STM JOURNALS
Journal of Artificial Intelligence Research & Advances
ISSN: 2395-6720 (Online)
Volume 5, Issue 3
www.stmjournals.com

Artificial Neural Network-Based Multifocus Image Fusion

Meenu Manchanda[1,]*, Deepak Gambhir[2]
[1]Associate Professor, Department of Electronics and Communication Engineering,
Vaish College of Engineering, Rohtak, Haryana, India
[2]Assistant Professor, Department of Electronics and Communication Engineering,
Amity School of Engineering and Technology, Bijwasan, New Delhi, India

Abstract
Fusion of multiple images is needed for combining information contained in images obtained using different imaging devices or different camera settings. Due to the physics of optical lenses present in these devices, it is difficult to capture an image that has all the relevant objects well in focus. Thus, artificial neural network (ANN)-based multifocus image fusion method in YCbCr color space is proposed. Input images are first mapped from RGB color space to YCbCr color space. Y-component of both input images is then divided into same-size blocks. Features that reflect the clarity of these blocks are extracted and used by ANN to select the input blocks that are used to produce a single all-in-focus fused image. The fused image is then mapped back into the RGB color space, thereby producing the final fused image. Both subjective and objective results prove that the proposed ANN-based multifocus image fusion method outperforms the traditional wavelet transforms-based image fusion algorithms.

Keywords: *Image fusion, Artificial neural network*

****Author for Correspondence*** E-mail: meenumanchanda73@gmail.com

INTRODUCTION

Digital image processing techniques are extensively being used in several applications with the help of imaging devices. However, the limitations and abilities of various imaging devices to capture the information are different from one another. Optical lenses used in imaging devices have limited depth of focus and, therefore, it is not possible to obtain an image with all relevant objects in focus. The objects which are in focus appear sharper and the objects which are not in focus appear blurred in the captured image. To overcome this problem, multifocus image fusion is one of the best possible solutions. Multifocus image fusion method [7,13] combines multiple images with different focus settings to produce an image with increased depth of focus. This newly produced image with "all-in-focus" objects is extensively suitable for human visual perception as well as for various computer processing tasks such as object detection, object segmentation, and object recognition.

In general, image fusion algorithms [6,10] can be divided into pixel-level, feature-level, and decision-level image fusion algorithms. Pixel-level algorithms assume correspondence between individual pixels in the input images, feature level assumes known correspondence between features and decision level requires correspondence between image descriptions to produce fused images. Since feature-level and decision-level fusion algorithms are more application-dependent, therefore, pixel-level algorithms are more popular among researchers. Pixel-level fusion algorithms [9] perform fusion either in the spatial domain or in the transform domain. Spatial domain algorithms include weighted average, component substitution, high pass filtering, and arithmetic-based fusion algorithms. These algorithms are relatively simple and time-efficient fusion techniques. However, these algorithms are successful in improving the spatial details but tend to produce spectral distortion in the fusion results. The limitations of spatial domain image fusion algorithms led to the development of transforms-based image fusion algorithms. Typical transforms include pyramid transforms and discrete wavelet transforms. However, pyramid-based fusion algorithms produce poor SNR and provide no directional information whereas discrete wavelet-transform-based algorithms suffer from shift variance.

Artificial neural network (ANN) has been found to be a more powerful and self-adaptive tool of pattern recognition. Since YCbCr color space exploits the characteristics of the human eye, thus ANN-based multifocus image fusion in YCbCr color space is proposed in the study.

The rest of the study is organized as follows: Section II gives an introduction to ANN, Section III describes the YCbCr color space, Section IV presents the proposed work, Section V gives results and discussions and finally conclusion is given in Section VI.

ARTIFICIAL NEURAL NETWORK

An ANN is an immensely parallel distributed structure vaguely inspired by the biological neural networks. The network gains knowledge through a learning process, stores this knowledge in synaptic weights, and finally produces reasonable outputs for inputs that were not encountered during learning. An important property that enhances the performance of a neural network is its learning ability. The basic architecture (Figure 1) of neural network consists of three or more layers. An input layer consisting of a set (say N) of input nodes, one or more layers consisting of another set (say K<N) of hidden computational nodes and an output layer consisting of at least one output node. The i^{th} node of the input layer is connected to the j^{th} node of the hidden layer by weight W_{ij} whereas the weight between the j^{th} node of the hidden layer and the k^{th} node of the output layer is V_{jk} (in case of one output node k=1). These weights are set explicitly during training by applying several teaching patterns to the network and allowing it to adjust its weight according to the predefined learning rule. The hidden layer implements a set of radial basis functions; this results in non-linear transformation from the input layer to the hidden neuron. Commonly used activation function of the hidden neuron is the Gaussian basis function defined as:

$$G(\| x-\mu_j \|) = e^{-\frac{\|x-\mu_j\|^2}{2\sigma_j^2}} \tag{1}$$

where, X is the input data, μ_j is the center and σ_j is the width of the basis function of the j^{th} hidden neuron.

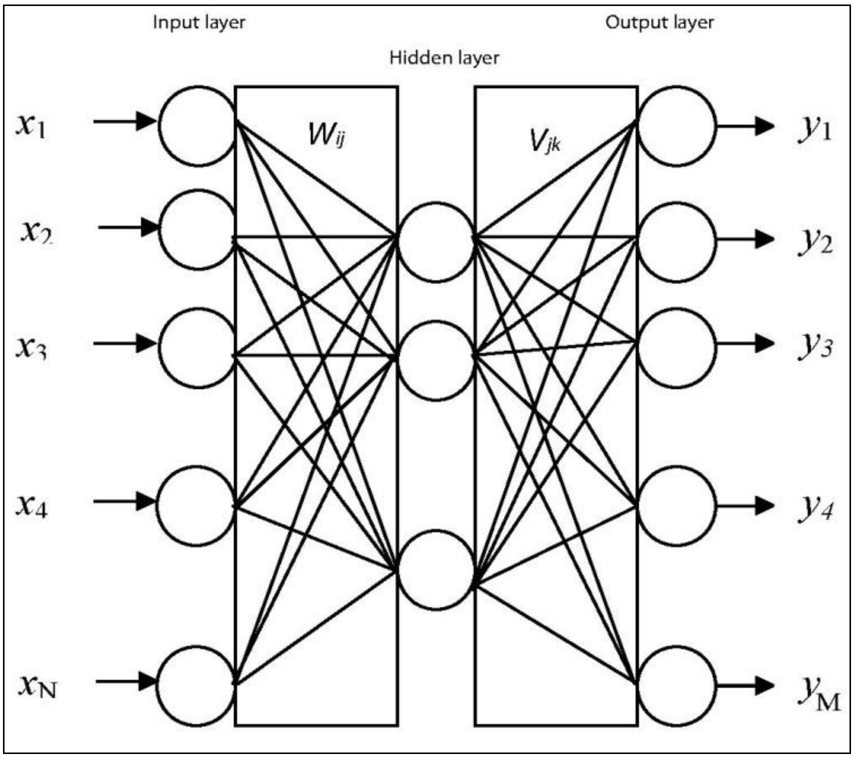


Fig. 1: *Basic Architecture of an Artificial Neural Network.*

The output neurons perform a weighted sum of outputs of individual hidden neurons, i.e.,

$$Y_k(x) = \sum_{j=1}^{N} v_{jk} H_j(x) \text{ for k=1,2,…,M} \quad (2)$$

Where, Y_k and H_j are the outputs of k^{th} output neuron and j^{th} hidden neuron. The number of neurons present in the hidden layer determines how well the network is successful in separating the data. A large number of neurons in the hidden layer ensure correct learning whereas a neural network with too low hidden neurons fails to learn the relationship among the data. The number of neurons and the activation functions used at the hidden layer and at the output layer describe the behavior of the neural network.

Radial basis function neural network (RBFN) [8] and Probabilistic neural network (PNN) [1] are the two closely related ANNs and have been used in the study for developing multifocus image fusion algorithm. PNNs are organized into a multilayered feed-forward network with four layers: input layer, hidden layer, summation layer and output layer. PNNs are much faster and more accurate but are require more memory to store the model.

YCbCr Color Space

In YCbCr color space, Y represents the luminance component whereas Cb and Cr represent the blue difference and red difference color components respectively. Since the YCbCr transformation is linear, its computational complexity is far lower than the RGB color space. The forward and reverse YCbCr transformations are achieved by means of Equation (3) and (4), respectively.

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.2990 & 0.5870 & 0.1140 \\ -0.1687 & -0.3313 & 0.5000 \\ -0.5000 & -0.4187 & -0.0813 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3)$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.0000 & 0.0000 & 1.4020 \\ 1.0000 & -0.3441 & -0.7141 \\ 1.0000 & 1.7720 & 0.0000 \end{bmatrix} \begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} \quad (4)$$

THE PROPOSED ALGORITHM

Two input images A and B to be fused are first transformed from RGB color space to YCbCr color space [2] (where Y represents the luminance component and Cb and Cr represent the blue difference and red difference chrominance component). This color space exploits the characteristics of human eye. Since most of the information is contained in the Y-components, therefore, the Y-component of these images is then divided into several blocks. Features from two of these blocks (one from each Y-component) are extracted and features vector incident to ANN is created. The vector so created helps ANN to determine which one of these blocks is clearer and should be selected for creating the fused image. The proposed algorithm uses five features [3,4]: spatial frequency (SF), phase congruency (PC), gradient magnitude (GM), visibility (V) and edge information (E) to represent the clarity of the image. The detail of these features is as follows:

(1) *Spatial Frequency* (SF) is utilized to determine the level of details present in an image block. A block with fine details and sharp edges has high spatial frequency. For a block B of size R×S,

$$SF = \sqrt{(R_{Freq})^2 + (C_{Freq})^2}$$

Where, R_{Freq} and C_{Freq} are the row frequency and the column frequency respectively and are obtained as:

$$R_{Freq} = \sqrt{\frac{1}{RS}\sum_{r=0}^{R-1}\sum_{s=0}^{S-1}[Blk(r,s) - Blk(r,s-1)]^2}$$

and

$$C_{Freq} = \sqrt{\frac{1}{RS}\sum_{r=0}^{R-1}\sum_{s=0}^{S-1}[Blk(r,s) - Blk(r-1,s)]^2}$$

(2) *Phase Congruency* (PC) provides information about local structures in an image and is defined as:

$$PC = \frac{1}{RS}\sum_{r=0}^{R-1}\sum_{s=0}^{S-1} \frac{\sum_j E_{\theta_j}(r,s)}{\sum_k \sum_j A_{k,\theta_j}(r,s) + \varepsilon}$$

Where $E_{\theta_j}(r,s)$ is the local energy along orientation θ_j and $A_{k,\theta_j}(r,s)$ is the local amplitude at scale k and orientation θ_j.

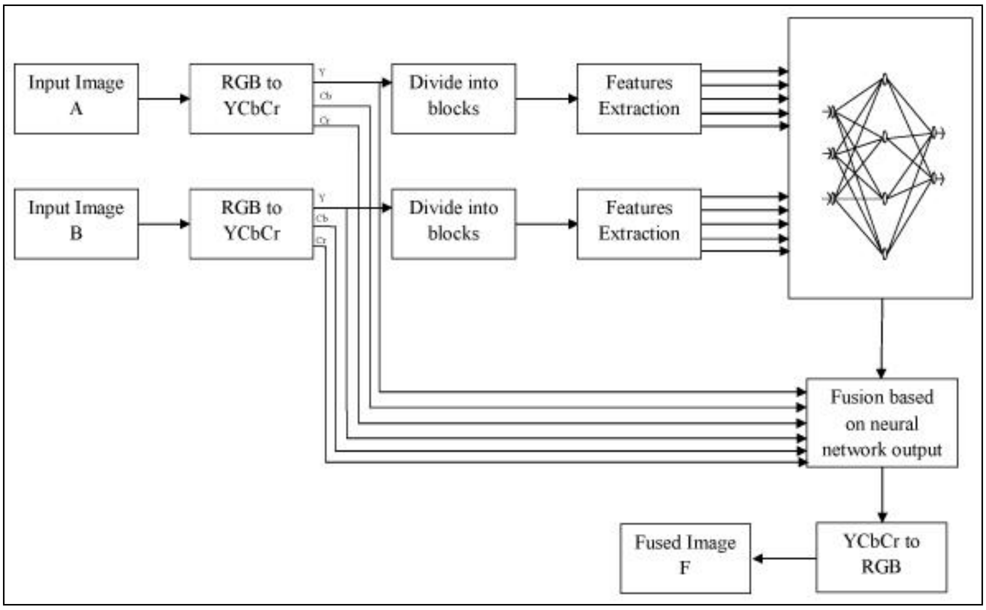


Fig. 2: *The Proposed Algorithm.*

(a) Left-blurred image *(b) Right-blurred image*
(c) DWT fused image *(d) LWT fused image*
(e) LPT fused image *(f) PCA fused image*
(g) RBFN fused image *(h) PNN fused image*

Fig. 3: *(a) to (h) Subjective Image Quality Assessment of Image Set 1.*

(3) *Gradient Magnitude* (GM) measures the sharpness, i.e., the clarity of an image and is defined as:

$$GM = \frac{1}{RS}\sum_{r=0}^{R-1}\sum_{s=0}^{S-1}\sqrt{G_r^2 + G_s^2}$$

Where, G_r and G_s are the gradients along the horizontal and vertical directions in a block respectively.

(4) *Visibility* (V) of block Blk is determined using the following equation:

$$V = \frac{1}{RS}\sum_{r=0}^{R-1}\sum_{s=0}^{S-1}\frac{|Blk(r,s) - \mu|}{\mu^{e+1}}$$

Where, μ is the mean value of pixel's intensity in image block Blk and e is a constant ranging from 0.6 to 0.7 [3].

(5) *Edge Information (E):* Since edges provide important visual information, therefore, an ideal fused image should contain all the edge details present in input images. The edge information can be extracted using either Canny or Sobel edge operator. The block with higher number of edge pixels is then used for constructing the fused image.

The proposed ANN-based image fusion algorithm (Figure 2) is as follows:

1. Input images A and B are mapped from RGB color space to YCbCr color space.
2. The Y-component of these images is then divided into several blocks of size $R \times S$.

Journal of Artificial Intelligence Research & Advances
Volume 5, Issue 3
ISSN: 2395-6720 (Online)

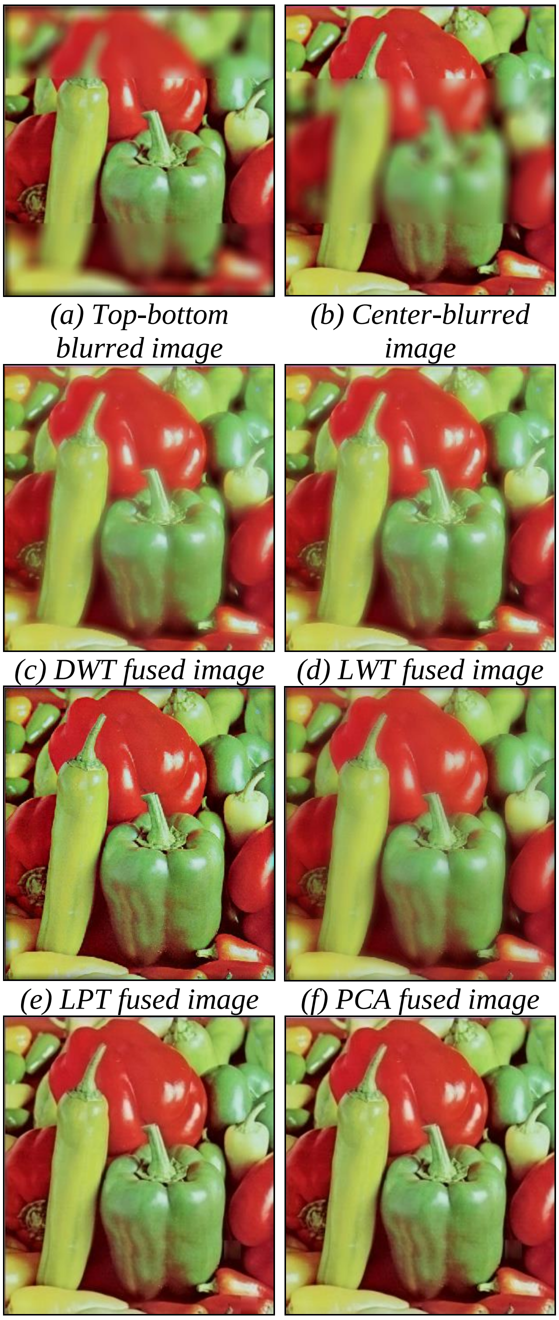
3. For each Y-component block pair (A_i,Bi), five important features: SF, PC, GM, V, and E are extracted. These features reflect blocks clarity and contribute towards image fusion. Create two feature vectors: $(SF_{A_i}, PC_{A_i}, GM_{A_i}, V_{A_i}, E_{A_i})$ and $(SF_{B_i}, PC_{B_i}, GM_{B_i}, V_{B_i}, E_{B_i})$ for both image blocks, A_i and B_i respectively.
4. Use these features vectors to train ANN in order to determine whether the block A_i and B_i is clearer. The difference of features vectors, i.e., $(SF_{A_i} - SF_{B_i}, PC_{A_i} - PC_{B_i}, GM_{A_i} - GM_{B_i}, V_{A_i} - V_{B_i}, E_{A_i} - E_{B_i})$ is used as input to ANN.
5. Perform testing of the trained ANN on all Y component pairs ($A_i ; B_i$). The output of trained ANN is then used to decide whether block of image A or of image B should be selected for constructing the fused image F.
6. Map the resultant fused image, F from YCbCr color space to RGB color space.

RESULTS AND DISCUSSION

The performance of the proposed multifocus image fusion algorithms based on ANN in YCbCr color space has been experimented on four sets of multifocus images, created artificially by blurring different parts of corresponding reference images of size 512×512. Blurring is achieved by using a Gaussian disc of different radius (10 for set 1 image and 15 for set 2 images). Twenty-five percent of blocks of size 32×32 of set 1 input image are selected for training the neural network. Trained neural network is then tested on the remaining blocks to obtain final fused image.

Table 1: *Quantitative Comparison for Image Set 1.*

Methods	MSE	PSNR	Entropy	SD
DWT	83.1472	28.9663	7.6737	56.3472
LWT	82.5045	29.0000	7.6769	56.4435
LPT	11.6901	37.4866	7.7341	58.2119
PCA	22.0640	34.7280	7.6698	56.4224
RBFN	6.4621	40.0611	**7.7452**	58.3466
PNN	**5.2886**	**40.9314**	7.7446	**58.3619**



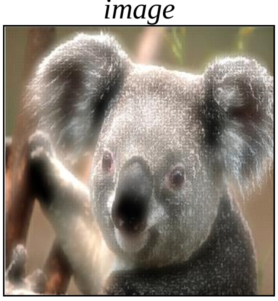
(a) Top-bottom blurred image *(b) Center-blurred image*
(c) DWT fused image *(d) LWT fused image*
(e) LPT fused image *(f) PCA fused image*
(g) RBFN fused image *(h) PNN fused image*

Fig. 4: *(a) to (h) Subjective Image Quality Assessment of Image Set 2.*

Table 2: *Quantitative Comparison for Image Set 2.*

Method	MSE	PSNR	Entropy	SD
DWT	290.3066	23.5362	7.7364	61.4677
LWT	280.7977	23.6809	**7.7489**	61.8172
LPT	20.8825	34.9670	7.7381	64.3512
PCA	43.0927	31.8208	7.7214	60.2147
RBFN	16.9951	35.8616	7.7258	**64.7452**
PNN	**16.8253**	**35.9052**	7.7250	64.7350


(a) Border blurred image *(b) Center blurred image*

(c) DWT fused image *(d) LWT fused image*

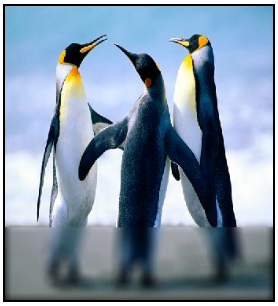

(e) LPT fused image *(f) PCA fused image*


(g) RBFN fused image *(h) PNN fused image*

Fig. 5: *(a) to (h) Subjective Image Quality Assessment of Image Set 3.*

The fusion results of the proposed algorithm have been compared both subjectively and objectively with the performance of DWT [12], LWT [11], LPT [14] and PCA [15]-based image fusion algorithms. From subjective results shown in Figures 3–6, it is observed that fused images obtained using proposed RBFN (Figures 3(g)–6(g)) and proposed PNN (Figures 3(h)–6(h)) are visually more pleasing with every part in focus. These images contain sharp features and better visibility then the fused images obtained using DWT (Figures 3(c)–6(c)), LWT (Figures 3(d)–6(d)), LPT (Figures 3(e)–6(e)) and PCA (Figures 3(f)–6(f)) algorithms.

For objective comparison, various parameters such as MSE, PSNR, entropy and standard deviation (SD) have been used. MSE and PSNR are the two error measures used to compare the original image and the fused image. MSE measures the cumulative squared error whereas PSNR measures the peak error.




(a) Bottom blurred image *(b) Top blurred image*


(c) DWT fused image *(d) LWT fused image*


(e) LPT fused image *(f) PCA fused image*


(g) RBFN fused image *(h) PNN fused image*

Fig. 6: *(a) to (h) Subjective Image Quality Assessment of Image Set 4.*

Table 3: *Quantitative Comparison for Image Set 3.*

Methods	MSE	PSNR	Entropy	SD
DWT	327.6638	23.0105	7.7985	57.1260
LWT	337.8898	22.8770	7.8014	57.5553
LPT	20.8732	34.9689	**7.8582**	60.6764
PCA	55.9882	30.6838	7.7625	54.7096
RBFN	**1.8244**	**45.5537**	7.8427	**60.8542**
PNN	1.8865	45.4082	7.8422	60.8255

Table 4: *Quantitative Comparison for Image Set 4.*

Methods	MSE	PSNR	Entropy	SD
DWT	239.6013	24.3699	7.7669	76.6220
LWT	249.0176	24.2025	7.7580	76.8732
PCA	29.8780	33.4113	**7.8161**	78.0731
RBFN	3.9756	42.1708	7.7154	80.4013
PNN	**3.9408**	**42.2089**	7.7155	**80.9074**

Mathematically,

$$MSE = \frac{1}{RS}\sum_{r=0}^{R-1}\sum_{s=0}^{S-1}\left[I(r,s) - F(r,s)\right]^2$$

Where, I is the original image. And

$$PSNR = 10\log_{10}\left[\frac{255^2}{MSE}\right]$$

Entropy and standard deviation [5] are the statistical parameters used for analyzing the texture of an image by quantifying the spatial variations in pixel intensities. Mathematically,

$$Entropy = -\sum_{l=0}^{255} p_l \log_2(p_l)$$

Where, p_l is the probability associated with l^{th} gray level of image. And

$$SD = \sqrt{\frac{\sum_{r=0}^{R-1}\sum_{s=0}^{S-1}(f(r,s) - \mu)}{RS}}$$

where, μ is the mean of image F. The values of these parameters for different algorithms are shown in Table 1, Table 2, Table 3 and Table 4. Before analyzing these values, it should be remembered that an image with high value of PSNR, high value of entropy and high value of SD whereas with low value of MSE reflects a better quality of image. The best value of these parameters is shown in bold in the Tables 1–4. From the Tables 1–4, it is observed that the PNN method of fusion has maximum parameters with their best values and, therefore, is considered to be a better method of multifocus image fusion.

CONCLUSION

This study proposes ANN-based multifocus image fusion method in YCbCr color space. Five different features reflecting the clarity of image blocks are extracted and used by ANN to determine which input block is clearer and is to be used for reconstructing the final fused image. Both the subjective and objective results presented in the study prove the superiority of the proposed algorithm for all sets of input images, in comparison to standard DWT, LWT LPT and PCA-based fusion algorithms.

REFERENCES

1. Manchanda M, Sharma R. Fuzzy transform-based fusion of multiple images. *International Journal of Image and Graphics*. 2017; 17(2): 1750008.
2. Zhang Q, Guo BL. Multifocus image fusion using the non-subsampled contourlet transform. *Signal Processing*. 2009; 89(7): 1334–46.
3. Shutao Li, James T Kwok, Yaonan Wang. Multifocus image fusion using artificial neural networks. *Pattern Recognition Letters*. 2002; 23(8): 985–97.
4. Zheng Liu, Lagani`ere R. Phase congruence measurement for image similarity assessment. *Pattern Recognition Letters*. 2007; 28(1): 166–72.
5. Manchanda M, Sharma R. A novel method of multimodal medical image fusion using fuzzy transform. *Journal of Visual Communication and Image Representation*. 2016; 40: 197–217.
6. Manchanda M, Sharma R. Fuzzy transform-based fusion of multiple images. *International Journal of Image and Graphics*. 2017; 17(2): 1750008.
7. Mertins A. Multifocus image fusion by combining curvelet and wavelet transform. *Optical Engineering*. 1999; 38(6): 991–1000.
8. Mark JL Orr et al. Introduction to radial basis function networks. *Centre for Cognitive Science University of Edinburg, Scotland*. 1996.
9. Piella G. A general framework for multiresolution image fusion from pixels to regions. *Information Fusion*. 2003; 4(4): 259–80.

10. Pohl C, Van Genderen JL. Review article multisensor image fusion in remote sensing: concepts, methods and applications. *International Journal of Remote Sensing*. 1998; 19(5): 823–54.
11. Xiaohong Xiao, Zhihong Wu. Image fusion based on lifting wavelet transform. *International Symposium on Intelligence Information Processing and Trusted Computing (IPTC)*. 2010; 659–62.
12. Yong Yang. A novel DWT based multi-focus image fusion method. *Procedia Engineering*. 2011; 24: 177–81.
13. Zhang Q, Guo BL. Multifocus image fusion using the non-subsampled contourlet transform. *Signal Processing*. 2009; 89(7): 1334–46.
14. Wang W, Chang F. A multi-focus image fusion method based on Laplacian pyramid. *Journal of Computers*. 2011; 6(12): 2559–66.
15. Naidu VPS, Raol JR. Pixel-level image fusion using wavelets and principal component analysis. *Defence Science Journal*. 2008; 58(3): 338.

Cite this Article
Meenu Manchanda, Deepak Gambhir. Artificial Neural Network-Based Multifocus Image Fusion. *Journal of Artificial Intelligence Research & Advances*. 2018; 5(3): 16–23p.