impacts of far-field diffraction. A CNN f: RnRn, trained to retrieve a high-quality estimate of x as $x=f(x)$, is given the approximation $x=Dy$ in the second stage. The use of a CNN for restoration jobs is recommended over employing end-to-end techniques that attempt to directly map a measurement space to an image space. We want to train a mapping f from a low-quality picture subspace $W\ Rn$ to a high-quality image subspace $V\ Rn$. To define these subspaces more accurately, a transducer with a specific aperture is used, made of an array of transducer components with a given shape, center frequency, and bandwidth. To reassemble US pictures using D, we define W as the subspace of measurements recorded by a single in sonification utilising the whole aperture. High SL and EW artefacts, as well as probable GL artefacts, may be seen in these photos (e.g., linear-array designs). In order to avoid GLs, a transducer with the same aperture and physical parameters as the one used for W is suggested for V. To create reference pictures from this array, we use the D operator for each in sonification and coherent compounding. SA is commonly regarded as the gold standard, generating pictures with great resolution while reducing SL and EW artefacts.

Moreover, SA photographs have better speckle patterns than "exotic" ground truth images. We ensure that both subspaces of the CNN include speckle patterns from sub-resolution scattering interferences. Due to the assumptions made in the physical measurement technique (2) and the following back projection operator (3) used to determine W and V, the trained CNN is not expected to repair artefacts from ignored physical events. With increased resolution, diffraction artefacts are reduced yet speckle is retained.

6. RESULTS AND DISCUSSION

The execution of the proposed algorithm is assessed using comparisons between the SSIM and PSNR of various post-processing denoising methods as well as deblocking methods for images such as Zhangs's, DicTv, Ren's, Sun's, WNNM, ARCNN and BM3D because of their excellent performance in producing output images. Also, the ComCNN and RecCNN are individually trained as well as trained simultaneously for comparison with the PSNR and SSIM values obtained using the proposed framework. The values of PSNR and SSIM obtained from the combination of ComCNN and RecCNN together, outperforms the values of PSNR and SSIM obtained using the 2 CNNs individually as seen from Tables I and II [9]. This result showcases the importance of training the 2 CNNs simultaneously and using them along with MS-ROI heat map, which represents numerous prominent areas of a picture in one pass and hence can be used to increase visual aspect of only the semantic objects in the picture, making the image visually attractive and thus giving excellent values of PSNR as well as SSIM for QF=5. In the experiments performed, the 2 CNNS (ComCNN and RecCNN) are trained simultaneously using a set of approximately 1000 images (256 x 256) of various different types like vehicles, people, scenery, plants, animals etc. and is trained for 200 epochs. The experiments are executed in Python Environment (Anaconda) on Laptop having Intel(R) Core (TM) i5-10210U CPU @1.60GHz 2.11 GHz and an Nvidia GeForce MX230 GPU. Around 12-15 hours of PC time is required for training the CNNs simultaneously up to 200 epochs on GPU [9].

In the experiments performed, 6 test images as that used in the paper [9] are used, so that the comparison becomes easier with all the Image Enhancement methods. From the PSNR and SSIM table for QF=5, it can be concluded that the framework attains PSNR values in the domain 31.28-35.26dB with a mean PSNR value of 32.9dB, achieving a gain of 3.52dB and achieves SSIM values in the domain 0.9065-0.9538dB with a mean SSIM value of 0.9262, achieving a gain of 0.0723dB over the "ComCNN+RecCNN [9]" method which gives optimal outputs of PSNR and SSIM in comparison to every method mentioned in Table 1 and 2. The framework proposed eclipses each and every image intensification method including the ARCNN [20], which is considered the landmark of CNN. The framework proposed not only eliminates a bulk of the antiquities notably but also maintains more attributes on the texture as well as the edges. This framework is able to achieve additional high-frequency details, retrieve salient edges as well as perfect textures in the reproduced image.

A total of 50 test images were used and some of the output images along with their outputs are displayed in Fig. 9 and Table 3 respectively. Various test images with QF=5 are shown in Table III and their corresponding input size (in KB), output size (in KB), Compression Ratio, PSNR, and SSIM are noted down. For these test images, the Compression Ratio values range between 1.23x-3.15x with an average compression rate of 1.62x, the PSNR values range between 31.28-45.37dB with an average PSNR of 37.26dB while the SSIM values range between 0.8889-0.9915dB with an average SSIM of 0.9556dB. These values of PSNR and SSIM suggest that the framework proposed for compression undergoes very minute changes from the input image to the output image with a considerable compression rate and hence, an output image is obtained which is very similar to the input image.

As can be seen from the pictures displayed in Fig. 6 and Fig. 9, there is not much visual difference observed between the input and the output images (the output image is the same as the input image) with a significant amount of compression achieved. This is because the MS-ROI heat map highlights the semantically notable areas of the image which are then encrypted at a superior grade in contrast to the other less notable areas and the 2 CNNs work simultaneously to rebuild

Table 1. Analogy of PSNR of the test inputs of various Image deblocking, denoising and compression methods for QF=5 with the framework proposed

TEST IMAGES	Parrots	Cameraman	Lena	Butterfly	House	Peppers	Mean
				PSNR			
JPEG	26.19	24.45	27.33	22.58	27.77	27.17	25.92
BM3D [11]	27.33	25.27	28.63	24.05	29.21	28.52	27.17
Ren's [10]	27.87	25.46	29.07	24.58	29.66	29.07	27.62
Zhang's [12]	27.78	25.39	29.00	24.20	29.24	29.07	27.45
Sun's [4]	27.45	25.25	28.87	23.83	29.09	29.05	27.26
DicTV [9]	26.83	24.54	28.07	23.10	28.45	27.95	26.49
Zhang's [13]	28.27	25.61	29.51	25.30	30.12	29.61	28.07
ARCNN [14]	28.13	25.27	29.31	25.64	29.68	29.02	27.84
WNNM [15]	27.80	25.49	28.95	24.75	29.62	28.99	27.60
ComCNN [5]	27.67	24.93	29.46	23.05	29.17	29.33	27.27
RecCNN [5]	28.52	26.33	29.63	25.99	30.13	29.81	28.40
ComCNN+ RecCNN [5]	30.12	26.53	31.14	26.23	31.45	30.84	29.38
Proposed	**32.58**	**33.19**	**33.28**	**31.28**	**31.79**	**35.26**	**32.90**

Table 2. Analogy of SSIM of the test inputs of various Image deblocking, denoising and compression methods for QF=5 with the framework proposed

TEST IMAGES	Parrots	Cameraman	Lena	Butterfly	House	Peppers	Mean
				SSIM			
JPEG	0.7581	0.7283	0.7367	0.7378	0.7733	0.7087	0.7404
BM3D [11]	0.8118	0.7607	0.7837	0.8184	0.8082	0.7639	0.7911
Ren's [10]	0.8310	0.7666	0.8010	0.8419	0.8197	0.7876	0.8080
Zhang's [12]	0.8308	0.7672	0.8035	0.8313	0.8141	0.7895	0.8061
Sun's [4]	0.8323	0.7687	0.8061	0.8321	0.8113	0.7931	0.8073
DicTV [9]	0.8005	0.6658	0.7744	0.7769	0.7963	0.7456	0.7599
Zhang's [13]	0.8460	0.7666	0.8169	0.8667	0.8285	0.8031	0.8213
ARCNN [14]	0.8446	0.7674	0.8142	0.8741	0.8209	0.7961	0.8196
WNNM [15]	0.8287	0.7674	0.7947	0.8445	0.8178	0.7827	0.8060
ComCNN [5]	0.8377	0.7662	0.8042	0.7488	0.8119	0.7966	0.7942
RecCNN [5]	0.8497	0.7945	0.8195	0.8760	0.8251	0.8004	0.8275
ComCNN+ RecCNN [5]	0.8951	0.8167	0.8486	0.8847	0.8456	0.8328	0.8539
Proposed	**0.9307**	**0.9065**	**0.9538**	**0.9155**	**0.9189**	**0.9315**	**0.9262**

Table 3. PSNR, SSIM and Compression Ratios of the various test images for QF=5

Test Inputs	Input Size(KB)	Output Size(KB)	Compression Ratio	PSNR	SSIM
Peppers	6.01	4.12	1.46x	35.26	0.9315
Parrots	6.20	4.32	1.44x	32.58	0.9307
House	4.56	3.69	1.24x	31.79	0.9189
Lena	7.81	5.39	1.45x	33.28	0.9538
Butterfly	10.2	5.40	1.89x	31.28	0.9155
Cameraman	6.81	4.66	1.46x	33.19	0.9065
TT1	760	499	1.52x	37.7	0.9813
TT2	771	472	1.63x	38.47	0.9905
TT3	333	138	2.41x	45.1	0.9828
TT4	194	155	1.25x	43.27	0.9915
TT5	711	550	1.29x	45.37	0.9875
TT6	1400	880	1.59x	35.96	0.9812
TT7	752	606	1.24x	38.81	0.9887
TT8	4530	1440	3.15x	37.98	0.8889
TT9	1540	1250	1.23x	38.82	0.9845
Mean	-	-	**1.62x**	**37.26**	**0.9556**

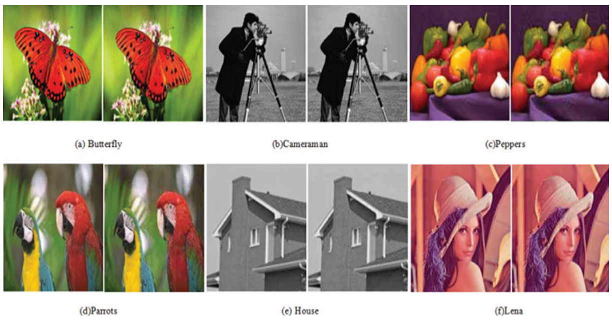

Fig. 6. Different Test Images and their Outputs

the original image, generating an image appearing similar to the original image, hence giving higher values of PSNR as well as SSIM. The greater the number of epochs used in the original image, hence giving higher values of PSNR as well as SSIM. The greater the number of epochs used in the training of CNNs concurrently, the better results the framework produces. Overall, a total of 50 test images were used to validate the quality of framework and the results show that this framework achieves a mean PSNR, SSIM and Compression Ratio of 38.45, 0.9602 and 1.75x respectively, thus proving to be an ideal framework for image compression.

The graphs in Fig. 7 show the mean PSNR and SSIM obtained for the 6 test inputs using various Image compression, deblocking, and denoising methods. From the graphs, it can be concluded that the mean PSNR and SSIM obtained for the 6 test inputs at QF=5 easily outperforms all the other image compression, deblocking, and denoising methods to achieve significantly high values of 32.90 and 0.9262 respectively.

The graphs in Fig. 8 show the PSNR, SSIM, and Compression Ratio values obtained for the various test images, used in this experiment of image compression using CNNs. From the graphs, it can be concluded that the mean PSNR of 37.26, mean SSIM of 0.9556, and an average compression of 1.62x for the test images at QF=5 make this framework ideal for use. The PSNR values obtained for the test images in Fig. 9 are greater than 30 while the SSIM values are greater than 0.88 which shows a significant improvement over other image compression as well as image deblocking and denoising methods. These graphs in Fig. 7 and 8 are direct evidence that the proposed framework gives an output image that undergoes very minute changes compared to the input image along with significant compression, achieved in the range of 1.18x-3.75x.

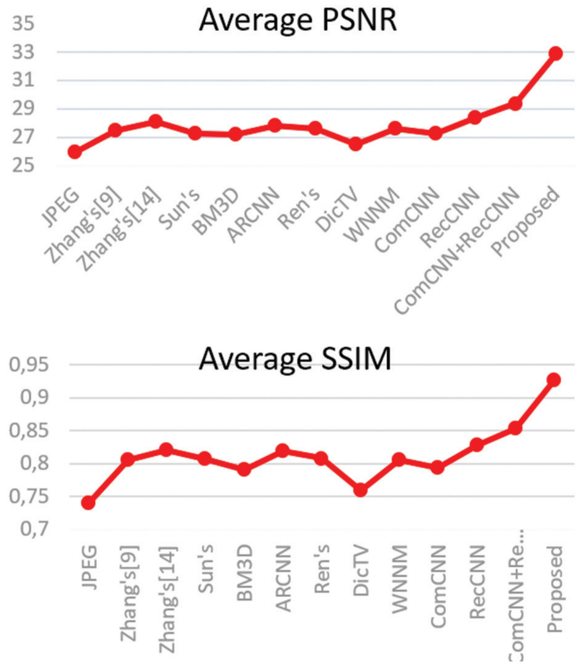

Fig. 7. Analogy of the mean PSNR and SSIM with various image compression, deblocking and denoising methods for QF=5

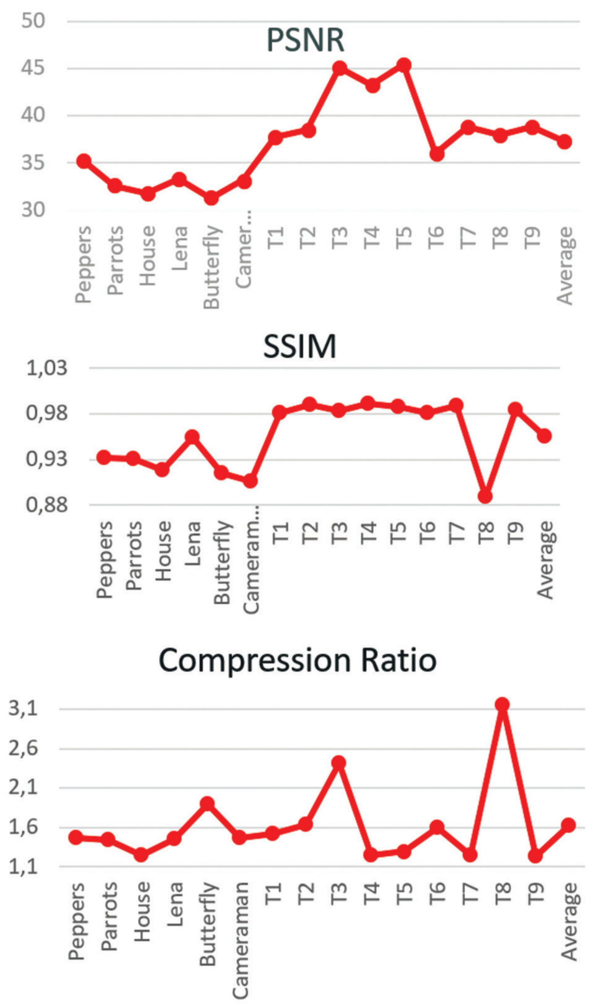

Fig. 8. Analogy of the mean PSNR, SSIM and Compression Ratio for the 15 Test Images for QF=5

7. CONCLUSION AND FUTURE WORK

The project consists of a novel compression framework involving CNN comprising of ComCNN and RecCNN where they are trained concurrently and ideally consolidated into a compression framework, along with MS-ROI mapping which highlights the semiotically notable portions of the image and JPEG2000 image codec for image compression. ComCNN gains an excellent compact characterization from the image at the input, hence conserving the structural details of the image, RecCNN has been utilized to rebuild the image with excellent quality, thus producing an output image of high quality and maintaining all the structural details of the input image whereas MS-ROI mapping is used to represent numerous prominent areas of a picture in one pass and hence can be used to boost the visual attributes of only the semantic objects, thus making the image visually attractive. The results of the experiments performed depict that this framework attains excellent values of PSNR and SSIM for QF=5, outperforming all the post-processing innovations, providing a suitable compression ratio of 1.18-3.75 times the original image and a mean PSNR and SSIM value obtained as 38.45 and 0.9602 respectively. For future work, the same framework can be used for examining the outputs (PSNR and SSIM) with different values of QF. Secondly, different architectures of CNNs can be tested for providing improved performances over the proposed method. Also, different image compression methods can be applied to this framework like JPEG, BPG, etc to check for their performance in comparison to the proposed JPEG2000 image.

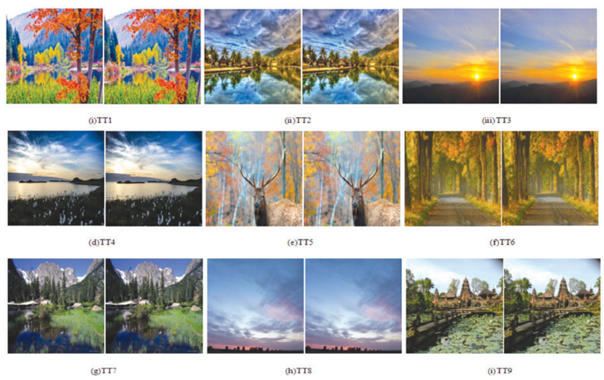

Fig. 9. Different Test Inputs And Their Outputs

8. REFERENCES

[1] D. Mishra, S. K. Singh, R. K. Singh, "Deep Architectures for Image Compression: A Critical Review", Signal Processing, Vol. 191, 2022.

[2] M. Gnacek, C. Heatwole, D. Fan, M. Hoffman, "Convolutional Neural Networks for Enhanced Compression Techniques", Proceedings of the IEEE National Aerospace and Electronics Conference, Dayton, OH, USA, 16-19 August 2021.

[3] Cheng Wang, Yifei Han, Weidong Wang, "An End-to-End Deep Learning Image Compression Framework Based on Semantic Analysis", Applied Sciences, Vol. 9, No. 17, 2019.

[4] Z. Zhang, G. Xu, J. Song, "CubeSat cloud detection based on JPEG2000 compression and deep learning", Advances in Mechanical Engineering, Vol. 10, No. 10, 2018.

[5] Z. Liu, T. Liu, W. Wen, L. Jiang, J. Xu, Y. Wang, G. Quan, "DeepN-JPEG: A Deep Neural Network Favorable JPEG-based Image Compression Framework", Proceedings of the 55th Design Automation Conference, San Francisco, CA, USA, 24-29 June 2018, pp. 1-6.

[6] O. H. Nagoor, J. Whittle, J. Deng, B. Mora, M. W. Jones, "Lossless Compression for Volumetric Medical Images using Deep Neural Network with Local Sampling", Proceedings of the IEEE International Conference on Image Processing, Abu Dhabi, UAE, 25-28 October 2020.

[7] D. Sun, W.-K. Cham, "Postprocessing of low bit-rate block DCT coded images based on a field of experts prior", IEEE Transactions on Image Processing, Vol. 16, No. 11, 2007, pp. 2743-2751.

[8] H. Fu, F. Liang, B. Lei, N. Bian, Q. Zhang, M. Akbari, J. Liang, C. Tu, "Improved Hybrid Layered Image Compression using Deep Learning and Traditional Codecs", Signal Processing: Image Communication, Vol. 82, 2020.

[9] F. Jiang, W. Tao, S. Liu, J. Ren, X. Guo, D. Zhao, "An End-to-End Compression Framework Based on Convolutional Neural Networks", Proceedings of the IEEE Transactions on Circuits and Systems for Video Technology, Snowbird, UT, USA, 2 August 2017.

[10] A. Prakash, N. Moran, S. Garber, A. DiLillo, J. Storer, "Semantic Perceptual Image Compression using Deep Convolution Networks", Proceedings of the Data Compression Conference, 4-7 April 2017, pp. 250-259.

[11] H. M. Yasin, A. M. Abdulazeez, "Image Compression Based on Deep Learning: A Review", Asian Journal of Research in Computer Science, Vol. 8, No. 1, 2021, pp. 62-76.

[12] A. Ramdani, A. Virgono, C. Setianingsih, "Food Detection with Image Processing using Convolutional Neural Network (CNN) method", Proceedings of the IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology, Bali, Indonesia, 7-8 July 2020.

[13] E. Jalilian, H. Hofbauer, A. Uhl, "Iris Image Compression Using Deep Convolutional Neural Networks", Sensors, Vol. 22, No. 7, 2022.

[14] J. Sarkhawas, P. D. Khandekar, A. A. Kulkarni, "Variable Quality Factor JPEG Image Compression Using Dynamic Partial Reconfiguration and Microblaze", Proceedings of the International Conference on Computing Communication Control and Automation, Pune, India, 26-27 February 2015, pp 620-624.

[15] H. Chang, M. K. Ng, T. Zeng, "Reducing artifacts in jpeg decompression via a learned dictionary", IEEE Transactions on Signal Processing, Vol. 62, No. 3, 2014, pp. 718-728.

[16] J. Ren, J. Liu, M. Li, W. Bai, Z. Guo, "Image blocking artifacts reduction via patch clustering and low-rank minimization", Proceedings of the Data Compression Conference, Snowbird, UT, USA, 20-22 March 2013, p. 516.

[17] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering", IEEE Transactions on Image Processing, Vol. 16, No. 8, 2007, pp. 2080-2095.

[18] M. Gnacek, C. Heatwole, D. Fan, M. Hoffman, "Convolutional Neural Networks for Enhanced Compression Techniques", Proceedings of the IEEE National Aerospace and Electronics Conference, Dayton, OH, USA, 16-19 August 2021.

[19] J. Zhang, R. Xiong, C. Zhao, Y. Zhang, S. Ma, W. Gao, "Concolor: constrained non-convex low-rank model for image deblocking", IEEE Transactions on Image Processing, Vol. 25, No. 3, 2016, pp. 1246-1259.

[20] C. Dong, Y. Deng, C. C. Loy, X. Tang, "Compression artifacts reduction by a deep convolutional network", Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7-13 December 2015, pp. 576-584.

[21] S. Gu, L. Zhang, W. Zuo, X. Feng, "Weighted nuclear norm minimization with application to image denoising", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23-28 June 2014, pp. 2862-2869.