


# Semantic Segmentation of Hippocampal Subregions With U-Net Architecture

Soraya Nasser, Université Oran 1, Algeria

Moulkheir Naoui, Université Oran 1, Algeria

Ghalem Belalem, Université Oran 1, Algeria

 <https://orcid.org/0000-0002-9694-7586>

Saïd Mahmoudi, Mons University, Belgium

## ABSTRACT

The automatic semantic segmentation of the hippocampus is an important area of research in which several convolutional neural networks (CNN) models have been used to detect the hippocampus from the whole cerebral MRI. In this paper, the authors present two convolutional neural networks. The first network (hippocampus segmentation single entity [HSSE]) segmented the hippocampus as a single entity, and the second was used to detect the hippocampal sub-regions (hippocampus segmentation multi-class [HSMC]), these two networks inspire their architecture of the U-net model. Two cohorts were used as training data from NITRC (neuroimaging tools and resources collaboratory) annotated by ITK-SNAP software. The authors analyze this network alongside other recent methods that do hippocampal segmentation. The results obtained are encouraging and reach dice scores greater than 0.84.

## KEYWORDS

Convolutional Neural Network (CNN), Hippocampus, ITK-SNAP Software, NeuroImaging Tools and Resources Collaboratory (NITRC), Pixel Classification, Semantic Segmentation, Training, Up-Convolution

## INTRODUCTION

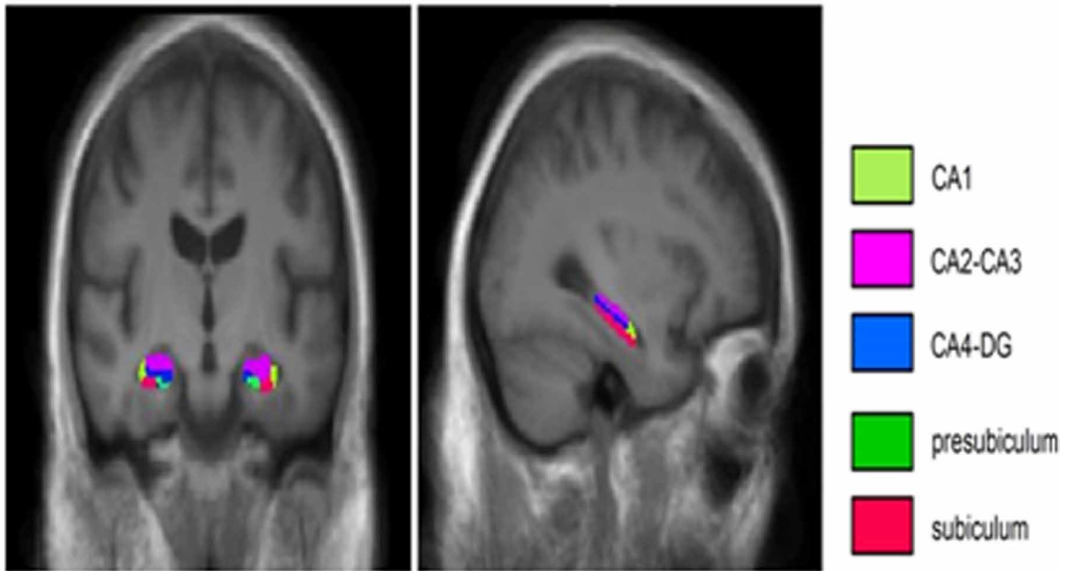
The hippocampus is a structure that is anatomically complex. This consists of two cortex coiled around each other with gray matter, cornu ammonis (CA) and dentate gyrus (DG), separated by the hippocampal groove (Henry, et al., 2011). The *Cornu ammonis* is also composed of four CA areas (1-4) each one is composed of different variable layers in cell composition: the Pyramidal Stratum (SP) is richer in the neuronal bodies than the Radiatum (SR) Strata, Lacunosum and Moleculare (SRLM) are poorer in the neuronal bodies (Cury., 2015); (Boutet, et al., 2014). The subiculum is the hippocampal formation's most subordinate component. This sits between the normal hippocampus entorhinal cortex and the normal hippocampus subfield CA1 (see Figure 1).

Hippocampus function is important in cognitive processes such as short-term processing of memory, thinking, spatial navigation and behavioral inhibition. The hippocampus is one of the first structures affected in Alzheimer's disease (Bobinski et al., 1999), epilepsy and schizophrenia (Koolschijn et al., 2010). It is characterized by atrophies in its shape, the volume of which would be a risk factor. This atrophy of the hippocampus can be assessed from magnetic resonance imaging (MRI) using manual segmentation like in (Kerchner, et al., 2010); (Henry, et al., 2011); (Boutet, et al.,

DOI: 10.4018/IJEHMC.20211101.oa4

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

Figure 1. Hippocampus anatomy (Györfi et al., 2017)



2014). The hippocampus was manually segmented according to the anatomy described by (Duvernoy, 2005). The sub-regions are the Ammon horn, specifically the CA1, CA2, CA3 fields, CA4 and the dentate gyrus (DG) (Figure 1).

Wisse et al. (2012, 2014) proposed a new manual segmentation protocol for hippocampal sub-regions.

However, a lot of work has been done to develop manual segmentation to an automatic segmentation that provides a quantitative and qualitative estimate of this complex structure who they include the notion of registration and deformation using atlas(es) (Chupin, et al., 2009; Kim et al., 2011).

In contrast, an effort was made to systematically analyze the anatomical variation of the hippocampus from large data sets by estimating the iterated barycenter of the population as a model for the analysis of anatomical variability by principal component analysis (PCA) on initial moment vectors or an approximate distance matrix (Cury, 2015).

Our research aims at achieving an automatic semantic segmentation evaluating the hippocampus as a single entity from the MRI on the one side and the hippocampal sub-regions on the other.

This paper will be organized as follows, we review two states-of-the-art, the first concerning the semi-automated and automated hippocampus segmentation and the second one in the field of medical images segmentation using CNNs model (Related work section). Then we detail our approach via two convolutional neural networks (Method section). Finally, the results are reported and discussed (Experiment, Results sections). The conclusion and perspective are described at the end (Conclusion section).

## RELATED WORK

The diagnosis is performed in a clinical context; the study of the morphological features of the hippocampus is based on a visual examination by an often biased neuro-radiologist.

On the other hand, semi-automated segmentation approaches are more advanced, such as those based on deformable models and the use of atlases. Both approaches use precompiled geometric

information that can be created from datasets on real subjects. User interaction with these strategies only occurs at the initial stage, which significantly eliminates manual activity.

(Ashton et al., 1997) proposed a method based on Active Contour Model (ACM).

The Active Shape Model (ASM) method was proposed by Ghanei et al. (1998), Kelemen et al. (1999), and Shen et al. (2002) for the hippocampus segmentation. Hybrids approach for the segmentation of the hippocampus combines the individual atlas (Single-Atlas) and the ACM, was developed by Schnabel et al. (1999). In the method presented by De Alejo et al. (2003), Multiple Layer Perceptron (MLP) neural networks are used in two stages of classification. In the first step, a neural network performs unsupervised voxel classification monitoring, and in the second step, the user performs manual segmentation of the image in a hippocampal section. The characteristics of this segmentation are used as parameters for the second neural network, which performs a supervised classification on the result of the previous step. On the other hand, so-called automatic methods have been proposed, such as the hybrid Single-Atlas approach with the deformable models as in Klemencic et al. (2004). The atlas is aligned with the target image. The regions of the hippocampus transferred to the target image function as the initial contour for an Active Appearance Model (AAM).

Heckemann et al. (2006) used 30 atlases of individuals. Certain approaches are based around the use of probabilistic atlas. The method consists in aligning a series of atlases in a standard space in this standard space, the statistics of the structure of interest, such as the probability of occurrence of image intensities of the global location and adjacent structures can be reported. However, these works require adequate atlases, and depend strongly on the quality of the registration process and takes around 8 hours in a single volume.

On the other hand, few methods have used artificial neural networks such as De Alejo (2003) and Nasser et al. (2018) which has classified through an artificial neural network that integrates the fuzzy ANFIS system (Adaptive Network Based Fuzzy Inference System). Learning was applied on 75% of the original image taken randomly; the test was done by the remaining part of the original image. The ANFIS remains limited by the input parameters (the size of the database) the thing that makes it inconvenient in the segmentation of 3D images.

It is advised to use convolutional neural networks (ConvNet or CNN) in order to provide a strong neural network that can automatically extract a maximum of characteristics from a large database so it can generalize many cases (hippocampal forms).

CNN is a type of machine learning method that has a methodology similar to the traditional supervised learning methods: they receive input images, detect the characteristics of each of them, and then form a classifier on them. They are fed by increasing amounts of data.

Unlike supervised learning techniques, convolutional neural networks automatically learn complex representative characteristics directly from the data themselves (Isin et al., 2016). This is where their strength lies: networks do all the work of automatic features extraction, unlike learning techniques. Nowadays, many works have used CNN precisely in image processing for classification reasons such as Kruthika et al. (2019) which proposed a Content Based Image Retrieval (CBIR) system using the 3D Network Capsule, the convolutional 3D neural network and the technology pre-trained 3D automatic coder for early detection of Alzheimer's disease.

Although deep learning methods for medical image analysis have been used successfully before it has been successfully applied on cardiac imaging (Margeta et al., 2015; Poudel et al., 2016), kidney (Thong et al., 2016); its specific use for 3D volume data has become popular, precisely in brain images such as in Cicek et al. (2016), which presents a volumetric segmentation network that results in low annotated volumetric images. The network extends the previous U-net architecture of Ronneberger et al. (2015). Xu et al. (2018) proposed a fast automatic method of segmentation of different, tissue 3D MRI images of new born brains; it uses a Fully Convolutional Neural networks (FCN) and the transfer of learning and pseudo-3D information.

In addition, the CNN has also made it possible to locate pathological brain lesions (Brosch et al., 2016; Ghafoorian, et al., 2016; Kamnitsas et al., 2017) or tumors (Havaei, et al., 2017). In this

work, they explore a variety of architectures by using the concatenation of feature maps from different layers as another operation when composing CNNs. This operation allows to construct architectures with multiple computational paths.

Ben Naceur et al. (2018) offer 3 Deep Learning 2CNet, 3CNet and EnsembleNet models that combine segmentation results obtained from 2CNet and 3CNet models using a non-parametric fusion function, they propose a new strategy called "ELOBA \_" to train the CNNs models, in which it takes into account the most influencing hyper-parameters (Epochs, Learning rate, Optimizer, Batch size), the most influencing hyper-parameters (Epochs, Learning rate, Optimizer, Batch size).

Roy et al. (2018) used a multiple CNN approach for whole brain segmentation. QuickNAT has an encoder/decoder like 2D F-CNN architecture with four encoders and decoders separated by a bottleneck layer. The final layer is a classifier block with softmax. Its methodology follows a consensus of three 2D F-CNN specialized in each slice views: coronal, axial, and sagittal.

Nowadays, the hippocampus segmentation using deep learning has known great success. Cao, et al. (2017) proposed a multi-task deep learning (MDL) method for joint hippocampus segmentation and clinical score regression using MR images, the proposed CNN network contains two subnetworks, including a U-Net with a Dice-like loss function for hippocampus segmentation, and a convolutional neural network with a mean squared loss function for clinical regression, this study used 407 subject.

Thyreau et al. (2017) proposed a deep learning to segment the 3D hippocampus IRMs as a single entity, the method begins by recording the images on a region of interest of fixed size, and then the hippocampus region is introduced to the network which produces a probabilistic segmentation map, this study used more than 10000 volumes with augmentation.

Carmo et al. (2020) used the U-net model modified with 2D IRMs, the modification consists in adding patch inputs and residual connections between the input and output of the double convolutional block.

Looking at these recent works, U-net is the popular choice used in the segmentation of cerebral (tumor) and hippocampal MRIs, whether it is in its original form or modified (Cicek et al., 2016; Roy et al., 2018; Thyreau et al., 2017; Carmo et al., 2020), which gives very encouraging results.

Nasse et al. (2020) have designed a network similar to the net without concatenation which aims to detect the hippocampal sub-regions the results were encouraging but not sufficient.

On the other hand, the works cited (Cao, et al., 2017; Thyreau et al; 2017; Carmo et al., 2020) segment the hippocampus as a single entity from a histological point of view, this is why we found that it is time to find a whole methodology to determine the hippocampal sub-regions through a semantic segmentation.

We propose in this article, two automatic semantic hippocampal segmentations using the U-net model (Ronneberger et al., 2015); the first one (HSSE: Hippocampus Segmentation Single Entity) is used to segment the hippocampus as a single entity, and the second one (HSMC: Hippocampus segmentation Multi-Class) detects hippocampal sub- regions in 2D MRI.

## METHOD

We have formed two models of convolutional neural networks to segment the hippocampus. In this section, we detail our training data and the preprocessing applied, then we describe the architectures networks and its training procedure.

## DATA SET

We evaluate our experiences with a real healthy and patient subject dataset, obtained from the NeuroImaging Tools & Resources Collaboratory (NITRC)<sup>1</sup>. It is a free web resource offering comprehensive information on software and neuroinformatics data. Our training data contains real images from subsets of several 3D MRI cohorts from (NITRC). As a supervised learning system, a

training data requires a pair of images (a hippocampal image and the same segmented image with a expert annotation), tagged images are made by ITK-SNAP<sup>2</sup>

The selected subsets databases from (NITRC) are:

A. **Nimg2009 Atlas:**

Contains Images of five 3D hippocampus samples (3 right, 2 left) were acquired on a 9.4 Tesla with their segmentation (Yushkevich et al., 2009).

B. **Atlas-UMCU-15-9-17:**

Contains multiple directories, we are interested by train. This is a set of atlases that can be used with the ASHS VERSION 1.0 software (September 2017; Yushkevich et al., 2015) to automatically segment the subregions of the 7Tesla MRI hippocampus. The scans were manually segmented by (Wisse, et al., 2016). Weighted scans in T2 have a resolution of  $0.35 \times 0.35 \times 0.35 \text{ mm}^3$  (26 subjects).

Figure 2 shows some samples of the cohorts cited.

## PRE-PROCESSING

For optimal use of CNN, it is recommended to go through a preprocessing stage.

Firstly, since our dataset do not possess an isotropic resolution and the spacing in the third dimension is not consistent, we consider performing the segmentation slice by slice (2D) from the axial view.

The choice of the size is taken according to the real size of the images while keeping the information by discarding some slices on the z-axis and we decrease the background size on the x and y axis (Clear Label) that do not contain information (Remove the 1% highest and lowest intensities). In addition, the reduced MRIs size minimizes the memory space and the learning time, so we resize the images at (128x128). Then, we apply an N4ITK bias correction to all MRIs (Avants, 2009).

## NETWORK ARCHITECTURES

We develop two multi-classes convolutional neural network based on U-net architecture (Ronneberger et al., 2015); the first one (Figure 3) aims to segment the hippocampus as a single entity while the second (Figure 4) is used to detect hippocampal sub-regions.

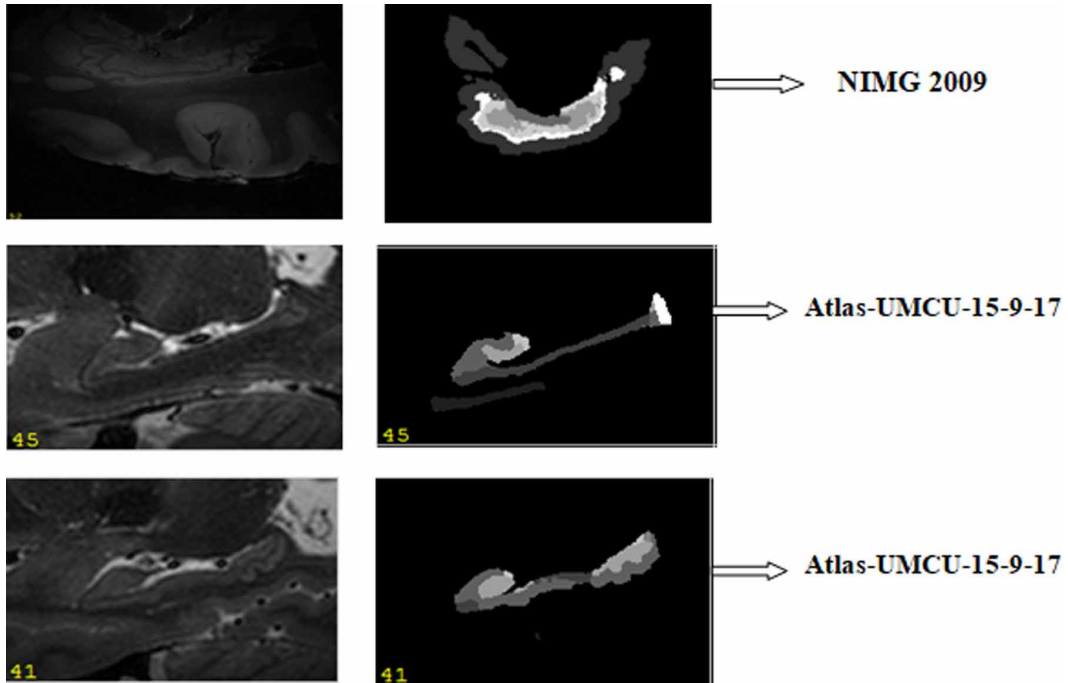
The U-net model is very clear in both figures (3 & 4). It has an encoding and a decoding path with four encoders and four decoders separated by a bottleneck layer.

### 1. The Encoding Path

It is constituted of four 2D convolutional layers each layer contains two(3x3) convolutions layers followed by a rectified linear unit (ReLU: Rectified Linear Unit), and then a (2x2) max pooling layers with strides of one in each dimension and auto padding(same). We choose to use small filters. Small filters (3x3) have proven in (Ben Naceur et al., 2018; Thyreau et al., 2017; Cicek et al., 2016) that they keep a lot of information for our hippocampus image, and at the same time they reduce the number of operations. Also, we double the number of filters (64, 128, 256, 512) after each pooling layer (max-pooling (2x2)), these techniques allow the model to be more deeper in terms of layers .

To speed convergence and accelerate learning, batch normalization was applied after each convolution layer

Figure 2. Some samples of hippocampus MRI from our cohorts and their segmentation. From left to right: original image and the second image is the segmented image.



## 2. The Decoding Path

It is consisted of 4 up-convolution layers (2x2) each one concatenates with the corresponding feature map from the encoding path, halving the number of filters (feature maps) in each stage, followed by two (3x3) convolutions. Similarly, each convolution is followed by a ReLu function. The concatenation is used to combine the location information of the encoding path with the contextual information in the decoding path to finally obtain general information combining location and context, which is necessary to predict a good segmentation map.

## 3. The Bottleneck Layer

It is consisted of two convolutional layers (3x3) with 1024 feature maps.

## 4. The Classifier Layer

In the last layer a (1x1) convolution reduces the number of output channels to the number of labels, after that, the output result are activated by an activation function (Sigmoid or softmax), the choice of this function depends on the objective of the network. In our case we have two objectives to achieve:

### Automatic Semantic Segmentation of the Hippocampus as a Single Entity (HSSE Network)

In this part, we consider the segmentation as a binary classification problem with two class (background and hippocampus region), we used the sigmoid function as classifier of the network output, this

function aims to limit the output to a range between 0 and 1. We then convert the probability mask into a binary map using a simple thresholding process where chose 0.5 as a threshold, whose values below the threshold of 0.5 are assigned to a background and values greater than or equal to 0.5 are assigned to the hippocampus.

### Automatic Semantic Segmentation of the Hippocampal Sub-Regions (HSMC Network)

In this part, we consider segmentation as a multi-class classification problem where it allows fine inference by making predictions inferring labels for each pixel. The number of neurons in this layer is the number of hippocampus sub-regions. The best choice for predicting the network output is the softmax activation function, this function will impose that the sum of the probabilities of our output classes is equal to one, so in order to increase the probability of a particular class, the model must accordingly decrease the probability of at least one of the other classes.

#### 5. Loss Function

The kernel parameters represent the weights that need to update to get the most suitable configuration, they are found by minimizing a loss function between prediction and ground truth labels.

We used in the HSSE network (Figure3) the Binary Cross Entropy as loss function which it is often accompanied with the sigmoid activation function.

In binary classification, where the number of classes  $M=2$ , cross-entropy can be calculated as:

$$-(y \log(p) + (1 - y) \log(1 - p))$$

- $M$  is a number of classes.
- $y$  is a binary indicator (0 or 1) if class label  $c$  is the correct classification for observation  $o$ .
- $p$  is a predicted probability observation  $o$  is of class  $c$ .

As for the HSMC network (Figure4) we used Categorical cross-entropy loss function which it is often accompanied with the softmax activation function, where  $M>2$  (multiclass classification), we calculate a separate loss for each class label per observation and sum the result.

$$-\sum_{c=1}^M y_{o,c} \log(p_{o,c})$$

The optimization of two networks weights is done using the Adam optimizer.

#### 6. Dropout Layer

The U-net model is a vast and complicated network of which it contains a lot of neurons and of weights, the thing which can pose problems of over-fitting, to overcome this we have to add two layers of dropout, the first layer located in the last convolution block in an encoding path before maxpooling and the second one was added in bottleneck layer. The dropout rate is set to 50%

The addition of dropout layers from the last encoding path layer is used to store the inputs data information.

## TRAINING

CNN's architectures usually require a big and varied dataset to achieve generalization, data augmentation increases the training dataset by simulating different variations of the same data, in our case the dataset size is increased two-fold by an augmentation (rotation and zoom), We use random magnitude of rotation within a range of  $\pm 20$  degrees and a resolution of 0.5 degrees, and the zoom operation is performed within a range  $[0.8, 1]$  and resolution of 0.05.

We implement HSSE and HSMC by creating a dataset of mini-batches of 32 (2D images) for each network paired with the corresponding center segmentation label as the target. We choose this small volume for mini-batches because it has been shown that small batch sizes tend to avoid settling on a local minimum (Keskar et al., 2016), the number of epochs is 100 epoch for HSSE network and 250 for HSMC network.

The core CNN was implemented on Tensorflow via the keras library using CuDNN 5.1, trained on a NVIDIA GeForce GTX 1050 GPU.

## EXPERIMENT

As indicated above (pre-processing section) each network has been trained by 90 IRM 3D of size (128x128x80) what it means about 7200 2D IRM (including data augmentation).

We divide the training and testing sets in the ratio of 80: 20% respectively.

The trend of binary cross-entropy loss for training and testing sets are shown in (Figure 3) for HSSE network, and categorical cross-entropy loss for training and testing sets are shown in (Figure4) for HSMC network.

We observe that the loss saturates after 30 epoch for HSSE network and after a sixtieth epoch for HSMC network, and the models exhibit comparable loss performance for both training and testing sets. We choose the HSSE and HSMC model with the lowest validation loss for our subsequent experiments.

## RESULTS

We have evaluated our experiments on real subject's data, in which we have got this dataset from the different cohorts from the NeuroImaging Tools & Resources Collaboratory (NITRC).

To obtain the quantitative evaluation of the models performance, we put the segmentation results under a series of different tests: accuracy, Dice (identical to F-measure), and recall (Sensitivity) for each hippocampus region, the evaluation metrics are calculated as follows:

The sensitivity is intuitively the ability of the classifier to find all the positive samples. The best value is 1 and the worst value is 0.

$$Sensitivity = \frac{TP}{TP + FN} \quad (3)$$

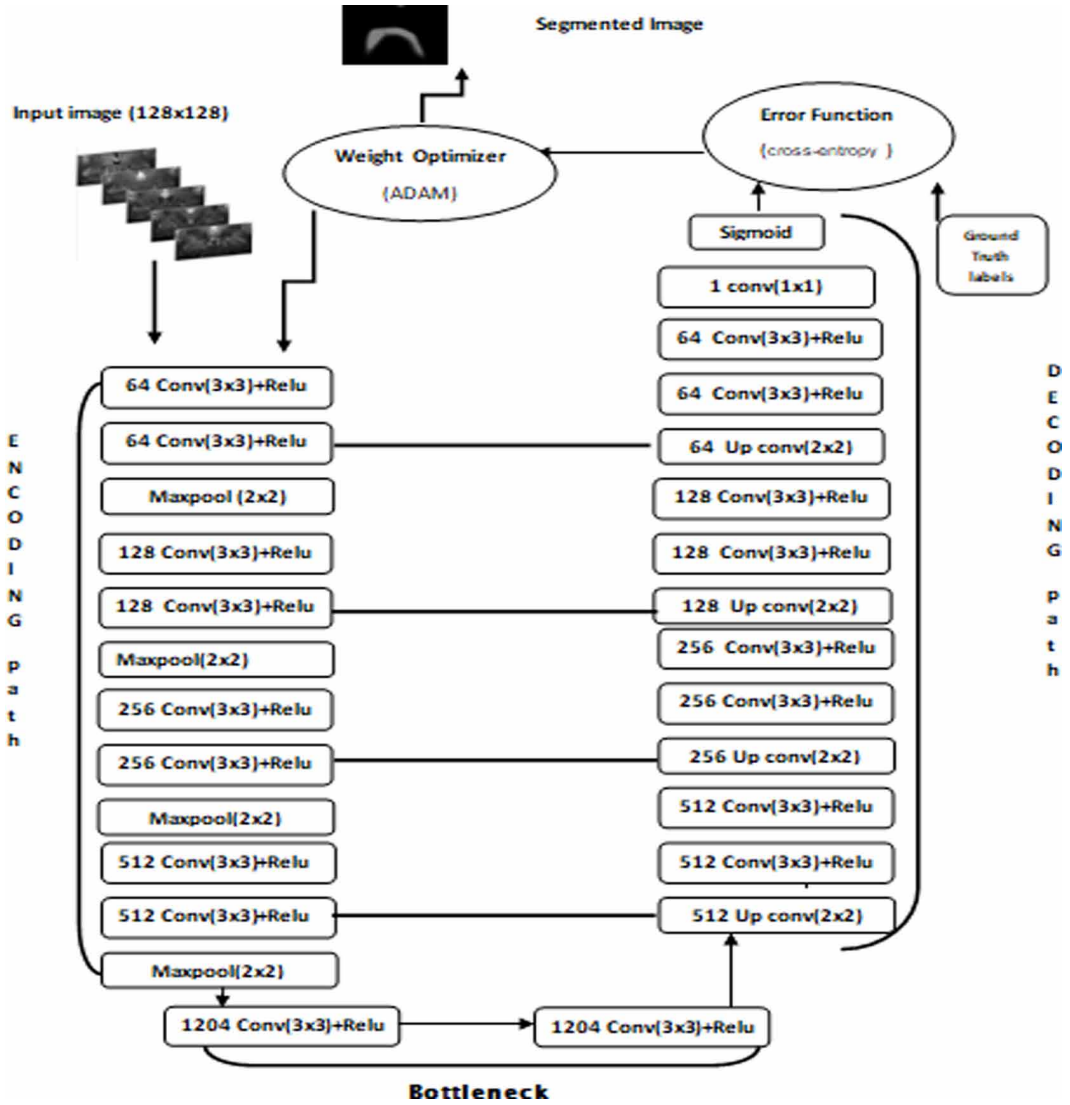
The Accuracy is a number of correct predictions per total number of predictions, is essentially a measure of overlap between two samples.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

The Dice coefficient was originally developed for binary data, is the harmonic mean of the accuracy and sensitivity, it can be calculated as:



Figure 3. HSSE network. the network consists in segmenting the hippocampus as a single entity by activating the neurons of the last layer by the sigmoid function . The symbols used inside the boxes indicate: Conv: convolution, Maxpool: max pooling, (#x#x):size of filter, Relu: linear function Relu, #:number of filters and the gray lines represent the concatenation(-)

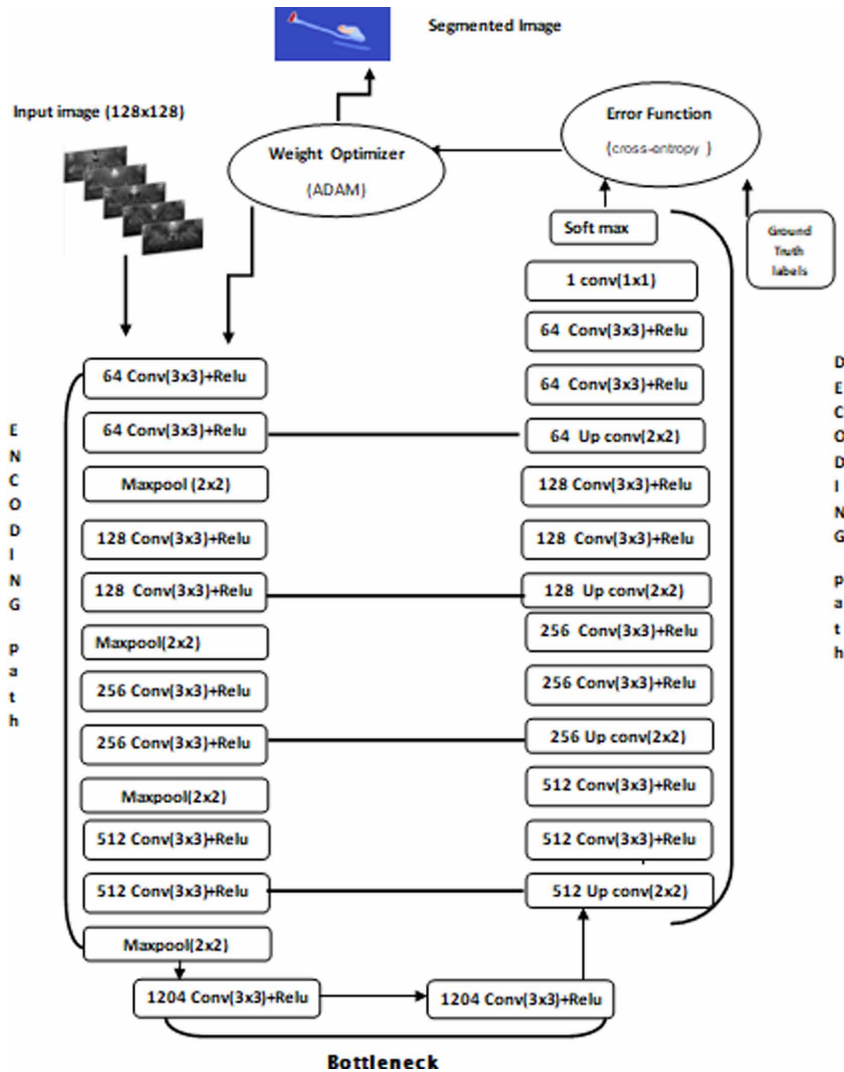


$$Dice = \frac{2TP}{2TP + FP + FN} \quad (5)$$

Where true positives (TP) as predicted hippocampus pixels inside positive regions in ground-truth, false positives (FP) as predicted hippocampus pixels outside positive regions in ground-truth, true negatives (TN) as predicted background pixels outside positive regions in ground-truth, and false negatives (FN) as predicted background pixels inside positive regions in ground-truth.

So as suggested by above equations, sensitivity is the proportion of true positives that are correctly identified by a diagnostic test. It shows how good the test is at detecting a disease. Specificity is the proportion of the true negatives correctly identified by a diagnostic test. It suggests how good the

Figure 4. HSMC network. the network detect hippocampal subregions by activating the neurons of the last layer by the softmax function . The symbols used inside the boxes indicate: Conv: convolution, Maxpool: max pooling, (#x#):size of filter, Relu: linear function Relu,#:number of filters and the gray lines represent the concatenation (-)



test is at identifying normal (negative) condition. Accuracy is the proportion of true results, either true positive or true negative, in a population. It measures the degree of veracity of a diagnostic test on a condition. Some of the most important hyperparameter experiments can be seen in Table 1.

Table 2 presents the quantitative results of the two networks. We note from table 2 that the scores obtained from the two networks are high, however the HSSE network generates low scores compared to the HSMC network where the dice = 0.78 and sensitivity = 0.70 specially for the Ning 2009 cohort thing that shows the high number of false negative. we tried to overcome this problem by increasing the number of epochs but we found identical scores

We compare the performance of our networks with current hippocampal segmentation algorithms. In the literature, there are no algorithms or methods based on CNN architectures that can segment the hippocampal subregions. In addition the training data for this article is not used in machine learning

Figure 5. Training and testing loss of HSSE model over 100 epoch

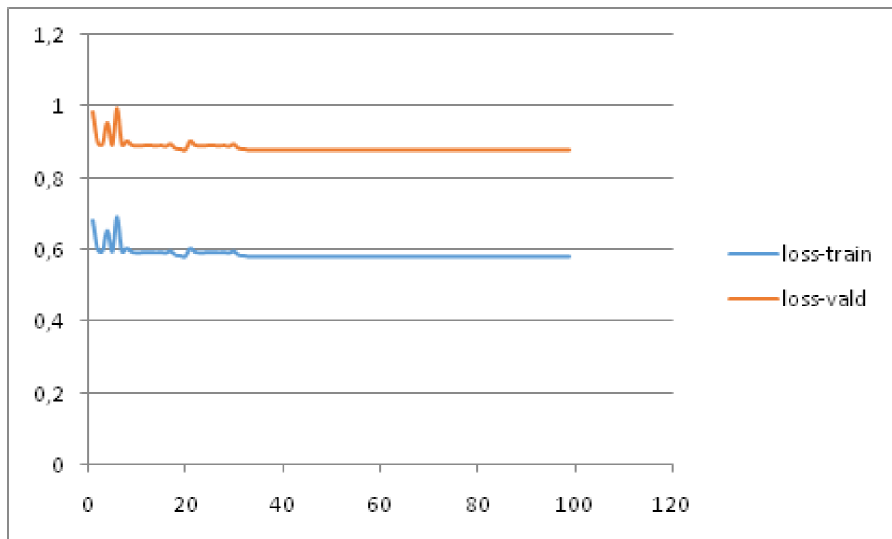


Figure 6. Training and testing loss of HSMC model over 250 epoch

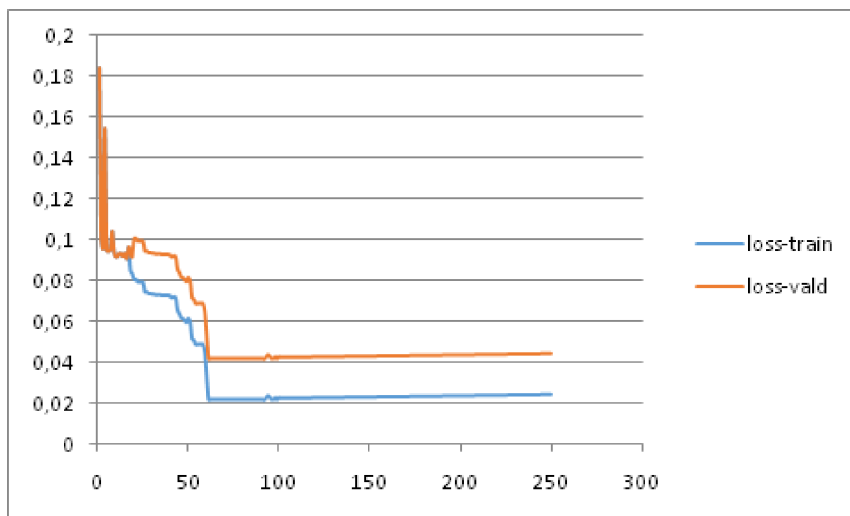


Table 1. Hyper-parameter experiments

Method	Epoch Number	Batch Size	Activation Function in the Classifier Layer	Optimizer
HSSE	100	32	Sigmoid	ADAM
HSMC	250	32	Softmax	ADAM

via a given CNN except in Nasser et al. (2020) for automatic segmentation, Yushkevich et al. (2009) for a manual segmentation, and in Yushkevich et al. (2015) who used a multi-atlas segmentation with joint label fusion.

The comparative works detect the hippocampus as a single entity from the complete cerebral MRI using a deep learning segmentation with other data. For this, we compared our results by these last works in a general way. These works used architecture similar to U-net.

We see in the architectures of the works cited in table 2 that the models of the networks are inspired by the U-net architecture with some modifications:

**Table 2. Performance of the HSSE and HSMC networks for Atlas UMCU and Nimg2009 data set**

Method		Dice	Sensitivity	Accuracy
HSSE	Atlas UMCU	0.78	0.70	0.92
	Nimg2009	0.91	0.89	0.95
HSMC	Atlas UMCU	0.90	0.87	0.96
	Nimg2009	0.92	0.90	0.96

- 1) The size of the input images where they are small compared to the original model of Ronneberger et al. (2015).
- 2) The number of feature maps which start with 32 at Cao, et al. (2017) until 128 in the last layer of encoder path whereas Carmo et al. (2020) initialized the network by 64 feature maps until 512, and Thyreau et al. (2017) chose 48 feature maps without increase in all network. Nasser et al. (2020) used 16 then 32 feature maps. These reductions in input parameters size can be justified by the presence of a complex network containing a large number of inputs images.

In Nasser et al. (2020) we tried to find a simple uncomplicated network which segments hippocampal 3D MRI, we inspired the network from U-net architecture but without concatenation and with few feature maps. 3D images require a delicate preprocessing phase which can pose a problem in learning if it has not been done correctly and consequently the results were encouraging but not sufficient with a dice of 0.64 the networks could not detect all hippocampal regions, for this we have modified the network architecture to add concatenation layers, increase the number of feature maps and use 2D images instead of 3D.

When we notice the results in Table 3, we find that these networks have produced good Dice. Concerning our proposed networks the dices obtained are greater than 0.80 with a U-net network containing up to 1024 feature maps and an image size of 128x128. Also, several more visual segmentation results of our method are shown in figure 7,8,9 and 10.

In Figure 7 and Figure 8 we report the Nimg 2009 cohort segmentation results achieved by HSSE and HSMC network successively .

From the original images we can clearly see the difficulty of detecting the hippocampus of neighboring tissues, the segmented images in figure 7 indicate that HSSE network can accurately describe the limit of the hippocampus; in addition, we also show the segmentation results of the same cohort by HSMC network in Figure 8 where it can detect hippocampal sub-regions. In the two images the network was able to detect 3 sub regions (Figure 8 (B) the first, the third and the fourth arrow) whereas it could not correctly classify the interior regions (second arrow) and some pixels surrounded by a circle.

Table 3. Comparison of HSSE and HSMC with state-of-the-art methods

Method	Input Data Size	Number Input Data	Testing Time	Dice
(Cao et al., 2017)	72x72x72	About 797 MRI volume	/	0.89
(Thyreau et al., 2017)	48x72x64	About 25000 MRI volume	30 second	0.85
(Diedre et al., 2020)	64x64	About (190+135) MRI volume	15s per volume(GPU) 3mn per volume(CPU)	0.90
(Nasser et al., 2020)	145x145x80	80 MRI volume	/	0.64
<b>HSSE</b>	128x128	90 MRI volume	20s	<b>0.84</b>
<b>HSMC</b>	128x128	90 MRI volume	23s	<b>0.91</b>

In Figures 9 and 10, we report the results of the segmentation of the Atlas UMCU cohort obtained by the HSMC network. This network was learned on 200 epoch, we took the progression of the model after 20 epochs for a subject from our validation set (Figure 9).

From figure 10, we can see that HSMC network was able to detect all the hippocampal sub-regions by comparing by ground truth which implies that the global segmentation of the performance of our network is reasonable.

## DISCUSSION

We designed two networks destined to segment the hippocampus, the first one segments the hippocampus as a single entity (HSSE) and the second detects hippocampal sub-regions (HSMC). The two networks inspire their architectures from U-net model. We have tried to reduce the size of the input images to 128x128 to save memory space and minimize learning time.

We generated the maximum of feature maps to detect the contextual and spatial information of the image, the high number of feature maps and layers weighed down the networks. It's prompted us to add drop layers to remove a certain connection and weights between neurons and to avoid a problem of over-adjustment.

Figure 4 shows that the loss saturates after a few epochs for HSSE network, and the model exhibits comparable loss performance for both training and testing sets.

From figures 5, we notice that the loss of the training data set is less than the validation data set.

The trace of the training loss decreases to a point of stability. The trace of the loss of validation decreases to a point of stability and presents a small deviation with a loss of training equal to 0.02.

We can deduce that the HSMC network avoided over-fitting and this is proven by a loss of training and validation which decreases to a point of stability with a minimum difference between the two final loss values.

The sigmoid function was used in HSSE to activate the neurons of the last layer and classify the pixels between background and hippocampus region, the network was learned on 100 epochs until the weight stabilization. On the other hand in HSMC network we used the softmax function to detect the hippocampal sub-regions in a higher epoch set 100 times than HSSE and this means by the number of classes to be classified. The two networks gave high scores in the two databases, However the HSSE network generates a weak dice for the atlas-UMCU database compared to the other dice generated.

The high number of false negatives can cause this kind of problem which is related in our opinion to the variety of data sets. We thought of using a single cohort to avoid this problem but the number of images in each cohort is probably low for a convolutional neural network. It is important to clarify that the data sets which contain the segmentation of the hippocampal sub regions is few the thing which obliges us to mix several small sets to learn the networks. We have not yet solved the problem

Figure 7. Visualization results of the pixels classification by HSSE network in the Nimg 2009 Atlas test set, above the original image on the left the segmented image on the right

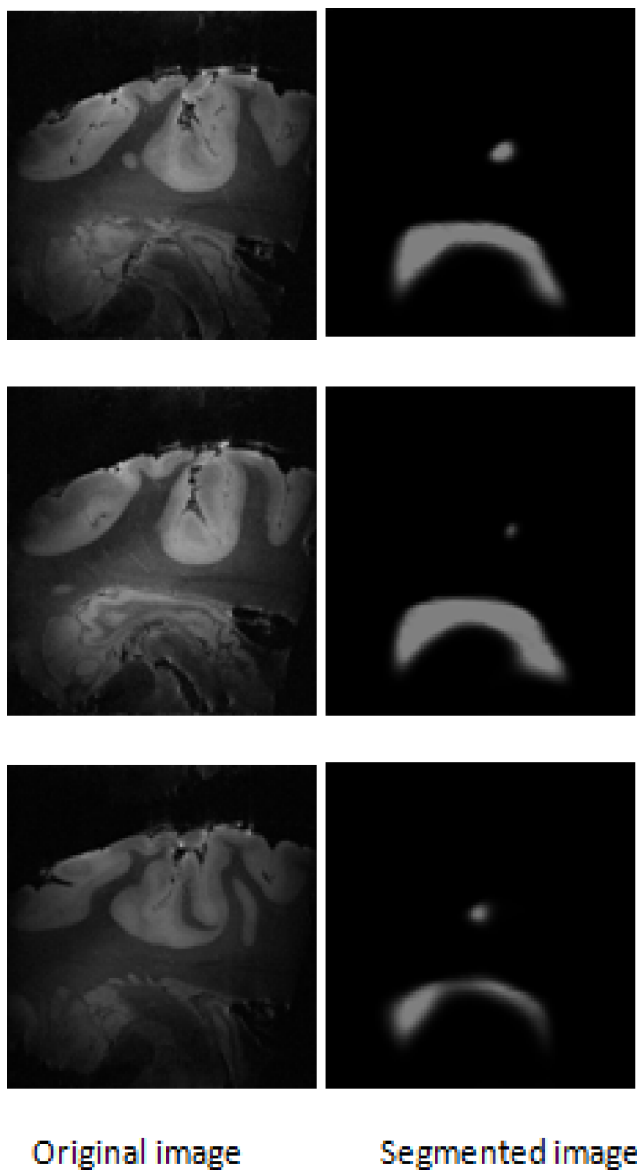


Figure 8. Results of visualization of the pixels classification by HSMC network in the Nimg 2009 Atlas test set, on the left the original image in the middle the ground truth on the right segmented image

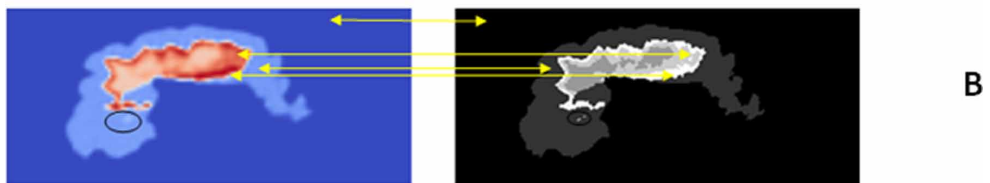
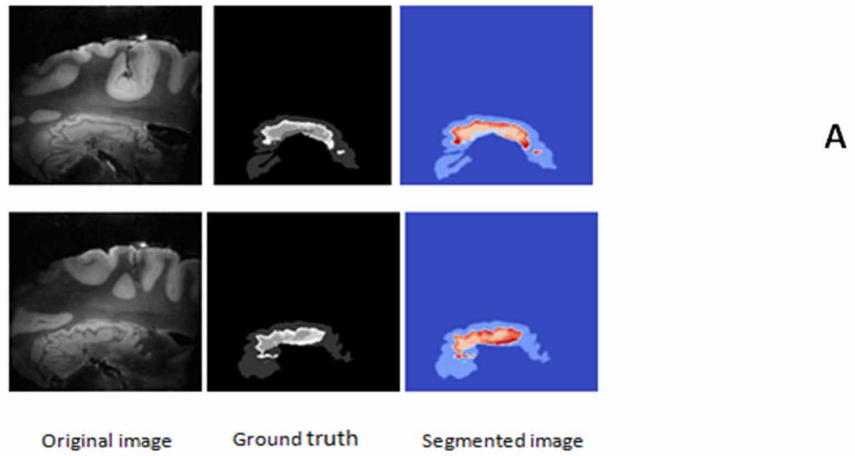
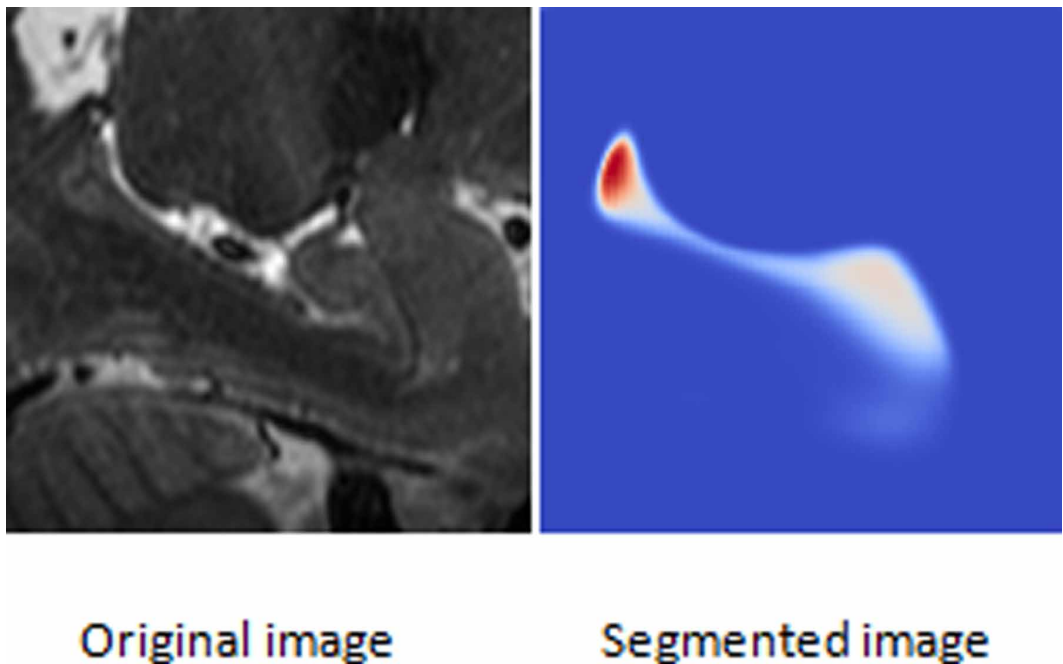
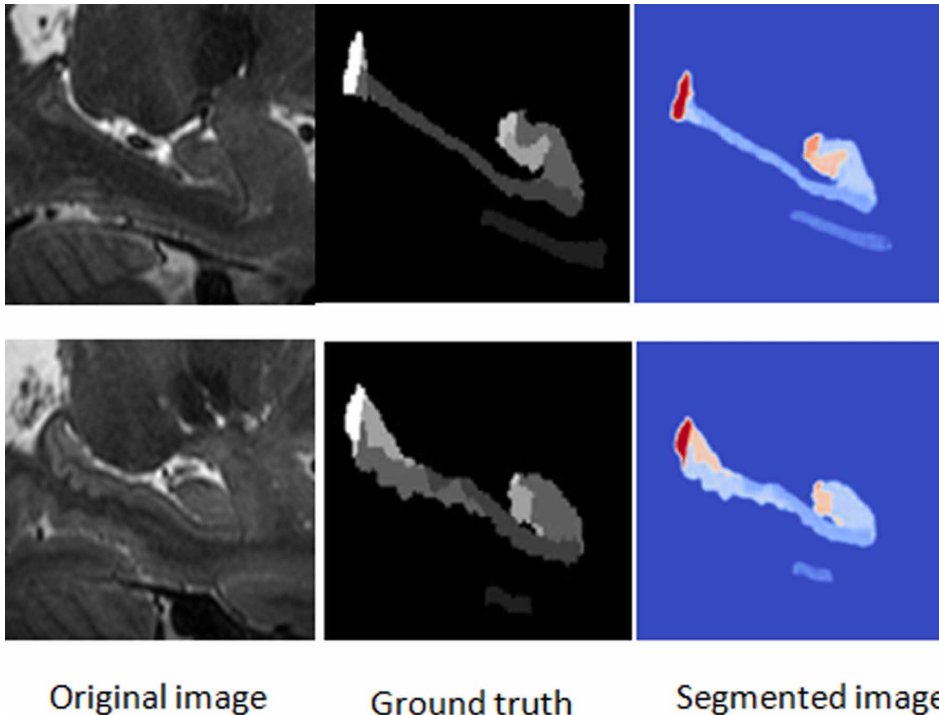


Figure 9. Results of visualization of the pixels classification by HSMC network in the Atlas-UMCU-15-9-17 test set after the fiftieth epoch, on the left the original image in the middle the ground truth on the right segmented image



**Figure 10.** Results of visualization of the pixels classification by HSMC network in the Atlas-UMCU-15-9-17 test set, on the left the original image in the middle the ground truth on the right segmented image



but we think of applying in the next works an oversampling or a sub-sampling (SMOTE), on the train data as a phase pretreatment.

## CONCLUSION

In this article, we have proposed two convolutional neural networks based on the U-net architecture aimed at segmenting the hippocampus, the first network (HSSE) segments the hippocampus into a single entity, whereas the second one (HSMC) detects hippocampal sub-regions.

Both networks have been applied to new databases used in the literature which contain hippocampus manually segmented using ITK-SNAP software to sub-regions, the reason why we chose them.

To ensure good generalization on hippocampal forms, we increased the size of the data set twice as much and to guarantee in-depth learning, we used thousands of feature maps and filters.

The methods presented allow to obtain encouraging performances on the hippocampus of Nimg 2009 and Atlas-UMCU-15-9-17 with more than 0.84 of dice and 0.92 of precision except in certain examples of Atlas-UMCU -15-9 -17 which cannot be detected correctly as a single entity the thing which we will treat in the next works by applying oversampling or sub-sampling (SMOTE), on the train data as phase preprocessing.

In addition, we intend to use other kinds of medical images in order to segment them via CNN architectures such as the images of chronic wounds.

In Chakraborty (2019), Chinmay Chakraborty used an automated classification via non-deep approaches; our goal is to introduce CNNs in this kind of medical problem of course in the case of having an annotated images database.



## REFERENCES

- Ashton, E., Parker, K., Berg, M., & Chen, C. W. (1997). A novel volumetric feature extraction technique with applications to MR images. *IEEE Transactions on Medical Imaging*, 16(4), 365–371. doi:10.1109/42.611343 PMID:9262994
- Avants, B. T. (2009). Advanced normalization tools (ants). *Insight J*. <http://hdl.handle.net/10380/3113>
- Ben Naceur, M., Saouli, R., Akil, M., & Kachouri, R. (2018). Fully Automatic Brain Tumor Segmentation using End-to-End Incremental Deep Neural Networks in MRI images. *Computer Methods and Programs in Biomedicine*. Advance online publication. doi:10.1016/j.cmpb.2018.09.007
- Bobinski, M., De Leon, M. W., Desanti, S., Rusinek, H., & Wisniewski, H. (1999). The histological validation of post mortem magnetic resonance imaging-determined hippocampal volume in Alzheimer's disease. *Neuroscience*, 95(3), 721–725. doi:10.1016/S0306-4522(99)00476-5 PMID:10670438
- Boutet, C., Chupin, M., Lehericy, S., Marrakchi-Kacem, L., Epelbaum, S., Poupon, C., Wiggins, C., Vignaud, A., Hasboun, D., Defontaine, B., Hanon, O., Dubois, B., Sarazin, M., Hertz-Pannier, L., & Collio, O. (2014). Detection of volume loss in hippocampal layers in Alzheimer's disease using 7 T: a feasibility study. *Neuroimage: Clinical*, 5, 341–348. doi:10.1016/j.nicl.2014.07.011
- Brosch, T., Tang, L., Yoo, Y., Li, D., Traboulsee, A., & Tam, R. (2016). Deep 3D Convolutional Encoder Networks With Shortcuts for Multiscale Feature Integration Applied to Multiple Sclerosis Lesion Segmentation. *IEEE Transactions on Medical Imaging*, 35(5), 1229–1239. doi:10.1109/TMI.2016.2528821 PMID:26886978
- Cao, L., Li, L., Zheng, J., Fan, X., Yin, F., & Shen, H. (2017). Multi-task neural networks for joint hippocampus segmentation and clinical score regression. *Multimed Tools Appl*. 10.1007/s11042-017-5581-1
- Carmo, D., Silva, B., Yasuda, C., Rittner, L., & Lotufo, R. (2020). *Hippocampus Segmentation on Epilepsy and Alzheimer's Disease Studies with Multiple Convolutional Neural Networks*. ArXiv, abs/2001.05058.
- Chakraborty, C. (2019). Computational approach for chronic wound tissue characterization. *Informatics in Medicine Unlocked*, 100162. doi:10.1016/j.imu.2019.100162
- Chupin, M., Gerardin, E., Cuingnet, R., Boutet, C., Lehericy, L., & Benali, H. (2009). Fully Automatic Hippocampus Segmentation and Classification in Alzheimer's Disease and Mild Cognitive Impairment Applied on Data From ADNI. *Hippocampus*, 19(6), 579–587. doi:10.1002/hipo.20626 PMID:19437497
- Cicek, O., Abdulkadir, A., Lienkamp, S., Brox, T., & Ronneberger, O. (2016). 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. In *19th International Conference Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016, Athens, Greece, October 17-21, 2016, Proceedings, Part II*. Springer.
- Cury, C. (2015). *Analyse statistique de la variabilité anatomique de l'hippocampe à partir de grandes populations* (Doctoral thesis). University of Pierre and Marie Curie.
- Davy, A., Havaei, M., Warde-Farley, D., Biard, A., Tran, L., & Jodoin, P. (2014). Brain Tumor Segmentation With Deep Neural Networks. *MICCAI - brats*. 10.1016/j.media.2016.05.004
- De Alejo, R. P.-C., Cortijo, M., Rodriguez, I., Echave, I., Regadera, J., & Graña, M. (2003). Computer-assisted enhanced volumetric segmentation magnetic resonance imaging data using a mixture of artificial neural networks. *Magnetic Resonance Imaging*, 21(8), 901–912. doi:10.1016/S0730-725X(03)00193-0 PMID:14599541
- Ghafoorian, M., Karssemeijer, N., Heskes, T., Bergkamp, M., Wissink, J., & Obels, J. et al.. (2016). Deep Multi-scale Location-aware 3D Convolutional Neural Networks for Automated Detection of Lacunes of Presumed Vascular Origin. *NeuroImage: Clinical*. Advance online publication. doi:10.1016/j.nicl.2017.01.033 PMID:28271039
- Ghanei, A., Soltanian-Zadeh, H., & Windham, J. P. (1998). A 3D deformable surface model for segmentation of objects from volumetric data in medical images. *Computers in Biology and Medicine*, 28(3), 239–253. doi:10.1016/S0010-4825(98)00007-9 PMID:9784962

Györfi, O., Nagy, H., Bokor, M., Moustafa, A., Rosenzweig, I., Kelemen, O., & Kéri, S. (2017). Reduced CA2–CA3 Hippocampal Subfield Volume Is Related to Depression and Normalized by L-DOPA in Newly Diagnosed Parkinson's Disease. *Frontiers in Neurology*, 8(84). Advance online publication. doi:10.3389/fneur.2017.00084 PMID:28367136

Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courvill, A., & Bengio, Y. (2017). Brain Tumor Segmentation with Deep Neural Networks. *Medical Image Analysis*, 35, 18–31. 10.1016/j.media.2016.05.004

Heckemann, R. A., Hajnal, J. V., Aljabar, P., Rueckert, D., & Hammers, A. (2006). Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. *NeuroImage*, 33(1), 115–126. <http://www.sciencedirect.com/science/article/pii/S1053811906006458>

Henry, T. R., Chupin, M., Lehericy, S., Strupp, J. P., Sikora, M. A., Sha, Z. Y., Ugurbil, K., & Van de Moortele, P. F. (2011). Hippocampal sclerosis in temporal lobe epilepsy: Findings at 7 T<sup>1</sup>. *Radiology*, 261(1), 199–209. doi:10.1148/radiol.11101651 PMID:21746814

Isin, A., Direkoglu, C., & Sah, S. (2016). Review of MRI-based brain tumor image segmentation using deep learning methods. *Procedia Computer Science*, 102, 317 – 324. 10.1016/j.procs.2016.09.407

Kamnitsas, K. L., Simpson, J., Kane, A., Menon, D., Rueckert, D., & Glocker, B. (2017). Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical Image Analysis*, 36, 61–78. doi:10.1016/j.media.2016.10.004 PMID:27865153

Kelemen, A., Szekely, G., & Gerig, G. (1999). Elastic model-based segmentation of 3-D neuroradiological data sets. *IEEE Transactions on Medical Imaging*, 18(10), 828–839. doi:10.1109/42.811260 PMID:10628943

Kerchner, G., Hess, C., Hammond-Rosenbluth, K., Xu, D., Rabinovici, G., Kelley, D., Vigneron, D. B., Nelson, S. J., & Miller, B. L. (2010). Hippocampal CA1 apical neuropil atrophy in mild Alzheimer disease visualized with 7-T MR. *Neurology*, 75(15), 1381–1387. doi:10.1212/WNL.0b013e3181f736a1 PMID:20938031

Keskar, N. S., Nocedal, J., Tang, P. T. P., Mudigere, D., & Smelyanskiy, M. (2017). *On large-batch training for deep learning: Generalization gap and sharp minima*. Paper presented at 5th International Conference on Learning Representations, ICLR 2017, Toulon, France.

Kim, H., Chupin, M., Colliot, O., Bernhardt, B., Bernasconi, N., & Bernasconi, A. (2011). Automatic hippocampal segmentation in temporal lobe epilepsy: Impact of developmental abnormalities. *NeuroImage*, 59(4), 3178–3186. doi:10.1016/j.neuroimage.2011.11.040 PMID:22155377

Klemencic, J., Pluim, J., Viergever, M., Schnack, H., & Valencic, V. (2004). Non-rigid registration based active appearance models for 3D medical image segmentation. *Journal of Imaging Science and Technology*, 48(2), 166–171. <https://research.tue.nl/en/publications/non-rigid-registration-based-active-appearance-models-for-3d-medi>

Koolschijn, P. C., van Haren, N. E., Cahn, W., Schnack, H. G., Janssen, J., Klumpers, F., Hulshoff Pol, H. E., & Kahn, R. S. (2010, June). Hippocampal volume change in schizophrenia. *The Journal of Clinical Psychiatry*, 71(6), 737–744. doi:10.4088/JCP.08m04574yel PMID:20492835

Kruthika, K., Rajeswari, H., Maheshappa, & Initiative, A. D. (2019). CBIR system using Capsule Networks and 3D CNN for Alzheimer's disease Diagnosis. *Informatics in Medicine Unlocked*, 14, 59–68. 10.1016/j.imu.2018.12.001

La Joie, R. P., Doeuvre, L., Belliard, S., Eustache, F., Desgranges, B., & Chételat, G. (2013). Hippocampal subfield volumetry in mild cognitive impairment, Alzheimer's disease and semantic dementia. *NeuroImage. Clinical*, 14(3), 155–162. doi:10.1016/j.nicl.2013.08.007 PMID:24179859

Margeta, J., Criminisi, A., Cabrera Lozoya, R., Lee, D., & Ayache, N. (2015). Fine-tuned convolutional neural nets for cardiac MRI acquisition plane recognition. *Computer Methods in Biomechanics and Biomedical Engineering. Imaging & Visualization*, 1–11. doi:10.1080/21681163.2015.1061448

Mueller, S., Schuff, N., Yaffe, K., Madison, C., Miller, B., & Weiner, M. (2010). Hippocampal atrophy patterns in mild cognitive impairment and Alzheimer's disease. *Human Brain Mapping*, 31(9), 1339–1347. 10.1002/hbm.2093420839293

Nasser, S., Naoui, M., & Belalem, G. (2018). *Automated segmentation of 7T MRI images of hippocampus*. 2d IEEE-2018 international Conference of control, automation and diagnosis, Marrakech, Morocco.

- Nasser, S., Naoui, M., Belalem, G., & Mahmoudi, S. (2020). Automated Segmentation of Brain Medical Image. *Journal of Signal Processing, Image Processing and Pattern Recognition*, 13(1), 17-26. DOI: doi:10.33832/ijsp.2020.13.1.02
- Poudel, R., Lamata, P., & Montana, G. (2016). Recurrent Fully Convolutional Neural Networks for Multi-slice MRI Cardiac Segmentation. *Conference: International Workshop on Reconstruction and Analysis of Moving Body Organs International Workshop on Whole-Heart and Great Vessel Segmentation from 3D Cardiovascular MRI in Congenital Heart Disease*, 10129, 83-94. doi:10.1007/978-3-319-52280-7\_8
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *MICCAI. LNCS9351*, 234–241. doi:10.1007/978-3-319-24574-4\_28
- Roy, A. G., Conjeti, S., & Navab, N. (2018). QuickNAT: A fully convolutional network for quick and accurate segmentation of neuroanatomy. *NeuroImage*, 186, 713–727. doi:10.1016/j.neuroimage.2018.11.042 PMID:30502445
- Schnabel, J. A., Lemieux, L., Wiesmann, U. C., & Arridge, S. R. (1999). Measurement of hippocampal volume changes in serial MRI scans. In K. M. Hanson (Ed.), *Medical Imaging. Image Processing*. SPIE. doi:10.1117/12.348535
- Shen, D., Moffat, S., Resnick, S., & Davatzikos, C. (2002). Measuring size and shape of the hippocampus in mr images using a deformable shape model. *NeuroImage*, 15(2), 422–434. doi:10.1006/nimg.2001.0987 PMID:11798276
- Thong, W., Kadoury, S., Piché, N., & Pal, C. (2016). Convolutional networks for kidney segmentation in contrast-enhanced CT scans. *Computer Methods in Biomechanics and Biomedical Engineering. Imaging & Visualization*, 1–6. doi:10.1080/21681163.2016.1148636
- Thyreau B, Sato K, Fukuda H, Taki Y. (2017). Segmentation of the hippocampus by transferring algorithmic knowledge for large cohort processing. *Med Image Anal.*, 43, 214-228. doi:10.1016/j.media.2017.11.004
- Valverde, S., Cabezas, M., Roura, E., González-Villà, S., Pareto, D., Vilanova, J., Ramió-Torrentà, L., Rovira, À., Oliver, A., & Lladó, X. (2017). Improving automated multiple sclerosis lesion segmentation with a cascaded 3D convolutional neural network approach. *NeuroImage*, 155, 159–168. Advance online publication. doi:10.1016/j.neuroimage.2017.04.034 PMID:28435096
- Wisse, L. E., Biessels, G. J., Heringa, S. M., Kuijf, H. J., Koek, D. H., Luijten, P. R., & Geerlings, M. I. Utrecht Vascular Cognitive Impairment (VCI) Study Group. (2014, September). Hippocampal subfield volumes at 7T in early Alzheimer's disease and normal aging. *Neurobiology of Aging*, 35(9), 2039–2045. doi:10.1016/j.neurobiolaging.2014.02.021 PMID:24684788
- Wisse, L.E., Gerritsen, L., Zwanenburg, J.J., Kuijf, H.J., Luijten, P.R., Biessels, G.J., & Geerlings, M.I. (2012). Subfields of the hippocampal formation at 7 T MRI: in vivo volumetric assessment. *Neuroimage*, 61(4), 1043-9. doi:10.1016/j.neuroimage.2012.03.023
- Wisse, L. E., Kuijf, H. J., Honingh, A. M., Wang, H., Pluta, J. B., Das, S. R., Wolk, D. A., Zwanenburg, J. J., Yushkevich, P. A., & Geerlings, M. I. (2016, June). Automated Hippocampal Subfield Segmentation at 7T MRI. *AJNR. American Journal of Neuroradiology*, 37(6), 1050–1057. doi:10.3174/ajnr.A4659 PMID:26846925
- Xu, Y., Geraud, T., & Bloch, I. (2018). Segmentation d'IRM de cerveaux de nouveau-nés en quelques secondes à l'aide d'un réseau de neurones convolutif pseudo-3D et de transfert d'apprentissage. *Conference RFIAP 2018*.
- Yushkevich, P. A.-H., Pickup, S., Liu, W., Gee, J. C., Grossman, M., & Detre, J. (2009). A High-Resolution Computational Atlas of the Human Hippocampus from Postmortem Magnetic Resonance Imaging at 9.4 Tesla. *NeuroImage*, 44(2), 385–398. doi:10.1016/j.neuroimage.2008.08.042 PMID:18840532
- Yushkevich, P. P., Xie, L., Ding, S. G., Klot, D., & Das, S. e. (2015). Automated Volumetry and Regional Thickness Analysis of Hippocampal Subfields and Medial Temporal Cortical Structures in Mild Cognitive Impairment. *Human Brain Mapping*, 36(1), 258–287. doi:10.1002/hbm.22627 PMID:25181316

## **ENDNOTES**

- <sup>1</sup> [https://www.nitrc.org/include/about\\_us](https://www.nitrc.org/include/about_us)
- <sup>2</sup> <http://www.itksnap.org/pmwiki/pmwiki.php>