International Conference on Inventive Systems and Control (ICISC 2019)
IEEE Xplore Part Number: CFP19J06-ART; ISBN: 978-1-5386-3950-4

PREDICTING MISSING IMAGE IN REMOTE SENSING TIME SERIES USING SPATIAL-TEMPORAL-SPECTRAL DATA

Deepa Palanisamy
Department of Information Technology
MIT Campus, Anna University
Chennai, India
deepapalanisamy25@gmail.com

Radha Senthilkumar
Department of Information Technology)
MIT Campus, Anna University)
Chennai, India
radhasenthilkumar@gmail.com

***Abstract*—Remote sensing is the acquisition of physical characteristics (reflecting radiation) of the remote object. It can be collected via special cameras or sensors in the satellite or aircraft or weather balloons. Each remote sensing images has multiple spectral bands. The remote sensing images analysis is used by multiple applications like metrological prediction, Land Cover and Land Usage prediction (LCLU), vegetation change detection. Missing image in the remote sensing time series produces a lot of glitches, causing serious upshot in the multi-temporal analysis, when the images at various time stamps are missing over a period of time.The existing work reconstructs missing image in remote sensing time series via spatial and temporal data. The proposed method Tensor-Deep Stacking Network Spatial-Temporal-Spectral (TDSN-STS) helps to reconstructs the missing image in remote sensing time series using spatial, temporal and spectral data. Thus the accuracy of the reconstructed image in TDSN-STS was increased substantially compared to the existing work.**

***Index Terms*—Remote Sensing, Time series, Tensor Deep Stacking Network, spatial, temporal, spectral.**

I. INTRODUCTION

Remote sensing is the process of obtaining information about the objects or areas from a distance, typically from aircraft or satellites. Cameras on satellites and airplanes can cover the wide area of Earths surface, allowing us to see much more than what we can see standing on the ground. It is possible to collect data from the inaccessible areas or dangerous area. Remote sensing is mostly used in Earth Science disciplines, the military and by various application such as metrological prediction, monitoring and movement of polar icebergs, spread of oil in the ocean etc [1].

Landsat 7 is one of the satellites used for remote sensing. It is launched on April 15, 1999. Its primary goal is to provide up to date cloud-free image of the Earth. Landsat program is maintained by the United States Geological Survey (USGS) [15] and is capable to collect and transmit 532 images of Earth per day.It uses Enhanced Thematic Mapper Plus (ETM+) sensor to capture the image. These image contains eight spectral bands of 30-meter spatial resolutions.

Spatial data or geospatial data is the data that are directly or indirectly referenced by specific location in the Earth. There is a relation between the images collected in the series, as all the images belong to the same co-ordinates [5]. Temporal data is the image of Earth acquired at different period of time. It will help for multi-temporal analysis like change detection, metrological prediction, detection of vegetation changes in an alpine protected Area [5].Spectral information is every satellite image contains multiple bands because, different object reflects and observes different wave-length. In the existing work by Monidipa Das et al. [11] uses only two bands (not include all the spectral bands) to predict missing image from the remote sensing time series data. Predicting missing image by fusing all the spectral bands gives more information hence, the accuracy of the predicted image was increased.

In the existing work converts the raw satellite image into Normalized Difference Vegetation Index (NDVI). NDVI is the standard way to measure the healthy vegetation. NDVI is calculated by using Near InfraRed (NIR) band and red band. The value of NDVI is between -1 to +1 [13][14]. Satellites like Landsat and Sentinel provide necessary bands (NIR and RED) to calculate NDVI [15].

A. Motivation

One of the common barriers frequently appearing in remote sensing time-series analysis are the non-availability of data in the temporal sequence. The origin of such missing images are produced by poor atmospheric condition, or a fault in the sensor used to capture the image of the Earth. Unable to acquire the image until the said issue is fixed. These missing image in the multi-temporal analysis may degrade the performance of the analytical process like LCLU prediction, urban sprawl detection etc. [1].

In Fig. 1. shows a sequence of Landsat -7 images, where the image at the time instant [t+3] is missing in the temporal sequence. This missing image reduce the accuracy of the analytical process. Therefore, it is necessary to reconstruct the missing image inorder to improve the accuracy of the analytical process.

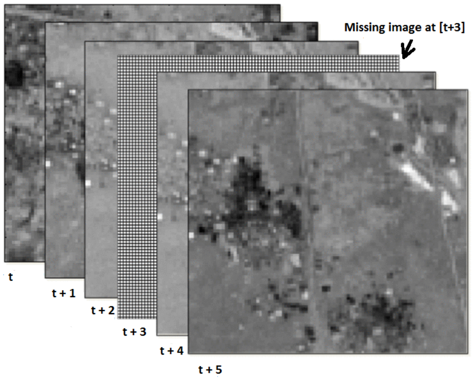

Fig. 1. Missing Image in the Sequence of Remote Sensing Time Series.

II. RELATED WORKS

In Deep learning based predicting missing data for remote sensing analysis [11] use spatial and temporal data to predict the missing image in the time series. In this work, raw satellite image is convert NDVI image.

In predicting remote sensing data using spatial-temporal technique [12]. In recent years remote sensing is used in multiple application and it is simple and easy to acquire high-resolution satellite image. By the same time, it creates a challenge for analysis community to interpretation from the images of the huge volume of data. Hence use deep stacking network to predict remote sensing data at the timestamp t. Deep stacking network is stacking of multilayer perceptron with one hidden layer.

In parallel training of deep stacking networks[7] and scalable training of deep stacking network [10] provide parallel and scalable training of deep stacking network. Parallelism reduces the memory and time requirement but arising cost for inter CPU communication. It also demonstrates the full-batch training significantly reduce the error rate by distributing the training dataset over the CPU cluster compare to the mini-batch training using single CPU cluster.Parallelism is achieved by parallelizing the matrix multiplies. The job requires to synchronize all parallel batches distributed over CPU cluster. It must wait for all the dependencies.

In Tensor Deep Stacking Networks (TDSN) [2] the complex functions are calculated from the stacking of multiple simple functions. Each module consists of linear input layer, and output layer, and a two non-linear hidden layer. Non-linear activation function sigmoid is used to map linear input to non-linear output to improve the performance. The input to the first module is corresponding to the first available image in the temporal sequence . For further modules, the output of the previous module along with the new input image is given as input. The final module of the TDSN predicts the actual value.

III. PROBLEM STATEMENT

In the sequence of remote sensing images $I_1, ..., I_{m1}, ?, I_{m+1}, ..., I_{x1}, ?, I_{x+1}, ..., I_t$, over t timestamps. Where the symbol ? indicates missing image in some period of timestamps. The images I_m and I_x is missing in the sequence of images over t timestamps. The existing method use only temporal and spatial data of the images to reconstructing missing image in the time series [11]. The proposed work use spatial-temporal-spectral data of the images to reconstruct the missing image and proved that the accuracy of predicting image was increased substantially.

IV. PREDICTING MISSING IMAGE USING SPATIAL-TEMPORAL-SPECTRAL DATA

A. Framework

In remote sensing, the missing image in time series data can be predicted using TDSN. In the proposed approach the TDSN-STS does not require GPU even for a large dataset and utilize supervised information at each module of TDSN and the output of each module is given as input to the next immediate module along with the new raw input except the final module of the prediction [9]. The final module gives the actual predicted output. The Fig. 2. depicts the overall architecture of the proposed work and the number of the module is equal to the number of training image available in the remote sensing time series. Each forecasting modules work with early available images in the time series of remote sensing and predict the next immediate missing image as a dummy image. This dummy image is also provided to predict the next missing image in the time series images. The final module predicts the image at the latest available timestamps. This predicted image and actual image is used to check the accuracy of the prediction. If the accuracy is below the threshold level the dummy image is again tuned. This process is continued until there is no improvement in the final predicting image

For spatial and temporal technique use only two bands in the Landsat -7 ETM+ satellite image. Spatial temporal, and spectral technique use all the eight available bands in the satellite image by fusion of all the bands. Fusion of all bands use each and every information of the location.hence, the accuracy of the predicted image by spatial, temporal, spectral obtained is higher compared to the accuracy of the image predicted by using the spatial and temporal technique.

The entire architecture composed of three modules i) feature set preparation ii) Tensor-Deep Stacking Network model training and prediction and iii) Dummy image tuning.

B. Feature Set Preparation

From the satellite image the spatial-temporal technique derive the NDVI image, it is enough to fuse two bands of that particular timestamp [6]. The sequence of the image collected for the particular location of same season gives the spatial and temporal information. In the spatial-temporal and spectral technique all the bands in the satellite image in the same spatial resolution were fused. If the image of different spatial resolution is fused it may collapse the values of the fused image.

NDVI image and spectral image are obtained from the raw satellite imagery and it is converted and stored into the dense

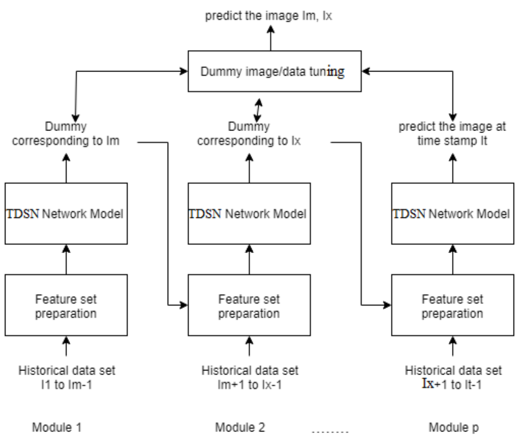

Fig. 2. TDSN-STS to Reconstruct Missing Image

file format in .binary extension. From the dense format the weight matrix of TDSN is calculated during the training phase. The dense file is in the format of N x M matrix and the values are stored in a column-major order. TDSN is a supervised learning model, we have provide both input and target features in the spare file format during training. The sparse matrix is in the form of N x M matrix of target labels. The number of rows in the dense and sparse file must be the same [4].

C. Tensor-Deep Stacking Netwrok Model for Training and Prediction

The architecture of TDSN for three module is depicted in the Fig. 3. Each module consists of a linear input layer, linear output layer, and two non-linear hidden layers. The weight matrices W are the maps between the input layer and hidden layers. The weight matrix U map between two hidden layers and the output layer. The central idea of TDSN is the concept of stacking simple modules to learn the complex function [8][9]. Each module is basically stacking of simple multi-layer perceptron with two hidden layers. The number of modules in the TDSN is equal to the number of training images. The input to the bottommost layer of TDSN is raw input image in the first available sequence. The input to the subsequent models are the output of the immediate previous module and the corresponding raw input image in the time series. The output of each module learns the spatial and spectral feature for the particular timestamp. The output of the final module gives the actual predicted image at the timestamp I_t. The images are converted into pixel and stored in a vector format.

Input vectors are in the form of I= [I_1,I_2,I_N], in which each vector is denoted by I_i=[I_{1i},.,I_{ji}, .I_{Di}], where D is the dimension of the input vector, which is sequence of N total number of training samples in the sequence of time seies.

While training provide training image as dense file and corresponding targets features as sparse file format, number of blocks, if we want to train the modules as parallel give parallel flag as a command line argument. Once the training gets over it generate weight matrices W and U for all the

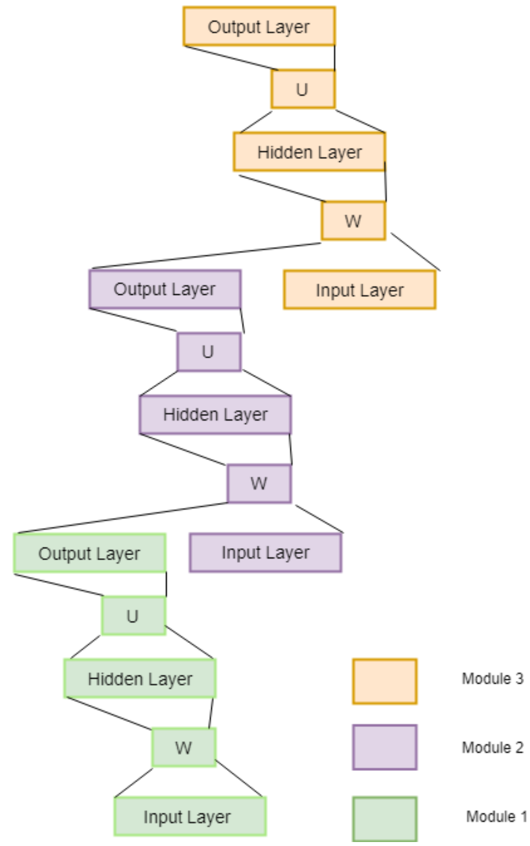

Fig. 3. Architecture of TDSN With Three Modules

blocks of TDSN. These weight matrices are used to predict the missing image.

The output of each module of TDSN is calculated as follows

$$Y_i = U^T h_i \tag{1}$$

In the equation (1) y_i is the output from each module of TDSN, U is the weight matrix between hidden layer and output layer, and h_i is the output from the hidden layers. The hidden layer output is calculated as follows.

$$hi = \sigma(W^T X_i) \tag{2}$$

In equation 2 W denotes the weight matrix between input layer and hidden layer, and X_i denotes the input vector. $\sigma(.)$ is a sigmoid function. Where, the dimension of upper matrix U is [L x C], and the dimension of lower matrix W is W [D x L]. L denotes by the number of hidden units and C denotes the dimension of the output vector.

D. Dummy Image Tunning

This module control overall forecasting process. Consider predicted missing images as a dummy image and these images are used for predicting further missing image in the sequence of time series. At the t^{th} forecasting module predict the image at the latest timestamp. This predicted image is compared to the original image I_t. If the prediction accuracy is below the threshold value the forecasting process is again triggered. This process is continued until there is no change in final predicted output or accuracy value reached the threshold level. While each iteration the weight parameter is tuned to improve the accuracy of the predicted output. At first weight matrix W and U are assigned to the random value [11].

V. RESULT AND ANALYSIS

A. Dataset and Study Area

In order to assess the proposed method, the Landsat 7 ETM+ satellite images are taken form Land Process Distributed Active Archive Center of the United States Geological Survey (USGS) [15]. These satellite image covers the area of Barddhaman district, West Bengal, India of co-ordinates, topmost $23.45°N, 87.47°E$, and bottommost $23.42°, 87.50°E$. The images are collected in the period of 2004 2011 in the month of March. Consider the images of the year 2007 and the year 2009 are missing. These missing images are choosing randomly. The raw satellite image contain multiple bands are fuesed into single multi-spectral image by using Raster package in R-language.

The accuracy of the predicted image can be analysis by using Root Mean Square Error (RMSE), Mean Average Error (MAE), Peak Signal to Noise Ratio (PSNR) and Mean Structural Similarity Index (MSSIM). RMSE and MAE are used to calculate the overall loss in the prediction process. PSNR and MSSIM are used to assess the quality of the predicted image with the original image [12]. The mathematical formulae for each of the process are given below. Equation 3 and 4 gives the formulae for RMSE and MAE respectively.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=0}^{n}(Vo - Vp)^2} \quad (3)$$

Where Vo denotes original image and Vp denotes Predicted image. n denotes total number of observations. Output of TDSN was the predicted image in a vector form. compare the original image after converted into vector form with the predicted vector. RMSE and MAE value gives the number of pixels that vary in original image and predicted image.

$$MAE = \frac{1}{n}\sum_{i=0}^{n} \mid (Vo - Vp) \mid \quad (4)$$

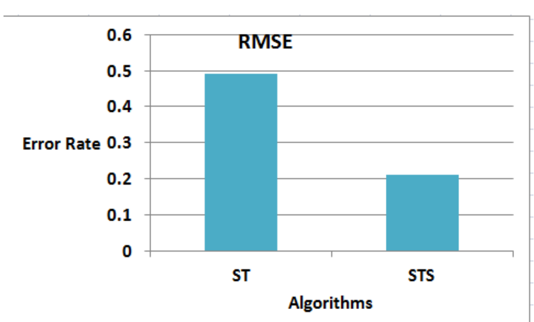

Fig. 4. RMSE

Fig. 4. shows the comparision between the RMSE value of proposed work spatial, temporal, and spectral (STS) and existing work spatial, temporal (ST). Fig. 5. shows the comparision

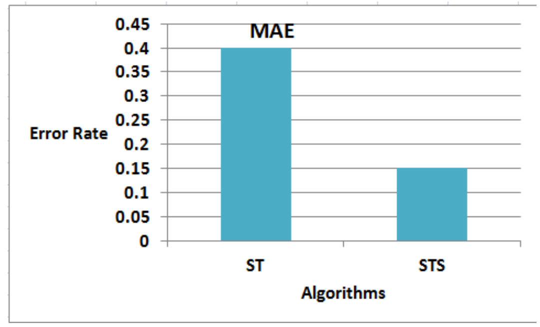

Fig. 5. MAE

between the MAE value of proposed work spatial, temporal, and spectral (STS) and existing work spatial, temporal (ST). From the Fig. 4. and Fig. 5. conclude that the overall loss of predicting missing image using spatial, temporal and spectral method is less.

The mathematical equation for PSNR and MSSIM are given as follows

$$PSNR = 20log_{10}\frac{MAX_i}{\sqrt{MAE}} \quad (5)$$

Where MAX_i is the maximum possible pixel value in the image.

$$MSSIM(r, x) = \frac{1}{n}\sum_{i=0}^{n} SSIM(r_i, x_i) \quad (6)$$

Where r and x are the referenced/original and predicted images respectively [12].

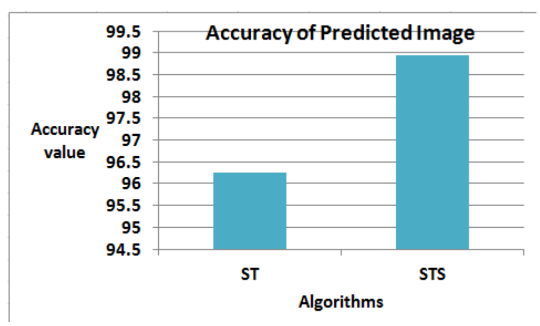

Fig. 6. Accuracy of the Predicted Image

Fig. 6. shows the accuracy of the predicted missing image calculated using PSNR and MSSIM. The accuracy of the predicted image using spatial, temporal and spectral method is high in compare to the existing spatial temporal method.

VI. CONCLUSION AND FUTURE WORK

The performance of the remote sensing time series analysis is frequently affected because of missing image. This problem can be solved by proposed work by reconstructing missing image using the spatial-temporal-spectral data in the image more accurately in compare to the existing work. The experiment is carried out during the period of 2004 to 2011 over the area of Bardhhaman district, West Bengal, India.

In future, this work can be extended to predicting missing images by using band selection and band fusion method. Choose the bandsHwhich provide more information and leave the bands that gives very few information. ence, the diff iculty in feature set preparation and computing process are reduced to predict the missing image and also get more accurate result and also time taken to predict the missing image is also reduced.

REFERENCES

[1] B. P. Salmon, K. J. Wessels, J. C. Olivier, F. van den Bergh , W. Kleynhans and K. C. Steenkamp, Unsupervised Land Cover Change Detection: Meaningful Sequential Time Series Analysis in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 4, no. 2, pp. 327-335.2011.

[2] Brian Hutchinson, Li Deng and Dong Yu, Tensor Deep Stacking Networks, IEEE Transactions on pattern analysis and machine intelligence, vol. 35, no. 8, 2013.

[3] C. C. Loy, C. Dong, K. He, and X. Tang, Image super-resolution using deep convolutional networks,IEEE Trans. Pattern Anal. Mach. Intell., vol. 38, no. 2, pp. 295307.2016.

[4] Brian Hutchinson, David Palzer The Tensor Deep Stacking Network Toolkit,Neural Networks (IJCNN) International Joint Conference on, pp.- 15.2015.

[5] M. Zorrilla, J. L. Crespo, P. Bernardos, and E. Mora, A new image prediction model based on spatio-temporal techniques,Vis. Comput., vol. 23, no. 6, pp. 419431.2007.

[6] Jose Carlos Neves Epiphanio, Julio Cesar de Oliveira , and Camilo Daleles Renno, Window Regression: A Spatial-Temporal Analysis to Estimate Pixels Classified as Low-Quality in MODIS NDVI Time Series,Remote Sensing, vol. no. 6, 3123-3142,2014.

[7] B. Hutchinson, L. Deng, and D. Yu,Parallel training for deep stacking networks, in Proc. 13th Annu. Conf. Int. Speech Commun. Assoc., pp. 14,2012.

[8] Xiaodong He, Li Deng, Jianfeng Gao,Deep stacking networks for information retrieval, Acoustics Speech and Signal Processing (ICASSP) IEEE International Conference on, pp. 3153-3157,2013

[9] D. Yu and L. Deng,Deep learning: Methods and applications, Found. Trends Signal Process., vol. 7, nos. 34, pp. 197387,2014

[10] Dong Yu , Li Deng , John Platt,Scalable stacking and learning for building deep architectures , , IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 129-132.2012.

[11] Soumya K. Ghosh, Monidipa Das, A Deep-Learning-Based Forecasting Ensemble to Predict Missing Data for Remote Sensing Analysis, Selected Topics in Applied Earth Observations and Remote Sensing IEEE Journal of, vol. 10, no. 12, pp. 5228-5236,2017

[12] Soumya K. Ghosh, Monidipa Das,Deep-STEP: A Deep Learning Approach for Spatiotemporal Prediction of Remote Sensing Data, IEEE Geoscience and Remote Sensing Letters, vol. 13, pp. 1984 1988,2016

[13] D. Jeevalakshmi, S. N. Reddy, and B. Manikiam,Land cover classification based on NDVI using LANDSAT8 time series: A case study Tirupati region, International Conference on Communication and Signal Processing (ICCSP), Melmaruvathur, pp. 1332-1335,2016

[14] S.Parthiban, NagarajThummalu, G. MeeraGandhi and A.Christy,NDVI: Vegetation Change Detection Using Remote Sensing and GIS A Case Study of Vellore District', volume 57, pg. no. 1199-1210,2015

[15] 15. USGS Earth Explorer: Land Processes Distributed Active Archive Center, accessed on Aug. 2018. [Online]. Available: https://lpdaac.usgs.gov/usgs earth explorer