Journal of Scientific & Industrial Research
Vol. 80, October 2021, pp. 905-914

# WeAbDeepCNN: Weighted Average Model and ASSCA based Two Level Fusion Scheme For Multi-Focus Images

Vineeta Singh* and Vandana Dixit Kaushik

Department of Computer Science and Engineering, Harcourt Butler Technical University, East Campus, Nawabganj, Kanpur, Uttar Pradesh 208 002, India

*Received 25 February 2021; revised 26 May 2021; accepted 19 September 2021*

Fusion of images is a strategy that merges various moderately focused images or non-focused images of a single scene to generate a fully focused, clear and sharp image. The goal of this research is to discover the focused regions and further combination of focused regions of different source images into solitary image. However, there exist several issues in image fusion that involves contrast reduction, block artifacts, and artificial edges. To solve this issue, a two level fusion scheme has been devised, which involves weighted average model along with Atom Search Sine Cosine algorithm-based Deep Convolutional Neural Network (ASSCA-based Deep CNN) and may be abbreviated as "WeAbDeepCNN" i.e. weighted average model and ASSCA based Deep CNN. In the study two images are fed to initial fusion module, which is performed using weighted average model. The fusion score are generated whose values are determined in an optimal manner. Thus, final fusion is performed using proposed ASSCA-based Deep CNN. The Deep CNN training is carried out with proposed ASSCA, which is devised by combining Sine Cosine Algorithm, abbreviated as SCA, as well as atom search optimization (ASO). The proposed ASSCA-based Deep CNN offers improved performance in contrast to current state of the art techniques with a highest value 1.52 of mutual information (MI), with a highest value of 32.55 dB of maximum Peak Signal to Noise Ratio i.e. PSNR as well as value of 7.59 of Minimum Root Mean Square Error (RMSE).

**Keywords**: Atom search optimization, Deep convolutional neural network, Image fusion algorithm, Optimization technique, Multi-focus image fusion

## Introduction

The majority of computer vision services like detection of object, and identification utilize focused images as key in spite of blur image. Due to a limitation of camera lens, objects situated in the focus area of the camera lens appears sharp and focused while other objects which are either at certain interval or behind the focus, appears blurred or un-focused.[1,2] Nonetheless, because of less depth-of-field (DOF), it becomes complex to detain all decisive imagery. Thus, image fusion methods are adapted for fusing several moderately-focused or un-focused regions of the source images for generation of focused fused image.[3] It is imperative in domains of computing, medical imaging, microscopic imaging, remote sensing, geographical area, classification and handling images and acquired immense consideration in current day to day lives.[4–9] Multi-focus image fusion is an effective techniques adapted for solving the issues. In this technique, the shooting of different images of similar scene using various settings is performed. Thereafter, the original images are combined for getting composite image wherein all objects from same scene are focused. The composite image termed as fused image is apposite for individual observation. The multi-focus images are employed in several domains that involve machine vision, digital imaging and microscopic imaging.[10]

The majority of representative techniques for fusing images are devised involving proportion of low-pass pyramid (RP)[11], curvelet transform (CVT)[12], non-sub-sampled contourlet transform (NSCT)[13], Laplacian pyramid (LP)[14], gradient pyramid (GP)[15], sparse representation (SR)[16] discrete wavelet transform (DWT).[17] In this technique, the choice of transform province and fusion rule discovers ocular insight of combined outcomes.[6] Various algorithms are devised for fusing image. These methods are categorized in different classes that involve sparse representation-based techniques, and CNN-based strategy. Other method adapted for multi-focus image fusion is morphological pyramid (MP). The aforementioned strategies pose the ability to

*Author for Correspondence
E-mail: vineeta.singh.cs@gmail.com

effectually mine features from focused areas. However, the majority of the techniques are complicated and lengthy process.[18]

The purpose is devising a novel image fusion technique, namely ASSCA-based Deep CNN to fuse effectively multi-focus images. In the proposed fusion technique, two input images are submitted to initial fusion phase, which is carried out using weighted average model. Here, the fusion score are produced whose values are discovered in an optimal manner. Thus, final fusion is carried out which is performed using is proposed ASSCA-based Deep CNN. The training of Deep CNN is done using ASSCA, which is devised by combining ASO and SCA. The key contribution is:

**Development of ASSCA-based Deep CNN for Fusing Multi-Focus Images**

The ASSCA-based Deep CNN is devised to fuse images. Here, deep CNN training is carried out with ASSCA and is obtained by integrating ASO and SCA.

Though several techniques has been devised for image fusion via Deep CNN model[19] but still there is a scope in this area for training of DCNN via enhanced optimized techniques. Our research includes an optimization based algorithm named as ASSCA to train the Deep CNN to combine multi-focus images.

We are adapting the two layers of fusion scheme in our proposed model. Initially fusion will be accompanied by weighted average model i.e. fusion factors will be generated and second level of fusion will include training of Deep CNN via ASSCA algorithm, where ASSCA is combination of ASO as well as SCA algorithm.

**Related Work**

Zhang *et al*.[4] devised a larger as well as realistic dataset consisting of realistic multi-focus images for training purpose in deep learning algorithms. Variety, diversity, size as well quality of the devised dataset is needed to be maintained and enhanced.

Xu *et al*.[6] devised a multi-focus image fusion framework by forwarding connected regions as well as gradients relied unsupervised technique. The concept of Mask-Net was proposed and Mask-Net was trained for generation of binary mask on direct basis and every coefficient there was indicating for a focused pixel or for a unfocused pixel. The integration of the loss functions were executing of connected regions as well as of gradients to deliver the mask representing above relevance. The devised method was an unsupervised technique where the synthesis of ground-truth data was a secondary term. But the size of dataset taken to implement the methodology was not the large one.

Hu *et al*.[12] developed a novel fusion scheme for fusion of multi-focus images, moreover use of curvelet transform was done. Here energy of two images was calculated via maximum local energy technique. Initially, curvelet transform was applied for generation of coefficients of two input images focused at different locations of the same scene. At next stage of fusion scheme, the selection of coefficients possessing low frequency was accompanied by maximum local energy and with the help of sliding window final outcome was generated as maximum energy pixel detail. At subsequent stage, selection of high frequency coefficients was fulfilled with the help of absolute maximum technique. At last stage, the melded image was generated via application of inverse curvelet transform (ICVT).

Chen *et al.*[14] devised a novel fusion strategy relied on detection of focused regions which was fulfilled with the help of Gaussian filter as well as point detection filter. These mentioned filter techniques proven to be much effective. Researchers devised a consistency verification method relied on super-pixel via integration of superpixels of image for enhancement of effectiveness of the fusion technique. The use of image superpixels was done to illustrate local features of images in a perceptual manner.

Zhang *et al*.[20] devised generative adversarial network using gradient joint and adaptive parameter for fusing multi-focus images. Here, an adaptive decision block was devised for discovering if source pixels are focused or not using disparity of repeated blur. The content loss factor was devised for dynamically guiding optimization inclination that force to generate fused outcome. To improve texture details, an adversarial game were incorporated on gradient map relied on input images. But, the technique was unsuccessful to include multi-modal deep learning for performing different fusion tasks.

Panigrahy *et al*.[21] devised a strategy considering adaptive DCPCNN (PA-DCPCNN) wherein attributes were evaluated in an adaptive manner considering the inputs. In addition, the linking potency of the PADCPCNN was evaluated with fractal dimension based focus measure i.e. FDFM. The fused images were generated by adapting inverse NSCT to fused sub-bands. However, the method failed to apply computationally effective multi-scale multi-direction transforms in order to minimize computation cost.

Guo *et al*.[22] devised fusion technique to fuse multi-focus images involving fully convolutional network for focus detection, in short, FD-FCN. For obtaining more accurate focus detection maps, the addition of skip layers in the network was incorporated for making both comprehensive and abstract visual data to produce maps. The fusion of image considering FD-FCN comprises three processes in which focus maps were generated considering FD-FCN and generation of decision map occurred with decision map. However, the method failed to involve improved-performing networks to generate precise focus maps.

Ma *et al*.[23] devised deep architecture for fusing images. Here, encoder–decoder network was trained with unsupervised way for acquiring deep features. Spatial frequencies were utilized considering gradient-assisted technique for measuring sharp alteration with deep features for reflecting the levels of activities. In addition, the consistency validation techniques were utilized for adjusting decision map to generate fused result.

Han *et al*.[24] devised a novel image fusion technique to fuse hyperspectral images via use of optimized twin dictionaries. Here HM and LH images were fused to gain the fused HH image extracted out of the same scene. The twin (spatial dictionary as well as spectral dictionary) dictionaries used here after the optimization were capable of combining total information including spectral information and spatial information leveraged with high frequency to generate the combined output HH images parallel. Optimization of spectral dictionary accompanied by minimization of augmented Lagrangian function with the help of ADMM excluding positive restrictions. Learning of spatial dictionary was accompanied by K-SVD algorithm by residue spatial detail remaining via the spectral dictionary. Spatial as well as spectral preservation was better in this fusion model. But the method was used only for hyper-spectral image fusion purpose.

**Challenges**

The Challenges faced by traditional image fusion techniques are illustrated as:

In the framework devised by Xu *et al*.[6] the dataset used did not have plenty and diverse images. The method of Zhang *et al*.[20] failed to include multi-modal deep learning for performing different fusion tasks. Panigrahy *et al*.[21] failed to apply computationally effective multi-scale multi-direction transforms in their method in order to minimize computation cost. The method of Guo *et al*.[22] failed to see improved-performing networks to generate precise focus maps.

**Proposed ASSCA-based Deep CNN for Multi-Focus Image Fusion**

Multi-focus image fusion method is termed as an imperative domain for obtaining composite image considering all the objects. The key role is to devise an effectual activity level measurement for computing the clarity of source images. The goal is combining group of images which are captured from similar scene but in different focus for generating other sharper image. The major problem in designing in image fusion technique for multi-focus images involves calculating information of local content from input images. Proposed fusion model ASSCA-based Deep CNN architecture has been demonstrated through Fig. 1. Here, two images are fed as an input to initial fusion wherein weighted average model is adapted for performing initial fusion. Thereafter, the output image is again fed to final fusion wherein Deep CNN[25,26] is employed for performing final multi-focus image fusion. Here, Deep CNN training is carried out by ASSCA algorithm, which is obtained by combining SCA[27] and ASO.[28]

Assume $C_{c,e}$ and $D_{c,e}$ be two multimodal images taken from the imaging repository $I$ one that contains $o$ images. These two input image undergoes weighted average modelfor performing initial fusion. The weighted average model is incorporated for performing initial fusion of images, and weight score for fusion is chosen in an optimal manner using proposed ASSCA algorithm.

**Initial Fusion With Weighted Average Model**

The weighted average model is incorporated to fuse both input images for acquiring fused images. In initial image fusion, the weighted average model fuses both input images considering weighted score. Here, the images undergo multiplication process in which the images are multiplied using fusion score $\alpha$ and $\beta$ that iteratively executes optimization process to discover optimum values. The fusion procedure with weighted average model is expressed as,

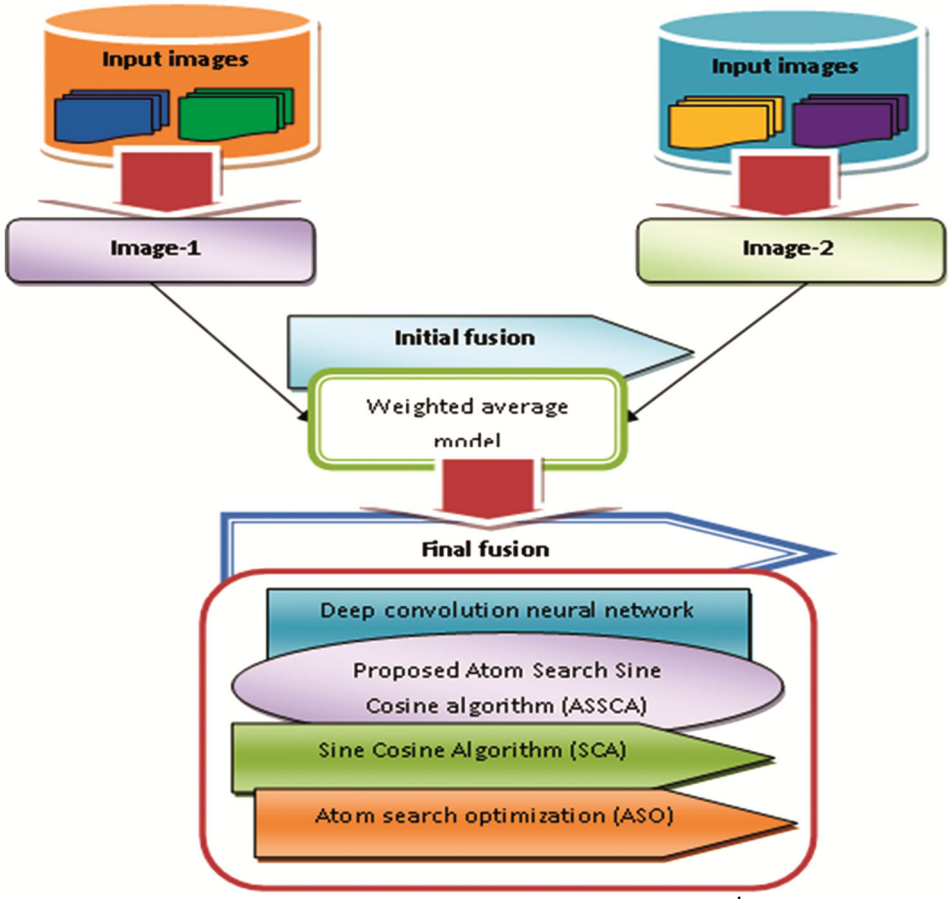


Fig. 1 — Architecture of Image Fusion Model using Proposed ASSCA-based Deep CNN

$$O = \alpha.C_{c,e} + \beta.D_{c,e} \quad \ldots (1)$$

**Final Image Fusion using Proposed ASSCA-Based DeepCNN**

The final image fusion is an effectual strategy employed for integrating pertinent information from set of images considering similar scene into comprehensive image. The fused images are generally informative than other source images. Here, the final fusion is performed by proposed ASSCA-based Deep CNN in which the optimal fusion score is computed. The ASSCA is employed for training Deep CNN that helps to tune optimal weights. In addition, the ASSCA is devised by combining SCA and ASO. The architectural design of Deep CNN and the training procedure of ASSCA algorithm are illustrated below:

**Architecture of Deep CNN**

A structural detail as well as elaboration of the DCNN is presented in Fig. 2. The Deep CNN comprises Fully Connected (FC), convolutional (conv), and pooling (POOL) layers, wherein each

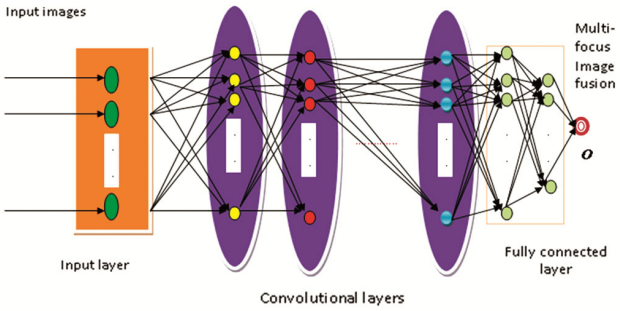


Fig. 2 — Structural Design of Deep CNN

layer is responsible for performing particular mission. The FC helps to perform final image fusion.

*Conv Layers:* This layer is utilized to generate features and provides extraction of pattern considering the feature maps which is produced with objects. Here, the neurons are associated using trainable weights and are convoluted with inputs for devising feature maps. Moreover, outcome is subjected to non-linear activation to enhance complicated functional mappings among response variables. The convolution layer input and count of conv layers is modeled as,

$$T = \{T_1, T_2, ..., T_h, ..., T_i\} \quad \ldots (2)$$

where, $i$ symbolize total conv layers and $T_h$ signifies $h^{th}$ conv layer. The units positioned in $(p, w)$ obtains output which is formulated as,

$$\left(T_{\vartheta}^{h}\right)_{x,w} = \left(v_{\vartheta}^{h}\right)_{x,w} + \sum_{x=1}^{W_1^{x-1}} \sum_{z=-r_1^i}^{r_1^i} \sum_{y=-r_2^i}^{r_2^i} \left(X_{f,x}^{i}\right)_{z,y} * \left(T_{\vartheta}^{h-1}\right)_{x+z,w+y} \quad \ldots (3)$$

where, $*$ represent conv operator, which facilitates extraction of pattern using outputs produced with adjacent conv layers, $\left(T_{\vartheta}^{h-1}\right)_{x+z,w+y}$ signifies fixed feature maps, while $W_1^{x-1}$ represent total feature maps, as well as $\left(X_{f,x}^{i}\right)_{z,y}$ denote weight, trained with ASSCA algorithm.

*ReLU:* ReLU depicts an activation module which facilitates elevated efficiency and straight forwardness and insinuation of ReLU layer is that it works in quick manner and is able to deal using huge networks. The outcome generated with ReLU layer is forwarded to feature maps and may be modeled as following,

$$T_f^u = fun\left(T_f^{u-1}\right) \quad \ldots (4)$$

where, $fun(\,)$ represents in $u^{th}$ layer, the activation function.

***FC Layers:***

The patterns generated using conv and pooling layers are subjected as an input to FC layers to commence image fusion. The output produced through FC is modeled as,

$S_f^u = Z\left(a_f^u\right)$ with

$$a_f^u = \sum_{p=1}^{W_1^{p-1}} \sum_{m=1}^{W_2^{p-1}} \sum_{n=1}^{W_3^{p-1}} \left(V_{f,p,m,n}^{u}\right).\left(T_f^{u-1}\right)_{m,n} \quad \ldots (5)$$

Here, $V_{f,p,m,n}^{u}$ represent weight associating $(m, n)$ at feature map $p^{th}$, for $u-1$ layer as well as $f^{th}$ unit for $u$ layer. Here, weights tuning is done in an optimal manner considering ASSCA algorithm.

**Training of Deep CNN**

The ASSCA is adapted to choose optimum value for generating fusion score in order to generate effective image fusion. ASSCA development is accompanied via combination of ASO+SCA and is adapted by acquiring the benefits of both techniques. SCA is devised considering sine and cosine functionalities. The algorithm provides a tradeoff between exploration and exploitation states to discover search space regions and to assist to attain global optimum. This technique is helpful for avoidance of local optima as well as involves high exploration that assists to solve real-world problems. Meanwhile, ASO is inspired from the molecular dynamics which assist to solve engineering issues in real-world. Atom search optimization algorithm searches for local regions and further it executes nicely along with unimodal functions/procedures as well as represents effective performance demonstration via utilizing unimodal functions/procedures. In addition, ASO offers a balance between exploitative and explorative nature that helps to handle complicated optimization functions. The sine cosine algorithm is combined in the ASO update equation for augmenting the ASO performance via choosing optimal fusion score.

On the other hand, the ASSCA algorithm offers improved balance allying exploitation as well as exploration, subsequently ASSCA efficiency improves. Here, the issues of SCA are addressed with ASO that offers enhanced rate of convergence that produce optimal solution. Steps series of ASSCA involves:

*Initialization*

Starting step involves solution initialization, whose positions are initiated in an arbitrary manner in parallel with other algorithmic attributes expressed as,

$$A = \{A_1, A_2, \cdots, A_x, \cdots A_y\}; 1 < x \le y \quad \ldots (6)$$

where, $A_x$ symbolizes $x^{th}$ solution and $y$ signifies total solutions.

*Determination of Error*

The generation of optimal solution with fitness is considered as problem of minimization. Hence, solution with least MSE is chosen as optimal solution. Hence, MSE are evaluated as,

$$MS_{err} = \frac{1}{o} \sum_{b=1}^{o} \left[\xi_b - \xi_b^*\right]^2 \quad \ldots (7)$$

where, $\xi_b$ denote output which is expected and $\xi_b^*$ indicate output which is to be predicted, $o$ symbolize number of data samples, where $1 < b \le o$.

*Evaluation of Update Position*

The ASO maintain property of improved self-adaptive convergence which assist each atom to interrelate with other and to increase algorithmic performance. For making it understandable, velocity as well as position of $g^{th}$ atom at $(d+1)^{th}$ iteration is updated and can be illustrated in following manner,

$$A_g(d+1) = A_g(d) + b_g(d+1) \quad \ldots (8)$$

where, $b_g(d+1)$ signifies at $(d+1)$ iteration , the velocity of $g^{th}$ atom, while, $A_g(d)$ depicts at $d^{th}$ iteration the location of $g^{th}$ atom , as well as, $A_g(d+1)$ signifies, at $(d+1)^{th}$ iteration, the update location of $g^{th}$ atom.

$$A_g(d+1) = A_g(d) + Rand_g b_g(d) + k_g(d) \quad \ldots (9)$$

where, $Rand_g$ indicates arbitrary number in $[0,1]$ $k_g(d)$ represents, at $d^{th}$ iteration, the acceleration of $g^{th}$ atom, and $k_g(d)$ signifies, at $d^{th}$ iteration, the velocity of $g^{th}$ atom.

$$\mathrm{A_g(d+1) = A_g(d) + Rand_g b_g(b) - \lambda\left(1-\frac{d-1}{R}\right)e^{\frac{-20d}{R}} \sum_{1\in Kbest} \frac{Rand_1\left[2\times(j_{gl}(d))^{13}\right]-(j_{gl})^7}{p_g(d)}}$$
$$\mathrm{\frac{A_1(d)-A_g(d)}{\|A_g(d),A_1(d)\|_2} + \eta e^{\frac{-20d}{R}} \frac{A_{best}(d)-A_g(d)}{p_g(d)}} \quad \ldots (10)$$

where, $\lambda$ represents weight of depth, and $R$ indicates highest number of iterations while Kbest represents best suitable fitness value as well as $p_g(d)$ denotes $g^{th}$ atom mass at $d^{th}$ iteration procedure/function, whereas $\eta$ represents multiplier weight.

Sine Cosine Algorithm[27] is utilized to solve problems in real-time, as it prevent local optima by revealing search space to attain global optima. As per SCA, the update of solution is represented as,

For $t_4 < 0.5$,

$$A_g(d+1) = A_g(d) + t_1 Sin(t_2) \,|\, (t_3 N_g(d) - A_g(d)) \,| \quad \ldots (11)$$

where, $t_1$ denotes direction of movement, $t_2$ depicts at what distance movement should be accompanied either in the direction of the target or away from the target, $t_3$ yields weight randomly for target, and $t_4$ utilizes switching between components of sine as well as cosine, $A_g(d)$ denotes solution at the present, while $N_g(d)$ indicate location of destination point as well as $||$ symbolize absolute value. Let $N_g(d) > A_g(d)$, the update location is modeled as,

$$A_g(d+1) = A_g(d) + t_1 Sin(t_2) t_3 N_g(d) - t_1 Sin(t_2) A_g(d) \quad \ldots (12)$$

$$N_g(d) = \frac{A_g(d+1) - A_g(d) + t_1 Sin(t_2) A_g(d)}{t_1 t_3 Sin(t_2)} \quad \ldots (13)$$

As $A_{best}(d) = N_g(d)$, the update position of ASSCA algorithm is obtained by substituting Eq. (13) in Eq. (10) expressed as,

$$\mathrm{A_g(d+1) = A_g(d) + Rand_g b_g(b) - \lambda\left(1-\frac{d-1}{R}\right)e^{\frac{-20d}{R}} \sum_{1\in Kbest} \frac{Rand_1\left[2\times(j_{gl}(d))^{13}\right]-(j_{gl})^7}{p_g(d)}}$$
$$\mathrm{\frac{A_1(d)-A_g(d)}{\|A_g(d),A_1(d)\|_2} + \eta e^{\frac{-20d}{R}} \frac{A_g(d+1)-A_g(d)+t_1Sin(t_2)A_g(d)}{t_1t_3Sin(t_2)} - A_g(d)}$$
$$\mathrm{\frac{}{p_g(d)}} \quad \ldots (14)$$

$$\mathrm{A_g(d+1) = A_g(d) + Rand_g b_g(b) - \lambda\left(1-\frac{d-1}{R}\right)e^{\frac{-20d}{R}} \sum_{1\in Kbest} \frac{Rand_1\left[2\times(j_{gl}(d))^{13}\right]-(j_{gl})^7}{p_g(d)}}$$
$$\mathrm{\frac{A_1(d)-A_g(d)}{\|A_g(d),A_1(d)\|_2} + \eta e^{\frac{-20d}{R}} \frac{A_g(d+1)-A_g(d)+t_1Sin(t_2)A_g(d)-A_g(d)t_1t_3Sin(t_2)}{p_g(d)t_1t_3Sin(t_2)}} \quad \ldots (15)$$

$$\mathrm{A_g(d+1) = A_g(d) + Rand_g b_g(b) - \lambda\left(1-\frac{d-1}{R}\right)e^{\frac{-20d}{R}} \sum_{1\in Kbest} \frac{Rand_1\left[2\times(j_{gl}(d))^{13}\right]-(j_{gl})^7}{p_g(d)}}$$
$$\mathrm{\frac{A_1(d)-A_g(d)}{\|A_g(d),A_1(d)\|_2} + \frac{\eta e^{\frac{-20d}{R}} A_g(d+1)}{p_g(d)t_1t_3Sin(t_2)} - \frac{\eta e^{\frac{-20d}{R}} A_g(d)[1-t_1Sin(t_2)+t_1t_3Sin(t_2)]}{p_g(d)t_1t_3Sin(t_2)}} \quad \ldots (16)$$

$$A_g(d+1)-\frac{\eta e^{\frac{-20d}{R}}A_g(d+1)}{p_g(d)t_1t_3Sin(t_2)}=A_g(d)+Rand_g b_g(b)-\lambda\left(1-\frac{d-1}{R}\right)e^{\frac{-20d}{R}}\sum_{l\in Kbest}\frac{Rand_l\left[2\times\left(j_{gl}(d)\right)^3\right]-(j_{gl})^7}{p_g(d)}$$
$$\frac{A_l(d)-A_g(d)}{\|A_g(d),A_l(d)\|_2}-\frac{\eta e^{\frac{-20d}{R}}A_g(d)\left[1-t_1Sin(t_2)+t_1t_3Sin(t_2)\right]}{p_g(d)t_1t_3Sin(t_2)} \quad \ldots (17)$$

$$\frac{p_g(d)t_1t_3Sin(t_2)A_g(d+1)-\eta e^{\frac{-20d}{R}}A_g(d+1)}{p_g(d)t_1t_3Sin(t_2)}=A_g(d)+Rand_g b_g(b)-\lambda\left(1-\frac{d-1}{R}\right)e^{\frac{-20d}{R}}$$
$$\sum_{l\in Kbest}\frac{Rand_l\left[2\times\left(j_{gl}(d)\right)^3\right]-(j_{gl})^7}{p_g(d)}\frac{A_l(d)-A_g(d)}{\|A_g(d),A_l(d)\|_2}-\frac{\eta e^{\frac{-20d}{R}}A_g(d)\left[1-t_1Sin(t_2)+t_1t_3Sin(t_2)\right]}{p_g(d)t_1t_3Sin(t_2)} \quad \ldots (18)$$

where, $\lambda$ represents weight of depth, and $R$ indicates highest number of iterations while Kbest represents best suitable fitness value as well as $p_g(d)$ denotes $g^{th}$ atom mass at $d^{th}$ iteration procedure/function, whereas $\eta$ represents multiplier weight, $Rand_g$, $Rand_i$ are the random numbers generated in [0,1]|| symbolize absolute value.

$$\frac{A_g(d+1)\left[p_g(d)t_1t_3Sin(t_2)-\eta e^{\frac{-20d}{R}}\right]}{p_g(d)t_1t_3Sin(t_2)}=A_g(d)+Rand_g b_g(b)-\lambda\left(1-\frac{d-1}{R}\right)e^{\frac{-20d}{R}}$$
$$\sum_{l\in Kbest}\frac{Rand_l\left[2\times\left(j_{gl}(d)\right)^3\right]-(j_{gl})^7}{p_g(d)}\frac{A_l(d)-A_g(d)}{\|A_g(d),A_l(d)\|_2}-\frac{\eta e^{\frac{-20d}{R}}A_g(d)\left[1-t_1Sin(t_2)+t_1t_3Sin(t_2)\right]}{p_g(d)t_1t_3Sin(t_2)} \quad \ldots (19)$$

The update is carried out with ASSCA, and is computed such that values associated to minimal fitness are adapted to choose optimal value towards fusion coefficient represented via following given equation,

$$A_g(d+1)=\frac{p_g(d)t_1t_3Sin(t_2)}{p_g(d)t_1t_3Sin(t_2)-\eta e^{\frac{-20d}{R}}}\left[\begin{array}{l}A_g(d)+Rand_g b_g(b)-\lambda\left(1-\frac{d-1}{R}\right)e^{\frac{-20d}{R}}\\ \sum_{l\in Kbest}\frac{Rand_l\left[2\times\left(j_{gl}(d)\right)^3\right]-(j_{gl})^7}{p_g(d)}\\ \frac{A_l(d)-A_g(d)}{\|A_g(d),A_l(d)\|_2}-\frac{\eta e^{\frac{-20d}{R}}A_g(d)\left[1-t_1Sin(t_2)+t_1t_3Sin(t_2)\right]}{p_g(d)t_1t_3Sin(t_2)}\end{array}\right] \quad \ldots (20)$$

In Eq. (20), the tradeoff between exploitation and exploration are done in which ranges for sine as well as cosine are updated to highlight utilization of search spaces when iteration counter becomes maximum in terms of value. The SCA provides tradeoff between exploration and exploitation to discover optimal solutions.

*Re-Evaluation of Solution with Error*

The re-evaluation of error is done with solution generated in Eq. (7). The solution producing least error is utilized in deep CNN for image fusion.

*Termination*

Criteria for termination are verified for spotting optimization procedure end. The criteria for termination is illustrated for enhanced convergence property of algorithm, involving maximal iterations and improves percentage, and time-taken for implementation. The pseudo code of ASSCA is displayed in Table 1.

## Results and Discussion

This portion evaluates efficiency of proposed technique by varying different images.

### Setup of Experiment

The experimentation of the proposed work has been carried out via MATLAB along with Windows 10 operating system where Intel i3 core processor with a RAM of 2GB was used. Dataset[29] is available publically online and involves 20 sets of colored multi-focus images with an image size of 520 × 520 pixels. And further four series of multi-focus images present in the dataset involving three sources.

### Evaluation Measures

The analysis of proposed ASSCA-based Deep CNN is adapted to analyze techniques with PSNR, RMSE and MI.

**a) *PSNR:*** It may be illustrated as the proportion of maximal possible signal power to power of affected noise. The greater value of PSNR depicts better system and it is measured in decibel (dB).

Table 1 — Pseudo code of ASSCA

***Input:*** *Solution set* $A$*, velocity* $k$

***Output:*** $A_{best}$

***Begin***

Initialize set of atoms in random manner.
while stopping criterion is unsatisfied do
for each atom $A_g$, do
Evaluate fitness with Eq. (7)
If $Fit_g < Fit_{best}$, then
$Fit_{best} = Fit_g$
$A_{best} = A_g$
End if
Update position with Eq. (20)
end for
end while
Discover best solution
Return $A_{best}$
end

$$PSNR = 10\log_{10}\left(\frac{m_{\max}{}^2}{MSE}\right) \quad \ldots (21)$$

where, the term $m_{\max}$ indicates the highest image pixel value, and MSE denotes mean square error. Greater value of PSNR depicts higher the performance of the image fusion method.

***b) RMSE:*** It is utilized to calculate the fusion model error for estimating the quantitative data, and is illustrated as,

$$L = \frac{1}{2}[L[G(y,z),V(y,z)] + L[D(y,z),V(y,z)]] \quad \ldots (22)$$

A minimum value of RMSE determines the better performance.

***c) MI:*** It is utilized to calculate the mutual dependence quantity in between the two images, it is computed using the below expression.

$$\mathrm{M} = \frac{1}{2}[\mathrm{M}[\mathrm{G}(\mathrm{y},\mathrm{z}),\mathrm{V}(\mathrm{y},\mathrm{z})] + \mathrm{M}[\mathrm{D}(\mathrm{y},\mathrm{z}),\mathrm{V}(\mathrm{y},\mathrm{z})]] \quad \ldots (23)$$

where, the two input images are illustrated as $\mathrm{G}(\mathrm{y},\mathrm{z})$, D(y,z) and the fused image is denoted as $\mathrm{V}(\mathrm{y},\mathrm{z})$. The higher value of mutual information shows the better performance.

**Experimental Results**

The experimental outputs obtained by the developed technique are discussed in this section. Simulated results obtained from the proposed ASSCA-based Deep CNN using Lytro Multi-focus Image Dataset has been demonstrated through Fig. 3.

**Comparative Methods**

Comparative methods involves: DWT+DCNN[19], DWT + ASSCA + Renyi entropy[30], Contourlet + Taylor ASSCA[31] and proposed ASSCA-based Deep CNN.

**Comparative Analysis**

The efficiency of proposed method using classical strategies with PSNR, RMSE and MI parameters is examined. Analysis has been carried out by fluctuating the values of alpha $\alpha$ from 0.2 to 0.3. For detailed analysis please refer Table 2 and Table 3. The analysis of methods considering $\alpha = 0.2$ with PSNR, RMSE and MI key performance evaluators have been demonstrated through Fig. 3 involving Fig. 3a), 3b), 3c) respectively. The analysis of methods considering $\alpha = 0.3$ with PSNR, RMSE and MI is described in Fig. 4 involving Fig. 4a), 4b) and 4c) respectively.

**Analysis with α = 0.2**

The analysis of methods considering $\alpha = 0.2$ with PSNR, RMSE and MI key performance evaluators have been demonstrated through Fig. 3

(involves Fig. 3a, Fig. 3b and Fig. 3c) while the analysis of methods considering $\alpha = 0.3$ with PSNR, RMSE and MI key performance indicators is described in Fig. 4 (involves Fig. 4a, Fig. 4b and Fig. 4c) respectively.

The maximal **PSNR** of **32.55dB** is evaluated by proposed ASSCA-based Deep CNN. The minimal **RMSE** of **7.59** and maximum **MI** of **1.52** is computed by proposed method.

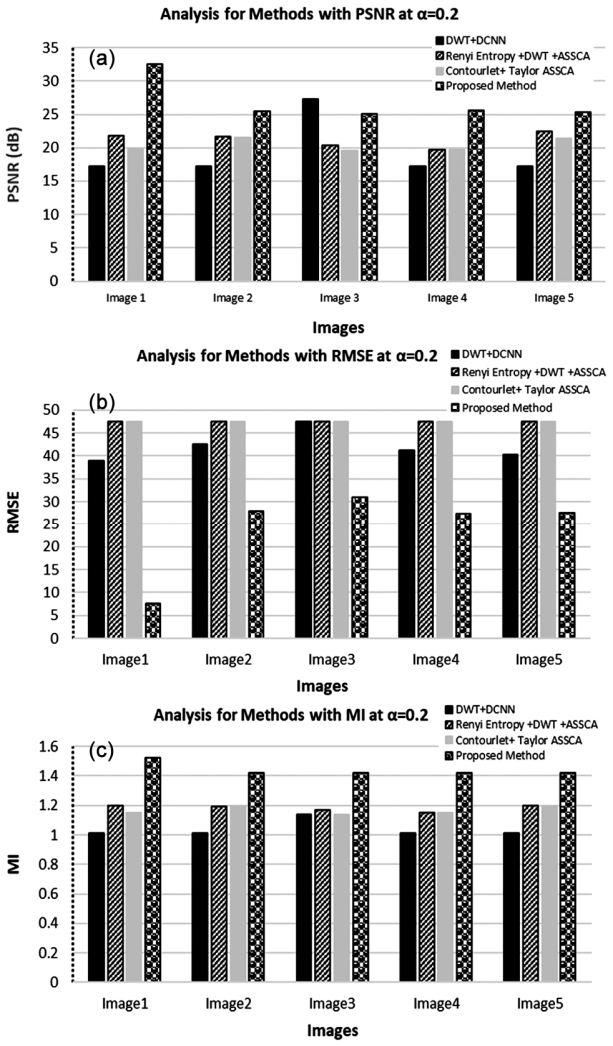


Fig. 3 — Analysis of Methods using $\alpha = 0.2$ Considering a) PSNR b) RMSE c) MI

Table 2 — Quantitative Assessment using evaluation metrics MI, PSNR, RMSE when α = 0.2 for different set of images

| Metrics | Images | Ref[19] | Ref[30] | Ref[31] | Proposed Method |
|---|---|---|---|---|---|
| MI | 1 | 1.01 | 1.20 | 1.15 | **1.52** |
| MI | 2 | 1.01 | 1.19 | 1.19 | 1.42 |
| MI | 3 | 1.14 | 1.17 | 1.14 | 1.42 |
| MI | 4 | 1.01 | 1.15 | 1.15 | 1.42 |
| MI | 5 | 1.01 | 1.20 | 1.19 | 1.42 |
| PSNR(dB) | 1 | 17.27 | 21.74 | 19.88 | **32.55** |
| PSNR(dB) | 2 | 17.27 | 21.70 | 21.50 | 25.49 |
| PSNR (dB) | 3 | 27.36 | 20.39 | 19.52 | 25.13 |
| PSNR(dB) | 4 | 17.27 | 19.74 | 19.66 | 25.57 |
| PSNR(dB) | 5 | 17.27 | 22.48 | 21.38 | 25.37 |
| RMSE | 1 | 39 | 47.59 | 47.59 | **7.59** |
| RMSE | 2 | 42.5 | 47.59 | 47.59 | 27.98 |
| RMSE | 3 | 47.59 | 47.59 | 47.59 | 31.01 |
| RMSE | 4 | 41.26 | 47.59 | 47.59 | 27.34 |
| RMSE | 5 | 40.26 | 47.59 | 47.59 | 27.65 |

Table 3 — Quantitative Assessment using evaluation metric MI, PSNR, RMSE when α=0.3 for different set of images

| Metrics | Images | Ref[19] | Ref[30] | Ref[31] | Proposed Method |
|---|---|---|---|---|---|
| MI | 1 | 1.01 | 1.20 | 1.20 | **1.51** |
| MI | 2 | 1.01 | 1.20 | 1.19 | 1.42 |
| MI | 3 | 1.01 | 1.15 | 1.15 | 1.42 |
| MI | 4 | 1.01 | 1.19 | 1.19 | 1.42 |
| MI | 5 | 1.01 | 1.20 | 1.17 | 1.42 |
| PSNR(dB) | 1 | 17.27 | 22.16 | 21.64 | **29.58** |
| PSNR(dB) | 2 | 17.27 | 21.97 | 20.99 | 25.93 |
| PSNR (dB) | 3 | 17.27 | 19.64 | 19.70 | 25.24 |
| PSNR (dB) | 4 | 17.27 | 21.32 | 20.99 | 25.60 |
| PSNR(dB) | 5 | 17.27 | 21.93 | 20.53 | 25.38 |
| RMSE | 1 | 42.15 | 47.59 | 47.59 | **10.73** |
| RMSE | 2 | 40.25 | 47.59 | 47.59 | 24.75 |
| RMSE | 3 | 38.56 | 47.59 | 47.59 | 30.07 |
| RMSE | 4 | 42.5 | 47.59 | 47.59 | 27.14 |
| RMSE | 5 | 47.58 | 47.59 | 47.59 | 28.89 |

**Analysis with α = 0.3**

**Comparative Discussion**

It is evident from Table 2 and Table 3, that the performance metrics are depicting best results for proposed method as compared to other listed methods. Analysis has been performed from images 1 to 5 by varying images and the obtained results are listed. Here value of α has been changed from 0.2 to 0.3 and outcomes are analyzed for different metrics such as PSNR, RMSE, MI for the listed methods. Bold font values denote better performance results. The proposed ASSCA-based Deep CNN offers improved performance in contrast to current state of the art techniques with a highest value **1.52** of *mutual information*, with a highest value of **32.55dB** of maximum *Peak Signal to Noise Ratio* and value of **7.59** of *Minimum Root Mean Square Error.*

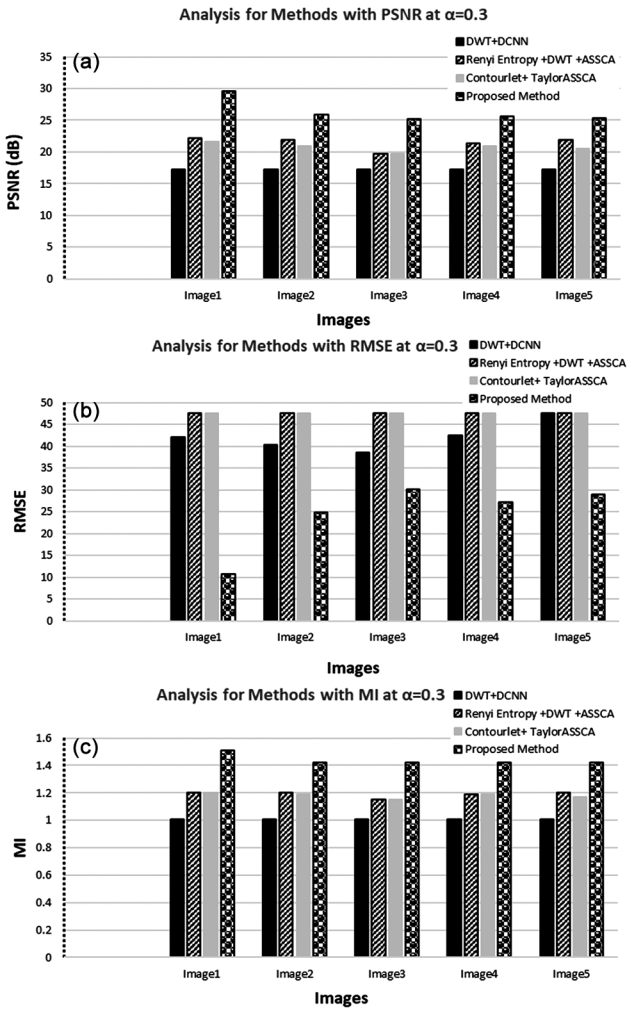


Fig. 4 — Analysis of Methods using $\alpha = 0.3$ Considering a) PSNR b) RMSE c) MI

## Conclusion and Future Scope

A novel image fusion method, namely ASSCA-based Deep CNN has been proposed through this research paper. The goal of determining the focused region and then combining focused region has been achieved. Here, the two images are fed to initial fusion phase using weighted average model. Further, the fusion score are computed and values were determined using an optimization model. Thereafter, final fusion is performed using proposed method. The Deep CNN training is carried out with ASSCA. The proposed method provides improved performance in

contrast to other methods with maximum *MI* of **1.52**, maximum *PSNR* of **32.55** dB, and minimum *RMSE* of **7.59**. In future, the method can be expanded to different multi-modal tasks that involve visible and infrared images.

## References


1 Guo X, Nie R, Cao J, Zhou D, Mei L & He K, Fusegan: Learning to fuse multi-focus image via conditional generative adversarial network, *IEEE Trans Multimed*, **21(8)** (2019) 1982−1996.

2 Kou F, Wei Z, Chen W, Wu X, Wen C & Li Z, Intelligent detail enhancement for exposure fusion, *IEEE Trans Multimed*, **20(2)** (2017) 484−495.

3 Xiao B, Ou G, Tang H, Bi X & Li W, Multi-focus image fusion by hessian matrix based decomposition. *IEEE Trans Multimed*, **22(2)** (2019) 285−297.

4 Zhang J, Liao Q, Liu S, Ma H, Yang W & Xue J H, Real-MFF: A large realistic multi-focus image dataset with ground truth, *Pattern Recognit Lett*, **138** (2020) 370−377.

5 Amin-Naji M & Aghagolzadeh A, Multi-focus image fusion in DCT domain using variance and energy of Laplacian and correlation coefficient for visual sensor networks, *J Artif Intell Data Min*, **6(2)** (2018) 233−250.

6 Xu H, Fan F, Zhang H, Le Z & Huang J, A Deep Model for Multi-Focus Image Fusion Based on Gradients and Connected Regions, *IEEE Access*, **8** (2020) 26316−26327.

7 Dou W, Image degradation for quality assessment of pan-sharpening methods, *Remote Sens*, **10(1)** (2018) 154.

8 Cao T, Dinh A, Wahid K A, Panjvani K & Vail S, Multi-focus fusion technique on low-cost camera images for canola phenotyping, *Sensors*, **18(6)** (2018) 1887.

9 Li Q, Yang X, Wu W, Liu K & Jeon G, Multi-focus image fusion method for vision sensor systems via dictionary learning with guided filter, *Sensors*, **18(7)** (2018) 2143.

10 Chen Q, Yang B, Li Y & Pang L, Multi-focus Image Fusion with Point Detection Filter and Superpixel Based Consistency Verification, *IEEE Access,* 2020.

11 Toet A, Image fusion by a ratio of low-pass pyramid, *Pattern Recognit Lett*, **9(4)** (1989) 245−253.

12 Hu X, Lu H, Zhang L, & Serikawa S, A new type of multi-focus image fusion method based on curvelet transforms, in Proceedings of *Int Conf Electr Contr Eng*, (2010) 172–175.

13 Li H, Chai Y & Li Z, Multi-focus image fusion based on nonsubsampled contourlet transform and focused regions detection, *Optik*, **124(1)** (2013) 40−51.

14 Wang W & Chang F, A Multi-focus Image Fusion Method Based on Laplacian Pyramid, *J Comput*, **6(12)** (2011) 2559−2566.

15 Petrovic V S & Xydeas C S, Gradient-based multi resolution image fusion, *IEEE Trans Image Process*, **13(2)** (2004) 228−237.

16 Zhang Q, Liu Y, Blum R S, Han J & Tao D, Sparse representation based multi-sensor image fusion for multi-focus and multi-modality images: A review, *Inf Fusion*, **40** (2018) 57−75.

17 Yang Y, Huang S, Gao J & Qian Z, Multi-focus image fusion using an effective discrete wavelet transform based algorithm, *Meas Sci Rev*, **14(2)** (2014) 102.

18 Duan Z, Zhang T, Tan J & Luo X, Non-Local Multi-Focus Image Fusion With Recurrent Neural Networks, *IEEE Access*, **8** (2020) 135284−135295.

19 Zhao W, Wang D & Lu H, Multi-focus image fusion with a natural enhancement via a joint multi-level deeply supervised convolutional neural network, *IEEE Trans Circuits Syst Video Technol*, **29(4)** (2018) 1102−1115

20 Zhang H, Le Z, Shao Z, Xu H & Ma J, MFF-GAN: An unsupervised generative adversarial network with adaptive and gradient joint constraints for multi-focus image fusion, *Inf Fusion*, **66** (2020) 40−53.

21 Panigrahy C, Seal A & Mahato N K, Fractal dimension based parameter adaptive dual channel PCNN for multi-focus image fusion, *Opt Lasers Eng*, **133** (2020) 106141.

22 Guo R, Shen X J, Dong X Y & Zhang X L, Multi-focus image fusion based on fully convolutional networks, *Front Inf Technol Electron Eng*, **21(7)** (2020) 1019−1033.

23 Ma B, Zhu Y, Yin X, Ban X, Huang H & Mukeshimana M, Sesf-fuse: An unsupervised deep model for multi-focus image fusion*, Neural Comput Appl*, (2020) 1−12.

24 Han X, Yu J, Xue JH & Sun W, Hyperspectral & multispectral image fusion using optimized twin dictionaries, *IEEE Trans Image Process*, **29** (2020) 4709−4720.

25 Babu GS, Zhao P & Li XL, Deep convolutional neural network based regression approach for estimation of remaining useful life, In *Int conf on DASFAA, Springer*, *Cham* (2016) 214−228.

26 Rajesh Kumar T, Suresh G R, Kanaga Subaraja S & Karthikeyan C, Taylor-AMS features and deep convolutional neural network for converting nonaudible murmur to normal speech, *Comput Intell*, **36(3)** (2020) 940−963.

27 Mirjalili S, SCA: a sine cosine algorithm for solving optimization problems, *Knowl Based Syst*, **96** (2016) 120−133.

28 Zhao W, Wang L & Zhang Z, Atom search optimization and its application to solve a hydrogeologic parameter estimation problem, *Knowl Based Syst*, **163** (2019) 283−304.

29 Lytro Multi-focus Image Dataset taken from," https://www.researchgate.net/publication/291522937_Lytro_Multi-focus_Image_Dataset", accessed on October 2019.

30 Singh V & Kaushik VD, Renyi entropy and atom search sine cosine algorithm for multi focus image fusion, *Signal Image Video Process*, **15** (2021) 903−912.

31 Singh V & Kaushik VD, HoEnTOA: Holoentropy and Taylor assisted optimization based a novel image quality enhancement algorithm for multi-focus image fusion, *J Sci Ind Res*, 2021 (In Press).