

AUTOMATIC SEGMENTATION OF MICROBLEEDS IN 3D MRI

Groothuis I., Sudre C., Pai A., Sørensen L., Nielsen M., Barkhof F., Ourselin S., Cardoso M. J., Modat M.



Abstract. Microbleed segmentation is a challenging task due to skewed foreground-background ratio in the dataset. Training a convolutional neural network with an selective sampling scheme overcomes this difficulty, but we still observe a high number of false positive per image. Various methods are explored to reduce the number of FPs per image, but this does seem to have a negative impact on the amount of true positives as well.

Introduction

Cerebral vascular microbleeds (CMB) are small haemorrhages that occur in healthy ageing people as well as individuals with dementia. Manually segmenting them on MRI images is a time consuming task and the results vary greatly between clinicians. By developing a tool based on convolutional neural networks (CNN) we hope to tackle the issues mentioned above.

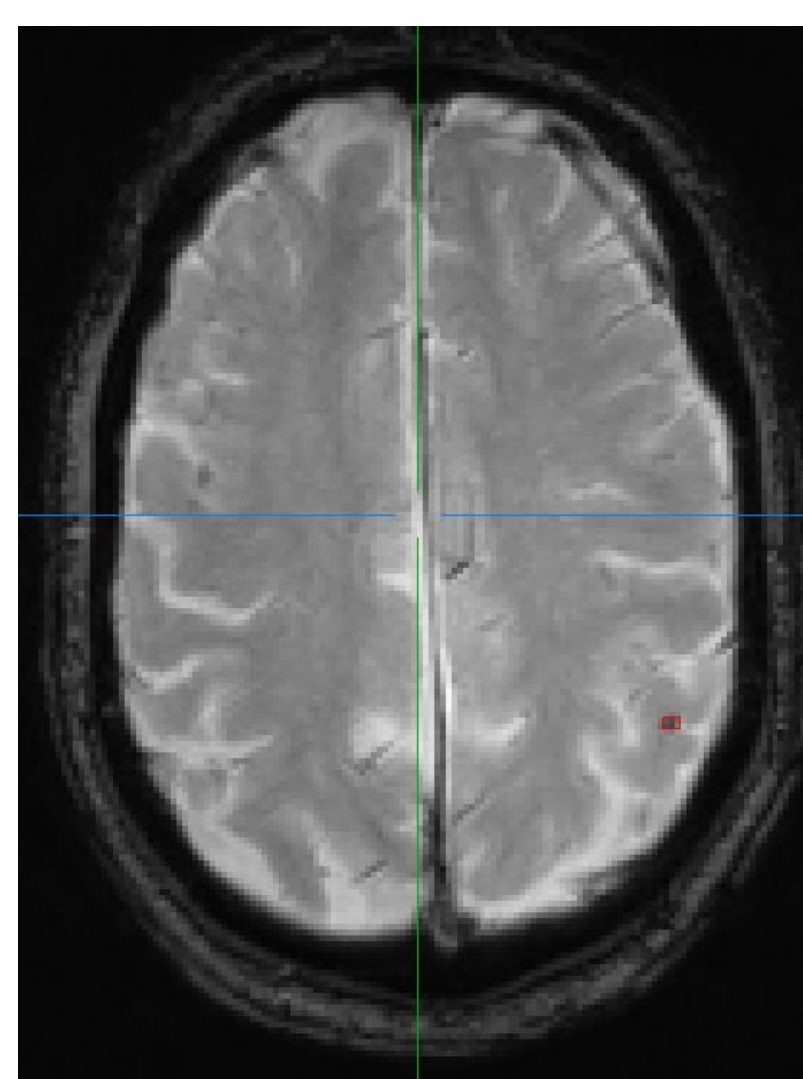


Fig 1. A microbleed (circled in red) in a T2S MRI images

Challenges

The data used is the ADNI dataset. A subset of patients with microbleeds was selected, such that in total we have a dataset of 496 microbleeds. Not each patient has the same amount of microbleeds, with the majority of the participants having one microbleeds. This makes the data very unbalanced, as per patch of 48 x 48 x 48 1 mm isotropic voxel with a microbleed will have a ratio of 0.2% foreground vs background.

Methodology

Deep learning is a subfield of machine learning. It focuses on 'deep' neural networks. Neural networks are a computational tool based on the neurons in the brain. One neuron on its own does not have a lot of processing power, but more together in layers are able to make linear decisions. Stacking these layers increases the computational and expressive powers of these networks. If many layers are stacked together a network is considered to be deep. CNNs are a type of neural networks that employ special filter for feature detections, which makes them excellent to use in (medical) image processing.

The neural network were implemented using NiftyNet [1], a deep learning tool for medical imaging. The current set up of the network consists of a series of residual block, and in each block the relu activation function is applied, after which batch normalisation is applied and a dilated convolution, which is repeated twice. This setup is chosen due to its excellent performance on brain segmentation, although it has been slightly altered. The network has been trained for 10 000 iterations, with a decaying learning rate of 0.01 with 0.001 decay. The images were sampled into patches of 48x 48 x 48, with a minimum 0.001 foreground/background ratio. The generalised Dice Score is used as a loss function, as it is robust to imbalanced data. [2]

Intermediate results

After training the network for 100 epoch using a selective sampling method based on lesion location, a high number of False Positive predictions per image was observed. This is caused by the selective sampling method, as sampling restricts the training set to contain mostly positive samples, and thus patches containing no microbleeds but having a similar appearance will be misclassified.

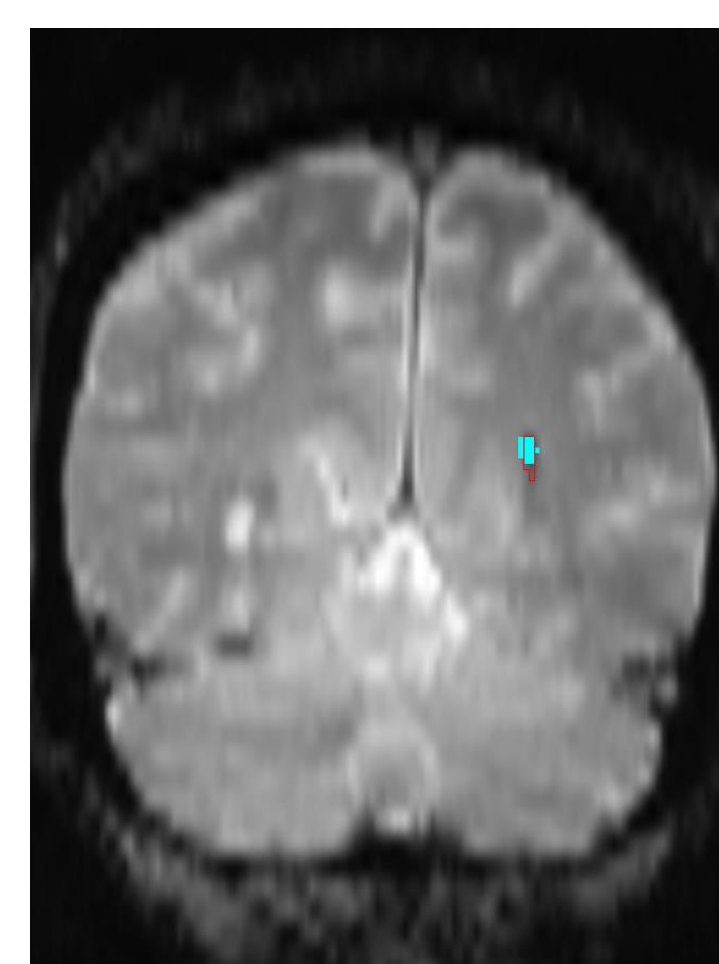
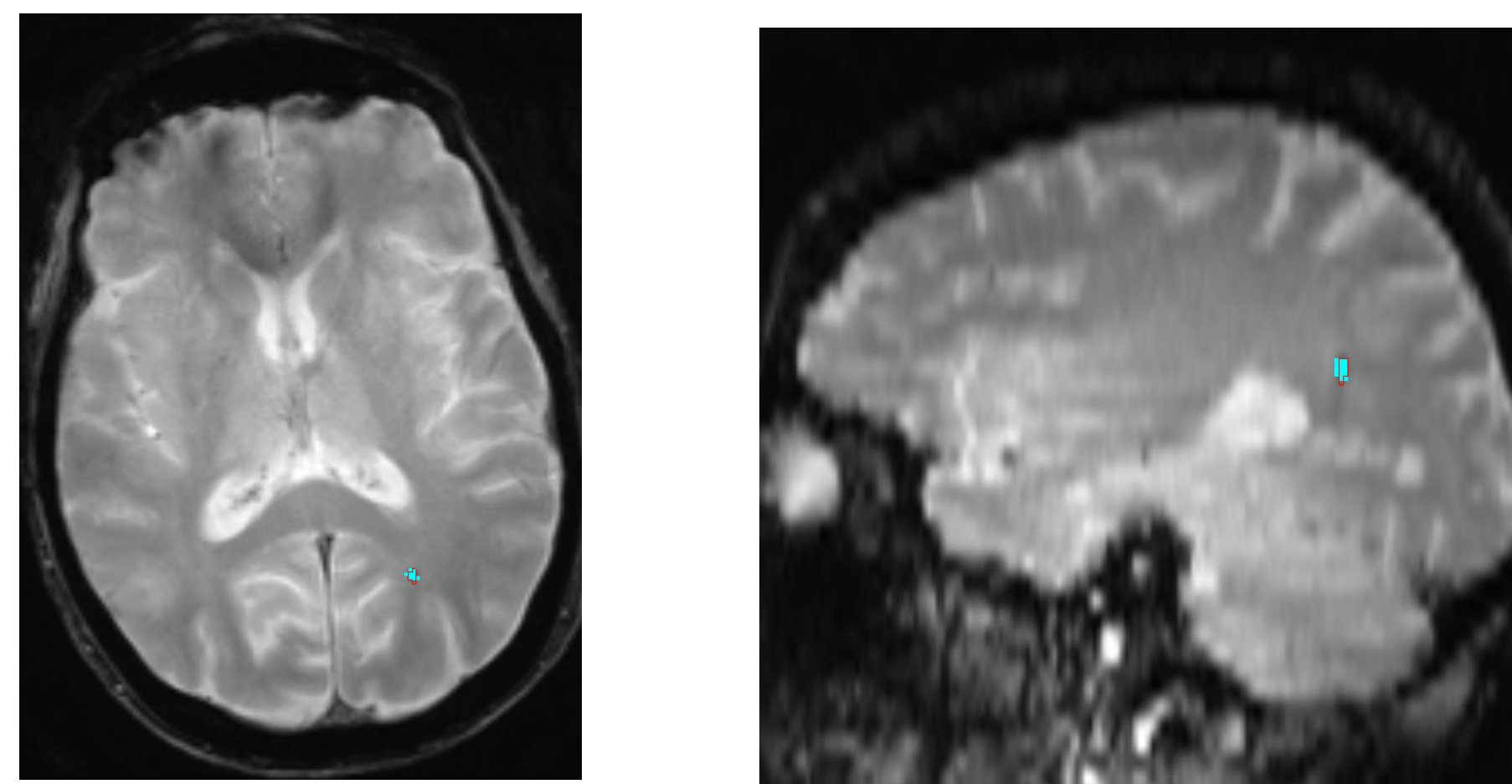


Fig 2 a, b and c. (Axial, coronal and sagittal view respectively) The above figures present a correct segmentation of the microbleeds, and a close up of the segmentation. The red outline is the segmentation used as ground truth, the blue the results obtained by the network

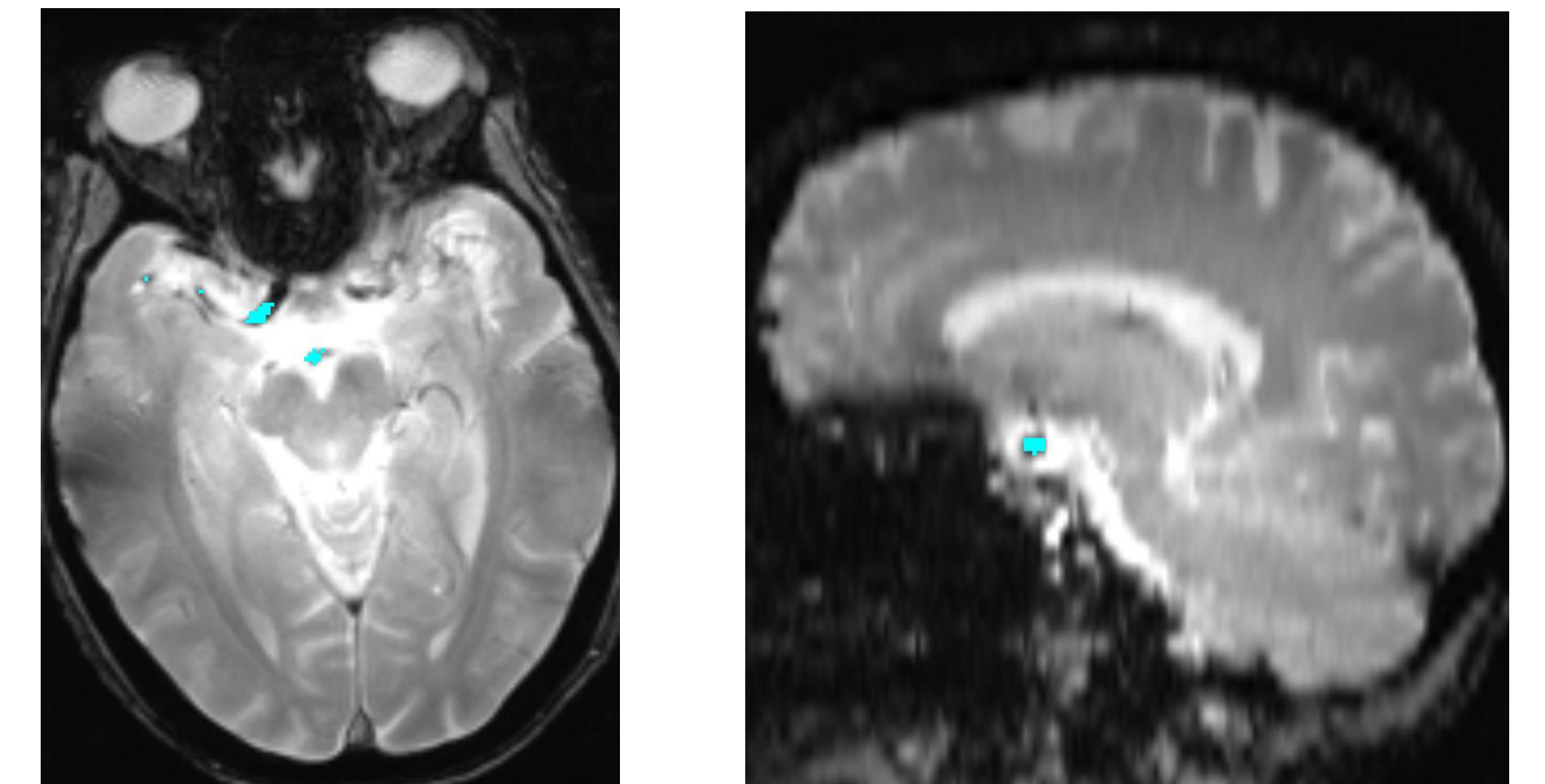
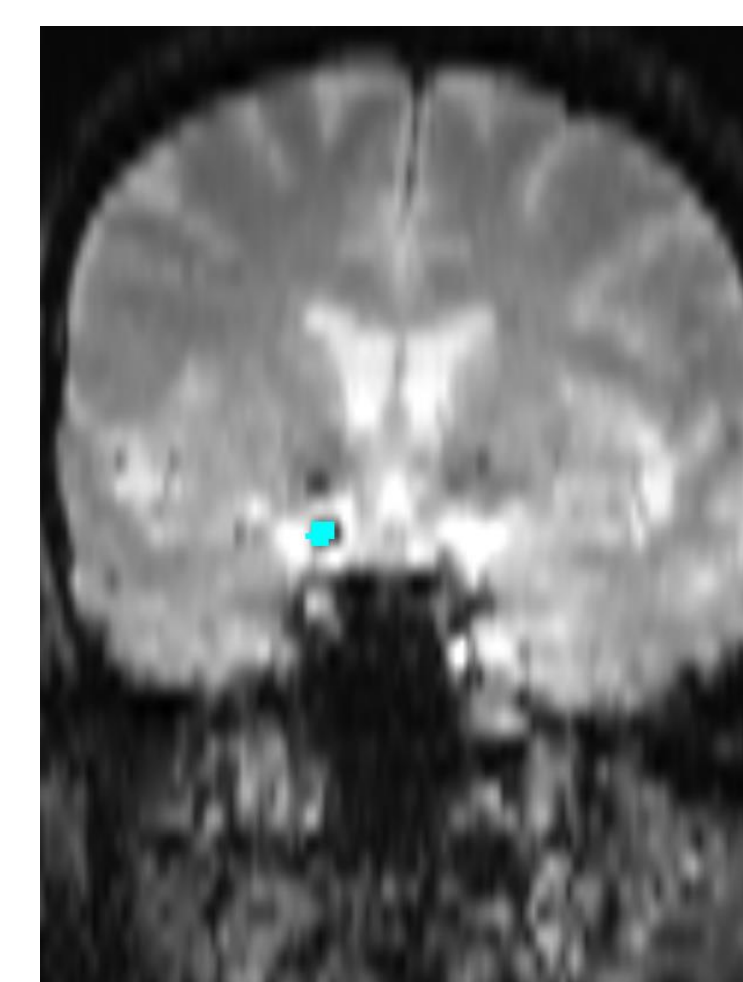


Fig 3 a and b. (Axial and sagittal view) The two figures above display the incorrect segmentation results. The network in labelling (in blue) vessels in the white matter as microbleeds. (Note the absence of the red line indicating microbleeds.)



Discussion

Looking at the results from the segmentation there are still many false positives. These seem to be caused by microbleeds being less hyperintense than other nearby structures. The main feature the network seems to use for distinguishing microbleeds is the intensity. Ideally, the network would learn higher level features.

The regularisation methods seem to reduce the false positives but sadly still at the cost of true positives.

Next steps

The results suggest that the training parameters are not optimal and this is certainly a point that needs to be improved upon. We plan to improve this network so that we have a baseline for microbleed segmentation. We can then use this baseline to develop more complex networks, or to develop a network that is able to handle input from multiple modalities. Having multiple modalities is expected to greatly improve performance because common false positives – such as vessels – will appear different on other modalities than microbleeds.

References

- [1] li Gibson, Wenqi Li, Carole Sudre, Lucas Fidon, Dzoshkun Shakir, Guotai Wang, Zach Eaton-Rosen, Robert Gray, Tom Doel, Yipeng Hu, Tom Whyntie, Parashkev Nachev, Dean C. Barratt, Sebastien Ourselin, M. Jorge Cardoso, and Tom Vercauteren. Niftynet: a deep-learning platform for medical imaging. arXiv preprint arXiv:1709.03485, 2017.
- [2] Sudre, C. et. al. (2017) Generalised Dice overlap as a deep learning loss function for highly unbalanced segmentations. DLMI 2017

Adaptive Sampling

In order to lower the rate of False Positives, the sampling scheme was changed to a dynamic scheme. This means that after training for a few initial epoch, training would be stopped and inference would be run on the training set, and the results would be used to update the sampling map, so that the network would also be presented with False positive patches during training.

Count regularisation

Whilst the adaptive sampling did reduce the number of FPs per image, it was still relatively high. By using the number of detected microbleeds in a patch we hoped to reduce the number of False positives, without sacrificing performance. The dice loss was changed to also accommodate a loss term based on the number of microbleed in the predicted patch versus the actual number of microbleeds.

	FP per image	TP	FN	Precision	Recall
Baseline	165	89 %	11 %	0.07	0.88
Adaptive sampling	28	74 %	32 %	0.26	0.89
Count	23	69 %	30 %	0.30	0.69