INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE

ISSN: 2057-5688

A Medical Image Fusion Method using Deep Learning Convolutional Neural Network

M S Ashish(168R1A0495), M Rishabh(168R1A0496), M Arif(168R1A04A1), M Sainath(168R1A04A2)
Department of ECE, C M R Enginnering college,Hyderabad, Telangana, INDIA,

Abstract-

In medical imaging using different modalities such as MRI and CT, complementary information of a targeted organ will be captured. All the necessary information from these two modalities has to be integrated into a single image for better diagnosis and treatment of a patient. Image fusion is a process of combining useful or complementary information from multiple images into a single image. In this Project, we present a new weighted average fusion algorithm to fuse MRI and CT images of a brain based on guided image filters and the image statistics. The proposed algorithm is as follows: detail layers are extracted from each source image by using a guided image filter. Weights corresponding to each source image are calculated from the detail layers with help of image statistics. Then a weighted average fusion strategy is implemented to integrate source image information into a single image. Fusion performance is assessed both qualitatively and quantitatively. Proposed method is compared with the traditional and recent image fusion methods. Results showed that our algorithm yields superior performance.

I. Introduction:

With the rapid development of sensor and computer technology, medical imaging has emerged as an irreplaceable component in various clinical applications including diagnosis, treatment planning and surgical navigation. To provide medical practitioners sufficient information for clinical purposes, medical images obtained with multiple modalities are usually required, such as X-ray, computed tomography (CT), magnetic resonance (MR), positron emission tomography (PET), single photon emission computed tomography (SPECT), etc. Due to the difference in imaging mechanism, medical images with different modalities focus on different categories of organ/tissue information. For instance, the CT images are commonly used for the precise localization of dense structures like bones and implants, the MR images can provide excellent soft-tissue details with high-resolution anatomical information, while the functional information on blood flow and metabolic changes can be offered by PET and SPECT images but with low spatial resolution. Multi-modal medical image fusion aims at combining the complementary information contained in different source images by generating a composite image for visualisation, which can help physicians make easier and better decisions for various purposes [1].In recent years, a variety of medical image fusion methods have been proposed [2]–[17]. Due to the difference in imaging mechanism, the intensities of different source images at the same location often vary significantly. For this reason, most of these fusion algorithms are introduced in a multi-scale manner to pursue perceptually good results. In general, these Multi-Scale transform (MST)-based fusion methods consist of three steps, namely, decomposition, fusion and reconstruction. Multi-scale transforms which are frequently studied in image fusion include pyramids [17]–[19], wavelets [9], [20], [21], multi-scale geometrical transforms like contour let and shear let [2], [6], [10], [16]. In image fusion research, sparse representation is another popular image modelling approach, which has also been successfully applied to fuse multimodal medical images [4], [5], [15], [22]. One of the most crucial issues in image fusion is calculating a weight map which integrates the pixel activity information from different sources. In most existing fusion methods, this target is achieved by two steps known as absolute value of a decomposed co-efficient (or the sum of those values within a small window) is employed to measure its activity, and then a “choose-max” or “weighted-average” fusion rule is applied to assign weights to different sources based on the obtained measurement. Clearly, this kind of activity measurement and weight assignment are usually not very robust resulting from many factors like noise, Miss- registration and the difference between source pixel intensities. To improve the fusion performance, many complex decomposition approaches and elaborate weight assignment strategies have been recently proposed in the literature [6], [8]–[13], [15], [16].However, it is actually not an easy task to design a ideal activity level measurement or weight assignment strategy which can comprehensively take all the key issues of fusion into account. Moreover, these two steps are designed individually without a strong association by many fusion methods, which may greatly limit the algorithm performance. In this paper, this issue is addressed from another viewpoint to overcome the difficulty in designing robust activity level measurements and weight assignment strategies. Specifically, a Convolutional neural network (CNN) [23] is trained to encode a direct mapping from source images to the weight map. In this way, the activity level measurement and weight assignment can be jointly achieved in an “optimal” manner via learning network parameters. Considering the different imaging modalities of multi-modal medical images, we adopt a multi-scale approach via image pyramids to make the fusion process more consistent with human visual perception. In addition, a local similarity-based strategy is applied to adaptively adjust the fusion mode for the decomposed Coefficients’ of source images.

activity level measurement and weight assignment. In conventional transform domain fusion methods, the

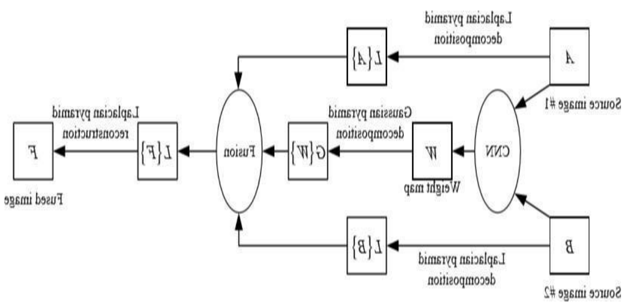
MAGNETIC RESONANCE IMAGING (MRI)-

MR imaging is a non-invasive imaging technology that produces three dimensional detailed anatomical images. It is often used for disease detection, diagnosis, and treatment monitoring. It is based on sophisticated technology that excites and detects the change in the direction of the rotational axis of protons found in the water that makes up living tissues.MRIs employ powerful magnets which produce a strong magnetic field that forces protons in the body to align with that field. When a radiofrequency current is then pulsed through the patient, the protons are stimulated, and spin out of equilibrium, straining against the pull of the magnetic field. When the radiofrequency field is turned off, the MRI sensors are able to detect the energy released as the protons realign with the magnetic field. The time it takes for the protons to realign with the magnetic field, as well as the amount of energy released changes depending on the environment and the chemical nature of the molecules. Physicians are able to tell the difference between various types of tissues based on these magnetic properties.To obtain an MRI image, a patient is placed inside a large magnet and must remain very still during the imaging process in order not to blur the image. Contrast agents (often containing the element Gadolinium) may be given to a patient intravenously before or during the MRI to increase the speed at which protons realign with the magnetic field. The faster the protons realign, the brighter the image.One kind of specialised MRI is functional Magnetic Resonance Imaging (FMRI.) This is used to observe brain structures and determine which areas of the brain “activate” (consume more oxygen) during various cognitive tasks. It is used to advance the understanding of brain organisation and offers a potential new standard for assessing neurological status and neurosurgical risk.Traditional MRI, unlike PET or SPECT, cannot measure metabolic rates. However, researchers funded by NIBIB have discovered a way to inject specialised compounds (hyperpolarized carbon 13) into prostate cancer patients to measure the metabolic rate of a tumour. This information can provide a fast and accurate picture of the tumour's aggressiveness. Monitoring disease progression can improve risk prediction, which is critical for prostate cancer patients who often adopt a wait and watch approach

II. II. Literature survey:

III.

In our recent work [24], a CNN-based multi-focus image fusion method which can obtain state-of-the-art results was proposed. In the method, two source images are fed to the two branches of a Siamese convolutional network in which the two branches share the same architecture and weights [25], respectively. Each branch contains three convolutional layers and the obtained feature maps essentially act as the role of activity level measures. The feature maps of two branches are concatenated and then pass through two fully-connected layers (they are converted into equivalent convolutional layers in the fusion process to allow arbitrary input size [26]), which can be viewed as the weight assignment part of a fusion method. As a result, the value of each co- efficient in the network output map indicates the focus property of a pair of source image patches at a corresponding location. By assigning the value as the weights of all the pixels within the patch location and then averaging the overlapped pixels, a focus map with the same size of source images is generated. The final fused image is obtained based on the focus map using the weighted-average rule along with two consistency verification techniques [20]. In [24], the feasibility and superiority of CNNs used for image fusion have been explicitly presented. Please refer to [24] or more details.The target of this paper is to extend he CNN model to medical image fusion. However, he method proposed in [24] cannot be directly used o fuse medical images primarily due to the ollowing two reasons.To address the first problem, we apply a pyramid-based multi-scale approach[27] o pursue perceptually better results. Specifically, ach source image is decomposed into a Laplacian pyramid while the weight map obtained from the network is decomposed into a Gaussian pyramid. The fusion procedure is conducted at every decomposition level.For the second issue, we adopt a local similarity-based fusion strategy to determine he fusion mode for the decomposed coefficients 18]. When the contents of source images have high similarity, the "weighted-average" fusion mode is applied to avoid losing useful information. In this situation, the weights obtained by the CNN are nore reliable than the coefficient based measure, so hey are employed as the merging weights. When he similarity of image contents is low, the "choose-nax" or "selection" fusion mode is preferred to preserve the salient details from source images. In his situation, the CNN output is not reliable, and he pixel activity is directly measured by the absolute values of the decomposed coefficients.Based on the above ideas, the CNN nodel presented in [24] can be applied to the fusion of medical images. It is worthwhile to note that



both the pyramid-based decomposition and the similarity-based fusion mode determination are just 'naive" techniques which are commonly used in the ield of image fusion. Nevertheless, it will be demonstrated that a reasonable usage of these

INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE
ISSN: 2057-5688

techniques incorporated with the CNN model can result in state-of-the-art fusion performance.

IV. V. III. Proposed methodology:

Fig shows the convolutional network used in the proposed fusion algorithm. It is a Siamese network in which the weights of the two branches are constrained to the same. Each branch consists of three convolutional layers and one max-pooling layer which is the same as the network used in [24]. To reduce the memory consumption as well as increase the computational efficiency, we adopt a much slighter model in this work by removing a fully connected layer from the network used in [24]. The 512 feature maps after concatenation are directly connected to a 2-dimensional vector. It can be calculated that the slight mode only takes up about 1.66 MB of physical memory in single precision, which is significantly less than the 33.6 MB model employed in [24]. Finally, this 2dimensional vector is fed to a 2-way Soft Max layer (not shown in Fig. 1), which produces a probability distribution over two classes. The two classes correspond to two kinds of normalised weight assignment results, namely, "First patch 1 and second patch 0" and "first patch 0 and second patch 1", respectively. The probability of each class indicates the possibility of each weight assignment. In this situation, also considering that the sum of two output probabilities is 1, the probability of each class just indicates the weight assigned to its corresponding input patch.

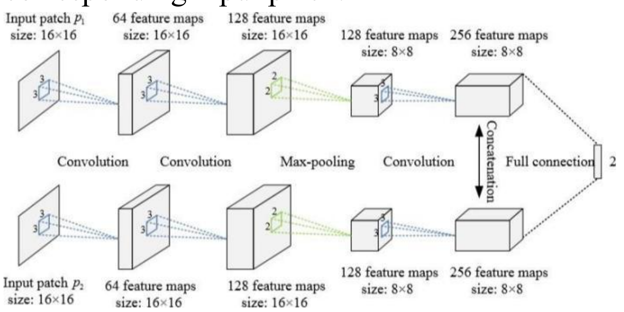


Fig 1: Architecture for CNN training

The network is trained by high-quality image patches and their blurred versions using the approach in [24]. In the training process, the spatial size of the input patch is set to 16 × 16 according to the analysis in [24]. The creation of training examples is based on multi-scale Gaussian filtering and random sampling. The SoftMax loss function is employed as the optimization objective and we adopt the stochastic gradient descent (SGD) algorithm to minimise it. The training process is operated on the popular deep learning framework Caffe [28]. Please refer to [24] for the details of example generation and network training.Since the network has a fully connected layer that has fixed dimensions (pre- defined) on input and output data,

the input of the network must have a fixed size to ensure that the input data of a fully connected layer is fixed. In image fusion, to handle source images of arbitrary size, one can divide the images into overlapping patches and input each patch pair into the network, but it will introduce many repeated calculations. To solve this problem, we first convert the fully connected layer into an equivalent convolutional layer containing two kernels of size 8 × 8 × 512 [26]. After the conversion, the network can process source images of arbitrary size to generate a dense prediction map, in which each prediction (a 2-dimensional vector) contains the relative clarity information of a source patch pair at the corresponding location. As there are only two dimensions in each prediction and their sum is normalised to 1, the output can be simplified as the weight of the first (or second) source. Finally, to obtain a weight map with the same size of source images, we assign the value as the weights of all the pixels within the patch location and average the overlapped pixels.

Only terms for which (x-m)/2 and (y-n)/2 are integers are included in the sum. Rather than encode $_{0,0}$ and $_{1}$ is encoded. This results in a net data compression because:$_0$is largely uncorrelated, and so may be represented pixel by pixel with many fewer bitsthan $_0$ $_1$islow pass filtered, and so may be encoded at a reduced sample rate. Further data

INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE

ISSN: 2057-5688

compression is achieved by iterating this process.By repeating these steps several times, a sequence of images $_0$, $_1$, $_2$, … , are obtained. If we now imagine these images stacked one above another, the result is a tapering pyramid data structure - hence the name. The Laplacian pyramid can thus be used to represent images as a series of band-pass filtered images, each sampled at successively sparser densities.

VI. IV. Simulation Results:

All the experiments have been done in MATLAB 2016b version under the high- speed CPU conditions for faster running time with test images shown in

Fig2:Proposedmedical imagefusion

Fig3:Testimagesdataset1(MR-Gad&CT)

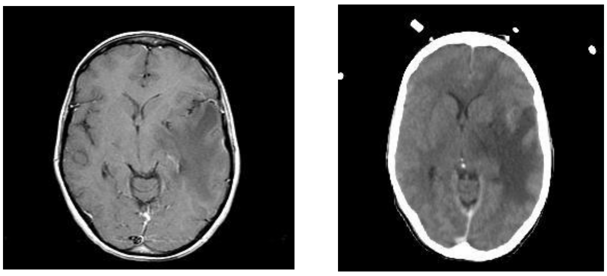
Aim of any fusion algorithm is to integrate required information from both source images in the output image. Fused images cannot be judged exclusively by seeing the output image or by measuring fusion metrics. It should be judged qualitatively using visual display and quantitatively using fusion metrics. In this section, we are presenting both visual quality and quantitative analysis of proposed and existing algorithms such as, Wavelet based methods discrete wavelet transform (DWT), stationary wavelet transform (SWT). Analysis of fusion metrics along with image quality assessment (IQA) metrics such as peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), correlation coefficient (CC), root mean square error (RMSE) and entropy (E) are considered to verify the effectiveness of the proposed algorithm. The objective of any fusion algorithm is to generate a qualitative fused image.For better quality, fused image should have optimal values for all these metrics. The fusion metric with best value is highlighted in bold letters.

thodology	NR dB)	ISE		n	ropy
'T	95	909	4	88	2
/T	69	93	44	8	1
posedmethod	**51**	**0047**			**7**

Table 1:Quantative analysis of fusion method for data set

V. Conclusion:

In this Project, a medical image fusion method based on convolutional neural networks is proposed. We employ a Siamese network to generate a direct mapping from source images to a weight map which contains the integrated pixel activity information. The main novelty of this approach is it can jointly implement activity level measurement



and weight assignment via network learning, which can overcome the difficulty of artificial design. To achieve perceptually good results, some popular techniques in image fusion such as multi-scale processing and adaptive fusion mode selection are appropriately adopted. Experimental results demonstrate that the proposed method can obtain high-quality results in terms of visual quality and objective metrics. Infi addition to the proposed algorithm itself, another contribution of this work is that it exhibits the great potential of some deep learning techniques for image fusion, which will be further studied in the future.

ISSN: 2057-5688

References:

[1]A. James and B. Dasarathy, "Medical image fusion: a survey of the state of the art," Information Fusion, vol. 19, pp. 4–19, 2014.

[2]L. Yang, B. Guo, and W. Ni, "Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform," Neurocomputing, vol. 72, pp. 203–211, 2008.

[3]Z. Wang and Y. Ma, "Medical image fusion using m-pcnn," Information Fusion, vol. 9, pp. 176–185, 2008.

[4]B. Yang and S. Li, "Pixel-level image fusion with simultaneous orthogonal matching pursuit," Information Fusion, vol. 13, pp. 10–19, 2012.

[5]S. Li, H. Yin, and L. Fang, "Group-sparse representation with dictionary learning for medical image denoising and fusion," IEEE Transactions on Biomedical Engineering, vol. 59, pp. 3450–3459, 2012.

[6]Z. L. G. Bhatnagar, Q. Wu, "Directive contrast based multimodal medical image fusion in nsct domain," IEEE Transactions on Multimedia, vol. 15, pp. 1014–1024, 2013

[7]S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," IEEE Transactions on Image Processing, vol. 22, no. 7, pp. 2864–2875, 2013.

[8]R. Shen, I. Cheng, and A. Basu, "Cross-scale coefficient selection for volumetric medical image fusion," IEEE Transactions on Biomedical Engineering, vol. 60, pp. 1069–1079, 2013.

[9]R. Singh and A. Khare, "Fusion of multimodal medical images using daubechies complex wavelet transform c a multiresolution approach," Information Fusion, vol. 19, pp. 49–60, 2014.

[10]L. Wang, B. Li, and L. Tan, "Multimodal medical volumetric data fusion using 3-d discrete shearlet transform and global-to-local rule," IEEE Transactions on Biomedical Engineering, vol. 61, pp. 197–206, 2014.

Z