# Isointense infant brain MRI segmentation with a dilated convolutional neural network

**Document status and date:**
Published: 09/08/2017

**Document Version:**
Author's version before peer-review

**Please check the document version of this publication:**

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

# Isointense infant brain MRI segmentation with a dilated convolutional neural network

Pim Moeskops and Josien P.W. Pluim

Medical Image Analysis Group, Eindhoven University of Technology, The Netherlands

**Abstract.** Quantitative analysis of brain MRI at the age of 6 months is difficult because of the limited contrast between white matter and gray matter. In this study, we use a dilated triplanar convolutional neural network in combination with a non-dilated 3D convolutional neural network for the segmentation of white matter, gray matter and cerebrospinal fluid in infant brain MR images, as provided by the MICCAI grand challenge on 6-month infant brain MRI segmentation.

## 1 Introduction

Brain MRI is important for quantitative analysis of neurodevelopment at all ages. A particularly difficult age in the context of brain MRI analysis is the age of 6 months, where the contrast between white matter and gray matter is limited because of the development of myelination [6,9].

Convolutional neural networks (CNNs) are an effective approach for infant brain MRI segmentation [9,3,4]. A specific type of CNNs proposed for image segmentation are dilated CNNs, which can achieve a large receptive field using a limited number of trainable weights [8]. In medical imaging, dilated CNNs have been used for cardiac MRI segmentation [7].

In this study, we investigate using a dilated triplanar CNN in combination with a non-dilated 3D CNN for the segmentation of brain MRI of infants at an age of 6 months.

## 2 Materials and Methods

### 2.1 Data

The data used in this study is provided by the MICCAI grand challenge on 6-month infant brain MRI segmentation[1]. 10 images are provided as training data, 13 images are used to evaluate the method in the challenge. For all patients, $T_1$- and $T_2$-weighted images are provided.

To allow evaluation of the performance in initial experiments, 3 training images were randomly selected as validation set, the remaining 7 images are used to train the method.

---

[1] This paper is a submission to the MICCAI grand challenge on 6-month infant brain MRI segmentation (`http://iseg2017.web.unc.edu/`).

The images were preprocessed by the organisers of the challenge, this included skull stripping, intensity inhomogeneity correction and removal of the cerebellum and brain stem. As a reference standard, the images were manually segmented into white matter (WM), gray matter (GM) and cerebrospinal fluid (CSF).

## 2.2 Method

A CNN with triplanar dilated convolutions and 3D convolutions in four network branches is used to segment the images into WM, GM and CSF (Table 1).

The dilated branches consist of the same architecture as proposed by Yu et al. [8], which uses layers of $3 \times 3$ kernels with increasing dilation factors. Input for the dilated branches is obtained from the axial, sagittal and coronal planes, as also proposed by Wolterink et al. [7]. In contrast, our network uses three network branches that are combined in the last layer, instead of processing in three directions and averaging of the output probabilities. Moreover, in addition to these three dilated triplanar network branches, we include an additional 3D network with non-dilated 3D convolutions. All four network branches use 2-channel input from the $T_1$- and the $T_2$-weighted images.

| Axial | | | Coronal | | | Sagittal | | | 3D | |
|---|---|---|---|---|---|---|---|---|---|---|
| Kernels | Size | Dilation | Kernels | Size | Dilation | Kernels | Size | Dilation | Kernels | Size |
| 32 | 3×3 | 1 | 32 | 3×3 | 1 | 32 | 3×3 | 1 | 32 | 3×3×3 |
| 32 | 3×3 | 1 | 32 | 3×3 | 1 | 32 | 3×3 | 1 | 32 | 3×3×3 |
| 32 | 3×3 | 2 | 32 | 3×3 | 2 | 32 | 3×3 | 2 | 32 | 3×3×3 |
| 32 | 3×3 | 4 | 32 | 3×3 | 4 | 32 | 3×3 | 4 | 32 | 3×3×3 |
| 32 | 3×3 | 8 | 32 | 3×3 | 8 | 32 | 3×3 | 8 | 32 | 3×3×3 |
| 32 | 3×3 | 16 | 32 | 3×3 | 16 | 32 | 3×3 | 16 | 32 | 3×3×3 |
| 32 | 3×3 | 1 | 32 | 3×3 | 1 | 32 | 3×3 | 1 | 32 | 3×3×3 |
| | | | | | | | | | 32 | 3×3×3 |
| | | | | | | | | | 32 | 3×3×3 |
| | | | | | | | | | 32 | 3×3×3 |
| | | | | | | | | | 32 | 3×3×3 |
| | | | | | | | | | 32 | 3×3×3 |
| Concatenation of $4 \times 32 = 128$ feature maps | | | | | | | | | | |
| | | | | | | | | | 3 | 1×1×1 |

Table 1: Combined 7-layer dilated triplanar CNN and 12-layer 3D CNN architecture. The dilated branches have receptive fields of $67 \times 67$ and the 3D branch has a receptive field of $25 \times 25 \times 25$. Because the background class is already defined by skull stripping, three output classes are used in the last layer (WM, GM and CSF).

Three versions of the network are evaluated in this study: (1) the dilated triplanar network with with shared weights between the three branches, (2) the dilated triplanar network with separate weights for each of the three branches, and (3) the dilated triplanar network with separate weights combined with the 3D network as fourth branch.

Batch normalisation [1] and ReLUs were used throughout. Dropout [5] was used before the output layer. The network is trained with Adam [2] based on the cross-entropy loss, using mini-batches of 200 or 300 samples in 10 epochs of 50,000 random samples per class per training image.

The fully convolutional nature of all four branches allows arbitrarily sized inputs during testing. The method takes about 1 minute to segment a full image on a NVIDIA Titan X Pascal GPU.
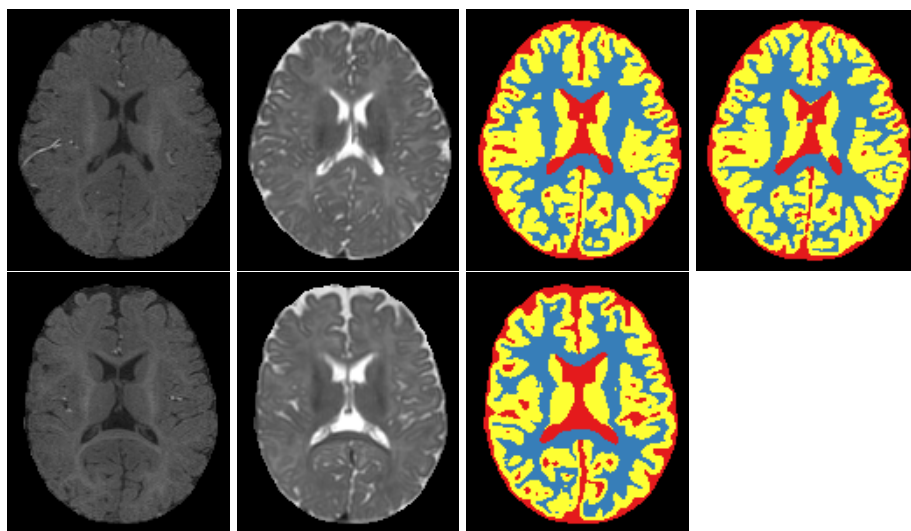


(a) $T_1$-w image    (b) $T_2$-w image    (c) Automatic    (d) Reference

Fig. 1: Example segmentations from the validation set (top), trained with 7 of the training images, and from the test set (bottom), trained with all 10 training images. From left to right: $T_1$-weighted image (a), $T_2$-weighted image (b), automatic segmenation (c), reference segmentation (d).

| **Triplanar network with shared weights** | | | |
| --- | --- | --- | --- |
| | WM | GM | CSF |
| Validation image 1 | 0.859 | 0.874 | 0.914 |
| Validation image 2 | 0.818 | 0.795 | 0.905 |
| Validation image 3 | 0.891 | 0.870 | 0.897 |
| Average | 0.856 | 0.846 | 0.905 |
| **Triplanar network with separate weights** | | | |
| | WM | GM | CSF |
| Validation image 1 | 0.863 | 0.889 | 0.927 |
| Validation image 2 | 0.860 | 0.858 | 0.903 |
| Validation image 3 | 0.881 | 0.866 | 0.895 |
| Average | 0.868 | 0.871 | 0.908 |
| **Combined triplanar and 3D network** | | | |
| | WM | GM | CSF |
| Validation image 1 | 0.885 | 0.909 | 0.948 |
| Validation image 2 | 0.830 | 0.821 | 0.921 |
| Validation image 3 | 0.905 | 0.900 | 0.925 |
| Average | 0.874 | 0.877 | 0.932 |

Table 2: Dice coefficients for the validation set, trained with 7 of the training images and evaluated on the remaining 3 images. From top to bottom: orthogonal triplanar dilated network with shared weights, orthogonal triplanar dilated network with separate weights, and orthogonal triplanar dilated network with separate weights and additional non-dilated 3D network.

## 3   Results and Discussion

Example segmentation results for the validation and test set are shown in Figure 1. The results on the validation set in terms of Dice coefficients are listed in Table 2.

The results show accurate segmentation performance in 6-month isointense infant brain MR images. On the validation set, the performance of the triplanar network with separate weights was better than the performance of the network with shared weights. The best performance on the validation set was obtained using the network that combines the dilated triplanar network with a non-dilated 3D network (Table 1). This combined network is therefore applied to the test set. The results on the test set will be available on the website of the MICCAI grand challenge on 6-month infant brain MRI segmentation[2].

## References

1. Ioffe, S., Szegedy, C.: Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: ICML (2015)
2. Kingma, D., Ba, J.: Adam: A method for stochastic optimization. In: ICLR (2015)
3. Moeskops, P., Viergever, M.A., Mendrik, A.M., de Vries, L.S., Benders, M.J., Išgum, I.: Automatic segmentation of MR brain images with a convolutional neural network. IEEE T Med Imaging 35(5), 1252–1261 (2016)
4. Nie, D., Wang, L., Gao, Y., Sken, D.: Fully convolutional networks for multi-modality isointense infant brain image segmentation. In: ISBI. pp. 1342–1345. IEEE (2016)
5. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: A simple way to prevent neural networks from overfitting. J Mach Learn Res 15(1), 1929–1958 (2014)
6. Wang, L., Gao, Y., Shi, F., Li, G., Gilmore, J.H., Lin, W., Shen, D.: LINKS: Learning-based multi-source integration framework for segmentation of infant brain images. NeuroImage 108, 160–172 (2015)
7. Wolterink, J.M., Leiner, T., Viergever, M.A., Išgum, I.: Dilated convolutional neural networks for cardiovascular MR segmentation in congenital heart disease. In: MICCAI HVSMR workshop. pp. 95–102. Springer (2016)
8. Yu, F., Koltun, V.: Multi-scale context aggregation by dilated convolutions. In: ICLR (2016)
9. Zhang, W., Li, R., Deng, H., Wang, L., Lin, W., Ji, S., Shen, D.: Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. NeuroImage 108, 214–224 (2015)

---

[2] `http://iseg2017.web.unc.edu/results/`