A plug-and-play framework for curvilinear structure segmentation based on a learned reconnecting regularization

Sophie Carneiro-Esteves[a,b,*], Antoine Vacavant[b], Odyssée Merveille[a]

[a]*Univ Lyon, INSA-Lyon, Université Claude Bernard Lyon 1, UJM-Saint Etienne, CNRS, Inserm, CREATIS UMR 5220, U1294,F-69100, Lyon, France*
[b]*Université Clermont Auvergne, CNRS, SIGMA Clermont, Institut Pascal, F-63000, Clermont-Ferrand, France*

Abstract

Curvilinear structures are present in various fields in image processing such as blood vessels in medical imaging or roads in remote sensing. Their detection is crucial for many applications. In this article, we propose an unsupervised plug-and-play framework for the segmentation of curvilinear structures that focuses on the preservation of their connectivity. This framework includes an algorithm for generating realistic pairs of connected/disconnected curvilinear structures and a reconnecting regularization operator that can be learned from a synthetic dataset. Once learned, this regularization operator can be plugged into a variational segmentation scheme and used to segment curvilinear structure images without requiring annotations. We demonstrate the interest of our approach on the segmentation of vascular images both in 2D and 3D and compare its results with classic unsupervised and deep learning-based approach. Comparative evaluations against unsupervised classic and deep learning-based methods highlight the superior performance of our approach, showcasing remarkable improvements in preserving the connectivity of curvilinear structures (approximately 90% in 2D and 70% in 3D).We finally showcase the good generalizability behavior of our approach on two different applications : road cracks and porcine corneal cells segmentations.

Keywords: segmentation, curvilinear structures, plug-and-play, connectivity.

1. Introduction

Curvilinear structures are complex objects that exhibit a thin, curved, and elongated shape. They can be observed in many image processing fields, such as remote sensing (rivers and roads networks) or medical imaging (blood vessels and neurons). The detection (*i.e.*, segmentation) of curvilinear structures in images is the first crucial task of many applications such as blood flow simulation, urban planning, or autonomous navigation. Despite more than thirty years of research, it remains an open problem because of the complexity of its geometry and appearance in images. Curvilinear structures are sparsely distributed in images, they usually exhibit a high level of tortuosity, and are often organized in networks. Additionally, curvilinear structures are most of the time low-contrasted to the point that they can easily be altered by noise. Consequently, generic segmentation methods lead to poor results when applied to curvilinear structures. In particular, thin curvilinear structures are often missed, and the connectivity of curvilinear structure networks is rarely preserved. Some segmentation approaches tackling curvilinear structure specificities have been developed over the years. Regularization terms for variational formulations were proposed to capture specific geometric properties. The Mumford and Shah functional [1] was used to favor piecewise smoothness along the curvilinear structure intensity profile and promote connectivity [2]. However, solving this functional poses computational challenges, necessitating approximations [3] and reducing its interest. Merveille *et al.* proposed a directional total variation [4] aiming at denoising the curvilinear structures specifically along their main direction in order to reduce disconnections coming from isotropic denoising. These strategies employ indirect manners to ensure connectivity, given that connectivity is a complex geometric property that is challenging to model explicitly.

Recently, deep learning has shown promise in detecting curvilinear structures. Deep learning models can impose properties relevant to curvilinear structures,

---

*Corresponding author
Email address:* sophie.carneiro@creatis.insa-lyon.fr (Sophie Carneiro-Esteves)

such as sparsity or connectivity. For example, attention mechanisms [5] or reinforcement learning with a varying-patch sampling process [6] were proposed to specifically focus on the hierarchical structure and local contextual dependencies of curvilinear structures. Shit *et al.* proposed a loss function called ClDice [7] to penalize results based on their skeleton instead of their volume in order to promote connectivity. However, these methods rely on a significant amount of annotated data, which are difficult to obtain, especially in the field of curvilinear structures. Indeed, the annotation of curvilinear structures is a laborious and time-consuming task. This holds especially true in medical imaging, where the annotation task demands the expertise of a trained professional, making time a valuable resource.To cope with this, some algorithms were proposed to generate binary curvilinear structures[8, 9, 10, 11, 12], which are later used to train supervised segmentation models [10, 13].

Unrolling and "plug-and-play" approaches combine deep learning and variational methods to harness the strengths of both approaches. The unrolling approach consists of unfolding the iterative scheme of a variational approach in layers of a neural network [14, 15], and learning optimization parameters or regularization terms. On the other hand, plug-and-play methods keep the iterative scheme of a variational approach but replace the operator related to the regularization term by a regularization term learned separately by a neural network [16, 17].

By keeping an explicit model of the problem (*i.e.*, the data fidelity term), these hybrid methods yield more interpretable and robust results than pure deep learning frameworks while taking advantage of its power.

Both approaches rely on annotated datasets to train the model, although they differ in their methodology. The unrolling approach uses these data to learn the task of interest. Therefore it is necessary to have annotations for this dataset. The plug-and-play approach requires annotations only once to learn the regularization term, which can then be used to solve various tasks of interest, on another dataset, without the need for further annotations.

Given the lack of annotated data for curvilinear structures and the diversity of modalities, a plug-and-play framework is a more suitable choice.

Therefore, we propose to develop a hybrid plug-and-play approach dedicated to the segmentation of curvilinear structures. Contrary to pure deep learning approaches, we aim at designing an approach that is interpretable and robust and does not require annotations on the dataset of interest, which we will refer to as the target dataset. To this end, we propose learning a curvilinear structure reconnecting regularization term based on synthetic data, which is plugged into a variational framework to segment the target, non-annotated, dataset.

In a preliminary work [18], we showed that such a regularization yields promising results by successfully reconnecting fragments of curvilinear structures in 2D images.

In this article, we improve and go beyond this work by:

- introducing a strategy to learn a network that effectively reconnects fragmented curvilinear structures in both 2D and 3D images;

- incorporating this reconnecting network as a regularization term into a segmentation variational approach, enhancing the connectivity of the segmentation results;

- outperforming the current state-of-the-art methods in curvilinear segmentation when no annotation of the target dataset is available;

- validating the method on several datasets both in 2D and 3D from different domains, and providing a comprehensive evaluation of its performance.

The paper is organized as follows. In Section 2, we propose a brief review of curvilinear structure processing and plug-and-play methods. In Section 3, we describe our plug-and-play method and the proposed reconnecting regularization term. In Section 4, we present experiments and comparisons on 2D and 3D images. A conclusion is provided in Section 5.

2. Related Works

Curvilinear structure processing and analysis have been studied for many years, and a wide range of methodologies has been developed [19, 20]. In particular, with the emergence of deep learning and new computational possibilities, many deep learning-based approaches have recently been proposed, outperforming the previous state-of-the-art [21]. An extensive literature review of curvilinear structure processing is out of the scope of this article. In the following, we will focus on the two main aspects of our proposed method: the preservation of the topology of curvilinear structures and the plug-and-play approach in the context of curvilinear structure analysis.

2.1. Topology-preserving curvilinear structure segmentation

Some paradigms such as tracking or minimal path methods are intrinsically designed to preserve the topology of curvilinear structures. Tracking [22] is based on recursively recruiting neighbors pixels according to a criterion (*e.g.*, the pixel intensity). Similarly, minimal path methods [23] aim to find the best path, according to a cost function, between two points. In the context of curvilinear structures, this path is supposed to be the structure centerline. Both paradigms show interesting results; however, they need a human interaction to define a set of seed points, which can be tedious in the context of large and complex curvilinear structure networks.

Many deep learning approaches have been proposed recently to tackle curvilinear structure segmentation. In particular, U-Net [24, 25], the gold standard for medical image segmentation, is frequently used for vascular segmentation [26, 27, 28]. Among these methods, two main ideas have been explored to preserve the topology of curvilinear structures: either constraining the loss function of the network or modifying its architecture.

First, Betti numbers, representing the number of k-dimensional holes in a structure (*i.e.*, a segmentation in our case), have been used to propose topology-preserving losses. In 2D, Hakim *et al.* [29] defined a loss based on the Euler characteristic (a topological invariant defined as a linear combination of the Betti numbers) to reduce the number of isolated objects. Similarly, Hu *et al.*[30] proposed a loss based on persistent homology to extract the first Betti number corresponding to the number of connected components. Clough *et al.* [31] extended the approach to 3D images and all Betti numbers. Another work on 3D images [32] proposed to predict the persistent homology along with the segmentation to estimate the topology error and therefore improve the model. These losses evaluate whether the topology of the ground truth and the predicted segmentation match without taking into account if the underlying shapes overlap. To address this problem, Stucki *et al.* [33] proposed a topological loss based on the induced matching of persistence barcodes, ensuring penalization based on the topology but also the geometry of the segmentations.

On the other hand, losses based on the curvilinear structure centerlines, such as the ClDice [7], were proposed to foster connectivity and reduce the well-known volumetric bias existing between thick and thin curvilinear structures, leading to the vanishing of the latter. However, its use in 3D has not been successful because

of the complexity to extract accurate centerlines. To tackle this issue, Rougé *et al.* [34] proposed a cascaded U-Net to learn the skeletonization operation in addition to the segmentation and use the ClDice based on these more accurate centerlines. Mosinska *et al.* [35] proposed to learn a VGG on the ImageNet dataset, which is then employed to extract features from both the segmentation prediction and its ground truth. They introduced a loss function that involves minimizing the distance between these features.

Architecture can also be adapted to take into account curvilinear structure features. Attention modules [36] have been added in the latent space of a U-Net based-architecture to promote the connectivity [5, 37]. Multitask architectures have also been proposed, adding one or several proxy tasks, such as centerline detection, that better take into account the topology of curvilinear structures [38]. Lin *et al.* [39] proposed a cascaded architecture of two U-Nets, the first one was employed to extract texture features, while the second one focused on topology correction with contrastive learning. Some works [40, 41] focused on adding a transformer architecture [36] into the bottleneck of a U-Net to make the model consider the overall context of the image. They claimed that it improved the detection of long-distance relationships between pixels, ultimately leading to the better preservation of curvilinear structures connectivity.

Many segmentation pipelines include a first step based on classic vesselness filtering [42] to enhance the contrast between curvilinear structures and the background. A few works investigated how to use these methods in a deep learning framework, in particular by using the directional information provided by these filters, and thereby promoting connectivity [43, 44]. Most of these methods were only developed for 2D images, and they require a large amount of annotated data, which limits their use in real applications.

Segmentations can also undergo post-processing techniques to reconnect interrupted segments. Different algorithms have been proposed based on the centerlines [45], graph-based approaches [46, 47], and contour completion processes [48]. These methods are complex to use, parameter-dependent, and were only proposed for 2D images.

2.2. Variational approaches for curvilinear structure segmentation

The variational approach for image restoration is defined as the minimization of two energies:

$$\hat{u} = \underset{u \in [0,1]^N}{\operatorname{argmin}} E_{\text{data}}(u, f) + \lambda E_{\text{reg}}(u), \tag{1}$$

where $f \in \mathbb{R}^N$ is the initial image of N pixels, $\hat{u} \in \mathbb{R}^N$ is the restored image, E_{data} is the data fidelity, E_{reg} the regularization term which promotes desirable solution properties and $\lambda \in \mathbb{R}$ is a regularization coefficient acting as a trade-off between both terms.

A classic regularization term used in image processing is the Total Variation (TV), which promotes a piecewise smooth solution [49]. Even though the term is efficient for denoising natural images, it tends to make thin curvilinear structures disappear, leading to a connectivity loss in many applications. To address this issue, a directional regularization adapted to curvilinear structures has been proposed [4] by incorporating prior knowledge on the curvilinear structure orientation. Even though this term promotes connectivity by denoising in an anisotropic manner, it does not lead to the reconnection of significantly separated curvilinear fragments. Moreover, this term requires the computation of a curvilinear structure estimator, such as the RORPO filter [50] or the Frangi vesselness [51], which can be time-consuming and/or complex to parametrize [42].

Outside the curvilinear structure scope, Heide *et al.* [52] proposed to replace or combine a classic regularization term, such as TV, with a denoising algorithm like BM3D [53]. BM3D was used as a self-similarity inducing denoising prior on the restored image. Following this work, a series of works [16, 54, 55, 56] has been proposed to replace the regularization term by a denoising neural network, which learns the noise distribution of a dataset of interest. Therefore, these learned regularization terms are more adapted to the application than classic regularization ones. These approaches were used to tackle image restoration applications such as demosaicking, deconvolution, inpainting, or deblurring. To the best of our knowledge, all these methods proposed to learn a prior only on the noise of the observed image. But no work has been proposed to learn more complex regularizations such as the connectivity of curvilinear structures.

This article presents the first regularization term aiming to add a more complex constraint on the solution of interest and a complete framework to perform segmentation.

3. Proposed method

In this section, we will first present a global overview of our proposed framework, then we will describe separately each step.

3.1. Overview

In this work, we aim at learning a regularization term promoting the connectivity of curvilinear structures, which is a property difficult to define explicitly. Subsequently, this term is incorporated into a segmentation plug-and-play approach.

To this end, we propose a three-step framework composed of:

- The generation of a disconnected binary curvilinear structure dataset.

  To learn a connectivity property, we propose to train a network using a synthetic dataset of paired images. This dataset comprises an input image containing disconnected curvilinear structures and its corresponding version with the structures connected (see Fig. 1, top left). To generate these images, we developed an algorithm simulating realistic disconnections in curvilinear structures.

- The reconnecting regularization term learning.

  From the previously generated dataset, we learn a network G_{reco} (Fig. 1 top right) with a residual U-Net aiming at reconnecting curvilinear structures in an image (see Section 3.3).

- The plug-and-play segmentation.

  The learned network G_{reco} is finally plugged into a segmentation variational framework, by replacing the proximity operator of the regularization term E_{reco} in the segmentation iterative scheme (Fig. 1 bottom).

The code of our framework is available at https://github.com/creatis-myriad/plug-and-play-reco-regularization.

3.2. Dataset creation

Our objective is to develop a regularization term that promotes the connectivity of curvilinear structures. To be generic and usable in many different applications, this term should not depend on the input image modality or the type of curvilinear structure (*e.g.* road, vessels, *etc.*) but rather focus on the geometric properties of these structures.

Such a reconnecting regularization would be too complex to express explicitly, so we have chosen to learn it based on a generated dataset of connected/disconnected curvilinear structure images. We have opted for a synthetic binary dataset to represent the connectivity property, as binary images not only exhibit

modality independence but also ease the generation of realistic disconnections.

From an initial image containing binary curvilinear structures, our algorithm creates random disconnections inside curvilinear structures according to several rules we have established based on our experience and observations of current automatic segmentation algorithm results.

More specifically, the following steps are performed to create an image containing disconnections:

1. Extraction of the curvilinear structure centerlines and its distance map. The distance map is defined for each pixel inside the curvilinear structures, and its value is the distance to the background.

2. Classification of each centerline pixel into $m \in \mathbb{N}$ classes based on their radius.

3. Selection of the p ($p \leqslant m$) thinnest curvilinear structure classes to avoid disconnections on the biggest structures.

4. Generation of disconnections:

   - Selection of a value $i \in [0, p]$ drawn from the probability distribution $P(i) = \frac{2^{p-i}}{2^p - 1}$. Here, i represents the radius class of the curvilinear structure which will be disconnected.

   - Random selection of a centerline pixel in the image corresponding to class i. This point will be the center of the disconnection.

   - Selection of the disconnection size s from a normal distribution with a mean of $\frac{C}{i+1}$ with $C \in \mathbb{N}$ a constant depending on the dataset.

   - Generation of the disconnection
     The disconnection is generated by removing random pixels in a disc (or a ball) centered on the centerline pixel.

Afterwards, additional fragments are added to the image to mimic non-curvilinear structure fragments. These fragments are introduced to make the network learn to differentiate fragments aligned with curvilinear structures, which should probably be reconnected, from noise and non-curvilinear structures, that should not be reconnected. In 2D, the fragments are generated by adding random pixels inside a disc area. In 3D, these fragments are generated using the algorithm proposed by Douarre *et al.* [57] to generate more complex patterns.

3.3. Reconnecting regularization term learning

We use the previously generated dataset to learn the reconnecting operator G_{reco}. Our model is designed to effectively reconnect fragments of curvilinear structures while removing the non-curvilinear fragments.
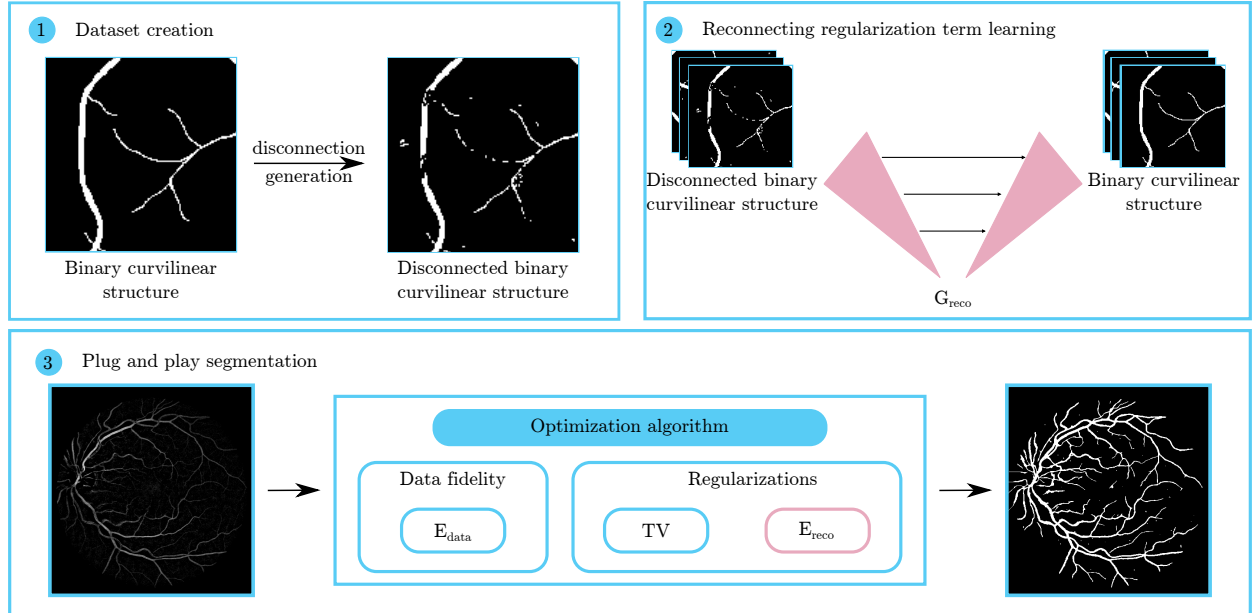

Figure 1: Pipeline of our proposed framework. First, a dataset is generated with pairs of synthetic connected/disconnected curvilinear structures. Second, this dataset is used to train a model G_{reco} with a residual U-Net architecture. Third, this model is finally plugged into the optimization scheme to solve the segmentation problem by replacing the proximity operator of E_{reco}. Note that for the plug-and-play segmentation, it is essential that the input has a homogeneous background. A preprocessing step, such as subtracting the median filter of the image, may be necessary.

We trained a residual nD UNet ($n = 2$ or $n = 3$ depending on the image dimension) [58]. U-Net is the gold standard for medical image segmentation, and adding residual units to its architecture improves its robustness to inputs that are dissimilar to training data. This property is essential because we aim at learning a generic regularization term which will be used on different target datasets.

Our U-Net architecture is 4-layer deep using batch normalizations and 16, 32, 64 and 128 features at each layer.

In the literature, the Dice loss [59] has been widely used for learning segmentation in unbalanced classes scenarios. This is the case for our problem, where the curvilinear structures only represent a small portion of the images pixels, and the missing fragments only a small fraction of the curvilinear structures themselves.

However, only using a Dice loss would focus the network attention on preserving the curvilinear structures in general, without focusing on the accurate identification of the missing fragments.

To address this issue, we propose to use a combination of two Dice losses (see Eq. 2). The first one is a classic Dice $\mathcal{D}$ on the image to guide the reconstruction of the whole curvilinear structures. The second Dice loss, called a weighted Dice $\mathcal{D}_w$, is computed only within a mask M of the missing fragments and its close neighborhood. The total loss $\mathcal{L}$ is defined as follows:

$$\mathcal{L}(x, y) = \mathcal{D}(x, y) + \mathcal{D}_w(x, y; M), \tag{2}$$

with x and y the image and its associated annotation.

The missing fragment mask is generated by a dilation of the missing fragments with a disc (or ball in 3D) structuring element of radius $r = 2$.

We train our models with patches of size 96^n voxels and performed an on-the-fly data augmentation with random rotations and flips. The learning rate is set to 10^{-3} and the networks is trained for 1000 epochs with an Adam optimizer. The model with the lowest validation loss is kept.

3.4. Plug-and-play segmentation

We express our segmentation problem as a variational approach (see Eq.1). In this section, we first present the choice of the data fidelity energy E_{data} and the regularization term E_{reg}, then we describe the iterative optimization scheme used to solve this segmentation problem.

3.4.1. Data fidelity and regularization terms

We have chosen the Chan *et al.* segmentation data fidelity term [60] defined as follows:

$$E_{\text{data}}(u, f) = \left\langle u, c_f \right\rangle_F,$$
$$c_f \mapsto (c_1 - f)^2 - (c_2 - f)^2, \tag{3}$$

with c_1 and c_2 constant values corresponding respectively to the mean value of the background and the foreground, and $\langle ., . \rangle_F$ is the Frobenius product.

In the previous sections, we have presented our reconnecting regularization term, E_{reco} which promotes a connected segmentation. In addition to this term, we need a classic regularization/denoising term promoting a smooth segmentation. To favor such a property, we add the classic total variation, $TV(u)$, defined by:

$$TV(u) = \|\nabla u\|_{2,1}, \tag{4}$$

with $\|.\|_{2,1}$ the l_1 norm of the l_2 norm and ∇ the gradient operator.

As our goal is to obtain a binary segmentation result, we need a last regularization term to constrain the values of the solution between 0 and 1. We use $\iota_{[0,1]^N}$, the indicator function of the set $[0, 1]^N$.

Our reconnecting regularization is defined based on binary images and cannot be applied directly on a grey-level image. During the iterative optimization scheme, the grey-level input image will progressively be transformed into a binary segmentation. When the image is close to a binary image, but the iterative scheme has not yet converged, we propose to introduce our learned reconnecting regularization E_{reco}. Our final regularization term E_{reg} is defined as follows:

$$E_{reg}(u) = \begin{cases} \lambda TV(u) + E_{\text{reco}}(u) & \text{if } u \text{ almost binary,} \\ \lambda TV(u) + \iota_{[0,1]^N}(u) & \text{else.} \end{cases} \tag{5}$$

The condition *almost binary* depends on the number of iterations of the solving algorithm and is set experimentally (see Section 4).

3.4.2. Optimization scheme

Our segmentation problem is a three-term energy : E_{data}, TV and $\iota_{[0,1]^N}$ or E_{data}, TV and E_{reco}. We propose to use a *Forward-Backward Primal-Dual* (FBPD) algorithm [61] which solves the following primal problem:

$$\hat{u} = \operatorname*{argmin}_{u} h(u, f) + g(Lu) + k(u), \tag{6}$$

where k, g and h are lower semi-continuous convex functions from $\mathbb{R}^N$ to $(-\infty, +\infty]$. h is differentiable with a β-Lipschitz continuous gradient ($\beta > 0$), and $L \in \mathbb{R}^{N^2}$ an operator and $\hat{u}$ the primal solution.

The FBPD algorithm solves the primal problem Eq.(6) by relying on its associated dual problem :

$$\hat{v} = \underset{v}{\operatorname{argmin}}(k^\star \square h^\star)(-L^T v) + g^\star(v), \qquad (7)$$

where $g^\star$ the conjugate of g, $\square$ the inf-convolution defined as $(f \square g)(x) = \inf_{y \in \mathbb{R}^N} f(y) + g(x - y)$, and $\hat{v}$ the dual solution.

The FBPD iterative scheme is defined in Eq. 8.

$$\begin{aligned} u_{i+1} &= \operatorname{prox}_{\tau k}(u_i - \tau(\nabla h(u_i) + L^T v_i)), \\ v_{i+1} &= \operatorname{prox}_{\sigma g^\star}(v_i + \sigma L(2u_{i+1} - u_i)), \end{aligned} \qquad (8)$$

with $\operatorname{prox}_{\sigma g}$ the *proximal* operator of σg, $g^\star$ the conjugate of g, $\tau \in \mathbb{R}^+$ and $\sigma \in \mathbb{R}^+$ scalar hyperparameters. τ and σ must be set according to Eq. 9 to ensure the convergence of the algorithm.

$$\tau^{-1} - \sigma\|L\|_S^2 \geq \frac{\beta}{2}, \qquad (9)$$

with $\|.\|_S$ the spectral norm.

To use the FBPD algorithm on our segmentation problem, we define the following terms:

$$\begin{aligned} h(u, f) &= E_{\text{data}}(u) = \left\langle u, c_f \right\rangle_F, \\ g(u) &= \lambda\|u\|_{2,1}, \\ L &= \nabla, \\ k(u) &= \begin{cases} E_{\text{reco}}(u) & \text{if } u \text{ almost binary,} \\ \iota_{[0,1]^N}(u) & \text{else.} \end{cases} \end{aligned} \qquad (10)$$

To apply the FBPD algorithm, closed-form solutions of the proximity operators of Eq. 8 are required :

$$\operatorname{prox}_{\sigma g^\star}(u) = \frac{\lambda \sigma^{-1}}{\max\left(\left\|\frac{u}{\sigma}\right\|_2, \lambda \sigma^{-1}\right)},$$

$$\operatorname{prox}_{\sigma \iota_{[0,1]^N}}(u) = \operatorname{proj}(u) = \begin{cases} u & \text{if } u \in [0,1] \\ 0 & \text{if } u < 0 \\ 1 & \text{otherwise} \end{cases}. \qquad (11)$$

When u is almost binary, following the classic plug-and-play paradigm, we replace the proximity operator $\operatorname{prox}_{\tau E_{\text{reco}}}$ of the FBPD iterative scheme by our reconnecting network G_{reco}.

Finally, our plug-and-play segmentation algorithm is presented in Algorithm 1.

Algorithm 1: Plug-and-play segmentation with the learned reconnecting operator

Data: $\alpha \in \mathbb{N}^{+*}$,
$\quad u_0 \in \mathbb{R}^{\mathbb{N}^2}, v_0 \in \mathbb{R}^{2\mathbb{N}^2}, (\tau, \sigma) \in (0, +\infty)^2$

for $i \geq 1$ **do**
$\quad p_i = (u_i - \tau(\nabla h(u_i) + L^T v_i)$
$\quad$ **if** $i < \alpha$ **then**
$\quad\quad u_{i+1} = \operatorname{prox}_{\tau \iota_{[0,1]^N}}(p_i)$
$\quad$ **else**
$\quad\quad u_{i+1} = G_{\text{reco}}(\operatorname{proj}(p_i))$
$\quad v_{i+1} = \operatorname{prox}_{\sigma g^\star}(v_i + \sigma L(2p_i - u_i))$

---

4. Experiments

In this section, after describing our experimental setup, we first demonstrate the interest of our reconnecting operator on 2D and 3D images of blood vessels. We compare its results with both classic and deep learning-based approaches. Then, we analyze it through an ablation study. Finally, we illustrate the interest of our approach on other applications.

4.1. Experimental setup
4.1.1. Datasets

We chose two applications to conduct our experiments: a 2D application on retinal images, and a 3D application on injected CT-scans of the liver. One of the main interest of our approach is that it does not require annotations of the target dataset to perform the segmentation, but only synthetic images for training G_{reco}. To illustrate this point, we needed for each application one synthetic dataset to learn our reconnecting term, and another to perform the segmentation.

In 2D, we chose to learn our reconnecting term on 20 binary curvilinear structures generated using the OpenCCO algorithm [9] (see Fig. 4(a)). Then, we generated our pairs of connected/disconnected curvilinear structure images using our algorithm to generate random disconnections (see Section 3.2) and applied random rotations, to obtain a dataset of 80 pairs of connected/disconnected curvilinear structure images. For the segmentation, we used the DRIVE dataset [62], which consists of 40 retinal images and their associated annotations.

In 3D, we trained our reconnecting term on a synthetic dataset of curvilinear structures generated with the Vascusynth software [1] [8]. We generated 315 images

---

[1]https://vascusynth.cs.sfu.ca/Welcome.html

(a) Initial image with background homogenisation

(b) Ground-truth

(c) Mean Teacher Synthetic

(d) Mean Teacher STARE

(e) Frangi

(f) RORPO

(g) Total Variation

(h) Directional Total Variation

(i) E_{remove}
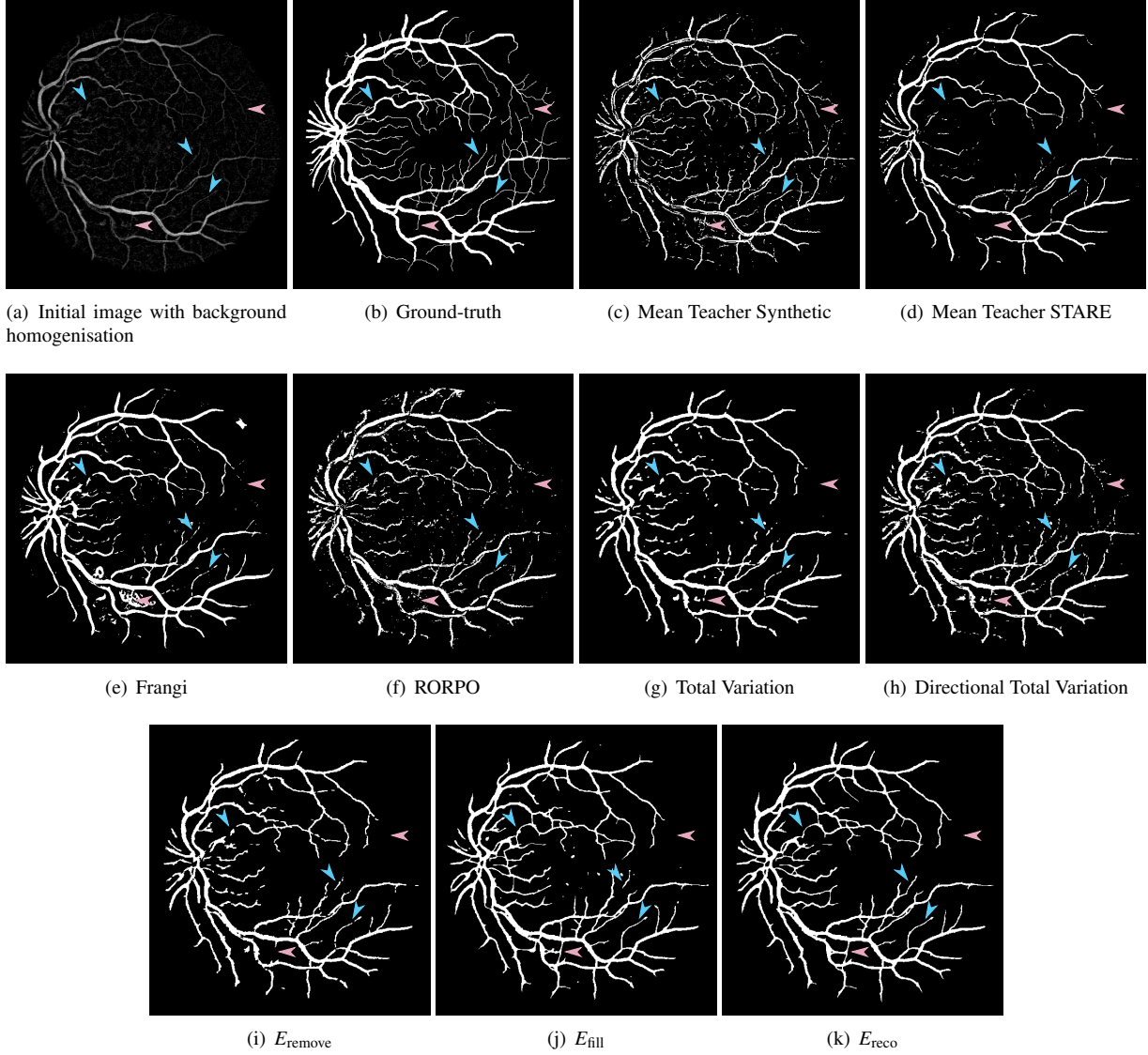
(j) E_{fill}

(k) E_{reco}

Figure 2: Comparison of segmentation results on one DRIVE image. Blue arrows point to some examples of successful reconnections of our reconnecting regularization term. Pink arrows show false reconnections or missing vessels.

that we disconnected with our proposed algorithm (see Section 3.2). As the target dataset for the segmentation, we used the IRCAD dataset[2], which is composed of 19 3D liver CT-scans and their associated annotations.

Note that the DRIVE and IRCAD annotations were only used for the evaluation of the method, and not for the training.

4.1.2. Compared methods

Firstly, we conducted a comparison between our learned reconnecting term E_{reco}, the classic TV regularization [49], and the directional TV [4] which is specifically designed to reconnect curvilinear structures. For a fair comparison, the segmentation with the classic TV (resp. the directional TV) is performed with a FBPD algorithm, where $g(Lu)$ is the TV (resp. the directional TV) and $k(u)$ is $\iota_{[0,1]^N}(u)$.

We set the same parameters c_1, c_2, σ and τ for the three methods. The background and foreground constants c_1 and c_2 were set experimentally. The parame-

(a) Ground-truth   (b) Frangi   (c) RORPO

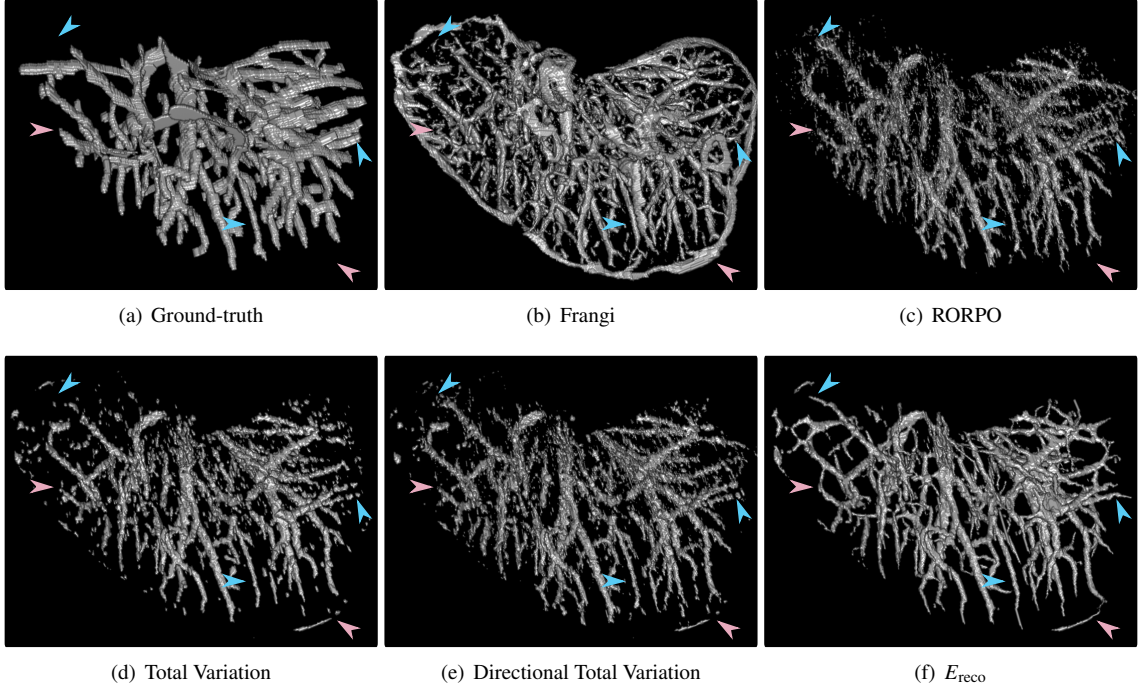(d) Total Variation   (e) Directional Total Variation   (f) E_{reco}

Figure 3: Comparison of segmentation results on one IRCAD volume. Blue arrows point to some examples of successful reconnections of our reconnecting regularization term. Pink arrows show false reconnections.

ters τ and σ were set respectively to 1.587 and 10^{-3} in 2D and to 1.493 and 10^{-3} in 3D based on Eq. 9.

To ensure a fair comparison, the values of λ were independently optimized for each image and each method. We selected the value maximizing the MCC in the range $[0.001, 0.080]$ in 2D, and $[0.001, 0.050]$ in 3D, with a step of 10^{-3}.

The FBPD scheme was applied until convergence or for a maximum of 1000 iterations. The switch between $\iota_{[0,1]^N}$ and our reconnecting regularization E_{reco} in our approach is done at $\alpha = 500$ iterations when we observed that u is usually almost binary. We observed that our approach is not very sensitive to this value, and it could be set approximately to half the maximum number of iteration performed.

The Chan $et\ al.$ data fidelity term assumes that the input image has a homogeneous background. We thus preprocessed each image by subtracting its median filter. The kernel of the median filter was set once for each dataset experimentally.

Furthermore, as we developed an unsupervised approach, we compared our framework with classic unsupervised method for blood vessels filtering and segmentation : RORPO [50] and Frangi [51]. Parameter optimization was carried out using a grid search, exploring a

predefined set of values for each hyperparameter to determine the optimal configuration of each method based on the MCC (see Sec 4.1.3).

Most recently proposed deep-learning-based vascular segmentation methods are primarily supervised, showcasing high performance when annotations are available. However, their performance drastically drop when annotations are lacking. Such supervised approaches are out of the scope of this article, however, for reference purposes, we provided average metrics reported in the literature, offering insights into the upper limits of performance for comparative analysis. In 2D, an average Dice of 0.836 was reported in [63], while in 3D, an average Dice of 0.838 was achieved according to [64]. In this article, we compared our results with a deep-learning based approach in similar conditions than our method, $i.e.$, in an unsupervised context, on an unlabeled target dataset. In this context, classic supervised deep learning-based segmentation strategies cannot be applied, and two scenarios may occur:

1. scenario 1: one has access to an annotated dataset similar to the target dataset.
2. scenario 2: one can generate a simulated dataset with annotation, similar to the target dataset.

We compared our results with both scenarios using

9

	Method	Volumetric Metrics				Geometric Metrics			Topological Metrics		
		TPR	PPV	MCC	Dice	ClDice	95HD	ASSD	ϵ_{β_0}	ϵ_{β_1}	ϵ_χ
2D	RORPO	0.650	0.713	0.636	0.678	0.626	11.789	2.293	52.275	0.924	3.872
		± 0.053	± 0.064	± 0.050	± 0.044	± 0.048	± 3.357	± 0.608	± 34.062	± 0.059	± 1.060
	Frangi	0.716	0.756	0.699	0.735	0.717	13.576	2.410	21.340	0.646	1.849
		± 0.051	± 0.042	± 0.044	± 0.041	± 0.047	± 4.441	± 0.846	± 15.959	± 0.126	± 0.794
	MT synth.	0.602	0.864	0.689	0.707	0.713	**7.268**	**1.509**	22.809	0.614	1.906
		± 0.070	± 0.039	± 0.047	± 0.051	± 0.053	± 3.230	± 0.511	± 16.696	± 0.204	± 0.531
	MT STARE	0.607	**0.903**	0.711	0.724	0.717	15.932	2.421	11.235	0.777	1.459
		± 0.061	± 0.033	± 0.041	± 0.044	± 0.051	± 3.728	± 0.580	± 7.725	± 0.100	± 0.323
	TV	0.690	0.821	0.719	0.747	0.730	13.489	2.265	24.220	0.749	2.135
		± 0.051	± 0.069	± 0.041	± 0.036	± 0.044	± 5.186	± 0.885	± 15.885	± 0.113	± 0.545
	Directional TV	0.694	0.819	0.720	0.748	0.728	12.143	2.071	25.833	0.741	2.219
		± 0.053	± 0.076	± 0.045	± 0.041	± 0.049	± 5.014	± 0.769	± 22.350	± 0.103	± 1.129
	E_{reco}	**0.725**	0.803	**0.729**	**0.759**	**0.744**	16.095	2.550	**2.685**	**0.316**	**0.417**
		± 0.049	± 0.069	± 0.041	± 0.036	± 0.045	± 5.235	± 0.924	± 2.767	± 0.187	± 0.232
3D	RORPO	0.408	0.640	0.487	0.488	0.525	13.145	2.917	1.029	57.818	66.991
		± 0.121	± 0.087	± 0.090	± 0.096	± 0.104	± 4.600	± 0.873	± 1.444	± 52.836	± 106.262
	Frangi	**0.455**	0.492	0.441	0.462	0.491	21.939	4.732	5.079	6.817	13.284
		± 0.081	± 0.127	± 0.087	± 0.079	± 0.109	± 7.465	± 1.491	± 4.980	± 16.565	± 27.567
	TV	0.344	**0.721**	0.473	0.450	0.533	**11.623**	2.954	2.253	**4.533**	4.789
		± 0.139	± 0.126	± 0.116	± 0.129	± 0.166	± 4.213	± 1.017	± 3.300	± 10.674	± 5.809
	Directional TV	0.362	0.695	0.477	0.462	0.562	11.734	2.666	1.682	41.966	32.495
		± 0.126	± 0.099	± 0.092	± 0.105	± 0.106	± 4.581	± 0.819	± 2.266	± 51.223	± 42.293
	E_{reco}	0.435	0.675	**0.513**	**0.507**	**0.585**	13.032	2.895	**0.746**	7.505	**7.265**
		± 0.153	± 0.105	± 0.087	± 0.102	± 0.079	± 4.169	± 0.877	± 0.432	± 10.088	± 8.939

Table 1: Quantitative segmentation results on the DRIVE database (2D) and on the IRCAD database (3D).

a mean teacher (MT) network [65], which is a popular approach used to learn a model with both annotated and non annotated data from related datasets.

In these experiments, we have chosen the STARE dataset [66] consisting of 20 retinophotographies, which is similar to the target DRIVE dataset. In scenario 1, called MT STARE, we trained the MT using both the STARE dataset (with annotations) and the DRIVE dataset (without annotation). In scenario 2, called MT synth., we extracted the image backgrounds from the STARE dataset [66] and added synthetic curvilinear structures generated by the CCO algorithm, similarly to what is proposed in Lin *et al.* [13]. An example of such generated images is shown on Fig 4. Then we trained the MT using both the generated image (with annotations) and the DRIVE dataset (without annotation).

4.1.3. Evaluation metrics

To evaluate the methods quantitatively, we computed metrics derived from the confusion matrix (True Positive TP, True Negative TN, False Positive FP, and False
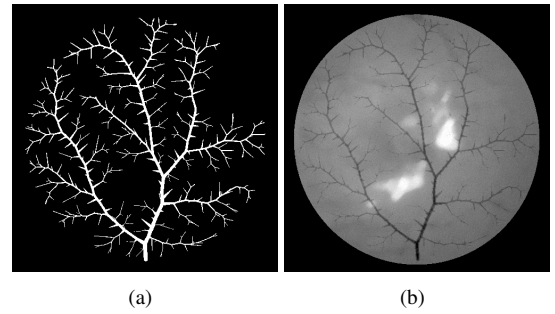

(a) (b)

Figure 4: Simulation of a retinophotography (b) using a binary curvilinear structure (a) generated by the OpenCCO algorithm and an image background extracted from the STARE dataset [66].

	Regularization	λ
2D	TV	$0.0096 \pm 0,0072$
	Directional TV	0.0176 ± 0.0145
	E_{remove}	0.0072 ± 0.0069
	E_{fill}	0.0114 ± 0.0078
	E_{reco} (with Diceloss)	0.0085 ± 0.0080
	E_{reco}	0.0089 ± 0.0075
3D	TV	0.0095 ± 0.0112
	Directional TV	0.0119 ± 0.0147
	E_{reco}	0.0084 ± 0.0098

Table 2: Mean ± standard deviation value of the regularization coefficient, λ, that were optimized for each approach.

Negative FN). As vessel segmentation is an unbalanced problem, we computed: the sensitivity TPR, the precision PPV, the Dice and the Matthew Correlation Coefficient MCC (see Eqs. 12). We will refer to these metrics as volumetric metrics. They were computed inside the field of view of the retinal images or inside the liver mask of the CT-scans.

$$
\begin{aligned}
\text{TPR} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\
\text{PPV} &= \frac{\text{TP}}{\text{TP} + \text{FP}}, \\
\text{Dice} &= \frac{2\text{TP}}{2\text{TP} + \text{FN} + \text{FP}}, \\
\text{MCC} &= \frac{\text{TP.TN} - \text{FP.FN}}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}}.
\end{aligned}
\tag{12}
$$

None of these metrics give insight on the geometry or the connectivity of the curvilinear structures. Indeed, the difference between a connected and a disconnected result is usually only a few pixels, which does not change significantly volumetric metrics, but drastically reduce its interest for follow-up applications. To achieve an evaluation centered on the geometry of the curvilinear structures, we computed the ClDice [7], the 95th percentile of the Hausdorff Distance 95HD, and the Average Symmetric Surface Distance ASSD. We will refer to these metrics as geometric metrics. To evaluate the connectivity of our results, we introduced classic topological metrics called Betti's numbers: β_0, β_1 and β_2. β_0 corresponds to the number of connected components, β_1 represents the number of tunnels, and β_2 denotes the number of cavities in a binary shape. We also added the Euler number, χ, defined as a linear combina-

tion of the Betti's numbers:

$$
\chi = \beta_0 - \beta_1 + \beta_2.
\tag{13}
$$

To evaluate the variation of these metrics in relation to their value in the annotated images, we computed the error ratio of β_0, β_1 and χ (see Eq. 14). We did not calculate the error ratio of β_2, as its value is 0 in the ground truth and in practice we did not observe cavities in the 3D segmentation results.

$$
\epsilon_{\text{M}} = \left| \frac{M - M_{\text{gt}}}{M_{\text{gt}}} \right|,
\tag{14}
$$

with M a topological metric (either β_0, β_1, or χ) computed on a segmentation and M_{gt} the same metric computed on its associated annotation. As these topological metrics are very sensitive to noise, we performed a post-processing before computing them. It consists in removing the small connected components (size less than 20 pixels in 2D and less than 90 pixels in 3D) and filling the small holes (size less than 10 pixels in 2D). This post-processing was neither applied to the qualitative results presented in this article nor before computing the volumetric and geometric metrics.

4.2. Results

In this section, we compare the results of our framework with the other strategies presented in section 4.1.2. Quantitative results are presented in Table 1, while some qualitative results are presented in Fig. 2 and in Fig. 3.

Qualitatively, the directional TV preserves more curvilinear fragments than TV, however these fragments are not connected, as shown by the similar values of ϵ_{β_0}. Our segmentation results are smoother and demonstrate better connectivity compared to the TV and directional TV approaches. Quantitatively, our approach only slightly increases the volumetric metrics. However, this small increase does not adequately represent the importance of the improvement of the results. Indeed, it comes from the addition of a few pixels to reconnect the vessels, which is crucial for ensuring better global coherence and connectivity of the vascular networks both in 2D and 3D. This improvement is visually clear, and is demonstrated by the significant decrease in topological metrics. In particular, the error on the number of connected components of our results has decreased by almost 90% in 2D and 70% in 3D. Despite the capability of our reconnecting term to remove non-curvilinear fragments, we observe a few false reconnections in the results (see pink arrows in Fig 2 and in Fig 3). This may happen when artefacts are aligned or in the same orientation as a curvilinear structure.

Method	Volumetric Metrics				Geometric Metrics			Topological Metrics		
	TPR	PPV	MCC	Dice	ClDice	95HD	ASSD	ϵ_{β_0}	ϵ_{β_1}	ϵ_χ
Before G_{reco}	0.967	0.982	0.971	0.974	0.959	0.350	0.221	107.367	0.113	4.410
	± 0.006	± 0.005	± 0.004	± 0.004	± 0.006	± 0.477	± 0.072	± 71.883	± 0.103	± 1.746
After G_{reco}	0.989	0.977	0.980	0.983	0.980	0.375	0.082	17.301	0.151	0.543
	± 0.003	± 0.004	± 0.003	± 0.003	± 0.005	± 0.484	± 0.014	± 12.690	± 0.135	± 0.234
E_{remove}	0.698	**0.825**	0.726	0.753	0.737	16.044	**2.524**	20.758	0.722	1.937
	± 0.054	± 0.068	± 0.039	± 0.035	± 0.045	± 5.431	± 0.915	± 12.537	± 0.122	± 0.438
E_{fill}	0.720	0.793	0.720	0.751	0.732	**15.428**	2.539	4.208	0.309	0.474
	± 0.050	± 0.077	± 0.047	± 0.041	± 0.048	± 5.605	± 1.041	± 4.476	± 0.173	± 0.271
E_{reco} (w/o $\mathcal{D}_w$)	0.716	0.810	0.728	0.757	0.745	16.351	2.583	2.979	0.371	0.531
	± 0.051	± 0.071	± 0.044	± 0.038	± 0.048	± 5.852	± 1.019	± 2.523	± 0.195	± 0.239
E_{reco}	**0.725**	0.803	**0.729**	**0.759**	**0.744**	16.095	2.550	**2.685**	**0.316**	**0.417**
	± 0.049	± 0.069	± 0.041	± 0.036	± 0.045	± 5.235	± 0.924	± 2.767	± 0.187	± 0.232

Table 3: Ablation study on the DRIVE database (2D). The first two rows of the table present quantitative results for the evaluation of G_{reco} alone applied to the 40 binary annotations of the DRIVE dataset, that were initially disconnected by our algorithm. The subsequent rows present the results of the ablation study of the segmentation framework with different versions of the regularization term, applied to the 40 images of the DRIVE dataset.

Moreover, our results present a slightly higher value for the distance-based metrics (95HD, ASSD) which may be due to the removal of small fragments that belong to vessels, but were detected as artefacts by our approach. However, the global vascular geometry is still better preserved as shown by the increase in the ClDice value.

The classic vesselness filters (RORPO and Frangi) yield segmentation volumes similar to the ones from the variational approaches, as shown by the close volumetric metrics. However, they exhibit much more artefacts, which increases the topological errors. Unsupervised deep learning methods achieved either an incomplete (MT STARE) or noisy (MT synth.) segmentation. However, the global architecture of the vascular network is relatively well preserved, as shown by the topological metrics. It is important to note that deep learning approaches show more promising results in a semi-supervised context where a few annotations are still available [67], which is out of the scope of this article.

4.3. Ablation study

To better understand the role of our reconnecting term G_{reco}, we conducted several experiments. First, we validated its behavior on disconnected binary curvilinear structures, independently of the segmentation variational scheme. Then, we analysed separately the design choices of this term. Finally, we analysed the influence of injecting earlier our term into the variational

scheme. The qualitative results of the ablation study are presented in Fig 2(i-k) and in Fig 6. The quantitative evaluation is presented in Table 3.

4.3.1. Validation of G_{reco}

To validate our reconnecting term, we applied it on the DRIVE manual annotations that we had disconnected with our algorithm. In Fig. 6, most disconnections were reconnected and artefacts removed, as confirmed by the quantitative results presented in the first two lines of Table 3. All metrics showed significant improvement after applying our reconnecting term, with particular emphasis on ϵ_{β_0}, highlighting a good reconnecting behavior as expected.

4.3.2. E_{fill} v.s. E_{remove} v.s. E_{reco}

Our reconnecting term performs two tasks. Firstly, it fills small disconnections between curvilinear fragments that are close and in the same orientation. We refer to this task as *fill*. Secondly, it removes small components that are just noise and are not supposed to be reconnected to the main structure. We refer to this task as *remove*. In this section, we study these two tasks separately, and together to show the interest of their synergy.

We trained three different regularization terms: G_{fill} which only fills disconnections between curvilinear fragments, G_{remove} which only removes small fragments that are not aligned with curvilinear structures, and G_{reco} which performs both tasks. These three models were trained with the same architecture presented in Section 3.3. The only difference is the ground truth that was

12

(a) Initial image   (b) Ground-truth   (c) Total Variation   (d) Directional Total Variation   (e) E_{reco}
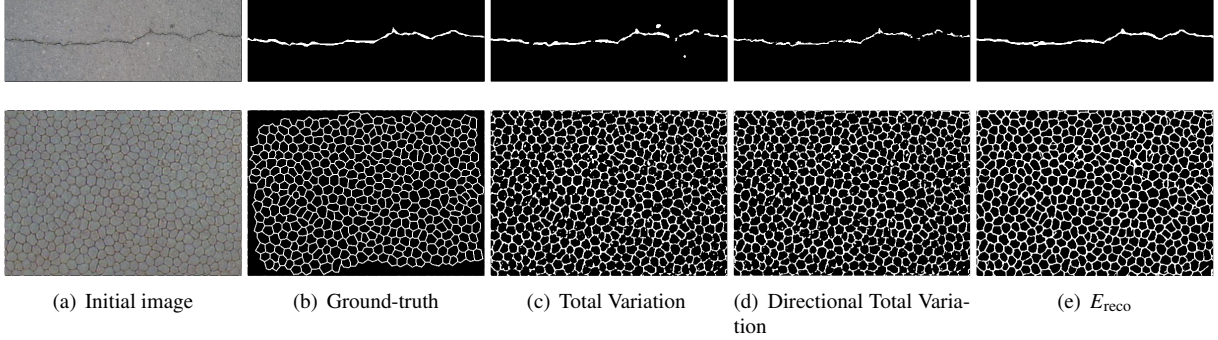
Figure 5: Segmentation of various curvilinear structures. The first row depicts an image from the road crack dataset [68, 69], and the second row shows an image from the porcine corneal cells dataset [70].
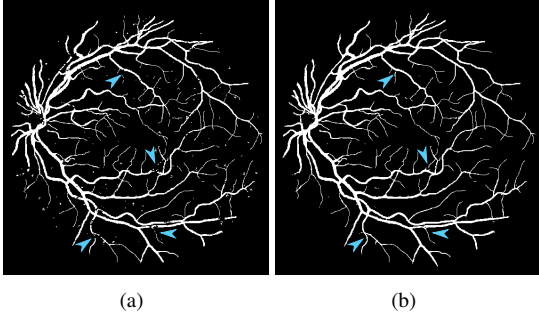

(a)   (b)

Figure 6: Test of our reconnecting term on the manual annotations of the Drive dataset.(a) the disconnected input image, (b) the result after applying G_{reco} on (a).

used to train the models, which we generated according to the task of interest.

Then, each term is added in the variational segmentation scheme, replacing the proximity operator of the regularization term (E_{fill}, E_{remove}, and E_{reco}). The results are presented in Table 3.

The results of E_{remove} are very similar, both qualitatively and quantitatively, to the results of the segmentation with TV only. Removing non-curvilinear structure fragments is a type of denoising, a task already performed by TV, so it does not significantly improve the results on its own.

The results of E_{fill} are similar based on the MCC to the results of TV and directional TV. However, a significant improvement is observed when looking at the topological metrics. In particular, the error on the number of connected components (ϵ_{β_0}) is almost divided by 5 compared with E_{remove} or TV. This shows that the reconnection worked as intended.

The mean and standard deviation of the topological errors of E_{reco} is improved compared to E_{fill} alone, and

in particular ϵ_{β_0} is significantly improved. Moreover, we observed that, by including a denoising targeting non-curvilinear structure fragments, the strength of the non-specific denoising TV can be reduced (see Table 2), allowing for more fragments to potentially be reconnected later by the reconnecting task.

4.3.3. Weighted Dice loss

As described in Eq. 2, we use a combination of two Dice losses to train our neural network. To prove the interest of this loss, we have trained 2D reconnecting terms both with and without the weighted Dice loss and plugged them into our segmentation framework. Quantitatively, we can observe that including the combination of Dice losses slightly improved the segmentation results, in particular regarding the topological metrics. While the difference is subtle in 2D, the 3D results show a much greater improvement with the addition of this extra loss.

4.3.4. Influence of α

As described in the section 3.4.2, our reconnecting term must be injected when its input is almost binary. In the previous experiments, we set $\alpha = 500$.

In this section, we investigate the influence of the parameter α on the results of our approach. We systematically varied α across the range of $[0, 500]$ with increments of 50. The results of this experiment are presented in Fig. 4.3.4.

When the reconnecting model is introduced at the beginning of the optimization scheme ($\alpha = 0$), the reconnecting term is applied to an image that deviates significantly from the distribution it has been trained on, *i.e.*, far from a binary image. Consequently, it produces

---

[2]https://github.com/cuilimeng/CrackForest-dataset

13

increasingly aberrant results, leading to poor segmentations.

Starting from $\alpha = 50$, the MCC and the ASSD stabilize, which coincides with the binarization of the image. Although the MCC and ASSD exhibit minimal variation beyond $\alpha = 50$, ϵ_{β_0} continues to decrease slowly. By $\alpha = 200$, the results reach a plateau. This observation suggests that the optimal α lies around 200. Moreover, it is noteworthy that increasing α beyond this value does not lead to degradation in results, underscoring the robustness of our approach with respect to this parameter.

4.4. Generalization on various curvilinear structures

By training our regularization operator on synthetic datasets, and successfully applying it to real vascular images, we have demonstrated that our learned reconnecting term is independent of the target dataset. In this section, we further illustrate this robust generalization behavior by applying our approach to non-vascular structures. Specifically, we evaluate our 2D reconnecting term, trained on the OpenCCO dataset, on two additional datasets: one comprising 118 images of road cracks [68, 69], and another containing 30 images of cells from porcine corneal endothelium [70]. We compare the performance of our term with that of the Total Variation and the directional Total Variation. Results are presented in Table 4 and Figure 5.

Despite being trained on synthetic binary vascular trees, our approach consistently produces high-quality results, notably enhancing the connectivity of the segmentation. This shows the effectiveness of our learned reconnecting term, which, by operating on binary structures and prioritizing the geometry of curvilinear structures, exhibits strong generalizability across diverse applications.

It is important to note that our approach rely on the Chan *et al.* segmentation data fidelity term [60] which is based on the assumption that the background of the input image is homogeneous. In our experiments, we consistently preprocessed the datasets to ensure this property. However, in some applications, this preprocessing is not enough to homogenize the image background. This is in particular the case when the image present multiple non-curvilinear objects with a high contrast. In this case, our framework can not be applied directly, and additional preprocessing should be considered such as vesselness filters [42].

5. Conclusion

In this article, we proposed a comprehensive plug-and-play framework for curvilinear structure segmenta-
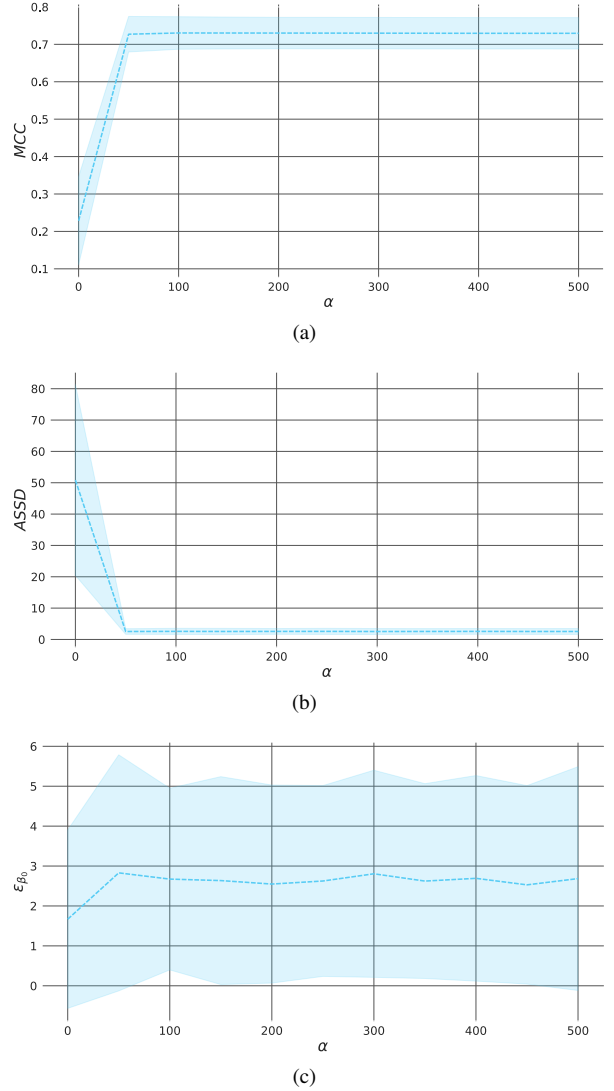

Figure 7: Variation of MCC (a), ASSD (b), and ϵ_{β_0} (c) on the DRIVE dataset in function of the iteration α from which the G_{reco} model is injected into the optimization scheme.

14

Method	MCC	ASSD	ϵ_{β_0}
Cracks			
TV	0.517	14.358	15.928
	± 0.171	± 12.322	± 26.183
Directional TV	0.557	**9.412**	10.043
	± 0.141	± 9.426	± 10.816
E_{reco}	**0.605**	9.691	**2.161**
	± 0.153	± 11.047	± 3.296
Cells			
TV	0.604	1.280	45.033
	0.038	± 0.271	± 33.291
Directional TV	0.606	1.273	46.700
	± 0.038	± 0.276	± 35.109
E_{reco}	**0.637**	**1.261**	**0.267**
	± 0.030	± 0.268	± 0.772

Table 4: Quantitative segmentation results on the cracks and porcine corneal cells.

tion that includes a dedicated constraint for the preservation of connectivity. It includes an algorithm for generating realistic pairs of connected/disconnected curvilinear structures and a reconnecting regularization operator that can be learned from a synthetic dataset. Contrary to classic deep learning segmentation approaches, our approach does not need annotations of the dataset of interest. Finally, we proposed a strategy to include this term into a variational segmentation scheme.

In our study, we demonstrated the efficacy of our regularization term through an ablation analysis. By learning a regularization term that can identify and remove non-curvilinear structure fragments, as well as reconnect true curvilinear structure fragments, significant improvements can be achieved in terms of global structure connectivity. This is evidenced by the topological metrics both in 2D and 3D.

In the future, it would be interesting to provide guarantees on the convergence of our plug-and-play approach. In particular, by learning our reconnecting term to be a maximally monotone operator [71]. This would enhance the robustness and reliability of our approach. Furthermore, we see potential in including our reconnecting term as a fixed layer of an end-to-end segmentation network to enforce connectivity in a supervised setting.

CRediT authorship contribution statement

Sophie Carneiro-Esteves: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Visualization. **Antoine Vacavant:** Conceptualization, Writing - Review & Editing, Supervision, Project administration, Funding acquisition. **Odyssée Merveille:** Conceptualization, Writing - Review & Editing, Supervision, Project administration, Funding acquisition.

Acknowledgment

References

[1] D. B. Mumford, J. Shah, Optimal approximations by piecewise smooth functions and associated variational problems, Communications on pure and applied mathematics (1989).

[2] E. Erdem, A. Sancar-Yilmaz, S. Tari, Mumford-shah regularizer with spatial coherence, in: International Conference on Scale Space and Variational Methods in Computer Vision, Springer, 2007, pp. 545–555.

[3] M. Foare, N. Pustelnik, L. Condat, Semi-linearized proximal alternating minimization for a discrete mumford–shah model, IEEE Transactions on Image Processing 29 (2019) 2176–2189.

[4] O. Merveille, B. Naegel, H. Talbot, N. Passat, n d variational restoration of curvilinear structures with prior-based directional regularization, IEEE Transactions on Image Processing 28 (8) (2019) 3848–3859.

[5] L. Mou, Y. Zhao, H. Fu, Y. Liu, J. Cheng, Y. Zheng, P. Su, J. Yang, L. Chen, A. F. Frangi, et al., Cs2-net: Deep learning segmentation of curvilinear structures in medical imaging, Medical image analysis 67 (2021) 101874.

[6] F. Wang, Y. Gu, W. Liu, Y. Yu, S. He, J. Pan, Context-aware spatio-recurrent curvilinear structure segmentation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 12648–12657.

[7] S. Shit, J. C. Paetzold, A. Sekuboyina, I. Ezhov, A. Unger, A. Zhylka, J. P. Pluim, U. Bauer, B. H. Menze, cldice-a novel topology-preserving loss function for tubular structure segmentation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 16560–16569.

[8] G. Hamarneh, P. Jassi, Vascusynth: Simulating vascular trees for generating volumetric image data with ground-truth segmentation and tree analysis, Computerized medical imaging and graphics 34 (8) (2010) 605–616.

[9] B. Kerautret, P. Ngo, N. Passat, H. Talbot, C. Jaquet, OpenCCO: An Implementation of Constrained Constructive Optimization for Generating 2D and 3D Vascular Trees, Image Processing On Line 13 (2023) 258–279, https://doi.org/10.5201/ipol.2023.477.

[10] T. Shi, X. Ding, L. Zhang, X. Yang, Freecos: self-supervised learning from fractals and unlabeled images for curvilinear object segmentation, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023, pp. 876–886.

[11] A. Runions, M. Fuhrer, B. Lane, P. Federl, A.-G. Rolland-Lagan, P. Prusinkiewicz, Modeling and visualization of leaf venation patterns, in: ACM SIGGRAPH 2005 Papers, 2005, pp. 702–711.

[12] M. Schneider, J. Reichold, B. Weber, G. Székely, S. Hirsch, Tissue metabolism driven arterial tree generation, Medical image analysis 16 (7) (2012) 1397–1414.

[13] L. Lin, L. Peng, H. He, P. Cheng, J. Wu, K. K. Wong, X. Tang, Yolocurvseg: You only label one noisy skeleton for vessel-style curvilinear structure segmentation, Medical Image Analysis 90 (2023) 102937.

[14] Y. Chen, T. Pock, Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration, IEEE transactions on pattern analysis and machine intelligence 39 (6) (2016) 1256–1272.

[15] J. Adler, O. Öktem, Learned primal-dual reconstruction, IEEE transactions on medical imaging 37 (6) (2018) 1322–1332.

[16] T. Meinhardt, M. Moller, C. Hazirbas, D. Cremers, Learning proximal operators: Using denoising networks for regularizing inverse imaging problems, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 1781–1790.

[17] H. Gupta, K. H. Jin, H. Q. Nguyen, M. T. McCann, M. Unser, Cnn-based projected gradient descent for consistent ct image reconstruction, IEEE transactions on medical imaging 37 (6) (2018) 1440–1453.

[18] S. Carneiro Esteves, A. Vacavant, O. Merveille, Learning a reconnecting regularization term for blood vessel variational segmentation, in: 2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI), IEEE, 2021, pp. 1–4.

[19] S. Moccia, E. De Momi, S. El Hadji, L. S. Mattos, Blood vessel segmentation algorithms—review of methods, datasets and evaluation metrics, Computer methods and programs in biomedicine 158 (2018) 71–91.

[20] P. Bibiloni, M. González-Hidalgo, S. Massanet, A survey on curvilinear object segmentation in multiple applications, Pattern Recognition 60 (2016) 949–970.

[21] T. A. Soomro, A. J. Afifi, L. Zheng, S. Soomro, J. Gao, O. Hellwich, M. Paul, Deep learning models for retinal blood vessels segmentation: a review, IEEE Access 7 (2019) 71696–71717.

[22] J. F. Carrillo, M. H. Hoyos, E. E. Dávila, M. Orkisz, Recursive tracking of vascular tree axes in 3d medical images, International Journal of Computer Assisted Radiology and Surgery 1 (2007) 331–339.

[23] W. Liao, S. Wörz, C.-K. Kang, Z.-H. Cho, K. Rohr, Progressive minimal path method for segmentation of 2d and 3d line structures, IEEE transactions on pattern analysis and machine intelligence 40 (3) (2017) 696–709.

[24] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation 234–241.

[25] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, O. Ronneberger, 3d u-net: learning dense volumetric segmentation from sparse annotation, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19, Springer, 2016, pp. 424–432.

[26] X.-F. Du, J.-S. Wang, W.-z. Sun, Unet retinal blood vessel segmentation algorithm based on improved pyramid pooling method and attention mechanism, Physics in Medicine & Biology 66 (17) (2021) 175013.

[27] V. Sathananthavathi, G. Indumathi, Encoder enhanced atrous (eea) unet architecture for retinal blood vessel segmentation, Cognitive Systems Research 67 (2021) 84–95.

[28] A. Affane, J. Lamy, M.-A. Lebre, A. Vacavant, Robust deep 3-d architectures based on vascular patterns for liver vessel segmentation, Informatics in Medicine Unlocked 34 (2022) 101111.

[29] L. Hakim, M. S. Kavitha, N. Yudistira, T. Kurita, Regularizer based on euler characteristic for retinal blood vessel segmentation, Pattern Recognition Letters 149 (2021) 83–90.

[30] X. Hu, F. Li, D. Samaras, C. Chen, Topology-preserving deep image segmentation, Advances in neural information processing systems 32 (2019).

[31] J. R. Clough, N. Byrne, I. Oksuz, V. A. Zimmer, J. A. Schnabel, A. P. King, A topological loss function for deep-learning based image segmentation using persistent homology, IEEE transactions on pattern analysis and machine intelligence 44 (12) (2020) 8766–8778.

[32] M. Haft-Javaherian, M. Villiger, C. B. Schaffer, N. Nishimura, P. Golland, B. E. Bouma, A topological encoding convolutional neural network for segmentation of 3d multiphoton images of brain vasculature using persistent homology, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020, pp. 990–991.

[33] N. Stucki, J. C. Paetzold, S. Shit, B. Menze, U. Bauer, Topologically faithful image segmentation via induced matching of persistence barcodes, in: International Conference on Machine Learning, PMLR, 2023, pp. 32698–32727.

[34] P. Rougé, N. Passat, O. Merveille, Cascaded multitask u-net using topological loss for vessel segmentation and centerline extraction (2023). arXiv:2307.11603.

[35] A. Mosinska, P. Marquez-Neila, M. Koziński, P. Fua, Beyond the pixel-wise loss for topology-aware delineation, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 3136–3145.

[36] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, I. Polosukhin, Attention is all you need, Advances in neural information processing systems 30 (2017).

[37] M. Cheng, K. Zhao, X. Guo, Y. Xu, J. Guo, Joint topology-preserving and feature-refinement network for curvilinear structure segmentation, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 7147–7156.

[38] D. Keshwani, Y. Kitamura, S. Ihara, S. Iizuka, E. Simo-Serra, Topnet: Topology preserving metric learning for vessel tree reconstruction and labelling, in: Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part VI 23, Springer, 2020, pp. 14–23.

[39] M. Lin, K. Zepf, A. N. Christensen, Z. Bashir, M. B. S. Svendsen, M. Tolsgaard, A. Feragen, Dtu-net: Learning topological similarity for curvilinear structure segmentation, in: International Conference on Information Processing in Medical Imaging, Springer, 2023, pp. 654–666.

[40] Y. Li, Y. Zhang, J.-Y. Liu, K. Wang, K. Zhang, G.-S. Zhang, X.-F. Liao, G. Yang, Global transformer and dual local attention network via deep-shallow hierarchical feature fusion for retinal vessel segmentation, IEEE Transactions on Cybernetics (2022).

[41] X. Shen, J. Xu, H. Jia, P. Fan, F. Dong, B. Yu, S. Ren, Self-attentional microvessel segmentation via squeeze-excitation transformer unet, Computerized Medical Imaging and Graphics 97 (2022) 102055.

[42] J. Lamy, O. Merveille, B. Kerautret, N. Passat, A benchmark framework for multi-region analysis of vesselness filters, IEEE Transactions on Medical Imaging PP (2022) 1–1. doi:10.1109/TMI.2022.3192679.

[43] T. Shi, N. Boutry, Y. Xu, T. Géraud, Local intensity order transformation for robust curvilinear object segmentation, IEEE Transactions on Image Processing 31 (2022) 2557–2569.

[44] Y. Peng, L. Pan, P. Luan, H. Tu, X. Li, Curvilinear object segmentation in medical images based on odos filter and deep learning network, arXiv preprint arXiv:2301.07475 (2023).

[45] H. Du, X. Zhang, G. Song, F. Bao, Y. Zhang, W. Wu, P. Liu, Retinal blood vessel segmentation by using the ms-lsdnet network and geometric skeleton reconnection method, Computers in Biology and Medicine 153 (2023) 106416.

[46] L. Mou, L. Chen, J. Cheng, Z. Gu, Y. Zhao, J. Liu, Dense dilated network with probability regularized walk for vessel detection, IEEE Transactions on Medical Imaging 39 (5) (2020) 1392–1403. doi:10.1109/TMI.2019.2950051.

16

[47] V. S. Joshi, M. K. Garvin, J. M. Reinhardt, M. D. Abramoff, Identification and reconnection of interrupted vessels in retinal vessel segmentation, in: 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, IEEE, 2011, pp. 1416–1420.

[48] J. Zhang, E. Bekkers, D. Chen, T. T. Berendschot, J. Schouten, J. P. Pluim, Y. Shi, B. Dashtbozorg, B. M. ter Haar Romeny, Reconnection of interrupted curvilinear structures via cortically inspired completion for ophthalmologic images, IEEE Transactions on Biomedical Engineering 65 (5) (2018) 1151–1165.

[49] L. I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D: nonlinear phenomena 60 (1-4) (1992) 259–268.

[50] O. Merveille, H. Talbot, L. Najman, N. Passat, Curvilinear structure analysis by ranking the orientation responses of path operators, IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) 40 (2) (2018) 304–317. doi:10.1109/TPAMI.2017.2672972.

[51] A. F. Frangi, W. J. Niessen, K. L. Vincken, M. A. Viergever, Multiscale vessel enhancement filtering, in: Medical Image Computing and Computer-Assisted Intervention—MICCAI'98: First International Conference Cambridge, MA, USA, October 11–13, 1998 Proceedings 1, Springer, 1998, pp. 130–137.

[52] F. Heide, M. Steinberger, Y. Tsai, M. Rouf, D. Pajak, D. Reddy, O. Gallo, J. L., W. Heidrich, K. Egiazarian, et al., Flexisp: A flexible camera image processing framework, ACM Transactions on Graphics 33 (6) (2014) 1–13.

[53] K. Dabov, A. Foi, K. Egiazarian, Video denoising by sparse 3d transform-domain collaborative filtering [c], in: Proc. 15th European Signal Processing Conference, Vol. 1, 2007, p. 7.

[54] K. Zhang, W. Zuo, S. Gu, L. Zhang, Learning deep cnn denoiser prior for image restoration, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[55] J. Rick Chang, C.-L. Li, B. Poczos, B. Vijaya Kumar, A. C. Sankaranarayanan, One network to solve them all–solving linear inverse problems using deep projection models, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 5888–5897.

[56] M. Le Pendu, C. Guillemot, Preconditioned plug-and-play admm with locally adjustable denoiser for image restoration, SIAM Journal on Imaging Sciences 16 (1) (2023) 393–422.

[57] C. Douarre, C. Crispim-Junior, A. Gelibert, L. Tougne, D. Rousseau, Novel data augmentation strategies to boost supervised segmentation of plant disease, Computers and Electronics in Agriculture 165 (2019).

[58] E. Kerfoot, J. Clough, I. Oksuz, J. Lee, A. King, J. Schnabel, Left-ventricle quantification using residual u-net, in: International Workshop on Statistical Atlases and Computational Models of the Heart, 2018, pp. 371–380.

[59] F. Milletari, N. Navab, S.-A. Ahmadi, V-net: Fully convolutional neural networks for volumetric medical image segmentation, in: 2016 fourth international conference on 3D vision (3DV), Ieee, 2016, pp. 565–571.

[60] T. Chan, S. Esedoglu, M. Nikolova, Algorithms for finding global minimizers of image segmentation and denoising models, SIAM journal on applied mathematics 66 (5) (2006) 1632–1648.

[61] N. Komodakis, J. Pesquet, Playing with duality: An overview of recent primal-dual approaches for solving large-scale optimization problems, Signal Processing Magazine 32 (6) (2015) 31–54.

[62] J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, B. Van Ginneken, Ridge-based vessel segmentation in color images of the retina, IEEE transactions on medical imaging 23 (4) (2004) 501–509.

[63] H. Boudegga, Y. Elloumi, M. Akil, M. H. Bedoui, R. Kachouri, A. B. Abdallah, Fast and efficient retinal blood vessel segmentation method based on deep learning network, Computerized Medical Imaging and Graphics 90 (2021) 101902.

[64] A. Affane, A. Kucharski, P. Chapuis, S. Freydier, M.-A. Lebre, A. Vacavant, A. Fabijańska, Segmentation of liver anatomy by combining 3d u-net approaches, Applied Sciences 11 (11) (2021) 4895.

[65] L. Yu, S. Wang, X. Li, C.-W. Fu, P.-A. Heng, Uncertainty-aware self-ensembling model for semi-supervised 3d left atrium segmentation, in: Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part II 22, Springer, 2019, pp. 605–613.

[66] A. Hoover, V. Kouznetsova, M. Goldbaum, Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response, IEEE Transactions on Medical imaging 19 (3) (2000) 203–210.

[67] C. Li, W. Ma, L. Sun, X. Ding, Y. Huang, G. Wang, Y. Yu, Hierarchical deep network with uncertainty-aware semi-supervised learning for vessel segmentation, Neural Computing and Applications (2022) 1–14.

[68] Y. Shi, L. Cui, Z. Qi, F. Meng, Z. Chen, Automatic road crack detection using random structured forests, IEEE Transactions on Intelligent Transportation Systems 17 (12) (2016) 3434–3445.

[69] L. Cui, Z. Qi, Z. Chen, F. Meng, Y. Shi, Pavement distress detection using random decision forests, in: International Conference on Data Science, Springer, 2015, pp. 95–102.

[70] A. Ruggeri, F. Scarpa, M. De Luca, C. Meltendorf, J. Schroeter, A system for the automatic estimation of morphometric parameters of corneal endothelium in alizarine red-stained images, British Journal of Ophthalmology 94 (5) (2010) 643–647.

[71] J.-C. Pesquet, A. Repetti, M. Terris, Y. Wiaux, Learning maximally monotone operators for image recovery, SIAM Journal on Imaging Sciences 14 (3) (2021) 1206–1237.

17