Scientific Journal of Impact Factor (SJIF): 5.71
e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 5, Issue 04, April -2018

Neural Network Based Multi-Focus Image Fusion

Riddhi Shukla[1] Pragnesh Patel[2]

[1,2] *Computer Engineering Department, Ipcowala Institute of Engineering & Technology, Dharmaj*

Abstract:- *The imaging equipment usually has difficulty in shooting the target object in which all the objects are effectively in focused. Image fusion plays a vital role in many applications. To overcome it multi-focus image fusion technology has emerged. An image is corrupted by noise blurring or limited focal length or due to different sensors and can have the poor visual quality. Image fusion is used to enhance the quality of a degraded image. Detection of the focused region is the key issue of the multi-focus image fusion algorithm. The proposed fusion method exploits the capabilities of artificial neural networks. Moreover, the learning capability of neural networks makes it feasible to customize the image fusion process. The experimental results show that the proposed method can perform better than the wavelet transform based method in some situations.*

Keywords: *Multi-focus image fusion, Image decomposition, CNN, neural network*

I. INTRODUCTION

Image fusion is an important research topic in many related areas such as computer vision, automatic object detection, remote sensing, image processing, robotics, and medical imaging. Image fusion is a sub-field of image processing in which more than one images are fused to create an image where all the objects are in focus. Image fusion is of significant importance due to its application in medical science, forensic and defence departments. The process of image fusion is performed for multi-sensor and multi-focus images of the same scene.
Multi-sensor images of the same scene are captured by different sensors whereas multi-focus images are captured by the same sensor. In multi-focus images, the objects in the scene which are closer to the camera are in focus and the farther objects get blurred. Contrary to it, when the farther objects are focused then closer objects get blurred in the image.

II. LITERATURE REVIEW

[1]CigdemTuran, Kin-Man Lam, Xiangjian[4]. In this paper, an emotion-based feature fusion method is using the Discriminant-Analysis of Canonical Correlations (DCC) for facial expression recognition. In our proposed method, four effective descriptors for facial expression representation, namely Local Binary Pattern (LBP), are considered. Supervised Locality Preserving Projection (SLPP) is applied to the respective features for dimensionality reduction and manifold learning. Experiments show that descriptors are also sensitive to the conditions of images, such as race, lighting, pose, etc.

[2] HannandeepKau, Jyoti Rani [10] have presented the survey of existing fusion schemes and a novel approach of panchromatic and multispectral images fusion using Discrete Wavelet Transform (DWT) in his paper entitled “Image Fusion on Digital Images using Laplacian Pyramid with DWT”.

[3] YukheLavinia, Holly H. Vo,AbhishekVerma [11] have studied on Fusion method that concatenates two and three deep convolutional neural networks (CNN). After examining the classification accuracy of each deep CNN candidates, we inserted a 100 dimensional fully-connected layer and extracted features from the new 100 dimensional and the last fully-connected layers to create a pool of candidate layers. We forwarded the concatenated features to Random Forest and SVM for classification. Compared to video, still images cost less memory footprint. The efficacy of still image in capturing the desired action, however, depends on the actions.

III. CONVOLUTIONAL NEURAL NETWORK

CNN is network architecture for deep learning. It is made of several layers that process and transforms the input image to produce an output image. The concept of CNN is obtained as a result of combining the two elements called neural networks and convolutions. The neural networks are composed of artificial neurons which simulate biological neurons in a limited way.

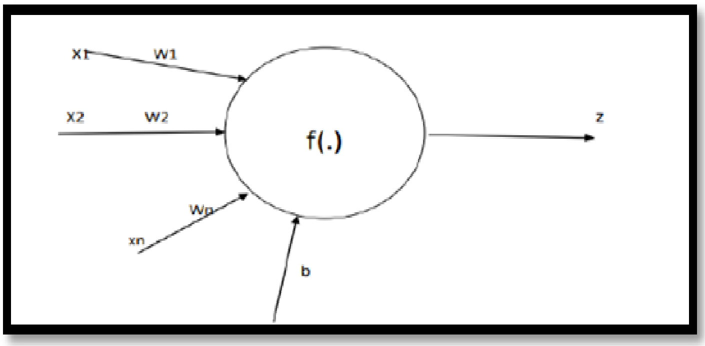

Figure 1. Artificial Neuron

CNN is an emblematical depth learning model that attempts to learn a hierarchical representation of an image at different abstraction levels [31]. As shown in Fig. 1, a typical CNN is mainly composed of an input layer, convolution layer, maxpooling(subsampling), fully connected layer and output layer. The input of the convolution neural network is usually the original image *X*. In this paper, we use *Hi* to represent the feature map of the *i*-th layer of the convolution neural network (*H*0 D *X*). Assuming that *Hi* is the convolution layer, the generation process of *Hi* can be described as follows:

$$Hi = f(Hi\text{-}1 * Wi + bi) \quad \text{-- -----------} \ (1)$$

where *Wi* is the convolutional kernel, *bi* is the bias. and indicates the convolutional operation. Here, *f* (0) is the nonlinear ReLU activation function.

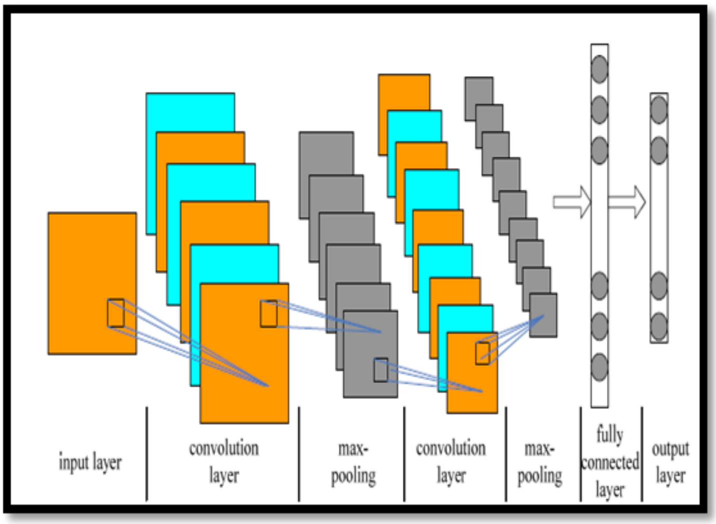

Figure 2. Typical Structure of a CNN.

A deep neural network combines multiple nonlinear processing layers, using simple elements operating in parallel and inspired by biological nervous systems. It consists of an input layer, several hidden layers, and an output layer. The layers are interconnected via nodes, or neurons, with each hidden layer using the output of the previous layer as its input.

***Salient* features of Deep Learning:**

1. Deep Learning can automatically learn features from the images. That means one can absolutely skip the manual feature extraction step.
2. Deep Learning is “black box” technique. In the sense it does not require to learn about the network at all. One can use the network just like a black box.
3. Deep learning updates learned weights at each layer.

FUSION ALGORITHM USING CNN:

1. The two source images are fed to a pre-trained CNN model to output a score map, which contains the focus information of source images. Particularly, each coefficient in the score map indicates the focus property of a pair of corresponding patches from two source images.
2. The focus map is segmented into a binary map with a threshold of 0.5.
3. We refine the binary segmented map with two popular consistency verification strategies, namely, small region removal and guided image filtering , to generate the final decision map.
4. In the last step, the fused image is obtained with the final decision map using the pixel-wise weighted-average strategy.

Data Set

We have used the data set of more than 20 inputs from different dimensions and different sizes. We have tested our Proposed work on these images and these images are collected personally from websites, online images etc.

Performance Metrics

1) **PSNR**: It means Peak signal to noise ratio. It IS an expression for the ratio between the maximum possible values (power) of a signal and the power of distorting noise that affects the quality of its representation. In this PSNR calculates the how much resultant image accurate.

$$PSNR = 10.\log_{10}\left(\frac{MAX_1^{\,2}}{MSE}\right)$$
$$= 20.\log_{10}\left(\frac{MAX_1}{\sqrt{MSE}}\right)$$
$$= 20.\log_{10}(MAX_1) - 10.\log_{10}(MSE)$$

2) **MSE**: Mean Square Error it means how much errors Occur in the system.

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - K(i,j)]^2$$

3) **Entropy**: Entropy is defined according order and disorder where entropy most probably defined the additional information extract from the physical present system or input images either from transformed images to further enhancement of fusion.
5) **SSI**: Structural similarity index depends on the human perception
6) **MI.:** Mutual Information tells about familiar information between input images.
7) **SD:** Standard deviation measures the contrast of images.

IV. DICUSSION AND RESULTS

1. One of the Important Quality metric in the Image and it was computed between source Images A, B and Fused Image.
2. Results are compared with the existing pyramid fused image results and our method gave the better results than the existing pyramid fusion technique.
3. For the construction of weight maps from the original source images we took the help of most popular convolution neural network model Alex net model.

This section analyzes the proposed system with the accuracy measures or quality measures. These parameters are performed in MATLAB. A. Test Images and Statistical Parameters in the analysis of gray scale and RGB images by using technique have more accurate results in quality measures like PSNR,MSE, Entropy, SSI, MI, SD. The following table Table I shows the test images and their resultant image after fusion.

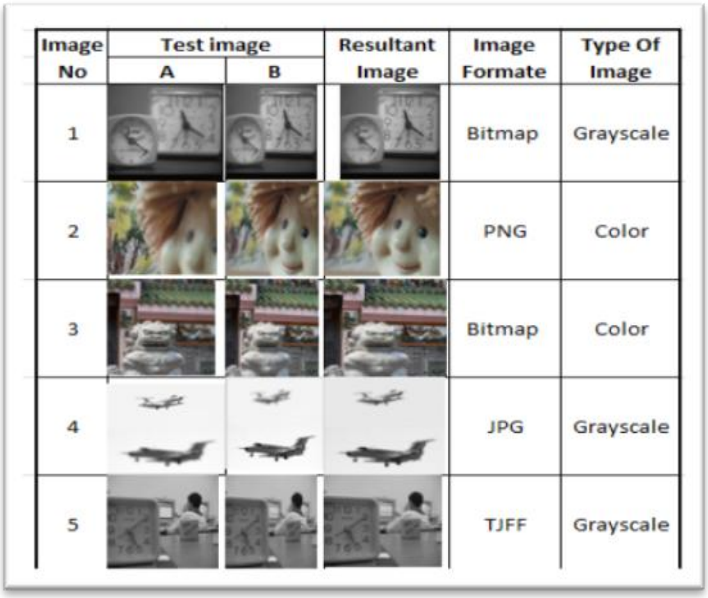

Image No	Test image A	Test image B	Resultant Image	Image Formate	Type Of Image
1				Bitmap	Grayscale
2				PNG	Color
3				Bitmap	Color
4				JPG	Grayscale
5				TJFF	Grayscale

Table I : List of Images

Analysis of proposed system

In this proposed system tested the images through this CNN method on images. We analyze the results of proposed system and calculate the all performance metrics under the mutual information (MI). Quality measures show much better accuracy than existing methods.

Table II Performance Parameters Of Images

Image No.	PSNR	MSE	Entro	SSI	M.I	S.D
1	60.59	0.0256	4.5	0.4589	0.123	45.56
2	63	0.0356	6.23	0.5654	0.0546	48.23
3	54.6	0.0125	7.12	0.4545	1.1112	23.45
4	64.89	0.0212	3.72	0.8862	1.8056	50.126
5	64.39	0.0458	8.92	0.6565	1.5689	48.35

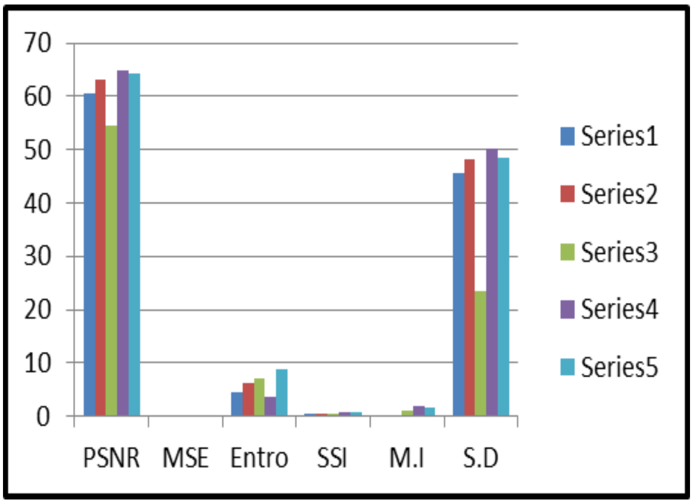

Figure 3 Quality measures of Images

The following Table III and Fig.4 shows the comparison of existing results and proposed result. Proposed results are more accurate compared to existing work parameters of methods.

TABLE III Comparison of Performance Parameters

Parameters	Existing	Proposed
PSNR	62.2506	64.89
MSE	0.0643	0.0212
Entropy	7.1365	3.72
SSI	0.9861	0.8862
MI	1.0219	1.8056
SD	19.8603	50.126

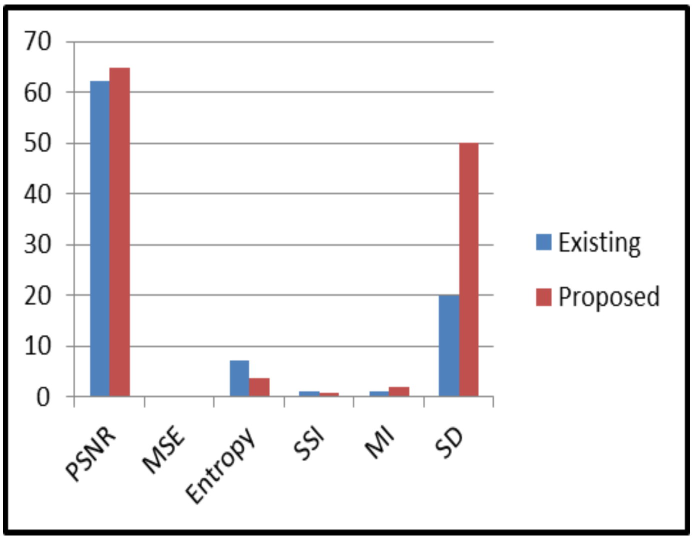

Figure 4. Comparison of Existing and Proposed System.

V. CONCLUSION AND FUTURE SCOPE

At the end we conclude that, By using CNN we can classified the best feature by providing training as well as we can also achieve better fused image that will not use during the training process. With that we can also achieve better fused image with less processing time. Image fusion provides an effective way to extract the relevant information from the two or more input images and merge them in to a single image which is more informatics. In this paper based on CNN, how it is helpful in extracting the image features was discussed. By using automated extracted features from the CNN model weight maps for different types of images were constructed. These weight maps are used in getting the final fused image with the help of CNN along with the original images. The experiments are carried out using mat lab and our proposed method results shown better PSNR values and batter MI value. In future, interesting issue is the size of training patches, which is set to a training patch pair. This may lead to inaccuracy around the boundaries between focused and defocused regions.

REFERENCES

[1] Matsuda, Yuji, Hajime Hoashi, and KeijiYanai. "Recognition of multiple-food images by detecting candidate regions." Multimedia and Expo (ICME), 2012 IEEE International Conference on. IEEE, 2015.

[2] P.F. Felzenszwalb, R.B. Girshick, D. McAllester, and D. Ramanan,"Object detection with discriminatively trained part-based models,"IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.32, no. 9, pp. 1627–1645, 2010.

[3] Y. Deng and B. S. Manjunath, "Unsupervised segmentation of colortextureregions in images and video," IEEE Transactions on PatternAnalysis and Machine Intelligence, vol. 23, no. 8, pp. 800–810, 2001.

[4] Turan, Cigdem, Kin-Man Lam, and Xiangjian He. "Facial expression recognition with emotion-based feature fusion." Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2015 Asia-Pacific. IEEE, 2015.

[5] Olding, Willem C., Jan C. Olivier, and Brian P. Salmon. "A Markov random field model for decision level fusion of multi-source image segments." Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International. IEEE, 2015.

[6] R. Nishii, "A markov random field-based approach todecision-level fusion for remote sensing image classification,"Geoscience and Remote Sensing, IEEE Transactionson, vol. 41, no. 10, pp. 2316–2319, Oct 2003.

[7] A. H. Solberg, T. Taxt, and A. K. Jain, "A markov randomfield model for classification of multisource satelliteimagery," Geoscience and Remote Sensing, IEEETransactions on, vol. 34, no. 1, pp. 100–113, 1996.

[8] T. Hoberg, F. Rottensteiner, and C. Heipke, “Contextmodels for crf-based classification of multitemporal remotesensing data,” ISPRS Annals of Photogrammetry,Remote Sensing and Spatial Information Sciences,vol. 25, pp. 128–134, 2012.

[9] T. Hoberg, F. Rottensteiner, R. QueirozFeitosa, and C. Heipke, “Conditional random fields for multitemporal and multiscale classification ofoptical satellite imagery,”Geoscience and Remote Sensing, IEEE Transactionson, vol. 53, no. 2, pp. 659–673, Feb 2015

[10]Kaur, Harmandeep, and Jyoti Rani. "Image fusion on digital images using Laplacian pyramid with DWT." Image Information Processing (ICIIP), 2015 Third International Conference on. IEEE, 2016.

[11]Lavinia, Yukhe, Holly H. Vo, and AbhishekVerma. "Fusion Based Deep CNN for Improved Large-Scale Image Action Recognition." Multimedia (ISM), 2016 IEEE International Symposium on. IEEE, 2016.

[12] Yu Liu, Xun Chen, Hu Peng, Zengfu Wang "Multi-focus image fusion with a deep convolutional neural network", INFORMATION FUSION, vol.36, pp.191-207, 2017.

[13] K. He , J. Sun , X. Tang , Guided image filtering, IEEE Trans. Pattern Anal. Mach. Intell. 35 (6) (2013) 1397–1409

