STRUCTURAL SIMILARITY BASED ANATOMICAL AND FUNCTIONAL BRAIN IMAGING FUSION

Nishant Kumar[1], Nico Hoffmann, Martin Oelschlägel, Edmund Koch, Matthias Kirsch, Stefan Gumhold[1]

[1]Computer Graphics and Visualisation, Technische Universtität Dresden, Germany

Abstract

- Multimodal medical image fusion helps in combining contrasting features from two or more input imaging modalities to represent fused information in a single image.
- In this paper, we present a novel end-to-end unsupervised learning based Convolutional neural network (CNN) for fusing MRI-PET image pairs.
- We exploit Structural Similarity Index (SSIM) as the loss function and then apply color coding for the visualization of the fused image by quantifying the contribution of each input image in terms of the partial derivatives of the fused image.
- We find that our fusion and visualization approach results in better visual perception of the fused image, while also comparing favorably to previous methods when applying various quantitative assessment metrics.

Introduction

- The integrated MRI-PET scanners results in high tissue contrast with significantly low radiation dose. But the development of a robust hybrid MRI-PET hardware is challenging due to compatibility issue of PET detectors in a high magnetic field environment of MRI.
- The post-hoc fusion of MRI-PET image pairs overcomes the challenges of fully integrated MRI-PET scanners and helps medical personnel to better diagnose brain abnormalities such as glioma and Alzheimer's disease.
- The past image fusion methods proposed a three step approach to the fusion problem. First, the source images were transformed using methods such as multi-scale decomposition. The transformed coefficients are combined using fusion strategy such as max selection and weighted-averaging. Finally, the fused image is reconstructed by taking the inverse of the transformation.
- Since, this approach is computationally inefficient, we propose a fast real time medical image fusion approach in an end-to-end unsupervised learning network trained on publicly available medical image pairs. Additionally, the fusion result is visualized based on the contribution of the input images to the fused output image.

Data acquisition

- Our training data consists of 500 MRI-PET image pairs publicly available at the Alzheimer's Disease Neuroimaging Initiative (ADNI). The MRI images were skull stripped T1 weighted N3m MPRAGE sequences while PET-FDG images were co-registered, averaged, standardized voxel sized with uniform resolution of the same subject. We aligned the MRI-PET image pairs using the Affine transformation tool of 3D Slicer registration library.
- Our testing data contain 90 MR-T1 and PET-FDG image pairs of 90 unique subjects from ADNI and 10 MR-T1 and PET-FDG image pairs from Harvard Whole Brain Atlas database2.
- Our approach has been implemented in Python 2.7 and Tensorflow 1.10.1 on a Linux Ubuntu 17.10 x86 64 system with 12 Intel Core i7-8700K CPU @ 3.70 GHz and 64-GB RAM

Method – Fusion framework

- The fusion architecture in Fig.1. takes two grayscale input images and generates a grayscale fused image. The network consist of feature extraction, fusion and reconstruction strategies to preserve most of the details from the input modalities. We train the parameters of the feature extraction and reconstruction layers by maximizing the structural similarity and minimizing the euclidean distance between fused image and the input images.

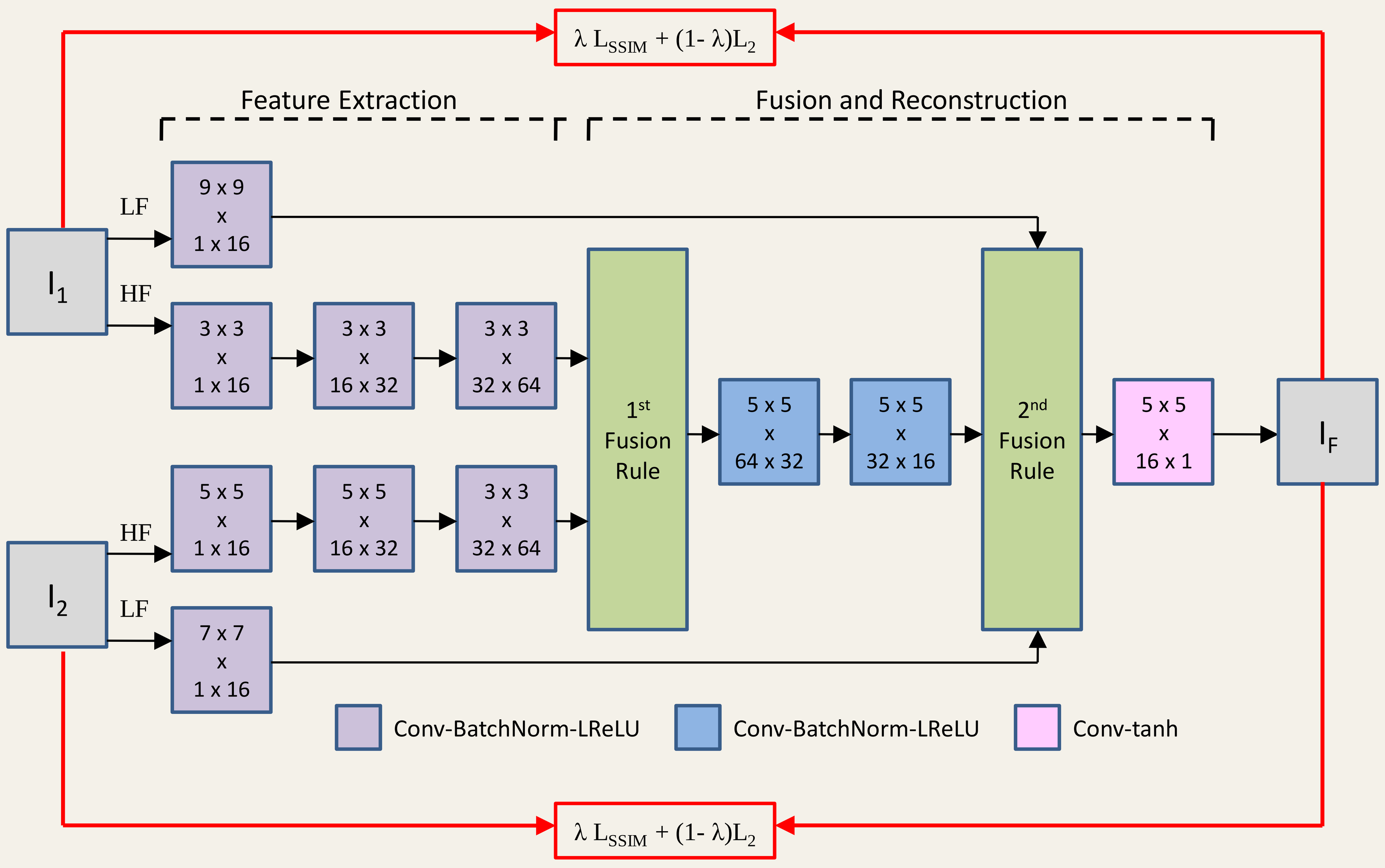

Figure 1. The proposed unsupervised learning based MRI-PET fusion framework.

$$Total\ Loss\ \ L_{total} = \lambda L_{ssim} + (1-\lambda)L_{L2}$$

$$L_{ssim} = 1 - SSIM(I_1, F)\ +\ 1 - SSIM(I_2, F) \qquad L_{L2} = ||F - I_1||_2 + ||F - I_2||_2$$

Method – Visualisation framework

- We visualized the functional and anatomical information in the fused grayscale image by calculating the partial derivative of each pixel of the fused image with respect to input images as shown in Fig.2. Assuming n*m as the image dimension, then the gradient ∇_{FI} of fused image with respect to the input image is given by:

$$\nabla_{FI}(n, m) = \sum_{i=0}^{k} \sum_{j=0}^{l} \frac{\partial F(i,j)}{\partial I(n,m)}$$

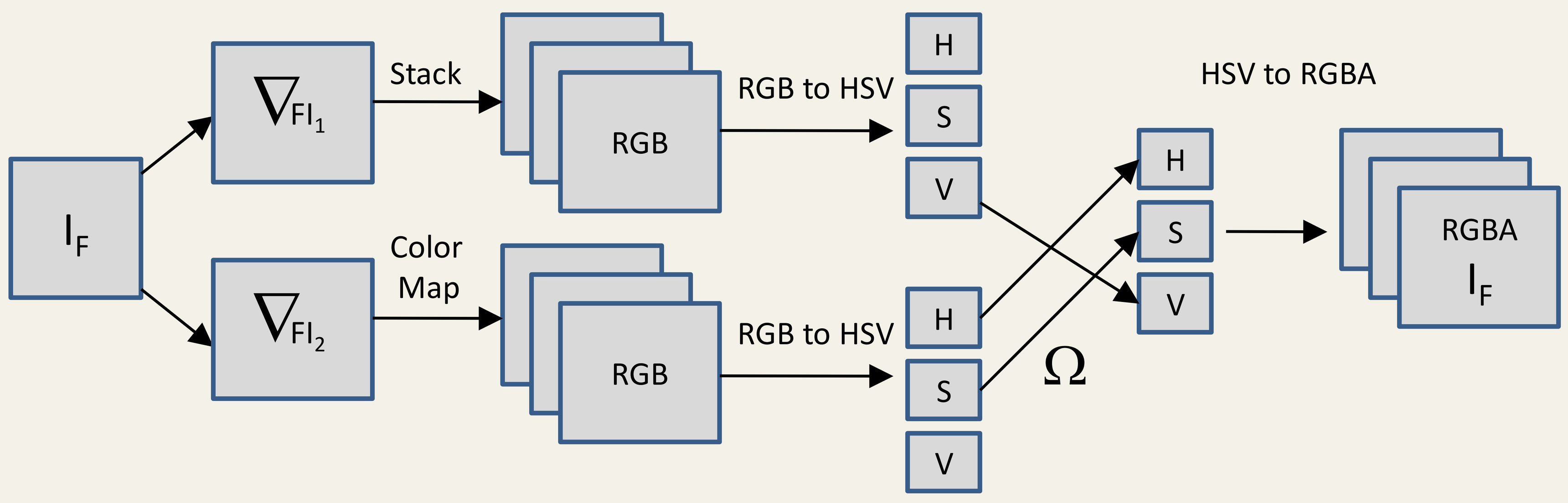

Figure 2. The proposed MRI-PET visualisation framework.

Fusion and Visualisation results

- The evaluated methods shows loss in edge and texture while our method conserve structural information better. The luminance of the proposed fusion results increases with greater λ values leading to brightness artifacts at corner cases of λ = 0 and λ = 1. The visualisation results at λ = 0.8 controlled by parameter Ω shows a shift in occlusion of the anatomical information with different values of Ω.

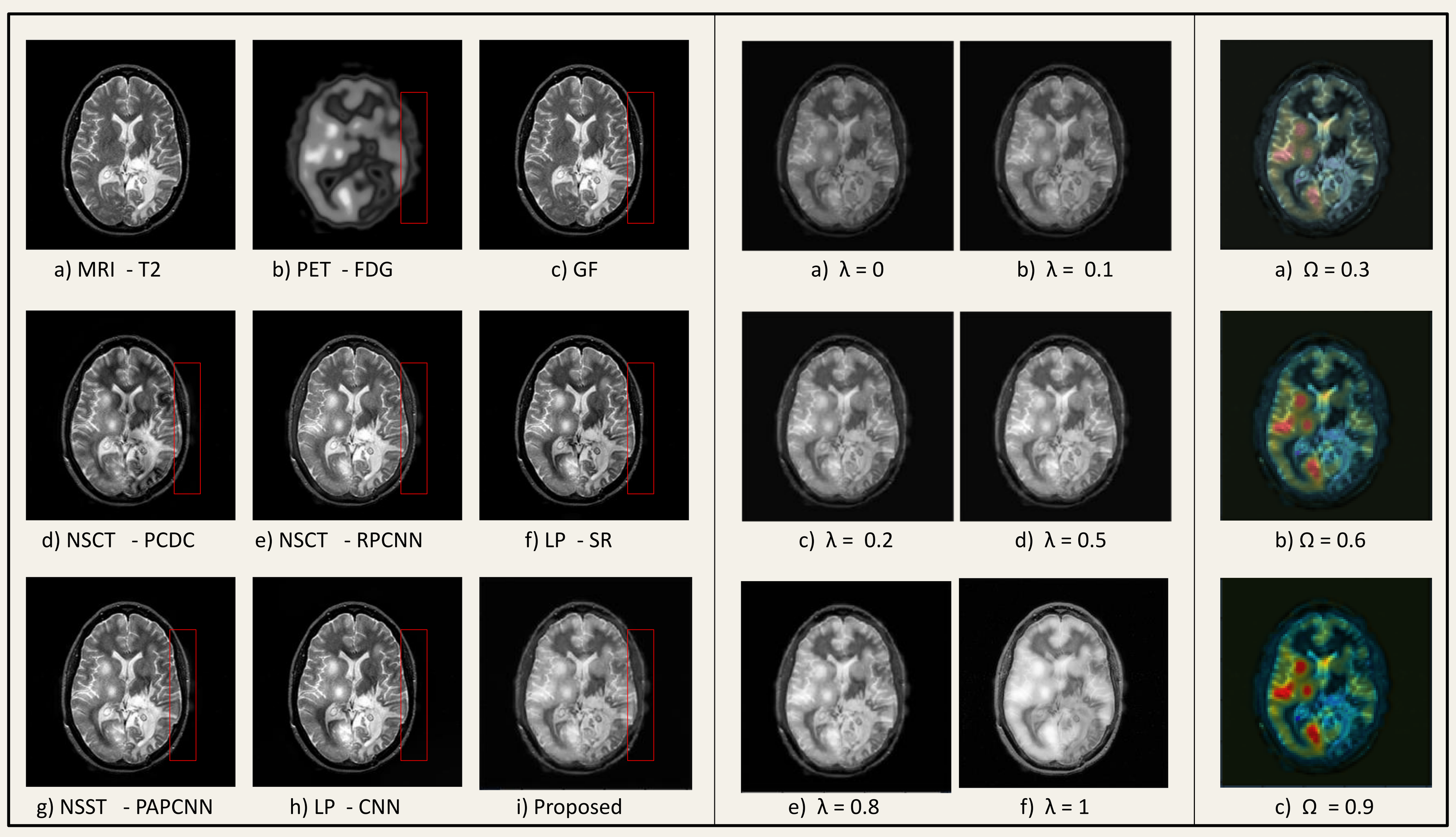

Figure 3. The visual results of our proposed framework.

Loss Curve Analysis

- The loss curves L_{SSIM} and L_2 convey rapid convergence for all λ values other than λ ≥0.9 where L_{SSIM} plays more important role than L_2
- L_{SSIM} has higher sensitivity to luminance variations in flat texture-less regions while L_2 is more sensitive to larger errors irrespective of the underlying regions within the image. This property leads to delayed convergence of L_{SSIM} for visually perceptive results at edges as well as flat regions of the fused image.

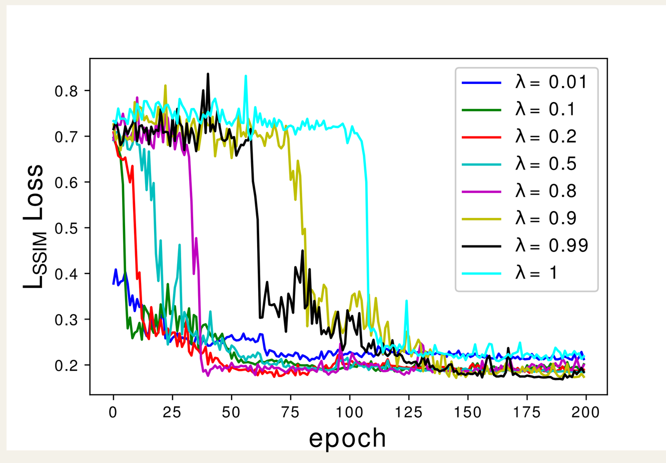

Figure 4. The L_{SSIM} loss curve.

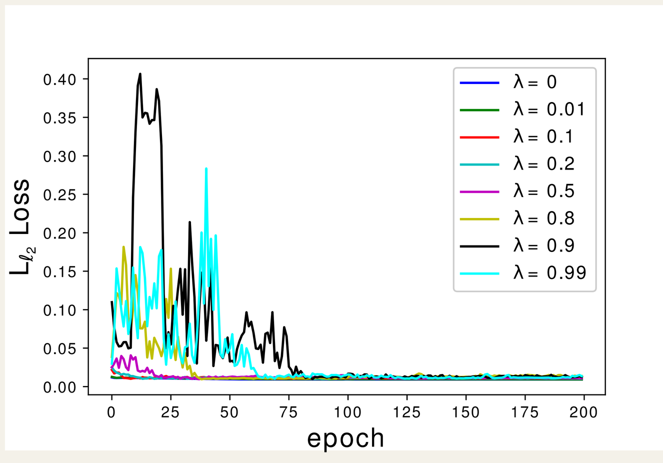

Figure 5. The L_{L2} loss curve.

Objective assessment

- Table 1. summarizes the average scores of 100 test image pairs computed for different fusion methods along with our proposed method at λ = 0.8 and Ω = 0.6. A method with a higher score performs better, a rule applicable for all the mentioned metrics.
- The results convey that our method performs better with the quality metric Q_{SSIM} and Q_{VIFF} since neural network optimizes a human perceptive loss function. Overall, the objective scores reflects the robustness of our method for medical image fusion.

Metrics	GF	NSCT-PCDC	LP-SR	NSCT-RPCNN	NSST-PAPCNN	LP-CNN	Proposed
Q_{IE}	0.8169	0.8080	0.8092	0.8132	0.8102	0.8076	0.8104
Q_G	0.7555	0.5457	0.6501	0.6702	0.6685	0.5665	0.5707
Q_{FMI}	0.9224	0.8754	0.8969	0.8941	0.8997	0.8958	0.8885
Q_{SSIM}	0.8260	0.7992	0.7837	0.8492	0.8318	0.7176	0.8610
Q_{VIFF}	0.2776	0.3415	0.5990	0.5430	0.6001	0.5326	0.6005
Time (s)	13.43	221.07	75.69	775.31	521.36	481.73	0.37

Table 1. The objective assessment results

Conclusion

- The proposed end-to-end learning based fusion model construct artifact free fusion images and the gradient based visualisation delineated the anatomical features of MRI from the functional features of PET in the fused image. The extensive evaluation of our approach conveyed significant improvements in human perceptive results compared to past methods.
- In future, our method could further be extended to include other combination of anatomical and functional imaging modalities by changing the fusion architecture especially the feature extraction layers. Additionally, we plan to immersively visualize the proposed results in an augmented reality based real time preoperative setup, thereby enabling medical experts to make robust clinical decisions.

Shen Zhen MICCAI 2019
Europe funds Saxony. ESF European Social Fund. European Union