May – June 2020
ISSN: 0193-4120 Page No. 15862 – 15868

Feasibility of Convolutional Neural Networks (CNN) for the Fusion of Temporal Medical Images

[1]M.A.Muthiah, [2]E.Logashanmugam, [2]N.M.Nandhitha, [3]Mohammad Azhar, [3]Shaik Irshad Ahamed
[1]Research scholar, Department of ECE, Sathyabama Institute of Science and Technology, Old Mamallapuram Road, Jeppiaar Nagar, Chennai 600119
Email: muthiah.m.a@outlook.com
[2]Professor, School of Electrical and Electronics Engineering, Sathyabama Institute of Science and Technology, Old Mamallapuram Road, Jeppiaar Nagar, Chennai 600119
[3]Final year students, Dept. of ECE, Sathyabama Institute of Science and Technology, Old Mamallapuram Road, Jeppiaar Nagar, Chennai 600119

Article Info
Volume 83
Page Number: 15862 – 15868
Publication Issue:
May - June 2020

Article History
Article Received*: 1May 2020*
Revised*: 11 May 2020*
Accepted*: 20 May 2020*
Publication*: 24May 2020*

Abstract

In medical applications, a decision is made based on the collective information obtained about an abnormality after analyzing a series of images instead of a single image. Analyzing a set of images and combining the information from each image is time consuming and tedious. Instead, if the features of interest in the set of images could be provided on a single image, decision making becomes easy. Process of combining images is termed as image fusion. This paper reports the image fusion techniques performed in spatial domain (Principal Component Analysis (PCA) based fusion), transform domain (Discrete Cosine Transform (DCT)) and Convolutional Neural Network (CNN) used for the fusion of MRI images. Convolutional layers and max pooling layers in VGG16, VGG19 and ALEXNET are used for extracting the features and fusion rules are used for obtaining the output images. Performance is measured in terms of Signal to Noise Ratio (SNR), Root Mean Square Error (RMSE), entropy and standard deviation. It is observed that VGG19 outperforms other image fusion techniques and provides consistently good performance which is evident from the performance evaluation.

Keywords: Fusion, MRI, features, PCA, DCT, VGG16, VGG19, ALEXNET

I. INTRODUCTION

In order to substantiate the results of medical diagnosis, the procedure is repeated on the same specimen at different intervals of time. Also, by virtue in Magnetic Resonance Imaging (MRI), the set of images are obtained at different slices but with the same equipment for the same patient. Hence the interpreter has to interpret the large set of images in order to arrive at a conclusion. It will be helpful, if all these images are combined and the interpreter can make decisions based on a single image rather than on a set of images. Image fusion is one such process where the input images are combined to form an output image. The output image contains both the complementary information and the redundant information. Various types of image fusion are cited in literature. Sumir RM et al (2019) proposed a method by deriving the gradients based on weights of each pixel in an image. R. Ahmmed and M. F. Hossain (2017) proposed a model that includes the template-based K means and improved fuzzy C means (TKFCM) algorithm for detecting human brain tumors in a magnetic resonance imaging (MRI) image.

Initially image fusion began with pixel-based image fusion, where the pixel intensities of the input images are combined using a specific set of rules to form an output image (K.P.Indira et al, 2015, Ella Madhava Babu et al., 2017). Later region level image fusion emerged wherein only the Regions of Interest are visible in the output image. Determining the Region of Interest is a subjective process and hence the output may not reflect the actual information. Hence the paradigm has shifted to feature level image fusion. In feature level image fusion, the features are extracted from the image using an appropriate transform and the features are then combined to form an output image. Features can be extracted through unitary transform decomposition, Stockwell decomposition, Walsh Hadamard decomposition etc. Also the principal components can be obtained and the features can be combined.

However these features greatly affect the output performance. Hence it is necessary to determine efficient feature extraction. Convolutional Neural Networks are pretrained networks which are proven to extract the features accurately and precisely. DCT and PCA based image fusion techniques are performed. Also, the feasibility of CNN for feature extraction and hence image fusion is also studied. Performance is measured in terms of Signal to Noise Ratio, Root Mean Square Error, Entropy and Standard deviation.

II. Research database

Research database used in this work is Whole Brain Atlas - Harvard Medical School http://med.harvard.edu/aanlib (Vidoni, Eric D, 2012) . Of the ten pairs of images (1.1a, 1.1b…1.10a-1.10b) used in this work, two pairs are shown in Figure 1. These images are gray scale images and the features are extracted from these images to perform fusion.

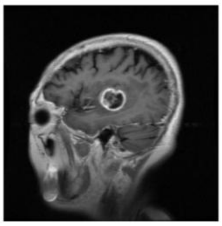


Figure 1.1a

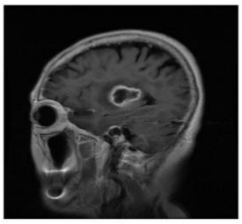


Figure 1.1b

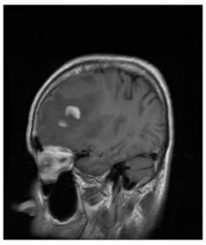


Figure 1.3a

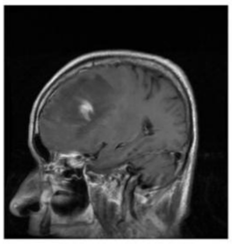


Figure 1.3b

Figure 1 Set of MI images used for mage fusion

III. Methodology for DCT based Image Fusion

Images to be fused are divided into non-overlapping blocks of size NxN. DCT coefficients are computed for each block and fusion rules are applied to get fused DCT coefficients. IDCT is then applied on the fused coefficients to produce the fused image/block. The procedure is repeated for each block. Image fusion is performed on the seven sets of input images and the performance is measured in terms of terms of Signal to Noise Ratio (SNR), Root Mean Square Error (RMSE) and Standard deviation. Signal to Noise ratio in image fusion refers to the amount of information retained in the output image from the input image. Root Mean Square Error is the measure of the amount of information (in the output image) missed from the corresponding input image. Standard deviation is a measure of how well the discontinuities are visible in the output image. Performance metrics for DCT based image fusion is shown in Table 1

Table 1 Performance metrics for DCT based Image fusion

Image #1	Image #2	SNR #1	SNR #2	RMSE #1	RMSE #2	Standard Deviation
1.1a	1.1b	41.00	42.00	21.72	27.70	0.414549
1.2a	1.2b	43.00	44.00	26.74	15.49	0.411569
1.3a	1.3b	45.00	46.00	22.15	32.28	0.407451
1.4a	1.4b	47.00	48.00	19.14	23.57	0.407961

May – June 2020
ISSN: 0193-4120 Page No. 15862 – 15868

1.5a	1.5b	49.00	50.00	20.42	15.31	0.406196
1.6a	1.6b	51.00	52.00	21.65	29.73	0.399529
1.7a	1.7b	53.00	54.00	16.79	19.30	0.409882

SNR can further be improved and the Root Mean Square Error can be reduced. In order to improve the performance of the fusion technique, PCA based image fusion is also performed.

IV. Methodology for PCA based Image Fusion

A weighted fusion algorithm using PCA consists of two source images. Initially images are smoothened. High frequency and the low frequency components are separated using Gaussian filter. The covariance matrix of the difference image data vectors is calculated. The principle eigenvector, corresponding to the maximum eigen value of the covariance matrix is determined. The components of the principle eigenvector are used as the weights. The fused image is composed by averaging the smoothed input images and adding the weighted, normalized sum of the deviations. Performance metrics for PCA based image fusion is given in Table 2.

Table 2 Performance metrics for PCA based image fusion

Image #1	Image #2	SNR #1	SNR #2	RMSE #1	RMSE #2	Standard deviation
1.1a	1.1b	41.00	42.00	17.46	17.31	0.424431
1.2a	1.2b	43.00	44.00	15.12	15.29	0.418431
1.3a	1.3b	45.00	46.00	12.88	19.46	0.404902
1.4a	1.4b	47.00	48.00	6.25	15.06	0.446588
1.5a	1.5b	49.00	50.00	12.41	12.57	0.412392
1.6a	1.6b	51.00	52.00	18.36	17.99	0.413451
1.7a	1.7b	53.00	54.00	12.01	12.48	0.418824

From the Table 2, it is evident that the performance has improved in terms of standard deviation. It implies that the contrast information is present in the output image. But the redundant and complementary information are not transferred from the input images to the output image. It is because of the fact that the features extracted by PCA or DCT could not capture the characteristics of the input images. Hence the paradigm has shifted to deep learning neural networks.

V. Feasibility of CNN for Image Fusion

Steps involved in the fusion of images using CNN are as follows: A set of 10 image pairs were given to the networks. Features are extracted by using the Convolutional and max pooling layers. These features are then combined using weight map and masking technique (M.A.Muthiah et al, 2019). These features are then combined and the output images are obtained.

In this work, VGG16, VGG19 and ALEXNET are used for extracting features from the set of input images. Though there are 13, 5 and 3 Convolutional layers, max pooling layers and dense layers respectively, only 16 weight layers are present and hence the name VGG16. Performance is measured in terms of Signal to Noise ratio, Mean, Standard Deviation, entropy and RMSE (Table 3).

Table 3 Performance Metrics of VGG16 for Image fusion

Images #1	Images #2	SNR 31	SNR #2	Mean	Standard deviation	Entropy	RMSE

TEST Engineering & Management

May – June 2020
ISSN: 0193-4120 Page No. 15862 – 15868

1.1a	1.1b	71.8084	72.613	0.4627	0.41828	4.91106	0.5745
1.2a	1.2b	84.0768	70.52	0.4493	0.42048	4.87583	0.56306
1.3a	1.3b	68.8922	72.784	0.4717	0.41363	4.8907	0.57958
1.4a	1.4b	65.288	74.32	0.4667	0.41235	4.94004	0.57355
1.5a	1.5b	90.4624	71.075	0.4501	0.41338	5.0449	0.56353
1.6a	1.6b	68.544	74.33	0.46217	0.41602	4.92773	0.57207
1.7a	1.7b	61.0831	72.803	0.46177	0.42062	4.8327	0.57563
1.8a	1.8b	60.631	70.266	0.45356	0.41946	4.8147	0.56631
1.9a	1.9b	53.4605	73.871	0.4605	0.41988	4.81321	0.57368
1.10a	1.10b	70.6927	71.317	0.466	0.42174	4.86817	0.58088

In order to improve the performance of the technique, VGG19 is used for feature extraction. In VGG19, there are 19 weight layers. VGG19 is also a pretrained network which accepts input 224x224 images. Performance metrics for VGG19 is shown in Table 4. From the Table 4, it is found that SNR #1 has improved than that of VGG16. However the remaining metrics remain the same.

Table 4 Performance Evaluation of VGG19 for Image fusion

Images #1	**Images #2**	**SNR #1**	**SNR #2**	**Mean**	**Standard deviation**	**Entropy**	**RMSE**
1.1a	1.1b	72.613	72.630	0.46278	0.418279	4.911059	0.574557
1.2a	1.2b	70.520	70.508	0.44939	0.420486	4.875830	0.563067
1.3a	1.3b	72.784	72.777	0.47177	0.413638	4.890792	0.579584
1.4a	1.4b	74.320	74.330	0.46673	0.412356	4.940043	0.573557
1.5a	1.5b	71.075	71.062	0.45615	0.413385	5.054489	0.563533
1.6a	1.6b	44.330	74.324	0.46217	0.416019	4.927732	0.572069
1.7a	1.7b	72.803	72.817	0.46177	0.420618	4.832701	0.575634
1.8a	1.8b	70.266	70.249	0.45356	0.419463	4.814699	0.566312
1.9a	1.9b	73.871	73.869	0.46050	0.419877	4.813207	0.573676
1.10a	1.10b	71.317	71.303	0.46666	0.421743	4.868165	0.580881

With an intention of improving the performance, ALEXNET is used for image fusion. ALEXNET uses Convolutional layers, max pooling layers, dropout, data augmentation, Rectified Linear Unit and fully connected layers (Nur Azida Muhammad et al. 2018). Performance is measured in terms of Signal to Noise ratio, Mean, Standard Deviation, entropy and RMSE and is shown in Table 5. From the third column of Table 5, it is

TEST Engineering & Management

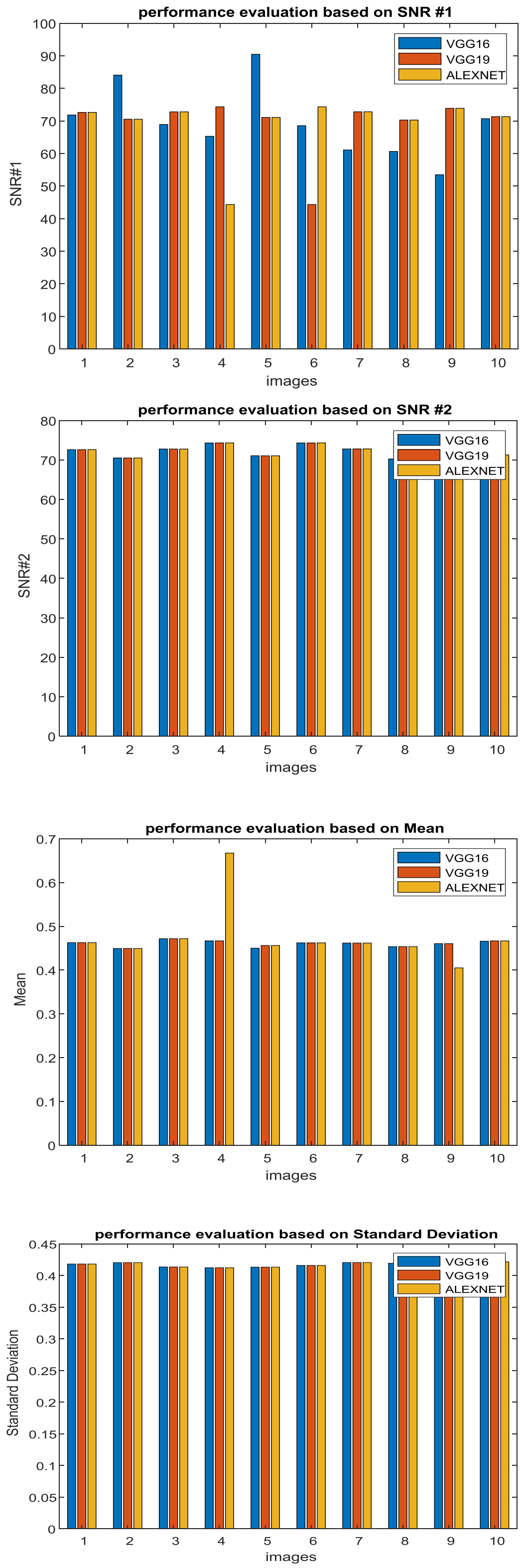
May – June 2020
ISSN: 0193-4120 Page No. 15862 – 15868

found that SNR has not improved than that of VGG19 but has decreased.

Table 5 Performance Evaluation of ALEXNET for Image fusion

Images #1	Images #2	SNR 31	SNR #2	Mean	Standard deviation	Entropy	RMSE
1.1a	1.1b	72.613	72.630	0.462783	0.418279	4.911061	0.574557
1.2a	1.2b	70.520	70.508	0.449392	0.420486	4.875240	0.563067
1.3a	1.3b	72.784	72.777	0.471777	0.413638	4.890790	0.57958
1.4a	1.4b	44.320	74.330	0.66732	0.412356	4.934707	0.573557
1.5a	1.5b	71.075	71.062	0.456156	0.413386	5.044893	0.563533
1.6a	1.6b	74.330	74.324	0.462173	0.416019	4.927733	0.57069
1.7a	1.7b	72.803	72.817	0.461774	0.420618	4.832699	0.57534
1.8a	1.8b	70.266	70.249	0.453563	0.412342	4.814698	0.566312
1.9a	1.9b	73.871	73.869	0.40507	0.419877	4.813207	0.573676
1.10a	1.10b	71.317	71.303	0.46666	0.42174	4.868166	0.580881

In order to understand the performance further, bar charts are drawn for comparing the performance of these three CNN (Figure 2). From the Figure 2, it is found that VGG19 consistently performs better than the other two networks for the given set of input images. By virtue to its architecture, VGG19 has determined the optimal features for these set of input images.



May – June 2020
ISSN: 0193-4120 Page No. 15862 – 15868

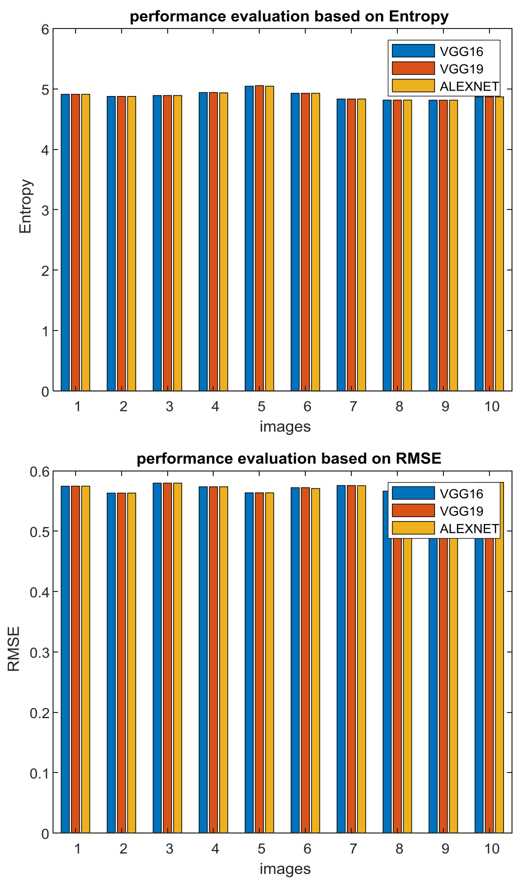


Figure 2 Performance Evaluation of CNN for fusion of medical images

VI. Conclusion and Future work

In this work, feature level image fusion is performed on ten set of MRI images depicting glioma. Three different methodologies are adopted for extracting features. Principal Component Analysis is used for extracting the features from the original intensity image and fusion is performed. Image is transformed from spatial domain to spectral domain using Discrete Cosine transform and the features of the four sub regions (Approximation co-efficients, horizontal detail, vertical detail and diagonal details) are combined to form the output image. Thirdly VGGNET and ALEXNET are used for extracting features and these features are combined to form the output image. Performance is measured in terms of Signal to Noise ratio (SNR), Root Mean Square Error (RMSE), entropy and standard deviation. From the performance evaluation, it is found that VGG19 provides the best output when compared to that of other networks.

However the image fusion is restricted to temporal images. Hence the work can be extended for spatiotemporal images but after performing image registration. Also the work has stopped at the fusion level which means that decision making is not performed. Hence decision making can also be performed. Variations in the CNN architecture for the improvement of feature extraction can also be performed.

REFERENCES

[1] S. R.M., S. Mishra and N. Shastry, "Segmentation of Brain Tumor from MRI Images using Fast Marching Method," 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore, India, 2019, pp. 1-5.

[2] M. A. Muthiah, E. Logashamugam and B. V. K. Reddy, "Fusion of MRI and PET Images Using Deep Learning Neural Networks," 2019 2nd International Conference on Power and Embedded Drive Control (ICPEDC), Chennai, India, 2019, pp. 283-287.

[3] Nur Azida Muhammad, Amelina Ab Nasir, Nurbaity Sabri, "Evaluation of CNN, Alexnet and GoogleNet for fruit recognition", Indonesian Journal of Electrical Engineering and Computer Science, Vol. 12, No. 2, November 2018, pp. 468-475

[4] R. Ahmmed and M. F. Hossain, "Tumor detection in brain MRI image using template based K-means and Fuzzy C-means clustering algorithm," *2016 International Conference on Computer Communication and Informatics (ICCCI)*, Coimbatore, 2016, pp. 1-6..

[5] KP Indira, R Rani Hemamalini, NM Nandhitha, Performance evaluation of DWT, SWT and NSCT for fusion of PET and CT Images using different fusion rules, Biomedical Research (2016) Volume 27, Issue 1, 2015, pp. 123-131.

[6] Ella Madhava Babu, S.Dushyanth Maniks, N.M.Nandhitha et al, Two-Dimensional Stockwell Transform Based Image Fusion for Combining Multifocal Images Proceedings of the International

Conference on Intelligent Sustainable Systems (ICISS 2017), IEEE Xplore Compliant - Part Number:CFP17M19-ART,ISBN:978-1-5386-1959-9

[7] Vidoni, Eric D., PT, PhD The Whole Brain Atlas, Journal of Neurologic Physical Therapy: June 2012 - Volume 36 - Issue 2 - p 108 doi: 10.1097/NPT.0b013e3182563795