XmoNet: a Fully Convolutional Network for Cross-Modality MR Image Inference

Sophia Bano[1], Muhammad Asad[2], Ahmed E. Fetit[3], and Islem Rekik[4]

[1] Wellcome/EPSRC Centre for Interventional and Surgical Sciences (University College London), [2] Imagination Technologies,
[3] Biomedical Image Analysis Group (Imperial College London), [4] BASIRA Lab, Computer Vision & Image Processing Group (University of Dundee), United Kingdom

UCL

1. Introduction

Magnetic Resonance (MR) modalities, e.g. T1-, T2-, and diffusion-weighted imaging, offer complementary contrast mechanisms for capturing visual characteristics of tissues of interest. In research settings and imaging clinical trials, factors such as cost, availability of scanning time, and patient discomfort impede the acquisition of multimodal scans. Cross-modality generation (synthesising a target modality given a separate source) may provide solutions to such challenges.

2. Architecture

We propose **XmoNet**, a deep learning architecture inspired by fully convolutional networks (FCNs) **[1, 3]**. XmoNet utlilises parallel pathways to encode low- and high-frequency visual features, thereby capturing rich hierarchies of feature representations. We present a preliminary study illustrating the utility of XmoNet in learning the mapping between heterogeneous T1- and T2-weighted MR scans rapidly and accurately.

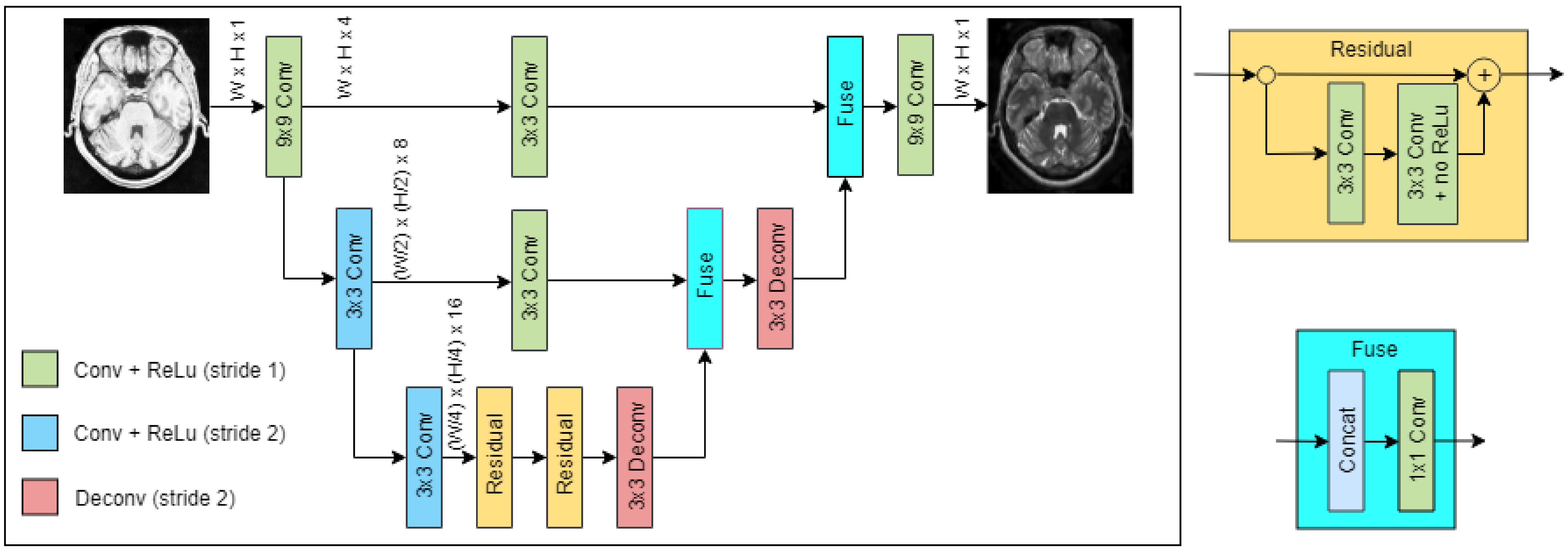

Fig 1. The proposed XmoNet deep learning architecture.

3. Experimental Analysis

Dataset

We used the publicly available MNI-HISUB25 dataset **[2]** of MR brain scans of 25 subjects. The data is heterogeneous as a few T2-weighted scans show missing regions.

Observations

XmoNet resulted in accurate and sharp MR image generation: (a) synthesised images better capture overall brain structures w.r.t source images. (b) regions with missing T2 signal successfully synthesised.

Limitations

The analysis requires a baseline, a thorough assessment by medical experts, and studying generalisation of the network under different acquisition settings.

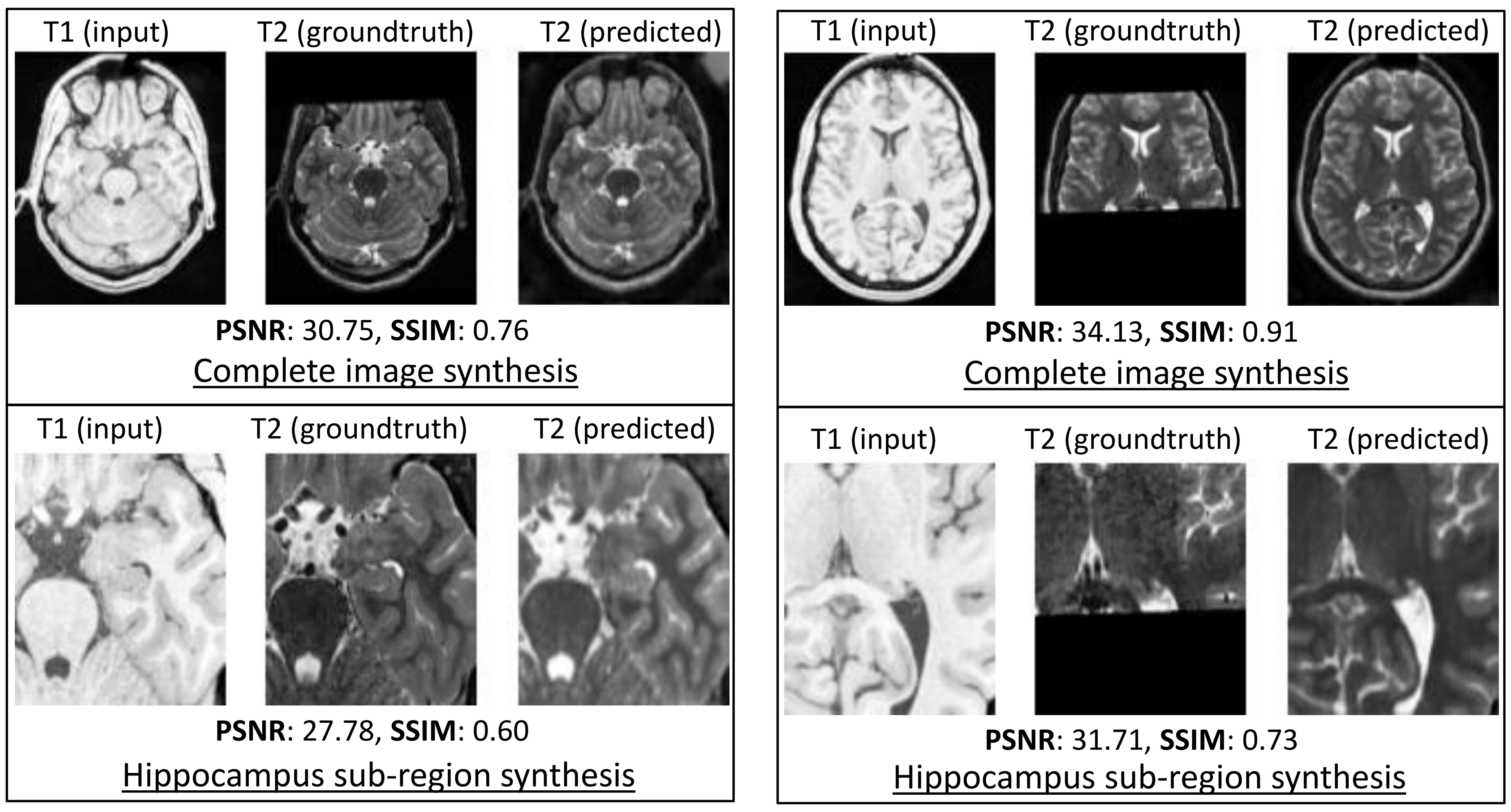

Fig 2. Experimental results for predicting T2-w from T1-w image. Scans with missing frontal lobe part in T2-w image (left). Section with missing frontal and parietal lobe parts in T2-w image (right).

4. Conclusion and Future Work

- We designed a deep learning network for the cross-modality MR image inference, utilising a fully convolutional architecture. Preliminary analysis illustrated the utility of XmoNet for accurate and realistic synthesis of T2-weighted images from source T1-weighted data.
- Future work involves validating XmoNet on larger datasets, comparison with the existing methods, adaptation to incorporate non-MR based modalities and extension to 3D image inference.

References

[1] Badrinarayanan, V., Kendall, A., Cipolla, R.: SegNet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE TPAMI, 2017.
[2] Kulaga-Yoskovitz, J., Bernhardt, B.C., Hong, S.J., Mansi, T., Liang, K.E., *et al.*: Multi-contrast submillimetric 3 tesla hippocampal subfield segmentation protocol and dataset. Scientic data 2, 2015.
[3] Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. IEEE CVPR, 2015.

wellcome EPSRC centre interventional+surgical sciences · Imagination · BioMedIA Imperial College London · University of Dundee · BASIRA Brain And SIgnal Research & Analysis Lab