Bayesian Skip Net: Building on Prior Information for the Prediction and Segmentation of Stroke Lesions

Julian Klug[1,2(✉)], Guillaume Leclerc[3], Elisabeth Dirren[1], Maria Giulia Preti[2], Dimitri Van De Ville[2], and Emmanuel Carrera[1]

[1] Stroke Research Group, Department of Clinical Neurosciences, University Hospital and Faculty of Medicine, Geneva, Switzerland
julian.klug@etu.unige.ch
[2] Medical Image Processing Laboratory, Institute of Bioengineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
[3] Massachusetts Institute of Technology, Cambridge, USA

Abstract. Perfusion CT is widely used in acute ischemic stroke to determine eligibility for acute treatment, by defining an ischemic core and penumbra. In this work, we propose a novel way of building on prior information for the automatic prediction and segmentation of stroke lesions. To this end, we reformulate the task to identify differences from a prior segmentation by extending a three-dimensional Attention Gated Unet with a skip connection allowing only an unchanged prior to bypass most of the network. We show that this technique improves results obtained by a baseline Attention Gated Unet on both the Geneva Stroke Dataset and the ISLES 2018 dataset.

Keywords: Prior · Stroke · Convolutional neural network · Medical image segmentation

1 Introduction

Ischemic stroke is a leading cause of mortality and disability worldwide [5]. In the last decade, advances in stroke imaging have enabled a more targeted approach to treatment and reperfusion. Perfusion CT (pCT) is widely used in acute ischemic stroke to determine eligibility for treatment by providing perfusion parameter maps informing about voxelwise cerebral blood flow, cerebral blood volume, transit time and time to maximum of the residue function (CBF, CBV, MTT and Tmax respectively) [8]. Necrotic tissue defining the ischemic core, as well as hypoperfused but still salvageable tissue, the ischemic penumbra, are commonly delineated by threshold-based tools [6]. These techniques have shown to be of great benefit to select patients for treatment in several clinical trials [2,21]. However, accurate segmentation of the ischemic core on pCT, as well as prediction of the final infarct is challenging, and diffusion weighed magnetic resonance imaging (MRI) remains the gold standard.

A. Crimi and S. Bakas (Eds.): BrainLes 2020, LNCS 12658, pp. 168–180, 2021.
https://doi.org/10.1007/978-3-030-72084-1_16

With the advent of deep convolutional neural networks (CNNs), significant improvements in stroke prediction and segmentation have been made [18]. Current methods take as input acute imaging sequences and learn a tissue outcome prediction function from labelled data obtained from follow-up imaging. Many currently presented models rely on a Unet backbone [9,10,15,18,28], allowing for multi-scale localisation and contextualisation through an encoder-decoder structure with skip-connections. In the setting of the ISLES 2018 challenge, the best performing submissions achieved a Dice score ranging from 0.48 to 0.51 for the segmentation of the ischemic core on perfusion CT [1,12,17]. Current methods for the prediction of the final lesion often use a combination of perfusion MRI and diffusion weighed imaging leading to a Dice score of up to 0.53 [20,30]. Predicting the final lesion from perfusion CT alone has been attempted through the use of baseline perfusion images and infarct growth estimations [16,26] with Dice scores of up to 0.48.

Incorporating prior knowledge has proven useful in many medical image segmentation tasks [22]. Indeed, the inclusion of a prior can simplify a model's task by reducing the amount of information to be learnt. Atlas models, distance boundaries, shape and topology specifications as well as edge polarity have been successfully used as regularisation terms in region growing segmentation methods. Recent work has used prior constraints in the form of adjacency, boundary or learned anatomical conditions [23].

In acute ischemic stroke, tissue of the penumbra is progressively recruited into the core, contributing to the growth of the lesion. Previous models have successfully used prior manual ischemic core and penumbra segmentations to obtain a representation of ischemic stroke growth directions [16]. Using standardized thresholds for the automated segmentation of the ischemic core could potentially leverage the strengths of a largely clinically validated model. Integrating this information as a prior can be used as a starting point to either refine the segmentation of the ischemic core or to predict the final lesion. In this work we aim to efficiently integrate prior information obtained by a standard threshold segmentation of the ischemic core directly into the commonly used Attention Gated 3D Unet [23,27,30].

2 Materials and Methods

2.1 Data

Geneva Stroke Dataset. This dataset comprises acute pCT images and final lesion labels of 144 patients who have benefitted from treatment by thrombectomy and/or thrombolysis as described in prior work [13]. For every subject a full-volume CBF, CBV, MTT, Tmax and non-contrast CT (NCCT) image is available. Manually annotated labels for the final infarct have been obtained from follow-up MRI in the sub-acute phase by expert neurologists. Briefly, a model has to learn the prediction of the final infarct after treatment.

ISLES 2018 Dataset. The Ischemic Stroke Lesion Segmentation 2018 dataset contains pCT images and acute lesion labels of patients before undergoing treatment [12,17]. Every subject has the main slices containing the acute lesion of CBF, CBV, MTT, Tmax and NCCT sequences. The gold-standard ischemic cores were defined on the subsequent MRI, performed in the acute setting and manually annotated. We used the publicly available training subset comprising 94 data points. In this setting, a model has to learn the segmentation of stroke lesions before treatment.

2.2 Data Pre-processing

We normalize by subtracting the mean and dividing by the standard deviation all input sequences, which are subsequently scaled to a 0–1 range. To ensure divisibility by 16 necessary for our Unet architecture we pad by 0 along every dimension. All data used for training is augmented ad hoc by applying random flip, random elastic transform, random shift, random scale and Gaussian noise [25]. Every individual transformation has a 50% probability of being applied on each batch.

Ischemic core is commonly defined as relative CBF < 0.3 (rCBF). In our work we have chosen a slightly more inclusive threshold at 0.38 for greater sensitivity. rCBF is defined as relative (by ratio) to the mean CBF of the contralateral hemisphere after smoothing with a 3D Gaussian kernel of 2 voxels width [7]. Cerebral spinal fluid (CSF) is segmented on NCCT images by applying a 5th percentile threshold on voxels bounded between 0 and 100 Hounsfield units (HU). A segmentation of the skull is obtained with the bet2 algorithm of the FMRIB Software Library (FSLv.5) [11,19] on NCCT images with the same bounds. Finally, major blood vessels are segmented by applying a 99th percentile threshold on CBF images. The resulting CSF, skull and vessel segmentations were then extended slightly by binary dilation with a spheroid structuring element of width 2 voxels. They were then removed from the initial ischemic core segmentations. A similar pipeline is implemented by the commercially available RAPID software package (RAPID, Ischemaview Stanford University, Stanford, USA) and widely used in clinical practice [29]. The final ischemic core segmentation is defined as prior segmentation (Fig. 1).

2.3 Network Architecture

Attention Gated Unet. A 3D Unet with attention gated skip connections is used as the baseline model. This encoder-decoder architecture makes use of contextual information by down-sampling input volumes. Spatial details are then recovered by up-sampling. The information bottleneck is classically overcome by skip connections spanning between down-sampling and up-sampling layers of corresponding scale to provide more contextual information to the model [31]. Attention gates applied to the skip connections extract features from coarse scale to highlight salient regions on a finer scale [24]. Every convolutional module is composed of two layers with a 3D-convolution with a $3 \times 3 \times 3$ kernel followed by

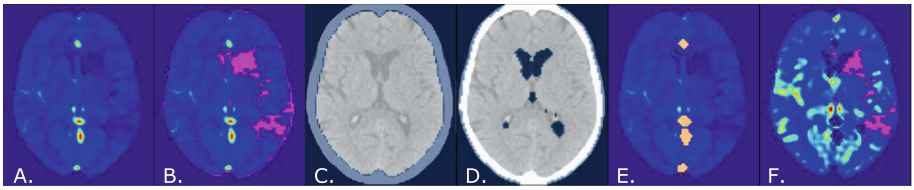

Fig. 1. From left to right: axial view of CBF (A.), thresholded relative CBF at 0.38 overlaid in magenta on CBF (B.), skull segmentation (C.) and ventricle segmentation (D.) overlaid in blue on NCCT, vascular noise filter in orange (E.) and the final ischemic core segmentation in magenta (F.) overlaid on CBF of a representative subject of the Geneva Stroke Dataset. The final threshold-based ischemic core segmentation is subsequently used as prior. (Color figure online)

batch-normalisation and a Rectified linear unit (ReLU) as activation function. Four convolutional modules are used respectively during down- and up-sampling with an additional central module. Deep supervision is used to enhance the transition between feature space and semantically relevant segmentation [14].

Bayesian Skip for Attention Gated Unet. The above described Unet creates large receptive fields by successive down-sampling of input information to model relationships at a gradually coarser scale. It remains however challenging to reduce false-positive rates for small and patchy segmentations of varying shape. To address this issue, we propose to allow selected input channels defined as prior to bypass the main part of the model, before being reintegrated at the final layer. The network thus effectively learns to model the difference relative to the prior given all inputs. To this end we add an additional skip connection spanning the network from its input to its final layer, termed bayesian skip. There, it is integrated by Method A) summation with the output or Method B) by convolution with the output. In Method B) a $1 \times 1 \times 1$ convolution is used to reduce computational overhead. The network can thus be described as the function *U(x)* with *u(x)* defined as the skipped part of the network and p the prior. Note that at any given step, $p \in x$ as the prior is not removed from the input. Thus, for the baseline model $U(x) = u(x)$. The Methods A) and B) can be respectively described by Eqs. 1 and 2.

$$U(x)_A = u(x) + p \tag{1}$$

$$U(x)_B = u(x) * p \tag{2}$$

Subsequently, *u(x)* learns to fit the difference between *U(x)* and p in Method A) and the deconvolution in Method B) (Fig. 2).

2.4 Experimental Setup

We rearrange our data in a three-way split, with 70% used for training, 15% for validation and 15% for testing. Volumes of size $96 \times 96 \times 96$ and $256 \times 256 \times 22$

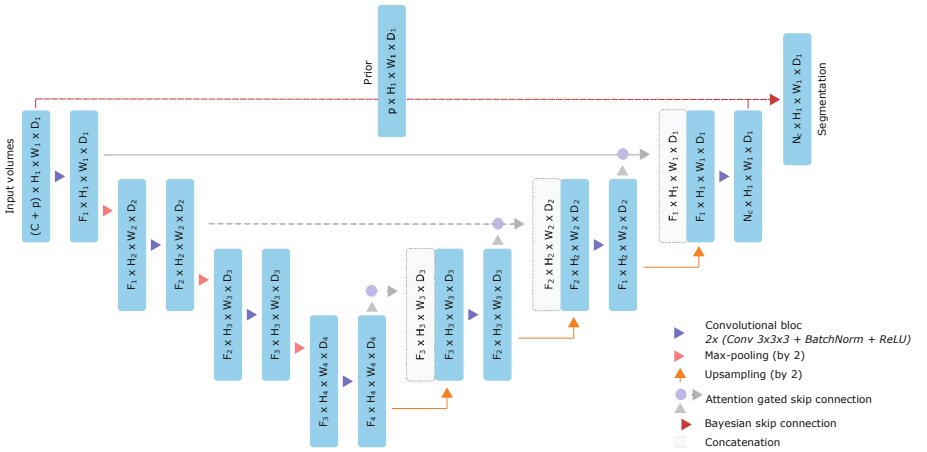

Fig. 2. A bloc diagram of the proposed Bayesian Skip Net, building on the previously designed Attention Gated Unet. The bayesian skip connection enables the prior to bypass the network unchanged. The prior is then reintegrated with the output of the final layer via Method A) summation or Method B) convolution.

are used respectively for the Geneva Stroke Dataset and the ISLES 2018 Dataset. We use a batch-size of 2 and 4 respectively for the datasets, dictated by computational limitations. Stochastic gradient descent is used with an initial learning rate of 0.0001 and Nesterov momentum. Weight decay is used as regularisation. The learning rate is reduced by a factor of 2 every 200 epochs. We optimize the commonly employed smooth Sorensen-Dice loss defined as follows:

$$L_{Dice} = 1 - \frac{2|X \cap Y| + \epsilon}{|X| + |Y| + \epsilon} \tag{3}$$

where a smoothing factor ϵ of 0.01 is used. Early stopping on the validation loss is used with a patience of 20 after 200 epochs. The best model is selected based on best Dice score on the validation set and is subsequently evaluated on the independent test set. No hyperparameter optimisation is attempted as absolute results are not the objective of this work. A baseline Attention Gated Unet, as well as two Attention Gated Unets with bayesian skip integrated through Method A) or B) respectively, have been trained and evaluated using CBF, CBV, MTT, Tmax and NCCT sequences, as well as the prior segmentation as input channels of the Geneva Stroke Dataset. We then validated the experiments for the baseline network and Method B) on the ISLES 2018 Dataset. All computations were done on a single Nvidia Tesla P100 GPU (Nvidia Corporation, Santa Clara, California, USA). Our Pytorch implementation for the proposed architecture is publicly available *here*.

Impact of Prior Quality. To evaluate the impact of the quality of the prior segmentation on the performance of the Bayesian Skip Net, we computed the

correlation between the Dice score obtained by the prior alone and the score obtained by the model making use of the prior using Pearson's correlation coefficient.

We further evaluated the performance of the proposed model given a degraded prior. To this aim we added random binary noise where $p(1) = 0.002$ on the prior segmentation used for training and testing. On average, this results in a noisy prior in which 25% of the total volume corresponds to randomly segmented voxels. For both datasets, a Bayesian Skip Net was trained and tested using the resulting noisy segmentation as prior.

3 Results

Experimental results on validation and test splits for the Geneva Stroke Dataset and ISLES 2018 Dataset are reported respectively in Tables 1 and 2. Overall, the Bayesian Skip Net with Method B) performs better than the baseline Attention Gated Unet on both datasets. It consistently achieves better Dice and precision scores on the Geneva Stroke Dataset as well as on the test split of the ISLES 2018 Datatset. On the ISLES 2018 validation split a performance similar to baseline is achieved. The baseline model mostly remains superior in terms of recall. All evaluated methods outperform the prior segmentation. Furthermore, the proposed method greatly speeds up the training process, as the Bayesian Skip Net with Method B) achieves convergence after about 300 epochs, compared to a baseline convergence of 450 epochs when evaluated on the ISLES 2018 dataset as shown in Fig. 3. The proposed model with Method A) does not yield any increase in performance and was therefore not further evaluated on the second dataset.

Correlation between the quality of the prior used and Dice scores obtained by the proposed model with Method B) on the test splits of both datasets are shown in Fig. 4. For both the Geneva Stroke Dataset (Pearson's $R = 0.82$, $p < 10^{-5}$) and the ISLES 2018 Dataset (Pearson's $R = 0.76$, $p < 10^{-2}$), Dice scores obtained by the prior and by the Bayesian Skip Net correlate strongly.

Evaluation of the Bayesian Skip Net with Method B) given a prior with and without added noise on the test splits for the Geneva Stroke Dataset and ISLES 2018 Dataset are reported in Table 3. On both tasks, adding noise to the prior slightly reduces model performance. When comparing with the baseline Unet, the proposed model achieves comparable or better results even when using a noisy prior.

Example predictions are shown in Fig. 5. A case-by-case visual analysis of the predicted lesions of all evaluated methods revealed that the proposed model consistently produced more precise segmentations and was less likely to produce false positive results than the baseline method. A selective analysis of cases where both models failed to segment the lesion revealed small infarcts along the midline, in the brainstem or the cerebellum. The prior segmentation did not include any of these lesions.

Table 1. Experimental results on test and validation splits for the Geneva Stroke Dataset. The results are reported in terms of mean and standard deviation for Dice score, precision and recall. The proposed Unet with bayesian skip with Method A) and B) is benchmarked against the baseline Attention Gated Unet. The prior segmentation's performance is reported as reference. Best model results in bold.

Data Split	Method	Dice	Precision	Recall
Validation	Prior	0.125 ± 0.135	0.149 ± 0.128	0.171 ± 0.194
	Unet	0.270 ± 0.215	0.265 ± 0.299	**0.404** ± 0.336
	Unet+Method A	0.246 ± 0.206	0.221 ± 0.247	0.300 ± 0.299
	Unet+Method B	**0.292** ± 0.211	**0.348** ± 0.333	0.294 ± 0.257
Test	Prior	0.099 ± 0.110	0.109 ± 0.116	0.119 ± 0.163
	Unet	0.192 ± 0.156	0.189 ± 0.235	**0.271** ± 0.284
	Unet+Method A	0.181 ± 0.154	0.132 ± 0.187	0.278 ± 0.324
	Unet+Method B	**0.212** ± 0.136	**0.289** ± 0.333	0.188 ± 0.219

Table 2. Experimental results on test and validation splits for the ISLES 2018 Dataset compared with the top three submissions in the ISLES 2018 challenge. The results are reported in terms of mean and standard deviation for Dice score, precision and recall. The proposed Unet with bayesian skip with Method B) is benchmarked against the baseline Attention Gated Unet. The prior segmentation's performance, as well as the top three submissions of the ISLES 2018 challenge are reported as reference. Best model results in bold.

Data Split	Method	Dice	Precision	Recall
Validation	Prior	0.189 ± 0.174	0.214 ± 0.181	0.236 ± 0.230
	Unet	**0.417** ± 0.072	**0.433** ± 0.304	**0.419** ± 0.264
	Unet+Method B	0.415 ± 0.073	0.431 ± 0.301	0.411 ± 0.264
Test	Prior	0.296 ± 0.256	0.251 ± 0.222	0.374 ± 0.307
	Unet	0.524 ± 0.182	0.532 ± 0.354	0.560 ± 0.312
	Unet+Method B	**0.552** ± 0.195	**0.561**± 0.238	**0.573** ± 0.292
ISLES 2018 Challenge	Song et al. [28]	0.51 ± 0.31	0.55 ± 0.36	0.55 ± 0.34
	Liu et al. [15]	0.49 ± 0.31	0.56 ± 0.37	0.53 ± 0.33
	Chen et al. [9]	0.48 ± 0.31	0.59 ± 0.38	0.46 ± 0.33

Table 3. Impact of noise added to the prior on the performance of the Bayesian Skip Net with Method B). Experimental results on the test splits of the Geneva Stroke Dataset and the ISLES 2018 Dataset with and without noise added to the prior are reported in terms of mean and standard deviation for Dice score, precision and recall.

Data set	Prior quality	Dice	Precision	Recall
Geneva Stroke Dataset	Standard	0.212 ± 0.136	0.289 ± 0.333	0.188 ± 0.219
	Noisy	0.191 ± 0.165	0.204 ± 0.277	0.208 ± 0.244
ISLES 2018 Dataset	Standard	0.552 ± 0.195	0.561 ± 0.238	0.573 ± 0.292
	Noisy	0.538 ± 0.205	0.565 ± 0.282	0.565 ± 0.271

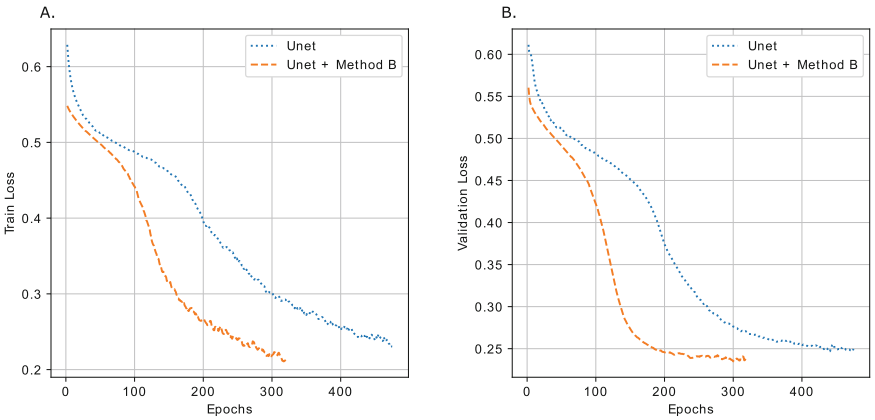

Fig. 3. Training (A.) and validation (B.) loss during training of the baseline Attention Gated Unet (blue) and the proposed Bayesian Skip Net with Method B) (orange) on the ISLES 2018 Dataset. Efficient integration of the prior ensures faster training. Both models tend to overfit on validation and test data. This is common in deep learning and generally does not prevent generalisation on test data [4]. (Color figure online)

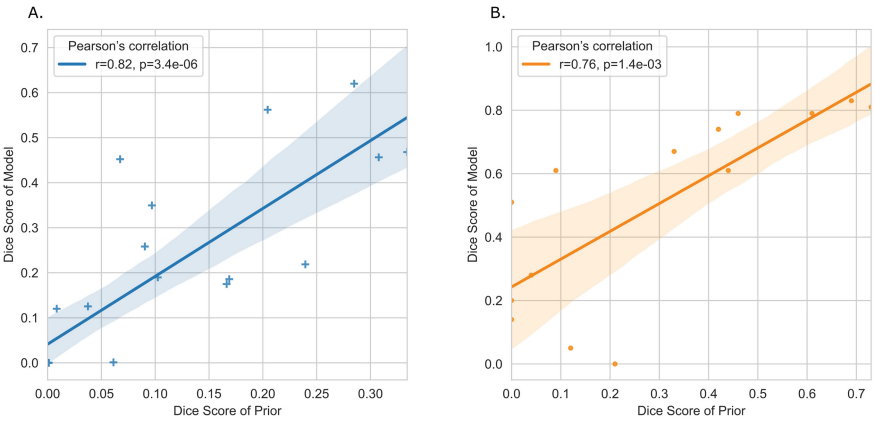

Fig. 4. Correlation of Dice scores obtained by the prior segmentation and the Bayesian Skip Net with Method B) on the test splits of the Geneva Stroke Dataset (A.) and the ISLES 2018 Dataset (B.) For both datasets, the performance of the model strongly correlates with the quality of the prior.

4 Discussion and Conclusion

In this paper, we present a novel bayesian skip connection as a way to take into account a prior in medical image segmentation. By reintegrating prior knowledge at the end of the proposed model, we explicitly let the rest of the network approximate the divergence from this prior. Having the same information at their disposal, both the baseline model and the model with the bayesian skip connec-

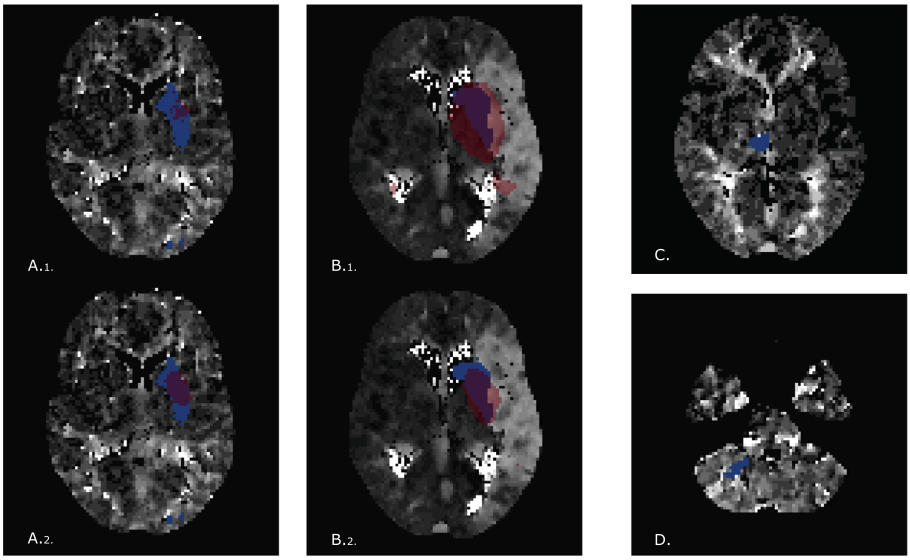

Fig. 5. Lesion labels, in blue, and model predictions, in red, projected on Tmax axial views for four subjects (A-D). The left and middle images show the predicted segmentation by the baseline Attention Gated Unet (upper subimage, 1.) and by the proposed Unet with bayesian skip with Method B) (lower subimage, 2.) for two patients (A and B). The two images on the right show subjects were neither method achieves to segment the lesion (C and D). Small lesions located along the midline and in the posterior fossa remain difficult for both models and account for most of the variability in performance. (Color figure online)

tion should be able to approximate the desired segmentation function. However, the amount of information to be learnt is smaller in the latter, improving its learning process. This results in faster convergence and better performance on both the Geneva Stroke Dataset and the ISLES 2018 dataset.

Within our experimental setting, bayesian skip connections with convolution (Method B) but not with summation (Method A) improve the performance of an Attention Gated Unet when a prior segmentation is given as input for a task of ischemic lesion prediction and segmentation. The prior used in our setting is associated with its own uncertainty. This can be modelled by the convolution but not by a summation explaining the superiority of Method B) relative to Method A).

The overall gain in performance is mainly driven by increased precision as the proposed model focuses on regions already segmented in the prior, thus reducing its false-positive rate. However, this comes at the cost of a lower recall value. The results reported here further show that the proposed method greatly speeds up model convergence which is crucial for rapid model iteration and reduces computational costs.

The Bayesian Skip Net compares favourably with the best performing models of the ISLES 2018 challenge. Although the analysis is somewhat limited as the test set used for the challenge is not publicly available and does not correspond exactly to the same subjects used for the test split in our experiments, this comparison on a same dataset suggests that our proposed model achieves state-of-the art performance for the segmentation of the ischemic core.

The comparison of results for the prediction of the final lesion from perfusion CT is more complicated as datasets referenced across the literature are heterogenous and not open-source. Unsurprisingly, the Bayesian Skip Net outperforms earlier general linear models with receptive fields on the Geneva Stroke Dataset [13]. Robben et al. use the raw perfusion CT signal as input to a CNN and report a Dice score of 0.48 on the dataset of the MRCLEAN study compromising globally bigger lesions (mean lesion volume [interquartile range]: 78 cm^3 [21-121] vs. 29.4 cm^3 [2.5-37.6]) [26]. This highlights the importance of lesion volume as a component of model performance and could suggest a potential loss of information due to the conversion from raw perfusion CT images to perfusion maps (Tmax, MTT, CBF, CBV). In a small sample of the TRAVESTROKE dataset, Lucas et al. obtain a Dice score of 0.46 by modulating a deformation model obtained from manual segmentations of the ischemic penumbra and core [16]. Although manual segmentations are not sustainable in clinical practice, this underscores the importance of the quality of the prior used and suggests that adding a threshold-based segmentation of the ischemic penumbra to the prior might yield further gains in performance.

We show that our model is relatively resistant to added noise as it maintains similar or greater performance than the baseline Unet when presented with a degraded prior. Nonetheless, the gain in performance of our model strongly correlates with the quality of the prior given as input. For the prediction of the final lesion, it also depends on the overlap between ischemic core and final lesion that strongly depends on clinical intervention. In this work, we choose to use our own open-source implementation of the ischemic core segmentation to ensure reproducibility. However, this could be improved by using commercially available software [3,29]. Moreover, for the prediction of the final lesion, a model segmenting the ischemic core such as the one proposed for the ISLES 2018 dataset, could be used to provide a prior of greater quality.

All models evaluated showed great inter-case variability and weak performance on small lesions located along the midline, as well as on lesions in the posterior fossa which remain difficult to detect. The prior segmentation failed to cover the same lesions and its integration could therefore not improve final performance. Moreover, results varied substantially between the two datasets, with greater scores achieved on the ISLES 2018 dataset. This can be explained by a more complicated task in the Geneva Stroke Dataset which consists of predicting a segmentation of the final infarct obtained several days after treatment. Greater inter-subject variability in this dataset mainly stems from a greater range of treatment administered to the patients, as well as highly divergent radiological outcomes.

In a clinical context, medical models rarely stand in isolation but are integrated with a prior representation of the problem. The segmentation and prediction of infarct evolution in acute ischemic stroke are challenging tasks for humans and machines alike. Threshold-based ischemic core segmentation has been largely validated and is commonly used as a starting point in this setting. Our proposed model can effectively leverage this prior segmentation to enhance the prediction of radiological outcome.

References

1. ISLES: Ischemic Stroke Lesion Segmentation Challenge (2018). http://www.isles-challenge.org/
2. Albers, G.W., et al.: Thrombectomy for Stroke at 6 to 16 Hours with Selection by Perfusion Imaging. New England J. Med. **378**(8), 708–718 (2018). https://doi.org/10.1056/NEJMoa1713973. Massachusetts Medical Society, publisher eprint
3. Shalini, A., et al.: Cerebral blood flow predicts the infarct core. Stroke **50**(10), 2783–2789 (2019). https://doi.org/10.1161/STROKEAHA.119.026640. https://www.ahajournals.org/doi/10.1161/STROKEAHA.119.026640. Publisher: American Heart Association
4. Belkin, M., Hsu, D., Ma, S., Mandal, S.: Reconciling modern machine learning practice and the bias-variance trade-off. arXiv:1812.11118 [cs, stat] (2019)
5. Benjamin Emelia, J., et al.: Heart disease and stroke statistics–2019 update: a report from the American Heart Association. Circulation, **139**(10), e56–e528 (2019). https://doi.org/10.1161/CIR.0000000000000659. https://www.ahajournals.org/doi/10.1161/CIR.0000000000000659. Publisher: American Heart Association
6. Campbell, B.C.V., Khatri, P.: Stroke. Lancet **396**(10244), 129–142 (2020). https://doi.org/10.1016/S0140-6736(20)31179-X. https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)31179-X/abstract. Publisher: Elsevier
7. Campbell, B.C.V., et al.: Cerebral blood flow is the optimal CT perfusion parameter for assessing infarct core. Stroke, **42**(12), 3435–3440 (2011). https://doi.org/10.1161/STROKEAHA.111.618355. https://www.ahajournals.org/doi/10.1161/strokeaha.111.618355. Publisher: American Heart Association
8. Carrera, E., Wintermark, M.: Imaging-based selection of patients for acute stroke treatment: is it ready for prime time? Neurology **88**(24), 2242–2243 (2017). https://doi.org/10.1212/WNL.0000000000004051
9. Chen, Y., Li, Y., Zheng, Y.: Ensembles of modalities fused model for ischemic stroke lesion segmentation, p. 1
10. Clèrigues, A., Valverde, S., Bernal, J., Freixenet, J., Oliver, A., Lladó, X.: Acute ischemic stroke lesion core segmentation in CT perfusion images using fully convolutional neural networks. Comput. Biol. Med. **115**, 103487 (2019). https://doi.org/10.1016/j.compbiomed.2019.103487. http://www.sciencedirect.com/science/article/pii/S0010482519303555
11. Jenkinson, M., Pechaud, M., Smith, S.: BET2 - MR-Based Estimation of Brain, Skull and Scalp Surfaces, p. 1
12. Kistler, M., Bonaretti, S., Pfahrer, M., Niklaus, R., Büchler, P.: The virtual skeleton database: an open access repository for biomedical research and collaboration. J. Med. Internet Res. **15**(11) (2013). https://doi.org/10.2196/jmir.2930. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3841349/

13. Klug, J., et al.: Integrating regional perfusion CT information to improve prediction of infarction after stroke. J. Cerebr. Blood Flow Metab. (2020). https://doi.org/10.1177/0271678X20924549. https://journals.sagepub.com/doi/10.1177/0271678X20924549. Publisher: SAGE PublicationsSage UK: London, England
14. Lee, C.Y., Xie, S., Gallagher, P., Zhang, Z., Tu, Z.: Deeply-supervised nets. arXiv:1409.5185 [cs, stat] (2014)
15. Liu, P.: Stroke lesion segmentation with 2D novel CNN pipeline and novel loss function. In: Crimi, A., Bakas, S., Kuijf, H., Keyvan, F., Reyes, M., van Walsum, T. (eds.) BrainLes 2018. LNCS, vol. 11383, pp. 253–262. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-11723-8_25
16. Lucas, C., Aulmann, L., Kemmling, A., Madany Mamlouk, A., Heinrich, M.: Estimation of the principal ischaemic stroke growth directions for predicting tissue outcomes, pp. 69–79 (2020). https://doi.org/10.1007/978-3-030-46640-4_7
17. Maier, O., et al.: ISLES 2015 - A public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI. Med. Image Anal. **35**, 250–269 (2017). https://doi.org/10.1016/j.media.2016.07.009
18. Kim, M., Patrick, T., Greg, Z.: Artificial intelligence applications in stroke. Stroke **51**(8), 2573–2579 (2020). https://doi.org/10.1161/STROKEAHA.119.027479. https://www.ahajournals.org/doi/10.1161/STROKEAHA.119.027479. Publisher: American Heart Association
19. Muschelli, J., Ullman, N.L., Mould, W.A., Vespa, P., Hanley, D.F., Crainiceanu, C.M.: Validated automatic brain extraction of head CT images. Neuroimage **114**, 379–385 (2015). https://doi.org/10.1016/j.neuroimage.2015.03.074
20. Nielsen, A., Hansen, M.B., Tietze, A., Mouridsen, K.: Prediction of tissue outcome and assessment of treatment effect in acute ischemic stroke using deep learning. Stroke **49**(6), 1394–1401 (2018). https://doi.org/10.1161/STROKEAHA.117.019740. https://www.ahajournals.org/doi/full/10.1161/strokeaha.117.019740. Publisher: American Heart Association
21. Nogueira, R.G., et al.: DAWN trial investigators: thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct. N. Engl. J. Med. **378**(1), 11–21 (2018). https://doi.org/10.1056/NEJMoa1706442
22. Nosrati, M.S., Hamarneh, G.: Incorporating prior knowledge in medical image segmentation: a survey. arXiv:1607.01092 [cs] (2016)
23. Oktay, O., et al.: Anatomically constrained neural networks (ACNNs): application to cardiac image enhancement and segmentation. IEEE Trans. Med. Imaging **37**(2), 384–395 (2018). https://doi.org/10.1109/TMI.2017.2743464. Conference Name: IEEE Transactions on Medical Imaging
24. Oktay, O., et al.: Attention U-Net: learning where to look for the pancreas, p. 10
25. Péez-García, F., Sparks, R., Ourselin, S.: TorchIO: a Python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning. arXiv:2003.04696 [cs, eess, stat] (2020)
26. Robben, D., et al.: Prediction of final infarct volume from native CT perfusion and treatment parameters using deep learning. Med. Image Anal. **59**, 101589 (2020). https://doi.org/10.1016/j.media.2019.101589. http://www.sciencedirect.com/science/article/pii/S136184151930129X
27. Schlemper, J., et al.: Attention gated networks: learning to leverage salient regions in medical images. Med. Image Anal. **53**, 197–207 (2019). https://doi.org/10.1016/j.media.2019.01.012. http://www.sciencedirect.com/science/article/pii/S1361841518306133

28. Song, T., Huang, N.: Integrated extractor, generator and segmentor for ischemic stroke lesion segmentation. In: Crimi, A., Bakas, S., Kuijf, H., Keyvan, F., Reyes, M., van Walsum, T. (eds.) BrainLes 2018. LNCS, vol. 11383, pp. 310–318. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-11723-8_31
29. Straka, M., Albers, G.W., Bammer, R.: Real-time diffusion-perfusion mismatch analysis in acute stroke. J. Magn. Reson. Imaging: JMRI, **32**(5), 1024–1037 (2010). https://doi.org/10.1002/jmri.22338. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2975404/
30. Yu, Y., et al.: Use of deep learning to predict final ischemic stroke lesions from initial magnetic resonance imaging. JAMA Netw. Open **3**(3), e200772–e200772 (2020). https://doi.org/10.1001/jamanetworkopen.2020.0772. https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2762679. Publisher: American Medical Association
31. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9901, pp. 424–432. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46723-8_49