International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2019): 7.583

Dense Fusion of Infrared and Visible Images

Gopika Ajay[1], H. Shihabudeen[2]

[1, 2]Department of Electronics and Communication Engineering, College of Engineering Kidangoor, Kerala, India

Abstract: ***This paper introduce a systematic learning architecture for the fusion of infrared and visible images. . Infrared and visible images are used as input images. Visible image capture reflected light and infrared capture thermal radiation. So the fused image of visible and infrared images provide better information and quality. Architecture have three main sections. The first section is encoding network. Inside the encoder network, conventional layers and dense block are present. The encoder section will extract rough as well as deep features from source images. Second section is Fusion layers. It follows two strategies, and designed to fuse these features. There are two fusion strategies, one is addition strategy and other one is l1-norm strategy is doing. Third section is decoder, it reconstruct the fused images. The final out is more informative and accurate than any single source image, and it consists of all the necessary information. The paper compare the proposed model with the existing fusion technology and also discuss the future scopes of proposed model.***

Keywords: dense fusion, visible images, infrared images, addition strategy, l1 – norm strategy

1. Introduction

The image fusion is the gathering of all the important information from multiple images, and inclusion into fewer images, usually a single image. This single image is more informative and accurate than any single source image or input image, and the final image consists of all the necessary information. The need of image fusion is not only to reduce the amount of data but also to construct images that are more appropriate and understandable for the human as well as machine perception.

A single image cannot provide all the objects in a scene. So in many situations multi-focus image fusion technique is used, which fuses several images of scene captured with focus on various objects using different sensors and then these images are fused to form a single image which provide all the information of objects in the scene.

Infrared and visible images are used as the source images. The task of visible and infrared image fusion is an important problem in image processing field. By fusion it attempts to extract salient features from input or source images, then integrated these features into a single image by appropriate fusion method.

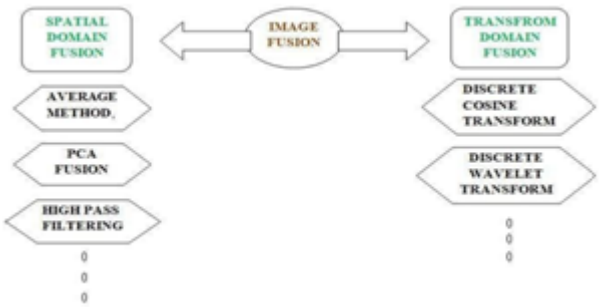

Figure 1: Classification of image fusion methods

2. Literature Survey

Evolution of Image Fusion and Different Image Fusion Techniques

A good image fusion method has the following properties. It must preserve the useful information of different images, it does not produce artifacts, which can mislead or distract a human observer or any subsequent image processing steps, it must be reliable and robust, it should not eliminate any salient information contained in any of the input images.

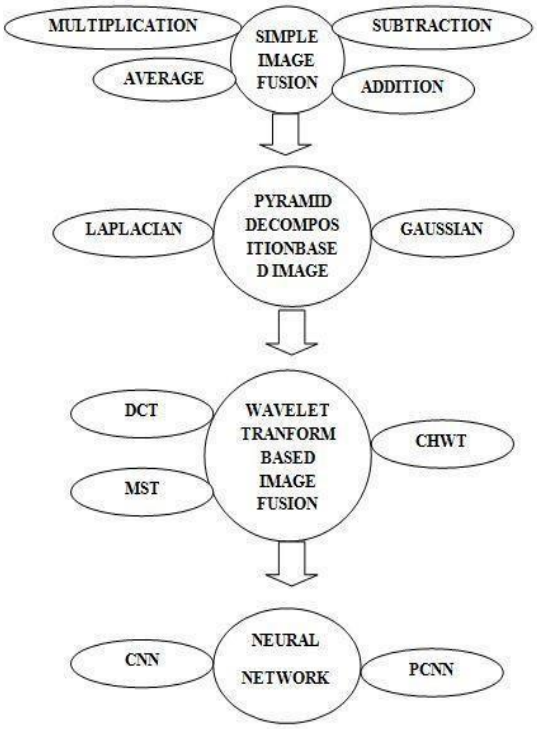

Figure 2: Evolution of image fusion

In [1] describes a survey of image fusion and the paper focusing mainly into the methods and techniques.

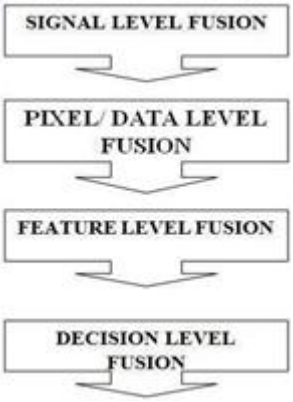

Figure 3: Evolution of image fusion technology

Pixel Based Image Fusion

In [2] introducing pixel level image fusion is the processing and combination of information gathered by various imaging sources to provide a better information of a scene. In this paper the image fusion is considered as an optimization problem and propose an information theoretic approach in a multi scale framework to get the solution. A biorthogonal wavelet transform of each input or source image is first calculated, and a new Jensen-Renyi divergence based fusion algorithm is developed to construct composite wavelet coefficients according to the measurement of the information patterns inherent in the source images.

Drawbacks of pixel level image fusion is Image noise, resolution difference between images, imperfect results, environmental conditions are badly affected to pixel level image fusion ect.,

PCNN Based Image Fusion

Many researchers have paid their time into image fusion technique based on pulse coupled neural network (PCNN). In [2] a systematically reviewed of PCNN is done. In order to make the researchers to understand the research development of image fusion based on PCNN, that will be very helpful. The main drawback of PCNN based image fusion is a deep study of PCNN is needed for image fusion and it having so many rules.

Cnn Based Image Fusion

Convolution Neural Network (CNN) is the most active field of research. Its architectures are used for the improvement of accuracy and efficiency in various fields. CNN is used in [4] in order to generate fusion of visible and thermal images to detect presents of persons in those images for the reliable surveillance application. There are various kinds of image fusion methods to achieve multi-sensor, multi-modal, multi-focus and multi-view image fusion. The proposed methodology includes Encoder-Decoder architecture for fusion of visible and thermal images and ResNet-152 architecture for classification of images to detect if there is a person present in the image or not. Korea Advanced Institute of Science and Technology (KAIST) multispectral dataset consisting of 95,000 visible and thermal images is used for training of CNNs. Drawback of CNN based image fusion is overfitting. Overfitting is the production of an analysis that corresponds too closely or exactly to a particular set of data, i.e. due to overfitting the system may fail to fit additional data or predict future observations reliably.

3. Proposed Work

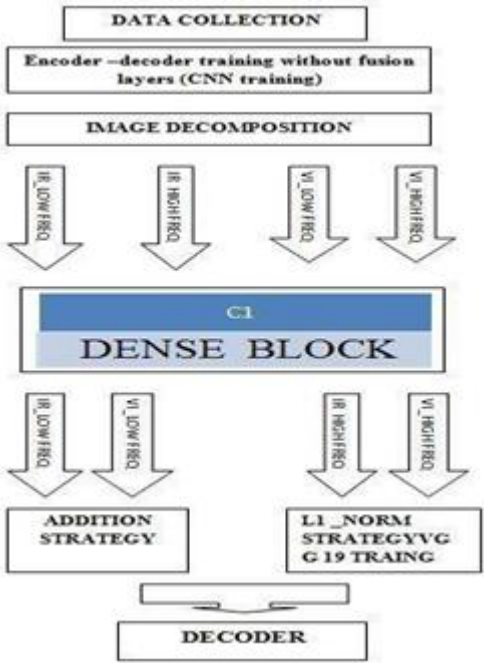

Figure 3: Flow chart of the project

$$f^m(x,y) = \sum_{i=1}^{k} \phi_i^m(x,y)$$

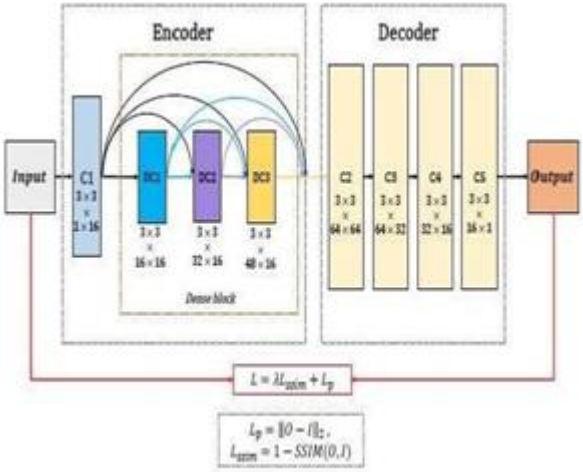

Figure 4: Block diagram of the proposed project

The block diagram have three sections, encoder part, fusion layers and decoder part. Encoder consist of two subsections called C1 and dense block. Both are convolutional layers. C1 extract the rough features and dense block extract the deep features from the source images. The fusion layer will fuse these features. There are two fusion strategies one is addition strategy and other one is l1_norm strategy.

Training Phase

In order to reconstruct the input image more accurately, loss function L should be minimum, to get minimum L train the encoder and decoder as shown in figure 5.

Figure 5: End to end training

Lp is pixel loss and SSIM is structural similarity and λ is weight.

The pixel loss Lp is calculated by the following equation .

$$L_p = ||O - I||_2$$

Here O is output and I is input. Lp will calculate the the euclidean distance between the output O and the input I.

The SSIM loss Lssim is obtained by the following equation.

$$L_{ssim} = 1 - SSIM(O, I)$$

Addition Strategy

The addition strategy is prepare methodically by following equation.

Here Q^mi is the feature maps obtained by encoder from source images. (x; y) indicates the position in feature maps and fused feature maps. And f^m will be the input to decoder.

L1 – Norm Strategy

L1 – norm strategy have two operations l1 – norm operation and soft-max operation. The initial activity level map Ci is calculated as

$$C_i(x, y) = ||\phi_i^{1:M}(x, y)||_1$$

final activity level map is calculated as

$$\hat{C}_i(x, y) = \frac{\sum_{a=-r}^{r} \sum_{b=-r}^{r} C_i(x + a, y + b)}{(2r + 1)^2}$$

Here r is 1.

Final out of l1 – norm is calculated as

$$f^m(x, y) = \sum_{i=1}^{k} w_i(x, y) \times \phi_i^m(x, y),$$

$$w_i(x, y) = \frac{\hat{C}_i(x,y)}{\sum_{n=1}^{k} \hat{C}_n(x,y)}$$

4. Comparison

Several typical fusion methods are existing. Some of them are cross bilateral filter fusion method (CBF), the joint sparse representation model (JSR), gradient transfer and total variation minimization (GGTF), the JSR model with saliency detection fusion method (JSRSD), deep convolutional neural network based method (CNN) and the Deep Fuse**.**

The fused images by CBF, JSR and JSRSD are having artificial noise and the features are not clear. But the fused images by GFT, CNN and deep fuse methods are preserves more information. The proposed model fuse the CNN and dense fuse model together to get maximum information.

5. Conclusion

The paper present systematic learning architecture based on CNN and dense block. The architecture focus on the image fusion problem of infrared and visible. And the proposed model can be used for the fusion of RGB images.

The proposed model has three sections encoder, fusion layer and decoder.

The feature maps are obtained in encoder. the input or the source images to the encoder are infrared and visible images. Since the encoder has C1 and dense block the chances of losing information is very less. In fusion layer features are fused by addition and l1-norm strategy. Then the feature maps are integrated into a single feature map which contains all salient features from input images.

The decoder will reconstructed the final out of the proposed model.

References

[1] **R.Johnson Suthakar, J.Monica Esther, D.Annapoorani, F.Richard Singh Samuel** (2014) Study of Image Fusion-Techniques, Method and Applications, *JCSMC, Vol. 3, Issue.* 11,November

[2] **Ben Hamza, Yun He, Hamid Krim, and Alan Willsky** (2005) A Multiscale Approach to Pixel-Level Image Fusion, *Integr. Comput. Aided. Eng*., vol. 12, pp.135146

[3] **Zhzaobin wang, Shuai Wang, Ying Zhu & Yide Ma** (2015)Review of image fusion based on pulse coupled neural network(PCNN) research gate

[4] **BUSHRA KHALID, ASAD MANSOOR KHAN, MUHAMMAD USMAN AKRAM, SHERIN BATOOL** (2019) PERSON DETECTION BY FUSION OF VISIBLE AND THERMAL IMAGES USING CONVOLUTIONAL NEURAL NETWORK, IEEE *2ND INTERNATIONAL CONFERENCE ON COMMUNICATION, COMPUTING AND DIGITAL SYSTEMS (C-CODE)*, ISLAMABAD, PAKISTAN, 2019, PP. 143-148, DOI: 10.1109/C-CODE.2019.8680991.

[5] **Li S, Kang X, Fang L, et al.** (2017). Pixel-level image fusion: A survey of the state of the art[J]. Information Fusion.

[6] **Ben Hamza A, He Y, Krim H, et al.** (2005)A multiscale approach to pixel-level image fusion[J]. Integrated Computer-Aided Engineering.

[7] **Yang S, Wang M, Jiao L, et al.** (2010) Image fusion based on a new contourlet packet[J]. Information Fusion.

[8] **Wang L, Li B, Tian L F.** (2014) EGGDD: An explicit dependency model for multi-modal medical image fusion in shift-invariant shearlet transform domain[J]. Information Fusion,

[9] **Pang H, Zhu M, Guo L.** (2012) Multifocus color image fusion using quaternion wavelet ransform[C]//*Image and Signal Processing (CISP), 5thInternational Congress on. IEEE,*