



**HAL**  
open science

# Cerebral AVM segmentation from 3D rotational angiography images by convolutional neural networks

Mounir Lahlouh, Raphaël Blanc, Michel Potin, Jérôme Szewczyk, Nicolas Passat, Yasmina Leroul

► **To cite this version:**

Mounir Lahlouh, Raphaël Blanc, Michel Potin, Jérôme Szewczyk, Nicolas Passat, et al.. Cerebral AVM segmentation from 3D rotational angiography images by convolutional neural networks. *Neuroscience Informatics*, 2023, 3 (3), pp.100138. 10.1016/j.neuri.2023.100138 . hal-04164909

**HAL Id: hal-04164909**

**<https://hal.science/hal-04164909v1>**

Submitted on 12 Aug 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NoDerivatives 4.0 International License

# Cerebral AVM segmentation from 3D rotational angiography images by convolutional neural networks

Mounir Lahlouh<sup>1,2</sup>, Raphaël Blanc<sup>3</sup>, Michel Pletin<sup>3</sup>, Jérôme Szewczyk<sup>4</sup>, Nicolas Passat<sup>1</sup>, Yasmina  
Chenoune<sup>2,5</sup>

<sup>1</sup>Université de Reims Champagne Ardenne, CReSTIC EA 3804, Reims 51097, France

<sup>2</sup>ESME Sudria Research Lab, Paris, France

<sup>3</sup>Fondation Ophtalmologique de Rothschild, Interventional Neuroradiology Department, Paris, France

<sup>4</sup>Sorbonne Université, CNRS UMR 7222, Inserm, U1150, ISIR, F-75005, Paris, France

<sup>5</sup>Université Paris-Est, LISSI (EA 3956), UPEC, F-94010, Vitry-sur-Seine, France

## *Abstract*

**Background and objective:** 3D rotational angiography (3DRA) provides high quality images of the cerebral arteriovenous malformation (AVM) nidus that can be reconstructed in 3D. However, these reconstructions are limited to only 3D visualization without possible interactive exploration of geometric characteristics of cerebral structures. Refined understanding of the AVM angioarchitecture prior to treatment is mandatory and vascular segmentation is an important preliminary step that allow physicians analyze the complex vascular networks and can help guide microcatheters navigation and embolization of AVM.

**Methods:** A deep learning method was developed for the segmentation of 3DRA images of AVM patients. The method uses a fully convolutional neural network with a U-Net-like architecture and a DenseNet backbone. A compound loss function, combining Cross Entropy and Focal Tversky, is employed for robust segmentation. Binary masks automatically generated from region-growing segmentation have been used to train and validate our model.

**Results:** The developed network was able to achieve the segmentation of the vessels and the malformation and significantly outperformed the region-growing algorithm. Our experiments were performed on 9 AVM patients. The trained network achieved a Dice Similarity Coefficient (DSC) of 80.43%, surpassing other U-Net like architectures and the region-growing algorithm on the manually approved test set by physicians.

**Conclusions:** This work demonstrates the potential of a learning-based segmentation method for characterizing very complex and tiny vascular structures even when the training phase is performed with

the results of an automatic or a semi-automatic method. The proposed method can contribute to the planning and guidance of endovascular procedures.

***Index Terms-*** 3D rotational angiography, cerebrovascular segmentation, convolutional neural networks, focal Tversky, fully automatic segmentation.

## 1. INTRODUCTION

Medical imaging plays a major role in the diagnosis of endovascular pathologies such as aneurysms, arteriovenous malformations (AVMs) or stenosis. It is also highly involved in the pre- and perioperative phases of endovascular treatments like embolization or mechanical thrombectomy. These techniques are considered minimally invasive and consist of navigating throughout the blood vessels using microcatheters and flexible guidewires with the help of imaging techniques such as conventional angiography or 3D Rotational Angiography (3DRA) [1] [2] [3]. Poor visualization, vessels tortuosity, or small size vessels can make these procedures very complex. The development of dedicated approaches for improved 3D visualization can help interventional neuroradiologists for their pre-operative planning or even reduce procedure time and thereby radiation dose.

These modeling tools could be useful in characterizing the geometry of the vessels (diameter, vessels length, tortuosity, bifurcation angles). Indeed, knowing the anatomy of these thin structures is essential as it allows a better understanding of the complex architecture of the vessels. However, the generation of a compliant 3D model of the vessels, allowing the extraction of precise geometric characteristics, strongly depends on segmentation quality, especially when dealing with small and tangled vessels, such as cerebral AVMs.

In the context of AVMs, a good quality segmentation enables precise anatomical localization, assessment of size and shape, and evaluation of treatment response. However, the segmentation of AVM structures presents several challenges. These include the complex, irregular and tangled shape of AVMs and the small size of the connected vessels. Overcoming these challenges requires the development of advanced segmentation algorithms and techniques, including deep learning approaches, that can effectively capture the details and accurately delineate AVM structures.

The aim of this study is to address challenges in the segmentation of cerebrovascular vessels on 3DRA images and improve the overall segmentation results. This was achieved through the development of a Dense U-Net architecture that combines Densenet and U-Net networks to enhance the segmentation of cerebrovascular vessels. A combination of the cross-entropy and focal Tversky loss functions was proposed to handle imbalanced data and improve the segmentation of small structures, such as AVMs. Comparative analysis was conducted to evaluate the performance of the approach against other segmentation methods, and the segmentation results were validated against a manual ground-truth.

### *1.1 Cerebral arteriovenous malformation*

Cerebral arteriovenous malformations (AVM) are one of the most complex neurovascular pathologies. they consist of an abnormal tangle of blood vessels that connect arteries and veins through a shunt which feeds into a vascular nidus that replaces the normal capillary network. Due to the resulting shunting, the AVM dilates over time. This dilation weakens the veins and makes them susceptible to aneurysms and hemorrhage. After accurate diagnosis by modalities such as Magnetic Resonance Angiography (MRA), Computed Tomography Angiography (CTA), a complete cerebral angiography imaging with 3DRA are required to propose an appropriate treatment [2]. Treatment options and modalities for these malformations can be single or multimodal therapy and include microsurgical resection, stereotactic radiotherapy, or endovascular embolization [4].

Endovascular embolization procedure is a minimally invasive technique to cure challenging AVM's cases[5]. This procedure requires the use of a microcatheter that navigates through blood vessels to the brain with the help of X-ray imaging, like 3DRA modality. An embolic agent is injected to occlude and reduce blood flow in the AVM by isolating the nidus from both feeding arteries and draining veins.

### *1.2 Clinical needs and objectives*

Semantic segmentation consists in classifying each pixel of the image in its correct category to generate a mask which highlights the areas of interest. In the medical field, vessel segmentation is used for the purposes of image-guided interventions, radiotherapy, or improved radiological diagnosis.

Previous research works have been carried out to assist neuroradiologists in the visualization, interpretation, and exploitation of 3DRA images for interventions planning in case of AVM [6] [7] [8]. One study [6] has principally focused on the development of a 3D region-growing-based segmentation method and a symbolic representation of the vessels for an advanced vascular visualization and embolization paths extraction. One of the limits of this work was that, in the absence of ground-truth, the validation was limited to a visual analysis of the segmentation quality although a partial quantitative validation had been carried out. In [7] a fully automated pipeline for vessel enhancement of tree-like structures on 3DRA, MRA, CTA was presented. In another study [8], the authors dealt with vessels delineation and localization on AVM patients based on a graph-like method to extract draining veins coming out of the nidus.

In order to improve the segmentation process, this work aims to experiment a deep neural network-based approach for the vascular structure segmentation from the 3DRA images. The idea is to exploit the binary masks automatically generated from a previous region-growing segmentation [6] to train and validate our model. The proposed model is based on a fully convolutional network adopting the U-Net-like architecture with a Densenet and a combination of focal Tversky (FT) and cross entropy (CE) loss functions.

To test our neural network and quantitatively validate our results, we used a database of 200 3DRA slices, manually segmented by an experienced clinician. The proposed segmentation method is described in Section 3. The obtained results are then presented in Section 4 and discussed in Section 5. Conclusions and perspectives are finally given in Section 6.

## **2. RELATED WORK ON SEGMENTATION BY DEEP LEARNING**

The following summarizes the contributions of deep neural networks to the segmentation of medical images, based on their architectures and loss functions.

### *2.1 Contributions in deep neural architectures*

Several works have been done to improve the architectures of deep neural networks in order to overcome the problems encountered during the learning phase, such as the vanishing or the exploding gradient descent problems and the depth of the network in terms of layers and filters.

The U-Net model [9], which is a fully convolutional neural network (FCN) designed in a U-like structure based on the encoder-decoder principle, is one of the architectures commonly employed in medical image segmentation and vessel extraction. The authors showed that this model gives good results in biomedical imaging on a restricted set of ground-truth segmentations of the input images. Thus, several variants of the model have been proposed and deployed for vessels and brain lesions extraction. For instance, Tetteh et al. [10] discussed the problems of low execution speed and high memory requirements associated with full 3D convolutional networks for cerebral vessel segmentation. They considered 3D angiographic volumes, including clinical Time-Of-Flight (TOF) MRA, as well as synchrotron radiation X-ray tomographic microscopy scans of the brain. Livne et al. [11] developed a specialized deep learning method based on U-Net model and a patching strategy for vessels segmentation. They considered patients with cerebrovascular disease like cortical laminar necrosis on TOF and Magnetization Prepared-Rapid Gradient Echo (MPRAGE) images. The developed method has shown an excellent performance in large

vessels and sufficient performance in small ones. Another work [12] has been conducted in which a 3D U-Net including inception modules has been developed in the context of cerebrovascular network segmentation of MRA images.

The previous architectures can be enhanced by integrating the ResNet [13] or Densenet [14] networks. This integration makes it possible to overcome the problem of vanishing gradient descent in very deep networks. As an example, we can cite the ‘‘One Hundred Layers Tiramisu’’ architecture [15] that combines U-Net and Densenet models in both the encoder and decoder paths.

The combination of these models has been widely used in the medical field, especially in the segmentation of small structures such as cerebrovascular network or cerebral lesions. To this end, Dolz et al. [16] proposed a novel architecture based on Dense U-Net for ischemic stroke lesion multi-modal scans segmentation. More recently, Meng et al. [17] adopted a Dense U-Net to segment cerebral vessels with different diameters on Digital Subtraction Angiography (DSA) images.

## 2.2 Contributions in loss functions

In addition to the improvements brought by deep architectures of Convolutional Neural Networks (CNN), the design of new loss functions has resulted in a better extraction of areas of interest, especially when the problem requires a detailed segmentation of structures occupying small regions in the image. Several loss functions have been used to deal with imbalanced data, where the area of interest occupies a small place compared to the whole image background. Among these functions, we find the overlapping ones like the Dice loss that calculates the similarity between the ground-truth and the predicted images. The binary class variant is defined as follows:

$$\text{Dice loss}_c = \sum_c \left( 1 - \frac{\sum_{i=1}^N p_{ic} g_{ic} + \epsilon}{\sum_{i=1}^N p_{ic} + g_{ic} + \epsilon} \right) \quad (1)$$

Where  $g_{ic} \in \{0,1\}$  and  $p_{ic} \in [0,1]$  represent the ground-truth label and the predicted label, respectively.  $N$  is the total number of pixels in the image.  $c$  represents the two complementary classes: background and the object of interest.  $p_{ic}$  is the probability that pixel  $i$  is of the class  $c$ . The same is true for the ground-truth label. The  $\epsilon$  prevents division by zero.

This loss based on the Dice coefficient [18] has been used for the extraction of vessels from a colored retinal image [19] or in the segmentation of 3D medical images of brain tumors on multimodal MRI scans

[20]. One of the limitations of the Dice loss is that it equally penalizes false positives and false negatives. This is particularly true for imbalanced datasets with small regions of interest such as cerebrovascular structures. For that, a generalized version of Dice loss called Tversky loss which allows flexibility in balancing false positives and false negatives has been frequently used according to the situation of an under- or over-segmentation. This function based on the Tversky index [21] has been tested in multiple sclerosis lesions segmentation on MR images [22]. The Tversky loss also struggles in segmenting small regions like vessels or small brain lesions. These structures occupy a very small place in the image, and they do not contribute significantly to the loss function. To solve this problem, Abraham and Khan [23] proposed the Focal Tversky (FT) loss function that allows the network to focus on less accurate predictions that have been misclassified. In their solution, the proposed FT loss decreases significantly for misclassified pixels with low Tversky index. Therefore, small regions of interest will contribute in minimizing the loss in the training phase. The binary class variant of this loss function is defined as follows:

$$FT_c = \sum_c \left(1 - \frac{\sum_{i=1}^N p_{ic} g_{ic} + \epsilon}{\sum_{i=1}^N p_{ic} g_{ic} + \alpha \sum_{i=1}^N p_{i\bar{c}} g_{ic} + \beta \sum_{i=1}^N p_{ic} g_{i\bar{c}} + \epsilon}\right)^{1/\gamma} \quad (2)$$

Symbols in (1) have the same meaning in (2).  $\alpha$  and  $\beta$  are parameters that weigh false negatives and false positives, respectively.  $\gamma$  is the focal parameter that varies in the range [1, 3].

In addition to the overlapping functions, the standard cross entropy (CE) loss and its weighted variants have also proven excellent performances. For instance, Tetteh et al. [10] introduced a class balancing Cross Entropy loss function with false positives rate correction to handle the high-class imbalance and high false positives rate problems for vessel segmentation.

In the context of brain AVMs, as reported in [24], there is no gold standard for 3D visualization. However, few learning-based studies have focused on the segmentation of AVM. Wang et al. [25] have developed a method based on the three-dimensional V-Net architecture with a combined loss function that includes cross entropy and Dice losses. The precision of their segmentation was evaluated on CT-scan images of 80 patients with AVMs manually labeled by physicians. For the delineation of the AVM nidus [26], a two-stage deep learning approach was developed to detect and then segment the malformation. The developed pipeline was trained on 179 contrasted TOF MRI. For the AVM segmentation on 3DRA modality [27], a U-Net-like architecture with a combination of Dice and soft centerline Dice losses was



used. The quantitative validation was assessed on a very small dataset (5 contrast-enhanced 3DRA scans) with a patched training strategy.

To the best of our knowledge, there are few studies in which deep neural architectures have been used in vessel segmenting of patients with brain AVMs, especially on 3DRA images. This gap is certainly due to the challenging tangled aspect of vessels and the difficulty of obtaining sufficiently annotated data in such pathology. Therefore, this work partly draws upon the idea of deep learning approaches to automatically extract vascular structures. Our contributions are the following:

- 1) A Dense U-Net architecture was developed to investigate the combination of Densenet and U-Net networks to fill the gap in the segmentation of cerebrovascular vessels on 3DRA images.
- 2) A combination of the cross-entropy and focal Tversky loss functions was proposed to address challenges related to imbalanced data and segmentation of small structures (AVMs).
- 3) Improved segmentation results for cerebrovascular structures were demonstrated, even when initiating the training phase with the results of an automatic or semi-automatic method.
- 4) A comparison was made between our approach, a region-growing method, and other learning-based architectures using a 3DRA dataset of 9 patients.
- 5) Segmentation results were validated against a manual ground-truth, consisting of 200 slices manually segmented and validated by a neuroradiologist.

### **3. MATERIALS AND METHODS**

#### *3.1 Data acquisition and preprocessing*

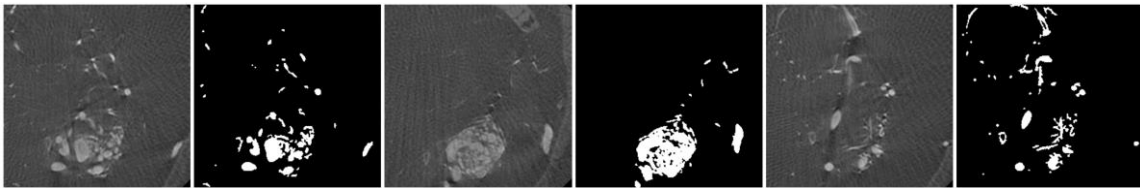
A set of images of 9 patients who underwent 3DRA exams were produced as part of a previous work [6]. The acquisitions were performed with a Philips Allura angiographic unit (Philips Healthcare, Netherlands) after injection of 28 ml of a contrast agent to highlight the vessels, at 4 ml/s with a rotation of 210° and a delay of 3 s between injection and acquisition. Each obtained volume was a 250×256×256 cubic matrix with a voxel size varying from 0.29 mm<sup>3</sup> to 0.49 mm<sup>3</sup>. The 3D images were centered on regions of interest containing the AVM, the veins and the arteries connected to the nidus. Thus, 2043 slices around the region of interest were selected and augmented applying random rotations, shifts and zoom factors during the training and validation phases to enlarge the set of examples seen by the neural network.

### 3.2 Data labeling

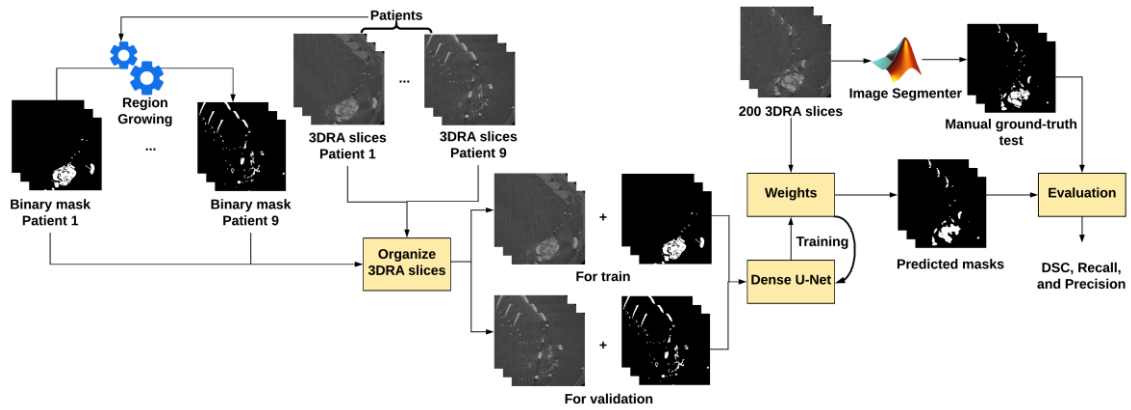
These volumes were organized into 2D slices. Each slice was labeled using a region-growing algorithm in its optimal configuration for 3DRA AVM images [6]. A set of 1843 3DRA masks underwent a qualitative validation by neuroradiologists. During the validation process, they visually compared the region growing results with the volume rendering and 3D MIP rendering obtained from the imaging station at the Hôpital Fondation Adolphe de Rothschild. The experts carefully analyzed the segmentations and recognized their similarity to the reference reconstructions provided by the imaging station. Notably, they emphasized the continuity of the vessels, the absence of holes, the smooth surfaces, and the presence of all vascular structures.

These automatically generated binary masks allowed us to avoid manual labeling of a large amount of 3DRA slices to train and validate our model. As the process of labeling is time-consuming, the training phase of our model was initiated using pairs of images and their binary masks generated by the region-growing algorithm. These masks were considered as the ground truth, despite a few imperfections observed in some segmented masks.

Finally, to exclusively use all the previous slices in the training and validation process, additional 200 slices of 3DRA images were manually segmented using the MATLAB "Image Segmenter" tool specifically for the test set (Fig. 1). This manual database was validated by a neuroradiologist with over 20 years of experience in endovascular AVM treatment and constitutes our ground-truth test set for the validation of our predictions. We summarize the 3DRA data organization for cerebrovascular vessel segmentation in the framework shown in Fig. 2; it contains training and testing phases.



**Fig 1.** Examples of 3DRA images and their manually segmented ground-truth used for the test phase of the neural network developed model.



**Fig 2.** Overview of the proposed framework: 3DRA images of 9 patients were segmented using the region-growing algorithm [6] in its optimal configuration. the images and their binary masks were organized to train and validate the model. The model was then tested on a manual ground-truth of 200 3DRA slices to evaluate its performance.

### 3.3 Proposed network architecture

CNNs struggle to learn useful feature information from low level as the architecture depth grows. Thus, the vanishing gradient descent problem occurs when the networks go deeper so that the information from the input layer until the output layer (and the backpropagation of the gradient) gets vanished before reaching the final destination. Therefore, the learning ability of the model is seriously affected. To ensure the reuse of features learned by the model, Dense networks were designed to connect layers with each other [14]. These networks were integrated to the U-Net [9] model for segmentation purposes.

Inspired by the “One Hundred Layers Tiramisu” network [15], a U-Net\_Densenet121 was developed in this study for cerebrovascular segmentation. This model combines the standard U-Net and the Densenet121 models [14] in which dense blocks are integrated into the encoder and decoder paths as shown in Fig. 3.

The Encoder (contraction path) is used to obtain the feature representations; it includes a succession of dense and transition blocks (Table I). The dimensions of the feature maps inside the dense blocks remain constant within a block, but the number of filters changes between them due to the growth rate which increases the channel dimension at every layer by 4 channels. This block contains batch normalization (BN), rectified linear unit (ReLU),  $1 \times 1$  convolution (Conv),  $3 \times 3$  convolution, and dropout layer. The number of convolutions inside a dense block changes from 6, 12, 24 to 16 convolutions as in a Densenet121 architecture. The layers between dense blocks are transition layers in charge of the downsampling. These transitions include batch normalization,  $1 \times 1$  convolution, that halves the number of feature channels, dropout, and  $2 \times 2$  pooling.

The Decoder (expansion path) consists of an upsampling of the feature map followed by a concatenation with the corresponding feature map from the encoder path, and a dense block that has the same structure as those in the contraction path (Table II). In this path, the number of convolutions inside a dense block changes in an inverse order compared to the encoder path from 16, 24, 12 to 6 convolutions. The vessel and background areas are segmented using the sigmoid activation function.

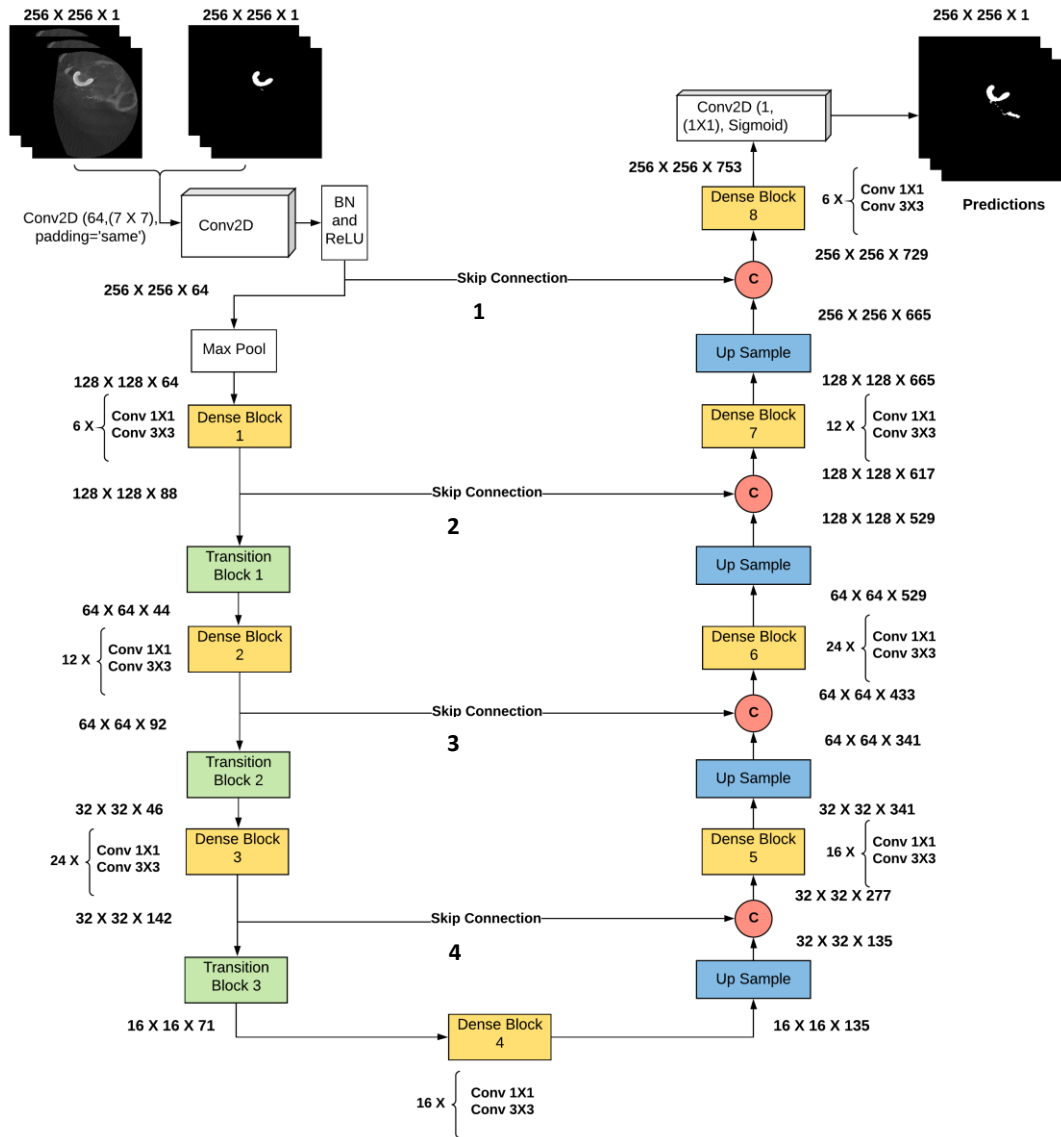


Fig 3. The network architecture of the proposed model U-Net\_Densenet121.

**Table I**

Encoder architecture of U-Net\_Densenet121: evolution of image dimensions and operations. Dense blocks maintain the same dimension while transition blocks halve them.

Image dimension	Block type	Operation
$256 \times 256 \times 1$	Input	$Conv((7 \times 7), 64)$ $BN + ReLU$
$256 \times 256 \times 64$	Pooling	$Max\ pooling$
$128 \times 128 \times 64$	Dense 1	$filters = 64; 6 \times \begin{cases} BN + ReLU \\ Conv((1 \times 1), filters) \\ BN + ReLU \\ Conv((3 \times 3), filters) \\ Dropout(0.2) \\ filters = filters + 4 \end{cases}$
$128 \times 128 \times 88$	Transition 1	$filters = 88;$ $\begin{cases} BN + ReLU \\ Conv((1 \times 1), filters/2) \\ Dropout(0.2) \\ Average\ pooling((2, 2), strides = (2, 2)) \end{cases}$
$64 \times 64 \times 44$	Dense 2	$filters = 44; 12 \times \begin{cases} BN + ReLU \\ Conv((1 \times 1), filters) \\ BN + ReLU \\ Conv((3 \times 3), filters) \\ Dropout(0.2) \\ filters = filters + 4 \end{cases}$
$64 \times 64 \times 92$	Transition 2	$filters = 92;$ $\begin{cases} BN + ReLU \\ Conv((1 \times 1), filters/2) \\ Dropout(0.2) \\ Average\ pooling((2, 2), strides = (2, 2)) \end{cases}$
$32 \times 32 \times 46$	Dense 3	$filters = 46; 24 \times \begin{cases} BN + ReLU \\ Conv((1 \times 1), filters) \\ BN + ReLU \\ Conv((3 \times 3), filters) \\ Dropout(0.2) \\ filters = filters + 4 \end{cases}$
$32 \times 32 \times 142$	Transition 3	$filters = 142;$ $\begin{cases} BN + ReLU \\ Conv((1 \times 1), filters/2) \\ Dropout(0.2) \\ Average\ pooling((2, 2), strides = (2, 2)) \end{cases}$
$16 \times 16 \times 71$	Dense 4 (The bottom of the U architecture)	$filters = 71; 16 \times \begin{cases} BN + ReLU \\ Conv((1 \times 1), filters) \\ BN + ReLU \\ Conv((3 \times 3), filters) \\ Dropout(0.2) \\ filters = filters + 4 \end{cases}$

**Table II**

Decoder architecture of U-Net\_Densenet121: evolution of image dimensions and operations. The feature maps are upsampled and concatenated with those generated by the encoder before being passed to the dense blocks.

Image dimension	Block type	Operation
$16 \times 16 \times 135$	Upsampling	$upsampling(2, 2) + concatenate(skip\ connection\ 4)$
$32 \times 32 \times 277$	Dense 5	$filters = 277; 16 \times \begin{cases} BN + ReLU \\ Conv((1 \times 1), filters) \\ BN + ReLU \\ Conv((3 \times 3), filters) \\ Dropout(0.2) \\ filters = filters + 4 \end{cases}$
$32 \times 32 \times 341$	Upsampling	$upsampling(2, 2) + concatenate(skip\ connection\ 3)$
$64 \times 64 \times 433$	Dense 6	$filters = 433; 24 \times \begin{cases} BN + ReLU \\ Conv((1 \times 1), filters) \\ BN + ReLU \\ Conv((3 \times 3), filters) \\ Dropout(0.2) \\ filters = filters + 4 \end{cases}$
$64 \times 64 \times 529$	Upsampling	$upsampling(2, 2) + concatenate(skip\ connection\ 2)$
$128 \times 128 \times 617$	Dense 7	$filters = 617; 12 \times \begin{cases} BN + ReLU \\ Conv((1 \times 1), filters) \\ BN + ReLU \\ Conv((3 \times 3), filters) \\ Dropout(0.2) \\ filters = filters + 4 \end{cases}$
$128 \times 128 \times 665$	Upsampling	$upsampling(2, 2) + concatenate(skip\ connection\ 1)$
$256 \times 256 \times 729$	Dense 8	$filters = 729; 6 \times \begin{cases} BN + ReLU \\ Conv((1 \times 1), filters) \\ BN + ReLU \\ Conv((3 \times 3), filters) \\ Dropout(0.2) \\ filters = filters + 4 \end{cases}$
$256 \times 256 \times 753$	Output	$Conv((1 \times 1), 1)$ $Sigmoid$
$256 \times 256 \times 1$		

### 3.4 Loss function

The combination of the U-Net and the Densenet121 aimed to enhance the transmission efficiency of features and enrich the feature information extracted by the network for a more stable training. Besides the choice of the network architecture, the selection of a suitable loss function is also very important.

The acquired 3DRA dataset has several problems limiting the training of the CNN to generate precise segmentations. These images suffer from class imbalance which means that the proportion occupied by the cerebral vessels, including AVMs is small compared to the image background. Thus, the network will have more ability to predict a background than a vascular structure. In addition, cerebral vessels dimensions are very small, therefore their segmentation must be very precise.

For these reasons, we were interested in the FT loss [23] and the effect of its combination with CE on the segmentation of small structures and imbalanced datasets. As mentioned in Section 2.2, this function allows the network to focus more on less accurate predictions that have been misclassified.

### 3.5 Experimental setup

The experiments were performed on an NVIDIA GeForce RTX 3090 (24 GB) graphic card with CUDA compute capability (V11.3) under python 3.6 using Keras based on Tensorflow GPU support. The models were trained for 150 epochs. The learning process was conducted with Adam optimizer [28] and a learning rate of  $10^{-4}$  decreasing by a factor of 0.9 with patience value set to 5 epochs if no improvement was noticed on the Dice Similarity Coefficient (DSC) of the validation set.

The training and validation phases of these experiments were conducted on the 3DRA slices of the 9 patients and tested on the additional 200 slices (see Section 3.1 and 3.2). The slices in each batch (of size 4) for each epoch were augmented on the fly during the learning process. The training and validation sets are free of data leakage, so patients involved in the training process are not included in the validation and vice versa. The dataset is organized as follows:

- *Train*: containing 1229 image slices and their binary masks obtained by region-growing.
- *Validation*: containing 614 slices of images and their binary masks obtained by region-growing.
- *Test (manual ground-truth)*: containing 200 slices of an AVM patient manually segmented and validated by a neuroradiologist.

Thus, to choose the best model that fits our 3DRA images, two experiments have been conducted, in which a set of hyperparameters and state-of-the-art architectures were evaluated:

- 1) *Experiment 1*: aimed to explore the effect of the Dice and FT [23] loss functions and their combination with the CE on the quality of the segmentation.
- 2) *Experiment 2*: aimed to benchmark our solution with state-of-the-art architectures.

Different pixel intensity thresholds were tested to binarize the predicted masks, the best threshold was chosen for the test phase. The experiments results are shown in Section 4.

## 4. RESULTS

In this section, the results of the experiments are presented to demonstrate the effectiveness of the proposed method and the choices made to enhance the segmentation of cerebral vessels. The evaluation of the models is done by calculating the mean of each evaluation metric such as: DSC, precision and recall for 3 runs with random initialization for the Adam optimizer. The mean absolute difference between precision and recall ( $|\text{Precision}-\text{Recall}|$ ) has also been reported on the test set.

### 4.1 Experiment 1

To choose the best function between the Dice and FT [23] losses to train the model, 3DRA images with their binary masks obtained by region-growing algorithm [6] have been used to feed our architecture.

As mentioned in Section 2.2, The  $\alpha$ ,  $\beta$  parameters control the balance between precision and recall. Thus, to achieve the best trade-off, different combinations of  $\alpha$ ,  $\beta$  and  $\gamma$  were tested, results are summarized in Table III. The best DSC of 79.38% is reached when  $\alpha = 0.6$ ,  $\beta = 0.4$  and  $\gamma = 4/3$ .

Subsequently, the FT loss, along with its optimal set of hyperparameters determined earlier, was combined with the CE loss. This combination also shows a very small difference between precision and recall (01.75) on the test set, in contrast to the Dice loss which gave a lower DSC with a significant difference of 17.16% between precision and recall (see Table IV).

As shown in Table IV, the absolute difference between recall and precision was reduced from 17.16 to 3.54 in the case of Dice loss. The focal Tversky component ensures a global overlapping rate between the ground-truth and the predictions and a better balance between precision and recall with its  $\alpha$  and  $\beta$  parameters. The focal parameter  $\gamma$  prevents premature convergence to a suboptimal solution when reaching the end of the training process where pixels are classified with high confidence. As shown in Fig. 4, the model trained with the Dice loss combined with CE function (see Fig. 4, third row) over-segments the 3DRA images (segmentation of additional anatomical structures), unlike the model trained with the FT and CE functions (see Fig. 4, last row) which focuses only on vessels and AVMs. Therefore, this model



generates masks (see Fig. 5, last row) that better delineate and separate the vessels compared to the 3D region-growing algorithm [6] used in the training phase (see Fig. 5, second row).

**Table III**

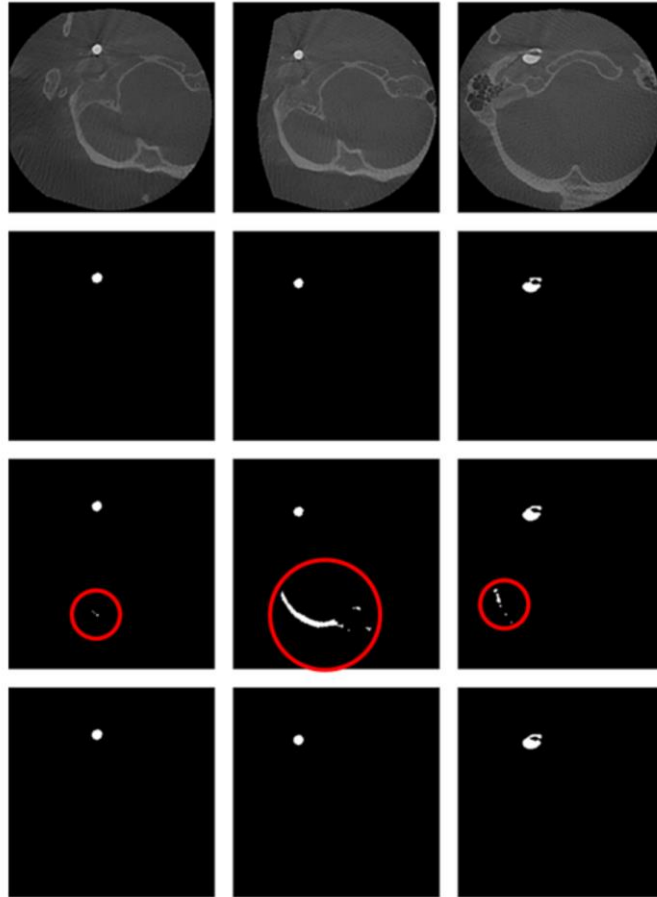
Segmentation results on the test set with different  $\alpha$ ,  $\beta$  and  $\gamma$  hyperparameters of the FT loss.

Best model	$\gamma = 4/3$	$\gamma = 2$
$\alpha = 0.6, \beta = 0.4$	<b>DSC: 79.38%</b>	DSC: 79.15%
	Precision: 82.75%	Precision: 79.00%
	Recall: 81.00%	Recall: 82.32%
	Precision-Recall : 01.75	Precision-Recall : 03.32
$\alpha = 0.7, \beta = 0.3$	DSC: 79.28%	DSC: 79.05%
	Precision: 83.98%	Precision: 85.69%
	Recall: 80.11%	Recall: 79.36%
	Precision-Recall : 03.87	Precision-Recall : 06.33
$\alpha = 0.8, \beta = 0.2$	DSC: 79.20%	DSC: 79.14%
	Precision: 89.57%	Precision: 87.63%
	Recall: 75.24%	Recall: 77.22%
	Precision-Recall : 14.33	Precision-Recall : 10.41
$\alpha = 0.9, \beta = 0.1$	DSC: 77.66%	DSC: 79.24%
	Precision: 91.04%	Precision: 88.24%
	Recall: 71.01%	Recall: 75.94%
	Precision-Recall : 20.03	Precision-Recall : 12.30

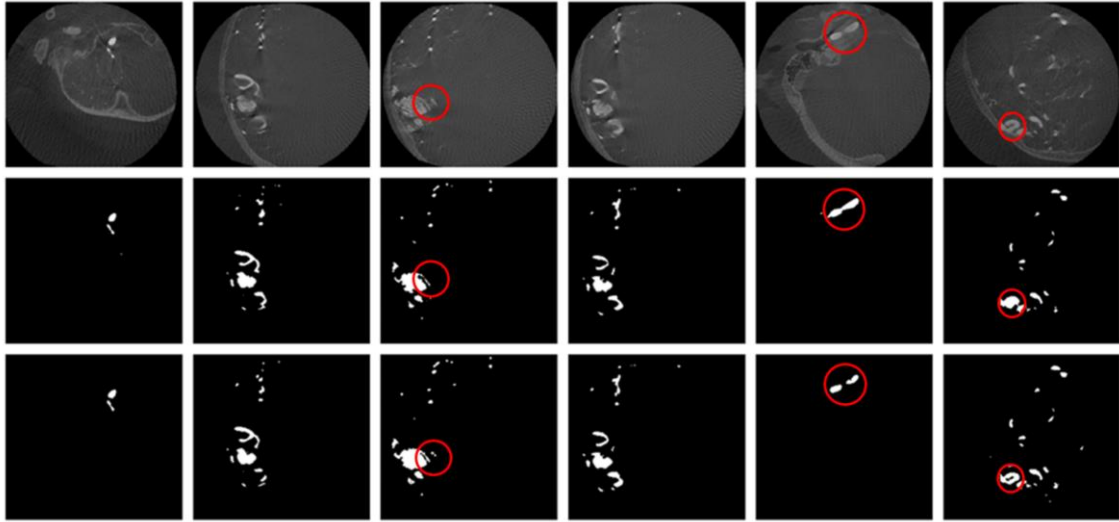
**Table IV**

Segmentation results with Dice and FT losses and their combination with CE loss.

Best model	DSC	Precision	Recall	Precision-Recall
Dice ( $\alpha = 0.5, \beta = 0.5$ )	<i>Train:</i> 84.35 %	<i>Train:</i> 85.61 %	<i>Train:</i> 86.84 %	17,16
	<i>Test:</i> 79.18 %	<i>Test:</i> 73.55 %	<i>Test:</i> 90.71 %	
CE + Dice	<i>Train:</i> 83.04 %	<i>Train:</i> 84.57 %	<i>Train:</i> 85.90 %	03.54
	<i>Test:</i> 78.87 %	<i>Test:</i> 80.16 %	<i>Test:</i> 82.04 %	
CE + FT ( $\alpha = 0.6, \beta = 0.4, \gamma = 4/3$ )	<i>Train:</i> 83.49 %	<i>Train:</i> 82.85 %	<i>Train:</i> 89.40 %	02.42
	<i>Test:</i> <b>80.43 %</b>	<i>Test:</i> 84.15 %	<i>Test:</i> 81.73 %	



**Fig 4.** Segmentation results of 3DRA images (first row): comparison between masks generated by region-growing [6] (second row) and predicted masks using U-Net\_Densenet121 model with both Dice + CE loss (third row) and FT + CE loss (fourth row). Extra segmented anatomical structures are circled in red.



**Fig 5.** Segmentation results of 3DRA images (first row): comparison between masks generated by region-growing [6] (second row) and predicted masks using U-Net\_Densenet121 model with FT + CE loss (third row). Red circles illustrate the improvement in segmentation brought by the neural network.

#### 4.2 Experiment 2

Once the appropriate loss function was selected based on the results of the previous experiment, our research proceeded by evaluating our model against state-of-the-art architectures including U-Net++[29], U-Net3+ [30], ResU-Net [31], Attention U-Net [32] and Inception U-Net [33].

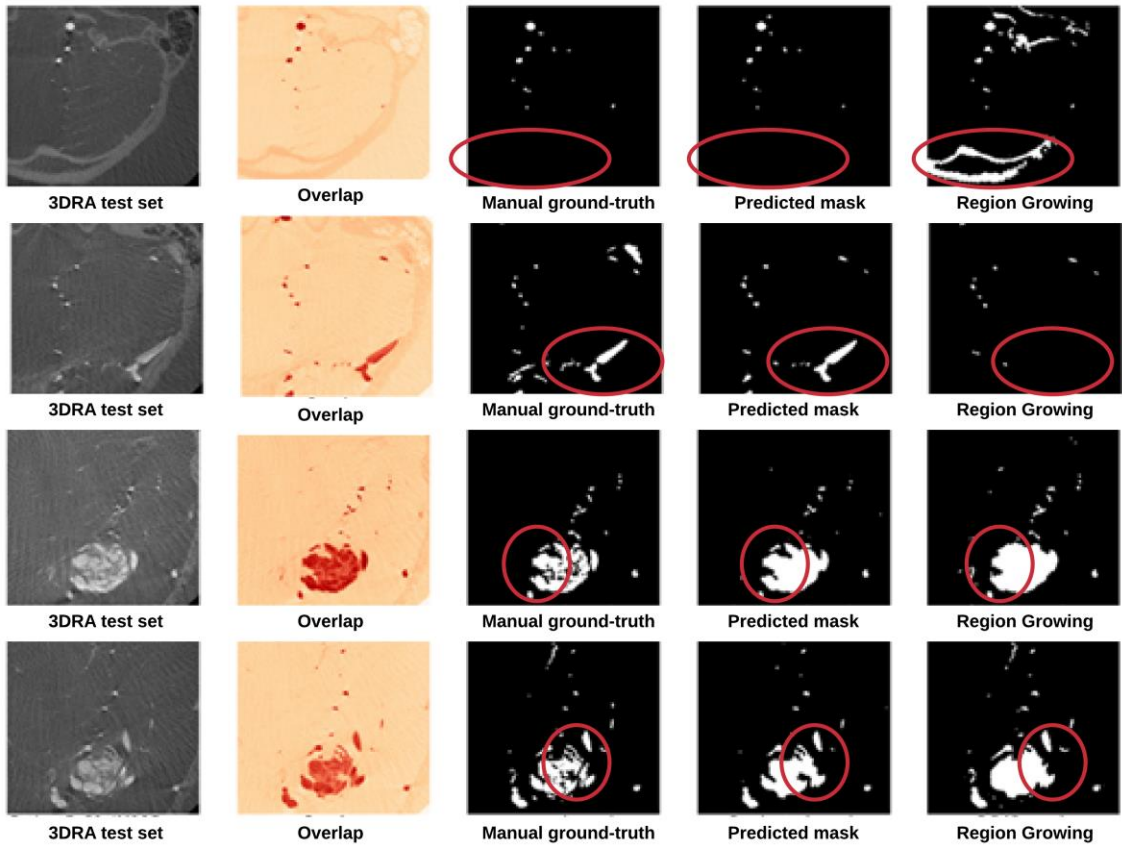
The results of the comparative analysis are summarized in Table V. As shown in the table, our network got a DSC of 80.43% compared to the region-growing algorithm which had a score of 72.35% in its optimal configuration detailed in [6]. Compared to other U-Net like variants, the U-Net\_Densenet121 achieved the best results in the segmentation metrics. Table V shows that the other architectures overfitted on the training set and did not generalize enough. Other than DSC, recall and precision metrics, the developed CNN minimizes better the gap between precision and recall; it shows a difference of 02.42% while the others had a difference of over 20%.

**Table V**

Segmentation results with FT loss using different U-Net-like architectures and the region growing algorithm [6].

Best model	DSC	Precision	Recall	Precision-Recall
U-Net_Densenet121	<i>Train:</i> 83.49%	<i>Train:</i> 82.85%	<i>Train:</i> 89.40%	02.42
	<b><i>Test:</i> 80.43%</b>	<i>Test:</i> 84.15%	<i>Test:</i> 81.73%	
ResU-Net	<i>Train:</i> 79.55%	<i>Train:</i> 70.14%	<i>Train:</i> 98.22%	21.45
	<i>Test:</i> 76.96%	<i>Test:</i> 90.92%	<i>Test:</i> 69.47%	
Attention U-Net	<i>Train:</i> 88.06%	<i>Train:</i> 86.69%	<i>Train:</i> 92.78%	24.22
	<i>Test:</i> 77.30%	<i>Test:</i> 92.63%	<i>Test:</i> 68.40%	
Inception U-Net	<i>Train:</i> 89.56%	<i>Train:</i> 88.19%	<i>Train:</i> 94.13%	23.65
	<i>Test:</i> 75.67%	<i>Test:</i> 90.95%	<i>Test:</i> 67.30%	
U-Net++	<i>Train:</i> 86.17%	<i>Train:</i> 85.23%	<i>Train:</i> 91.27%	20.57
	<i>Test:</i> 75.43%	<i>Test:</i> 88.86%	<i>Test:</i> 68.29%	
U-Net3+	<i>Train:</i> 81.36%	<i>Train:</i> 72.54%	<i>Train:</i> 98.48%	21.76
	<i>Test:</i> 75.46%	<i>Test:</i> 90.20%	<i>Test:</i> 68.45%	
Region-growing	<i>Test:</i> 72.35%	<i>Test:</i> 63.00%	<i>Test:</i> 92.85%	29.85

To sum up, some results of the predicted masks on the test set are illustrated in Fig. 6 along with their comparison with masks of the region-growing algorithm and the manual ground-truth.



**Fig 6.** Segmentation results on test set of 3DRA images (first column): comparison between the manual ground-truth (third column), predicted masks using U-Net\_Densenet121 model with FT and CE losses (fourth column) and masks generated by region-growing [6] (fifth column). Also, overlap between 3DRA images and predicted masks using our model are shown in the second column. Differences between manual annotations and predicted masks with both neural network and region-growing are circled in red.

## 5. DISCUSSION

This study highlights the potential of deep learning in accurately segmenting vessels in 3DRA images of patients diagnosed with cerebral AVM. The following neural network: Dense U-Nets were trained with Dice and FT loss functions and their combination with CE loss.

The initial findings of our study confirmed that incorporating the FT [23] resulted in a more precise model, enhancing both precision and recall. This improvement directly contributed to the enhanced extraction of cerebral vessels. The best models were those that combined the CE loss that captures local context and precision at a pixel level with the FT loss.

Besides the improvement of the DSC metric, in our case, the CE component in the loss function provided more stability in training, so that the difference between precision and recall was smaller (see Table III and IV). Therefore, this model generated masks that delineate and separate better the vessels compared to the region-growing algorithm.

As a result of the previous experiments, the loss function that combines FT and CE methods was selected for comparing our architecture with state-of-the-art U-Net-like models. From the different hyperparameters, the model that best fitted the 3DRA slices was the U-Net\_Densenet121 (as shown in Table V).

Different pixel intensity thresholds were tested to binarize the predicted masks, the best threshold was chosen each time for the test phase for each model to improve the quality of the segmentation and find a common ground between precision and recall metrics.

Our architecture based on Densenet121 added to both paths of the U-Net improved the segmentation results and unlike the other state-of-the-art U-Net like architectures, the U-Net\_Densenet121 generalized better showing a good understanding of the binary masks generated by the region growing algorithm used to initiate the training process.

The segmented 3DRA slices generated by the developed neural network were compared to those produced by the region-growing algorithm. Compared to other U-Net like variants, the U-Net\_Densenet121 achieved the best results in the segmentations metrics (see table V).

For all the previous experiments, the developed variant of U-Net (Fig. 3) was trained on 3DRA images and their binary masks obtained by the region-growing-based method [6]. As in the work of Ghafoorian et al. [34], our findings demonstrate that the learning of a neural network can be initiated with the results of an automatic or a semi-automatic method and manages to outperform it by being more precise in its classification of pixels. This is indeed an interesting result, since the training of CNNs needs a very large number of annotated images which is time consuming, especially in pathologies like AVM where vessels are tangled.

Finally, the results were validated on a set of 200 3DRA images manually annotated. This manual ground-truth was used as a reference to compare the developed deep learning approach with the region-growing algorithm and the U-Net-like variants in literature. It is difficult to acquire more images because cerebral AVMs are rare and often discovered accidentally. However, this small series of tests allowed us to evaluate our segmentation of the vessels and the AVM nidus. This segmentation is the first step of a complete geometrical/anatomical characterization aiming at improving the 3D visualization of the malformed vessels: veins, arteries and AVM nidus. This 3D visualization reduces the complexity of this rare pathology and helps physicians better plan embolization procedures.

## 6. CONCLUSION

This paper focuses on the automatic extraction of cerebral vascular structures in patients with AVM using deep learning approaches. The segmentation of 3D rotational angiography images was accomplished through the development of a U-Net-like architecture. The combination of loss functions as well as the developed deep architecture allowed us to have a robust model that segments small vessels.

Improved segmentation of small vessels is crucial for automating the characterization of vessel geometry. The knowledge gained from this process plays a key role in strengthening our understanding of the arterial supply and venous flow associated with arteriovenous malformations (AVMs). These insights have the potential to reduce procedural complications, optimize patient outcomes and help neuroradiologists identify the most effective embolization path. By providing detailed information on the complex vascular structures of AVMs, deep learning segmentation enables precise treatment planning, targeted interventions, and improved patient care.

To further improve this work, the constitution of a robust and shareable ground-truth database of AVM cases is suggested. This extensive dataset will not only facilitate the evaluation of various deep learning approaches, such as transformers or 3D U-Net-like architectures, but also enable multiclass segmentation of AVMs. The automatic extraction of the entire vascular tree, including feeding arteries, draining veins, and the AVM nidus, is of great importance. This extraction will assist physicians in gaining a better understanding of the angioarchitecture, which in turn will help them anticipate potential challenges during interventions. It is also important to incorporate biomechanical measurements, such as intravascular blood pressure and blood flow velocities, into the analysis. By integrating these measurements with the deep learning-based segmentation of AVMs, a more comprehensive understanding of the hemodynamic characteristics may be achieved. This integration would provide valuable insights into the behavior of blood flow within the AVM, helping in the assessment of its functional impact and guiding treatment planning. Furthermore, the combination of biomechanical measurements and advanced segmentation techniques can facilitate the development of personalized interventions.

### **Conflict of interest**

None Declared.

## References

- [1] S. Bash, J.P. Villablanca, R. Jahan, G. Duckwiler, M. Tillis, C. Kidwell, J. Saver, J. Sayre, Intracranial vascular stenosis and occlusive disease: evaluation with CT angiography, MR angiography, and digital subtraction angiography, *AJNR Am. J. Neuroradiol.* 26 (2005) 1012–1021.
- [2] R. Blanc, A. Seiler, T. Robert, H. Baharvahdat, M. Lafarge, J. Savatovsky, J. Hodel, G. Ciccio, D. Chauvet, S. Pistocchi, B. Bartolini, H. Redjem, M. Piotin, Multimodal angiographic assessment of cerebral arteriovenous malformations: a pilot study, *J. NeuroInterventional Surg.* 7 (2015) 841–847.
- [3] A. Feddal, S. Escalard, F. Delvoeye, R. Fahed, J.P. Desilles, K. Zuber, H. Redjem, J.S. Savatovsky, G. Ciccio, S. Smajda, M. Ben Maacha, M. Mazighi, M. Piotin, R. Blanc, Fusion Image Guidance for Supra-Aortic Vessel Catheterization in Neurointerventions: A Feasibility Study, *Am. J. Neuroradiol.* 41 (2020) 1663–1669.
- [4] C.S. Ogilvy, P.E. Stieg, I. Awad, R.D. Brown, D. Kondziolka, R. Rosenwasser, W.L. Young, G. Hademenos, Recommendations for the Management of Intracranial Arteriovenous Malformations: A Statement for Healthcare Professionals From a Special Writing Group of the Stroke Council, *American Stroke Association, Stroke.* 32 (2001) 1458–1471.
- [5] H. Baharvahdat, R. Blanc, R. Fahed, A. Pooyan, A. Mowla, S. Escalard, F. Delvoeye, J.P. Desilles, H. Redjem, G. Ciccio, S. Smajda, M. Hamdani, M. Mazighi, M. Piotin, Endovascular treatment as the main approach for Spetzler-Martin grade III brain arteriovenous malformations, *J. NeuroInterventional Surg.* 13 (2021) 241–246.
- [6] Y. Chenoune, O. Tankyevych, F. Li, M. Piotin, R. Blanc, E. Petit, Three-dimensional segmentation and symbolic representation of cerebral vessels on 3DRA images of arteriovenous malformations, *Comput. Biol. Med.* 115 (2019) 103489.
- [7] C.-Y. Hsu, M. Ghaffari, A. Alaraj, M. Flannery, X.J. Zhou, A. Linninger, Gap-free Segmentation of Vascular Networks with Automatic Image Processing Pipeline, *Comput Biol Med.* 82 (2017) 29–39.
- [8] D. Babin, A. Pižurica, L. Velicki, V. Matić, I. Galić, H. Leventić, V. Zlokolica, W. Philips, Skeletonization method for vessel delineation of arteriovenous malformation, *Comput. Biol. Med.* 93 (2018) 93–105.
- [9] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: *Int. Conf. Med. Image Comput. Comput.-Assist. Interv.*, Springer, 2015: pp. 234–241.
- [10] G. Tetteh, V. Efremov, N.D. Forkert, M. Schneider, J. Kirschke, B. Weber, C. Zimmer, M. Piraud, B.H. Menze, Deepvesselnet: Vessel segmentation, centerline prediction, and bifurcation detection in 3-d angiographic volumes, *ArXiv Prepr. ArXiv180309340.* (2018).
- [11] M. Livne, J. Rieger, O.U. Aydin, A.A. Taha, E.M. Akay, T. Kossen, J. Sobesky, J.D. Kelleher, K. Hildebrand, D. Frey, A U-Net deep learning framework for high performance vessel segmentation in patients with cerebrovascular disease, *Front. Neurosci.* 13 (2019) 97.
- [12] P. Sanches, C. Meyer, V. Vigon, B. Naegel, Cerebrovascular Network Segmentation of MRA Images With Deep Learning, in: *IEEE 16th Int. Symp. Biomed. Imaging ISBI*, 2019: pp. 768–771.
- [13] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: *IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016: pp. 770–778.
- [14] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely connected convolutional networks, in: *IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017: pp. 4700–4708.
- [15] S. Jégou, M. Drozdal, D. Vazquez, A. Romero, Y. Bengio, The one hundred layers tiramisù: Fully convolutional densenets for semantic segmentation, in: *IEEE Conf. Comput. Vis. Pattern Recognit. Workshop*, 2017: pp. 11–19.
- [16] J. Dolz, I.B. Ayed, C. Desrosiers, Dense multi-path U-Net for ischemic stroke lesion segmentation in multiple image modalities, in: *Int. MICCAI Brainlesion Workshop*, Springer, 2018: pp. 271–282.
- [17] C. Meng, K. Sun, S. Guan, Q. Wang, R. Zong, L. Liu, Multiscale dense convolutional neural network for DSA cerebrovascular segmentation, *Neurocomputing.* 373 (2020) 123–134.
- [18] L.R. Dice, Measures of the amount of ecologic association between species, *Ecology.* 26 (1945) 297–302.
- [19] T. A. Soomro, A. J. Afifi, J. Gao, O. Hellwich, M. Paul, L. Zheng, Strided U-Net Model: Retinal Vessels Segmentation using Dice Loss, in: *Digit. Image Comput. Tech. Appl. DICTA*, 2018: pp. 1–8.



- [20] R. Brügger, C.F. Baumgartner, E. Konukoglu, A partially reversible U-Net for memory-efficient volumetric image segmentation, in: *Int. Conf. Med. Image Comput. Comput.-Assist. Interv.*, Springer, 2019: pp. 429–437.
- [21] A. Tversky, Features of similarity, *Psychol. Rev.* 84 (1977) 327.
- [22] S.S.M. Salehi, D. Erdogmus, A. Gholipour, Tversky loss function for image segmentation using 3D fully convolutional deep networks, in: *Int. Workshop Mach. Learn. Med. Imaging*, Springer, 2017: pp. 379–387.
- [23] N. Abraham, N.M. Khan, A novel focal Tversky loss function with improved attention U-net for lesion segmentation, in: *IEEE 16th Int. Symp. Biomed. Imaging ISBI*, 2019: pp. 683–687.
- [24] E. Colombo, T. Fick, G. Esposito, M. Germans, L. Regli, T. van Doormaal, Segmentation techniques of brain arteriovenous malformations for 3D visualization: a systematic review, *Radiol. Med. (Torino)*. 127 (2022) 1333–1341.
- [25] T. Wang, Y. Lei, S. Tian, X. Jiang, J. Zhou, T. Liu, S. Dresser, W.J. Curran, H. Shu, X. Yang, Learning-based automatic segmentation of arteriovenous malformations on contrast CT images in brain stereotactic radiosurgery, *Med. Phys.* 46 (2019) 3133–3141.
- [26] W.C. You, Y.T. W, J.S. Hong, Y. Lin, M.S. Sun, Y.F. Lu, K.M. Chen, T.H. Huang, W.K. Lee, Detection and Segmentation of Arteriovenous Malformation Lesions Using a Two-Stage Deep Learning Strategy, *Int. J. Radiat. Oncol. Biol. Phys.* 114 (2022) e108.
- [27] C. García, Y. Fang, J. Liu, A.P. Narata, J.I. Orlando, I. Larrabide, A deep learning model for brain vessel segmentation in 3DRA with arteriovenous malformations, *ArXiv Prepr. ArXiv221002416*. (2022).
- [28] D.P. Kingma, J. Ba, Adam: A Method for Stochastic Optimization, *ArXiv Prepr. ArXiv14126980*. (2017).
- [29] Z. Zhou, M.M.R. Siddiquee, N. Tajbakhsh, J. Liang, UNet++: A Nested U-Net Architecture for Medical Image Segmentation, *ArXiv Prepr. ArXiv180710165*. (2018).
- [30] H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, Y. Iwamoto, X. Han, Y.-W. Chen, J. Wu, Unet 3+: A full-scale connected unet for medical image segmentation, in: *IEEE Int. Conf. Acoust. Speech Signal Process. ICASSP, IEEE*, 2020: pp. 1055–1059.
- [31] F.I. Diakogiannis, F. Waldner, P. Caccetta, C. Wu, ResUNet-a: a deep learning framework for semantic segmentation of remotely sensed data, *ISPRS J. Photogramm. Remote Sens.* 162 (2020) 94–114.
- [32] O. Oktay, J. Schlemper, L.L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N.Y. Hammerla, B. Kainz, Attention u-net: Learning where to look for the pancreas, *ArXiv Prepr. ArXiv180403999*. (2018).
- [33] I. Delibasoglu, M. Cetin, Improved U-Nets with inception blocks for building detection, *J. Appl. Remote Sens.* 14 (2020) 044512.
- [34] M. Ghafoorian, J. Teuwen, R. Manniesing, F.-E. de Leeuw, B. van Ginneken, N. Karssemeijer, B. Platel, Student beats the teacher: deep neural networks for lateral ventricles segmentation in brain MR, in: *Med. Imaging 2018 Image Process.*, 2018: p. 105742U.