A Novel of Pansharpening using the DCSTFN

V. Sruthi
Master of Technology
Department of Electronics and Communication Engineering
Sri Venkateswara University College of Engineering,
Tirupati, India

Dr. G. Umamaheswara Reddy
Professor,
Department of Electronics and Communication Engineering
Sri Venkateswara University College of Engineering,
Tirupati, India

Abstract— **The earth observing satellites provides two kinds of images: a low spectral and high spatial PAN image and various high spectral/low spatial resolution Multispectral (MS) images. We encounter some trade-offs in designing the remote sensing instruments due to some technical and budget constraints which makes it difficult to achieve high spatio-temporal resolution images. To overcome this problem, high spatio-temporal resolution images are derived with low temporal but high spatial resolution (LTHS) and high temporal but low spatial resolution (HTLS) from remote sensing images, is proposed in this paper. The DCSTFN model consists of three main parts: HTLS images expansion, high frequency components derived from LTHS images are extracted, and the extracted features fusion.**

Keywords— *DCSTFN, HTLS, LTHS, Panchromatic.*

I. INTRODUCTION

For the trees growth, the urban environment presents a various challenges. Higher level of pollutions are incurred due to the hydrological systems, soil compositions etc. the extensive impervious surfaces was characterized through the urban environments and prevents water filtration in some artificial models. The impervious surfaces are related with urban microclimates ,hydrological rerouted networks and diminished soil conditions, the major impact of these issues on tree development has high variability.

Based on this study, shows the relation at two spatial scales: inside the individual tree's local environment and along the broad-scale urban landscape. Soil aeration can be reduced by paved surfaces and we can modify the underlying moisture in the vertical distribution. Provision of sensible heat by transforming and absorbing the solar energy, microclimates are created due to surface imperviousness that contributes to the effects of urban heat island. The hyper spectral and multi spectral sensors are developed through the satellite imaging technology. The airborne and satellite sensors chemicals and material bonds are detected through this technology. Without being physical contact, it is an art of making scientific measurements in remote sensors. It is also used to collect the data from dangerous diseases.

Many earth-observing satellites, such as , IKONOS, Gaofen-1, Quick-Bird, Land sat can simultaneously shoot a panchromatic image and a multispectral image within the same coverage areas. Since the reflectance value varies by land covers ,and spectral bands, the recorded information of earth surfaces in multispectral(MS) images is more compared to the Panchromatic(PAN) images. Although, for the signal-to noise ratio(SNR) and the resolutions of sensors given, PAN images will have high spatial resolution than the multispectral images.

Hence, Image fusion is designed to get the most spatial and spectral data by fusing(transforming) Multispectral images and Panchromatic images. The spectral data and the PAN images resolution both are present in the fused images. Many remote sensing image fusion techniques have been implemented and can be graded into three types: component substitution, multi-resolution analysis, and sparse representation. . Redundant and complementary information integration can be dealt by image fusion methods for creating composite image contains a better understanding of the scene. The resulting fused image is a product generated by multiple sensors or provided by the same sensor, that can be useful for human visual perception, provides faster interpretation which can be helpful in extracting more features. The uncertainty related with the data that are taken from different sensors or by the same sensor, can be reduced using data fusion methods with temporal variation. Fusion of two data sets is done to get one single data set Having qualities of both data sets. The fused data product has the advantages of structural information, high spatial resolution, and spectral resolution.

II. EXISTING METHOD

In the existing method, fusion method on the remote sensing images in detail is demonstrated. The network architecture, which uses the acronym RSIFNN in the following, meaning CNN-based image fusion of remote sensing images. The block diagram for RSIFNN method is given below:
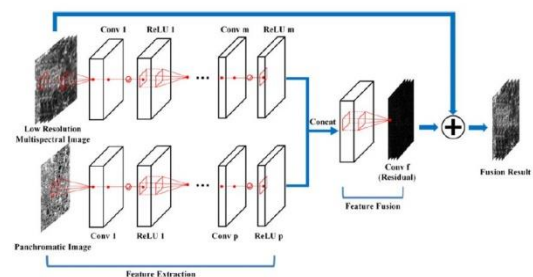

Fig. 1. Architecture of RSIFNN for remote sensing image fusion.

1) Procedure:

The MS image and PAN image are taken, then the features are obtained from the each image using the CNN model. Then these features from two images are concatenated and final features are obtained. From these features the fine features are obtained. Finally, these features are combined with the input Multispectral image and Final transformed (fused) image is obtained.

The layers of CNN model is given below:

- **CONV layer:** This layer will calculate the yield of neurons linked to local input areas, each calculating a dot product of tiny region linked to the quantity of input and their weights. This can lead in quantity like [115x115x1] for PAN image and [115x115x4] for MS image if we chose to use 4 filters for the first layer, 64filters for the next two layers.

  - **RELU layer:** A smart activation feature like max (0, x) max (0, x) limit at zero will be applied. This leaves the volume size unchanged [[115x115xX]). Where, X is 1 for PAN image and 4 for MS image.

  - **Fully Connected layer:** It provides a bias vector by multiplying the matrix weight of the input. As the name indicates it connects with the previous layer's neurons. This layer incorporates all the characteristics (local data) learned throughout the picture by the past layers to recognize the bigger patterns. The last completely linked layer combines the characteristics to classify the pictures for classification issues to the data set's no. of classes. The output size equal to the amount of response variable for regression issues.

2) Network Training:

Here by considering the CNN layers we take the information set. In this part, we will show the training procedure which aims to find optimal parameters to express the whole network sufficiently.

$$L = \frac{1}{n} \sum_{i=1}^{n} \left\| y^{(i)} - f(x_1^{(i)}, x_2^{(i)}) \right\|^2$$

In this existing method, the residual learning layer can solve this problem. Because there are many similarities between low- and high-resolution MS images, we can create a residual imager = y − x₁to describe their differences. Pixel values in the residual image will mostly be zero or small, and thus the whole residual image is sparse. In this way, we can ignore redundant information and just focus on the feature which mainly improves the spatial resolution for MS images. By adding the residual image to the low-resolution MS image, the high-resolution MS image is generated. The loss function now is modified as

$$L = \frac{1}{n} \sum_{i=1}^{n} \left\| r^{(i)} - g(x_1^{(i)}, x_2^{(i)}) \right\|^2$$

To realize residual learning, we can do a small modification to loss layer of the network. The loss layer consists of following parts: Input from the branch of MS image, residual image, and label (high-resolution MS image), and the predicted fusion result is generated by adding first two parts together. The mini-batch gradient descent algorithm with back-propagation is used to optimize the loss function. Besides, the weights are updated by the following equations:

$$\Delta_{i+1} = \beta \cdot \Delta_i - \omega \cdot \alpha \cdot \omega_i^l - \alpha \cdot \frac{\partial L}{\partial \omega_i^l}$$

And,

$$\omega_{i+1}^l = \omega_i^l + \Delta_{i+1}$$

III. PROPOSED METHOD

The block diagram of DCSTFN method is given below. The DCSTFN method consists of the deep convolutional spatio-temporal network which uses deconvolution filter and maxpooling filters in order to achieve high metric values.
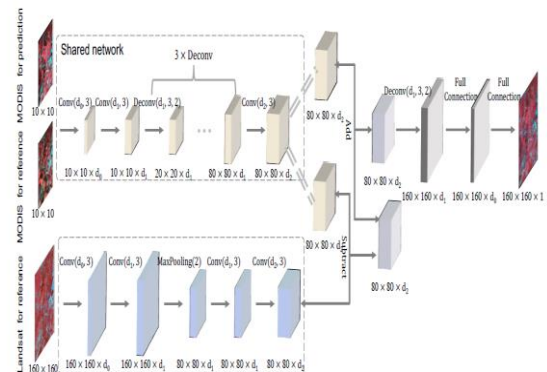

Fig.2. The architecture of the deep convolutional spatiotemporal fusion network (DCSTFN) model.

A deep convolutional fusion network was designed to derive high spatiotemporal resolution remote sensing images. The whole architecture of the DCSTFN can be divided into three parts: the expansion of the HTLS images, the extraction of high frequency components from LTHS images, and the fusion of extracted features.

Procedure of the proposed method is same as the existing method. But the network used for the extracting the features is different. The layers of the network used in this method also consists of the deconvolution layer, max pooling layers. This method increases the SNR of the final transformed(fused) image with high quality.

1) Deep Convolutional Spatio-temporal network:

In this network the extraction of features process of the MS image consists of the 2 layers of convolution filters, 3 layers of the deconvolution layers, again 2 layers of the convolution filters with full connected layer.

Whereas the extraction of features process consists of the PAN image has 2 convolutional layers, maxpooling layers, 2 convolutional layers with fully connected layers.

1) Deconvolution layer:
The Deconvolution in the framework of convolutional neural networks is more frequently used to indicate a sort of reverse convolution, which majorly and distractingly are not a proper mathematical deconvolution. In variation to unpooling, the up sampling of an image can be learned with the use of 'deconvolution'. The deconvolutional layer is quite often useful for up sampling the output image of a convent to the original image resolution.

2) POOL layer:
In pooling layer the input is divided into many rectangular regions using down sampling technique. The pooling layer is of two types they are average pooling layer and maximum pooling layer. These layers representation can be given as maxPooling2dLayer and averagePooling2dLayer. By using down sampling process the connections to neighbor layers will be reduced. These layers are not used for learning operation but reduce the features that are processed for other layers. By this process the over fitting can also be reduced.

3) Image Fusion Quality Metrics:
The fusion image quality is calculated by using the following quality metrics when in the absence of reference image such as Signal-to-Noise Ratio (SNR), Spectral Angular Mapper (SAM), Relative Dimensionless Global Error (ERGAS), Universal Image Quality Index (UIQI), Root Mean Square Value (RMSE), Correlation Coefficient (CC).

i. **Signal to Noise Ratio (SNR):** It is the ratio calculated between the information and noise of the fused image. Higher the SNR , higher the similarity between the reference and the fused images.

$$SNR = 10 \log_{10} \left[\frac{\sum_{i=1}^{M} \sum_{j=1}^{N} (I_r(i,j))^2}{\frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (I_r(i,j) - \Gamma_f(i,j))^2} \right]$$

ii. **Root Mean Square Value (RMSE):** The difference between the reference and fused images can be known by calculating the variation in the pixel values known as RMSE and is given by,

$$RMSE = \sqrt{\frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \left(I_r(i,j) - \Gamma_f(i,j) \right)^2}$$

iii. **Spectral Angular Mapper (SAM):** It calculates the angle in spectral space between pixels and a set of reference spectra(end members) for image classification based on spectral similarity.

$$SAM(V, \hat{V}) \triangleq arccos \left[\frac{\langle V, \hat{V} \rangle}{\|V\|_2 \|\hat{V}\|_2} \right]$$

Zero SAM denotes the absence of spectral distortion.

iv. **Universal Image Quality Index (UIQI):** It indicates the amount of relevant data transformation from reference into the fused image.

$$UIQI = \frac{4\sigma_{I_r I_f} \left(\mu_{I_r} + \mu_{I_f} \right)}{\left(\sigma_{I_r}^2 + \sigma_{I_f}^2 \right) \left(\mu_{I_f}^2 + \mu_{I_r}^2 \right)}$$

v. **Correlation Coefficient (CC):** It tells the similarity of the spectral features between the reference and the fused images.

$$CC = \frac{2C_{rf}}{C_r + C_f}$$

IV. RESULTS

The results attained from the two methods i.e. RSIFNN and DCSTFN with two different datasets are shown below:
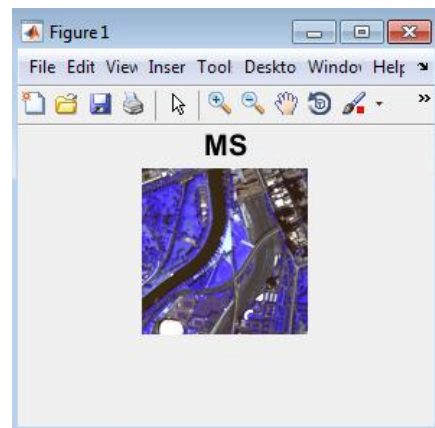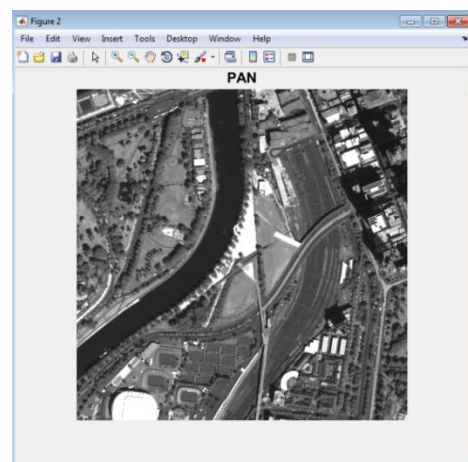

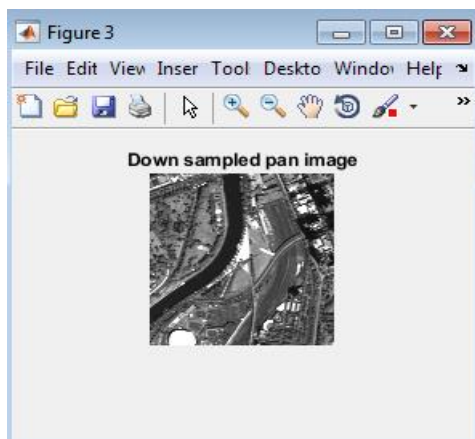Fig. 3. Multispectral In (6)

Fig. 4. Panchromatic Image-1

Fig. 5. Down sampled PAN Image-1

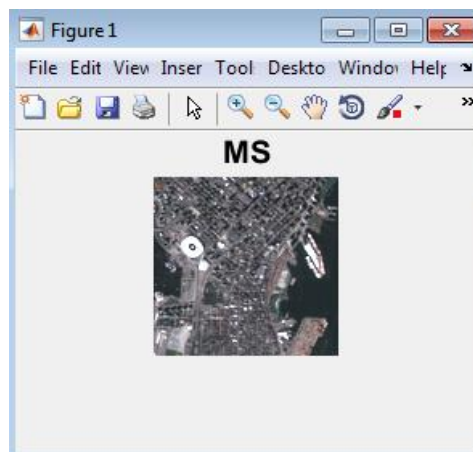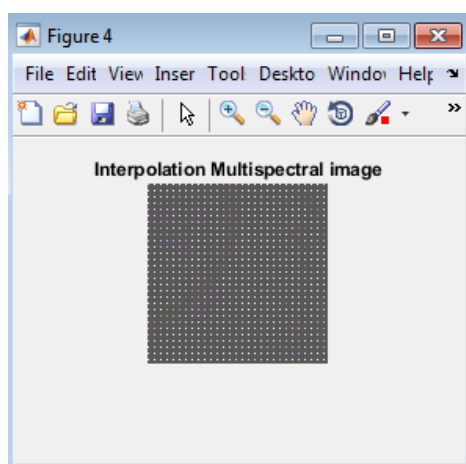
Fig. 8.Multispectral Image -2
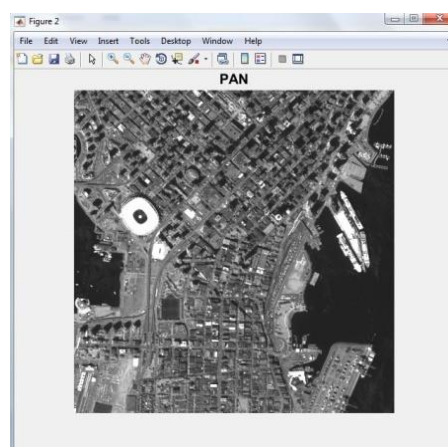

Fig. 6. Interpolation MS Image-1
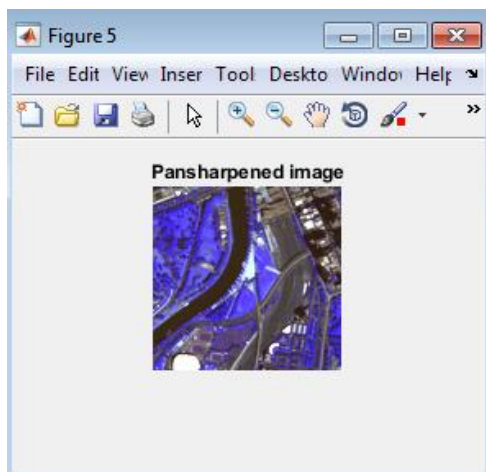

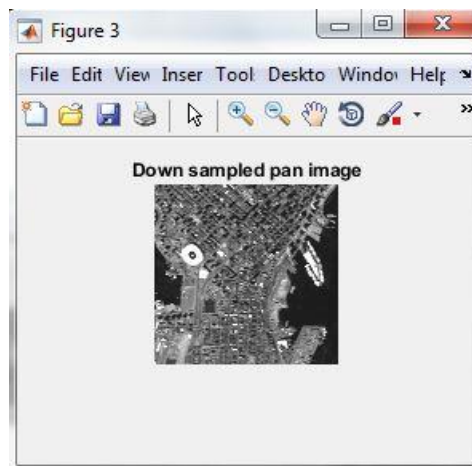Fig. 9.PAN Image-2

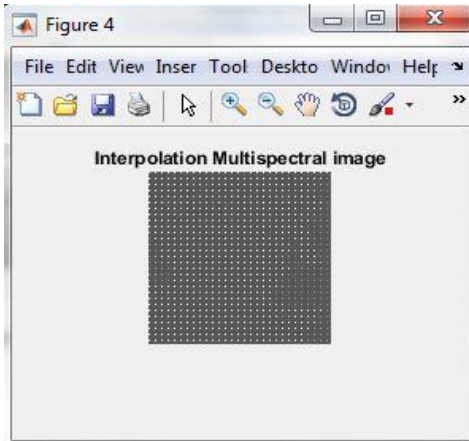Fig. 7. Pansharpened Image-1

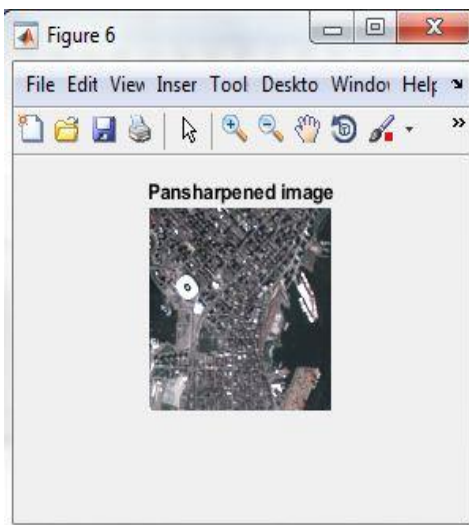Fig. 10. Down Sampled PAN Image -2

Fig:11. Interpolation of MS Image-2

Fig.12. PAN sharpened Image-2

TABLE1. Comparison of existing and proposed methods

Parameter	Existing method		Proposed method	
	Data set-1	Data set-2	Data set-1	Data set-2
SNR	27.9870	30.9512	50.3856	53.712
RMSE	0.0749	0.0362	0.0025	0.0018
UIQI	0.7986	0.7854	0.9829	0.9999
SAM	0.0248	0.0031	1.0450	0.1499
ERGAS	0.5021	0.2294	0.0249	0.0165
CC	0.0005	0.0002	0.0230	0.0135

IV. CONCLUSION

This paper presents remote sensing spatio-temporal data fusion realm analyzed with the deep learning approaches and produces a significant results. The advantages of the CNN-based fusion techniques has two equally important parts: (1) a series of high spatiotemporal multispectral images with high accuracy is generated and compared to the conventional approaches, the network is more robust and is less-sensitive to the input data quality; and (2) the DCSTFN model requires very less time for executing data in large amounts for a long-time series analysis. Once the network is established, it can be implemented for the entire dataset. In contrast, conventional methods are more suitable for tasks where input data are of relatively good quality and the data volumes to be processed are not too large.

REFERENCES

[1] A. A. Plowright, N. C. Coops, C. M. Chance, S. R. J. Sheppard, and N.W. Aven, "Multi-scale analysis of relationship between imperviousness and urban tree height using airborne remote sensing," Remote Sens. Environ., vol. 194, pp. 391–400, Jun. 2017.

[2] P. Zhang, M. Gong, L. Su, J. Liu, and Z. Li, "Change detection based on deep feature rep6resentation and mapping transformation for multi-spatialresolution remote sensing images," ISP6RS J. Photogramm. Remote Sens., vol. 116, pp. 24–41, Jun. 2016.

[3] L. Matikainen, K. Karila, J. Hyypp¨a, P. Litkey, E. Puttonen, and E. Ahokas, "Object-based analysis of multispectral airborne laser scanner data for land cover classification and map updating," ISPRS J. Photogramm. RemoteSens., vol. 128, pp. 298–313, Jun. 2017.

[4] T.-M. Tu, S.-C. Su, H.-C. Shyu, and P. S. Huang, "A new look at HIS-like image fusion methods," Inf. Fusion, vol. 2, no. 3, pp. 177–186, Sep. 2001.

[5] S. K. Pal, T. J. Majumdar, and A. K. Bhattacharya, "ERS-2 SAR and IRS-1C LISS III data fusion: A PCA approach to improve remote sensing based geological interpretation," ISPRS J. Photogramm. Remote Sens., vol. 61, no. 5, pp. 281–297, Jan. 2007.

[6] T.-M. Tu, P. S. Huang, C.-L. Hung, and C.-P. Chang, "A fast intensity–hue–saturation fusion technique with spectral adjustment for IKONOS imagery, " IEEE Geosci. Remote Sens. Lett., vol. 1, no. 4, pp. 309–312, Oct. 2004.

[7] S. Rahmani, M. Strait, D. Merkurjev, M. Moeller, and T. Wittman, "An adaptive IHS pan sharpening method," IEEE Geosci. Remote Sens. Lett., vol. 7, no. 4, pp. 746–750, May 2010.