ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 5, Issue 5, May 2018

Implementation of Convolution Neural Network in Processing of Satellite Video

[1] Ankitha Venkatesh, [2] Jami Sai Bandhavi, [3] Asha R, [4] B Nagarushitha, [5] Vinod B Durdi
[1][2][3][4][5] Department of Telecommunication, Dayananda Sagar College of Engineering Bangalore, India.

Abstract- **The video sent by satellite usually suffers severe degradation due to hardware imperfections or uncontrollable acquisition conditions. So there is a lot of scope in video processing-enhancement and reconstruction.The quality of image is improved by using Super-resolution. The main aim of satellite video processing is to obtain good Peak Signal to Noise Ratio (PSNR) and resolution of the stream of images of the video. To achieve this, Neural Networks are employed. Convolution Neural Networks is a part of Neural Networks which has been proven to be effective in areas of image processing.**

Index Terms— **Convolutional Neural Networks (CNN), Super Resolution (SR), Peak Signal to Noise Ratio (PSNR).**

I. INTRODUCTION

A Satellite captures continuous images of targets in a certain time period which has to be dynamically monitored in real time. This information helps in various fields like agriculture, weather forecast, remote sensing, floods, volcanic eruption etc. identifying changes in the landscape, or forecast future crop yields, remote sensing offers the capability to derive such information over large areas and with minimal expense compared to conventional methods.[6] Every feature on our planet exhibits a unique spectral character and remote sensing allows us to detect these signatures to derive information about the land. Super-resolution is employed to obtain high-quality images.[1] Super-resolution algorithm enhances high-frequency information to improve the image quality.[8] The obtained quality of the image would be degraded by illumination, angle, distance and other conditions, and it would result in failing to recognize the content. As a result, various methods are implemented to increase the resolution. Few of them are discussed -

Interpolation is a traditional method for super resolution, but this method does not ignore high frequency information. Besides, deep learning has a great performance in many applications. Convolution Neural Network consists of one or more convolution layers followed by one or more fully connected layers. A large dataset is available for training the neural network. In this paper SRCNN (Super Resolution Convolution Neural Network) is employed. In this method, the neural network learns an end-to-end mapping between the low-resolution input images with high-resolution output images. This method ensures state-of-the-art restoration quality and achieves fast speed for practical on-line usage hence this is chosen.

II. CONVOLUTION NEURAL NETWORKS

A CNN is an architecture commonly used for deep learning. CNN are feedforward networks where the information flow is unidirectional (from input to output) [9]. CNN has been proven to master in knowing the structure of raw image data. Image modeling field has to undergo many pre-processing techniques in order to get the input images aligned and transform it into which the modeling techniques can handle it better. However, scaling and slight rotations in the images make image processing a hard task. CNN has made it possible for the network to handle the raw image data [1]. CNN is used for recognition of objects and scenes and perform object detection and segmentation. Though there are a lot of variations in CNN architecture, generally it is considered to consist of convolutional and pooling or subsampling layers, which are grouped into modules. As in standard feedforward neural networks, either one or more fully connected layers follow these modules. To form a deep model these modules are often stacked on top of each other. The depth of the filter increases from left to right in the network. [9] The use of CNN for deep learning is popular due to three reasons- eliminates the need for manual feature extraction, produces state-of-the-art recognition results and CNN's can be retrained for new recognition tasks and also allow building on pre-existing networks.There are three ways to train CNN's for analysis of image- training the model from scratch, transfer learning and using a pre-trained CNN to extract features to train machine learning model.

III. SRCNN

A high-resolution image can be decomposed into a low-frequency information (corresponding to low-resolution

ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 5, Issue 5, May 2018

image) and high-frequency information (residual image or image details) using SRCNN [2]. Input and output images share the same low-frequency information. This is used to model the high-resolution image. Carrying the input to the end layer and reconstruction residuals is achieved by SRCNN. Moreover, SRCNN uses different learning rates for different layers in order to achieve stable convergence. SRCNN has been proved to be more effective than other learning-based SR methods, it lacks high-frequency details, as the architecture must preserve all input details since the image is discarded and the output is generated based on the learned features only [9]. Model SRCNN consists of three layers namely patch extraction/representation, non-linear mapping, and reconstruction. Few methods employ bicubic interpolation as its pre-processing step followed by the extraction of overlapping patches, via convolutional, as high dimensional vectors with many feature maps as their dimensions. SRCNN has only convolutional layers which can be advantageous as input images can be of any size and the algorithm is not patch-based. The output obtained through SRCNN model will be of a smaller size when compared to the input image. The output obtained through SRCNN model will be of a smaller size when compared to the input image.

III. EXISTING METHOD

Remote sensing systems that acquire images with large spatial extents will generally have a lower resolution, and thereby capture less detail, and hence has to be processed. One of the processing methods employed is Data fusion.[10] The integration of data and knowledge from several sources is known as data fusion. It can also be defined as a multi-level process dealing with the association, correlation, combination of data and information from single and multiple sources to achieve the refined position, identify estimates and complete and timely assessments of situations, threats and their significance. Different categories of Data fusion are (i) data association, (ii) state estimation, and (iii) decision fusion. Data fusion finds wide application in many areas of robotics such as object recognition, environment mapping, and localization.[10] Data fusion systems are complex combinations of sensor devices, processing, and fusion algorithms.

IV. PROPOSED METHOD

In paper proposed in IEEE journal, the semantic relationship between the objects in the image can be constructed within a hierarchy network. The method includes two modules namely segmentation and classification. The segmentation is grouping image pixels into the meaningful objects with the attribution of spectral, spatial and semantic information. The classification is performed in the semantic hierarchy with different scales or the image resolution, in which the hierarchy network is constructed by the classifying system or semantic relationship.[8] Then the classified objects are aggregated into the semantic groups called classification-based objects. In [2], various methods are proposed for increasing the resolution of the images. the difficulties encountered during the recovery of the super-resolved version of the given input signal (low resolution) is mainly focused. Some of them include Reconstruction constraint here the observed low-resolution image which is blurred and downsampled version of the high-resolution image is passed through a blurring filter to get the high-resolution output image. Sparsity prior is another method where, the sparse representation for each local patch, respecting spatial compatibility between neighbors is found. Later using the result from this local sparse representation, further regularization, and refinement of the entire image using the reconstruction constraint. Various methods have been proposed to solve image processing issues by using patch-based matching, coupled subspace learning, coupled dictionary learning techniques, etc.

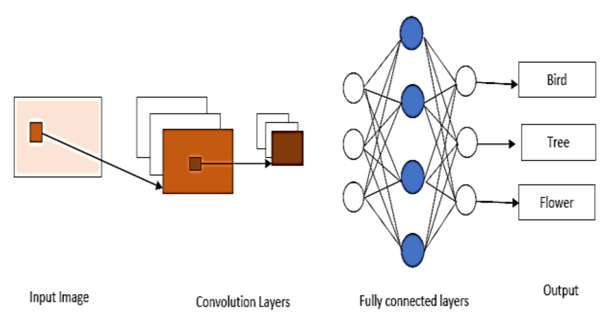

Fig 1. Neural Network.

In [3], a semi-coupled dictionary learning (SCDL) model is proposed to convert the image from one style to another will cut the error by half. Investigation of Very Deep CNN and Advanced Training Strategies for Document Image Classification, proposed under IEEE explains various deep

IFERP
connecting engineers... developing research

ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 5, Issue 5, May 2018

learning convolutional neural network architecture, Some of them are AlexNet, VGG-16, GoogLeNet, Resnet-5 among others.In Video Super-resolution via Convolutional Neural Network, published under IEEE journal, image stream classifies each patch into the changing patch or unchanging patch.A three step three-step search algorithm is employed, where the unmatching patches and the changing patches, between two consecutive frames is found. In [4], Super Resolution using Random Forests (SRF) is explained.This method is a combination of Single Image Super Resolution (SISR) and Random Forests. SISR is employed to generate a high-resolution output from a single low-resolution input image. The algorithm Random forests are highly non-linear learners that are extremely fast during both learning and evaluation. They can easily handle high-dimensional noisy inputs, which has led to their broad dissemination in many computer vision domains. The experimental setup is outlined including details on data sets and evaluation metrics, which are compared with different random forest variants as well as several state-of-the-art SISR approaches.
In the proposed system, the video captured from the satellite is received by the base station. Here the video is splitted into multiple images, each image is passed through a Three-layer convolution network. The resolution of images are increased here. The neural network is trained, tested and implemented to increase the resolution. Later the high resolution images are combined to produce the required video. The PSNR of input images and the output images of high resolution are compared.

EXPECTED OUTCOME

To obtain higher resolution of the input video signal and to obtain a better PSNR of at least 30dB.

ACKNOWLEDGEMENT

We thank Head of our Department for giving us this opportunity, we also thank our guide for helping us and guiding us through this endeavour.

REFERENCES

[1] Yimin Luo, Liguo Zhou, Shu Wang, and Zhongyuan Wang," Video Satellite Imagery Super Resolution via Convolutional Neural Networks," IEEE,VOL. 14, NO. 12, Dec 2017.

[2] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," IEEE Trans. Image Process., vol. 19, no. 11, pp. 2861–2873, Nov. 2010.

[3] S. Wang, L. Zhang, Y. Liang, and Q. Pan, "Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis," in Proc. CVPR, Providence, RI, USA, Jun. 2012,pp. 2216–2223.

[4] S. Schulter, C. Leistner, and H. Bischof, "Fast and accurate image upscaling with super-resolution forests," in Proc. CVPR, Boston, MA, USA, 2015, pp. 3791–3799.

[5] S. Ji, W. Xu, M. Yang, and K. Yu, "3D convolutional neural networks for human action recognition," IEEE Trans. Pattern Anal. Mach. Intell, vol. 35, no. 1, pp. 221–231, Jan. 2013.

[6] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in Proc. ICLR, San Diego, CA, USA, 2015, pp. 1–14.[4] Y. Wen, K. Zhang, Z. Li, and Y. Qiao, "A discriminative feature learning approach for deep face recognition," in Proc. ECCV, Amsterdam, The Netherlands, 2016, pp. 499–515.

[7] P. E. Dennison, A. R. Brunelle, and V. A. Carter, "Assessing canopy mortality during a mountain pine beetle outbreak using GeoEye-1 high spatial resolution satellite data," Remote Sens. Environ., vol. 114, no. 11, pp. 2431–2435, 2010.

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in Proc. NIPS, Lake Tahoe, NV, USA, 2012, pp. 1097–1105.

[9] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in Proc. CVPR, Boston, MA, USA, 2015, pp. 1646–1654.

[10]Q. Cheng, H. Liu, H. Shen, P. Wu, and L. Zhang, "A spatial and temporal nonlocal filter-based data fusion method," IEEE Trans. Geosci. Remote Sens., vol. 55, no. 8, pp. 4476–4488, Aug. 2017.

