Image Denoising by Hybridizing Preprocessed Discrete Wavelet Transformation and Recurrent Neural Networks

Raj Kumar Goel, Shweta Vishnoi, Saurabh Shrivastava

Abstract: In this paper, we are analyzing the performance of Recurrent Neural Network (RNN) for image recalling with improved training sets by Discrete Wavelet Transformation (DWT). DWT has been used for decomposing the images into four parts for low level feature extraction, to build the pattern information, encoding the pattern. When all the patterns of these four level training sets are encoded, and given as input to RNN to analyze the performance. This analysis is carried out in terms of successful and correct recalling of the images by hybridizing of DWT and RNN. Now we introduce salt and pepper noise so that the distorted feature vectors presented to the network. This gives a prototype pattern of noisy image and requires filtering of the training set. This leads to recalled output of the network that produces the pattern information for each part of the images. Now the integration is made possible if inverse discrete wavelet transformation (IDWT) to amalgamate the recalled outputs corresponding to each part of the image and final image is recognized.

Index Terms: Discrete wavelet transformation, auto associative memory, pattern storage networks, pattern classification.

I. INTRODUCTION

Pattern association is one of prominent technique for the pattern recognition task that one would like to cognize using an artificial neural network (ANN) as associative memory feature. Auto associative memory is a widely used network for storing and retrieving patterns. Recurrent neural networks (RNN) such as Hopfield, which can locally store patterns in the form of connection strengths between the processing units, are commonly known as auto-associative memory networks. RNN with bipolar processing unit, symmetric weight and asynchronous activation dynamics with deterministic or stochastic updating is widely used for pattern storage and recalling [1].

Over a very long period of time, pattern recognition and its applications have been studied and examples of these methods are linear and nonlinear [2], quantum pattern classification [3] and neural network [4] which have been proposed to accomplish the task of pattern classification. The recognition of handwritten curve scripts in the form of a character classification or the character association has been regarded as a dominant field in the area of pattern recognition with the techniques of machine learning [5].

The most significant development in the field was probably seen in the 1980s when Hopfield and Tank started working in the area and since then, especially the last 35 years, there has been a consistent improvement. The recurrent neural network (RNN) can obviously deal with temporal information directly and neutrally, however other network such as feed forward, they have some intermediate processes like conversion of pattern form temporal to spatial domain for additional developments. A Hopfield neural network (HNN) can retrieve correctly any of the learned patterns just by an exposure to only partial information about the learned pattern [6-8].

The use of the HNN for image restoration was first reported by Zhou et al. [9]. Paik and Katsaggelos [10] proposed the modified HNN (MHHN) models for the restoration of gray images, and by Sun and Yu [11] for the restoration and reconstruction of binary images. It has been observed that performance of pattern storage networks depends also on the feature extraction methods. Most of the work in this field focusing on adopting different feature extraction methods to reduce the probability of error in pattern recalling [12] with different learning rules, but the problem of false minima still occurs as the number of patterns and size of the network increase.

Numerous researchers have proposed different approaches for image denoising since 1970s. The wavelet denoising techniques is very effective and provides very good results of the distorted or blurred image. It has become powerful signal and image processing tool that has been used successfully in many scientific disciplines, including image compression, signal processing, and pattern recognition and computer graphics [13]. A novel nonlinear noise reduction method named undecimated wavelet denoising (UWD) is presented that uses the discrete wavelet transform. All sorts of signals coming from the physical environment have more or less disturbing noise. Therefore, wavelet denoising procedure are applied that employs hard thresholding of the shift invariant DWT [14].

Cheng and Chen introduced a new way to detect and track many moving objects on the basis of DWT and to identify the moving objects based on their color and spatial information. The LL sub-image of the wavelet transform has the capacity of a low-pass filter, so the noise caused by the weaving trees is removed. [15]. Hussain et.al [16] suggested a Hybrid Neural Network Predictive Wavelet Image Compression System (HNNPW). Using DWT high decomposition.
Image Denoising by Hybridizing Preprocessed Discrete Wavelet Transformation and Recurrent Neural Networks

levels is achieved, with the proposed system exhibiting high quality compressed images, while at low decomposition levels the visual quality of the compressed images using the proposed technique is similar in quality to JPEG2000.

A new multiresolution recognition scheme for recognizing unconstrained handwritten numerals was proposed by Lee et al. [17]. In this scheme, input pattern is recognized by using the coefficients of wavelet transform. The proposed scheme provides very good results and it is very robust in terms of various writing styles and sizes. A. Jaiswal et al. [18] Performed Median and Wiener filtering technique on a noisy image and subsequently soft and hard thresholding has been applied. The denoization is based on performance measures such as PSNR, MSE and visual quality of the image. After examining various techniques, they have been found that the filter and wavelet thresholding techniques together gave a good match of PSNR and MSE.

Here we used the efficacious feature extraction method, i.e. the DWT to make up the performance of the RNN more adept for pattern association. We look at the scanned images of various symbols. These images are decomposed into four parts using the wavelet transform. Each part of the transformed image is filtered, preprocessed and the feature vectors of these images are obtained. Therefore, for each part of the image pattern vector of size 1024×1 is constructed. The training set is built up which possess pattern information of all the sample images preprocessed with DWT. Once each patterns of the training set have been coded, by adding salt and pepper noise, we introduced noisy image and emulate the performance of the RNN for the presented noisy patterns of the scanned images. These noisy test images are also preprocessed with DWT. The performance of RNN is analyzed for recalled patterns information of each segment of the preprocessed image and then after its reconstruction. The reconstructed image is compared to the original image to evaluate the performance of RNN. The Rest of this paper has been structured in the under-mentioned way. Section 2 discusses DWT as a feature extraction method to construct feature vectors for input stimuli. RNN and its learning mechanisms have been discussed in section 3. Section 4 describes the Implementation detail and experiment design. Section 5 includes the simulated results with discussion. Section 6 contains the conclusion followed by the references.

II. FEATURE EXTRACTION

The feature extraction is a specific and salient feature for pattern recognition task. Therefore, to build the pattern information, encoding the pattern, the preprocessing is indispensable. Low-level feature extraction is based on finding points, edges, lines, etc., while high level feature extraction methods utilize low-level feature for providing meaningful information for further processing of image analysis. In most cases, the high-level Feature extraction uses the ANN for extracting the features in multiple layers. Extraction of low-level features from the image was performed using DWT [19]. In this procedure, the pattern set used for this study are the scanned handwritten Greek symbol images of 100 different individuals collected on a white paper. Due to differences in the quality of ink, pen, paper, and scanner, the unconstrained handwritten text image obtained as a scanner output may contain some impurities. Preprocessing plays a vital role in HCR for removing these types of impurities. Scanned image passes through a series of operations such as filtering, binarization, edge dilation and smoothing to produce a cleaned image for segmentation and recognition. A sample of these scanned images as shown in Figure 1.

Figure 1: Scanned images of handwritten symbols

Image segmentation is primarily based on the gray value of the image due to identify important edges and feature, so the first step is to convert the scanned RGB images into gray level. In this process, the basic information of the image is retained, and does not affect the contour of the image. These gray images are turned into binary images of size 60×60. In this method, the quality of the images deteriorates, so we use the edge detection and dilation method to show the defined base of the image as shown in Figure 2. Edge detection involves a process to discriminate, extract, emphasize useful image feature [20].

Figure 2: Various stages of images for feature processing.

Here after binarization, the necessity was to convert the binary image to bipolar image, since RNN work efficiently with bipolar pattern. Now the image has been changed to bipolar vectors. All the image vectors are put into a comprehensive matrix.

Discrete wavelet transform

DWT calculates the wavelet coefficients at discrete interval of time and scale. DWT is a multi-resolution analysis which is a desired feature of any signal decomposition technique. Video signal and real word digital images are non-stationary in nature so a wavelet transform decomposes a non-stationary image into a set of multi-scaled wavelet, so that each component becomes relatively more stationary and hence easy for coding.
Multi-scale signaling as input data may produce more accurate forecasting rather than a single pattern input. In the literature, it can be seen that in all cases wavelet progressively emerged as a unifying framework for dealing with many different types of scaling processes, and were identified as versatile generation of previous technique which may have been used in specific contexts [21, 22].

It has built-in multi-resolution structure, wavelet transforms were natural tools for uncovering self-similar features in signals, image or processes. Continuous Wavelet transformation (CWT) of 1-D signal f(x) is defined as

$$W_{\psi}f(x) = (f(x), \psi_{(x,t)}(x)) = \frac{1}{\sqrt{t}} \int_{-\infty}^{+\infty} f(x) \psi^{*}_{(x,t)}dx$$

where $s$ is the scale parameter, $t$ is the translation parameter and the * denote the complex conjugate of mother wavelet($\psi_{(x,t)}(x)$). The mother wavelet is the prototype for generating the other windows function [23].

A. Pattern formation using discrete wavelet transformation

DWT is the algorithm used to image compression, dimensionality reduction and feature extraction process. DWT is applied to these images to partition the images into four parts as depicted in Figure 3. The Low-Low (LL) sub-band is the approximate image of the input image, Low-High (LH) extracts the horizontal feature of the original image, High-Low (HL) gives vertical characteristics and HH sub-band provides diagonal features.

![Figure 3: Partition of the bipolar image.](Image)

The images were first scanned as gray images and then transformed with wavelet to retain the fine details in the images. After that, each part of the image is reshaped into 3600x1 pattern vector. Therefore, from each image we obtain the four pattern vectors of size 3600x1. Hence, we have 24 different handwritten symbols, so that we could construct the four training sets, each of them is of size 3600x24.

III. RECURRENT NEURAL NETWORK

The proposed RNN model consists of $N$ neurons and $N \times N$ link strengths. Every neuron can be reside in one of the two states i.e. ±1. Let L bipolar patterns are encoding in the NN for pattern association. Therefore, to store the L pattern on the network, Weight Metrics W is computed by using the Hebbian law as given:

$$W_{ij} = \frac{1}{N} \sum_{l=1}^{L} a_i^l a_j^l \quad (i \neq j)$$

Where $\{ a_i^l, i=1,2,3,\ldots,N; l=1,2,3,\ldots,L \}$ and $i \neq j$.

The network can be initialized as:

$$s_i'(t) = a_i'(t), \quad i = 1,2,3,\ldots,N; \quad a_i'(t) \neq j$$

In the Hopfield model, each unit can represent the activation value and output as follows:

$$y_i = \sum_{j=1}^{N} w_{ij} s_j(t), \quad i, j = 1,2,3,\ldots,N; i \neq j$$

and

$$s_i(t+1) = \text{sgn}(y_i)$$

Therefore, in order to memorize L scanned and preprocessed images in an N-unit RNN, there should be one steady state corresponding to each stored pattern. Hereby, the memory pattern in the end should be fixed-point attractors of the network and meet the fixed-point condition as:

$$y_i^j = \sum_{j=1}^{N} w_{ij} s_j^j \quad \text{Where } y_i^j \geq 0$$

So, the activation dynamics equation should satisfy for pattern storage and retrieving:

$$f(\sum_{j=1}^{N} w_{ij} s_j^j) = s_i^j;$$

Where, $i, j = 1,2,\ldots,N; i \neq j$

The initial weights are considered as $w_{ij} = 0$ (near to zero) \( \forall i \)’s and \( j \)’s. Therefore from the equation 2 we can obtain final weight matrix which represents the encoding of patterns information as:

$$W^L = \frac{1}{N} \begin{bmatrix}
0 & s_1 s_2 & s_1 s_3 & \cdots & s_1 s_N \\
s_2 s_1 & 0 & s_2 s_3 & \cdots & s_2 s_N \\
s_3 s_1 & s_3 s_2 & 0 & \cdots & s_3 s_N \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
s_N s_1 & s_N s_2 & s_N s_3 & \cdots & 0
\end{bmatrix}$$

The pattern retrieval involves setting the initial position of the network equal to an input vector $S_i$. The states of individual unit are updated frequently until the overall condition of the network is stable [25, 26].

IV. IMPLEMENTATION DETAIL AND EXPERIMENT DESIGN

The patterns in the form of bipolar vectors created in section 2 were then encoded in the RNN with Hebbian...
learning rule. Since, we have the four training sets for the sample images so as to construct the four independent RNNs each of 3600 processing units. Each neural network will be used for storage of one band of the images, i.e. LL, LH, HL and HH. Therefore, after the pattern storage, the recalling process from these neural networks starts. In this process the pattern vectors of individual parts of the image, i.e. LL, LH, HL and HH are presented separately to respective networks. The networks simulate to recall the pattern vectors as shown in Figure 4.

![Figure 4: Recalled performance for LL, LH, HL and HH part of the image.](image)

Further, we combine the simulated output of all the networks and reconstruct the pattern. The recalled performance has been evaluated for the original image pattern with combined simulated output of all the four individual networks. The simulation results for recalling of original image pattern are presented in Figure 5. Here, the original pattern image is considered after the combination of LL, LH, HL and HH of the image. The simulated results shows that the recalled pattern is more efficient and close to the original image with respect to the recombined original image from its parts.

![Figure 5: Recalled and original recombined image.](image)

Algorithm: Pattern Storage and Recalling

The algorithm for pattern recalls in RNN for storing L patterns is as follows:

1. Determine the value of the weight to store the pattern (using the Hebb’s rules) for pattern storage according to equations 2 and 8.
2. For each input vector \( x \), repeat steps 4 to 8.
3. Set initial activations of the net equal to the external input vector \( x \).
4. Perform steps 6 to 8 for each unit \( y_i \).
5. Compute the net input
   \[
   Y_j = x_i + \sum_{j=1}^{n} Y_j * W_{ji}
   \]
6. Determine the activation (output signal):
   \[
   Y_j = \begin{cases} 
   +1, & \text{if } Y_j > 0 \\
   -1, & \text{if } Y_j <= 0 
   \end{cases}
   \]
   Broadcast the value of \( Y_j \) to all other units.
7. Test for convergence as per equation 7.

Salt and Pepper noise removal is an active research area in digital image processing. In this noise, the black points and white dots scattered throughout the image usually look like salt and pepper. This is also called data drop noise because statistically its drop the original data values [24]. In noise free document images, we introduced 10% to 50% salt and pepper noise with step size of 10% so that we can compare our results with the noise free images by DWT combined sub bands approach with RNN. Now, we take the distorted form which is already put in network as discussed in section 2. The 50% noisy form of the image is shown below in Figure 6.

![Figure 6: Noisy form of the stored image](image)

The distorted image was preprocessed and presented like as the prototype input pattern vector to the RNN for retrieving. The preprocessing step employs the dilation method and then the DWT has been applied to partition the image into four parts. A sample of 50% distorted image as shown in Figure 7.

![Figure 7: Partition of the preprocessed 50% noisy sample image](image)
These images are presented in the form of 3600×1 bipolar pattern vectors. These pattern vectors are presented to the corresponding four recurrent networks as the prototype input pattern vectors. The simulated outputs from these networks are obtained. The performance of each neural network for each part of the noisy image is given in Figure 8.

Figure 8: Recalled performance for Noisy LL, LH, HL and HH part of the fifty percent noisy sample image.

Then, we combine the simulated outputs corresponding to each part of the noisy pattern vector of all networks and reconstruct the pattern. Performance of RNN for retrieving corresponding to the noisy image pattern has been assessed with integrated simulated output of all the four individual networks. The simulation result for recalling corresponding to the noisy image pattern can be seen in Figure 9. Here, the pattern image is constructed after the combination of parts (LL, LH, HL and HH) of the image. This reconstructed noisy image has been compared with the combination of recalled images corresponding to noisy input patterns of LL, LH, HL and HH from all the networks. Results clearly show that the recalled image is more efficient and close to the original image with respect to recombined noisy image from its parts.

Figure 9: Recalled and Noisy recombined images with 50% distortion.

V. RESULTS AND DISCUSSION

The results in Figure 8 and 9 are interesting and observable. The individual performance of the networks for corresponding parts of the image is not effective but when we combine the simulated outputs of these networks the recalled image for noisy image is approximate to original image. It shows the improvement in recalling competency of the RNN by using DWT method as preliminary processing.

The RNN has the ability not merely to retrieve the pattern for the presented prototype input pattern and to recover the encoded pattern as well corresponding to current noisy input pattern which has already been encoded. Thus, the network can retrieve the exact pattern when the noisy or part of the information of that pattern is presented to the network. The network can then be presented with either a portion of one of the images or an image corrupted with noise via multiple iterations it will attempt to reconstruct one of the stored images [25]. The efficiency of the selected feature extraction method determines storage at the level of the image quality for further processing. Therefore, the feature extraction algorithms discussed here have been considered to extract the effective and useful features from the images for their storage. The results of the simulations in this article show an improvement over the conventional limits on recall efficiency of the Recurrent Network. The results when compared to that in Singh et al., [26] show that the network is able to perform for both the noiseless and up to 50% of distorted patterns. It can be observed that the results produced by the network receiving DWT combined sub bands pattern are better and reliable up to 50% noisy level as compared with Singh et al. The Recalled Performance of RNN using preprocessing DWT method from 10 to 50% noise with step size of 10% of a sample image is shown in Table 1. Table 2 shows the average recalled performance of RNN using preprocessing DWT method at 50% noise of all sets of Greek symbols. The value of R indicates the relationship between the outputs and targets. If R equals to 1 that means the relationship between outputs and targets is exactly linear. If it is close to zero, the relationship between outputs and targets is not linear.

Table 1: Recalled Performance of RNN using preprocessing DWT method from 10 to 50% noise with step size of 10%.

<table>
<thead>
<tr>
<th>Greek symbols</th>
<th>Noise level</th>
<th>Regression (R) at LL, LH, HL and HH Sub-band</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0%</td>
<td>LL(Approximate) LH(horizontal) HL(vertical) HH(diagonal)</td>
</tr>
<tr>
<td></td>
<td>10%</td>
<td>.72484 .73685 .73848 .9118</td>
</tr>
<tr>
<td></td>
<td>20%</td>
<td>.68305 .72440 .73439 .9104</td>
</tr>
</tbody>
</table>
Image Denoising by Hybridizing Preprocessed Discrete Wavelet Transformation and Recurrent Neural Networks

<table>
<thead>
<tr>
<th>30 %</th>
<th>0.67857</th>
<th>0.72152</th>
<th>0.72921</th>
<th>0.90374</th>
</tr>
</thead>
<tbody>
<tr>
<td>40 %</td>
<td>0.65413</td>
<td>0.71819</td>
<td>0.72408</td>
<td>0.8976</td>
</tr>
<tr>
<td>50 %</td>
<td>0.64473</td>
<td>0.71657</td>
<td>0.72314</td>
<td>0.89235</td>
</tr>
</tbody>
</table>

Table 2: Average Recalled Performance of RNN using preprocessing DWT method at 50% noise

<table>
<thead>
<tr>
<th>Greek symbols</th>
<th>LL (Approximate)</th>
<th>LH (Horizontal)</th>
<th>HL (Vertical)</th>
<th>HH (Diagonal)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Λ</td>
<td>0.623355</td>
<td>0.705845</td>
<td>0.70993</td>
<td>0.88456</td>
</tr>
<tr>
<td>Β</td>
<td>0.58393</td>
<td>0.681935</td>
<td>0.683805</td>
<td>0.88132</td>
</tr>
<tr>
<td>Γ</td>
<td>0.609595</td>
<td>0.695055</td>
<td>0.6979</td>
<td>0.897025</td>
</tr>
<tr>
<td>Σ</td>
<td>0.612535</td>
<td>0.69615</td>
<td>0.69931</td>
<td>0.885695</td>
</tr>
<tr>
<td>Ε</td>
<td>0.594425</td>
<td>0.69049</td>
<td>0.695495</td>
<td>0.88131</td>
</tr>
<tr>
<td>Ψ</td>
<td>0.609455</td>
<td>0.695435</td>
<td>0.70025</td>
<td>0.89173</td>
</tr>
<tr>
<td>η</td>
<td>0.607655</td>
<td>0.6978</td>
<td>0.70671</td>
<td>0.874725</td>
</tr>
<tr>
<td>Ω</td>
<td>0.630395</td>
<td>0.70691</td>
<td>0.62972</td>
<td>0.89594</td>
</tr>
<tr>
<td>ι</td>
<td>0.60583</td>
<td>0.700045</td>
<td>0.70513</td>
<td>0.886155</td>
</tr>
<tr>
<td>κ</td>
<td>0.591525</td>
<td>0.67529</td>
<td>0.6756</td>
<td>0.88093</td>
</tr>
<tr>
<td>λ</td>
<td>0.62233</td>
<td>0.707775</td>
<td>0.711515</td>
<td>0.909935</td>
</tr>
</tbody>
</table>

| 0.60092 | 0.67922 | 0.682325 | 0.89966 |
| 0.609305 | 0.674655 | 0.676005 | 0.878955 |
| 0.60725 | 0.70094 | 0.707495 | 0.89298 |
| 0.597535 | 0.69027 | 0.690525 | 0.885785 |
| 0.616205 | 0.697705 | 0.68742 | 0.90536 |
| 0.602095 | 0.695895 | 0.697595 | 0.884185 |
| 0.58708 | 0.681 | 0.69343 | 0.88654 |
| 0.619815 | 0.69953 | 0.70192 | 0.89724 |
| 0.596045 | 0.667895 | 0.66924 | 0.886455 |
| 0.60038 | 0.70928 | 0.711095 | 0.891225 |
| 0.53925 | 0.659815 | 0.66118 | 0.87445 |
| 0.545855 | 0.667115 | 0.66905 | 0.882235 |
| 0.62054 | 0.71562 | 0.717975 | 0.90814 |

| 0.60092 | 0.67922 | 0.682325 | 0.89966 |
| 0.609305 | 0.674655 | 0.676005 | 0.878955 |
| 0.60725 | 0.70094 | 0.707495 | 0.89298 |
| 0.597535 | 0.69027 | 0.690525 | 0.885785 |
| 0.616205 | 0.697705 | 0.68742 | 0.90536 |
| 0.602095 | 0.695895 | 0.697595 | 0.884185 |
| 0.58708 | 0.681 | 0.69343 | 0.88654 |
| 0.619815 | 0.69953 | 0.70192 | 0.89724 |
| 0.596045 | 0.667895 | 0.66924 | 0.886455 |
| 0.60038 | 0.70928 | 0.711095 | 0.891225 |
| 0.53925 | 0.659815 | 0.66118 | 0.87445 |
| 0.545855 | 0.667115 | 0.66905 | 0.882235 |
| 0.62054 | 0.71562 | 0.717975 | 0.90814 |

VI. CONCLUSION

The experiential results show that RNN is performing better and efficiently for recalling process with the DWT technique of preprocessing. The scanned images were preprocessed with edge detection and dilation method. Further these preprocessed images were partitioned in four parts (LL, LH, HL and HH) by using DWT method. Each part of the image has been converted into bipolar pattern vector of size 3600×1 and presented to RNN for encoding using Hebbian correlation rule. The following points are observed form the experimental results:
1. The individual performances of networks corresponding to noisy images were not effective but when we mingle the simulated outputs of the networks and reconstruct the image the performances improve. The reconstructed images found more near to original image.

2. It has been noticed that recalling of reconstructed image is efficacious up to 50% of the noise in the original image.

3. The performance in recalling of individual networks for original images shows the effectiveness and the same is maintained for the reconstructed image.

4. The performance of networks degrades if more complicated and large size network is used.

RNN’s services can also perform an analysis on the larger number of images using some of the more sophisticated methods of feature extraction. The performance of RNN for pattern fetching can be further enhanced through the use of evolutionary algorithms.

ACKNOWLEDGEMENTS

This work is supported by the University Grants Commission (UGC), New Delhi, India, under grant number UGC-Ref No-3485.

REFERENCES


