Digital Watermarking using Machine Learning

Prof Jaya Jeswani[1], Savita Rajput[2], Karan Umredkar[3,] Apurva Ware[4,]

[1, 2, 3, 4]*Department of Information Technology, Xavier Institute of Engineering, Mumbai University, Maharashtra*

([*]Corresponding author's e-mail: akshadumredkar@gmail.com)

Abstract

Digital watermarking is a technique used for the information of the images that provides security for the confidentiality. The repetitions of the multimedia objects (i.e. audio, video, text, etc.) have been protected by some of the developed digital watermarking techniques. Digital Watermarking is the procedure of concealing messages in virtual contents with a view to affirm the rightful proprietor of the copyright protection. We have proposed a method that might assist its users to embed a watermark to the quilt photograph primarily based on an adaptive method in a miles robust way even as maintaining the quality of the cover image. The implementation of this algorithm is based on cascading DWT and PCA functions using Bhattacharya distance and kurtosis.

PCA decompresses the watermark to produce better PSNR and NCC values for the tested images. The proposed algorithm uses Bhattacharyya distance and Kurtosis to detect the scaling and embedding factors making it adaptive to the input image rather than providing constant value. It also explores Convolutional Neural Networks for Denoising (DnCNN) designs to go one step further and capture the progress of very deep architectures, learning algorithms, and regularization methods in image denoising.

Keywords: Digital watermarking, DWT-PCA, PSNR, Image denoising, convolutional neural network, DnCNN.

Introduction

The development of effective methods for copyright protection of digital images has recently become an urgent and necessary requirement in the multimedia industry due to the unauthorized manipulation and copying of digital images. Original digital objects on the rise. The new digital watermark technology has been widely supported and will have many practical applications such as digital cameras, medical imaging, image databases and on-demand video systems, and more. For a digital watermarking method to be effective, it must be imperceptible and resistant to common image manipulations such as compression, filtering, rotation, scaling, and conventional attacks. Current digital image watermarking techniques can be grouped into two main classes: spatial and frequency domain watermarking techniques. The current method of the digital image of the watermark can be grouped into two main classes. Watermark and reaching frequency of spatial properties. Compared to the spatial domain method, the frequency domain watermarking method has proven to be more efficient in achieving the concealment and robustness requirements of digital watermarking algorithms. A digital watermark is the application of a popular and well-known paper watermark to the digital world. Digital watermarks describe methods and techniques for hiding information such as numbers or text in digital media such as images, video, or audio. Embedding occurs by manipulating the contents of digital data. In other words, information does not surround data. The stash process should make the modifications to the media invisible. For images this means that the modifications of the pixel values have to be invisible.

Watermark embedding can be of two types, spatial domain or frequency domain. Frequency domain techniques are considered better as they increase imperceptibility. Discrete Cosine Transform (DCT) is a lossy compression technique in frequency domain where data is lost when the original image is reconstructed from compressed image. Discrete Wavelet Transform is a lossless compression technique. In this method the cover image is divided into four sub-frequency bands: LL, HL, LH, and HH using the Haar wavelet transform. Haar wavelet is used to convert given frequencies; in this case, image matrix into square spectral signal. Similarly, the watermark image is also divided into these sub-frequency bands. Now cover image's LL-band is embedded with watermark image LL-band with the help of embedding and scaling factors. Inverse DWT is applied using this embedded LL-band with host image's LH, HL, and

HH bands to form the embedded image. Now watermark image is extracted from this image by applying the process in a reverse manner using Inverse DWT.
As the watermark image is embedded in the cover image, it tends to affect the visual quality of the cover image. Hence, dimensionality reduction is required so that only the significant components required to generate an image (in this case, watermark image) are selected, and all other insignificant components are removed. This process is known as Principal Component Analysis (PCA). In this process, only significant components are preserved. Now, these components are embedded in the cover image to form a watermarked image by maintaining their visual behavior. For the adaptive watermarking technique, the embedding factor should be adaptive. In DWT algorithm the embedding equation requires a constant coefficient to be multiplied, in order to get that perfect embedding factor trial and error was used. So, in order to reduce any error in calculation of embedding factor we need to use adaptive technique that determines the embedding factor based on the cover image. This was achieved by using Bhattacharyya Distance and Kurtosis. Bhattacharyya Distance is used to measure the similarity between two signals, whereas Kurtosis is used to check the probability distribution of the signal.

For checking imperceptibility and robustness, the PSNR ratio and NCC ratio are calculated. Imperceptibility is determined by PSNR (Peak Signal to Noise Ratio), which is calculated using watermarked image and the cover image. NCC (Normalized Cross-Correlation), which is calculated using an original watermark image and extracted watermark image determines the robustness. For checking robustness, different attacks were performed on the embedded images including Noise attacks (Salt and Pepper, Gaussian) and Geometric attacks (Cropping attack and Rotation). Most of the watermarking algorithms developed earlier have a static embedding factor. In this paper, we have proposed a digital watermarking algorithm, which has an adaptive embedding factor. Thus, any image can be used as a cover image, and any image can be used as a watermark image. Adaptive Embedding factor makes the system more flexible, robust and imperceptible against attacks.

Related Work

Name of paper: "*DCT-DWT Based Digital Watermarking and Extraction using Neural Networks*"
Author: R S Kavitha, U Eranna, and M N Giriprasad Research Scholar, Ballari Institute of Technology and Management, Ballari, KTK. Department of ECE, JNTUA Ananthapuramu, AP, India. 2020 International Conference on Artificial Intelligence and Signal Processing (AISP)
A hybrid algorithm was developed and proposed to digitally watermark image. The systems uses two techniques to produce a strong watermark. Discrete cosine transform is first applied to obtain the visible watermarked image. Further, the discrete wavelet transform is applied to embed an invisible watermark into the image. This hybrid technique makes the watermarks strong and effective. On the other hand, extraction of watermark and the original image was also carried out using an extraction algorithm using a two stage neural network.

Name of paper: "*A Proposed Digital Image Watermarking Based on DWT-DCT-SVD*"
Author: Yuqi He, Yan Hu Computer science and technology, Wuhan University of Technology 2018 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC 2018)
In this paper the survey of the digital watermarking, its frame work, their requirements and applications are presented. Apart from it there is a brief and comparative discussion on the various techniques of the digital image water-marking along merits and demerits. In the digital watermarking techniques the emphasis is on fragile and robust watermarking because the watermarking in the robust watermarking are designed in such a way that it can be detected evenafter a various attempts made for the removal of watermark and in fragile water-marking watermark brings for the purpose of authentication of the digital data.

Name of paper: "*Image Watermarking Based On DWT, DCT and SVD Technique*"
Author: Prof. R. D. Salagar Miss. Akshata S. Kamatagi International Journal Of Engineering And Computer Science ISSN: 2319-7242 Volume 4 Issue 6 June 2015
In this paper, three watermarking techniques DWT, DCT and SVD have been used. In the image watermarking, the useful implementation of image with respect to embedding the watermark into host image using the DWT, DCT and SVD are most significant part to achieve imperceptibility and better

visibility. Therefore the proposed algorithm provides the good quality of watermarked image. Higher PSNR results in better quality of image. Lower MSE results in low errors.

Name of paper: "*Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising*"
Author: Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang
In this paper, a deep convolutional neural network was proposed for image denoising, where residual learning is adopted to separating noise from noisy observation. The batchnormalization and residual learning are integrated to speed up the training process as well as boost the denoising performance. Unlike traditional discriminative models which train specific models for certain noise levels, our single DnCNN model has the capacity to handle the blind Gaussian denoising with unknown noise level. Moreover, the paper showed the feasibility to train a single DnCNN model to handle three general image denoising tasks, including Gaussian denoising with unknown noise level, single image super-resolution with multiple up-scaling factors, and JPEG image deblocking with different quality factors.

Proposed Method

The Proposed algorithms tries to implement DWT-PCA with Bhattacharya distance and Kurtosis, to create robust and adaptive Digital Watermarking. Kurtosis and Bhattacharyya distance are combined for the calculation of scaling and embedding factors. These two factors have an impact on the shape of the probability distribution. The strength of the watermark to be embedded into the cover image is determined by the similarity between the input images as result perceptibility of the cover image is retained to the maximum. Using Kurtosis and Bhattacharyya distance we are able to determine right values for the scaling and embedding factors thus making the task simple rather than random guessing which may not result in better output.

Further in the embedding process DWT is applied to the cover image. The watermark is embedded into the approximation sub-band i.e. the LL sub-band of the cover image and then combined with the remaining subbands to form the final embedded watermarked image. Here the LL sub-band is chosen as it contains the maximum information of the cover image. The extracted LL sub-band is then modified by adding the scaling and embedding factors. PCA is a technique that is used for various application like dimension reduction, feature extraction and visualization. Using PCA in Digital Watermarking compress the watermark image to get principle components which reduce the number of pixels to be embedded. As a result, only the principle components of the watermark image get embedded into the cover image which further helps to recover the watermark image more precisely.

A. GENERATION OF EMBEDDING FACTOR

1. BHATTACHARYYA DISTANCE

In statistics, the Bhattacharyya distance measures how similar two probability distributions are. It is closely associated with the Bhattacharyya coefficient, which is a measure of the amount of overlap between two statistical samples or populations. The factor that determines the relative proximity of the two considered examples. It is used to quantify the distinctness of classes in the classification and it is viewed as more solid than the MahalaNobis distance.
For multivariate normal distributions:

$$D_B = \frac{1}{8}(\mu_1 - \mu_2)^T \Sigma^{-1}(\mu_1 - \mu_2) + \frac{1}{2} ln\left(\frac{det\Sigma}{\sqrt{det\Sigma_1 det\Sigma_2}}\right) \quad \ldots\ldots\ldots(1)$$

where μ_1 and Σ_2 are the means and covariances of the distributions, and

$$\Sigma = \frac{\Sigma_1 + \Sigma_2}{2} \quad \ldots\ldots\ldots(2)$$

2. KURTOSIS

In probability theory and statistics, kurtosis is described as the "tail" of the probability distribution of a real-valued random variable. Kurtosis measures the combined weight of the tails of a distribution relative to the center of the distribution. Like skewness, kurtosis describes the shape of a probability distribution, and like skewness, there are different methods to measure it for a theoretical distribution and related methods to evaluate it from a sample. Skewness differentiates extreme values in one versus the other tail, while kurtosis measures extreme values in either tail.
The kurtosis is the fourth institutionalized minute, characterized as

$$K = \frac{\mu_4}{\sigma_4} \text{.....(3)}$$

where μ_4 is the fourth central moment and σ is the standard deviation.

3. *ALGORITHM: WATERMARK EMBEDDING*

Input: Cover Image(I), Watermark(W)
Output: Embedded Watermarked Image(W')
Step-1: Take the Watermark image W and perform PCA to compress the image w.
Step-2: Apply DWT using HAAR wavelet on the cover image and watermark image.
Step-3: Band selection.
Step-4: Calculate the scaling and embedding factors (α, β).
Step-5: Insert the watermark using the following equation (let's say LL band):

LL' = LL + (β /α) x w ……..(4)

Where, α = Scaling Factor, β = Embedding Factor.

Step-6: Combine the modified LL sub band (LL') with other LH, HL and HH bands of the cover image.
Step-7: Apply inverse DWT to get watermark image.

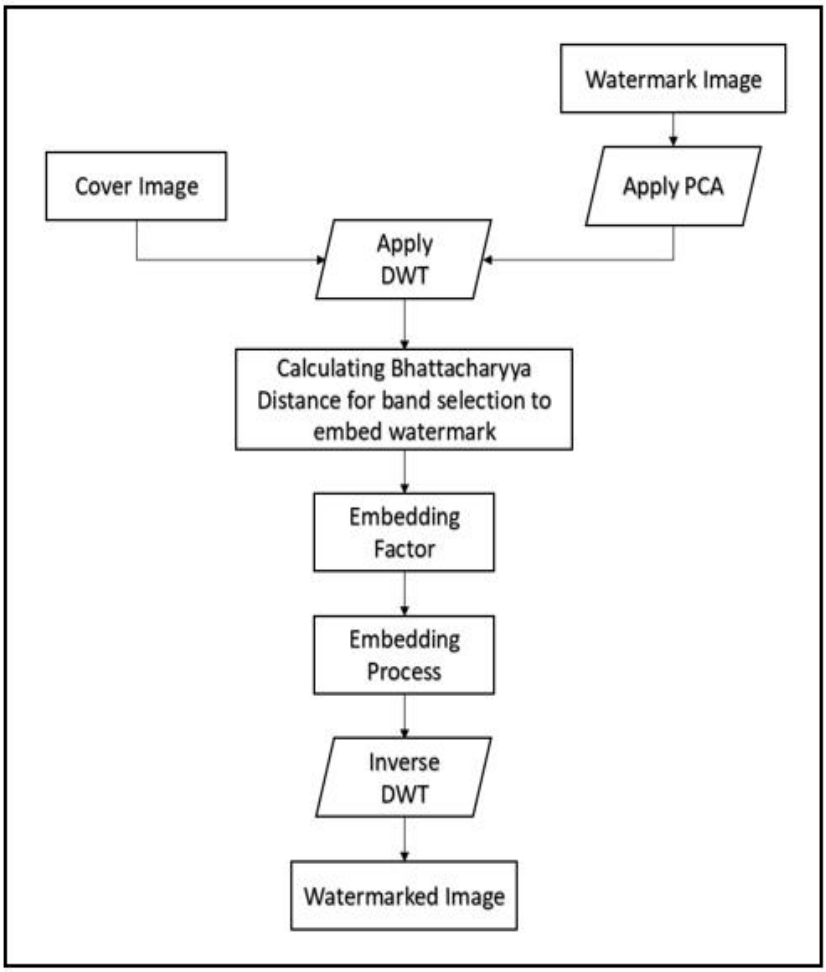


Fig 1: Watermark Embedding

4. *ALGORITHM: WATERMARK EXTRACTION*

Input: Cover Image(I) and Watermarked Image(I').
Output: Watermark Extracted (wm).
Step-1: Apply DWT using Haar wavelet on the watermarked image.
Step-2: Extract low frequency bands of both images.

Step-3: Obtain scaling and embedding factors (α, β) calculated during the embedding.
Step-4: Extract watermark using following equation:

$$wm = (LL' - LL) \times \alpha / \beta \quad \ldots\ldots\ldots(5)$$

Step-5: Reconstruct the watermark and image by applying IDWT (Inverse DWT).

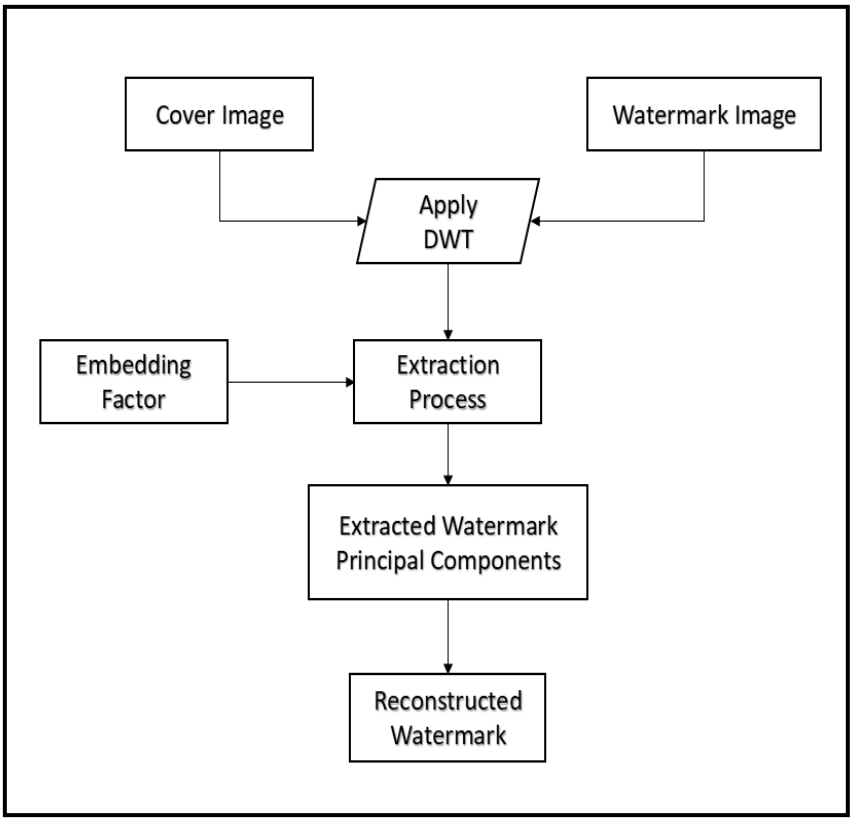


Fig 2: Watermark Extraction

5. *THE PROPOSED DENOISING CNN MODEL*

In this section, we present the proposed noisy CNN model, i.e. DnCNN, and extend it to handle some general image reduction tasks. In general, training a deep CNN model for a particular task usually involves two steps: (i) designing the network architecture and (ii) training the model from the training data. For the design of the network architecture, we modify the VGG network to make it suitable for image reduction and determine the depth of the network according to the effective patch sizes used in the methods. modern noise reduction method. For model training, we apply residual learning formula and combine it with batch normalization for fast learning and improve performance degrading performance. The input to our DnCNN is a noise observation y = x + v. Discriminant models such as MLP and CSF to learn a mapping function F(y) = x to predict the latent magnitude.

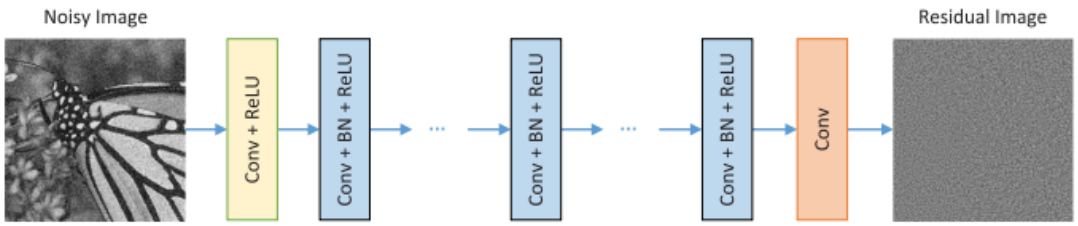


Fig 3: The architecture of the proposed DnCNN network.

For DnCNN, we adopt the residual learning formulation to train a residual mapping R(y) ≈ v, and then we have x = y − R(y). Formally, the averaged mean squared error between the desired residual images and estimated ones from noisy input

$$\ell(\Theta) = \frac{1}{2N}\sum_{i=1}^{N} \|\mathcal{R}(\mathbf{y}_i;\Theta) - (\mathbf{y}_i - \mathbf{x}_i)\|_F^2 \quad \text{…..(6)}$$

can be used as a loss function to learn trainable parameters in DnCNN. Here {(yi, xi)} N i = 1 represents N pairs of noisy training images (patches). Figure 3 illustrates the architecture of the proposed DnCNN for training R(y).

In the following, we explain the architecture of DnCNN and the strategy to reduce boundary artifacts.
1) Deep Architecture: With DnCNN of depth D, there are three types of layers.

(i) Conv+ReLU: for the first layer, 64 filters of size 3 × 3 × c are used to generate 64 feature maps, and rectified linear units (ReLU, max(0,·)) are then utilized for nonlinearity. Here c represents the number of image channels, i.e., c = 1 for gray image and c = 3 for color image.

(ii) Conv+BN+ReLU: for layers 2 ∼ (D − 1), 64 filters of size 3×3×64 are used, and batch normalization is added between convolution and ReLU.

(iii) Conv: for the last layer, c filters of size 3 × 3 × 64 are used to reconstruct the output. In summary, our DnCNN model has two main features: residual learning formula is applied to learn R(y), and batch normalization is combined to speed up training as well as improve working performance. performance reduction. By combining the convolution with Relu, DnCNN can gradually separate the image structure of noisy observation through hidden layers. Such a mechanism is similar to the repeated noise removal strategy passed in processes such as EPLL and WNM, but our DnCNN is formed in Endo-End fashion. After that, we will give many discussions about the reason for the combination of the remaining learning and batch standards.

2) Reduce the boundary artifacts: In many low-end vision applications, it usually requires the size of the output image kept like the entry. This can lead to limited artifacts. In MLP, the boundaries of noisy input images are buffered symmetrically in the preprocessing step, while the same buffering strategy is performed before each step in CSF and TNRD. Unlike the above methods, we fill zeros directly before the convolution to ensure that each middle layer feature map is the same size as the input image. We find that the simple no-buffer strategy does not lead to any boundary artifacts. This good property is probably due to the powerful capabilities of DnCNN.

The network can be used to form an origin map F(y) to predict x, or a residual map R(y) to predict v. Accordingly, when the original map looks like an identity map, the residual mapping is much easier to optimize. Note that the y noise observation looks more like the hidden image x than the residual image v (especially when the noise level is low). Therefore, F(y) will be closer to the recognition mapping than R(y), and the residual learning formula is more suitable for degrading the image.

Results and discussion

fdf For the proposed algorithm implementation, we have used Matlab version R2019a, cover image used (Barbara [512 x 512], Lena [220 x 220], Butterfly [256 x 256], Peppers [256 x 256], Family [256 x 256], Starfish [256 x 256]), watermark used (Cameraman [256 x 256]).

A.) ***PSNR*** *(Peak Signal to Noise Ratio)* is a quality metric used to measure distortion in the image after the embedding process. It is the ratio between the maximum possible value (power) of the signal and the power of the distorting noise affecting the quality.

$$PSNR = 20\log_{10}\left(\frac{MAX_f}{\sqrt{MSE}}\right)$$

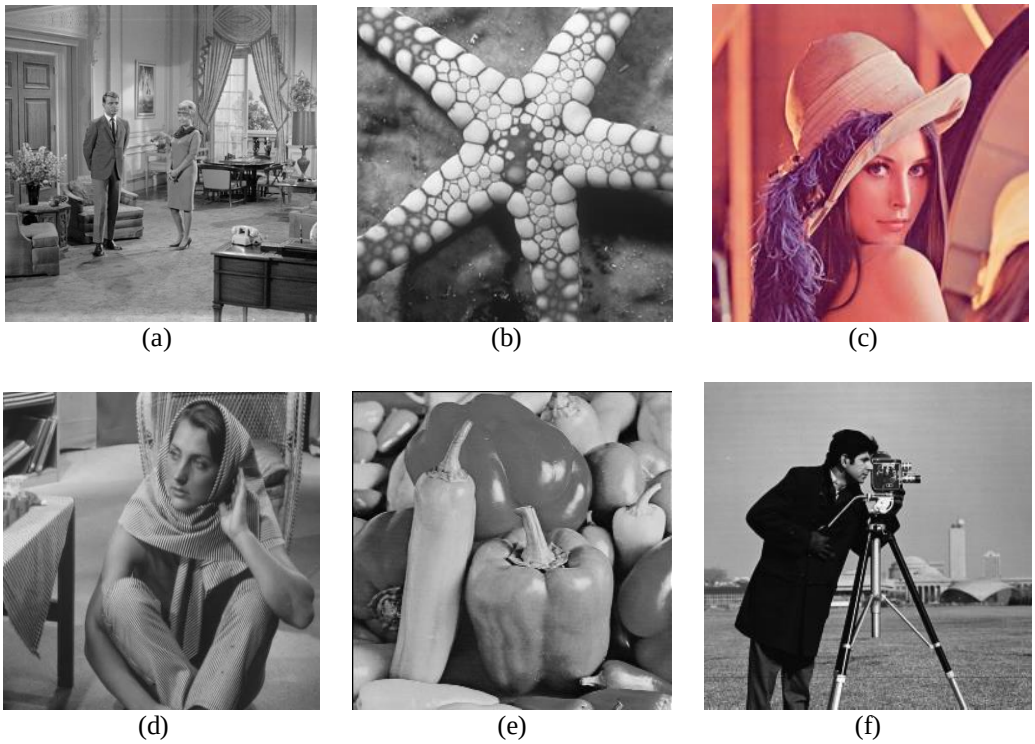
$$MSE = \frac{1}{mn}\sum_{0}^{m-1}\sum_{0}^{n-1}\|f(i,j) - g(i,j)\|^2 \quad \text{……(7)}$$

f represents the matrix data of our cover image
g represents the matrix data of embedded image
m represents the numbers of rows of pixels of the images and
i represents the index of that row
n represents the number of columns of pixels of the image and
j represents the index of that column
MAXf is the maximum signal value that exists in our cover image.

B.) **NC** (Normalised Correlation) is a metric used to verify the strength of the watermarking algorithm. It is used to find correlation coefficient between watermark and extracted watermark after attack/no attack.

$$NC=\frac{\sum_i \sum_j w(i,j)w'(i,j)}{\sum_i \sum_j w(i,j)^2} \quad \ldots\ldots(8)$$

W(i,j) = pixel values at location (i, j) of the original watermark,
W'(i,j) =pixel values at location (i, j) of the extracted watermark



(a) (b) (c)

(d) (e) (f)

Fig 4: Test cover Images (Family, Starfish, Lena, Barbara, Peppers), Watermark used (Cameraman).

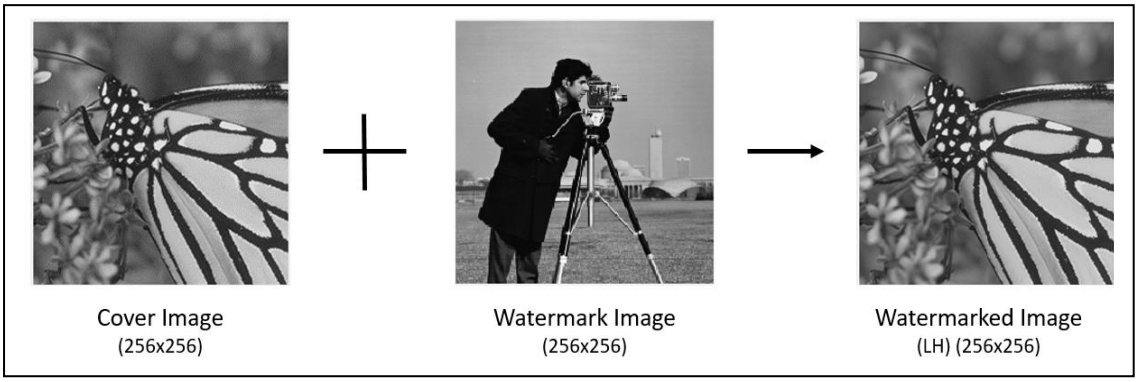


Fig 5: Watermark Embedding process

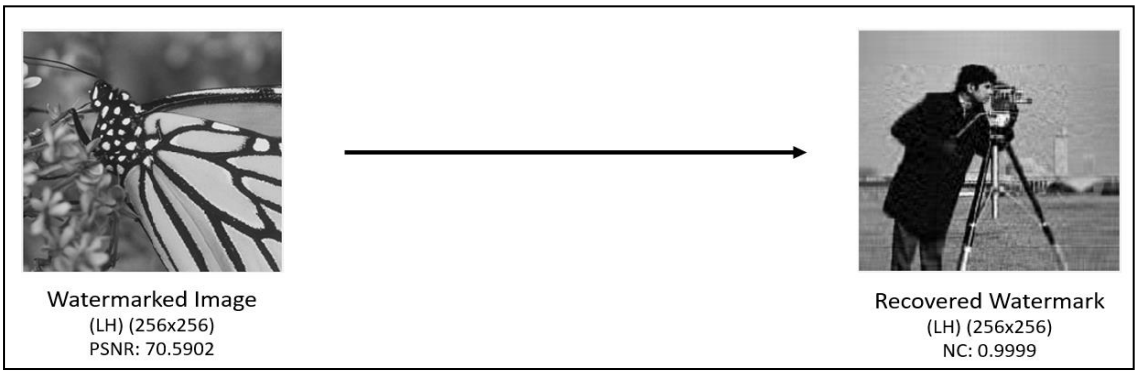


Fig 6: Watermark Extraction process

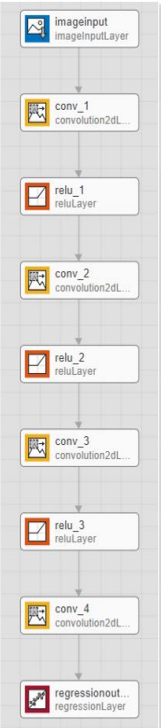
Table 4: Experimental results of some common image processing attacks and CNN attack on Butterfly image

Types of Attacks	Salt and Pepper Noise	Speckle Noise	Sharpen Image	Image Denoising using CNN
Attacked Images				
PSNR (dB)	45.9872	45.9978	49.2672	43.2536
Extracted Watermark				
NC	0.9543	43.2536	0.99158	0.91623

Table 5: Experimental results of common image processing attacks and CNN attack on all test cover images

Types of Attacks	Performance Evaluators	Family (LH Band)	Starfish (HL Band)	Lenna (LH Band)	Barbara (LH Band)	Peppers (HH Band)
No Attack	PSNR (dB)	73.9403	65.5183	76.9175	93.2853	90.275
	NC	1	0.99886	0.9999	1	0.9999
Salt - Pepper Noise	PSNR (dB)	42.4773	43.6348	42.1978	41.3455	45.7278
	NC	0.96949	0.94836	0.83113	0.87939	0.95368
Speckle Noise	PSNR (dB)	45.7038	45.0755	45.465	45.6824	45.4107
	NC	0.98521	0.969244	0.90935	0.95011	0.9504
Sharpen Image	PSNR (dB)	51.1259	49.9912	50.7728	47.4024	50.8657
	NC	0.99855	0.99424	0.98443	0.97971	0.99408
Image Denoising using CNN	PSNR (dB)	43.3738	45.0002	43.457	37.8126	41.4694
	NC	0.97214	0.96401	0.84646	0.7274	0.86126

Table 6: Model Analysis Architecture



ANALYSIS RESULT

	Name	Type	Activations	Learnables
1	imageinput 50x50x1 imag...	Image Input	50×50×1	-
2	conv_1 64 3x3x1 con...	Convolution	50×50×64	Weights 3×3×1×64 Bias 1×1×64
3	relu_1 ReLU	ReLU	50×50×64	-
4	conv_2 64 3x3x64 co...	Convolution	50×50×64	Weights 3×3×64×64 Bias 1×1×64
5	relu_2 ReLU	ReLU	50×50×64	-
6	conv_3 64 3x3x64 co...	Convolution	50×50×64	Weights 3×3×64×64 Bias 1×1×64
7	relu_3 ReLU	ReLU	50×50×64	-
8	conv_4 1 3x3x64 con...	Convolution	50×50×1	Weights 3×3×64 Bias 1×1
9	regressiono... mean-square...	Regression Output	-	-

Conclusions

The proposed algorithm combines adaptive watermarking approach the use of Bhattacharyya Distance, Kurtosis with DWT and PCA to create a more reliable, sturdy and imperceptible watermarking technique. With a purpose to provide adaptability for canopy picture, Scaling and Embedding elements are calculated by using combining Kurtosis and Bhattacharyya distance, thus lowering any possibilities of error in calculation of embedding factor. The usage of PCA we compress the watermark image to get essential additives and reduce the number of pixels to be embedded. As a consequence, simplest most important additives of watermark image get embedded in cover picture which allows to recover the watermark image greater precisely. As consequence, adaptability and use of hybrid algorithm help to make the existing algorithm more robust.

A deep convolutional neural network was proposed for image denoising, where residual learning is adopted to separating noise from noisy observation. Unlike traditional discriminative models which train specific models for certain noise levels, our single DnCNN model has the capacity to handle the blind Gaussian denoising with unknown noise level. Furthermore, we show the ability to train a single DnCNN model to handle three general image noise tasks, including Gaussian noise reduction with unknown noise level, single superpixel resolution with multiple scaling factors and debug JPEG images with different quality factors. Many test results have demonstrated that the proposed method not only gives favourable images that degrade performance quantitatively and qualitatively, but also has a promising runtime by implementing GPUs. In future, we will investigate proper CNN models for denoising of images with real complex noise and other general image restoration tasks.

Acknowledgements

We would like to thank Prof. Jaya Jeswani (Department of Information Technology) for her patience, encouragement, guidance and cooperation provided.

References

[1] R S Kavitha , U Eranna , and M N Giriprasad “DCT-DWT Based Digital Watermarking and Extraction using Neural Networks” , 2020 International Conference on Artificial Intelligence and Signal Processing (AISP)

[2] Purnima Pal, Harsh Vikram Singh, Sarvesh Kumar Verma “Study on Watermarking Techniques in Digital Images”, Proceedings of the 2nd International Conference on Trends in Electronics and Informatics (ICOEI 2018)

[3] Prof. R. D. Salagar1 , Miss. Akshata S. Kamatagi “Image Watermarking Based On DWT, DCT and SVD Technique”, International Journal Of Engineering And Computer Science Volume 4 Issue 6 June 2015

[4] Yuqi He , Yan Hu “A Proposed Digital Image Watermarking Based on DWT-DCT-SVD”,2018 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference(IMCEC 2018)

[5] Ahmed A Mohammed1, Bilal A Jebur1 and Karam M Younus “Hybrid DCT-SVD Based Digital Watermarking Scheme with Chaotic Encryption for Medical Images”, International Ninevah Conference on Engineering and Technology (INCET 2021)

[6] Hai Tao, Li Chongmin, Jasni Mohamad Zain, Ahmed N. Abdalla “Robust Image Watermarking Theories And Techniques: A Review” , Journal of Applied Research and Technology, Vol. 12, February 2014

[7] Jaya Jeswani, Tanuja Sarode, “A Hybrid DCT and DWT Color Image Watermarking in RGB Color Space” IJSCIT, vol. III, 2014.