Patch-based 3D U-Net and transfer learning for longitudinal piglet brain segmentation on MRI

Patty Coupeau, Jean-Baptiste Fasquel, E. Mazerand, P. Menei, C.N. Montero-Menei, Mickaël Dinomais

Patch-based 3D U-Net and transfer learning for longitudinal piglet brain segmentation on MRI

P. Coupeau[a], J.-B. Fasquel[a], E. Mazerand[b,c], P. Menei[b,c], C.N. Montero-Menei[b] and M. Dinomais[a,d]

[a]LARIS, SFR MATHSTIC, Université d'Angers, France
[b]CRCINA, UMR 1232 INSERM, Université de Nantes, Université d'Angers, France
[c]Département de neurochirurgie, Centre Hospitalier Universitaire d'Angers, France
[d]Département de médecine physique et de réadaptation, Centre Hospitalier Universitaire d'Angers, France

ARTICLE INFO

Keywords:
Piglet brain segmentation
3D U-Net
Patches
Transfer learning
MRI
Brain development

ABSTRACT

Background and Objectives In order to study neural plasticity in immature brain following early brain lesion, large animal model are needed. Because of its morphological similarities with the human developmental brain, piglet is a suitable but little used one. Its study from Magnetic Resonance Imaging (MRI) requires the development of automatic algorithms for the segmentation of the different structures and tissues. A crucial preliminary step consists in automatically segmenting the brain.
Methods We propose a fully automatic brain segmentation method applied to piglets by combining a 3D patch-based U-Net and a post-processing pipeline for spatial regularization and elimination of false positives. Our approach also integrates a transfer-learning strategy for managing an automated longitudinal monitoring evaluated for four developmental stages (2, 6, 10 and 18 weeks), facing the issue of MRI changes resulting from the rapid brain development. It is compared to a 2D approach and the Brain Extraction Tool (BET) as well as techniques adapted to other animals (rodents, macaques). The influence of training patches size and distribution is studied as well as the benefits of spatial regularization.
Results Results show that our approach is efficient in terms of average Dice score (0.952) and Hausdorff distance (8.51), outperforming the use of a 2D U-Net (Dice: 0.919, Hausdorff distance: 11.06) and BET (Dice: 0.764, Hausdorff distance: 25.91). The transfer-learning strategy achieves a good performance on older piglets (Dice of 0.934 at 6 weeks, 0.956 at 10 weeks and 0.958 at 18 weeks) compared to a standard training strategy with few data (Dice of 0.636 at 6 weeks, 0.907 at 10 weeks, not calculable at 18 weeks because of too few training piglets).
Conclusions In conclusion, we provide a method for longitudinal MRI piglet brain segmentation based on 3D U-Net and transfer learning which can be used for future morphometric studies and applied to other animals.

1. Introduction

Cerebral Palsy (CP) is the most common motor disability in childhood. It results from a permanent early brain lesion occurring in developmental brain [1]. Presently, no curative therapies are available to cure the brain damage. In the past decades, interest has grown on the mechanisms of brain plasticity after early brain lesion. Understanding the complex process of such post-lesional neuroplasticity occurring in an immature brain is a fast-growing field of neuroscientific research that has the potential to prompt more targeted and evidence-based interventions in order to enhance reorganization of neural functions (i,e. neural plasticity). Animal models [2] are needed to better understand in vivo mechanism of early brain lesion and neuroplasticity as they offer controlled experimental conditions. Whereas CP rodent model exists [3], compare to human, these latter may have potential bias and differences and omit important context, possibly hampering reliable translational outcomes. Thus, other animal models are required [2]. Piglet model seems to be a good candidate to examine early brain plasticity in a context of CP.

The use of the piglet as an animal model for studying infant brain development is increasing [4]. Indeed, the anatomical characteristics of the piglet brain, including its growth and maturation of myelin, are very similar to those of the human-being [5]. Thus, the piglet model makes it possible to better understand the process of brain maturation and to analyze future neurodevelopmental disorders [6]. To this end, a piglet model reproducing the lesion has been developed [7].

In order to quantify the consequences of the lesion on brain development at a macroscopic scale, magnetic resonance imaging (MRI) is a suitable tool [4]. Nevertheless, it is necessary to develop MRI image analysis algorithms such as the segmentation of brain structures and tissues or some morphometric analyzes which will allow quantitative longitudinal monitoring of brain development. The development of this type of automatic algorithms is essential to facilitate the study of large animal models for which little data is available. The goal is to set up automatic MRI image processing techniques on piglets equivalent to those existing for the human or the infant brain [8]. The absence of such tools for the piglet is explained by the few existing studies on this model. Some works were conducted on the minipig (Göttingen [9], Yucatan [10]) and the domestic pig [11] but only two studies dealing with the creation of a template of the piglet brain appear in the literature [4, 12]. Their aim was to automate the segmentation of brain structures using an atlas

patty.coupeau@etud.univ-angers.fr (P. Coupeau)
ORCID(s):

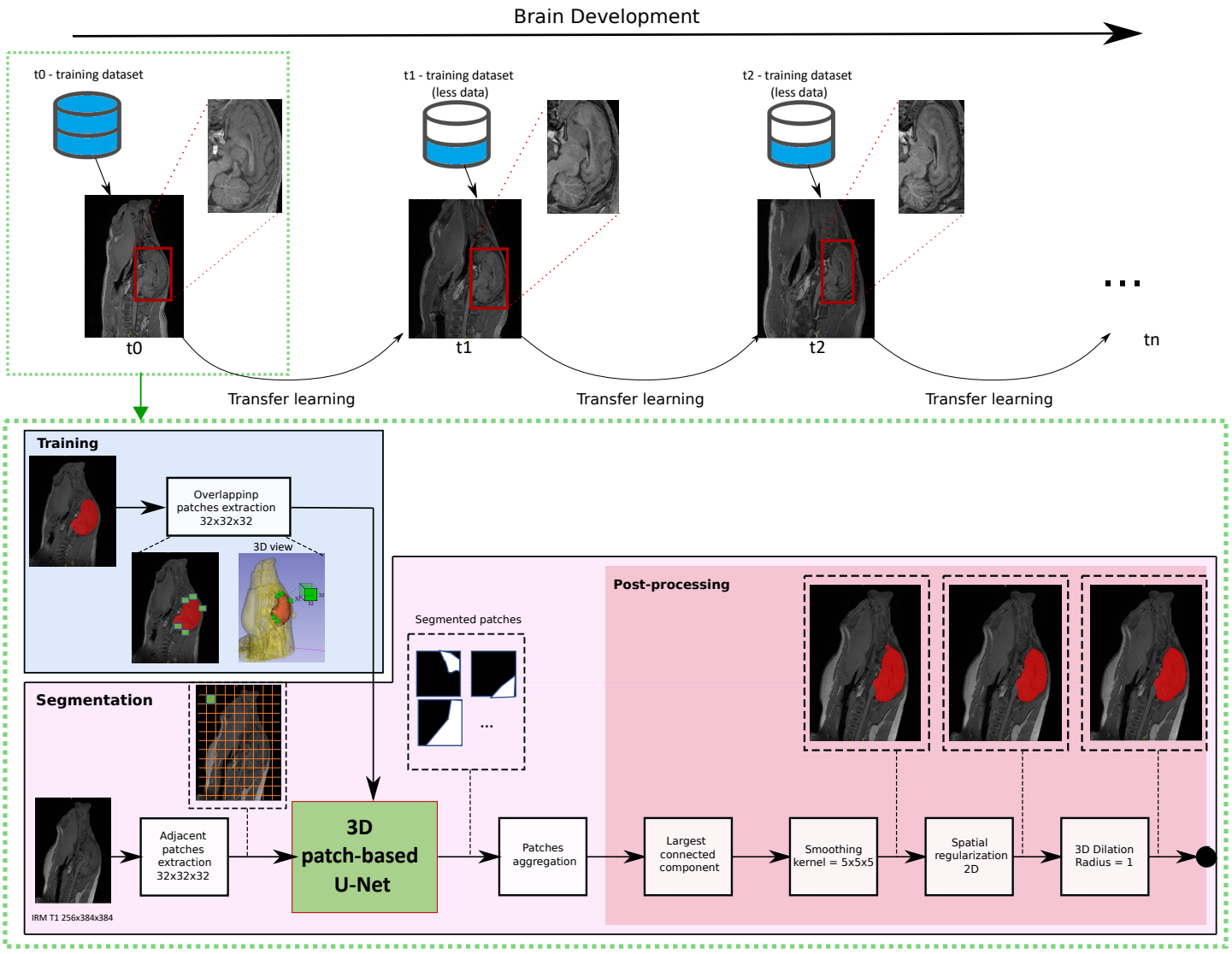

Figure 1: Overview of the proposed automatic piglet brain segmentation method on T1 MRI. The upper part illustrates the principle of transfer learning applied as the brain develops (intensification of contrasts, fat formation, growth). The lower part details the proposed segmentation pipeline (deep learning and post-processing).

approach. However, a preliminary step to these volumetric analyzes remains a challenge: piglet brain segmentation.

Currently, the brain mask is entirely manually drawn by an operator using dedicated software (ITKSnap[1], MITK[2]). This is a long operation (approximately 1h30 per MRI) unsuitable for building large datasets of segmented piglet brains. Therefore, it is necessary to find an alternative to automate the segmentation of the piglet brain on MRI. Our purpose is to propose an efficient method on the piglet but also generic enough to be transposed to other large animal models. Moreover, one expects that this method remains efficient even with relatively small training dataset, without requiring any atlas (i.e. atlas-free approach), compared to other related works [13].

The proposed approach is based on a deep learning semantic segmentation (U-Net [14]). The use of deep neural networks for brain segmentation has already shown its performance on child's brain MRI especially on the fetus MRI. Thus, *Ebner et al.* [15] proposes a pipeline for the reconstruction of the fetal brain starting with a skull stripping using 2D Convolutional Neural Network (CNN) while *Salehi et al.* [16] considers brain segmentation with a 2D U-net. However, the use of 3D Convolutional neural networks for the segmentation of animal brains is still little exploited despite good results in rodents (PCNN3D [17]) or macaques (DIKA-Nets [18]). It has never been used to segment a piglet brain.

In addition, in order to connect general context and local information during deep learning training, we propose a 3D

[1] http://www.itksnap.org/pmwiki/pmwiki.php

[2] https://www.mitk.org/wiki/The_Medical_Imaging_Interaction_Toolkit_(MITK)

patch-based approach, as recently proposed for the analysis of brain human MRI [19, 20]. The 3D patches U-Net with patch selection strategy focusing on the brain contour areas constitutes the first originality of this contribution.

Investigating about the impact of early brain lesion on neurodevelopment and neural plasticity requires a longitudinal monitoring of brain development. For this, it is necessary to provide a solution allowing to accurately segment the brain from MRI at any stage of the maturation process (longitudinal brain segmentation). However, the brain undergoes very rapid and significant transformations in the first months of its existence (intensification of tissue contrasts, formation of fat on the periphery, growth in brain volume) which may reduce the performance of a network trained at a young age [4]. Furthermore, the experimental conditions can sometimes lead to a smaller dataset over time, making it difficult to train a deep neural network at each stage of development. To overcome this challenge, we consider a transfer-learning strategy that leverages the young piglets dataset (2-week-old) to pretrain our 3D U-Net and then transfer this to the small training sample of older piglets. In this way, the brain segmentation at each stage of development is achieved through a pretrained network with the brain characteristics of the previous stage of development. The use of transfer learning for brain segmentation has shown its ability, in particular, to transfer good performance from the human brain to the brains of non-human primates or adult pigs [21]. However, it has never yet been tested for a longitudinal study of the porcine brain, this constituting the second originality of this contribution.

Many techniques already exist for the segmentation of the human brain like BET (Brain Extraction Tool - based on deformable models) [22]. Compared to human brain, being a large structure surrounded by highly contrasted skull, one considers piglet brains, being a small structure surrounded by muscles and fats [12]. Concerning the brain segmentation of other animals, one can mention works related to pigs [21], macaques (DIKA-Net [18], atlasBREX [13]) and rodents (2D U-Net [23], PCNN3D [17], RATS [24] and atlasBREX). Our proposal differs from existing techniques because it considers a 3D patch-based approach (only 2D in [23], 3D without patches in [21, 17]), it is atlas-free (unlike [13]) and addresses the issue of longitudinal monitoring (out of the scope of many related works [17, 24, 21, 23]) by using transfer learning (ignored in [18]).

The dataset and methods are detailed in Section 2. Experiments and results are presented in Section 3 and discussed in Section 4, before concluding this paper in Section 5.

2. Material and methods

Figure 1 gives an overview of the proposed method using an example of T1 MRI of a piglet considered in the study. A first network is trained at 2 weeks (t_0) from the available MRI dataset (see Figure 1 - bottom).

From an annotated training dataset of piglets MRI (depicted on a blue background), 3D overlapping patches of dimension 32^3 voxels are extracted. The extracted patches are used to train a neural network to segment the piglet's brain on MRI. The model considered is a U-Net network receiving as input 3D patches of dimension 32^3 voxels of the T1 MRI and giving as output patches of the same size with the segmentation of the brain. Thus, to infer on a new piglet MRI (depicted on a pink background), we first divide it into adjacent patches of this dimension which will pass through the U-Net network in order to be segmented. The segmented patches constituting the MRI are then aggregated to reconstitute the whole volume and obtain a first version of the brain mask. The latter then goes through a post-processing chain detailed in Section 2.4 consisting of a conservation of the largest connected component, a smoothing and a 2D spatial regularization by conditional random fields as presented in [25, 26]. It is followed by a morphological dilation to provide a precautionary margin for the subsequent stages of segmentation of internal brain structures.

A second U-Net network is trained 4 weeks later (t_1) from a smaller dataset (see Figure 1 - top). To compensate for this little data, a transfer learning mechanism is used to transfer the characteristics learned at 2 weeks (t_0). This logic is repeated at each stage of development to ensure a longitudinal monitoring. This means that at stage t_2 (10 weeks), another U-Net is trained from the available data and the transfer of characteristics learnt at stage t_1 (6 weeks).

We then detail our dataset (Section 2.1), the architecture of the considered U-Net (Section 2.2), the patch extraction strategy chosen for network training and inference (Section 2.3), the post-processing pipeline (Section 2.4) as well as the transfer learning used for longitudinal study (Section 2.5).

2.1. Dataset

Our method is evaluated using 27 piglets (26 males and 1 female). These piglets were followed during their first 4 months of life as part of the REPAR project [7] which seeks to characterize the consequences of cerebral palsy on neurodevelopment using the piglet as an animal model. In this context, an early brain lesion is induced on these piglets at 2 weeks by injection of endothelin. To characterize the impact of the lesion at the macroscopic level, a 3D T1 MRI is acquired at different stages of development (2 weeks, 6 weeks, 10 weeks and 18 weeks) for each piglet. In order to reduce the suffering of the animal linked to the MRI scan, a general anesthesia of the piglets by intramuscular injection of 20 mg/kg of ketamine (Imalgene 1000) and 2 mg/kg of xylazine (Rompun 2%) is implemented. It is then maintained by inhalation of 1% isoflurane conveyed by oxygen and put on respiratory assistance for the animal with a volume of 30 mL and at a frequency of 20/min. In addition, a weaning of the piglet adapted to its young age based on "Friandine" granules is implemented as it is done in breeding.

Due to the induced lesion and sacrifices made to complete the study at a microscopic level (histological sections), some piglets died during the study period. Therefore, the

Table 1
Dataset for each developmental stage studied

Age (weeks)	Number of 3D T1 MRI	Mean brain volume
2	27	53.25 (± 7.20) cm^3
6	17	71.85 (±10.71) cm^3
10	14	92.59 (±5.79) cm^3
18	5	118.5 (±13.6) cm^3

amount of MRI data available decreases with the age of the piglets. Thus, the data used at each stage of development are as detailed in Table 1.

The 3D T1 images are of size 256x384x384 voxels with a spatial resolution of 0.5x0.5x0.5 mm^3. The brain volume is evaluated with the manual delineation carried out with the ITKSnap software.

2.2. U-Net network architecture

The U-Net network is a standard architecture for the segmentation of medical images [14]. The U-Net model is an extension of fully convolutional networks (FCN) [27].

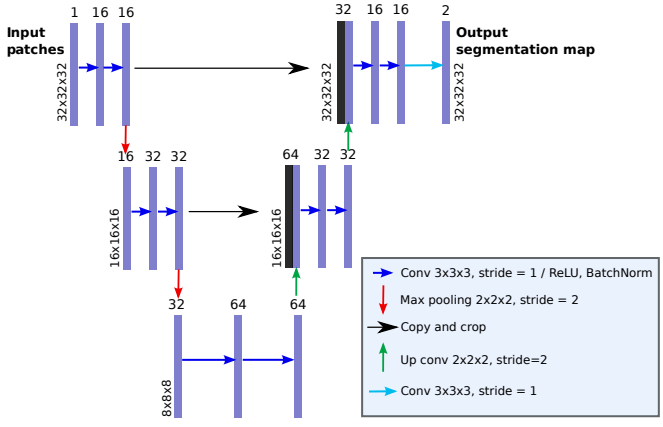

Figure 2: Architecture of the implemented U-Net

The FCN architecture consists of a so-called encoding path where the input is down-sampled using successive convolution and maxpooling operations. The resulting feature map is fed into an activation map to predict the class of each pixel. The U-Net network (Figure 2) also has this encoding path but it is followed by another so-called decoding path almost similar to the encoder path, thus giving the U-shaped architecture. During decoding path, pooling operations are replaced by oversampling operations.

Figure 2 represents the architecture of the implemented U-Net. It receives as input patches of dimension 32^3 and produces two segmentation maps of dimension 32^3 corresponding to the two classes (background and brain) after applying a Softmax function. All convolution operations are done with padding.

2.3. 3D patches strategy

U-net architecture has shown great performance for image segmentation [19]. However, the larger the input image size, the more GPU (Graphics Processing Unit) memory capacity is required for training. Additionally, when the architecture takes the entire MRI or a slice of the MRI as input, the model tends to miss detail in some areas of the image (issue of unbalanced classes [19]). To overcome this problem, we propose a patch-based approach which has the advantage of being more precise (the network can focus on information local to the patch) while requiring less memory for training and inference, thus reducing computation times.

In order to enrich the context information contained in the patch, a good solution consists in using 3D patches which benefit from the context in the 3 directions (axial, coronal and sagittal) [20], compared to a 2D approach focusing on 2 directions only and requiring an additional post-processing step to merge 2D decisions.

Hereafter, we present the 3D patch extraction strategy for both training and inference steps.

2.3.1. Training

For training, we consider overlapping patches (Figure 1) because overlapping patches help the convolutional network to see both the local and global context [20]. In fact, the use of small 3D patches focuses on local information (e.g. 32x32x32 local neighborhood). But the multiplicity of patches covering a large area of the brain brings global information to the network since each area is perceived through several patches, each one providing its local information. Indeed, the prediction of each voxel during training comes from the decision of all the overlapping patches containing this voxel (Eq.2).

For each voxel in each patch, the U-Net model provides a probability of belonging to the brain. The network parameters are adapted to minimize the binary cross-entropy loss function [28].

$$L(Y, \hat{y}) = \sum_i -(Y_i \log(\hat{y}_i) + (1 - Y_i) \log(1 - \hat{y}_i)) \quad (1)$$

where Y_i is the real class of voxel x_i (0 for background, 1 for brain) and $\hat{y}_i$ is the probability of voxel x_i of belonging to the brain. If patches overlap, $\hat{y}_i$ is averaged over all the patches concerned according to the relation:

$$\hat{y}_i = \frac{\sum_{p \in P_i} \text{Softmax}_p(y_i)}{|P_i|} \quad (2)$$

with P_i the set of patches containing x_i and $\text{Softmax}_p(y_i)$ the prediction of belonging to the brain of the voxel x_i in the patch p defined as such:

$$\text{Softmax}_p(y_i) = \frac{\exp^{y_i}}{\exp^{(1-y_i)} + \exp^{y_i}} \quad (3)$$

where y_i means classifying the voxel x_i as brain and $1-y_i$ as background.

Moreover, due to the fact that the brain occupies a relatively small volume of the MRI (issue of the representativity of the classes underlined in [20]), we impose the following distribution of the training patches (Figure 3):

- 85% of patches contain brain
- 15% of patches contain head parts external to the brain

In order to learn more precisely the borders of the brain, 50% of brain patches are located at brain edges.

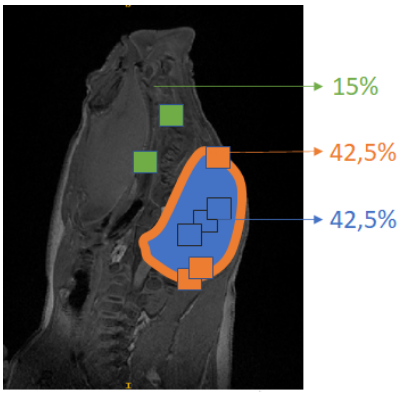

Figure 3: Training patches distribution (*green*: patches outside the brain, *blue*: patches inside the brain, *orange*: brain contours patches)

The selection of patches for each region (from annotated images) is automatic and made according to the following strategy:

- **Patches outside the brain**: For each patch, one randomly select patches center $O(x, y, z)$ in the image. A patch of dimension 32^3 is extracted around O. If the patch contains voxels of either the brain or the background (outside the piglet head), it is rejected. The random selection is repeated until having $0.15 * N$ patches, N being the total number of patches extracted from one MRI.
- **Patches inside the brain**: random selection (uniform distribution) of patches with a central voxel within 30 voxels of the barycenter of the brain. Note that the first extracted patch has the barycenter of the brain as central voxel. The random selection is repeated until having $0.425 * N$ patches.
- **Brain contours patches** : For this part, the brain mask boundary is retrieving on the ground truth to have a binary image of the brain contour. For each patch, one randomly select patches center $O(x, y, z)$ in the image. A patch of dimension 32^3 is extracted around O. The patch is considered as correct if the proportion of voxels located on the brain contour is higher than 1% i.e. if the patch is inside of a belt of the edge. This low percentage is because the contour of the brain is only one voxel in width so the proportion of edge voxels can't be high. The random selection is repeated until having $0.425 * N$ patches.

2.3.2. *Inference*

For inference, we consider adjacent patches in order to simplify the computations. Many studies have shown the relevance of adjacent patches (Figure 1) for segmentation purposes [19, 29].

Each MRI is divided into adjacent 3D patches covering the whole image as illustrated in Figure 1 (bottom). For each voxel x_i of each patch p, the U-Net network returns a probability vector v_i of belonging to each class (background or brain). Each voxel is assigned to the class of greatest probability:

$$\hat{y}_i = \text{argmax}\, v_i \tag{4}$$

The adjacent 3D patches are then aggregated to rebuild the entire image (see Figure 1 - bottom).

2.4. Post-processing

The aggregation of the segmented patches after deep learning produces a first mask of the piglet brain. Nevertheless, this mask is often irregular and may contain false positives. Therefore, some post-processing operations are applied.

2.4.1. *Largest connected component*

This operation eliminates potential false positives from the image by considering the brain as the largest structure detected during deep learning.

The segmentation obtained from deep learning is binary, i.e. it classifies each voxel of the MRI into background and brain, leaving the brain into a number of connected components. Connected component labelling is used to detect connected regions in the binary MRI. A 18-connectivity is considered (Figure 4). Once all the regions classified as brain are detected, the size of each one is determined by counting its number of voxels. Only the largest one is kept, the other regions are assigned to the background class.

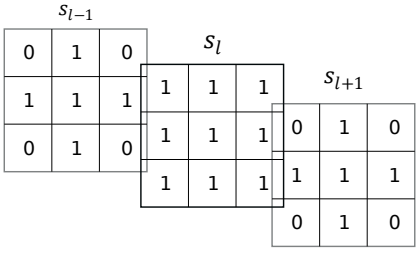

Figure 4: 3D Connectivity/Structuring element considered for connected component labelling/morphological dilation.

2.4.2. *Smoothing*

A first step to correct irregularities of the mask consists in smoothing it. Thus, a median filter of size 5^3 is applied

to the brain segmentation. To deal with edge problems, the segmented 3D volume is zero-padded.

2.4.3. Spatial regularization

The second step to correct spatial irregularities of the mask consists in applying conditional random fields (CRF) as presented in [25]. CRF is a discriminative statistical modelling method used when the class labels for different inputs are not independent. It is therefore interesting to use it in our case because the class label for the voxels depends on the label of its neighboring voxels also. Indeed, if the neighbors of a voxel belong to the brain, there is a high probability that it too.
We consider a 2D spatial regularization, i.e. we apply the conditional random fields slice by slice in order to reduce the processing time. We consider that the previous smoothing is sufficient to regularize the mask from one slice to another.

Let X_i be the variable associated to voxel i which represents the class assigned to voxel i (1 for brain, 0 otherwise). Let X be the vector formed by $X_1, ..., X_N$ with N the total number of voxels of the MRI. Given the image I, the pair (I, X) can be modelled as a CRF characterized by a Gibbs distribution:

$$P(X = x|I) = \frac{1}{Z(I)} \times e^{-E(x|I)} \tag{5}$$

$E(x)$ is the energy of the configuration x and $Z(I)$ the partition function defined in [26]. According to [30], the energy of a label assignment x can be formulated as:

$$E(x) = \sum_i \phi_u(x_i) + \sum_{i,j \in \mathcal{N}_i} \phi_p(x_i, x_j) \tag{6}$$

$\phi_u(x_i)$ is the unary energy component which measures the inverse likelihood of voxel i taking the class x_i and $\phi_p(x_i, x_j)$ is the pairwise energy component measuring the cost of assigning labels x_i, x_j to voxels i and j simultaneously ($\mathcal{N}_i$ being the neighborhood of the voxel i). This latter is modelized as weighted Gaussians [30].
The unary energy ϕ_u is defined from the labelling generated by deep learning. For the pairwise energy ϕ_p defined in [30], only the locations are considered, appearing sufficient for regularization in this context. A Gaussian kernel is considered with a symmetric normalization. The number of inferences is set to 10.

2.4.4. Dilation

Once the false positives eliminated and the mask regularized, a dilation is applied to ensure a margin. More concretely, it is preferable to have a brain mask slightly larger than the brain rather than losing certain regions of edge. Indeed, any loss will result in underestimated volumes of brain tissue in future morphometric analyses.

The 3-dimensional binary dilation is realised with the structuring element represented in Figure 4.

2.5. Transfer learning and longitudinal analysis

Transfer learning has exploded in popularity in the field of machine learning [31] and more recently for brain segmentation issues [21]. One of the main reasons for its success regards its ability to use the knowledge acquired from performing a task to solve a similar one (Figure 5). In fact, transfer learning allows to leverage knowledge (features, weights) from previously trained models for training newer models and thus tackle problems like having less data for the newer similar task. This fits perfectly with our case since we have trained a 3D U-Net model with data from 15-day-old piglets and we want to use it to automate brain segmentation on older piglets, a population for which we have less data. This lack of data often leads to poor performance of traditional machine learning algorithms [31]. Note that one also considers data augmentation as a complementary technique to overcome this lack of representativity [32].

Let's assume that the brain segmentation of piglets at stage $t-1$ corresponds to a T_s task. Let's call T_t the task of segmenting the brains of piglets at stage t. By applying the model trained for T_s to the older piglets, we should face performance degradation for different reasons related to the model's bias towards training data and the fast changes of the piglet brain.

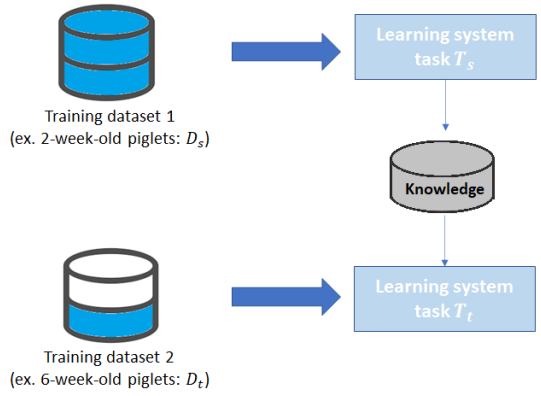

Figure 5: Transfer learning: illustrated principle (s is relative to the source domain (in our case the young piglets) and t to the target domain (the older piglets))

We are going to utilize the learning of T_s, and generalize this knowledge (weights) for task T_t (which has less training data) as illustrated in Figure 5.

Following the notation of [33], we define a domain D as a feature space $X = \{x_1, ..., x_n\}$ (set of grey levels of the image voxels) . A task T is composed of a label space $Y = \{0, 1\}$ (0 corresponds to the background, 1 to the brain label) and an objective predictive function f_t which is learned from the training data at stage t (i.e. pairs of $\{x_i, y_i\}$ where $x_i \in X$, $y_i \in Y$ and index i denotes voxel i). The transfer is done from a source domain and task to a target domain and task.

We will consider the source domain D_s as being that of T1 MRI of piglets at stage $t-1$ (for instance 2 weeks). The target domain D_t will be composed from T1 MRI of piglets at stage t (for instance 6 weeks). According to [34], we face a problem of transductive transfer learning because the domains D_s and D_t are different (different stages of development) but the tasks T_s and T_t are the same (brain segmentation).

To perform transfer learning, some approaches select weights from first layers and adjust only the last layers parameters [35]. This approach is explained by the fact that, according to [36], features are more generic in early layers and more specific in later layers. However, in this case, to avoid arbitrarily splitting the network (first/last layers), all weights Θ_{t-1} are considered during the transfer from one stage to another: the predictive function f_t is initialized with weights Θ_{t-1}. The learning will then consist in adapting the network parameters Θ_t to minimize the loss function:

$$\Theta_t = \underset{\theta}{\operatorname{argmin}} \frac{1}{|D_t|} \sum_{i=1}^{|D_t|} L(f_t(\theta, \Theta_{t-1}, x_{i,t}), y_{i,t}) \qquad (7)$$

with $|D_t|$ the dimension of the domain at stage t and L the loss function namely in this case the binary cross-entropy loss function defined by Eq.1.

Only f_{t_0} is initialized with random weights since it does not have a previous study stage from which to initialize its weights. Then, at stage t_1 (6 weeks), the transfer learning procedure presented above is applied from data of 2-week-old piglets. For the t_2 stage, D_s will consist of the 6-week-old piglets and D_t will consist of the 10-week-old piglets. The network trained for the 6-week-old piglets (from the 6-week-old data and the feature transferred from the 2-week-old piglets) is loaded and re-trained with the 10-week-old piglet MRI. This means that the network trained at 10-week-old will benefit from the transfer of features from the previous two developmental stages. The mechanism is repeated à t_3 (18 weeks) thus ensuring an automatic longitudinal monitoring.

3. Experiments

We present in this section, the evaluation protocol (Section 3.1) as well as the results of our experiments (Section 3.2). In section 3.2, results related to the use of 3D patches, spatial regularization, transfer learning and comparison with state of the art are addressed.

3.1. Evaluation protocol

All the experiments are carried out in a Python environment on 64-bit Windows with an Intel Core i7 @ 2.70 GHz CPU with 32GB of RAM and an Nvidia Quadro RTX 3000 GPU. To implement the U-Net architecture and perform the deep learning experiments, we use the *PyTorch* library[3]. The detailed parameters of the considered network are given in Section 2.2. Spatial regularization with CRF is carried out using the *pydensecrf* library [4].

For the experiments at the first stage of development, that is the 27 two-week-old piglets, we divide our dataset as follows: 7 piglets (and therefore 7 MRI) are used for training, 2 piglets for validation and 18 piglets for test. The model is trained with "Stochastic Gradient Descent (SGD)" on 75 epochs from mini-batches of size 8. A strategy of reduction of the learning rate on plateau is used with an initial learning rate $Lr_0 = 0.1$ and a reduction factor $\sigma = 0.01$. The cross-entropy loss function mentioned in Section 2.3.1 (Eq.1) is considered.

To train the U-Net network and assess the quality of the segmentations obtained, we consider as a reference a manual segmentation of the brain carried out with the ITKSnap software. From a quantitative point of view, we calculate the Dice [37] measuring the similarity between the ground truth I and the predicted segmentation I':

$$\text{Dice}(I, I') = \frac{2|I \cap I'|}{|I| + |I'|} \qquad (8)$$

where $|.|$ represents the cardinality of the set.

We also break down the Dice into precision and recall for some experiments. In fact, in order to subsequently segment the brain structures and tissues, an important criterion is not to loose any part of the brain. For this, we make sure to maximize the recall while ensuring acceptable precision.

Furthermore, the Hausdorff distance (HD) [38] and the average symmetric surface distance (ASSD) [39] between the ground truth and the predicted mask is evaluated. The Hausdorff distance between two sets I and I' is the maximum distance between a point of I and the closest point of I'. It is defined as:

$$\text{HD}(I, I') = \max(\max_{a \in I} \min_{b \in I'} |b - a|, \max_{b \in I'} \min_{a \in I} |a - b|) \quad (9)$$

The average symmetric surface distance between two sets I and I' is the average of all distances between the points of I and the closest point of I' and vice versa (I' to I). The distance of a voxel a to the surface of I' is given by $D_{I'}(a)$.

$$\text{ASSD}(I, I') = \frac{1}{|I| + |I'|} \times (\sum_{a \in I} D_{I'}(a) + \sum_{b \in I'} D_I(b)) \quad (10)$$

The size of the patches used to train the model is important because it affects the amount of considered local information [20]. Thus, we study the impact of the size of the patches considered by comparing the use of patches of size 32^3 and 64^3 voxels.

Likewise, we compare our training patch selection strategy with others in order to define the importance of patch

[3] https://pytorch.org/

[4] https://github.com/lucasb-eyer/pydensecrf

distribution between the brain and the rest of the image as well as the impact of a specific over-training of region boundaries. Three strategies are considered:

- **Patch strategy 1**: Fully random distribution of patches.
- **Patch strategy 2**: 15% of patches outside the brain / 85% of patches uniformly distributed in the region of the brain around the barycenter.
- **Patch strategy 3**: Our strategy (detailed in Section 2.3) that is 15% of the patches outside the brain / 42.5% of the patches uniformly distributed in the region of the brain around the barycenter / 42.5% of the patches focused on the contours of the brain.

Secondly, we study the benefits of considering final post-processing operations (spatial regularization with CRF and morphological dilation).

To show the relevance of the 3D patch-based strategy, we compare the performance of our patch-based 3D U-Net to the use of a 2D U-Net like the one considered for the segmentation of the fetal brain in [16] or the rodents in [23]. As for our 3D patch-based U-Net, the 2D U-Net is trained with "Stochastic Gradient Descent" (SGD) on 75 epochs from mini-batches of size 8. To compensate for the few training data, a data augmentation strategy is implemented with random bias field and random motion to create "realistic" data. This comparison is done with and without the use of conditional random fields in order to highlight the benefit of the patch strategy independently of any spatial regularization.

Then, we compare our method with the Brain Extraction Tool (BET) provided in FSL [5].We justify the choice of BET by the fact that it is a standard for the automatic extraction of the human brain on MRI offering good results [22]. In addition, the tool is based on deformable models and therefore does not depend on the amount of available data, unlike other brain segmentation methods based on a template [13].

We complete our comparison with state of the art by positioning our performances in relation to methods applied to other animal models: macaques [18, 13], rodents [17, 23, 24, 13] and pigs [21]. This report of performance of related works is for information purposes since dataset are different and all third party codes are not freely available [18, 21, 23].

Finally, we address the issue of the fast development of the porcine brain during the first weeks by evaluating our algorithm and the impact of transfer learning on older brains (6 weeks, 10 weeks and 18 weeks). At each stage of development t_i, we compared 3 different strategies in order to find the most efficient one for learning the age-related variability of the brain while taking into account the fact that few MRI are available :

[5] https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/BET/UserGuide

Table 2
Dataset splitting for deep learning. The numerical values correspond to the number of piglets used for each process.

Age	Train	Validation	Test
2 weeks	7	2	18
6 weeks	4	2	11
10 weeks	4	2	8
18 weeks	2	1	2

- **Train strategy 1**: Application of the network trained at stage t_{i-1} (network trained on 15-day-old piglets applied to 6-week data, network trained on 6-week-old piglets applied to 10-week data and so on). It is the strategy to consider that the requested task is the same and that the network trained at stage t_{i-1} will be more efficient due to the larger amount of data. If this strategy works it means that between stages t_{i-1} and t_i, piglet brain varies little.
- **Train strategy 2**: Training of a new network with the few data available at stage t_i. It is the strategy of training a network specific to each stage of development to face age-related variability despite the decreasing number of MRI data as the piglets get older.
- **Train strategy 3**: Our strategy (detailed in Section 2.5) which consists in using transfer learning to supplement the features learned at t_{i-1} by the data of piglets at stage t_i.

For each stage of development, we divide the available dataset (presented in Section 2.1) as detailed in Table 2.

3.2. Results

In this section, we present the results obtained for the experiments according to the previously described protocol. Results for 3D patches, final post-processing benefits and comparisons with state of the art including 2D U-Net and BET are obtained at stage t_0 corresponding to 2-week-old piglets.

3.2.1. Training patches size and distribution

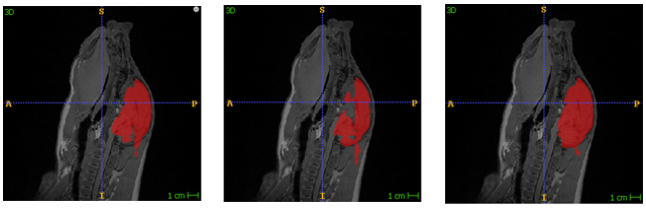

Figure 6: Examples of segmentations obtained on one piglet depending on the size and number of patches used (from left to right 300x64^3, 300x32^3, 700x32^3)

Table 3
Comparison of the training runtime and performance of the U-Net network according to the number and size of patches used (values averaged over 18 piglets). na: not available.

Patches	Dice	HD	ASSD	Training
$300x64^3$	0.918	30.23	2.68	4hrs30
$300x32^3$	0.879	21.24	3.83	**1hrs10**
$700x64^3$	na	na	na	more than 7hrs
$700x32^3$	**0.951**	**20.24**	**1.49**	**2hrs30**

Table 4
Comparison of the performance of the U-Net according to the distribution of training patches (average over 18 piglets)

	Precision	Recall	Dice	HD	ASSD
Strategy 1	0.861	0.653	0.757	24.42	5.81
Strategy 2	0.899	0.911	0.905	45.70	4.31
Strategy 3	**0.959**	**0.943**	**0.951**	**20.24**	**1.49**

Using small patches (32^3) significantly reduces training runtimes compared to larger patches (64^3) as shown in Table 3 (average over 18 piglets). However, they contain less information on the context, which degrades the results as it can be observed in Table 3 and Figure 6. Only the Hausdorff distance is improved by using smaller patches since they concentrate local information with greater precision.

Table 3 also shows that the results are better when increasing the number of small patches per MRI to 700 (probably due to more frequent patch overlap). Patches of size 32^3 allow to reach an average Dice of 0.951 compared to 0.918 with the patches of size 64^3 while almost halving training runtime. The average Hausdorff distance is reduced to 20.24. Finally, the average surface distance is improved (1.49 with 32^3 patches versus 2.68 with 64^3 patches). These results are illustrated by the image to the right in Figure 6 on which the piglet brain is segmented more accurately.

Table 4 compares the performance of the network according to the distribution of the 700 training patches of size 32^3.

We can see that by ensuring that 85% of the training patches contain parts of the brain (strategy 2), we can significantly improve, compared to strategy 1, the performance of the network (+ 0.148 for Dice, -1.50 for ASSD). The recall (since our main goal is to maximize it) is increased by 0.258. However, the average Hausdorff distance is significantly degraded due to the lack of training patches taken outside of the brain area or at its boundary (lack of balanced representativity). This significant degradation of the Hausdorff distance can be explained by the gross errors related to two specific piglets (piglets 15 and 17) as illustrated in Figure 7 - bottom (HD graph). The errors on these two piglets are also reflected in the graph by a smaller Dice.

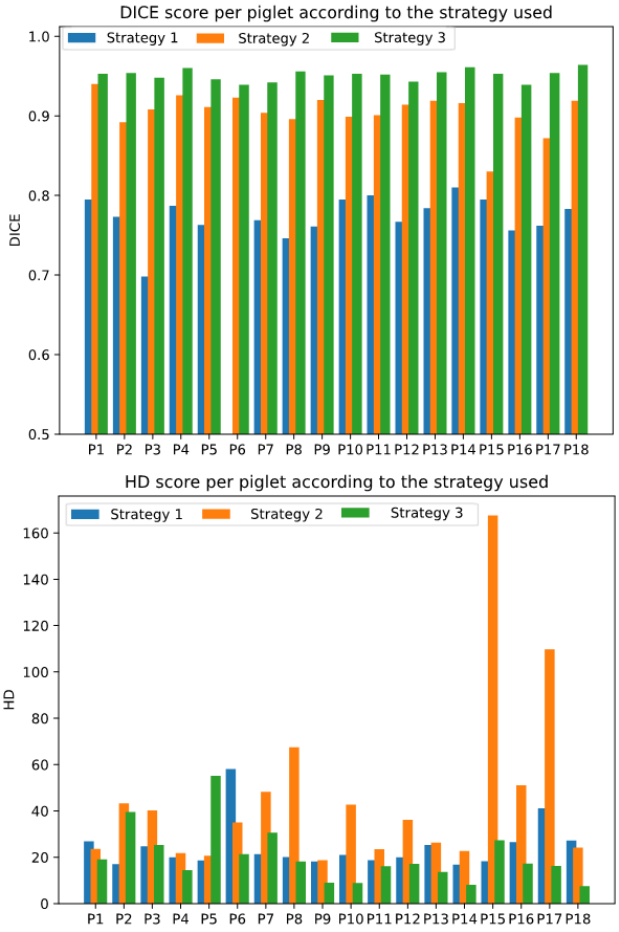

Figure 7: Detailed performance results (Dice and HD) obtained on the eighteen 2-weeks test piglets (P1 to P18) with the three strategies of training patches distribution.

According to Table 4, best results are obtained by considering the contour approach (strategy 3) which further improves segmentation performance (+0.046 for Dice, -2.82 for ASSD, -25.46 for HD compared to strategy 2). In Figure 7, we can clearly see that strategy 3 (our strategy) offers the best Dice for each piglet of the test dataset. Regarding the Hausdorff distance, it gives, on average, the best results even if, for some specific piglets, performance is not so good than with strategy 1 (piglets 2, 5, 7 and 15).

3.2.2. Post-processing: spatial regularization and dilation

In Figure 8, we can observe that spatial regularization using conditional random fields (CRF) corrects the irregularities of the segmentation at brain boundaries. According to the two first lines of Table 5, it appears that the use of CRF is relevant : it halves the Hausdorff distance and reduces the average symmetric surface distance. Adding this operation reasonably increases the MRI processing runtime. However, spatial regularization also degrades the recall making the Dice slightly lower (0.948 vs 0.951 without CRF).

Improving recall is the goal of morphological dilation applied at the end of the process as explained in Section

Table 5
Comparison of brain segmentations obtained with spatial regularization and dilation (IPP: initial post-processing, including largest connected component and smoothing)

	Precision	Recall	Dice	HD	ASSD	Processing runtime
U-Net + IPP	0.959	0.943	0.951	20.24	1.49	1min
U-Net + IPP + CRF	**0.978**	0.918	0.948	9.33	1.43	1min46s
U-Net + IPP + CRF + Dilation	0.953	**0.951**	**0.952**	**8.51**	**1.34**	1min48s

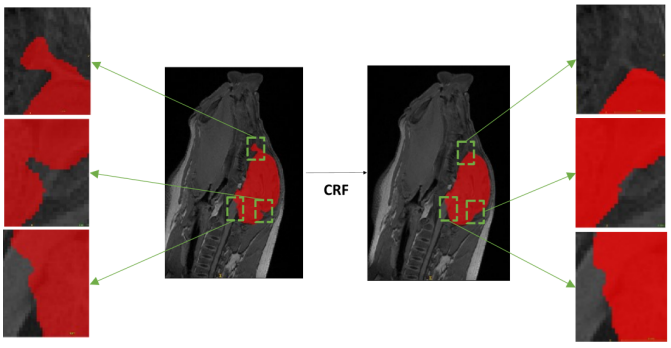

Figure 8: Impact of conditional random fields on brain mask

2.4. In Table 5, we notice that using dilation increases the recall from 0.918 to 0.951 and reduces the HD (8.51) and the ASSD (1.34). This reflects a recovery of certain areas of the brain that is initially lost. We can also note the good balance between precision and recall resulting from morphological dilation.

3.2.3. Comparison with state of the art

All the results are reported in Table 6. Only the 2D U-Net and BET (2nd and 3rd lines of Table 6) have been applied to our dataset (2-week-old piglets are considered) and provide a real comparison with our method. The dataset size of 27 piglets includes the 7 training piglets, the 2 validation piglets and the 18 test piglets. The other results (lines 4 to 9) were obtained on different datasets (less reliable comparison).
The first lines of the table show that our method outperforms 2D U-Net and BET for the segmentation of 2-week-old piglet brains. These results are detailed hereafter.

Table 7 reports the average results obtained on the 18 test piglets with our strategy and with a classic 2D U-Net as used in [16, 23]. The use of a 2D approach without post-processing produces a better Hausdorff distance (13.48) than with a patch-based 3D U-Net (20.24). However, the 3D approach offers a higher Dice (0.951 versus 0.919) and a smaller average symmetric surface distance (1.49 versus 2.16). These results can also be visualized in Figure 9. In fact, we observe with piglet 1 that the 2D approach can lose certain parts of the brain area (the upper part here) unlike the patched-based 3D approach. Nevertheless, both piglets

Table 6
Comparison of the proposed method with state of the art skull-stripping algorithms (T.L.: transfer learning from humans)

Animal	Method	Size dataset	Dice	HD
Piglets	Our method	27	**0.952**	**8.51**
Piglets	BET	27	0.764	25.91
Piglets	2D U-Net	27	0.919	11.06
Macaques	DIKA-Net [18]	155	**0.964**	**1.47**
Macaques/ Rodents	atlasBREX [13]	8/6	0.950	/
Rodents	PCNN3D [17]	6	0.930	/
Rodents	2D U-Net [23]	132	0.940	**6.81**
Rodents	RATS [24]	22	0.920	13.6
Pigs	T.L. [21]	3	0.930	/

Table 7
Comparison of the performance of the 2D U-Net network and our patch-based 3D U-Net (mean calculated over the 18 test piglets). FPP: final post-processing including spatial regularization and dilation. Conservation of the largest connected component and smoothing is performed in all cases.

	Dice	HD	ASSD
2D U-Net	0.919	13.48	2.16
patch-based 3D U-Net	0.951	20.24	1.49
2D U-Net with FPP	0.919	11.06	2.14
patch-based 3D U-Net with FPP	**0.952**	**8.51**	**1.34**

depict the irregularities produced by the 3D approach (at the front of the brain in particular and also at the back for piglet 1) which can considerably alter the Hausdorff distance as seen in the numerical results.

Adding the last steps of post-processing made up of a spatial regularization by CRF and a unit radius dilation modifies only very slightly the performances of the strategy in 2D with a reduction of the Hausdorff distance to 11.06. Conversely, the application of this post-processing to the 3D patch-based approach halves the average Hausdorff distance and reduces the ASSD. In Figure 9, we can visualize in the last line the effects of post-processing. We can clearly see that the latter does not correct the segmentation errors of

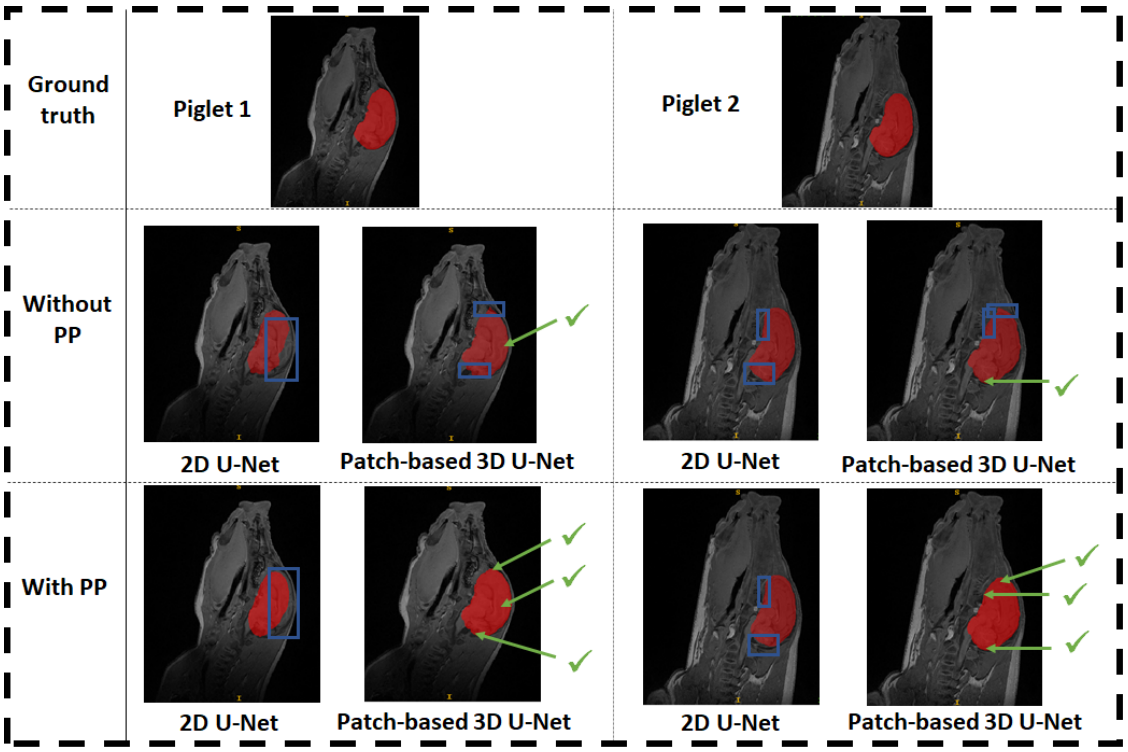

Figure 9: Examples of segmentations obtained on two piglets with a 2D U-Net and a patch-based 3D U-Net (PP : final post-processing including spatial regularization and dilation)

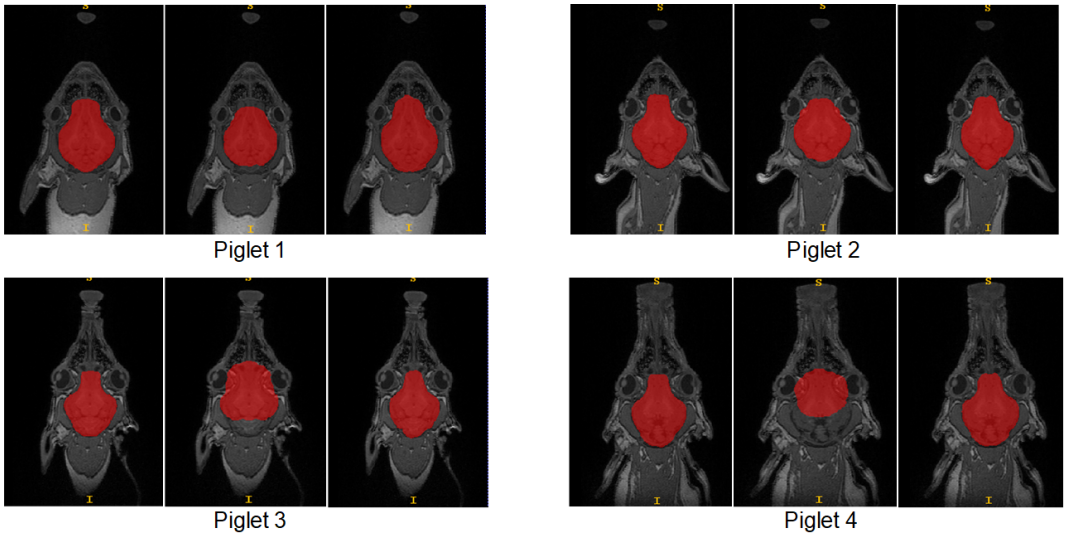

Figure 10: Examples of segmentations in coronal view for the first 4 piglets. From left to right: ground truth, BET, our method.

the 2D approach on piglet 1. On the contrary, it smooths the irregularities of the segmentation obtained by our method leading to a clear reduction in the Hausdorff distance. As a result, our proposal clearly surpasses the 2D approach, considered for rodents (line 7 of Table 6).

Table 8 gives the results obtained with our method and with the BET tool for the 18 test piglets. We note that our method outperforms BET for most piglets with an improvement of all considered metrics (Dice, HD and ASSD). BET obtains a better Hausdorff distance than our method only for one piglet (piglet 13) with a slight difference (13.96 with BET versus 14.32 with our method). Moreover, except for piglet 3, the estimated brain volume is closer to the actual brain volume (measured from the ground truth).

Figure 10 gives examples of segmentation obtained. We note that our method manages to segment the piglet's brain with more precision and consistency than BET, which tends to lose parts or to incorporate external areas into the brain mask (piglets 3 and 4).

Table 8
Results obtained with BET and our method on the 18 test piglets (the exact ages are given only for information).

	Age	Brain volume	Estimated brain volume		Dice		HD		ASSD	
			BET	Our method	BET	Our method	BET	Our method	BET	Our method
Piglet 1	16 days	$53.45cm^3$	$45.20cm^3$	**$55.03cm^3$**	0.890	**0.959**	13.64	**7.28**	2.91	**1.16**
Piglet 2	13 days	$48.81cm^3$	$46.04cm^3$	**$49.57cm^3$**	0.870	**0.955**	14.18	**9.43**	3.44	**1.25**
Piglet 3	13 days	$46.05cm^3$	**$45.83cm^3$**	$44.76cm^3$	0.680	**0.956**	29.07	**8.48**	8.52	**1.21**
Piglet 4	15 days	$52.40cm^3$	$38.76cm^3$	**$54.70cm^3$**	0.409	**0.959**	49.16	**5.38**	16.49	**1.19**
Piglet 5	15 days	$50.82cm^3$	$32.94cm^3$	**$51.99cm^3$**	0.675	**0.948**	37.01	**8.60**	8.33	**1.45**
Piglet 6	16 days	$54.94cm^3$	$48.54cm^3$	**$52.89cm^3$**	0.892	**0.956**	11.04	**6.40**	2.90	**1.25**
Piglet 7	15 days	$48.72cm^3$	$43.14cm^3$	**$49.61cm^3$**	0.861	**0.946**	16.03	**8.54**	3.55	**1.48**
Piglet 8	16 days	$53.82cm^3$	$26.51cm^3$	**$52.45cm^3$**	0.639	**0.944**	48.88	**9.43**	9.72	**1.59**
Piglet 9	16 days	$52.34cm^3$	$42.76cm^3$	**$52.40cm^3$**	0.846	**0.952**	15.26	**6.71**	4.02	**1.33**
Piglet 10	14 days	$50.68cm^3$	$31.07cm^3$	**$49.67cm^3$**	0.713	**0.949**	36.50	**10.30**	7.14	**1.38**
Piglet 11	14 days	$51.16cm^3$	$42.42cm^3$	**$50.04cm^3$**	0.868	**0.955**	15.56	**8.06**	3.38	**1.23**
Piglet 12	14 days	$50.11cm^3$	$42.87cm^3$	**$50.88cm^3$**	0.876	**0.950**	13.00	**8.31**	3.11	**1.33**
Piglet 13	15 days	$55.67cm^3$	$47.64cm^3$	**$53.72cm^3$**	0.883	**0.936**	**13.96**	14.32	3.09	**1.78**
Piglet 14	15 days	$53.98cm^3$	$33.94cm^3$	**$54.66cm^3$**	0.632	**0.950**	44.15	**9.85**	9.57	**1.45**
Piglet 15	16 days	$53.67cm^3$	$42.78cm^3$	**$54.70cm^3$**	0.842	**0.960**	24.06	**5.92**	4.10	**1.15**
Piglet 16	15 days	$54.23cm^3$	$39.68cm^3$	**$51.83cm^3$**	0.811	**0.945**	19.44	**11.57**	4.86	**1.56**
Piglet 17	15 days	$59.64cm^3$	$30.22cm^3$	**$57.07cm^3$**	0.581	**0.958**	39.32	**8.66**	10.99	**1.22**
Piglet 18	15 days	$52.38cm^3$	$37.11cm^3$	**$53.09cm^3$**	0.785	**0.963**	26.09	**6.00**	5.49	**1.06**

Regarding the other algorithms whose performances were obtained on other datasets (lines 4 to 9 of Table 6), only the DIKA-Net developed for the macaque significantly outperforms our method (Dice of 0.964 and HD of 1.47). Nevertheless, the dataset used is much larger (155 MRI of macaques) than the available one for piglets (27 MRI only). The only other method that surpasses our proposal, in terms of Hausdorff distance, is the 2D U-Net applied to rodents (HD of 6.81). However, the latter also has a large dataset of 132 rodents. Moreover, we shown previously that the 2D approach is less efficient on piglets than the patch-based 3D one.
Our approach achieves performances equivalent to those based on an atlas (atlasBREX) and exceeds RATS with a comparable dataset size (22 MRI). Regarding the application on pigs, our method offers better performances than the solution proposed by [21].

3.2.4. Transfer learning for longitudinal study

Table 9 reports performances obtained for the three training strategies (introduced in Section 3.1) to segment piglet brain at each stage of development i.e. 6, 10 and 18 weeks. The first row corresponding to stage t_0 (2 weeks) is given as a reference. Obviously, because it is the first stage studied, strategies 1 and 3 are not applied.

The first column (Strategy 1) shows the average results obtained at each stage of development t_i by applying the 3D U-Net trained at stage t_{i-1}. For stages t_1 and t_2, there is an alteration in segmentation performance compared to the reference results obtained at t_0 (first row). Indeed, all the metrics are weakened (-0.061 for Dice, +6.91 for HD, +1.8 for ASSD at t_1 and -0.058 for Dice, +22.58 for HD, +2.26 for ASSD at t_2). This drop in performance is confirmed by Figure 11 applying the U-Net trained at t_0 to the same piglet at t_0 (15 days) and t_1 (6 weeks) respectively. We can see that our network is still able to recognize the location of the brain in the 6-week MRI. Nevertheless, although it delineates it correctly, it loses significant parts of brain area (in particular the back of the cerebellum here, surrounded in Figure 11).

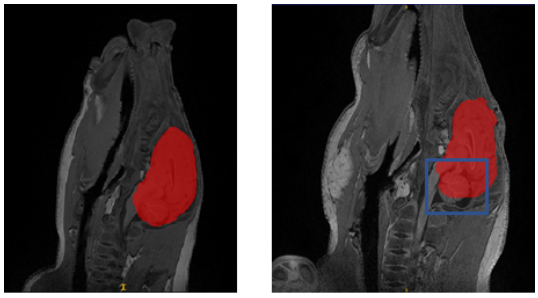

Figure 11: Examples of segmentations obtained for the same piglet at 15 days (left) and 6 weeks (right) with the patch-based 3D U-Net trained on the 15-day data.

These results confirm our hypothesis according to which characteristics learned at one stage (tissue intensity, size)

Table 9
Quality of the brain segmentations at each stage of development according to the training strategy used (Strategy 1: U-Net trained at stage t_{i-1}, Strategy 2: U-Net only trained at stage t_i , Strategy 3: transfer learning of the network trained at t_{i-1} and supplemented by data of stage t_i). Mean values are calculated over piglets of the test dataset (i.e. 18 piglets at t_0, 11 piglets at t_1, 8 piglets at t_2 and 2 piglets at t_3). Strategy 2 is not applied at stage t_3 because of the few number of available data.

	Train strategy 1			Train strategy 2			Train strategy 3		
	Dice	HD	ASSD	Dice	HD	ASSD	Dice	HD	ASSD
t_0: 2 weeks (reference)	-	-	-	**0.952**	**8.51**	**1.34**	-	-	-
t_1: 6 weeks	0.891	15.42	3.14	0.636	82.58	42.34	**0.934**	**10.72**	**1.99**
t_2: 10 weeks	0.894	31.09	3.60	0.907	44.76	4.15	**0.956**	**16.17**	**1.53**
t_3: 18 weeks	0.954	12.86	1.62	-	-	-	**0.958**	**12.20**	**1.47**

vary fast for the same piglet with it brain development reducing network performance. Such variation in terms of contrast is visible (especially in the region of the cerebellum) in the upper part of Figure 1 through the zooms of the brain at each stage of development.
However, the results obtained at 18 weeks (stage t_3) with strategy 1 are equivalent to the performance achieved at t_0. These good results may traduce a slowing down of the piglet's brain development between 10 and 18 weeks.

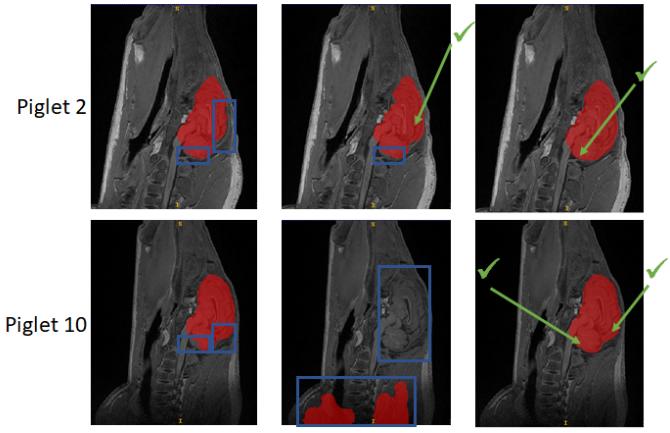

Figure 12: Comparison of segmentations obtained on two piglets according to the training strategy considered (from left to right : 2-week network, 6-week network, transfer learning)

We will then exploit the data available at stage t_i to train the network (columns of strategies 2 and 3 in Table 9). Several observations can be made from this table.

First of all, strategy 2 fails, considerably degrading all the metrics (-0.255 for Dice, +67.16 for HD, +39.20 for ASSD at t_1 and -0.045 for Dice, +36.25 for HD, +2.81 for ASSD at t_2). The very poor results at 6 weeks are explained by a false identification of the brain region on 3 of the 11 inference piglets as shown in Figure 13 (piglets 4, 9 and 10). Indeed, these 3 piglets have a zero Dice and an ASSD greater than 100, worsening the average results. We can see on the second line of Figure 12 corresponding to piglet 10, that the network trained at 6 weeks detects a brain in the bottom of the image due to the presence of similar contrasts. These gross errors are most likely due to a lack of training data (only 4 piglets).

Then, we can still observe that when the 6-week network does not make a gross error in locating the brain, the results are half the time improved (piglets 1, 2, 6 and 8) as shown in Figure 13. In Figure 12, we can see that the network trained at 6-weeks recovers some brain regions of piglet 2 compared to the network trained at 15 days. These results confirm the importance of exploiting the few number of 6-week data to take account of age-related variability in the learning process.

Table 9 shows that the use of transfer learning improves the average of all the metrics compared to strategies 1 and 2. This improvement is valid for all stages of development and provide a satisfactory longitudinal monitoring to correctly segment the piglet's brain at each of its developmental stages as shown in Figure 14. Let's take the stage t_1 to illustrate these improvements. Figure 13 evidenced the improvement of all metrics on each of the test piglets. We can observe the clear improvement in segmentations with the help of Figure 12 where, for the two piglets considered, the segmentation obtained by transfer learning (right image) is the one including the brain with the most precision.

We can end this study by highlighting the slightly poorer results obtained on the older piglets (despite the use of transfer learning) compared to those obtained at 2 weeks. For each stage, the Hausdorff distance and the average symmetric surface distance are slightly reduced (HD of 10.72 at t_1, 16.17 at t_2 and 12.20 at t_3 against 8.51 at t_0, ASSD of 1.99 at t_1, 1.53 at t_2 and 1.47 at t_3 against 1.34 at t_0). Nevertheless, these are satisfactory results that could be improved with more MRI data at each stage of development. One can note the very slight improvement of the Dice at t_2 and t_3 encouraging the significance of the proposed method.

4. Discussion

This work illustrates the ability of the patch-based 3D U-Net network to accurately segment the piglet's brain on T1 MRI, in particular by focusing the training patches at brain boundaries. Furthermore, the proposed method has

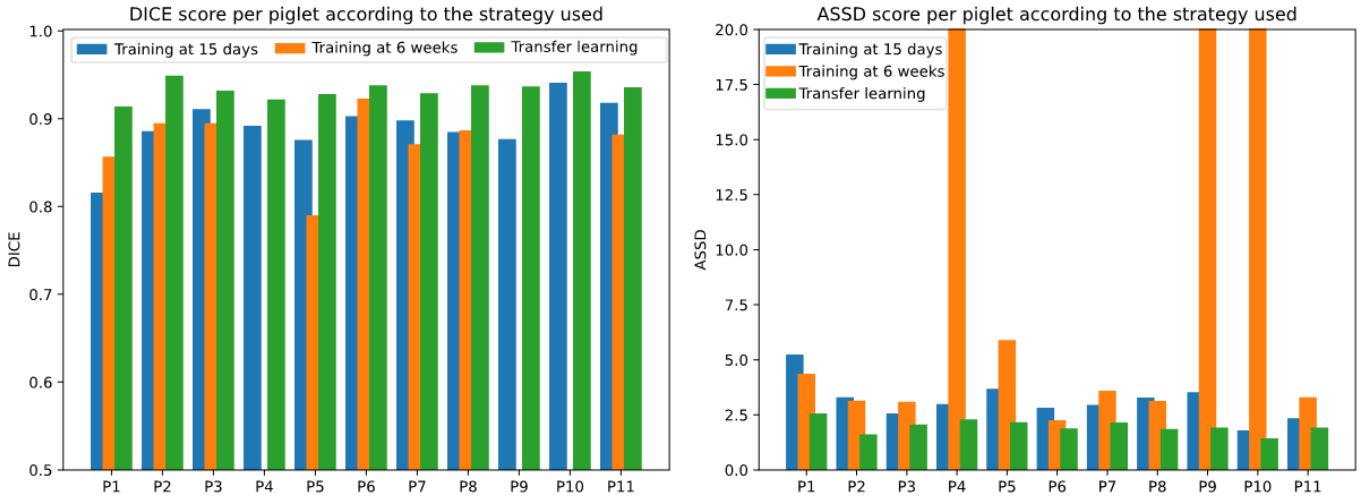

Figure 13: Detailed performance results (Dice and ASSD) obtained on the eleven 6-week test piglets with the three strategies of training

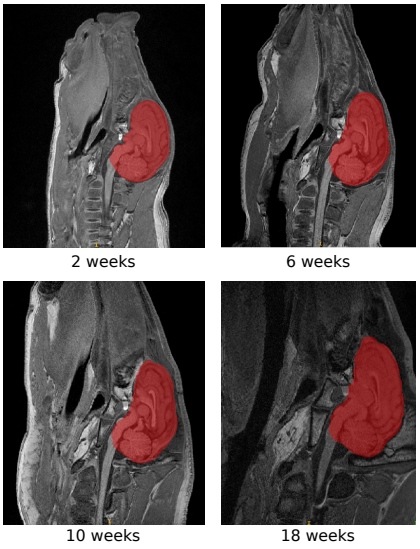

Figure 14: Segmentation of the piglet brain at each stage of development obtained with transfer learning (strategy 3)

.

the advantage of operating from a relatively small training dataset (only 7 training piglets at 2 weeks), a considerable advantage for studies on the large animal model for which little data is available. We also find a balance in the size of the patches in order to combine information related to the global context (set of patches covering a large area of the MRI) and local characteristics in an acceptable computation time. The large number of training patches extracted by MRI (700) compensates for the modest dataset size while ensuring more systematic overlaps making the prediction more precise. Unlike other solutions, no dedicated preprocessing step is required (e.g. image noise, bias field in [21]) and no a-priori information are required (e.g. center of gravity distance map, signed distance map in [18]).

In addition, our method is able to correct segmentation errors linked to the small number of data thanks to a post-processing pipeline including image processing and spatial regularization with conditional random fields (CRF). It should be noted that CRFs could be fully integrated into the U-Net model as recently proposed [40] in order to benefit from backpropagation during training. The post-processing would thus be directly applied on a regularized mask. This could also probably allow to get rid of the first post-processing steps (largest connected component and smoothing). In addition, using overlapping patches instead of adjacent ones during inference could compensate for the uncertainty of prediction related to edge effects and thus reduce irregularities in segmentation [20]. This would constitute an alternative to the proposed post-processing.

We have shown that the association of the patch-based 3D U-Net and the final post-processing provides better results than a fully 2D approach as often used [16, 23].

From all of these properties, we see that our solution undoubtedly exceeds, in the case of piglets, human-specific brain segmentation tools such as BET. More generally, we have seen that the proposed method exceeds other atlas-free methods proposed for different animals such as PCNN3D or RATS. Moreover, it offers slightly better performances than atlas-based techniques like atlasBREX. Nevertheless, the difference in results with DIKA-Net, mainly due to the significantly smaller size of our dataset in our sense, suggests that improvements (e.g. increasing the size of the training dataset) can be made to reduce the Hausdorff distance. It would be interesting to evaluate the relevance of the proposed transfer learning strategy on DIKA-Net (ignored in their work), for longitudinal monitoring.

In addition, we propose a solution for the longitudinal study of the piglet brain through transfer learning. The reduced performance of the network trained at state $t-1$ on piglets at the development stage t (Strategy 1 of Section

3.2.4) confirm the rapid change in the porcine brain during the first months of existence [4]. Note that this dynamic of brain development is already observed in two-week-old piglets whose ages fluctuate from 13 to 16 days with variations in volume from 46.05cm^3 to 59.64cm^3 (see Table 8). We observe that on average we obtain better results on piglets with a developmental stage closer to the training data (average brain volume of $53.25cm^3$). Obviously, these developmental variations are also unique to each individual.

Despite a very small training dataset (only 4 piglets at 6 weeks and 10 weeks, 2 training piglets at 18 weeks), we succeeded in obtaining good brain segmentation performance on older piglets from the network trained at previous stages. The success of the transfer learning approach presented in this paper can be explained by the similarity of the overall structure and the proportion of tissues in the piglet brain regardless of age. The features learned at state $t-1$ remain partly valid. The morphological variabilities linked to growth (intensities, volume) are learned from the small number of training piglets available at the studied stage of development. Results from Table 9 at 18 weeks suggest that the piglet brain slows down its development between 10 and 18 weeks. Indeed, the application of the trained network at 10 weeks on 18-week-old piglets offers good performances (Dice of 0.954, HD: 12.86, ASSD: 1.62) almost equivalent to those obtained with transfer learning. It would be interesting to continue the study at later developmental stages in order to confirm this hypothesis and to define more precisely the limit age from which the brain development of the piglet slows down and no longer necessarily requires new training data.
In case we have much less data for one of the developmental stages (e.g. 2 weeks), the strategy of transfer learning by increasing age could be discussed. Indeed, it could be possible to train the network at state t (e.g. 6 weeks) by transferring the characteristics of state $t+1$ (e.g. 10 weeks) for which more data are available. Nevertheless, from a medical point of view, this would not make sense since we study brain development, i.e. we know the brain characteristics at a given stage and we complete this knowledge with the new characteristics that appear with the growth. In fact, preliminary studies testing the reverse approach (by decreasing age) on 6-week-old piglets confirmed the unnaturalness of the approach with poorer results (Dice of 0.856 when transferring characteristics from 10-week piglets to 6-week piglets).

Although our approach is promising, certain limitations may be pointed out. First, the performance was evaluated on a relatively small dataset of 39 MRI used for testing purposes (18 at 2 weeks, 11 at 6 weeks, 8 at 10 weeks and 2 at 18 weeks). A greater number of MRI, especially at advanced stages, may have enabled us to obtain better results. Nevertheless, our complete dataset is of the same order of magnitude as other works dealing with piglets. Thus, the study of *Gan et al.* [4] involves 40 piglets and *Conrad et al.* [12] uses only 15 piglets while we use in total 27 MRI at 2 weeks, 17 MRI at 6 weeks, 14 MRI at 10 weeks and 5 MRI at 18 weeks. As neonatal studies on large animals are restrictive, we can consider our dataset to be satisfactory.

Our method is applied on piglets here. It would be interesting to use it on other animal models increasingly studied such as sheep[41], macaque[42], ferret[43] or dog[44].

5. Conclusions

In this paper, we have proposed an original method for segmenting piglet brain on T1 MRI using patch-based 3D deep learning, with a relevant patch selection strategy. Results show how this approach obtains a good quality of segmentation (average Dice of 0.952, HD of 8.51 and ASSD of 1.34), outperforming 2D approaches and atlas-free techniques specific to other species (humans (BET), rodents (PCNN3D, RATS)). We also proposed an efficient solution to the longitudinal study of the brain based on the use of transfer learning, tested at three developmental stages (a fourth one being the reference from which transfer is achieved).

Perspectives for this work are numerous. Indeed, from the segmented brain, we can consider the segmentation of the piglet's brain structures and tissues in order to set up various studies (morphometric analyzes, analyzes of structural information [45, 46]) on healthy and injured brains that will characterize the impact of early brain lesion related to cerebral palsy.

The generic methodology based on iterative transfer learning would be useful for studying the longitudinal brain development. Moreover, this method would be transposable to other types of brains and thus facilitate the study of various animal models essential to medical research [47].

Acknowledgements

This work was supported by the "Institut National de la Santé et de la Recherche Médicale" (INSERM) and by the University of Angers (Angers, France). This research is part of the REPAR project funded by the "Angers-Loire Métropole" and by "INSERM Transfert". We are thankful to the "INRA Centre Val de Loire" for assistance and care of the animals during surgery and throughout the experiments. We thank Laurence Sindji for her technical assistance.

References

[1] S. Chabrier, M. Pouyfaucon, A. Chatelin, Y. Bleyenheuft, J. Fluss, V. Gautheron, C. J. Newman, G. Sébire, P. Van Bogaert, C. Vuillerot, S. Brochard, M. Dinomais, From congenial paralysis to post-early brain injury developmental condition: Where does cerebral palsy actually stand?, Annals of Physical and Rehabilitation Medicine 63 (2019) 431–438. doi:10.1016/j.rehab.2019.07.003.

[2] C. Cavarsan, M. Gorassini, K. Quinlan, Animal models of developmental motor disorders: parallels to human motor dysfunction in cerebral palsy, Journal of Neurophysiology 122 (2019) 1238–1253. doi:10.1152/jn.00233.2019.

[3] J. Coq, M. Kochmann, D. Lacerda, H. Khalki, M. Delcour, A. Toscano, F. Cayetanot, M. Canu, M. Barbe, M. Tsuji, From cerebral palsy to developmental coordination disorder: Development of preclinical rat models corresponding to recent epidemiological changes, Annals of Physical and Rehabilitation Medicine 63 (2020) 422–430. doi:10.1016/j.rehab.2019.10.002.

[4] H. Gan, Q. Zhang, H. Zhang, Y. Chen, J. Lin, T. Kang, J. Zhang, F. A. Troy, B. Wang, Development of new population-averaged standard templates for spatial normalization and segmentation of MR images for postnatal piglet brains, Magnetic Resonance Imaging 32 (2014) 1396–1402. doi:10.1016/j.mri.2014.08.036.

[5] K. L. Thibault, S. S. Margulies, Age-dependant material properties of the porcine cerebrum: effect on pediatric inertial head injury criteria, Journal of Biomechanics 31 (1998) 1119–1126. doi:10.1016/s0021-9290(98)00122-5.

[6] C. N. Devi, A. Chandrasekharan, V. K. Sundararaman, Z. C. Alex, Automatic segmentation of infant brain MR images: With special reference to myelinated white matter, Biocybernetics and Biomedical Engineering 37 (2017) 143–158. doi:10.1016/j.bbe.2016.11.004.

[7] E. Mazerand, C. Gallet, L. Le Fournier, M. Dinomais, C. Montero-Menei, P. Menei, Development of a porcine model of cerebral palsy: first step before a study on the efficacy of a stereotactic stem cell therapy, in: European Academy of Childhood Disability, 2019.

[8] A. Makropoulos, S. J. Counsell, D. Rueckert, A review on automatic fetal and neonatal brain MRI segmentation, NeuroImage 170 (2018) 231–248. doi:10.1016/j.neuroimage.2017.06.074.

[9] H. Watanabe, F. Andersen, C. Z. Simonsen, S. M. Evans, A. Gjedde, P. Cumming, MR-Based Statistical Atlas of the Göttingen Minipig Brain, NeuroImage 14 (2001) 1089–1096. doi:10.1006/nimg.2001.0910.

[10] C. Norris, J. Lisinski, E. McNeil, J. W. VanMeter, P. VandeVord, S. M. LaConte, MRI brain templates of the male Yucatan minipig, NeuroImage 235 (2021) 118015. doi:10.1016/j.neuroimage.2021.118015.

[11] S. Saikali, P. Meurice, P. Sauleau, P. A. Eliat, P. Bellaud, G. Randuineau, M. Vérin, C. H. Malbert, A three-dimensional digital segmented and deformable brain atlas of the domestic pig, Journal of Neuroscience Methods 192 (2010) 102–109. doi:10.1016/j.jneumeth.2010.07.041.

[12] M. S. Conrad, B. P. Sutton, R. N. Dilger, R. W. Johnson, An In Vivo Three-Dimensional Magnetic Resonance Imaging-Based Averaged Brain Collection of the Neonatal Piglet (Sus scrofa), PLoS One 9 (2014) 1396–1402. doi:10.1371/journal.pone.0107650.

[13] J. Lohmeier, T. Kaneko, B. Hamm, M. R. Makowski, H. Okano, atlasBREX: Automated template-derived brain extraction in animal MRI, Scientific Reports 9 (2019) 12219. doi:10.1038/s41598-019-48489-3.

[14] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional networks for biomedical image segmentation, Lecture Notes in Computer Science 9351 (2015) 234–241. doi:10.1007/978-3-319-24574-4_28.

[15] M. Ebner, G. Wang, W. Li, M. Aertsen, P. A. Patel, R. Aughwane, A. Melbourne, T. Doel, S. Dymarkowski, P. De Coppi, A. L. David, J. Deprest, S. Ourselin, T. Vercauteren, An automated framework for localization, segmentation and super-resolution reconstruction of fetal brain MRI, NeuroImage 206 (2020) 116324. doi:10.1016/j.neuroimage.2019.116324.

[16] S. S. M. Salehi, S. R. Hashemi, C. Velasco-Annis, A. Ouaalam, J. A. Estroff, D. Erdogmus, S. K. Warfield, A. Gholipour, Real-time automatic fetal brain extraction in fetal MRI by deep learning, in: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), 2018, pp. 720–724. doi:10.1109/ISBI.2018.8363675.

[17] N. Chou, J. Wu, J. Bai Bingren, A. Qiu, K. Chuang, Robust automatic rodent brain extraction using 3-D pulse-coupled neural networks (PCNN), IEEE Transactions on Image Processing (2011) 54–64doi:10.1109/TIP.2011.2126587.

[18] T. Zhong, F. Zhao, Y. Pei, Z. Ning, L. Liao, Z. Wu, Y. Niu, L. Wang, D. Shen, Y. Zhang, G. Li, DIKA-Nets: Domain-invariant knowledge-guided attention networks for brain skull stripping of early developing macaques, NeuroImage 227. doi:10.1016/j.neuroimage.2020.117649.

[19] B. Lee, N. Yamanakkanavar, J. Y. Choi, Automatic segmentation of brain MRI using a novel patch-wise U-net deep architecture, PLoS One 15 (2020) 20 pages. doi:10.1371/journal.pone.0236493.

[20] M. Ben naceur, M. Akil, R. Saouli, R. Kachouri, Fully automatic brain tumor segmentation with deep learning-based selective attention using overlapping patches and multi-class weighted cross-entropy, Medical Image Analysis 63 (2020) 101692. doi:10.1016/j.media.2020.101692.

[21] X. Wang, X. H. Li, J. W. Cho, B. E. Russ, N. Rajamani, A. Omelchenko, L. Ai, A. Korchmaros, S. Sawiak, R. A. Benn, P. Garcia-Saldivar, Z. Wang, N. H. Kalin, C. E. Schroeder, R. C. Craddock, A. S. Fox, A. C. Evans, A. Messinger, M. P. Milham, T. Xu, U-net model for brain extraction: Trained on humans for transfer to non-human primates, NeuroImage 235 (2021) 118001. doi:10.1016/j.neuroimage.2021.118001.

[22] S. M. Smith, Fast robust automated brain extraction, Human Brain Mapping 17 (2002) 143–155. doi:10.1002/hbm.10062.

[23] L.-M. Hsu, S. Wang, P. Ranadive, W. Ban, T.-H. H. Chao, S. Song, D. Cerri, L. Walton, M. Broadwater, S.-H. Lee, D. Shen, Y.-Y. I. Shih, Automatic Skull Stripping of Rat and Mouse Brain MRI Data Using U-Net , Frontiers in Neuroscience 14 (2020) 935. doi:10.3389/fnins.2020.568614.

[24] I. Oguz, H. Zhang, A. Rumple, M. Sonka, RATS: Rapid Automatic Tissue Segmentation in rodent brain MRI, Journal of Neuroscience Methods 221 (2014) 175–182. doi:10.1016/j.jneumeth.2013.09.021.

[25] P. Krähenbühl, V. Koltun, Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials, Advances in Neural Information Processing Systems 24 (2011) 109–117.

[26] J. Lafferty, A. McCallum, F. Pereira, Conditional random fields: Probabilistic models for segmenting and labeling sequence data, in: Proceedings of the Eighteenth International Conference on Machine Learning, 2001, pp. 282–289.

[27] E. Shelhamer, J. Long, T. Darrell, Fully Convolutional Networks for Semantic Segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence 39 (2017) 640–651. doi:10.1109/tpami.2016.2572683.

[28] S. Mannor, D. Peleg, R. Rubinstein, The cross entropy method for classification, in: In Proceedings of the 22nd international conference on Machine learning, 2005. doi:10.1145/1102351.1102422.

[29] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P. M. Jodoin, H. Larochelle, Brain tumor segmentation with Deep Neural Networks, Medical Image Analysis 35 (2017) 18–31. doi:10.1016/j.media.2016.05.004.

[30] S. Zheng, S. Jayasumana, R.-P. B., V. Vineet, Z. Su, D. Du, C. Huang, P. H. Torr, Conditional random fields as recurrent neural networks, IEEE International Conference on Computer Vision (2015) 1529–1537doi:10.1109/ICCV.2015.179.

[31] K. Weiss, T. Khoshgoftaar, D. Wang, A survey of transfer learning, Journal of Big Data 3 (2016) 40 pages. doi:10.1186/s40537-016-0043-6.

[32] A. Garcia-Garcia, S. Orts-Escolano, S. Oprea, V. Villena-Martinez, P. Martinez-Gonzalez, J. Garcia-Rodriguez, A survey on deep learning techniques for image and video semantic segmentation, Applied Soft Computing 70 (2018) 41–65. doi:10.1016/j.asoc.2018.05.018.

[33] S. J. Pan, Q. Yang, A Survey on Transfer Learning, IEEE Transactions on Knowledge and Data Engineering 22 (2010) 1345–1359. doi:10.1109/TKDE.2009.191.

[34] H. Bertrand, Hyper-parameter optimization in deep learning and transfer learning : applications to medical imaging, Ph.D. thesis, Université Paris-Saclay (2019).

[35] K. You, Z. Kou, M. Long, J. Wang, Co-Tuning for Transfer Learning, in: Advances in Neural Information Processing Systems, Vol. 33, 2020, pp. 17236–17246.

[36] J. Yosinski, J. Clune, Y. Bengio, H. Lipson, How Transferable Are Features in Deep Neural Networks?, in: Proceedings of the 27th International Conference on Neural Information Processing Systems, Vol. 2, 2014, pp. 3320–3328.

[37] A. P. Zijdenbos, B. M. Dawant, R. A. Margolin, A. C. Palmer, Morphometric analysis of white matter lesions in MR images: method and validation, IEEE Transactions on Medical Imaging 13 (1994) 716–724. doi:10.1109/42.363096.

[38] M. Beauchemin, K. P. B. Thomson, , G. Edwards, On the Hausdorff distance used for the evaluation of segmentation results, Canadian Journal of Remote Sensing 24 (1998) 3–8. doi:10.1080/07038992.

1998.10874685.

[39] V. Yeghiazaryan, I. Voiculescu, Family of boundary overlap metrics for the evaluation of medical image segmentation, Journal of Medical Imaging 5 (2019) 015006. doi:10.1117/1.jmi.5.1.015006.

[40] L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A. Yuille, DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs, IEEE Transactions on Pattern Analysis and Machine Intelligence 40 (2018) 834–848. doi:10.1109/TPAMI.2017.2699184.

[41] A. Ella, M. Keller, Construction of an MRI 3D high resolution sheep brain template, Magnetic Resonance Imaging 33 (2015) 1329–1337. doi:10.1016/j.mri.2015.09.001.

[42] Y. Balbastre, D. Rivière, N. Souedet, C. Fischer, A. S. Herard, S. Williams, M. Vandenberghe, J. Flament, R. Aron-Badin, P. Hantraye, J. F. Mangin, T. Delzescaux, A validation dataset for Macaque brain MRI segmentation, Data in Brief 16 (2018) 37–42. doi:10.1016/j.dib.2017.11.008.

[43] E. B. Hutchinson, S. C. Schwerin, K. L. Radomski, N. Sadeghi, J. Jenkins, M. E. Komlosh, M. O. Irfanoglu, S. L. Juliano, C. Pierpaoli, Population based MRI and DTI templates of the adult ferret brain and tools for voxelwise analysis, NeuroImage 152 (2017) 575–589. doi:10.1016/j.neuroimage.2017.03.009.

[44] X. Liu, R. Tian, Z. Zuo, H. Zhao, L. Wu, Y. Zhuo, Y. Q. Zhang, L. Chen, A high-resolution MRI brain template for adult Beagle, Magnetic Resonance Imaging 68 (2020) 148–157. doi:10.1016/j.mri.2020.01.003.

[45] J. B. Fasquel, N. Delanoue, A graph based image interpretation method using a priori qualitative inclusion and photometric relationships, IEEE Transactions on Pattern Analysis and Machine Intelligence 41 (2019) 1043–1055. doi:10.1109/TPAMI.2018.2827939.

[46] J. Chopin, J. B. Fasquel, H. Mouchère, R. Dahyot, I. Bloch, Semantic image segmentation based on spatial relationships and inexact graph matching, in: International Conference on Image Processing Theory, Tools and Applications, 2020, pp. 1–6. doi:10.1109/IPTA50016.2020.9286611.

[47] F. Barré-Sinoussi, X. Montagutelli, Animal models are essential to biological research: issues and perspectives, Future Science OA 1 (2015) 3 pages. doi:10.4155\%2Ffso.15.63.