Full name of the applicant	Victor Joos
Reference	

SCIENTIFIC SECTION OF THE PROPOSAL

MAIN LANGUAGE CHOSEN = ENGLISH

This part includes the following elements:

1. Description of the research project
2. Comments on changes made in the research project in case of resubmission (optional)
3. Activities report on the first year of doctorate (**ONLY** for 1st grant - **2nd year** applicants)*
4. Potential interdisciplinary approach of the research project (optional)
5. Description of the work environment
6. Summary of the master's thesis or equivalent
7. Additionnal comments (optional)
8. Ph.D. work calendar per month

*** "1st grant - 2nd year" applicants have already worked on a full-time basis for one year full time equivalent on the Ph.D. project submitted to the FRIA.**

The applicant must fill in the sections below and convert the file into an unprotected PDF before appending it to the online application form.

The F.R.S.-FNRS insists on **strict compliance with the instructions given for each part of the proposal** (scientific section relevant to the instrument selected, number of pages allowed for documents to be enclosed with the application form…) and stresses again the sovereign consideration of the juries in case the file would exceed the applicable page limit.

1. DESCRIPTION OF THE RESEARCH PROJECT

The written project must be made up of 4 parts (max. 4 pages) according to the structure below, accompanied by a reference bibliography (max. 1 page besides the 4 pages dedicated to the project) listed by order of appearance within the text.

Graphs and tables may be added (max. 2 pages).

1.1 Goals of the research

In the context of biomedical imaging, and in particular computed tomography, deep learning methods have shown promise for a number of tasks, from disease classification and early detection, to segmentation of tumors, cartilage, and vessels. However, the use of convolutional neural networks have not shown the same overwhelming success and widespread use as in natural images. This stems from two main difficulties 3D CT scans, shown in Fig. 1, face in comparison to natural images:

- CT scans are difficult and time-consuming to annotate, and require (sometimes multiple) expert annotators to have a faithful and correct segmentation
- Noisy scans, and staining differences during acquisition make general models difficult to train without a large annotated database.

With access to computed tomography (CT) scans, as 3D volumes, from a 4 year project in collaboration with Prof. Greet Kerckhofs, as well as existing datasets from earlier work [1], we aim in particular to segment branched structures like vascular networks or nerves. This will have impact on the following applications:

- Tissue engineering need fully segmented vascular networks and nerves. The segmentation of small vessels or capillaries in particular is challenging.
- Pulmonary embolism detection requires the segmentation of vessels in the lungs.

This project will attempt to bypass the need for large annotated datasets by using self-supervised learning methods, which use the content of the image itself, and prior knowledge on the branch-like structures of the segmentation task that are exhibited in vascular and nervous networks.

1.2 State of the art

Convolutional neural networks (or CNN) have surpassed traditional methods on almost all vision tasks requiring supervision. They have successfully been used for classification [??], object detection [??], image generation [??], super-resolution [??], and semantic segmentation [??]. The last task is best served by an auto-encoder structure, with skip-connections, as shown in U-Net [2] for 2D biomedical images or 3D U-Net [3] and V-Net [4] for their 3D equivalent.

All the methods mentioned previously use supervision in order to learn a representation, with the need to annotate a lot of data. This need is prohibitive, especially for vascular segmentation, where experts annotations are needed, and accurate segmentation is time-consuming ([5] extrapolates more than a year of annotation for the vascular structure of one mouse brain from the annotation of 0.02% of a single brain). Previous work, including from UCLouvain, has studied the use of synthetic data [Simone][5], [6], but problems in domain transfer between the synthetic data and the real data still necessitate a not-small subset of data from expert annotation (11% in [5]).

To alleviate the need of expert annotation, this project will look at self-supervised learning for vascular networks. In order for a network to learn with no supervision, it needs a pretext task, called self-supervision. Completing a jigsaw puzzle [7], predicting rotation [8], or predicting

relative placement of a part of the image [9] have shown various levels of success in training the network for the subsequent task. In recent years, contrastive methods [10]–[14] have shown strong performance on classification tasks. Contrastive methods cluster related samples (the same image with various kinds of data augmentation, as shown in Fig.??), while maximizing distance between different images. This contrastive framework can be combined with pretext tasks [15], [16]. Due to the use of strong data augmentation, such contrastive methods are not directly suited for segmentation. **Our project strives to define contrastive losses and pretext tasks designed for segmentation, and in particular for vascular networks.**

To define an appropriate pretext task, we might look at the use of joint learning of orientation and segmentation [17], or branching detection [18], or at the design of anatomically constrained neural networks [19]. The numerous different designs of synthetic datasets using simple rules for the creation of vascular structure [6], [18], [20], [21] show an advantage from looking at this prior information in order to help self-supervised learning.

1.3 **Research project**

In order to perform successful segmentation of vascular and nervous networks, we plan to:

- define self-supervised mechanisms to reduce the need of expert annotations. Different kinds of pretext tasks and contrastive losses will be developed for vascular network segmentation.
- account for vascular or nervous structure as prior for the self-supervision task. We will first investigate the interaction between prior definition and self-supervised learning: joint learning of the main segmentation task and either branching detection or vascular orientation can lead to better results in the different sub-tasks. We will also define a self-supervised loss function that takes branching structures into account.

1.4 **Work plan**

As explained in the previous section, the work plan of this project is based on two parts: self-supervised learning for segmentation networks, and for the segmentation of vascular networks.

Transversal to the following tasks is the collection and annotation of a contrast-enhanced CT dataset, through close collaboration with Prof. Greet Kerckhofs' team as part of the 4 year Bio-Blueprints effort. In addition to this, we deem it wise to start working on synthetic datasets (as previously developed by [Simone]), in order to work out an upper bound to possible self-supervised accuracy.

Objective 1: Self-supervised Learning for segmentation

Task 1. <u>Supervised CNN implementation for segmentation</u>

State-of-the-art deep learning segmentation architectures are fully convolutional neural networks. For biomedical applications, the U-Net architecture (2D [2] or 3D [3] implementation) is particularly suited and is commonly used in the community. The U-Net architecture consists in a contracting encoder part to analyze the whole input image and a successive expanding part to produce a full-resolution segmentation. Skip connections between the contracting and expanding paths allow to combine high resolution features extracted in the contracting path with the up-sampled output. 2D and 3D U-Net have been implemented in UCLouvain to segment the bladder on CT and CBCT image slices [22], [23], but also the mineralized cartilage on micro-CT

images [1]. Those two networks serve as a baseline reference to compare with the alternatives studied in the rest of the project.

Task 2. Self-supervised segmentation using dictionary learning

We plan to investigate dictionary learning tasks as proxy tasks. Unsupervised representation learning has been largely used in natural language processing, but has been less successful for vision tasks. Contrastive loss [24], illustrated in Fig. 2 (a), has very recently been envisioned to learn image representations that effectively transfer to a variety of natural vision tasks. We plan to investigate a similar approach in the context of 3D biomedical image segmentation.

The contrastive loss commonly used for self-supervised learning is defined on a comparison between 2 images, which is ill-suited for the end-to-end training needed by U-Net, which doesn't rely on a shallow decoder. As Fig. 2 (b) shows, we will investigate a pixel-level approach, where pixels from the same image, but different data augmentation schemes will be grouped in feature space, while different pixels from the same or an other image will be separated.

Task 3. Proxy task for 3D segmentation using dictionary learning

As shown in [15], it might make sense to combine a proxy task with the use of a contrastive loss, instead of working at the image level. Few works have investigated the definition of proxy tasks that are suited for 3D segmentation, where the implementation of tasks like jigsaw solving [7] or rotation prediction [8] seem ill-suited to fine-grained segmentation needed in our practical context, either for their definition on images as a whole, or on patches. Our goal is to define pixel-oriented proxy-tasks.

In order to validate and understand the inner-working of the designed tasks, we will look at the weight and gradient characteristics during training, prompted by work done inside our research team [Simon].

Objective 2: Branch-like structures as prior for self-supervised learning

Task 4. Favoring branching priors through CNN regularization

When CNNs are only trained with the classical binary cross-entropy or the Dice loss as it is the case in Figure 1, the network prediction and the ground truth (i.e. the manual annotations) are compared on a pixel-wise manner and do not necessarily incorporate local geometry such as smoothness and shape features (e.g. tree-based topology of vascular elements). In this project, we consider the regularization of internal CNN representations with respect to the output domain.
The purpose is to favor the emergence of a segmentation mask that shares the structure exhibited by the ground truth output labels. Hence, the proposed CNN regularization scheme relies on the assumption that the ground truth labels exhibit such an internal structure. This is the case in the considered datasets since the organs in CT/micro-CT images contain connected structures (which implies that a shape such as the one segmented in the right image of Figure 1 is definitely not relevant).
Our proposed regularization scheme works in two steps. First, an auto-encoder will be trained to model the internal structure of ground-truth segments. In a second phase, the actual segmentation network of interest (i.e. the U-Net model) is trained by including the auxiliary task of predicting the output via the decoder learned in the first phase [25]. This process is illustrated in Fig. 3. In our case, the auto-encoder will be trained on manual or synthetic segmentation labels embedding models for ramifications.

Alternative approaches to connect the segmentation and the structure encoding networks have been proposed [19], [26] as well. Hence, determining the appropriate way to connect them and to favor the joint training of multiple (i.e. auxiliary and main) tasks by a network remains open and is fully part of this research question [27], [28].

Task 5. Prior information in contrastive losses

The contrastive loss formulation for segmentation of Task 2 gives us another angle at which to solve the problem of prior information. Instead of separating all pixels from an image in the feature space indiscriminately, we can look at local information, using for example graph cut [29], or watershed [30] algorithms, or using predefined filters to describe the natural branch-like structure of the networks. All methods widely used in interactive segmentation of vascular networks [31]–[33].

In this same context, we might try to define pretext tasks more suited to the segmentation of vascular networks.

[1] J. Léger, L. Leyssens, C. De Vleeschouwer, and G. Kerckhofs, “Deep Learning-Based Segmentation of Mineralized Cartilage and Bone in High-Resolution Micro-CT Images,” in *Computer Methods, Imaging and Visualization in Biomechanics and Biomedical Engineering*, Cham, 2020, pp. 158–170, doi: 10.1007/978-3-030-43195-2_12.

[2] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Cham, 2015, pp. 234–241, doi: 10.1007/978-3-319-24574-4_28.

[3] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, “3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, Cham, 2016, pp. 424–432, doi: 10.1007/978-3-319-46723-8_49.

[4] F. Milletari, N. Navab, and S.-A. Ahmadi, “V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation,” in *2016 Fourth International Conference on 3D Vision (3DV)*, Oct. 2016, pp. 565–571, doi: 10.1109/3DV.2016.79.

[5] M. I. Todorov *et al.*, “Machine learning analysis of whole mouse brain vasculature,” *Nature Methods*, vol. 17, no. 4, Art. no. 4, Apr. 2020, doi: 10.1038/s41592-020-0792-1.

[6] B. Ma, S. Liu, Y. Zhi, and Q. Song, “Flow Based Self-supervised Pixel Embedding for Image Segmentation,” Jan. 2019, [Online]. Available: http://arxiv.org/abs/1901.00520.

[7] M. Noroozi and P. Favaro, “Unsupervised Learning of Visual Representations by Solving Jigsaw Puzzles,” *arXiv:1603.09246 [cs]*, Aug. 2017, Accessed: Jul. 28, 2020. [Online]. Available: http://arxiv.org/abs/1603.09246.

[8] S. Gidaris, P. Singh, and N. Komodakis, “Unsupervised Representation Learning by Predicting Image Rotations,” *arXiv:1803.07728 [cs]*, Mar. 2018, Accessed: Jul. 28, 2020. [Online]. Available: http://arxiv.org/abs/1803.07728.

[9] M. Blendowski, H. Nickisch, and M. P. Heinrich, “How to Learn from Unlabeled Volume Data: Self-supervised 3D Context Feature Learning,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, Cham, 2019, pp. 649–657, doi: 10.1007/978-3-030-32226-7_72.

[10] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, “A Simple Framework for Contrastive Learning of Visual Representations,” 2020, [Online]. Available: http://arxiv.org/abs/2002.05709.

[11] M. Caron, I. Misra, J. Mairal, P. Goyal, P. Bojanowski, and A. Joulin, “Unsupervised Learning of Visual Features by Contrasting Cluster Assignments,” *arXiv:2006.09882 [cs]*, Jul. 2020, Accessed: Jul. 28, 2020. [Online]. Available: http://arxiv.org/abs/2006.09882.

[12] X. Chen, H. Fan, R. Girshick, and K. He, “Improved Baselines with Momentum Contrastive Learning,” *arXiv:2003.04297 [cs]*, Mar. 2020, Accessed: Jul. 28, 2020. [Online]. Available: http://arxiv.org/abs/2003.04297.

[13] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, “Momentum Contrast for Unsupervised Visual Representation Learning,” 2019, [Online]. Available: http://arxiv.org/abs/1911.05722.

[14] Y. Tian, D. Krishnan, and P. Isola, “Contrastive Multiview Coding,” *arXiv:1906.05849 [cs]*, Mar. 2020, Accessed: Jul. 28, 2020. [Online]. Available: http://arxiv.org/abs/1906.05849.

[15] A. van den Oord, Y. Li, and O. Vinyals, “Representation Learning with Contrastive Predictive Coding,” *arXiv:1807.03748 [cs, stat]*, Jan. 2019, Accessed: Jul. 28, 2020. [Online]. Available: http://arxiv.org/abs/1807.03748.

[16] O. J. Hénaff *et al.*, “Data-Efficient Image Recognition with Contrastive Predictive Coding,” *arXiv:1905.09272 [cs]*, Jul. 2020, Accessed: Jul. 28, 2020. [Online]. Available: http://arxiv.org/abs/1905.09272.

[17] A. Batra, S. Singh, G. Pang, S. Basu, C. V. Jawahar, and M. Paluri, “Improved road connectivity by joint learning of orientation and segmentation,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2019, vol. 2019-June, pp. 10377–10385, doi: 10.1109/CVPR.2019.01063.

[18] G. Tetteh *et al.*, “DeepVesselNet: Vessel Segmentation, Centerline Prediction, and Bifurcation Detection in 3-D Angiographic Volumes,” *ArXiv*, 2018.

[19] O. Oktay *et al.*, “Anatomically Constrained Neural Networks (ACNNs): Application to Cardiac Image Enhancement and Segmentation,” *IEEE Transactions on Medical Imaging*, vol. 37, no. 2, pp. 384–395, 2018, doi: 10.1109/TMI.2017.2743464.

[20] M. Schneider, S. Hirsch, B. Weber, G. Székely, and B. H. Menze, “Joint 3-D vessel segmentation and centerline extraction using oblique Hough forests with steerable filters,” *Medical Image Analysis*, vol. 19, no. 1, pp. 220–249, Jan. 2015, doi: 10.1016/j.media.2014.09.007.

[21] Q. Zhang and A. C. S. Chung, “3D Vessel Segmentation Using Random Walker with Oriented Flux Analysis and Direction Coherence,” in *Medical Imaging and Augmented Reality*, Cham, 2016, pp. 281–291, doi: 10.1007/978-3-319-43775-0_25.

[22] J. Léger, E. Brion, U. Javaid, J. Lee, C. De Vleeschouwer, and B. Macq, “Contour Propagation in CT Scans with Convolutional Neural Networks,” in *Advanced Concepts for Intelligent Vision Systems*, Cham, 2018, pp. 380–391, doi: 10.1007/978-3-030-01449-0_32.

[23] E. Brion, J. Léger, U. Javaid, J. Lee, C. D. Vleeschouwer, and B. Macq, “Using planning CTs to enhance CNN-based bladder segmentation on cone beam CT,” in *Medical Imaging 2019: Image-Guided Procedures, Robotic Interventions, and Modeling*, Mar. 2019, vol. 10951, p. 109511M, doi: 10.1117/12.2512791.

[24] M. Gutmann and A. Hyvärinen, “Noise-contrastive estimation: A new estimation principle for unnormalized statistical models,” in *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, Mar. 2010, pp. 297–304, Accessed: Jul. 28, 2020. [Online]. Available: http://proceedings.mlr.press/v9/gutmann10a.html.

[25] M. Mostajabi, M. Maire, and G. Shakhnarovich, “Regularizing Deep Networks by Modeling and Predicting Label Structure,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2018, pp. 5629–5638, doi: 10.1109/CVPR.2018.00590.

[26] H. Ravishankar, R. B. Venkataramani, S. Thiruvenkadam, and P. Sudhakar, “Learning and Incorporating Shape Models,” *Miccai 2017*, vol. 10433, no. 2, pp. 203–211, 2017, doi: 10.1007/978-3-319-66182-7.

[27] S. Carbonnelle and C. De Vleeschouwer, “Layer rotation: a surprisingly simple indicator of generalization in deep networks?,” presented at the ICML, 2019.

[28] Y. Sun, X. Wang, Z. Liu, J. Miller, A. A. Efros, and M. Hardt, “Test-Time Training with Self-Supervision for Generalization under Distribution Shifts,” *arXiv:1909.13231 [cs, stat]*, Jul. 2020, Accessed: Jul. 28, 2020. [Online]. Available: http://arxiv.org/abs/1909.13231.

[29] Y. Boykov, O. Veksler, and R. Zabih, “Fast approximate energy minimization via graph cuts,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222–1239, Nov. 2001, doi: 10.1109/34.969114.

[30] L. Vincent and P. Soille, “Watersheds in digital spaces: an efficient algorithm based on immersion simulations,” *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 6, pp. 583–598, 1991.

[31] H. P. Ng, S. H. Ong, K. W. C. Foong, P. S. Goh, and W. L. Nowinski, “Medical Image Segmentation Using K-Means Clustering and Improved Watershed Algorithm,” in *2006 IEEE Southwest Symposium on Image Analysis and Interpretation*, Mar. 2006, pp. 61–65, doi: 10.1109/SSIAI.2006.1633722.

[32] Y. Boykov and M.-P. Jolly, “Interactive Organ Segmentation Using Graph Cuts,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2000*, Berlin, Heidelberg, 2000, pp. 276–286, doi: 10.1007/978-3-540-40899-4_28.

[33] D. Freedman and Tao Zhang, “Interactive graph cut based segmentation with shape priors,” in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*, Jun. 2005, vol. 1, pp. 755–762 vol. 1, doi: 10.1109/CVPR.2005.191.

[34] G. Kerckhofs *et al.*, “Contrast-enhanced microCT to visualize and quantify the 3D vasculature in biological tissues without the need for perfusion,” 20170601, Accessed: Aug. 03, 2020. [Online]. Available: https://lirias.kuleuven.be/retrieve/451736.

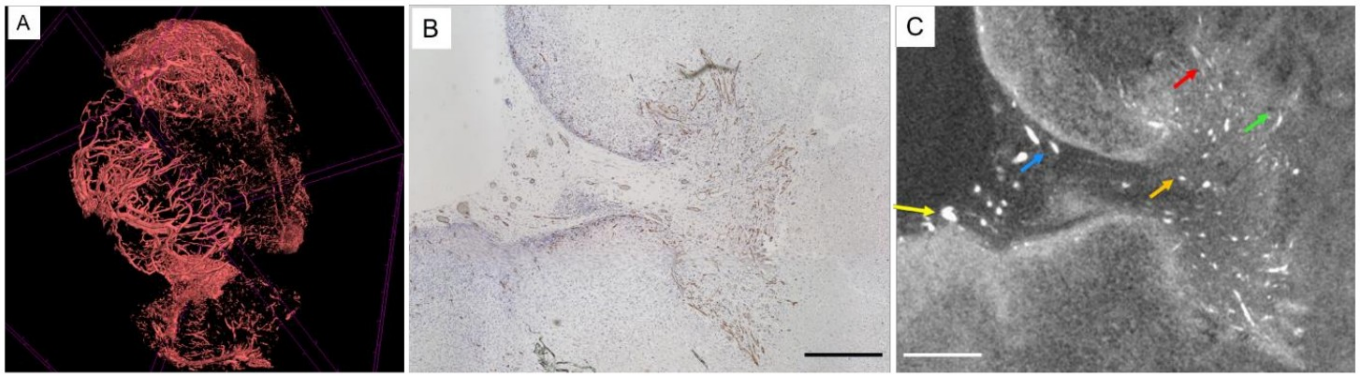


Figure 1: CE-microCT images of vascularization in a tumour xenograft sample, adapted from the work by Kerckhofs et al. [34] (A) 3D rendering of the vasculature in the xenograft, stained with Hf-WD POM; 3D scale bar = 100 µm. (B) The CD31 stained section and (C) the corresponding CE-microCT cross-section through the tumour xenograft. The brown colour in the histological section indicates CD31 positive blood vessels. The white colour in the CE-microCT image represents red blood cells in the blood vessels. Scale bars = 100 µm.

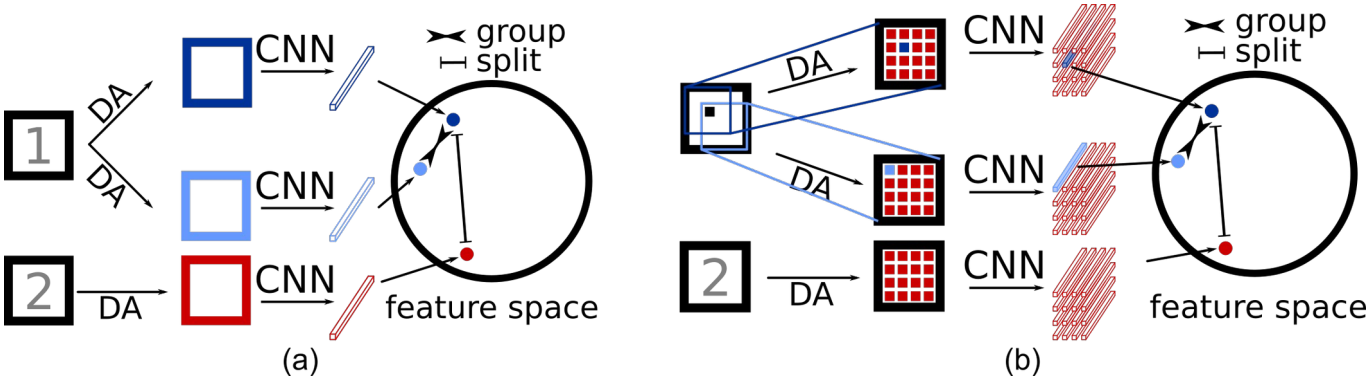
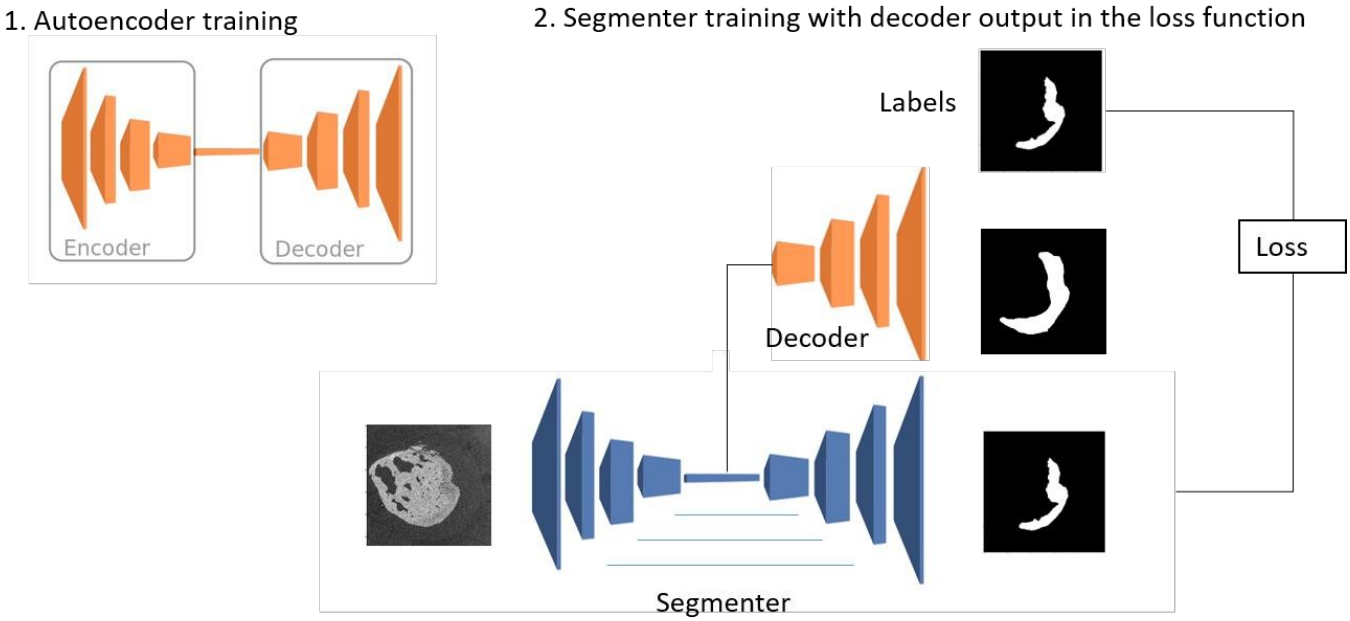


Figure 2: Illustration of contrastive loss for self-supervised learning. (a) Strategy for self-supervised classification. Each image undergoes 2 forms of data-augmentation (only shown for Im1). Once each are passed through a model the features from the same image are grouped, while isolated from other images. (b) Strategy for self-supervised segmentation. The same method as in (a) is used, pixel by pixel, taking zooming and cropping into account.

2. COMMENTS ON CHANGES MADE IN THE RESEARCH PROJECT IN CASE OF RESUBMISSION (OPTIONAL)

In case of former application submitted to the F.R.S.-FNRS via the same funding instrument, please specify the main changes made in your funding application following previous submission, identifying comments from experts that you may have taken into account (max. 1 page).



Not Applicable

3. ACTIVITIES REPORT ON THE FIRST YEAR OF DOCTORATE

ONLY FOR "1ST GRANT - 2ND YEAR" APPLICANTS

Please write a brief report (max. 2 pages) underlining the progress of your research during the first year of your doctorate.

Not Applicable

4. POTENTIAL INTERDISCIPLINARY APPROACH OF THE RESEARCH PROJECT (OPTIONAL)

If applicable, please identify the interdisciplinary approach of your research project (max. 1 page).

Thanks to a 4-year inter-university project, called Bio-blueprints, in partnership with Prof. Greet Kerckhofs, this project enjoys support in the creation of a dataset and the definition of success metrics.

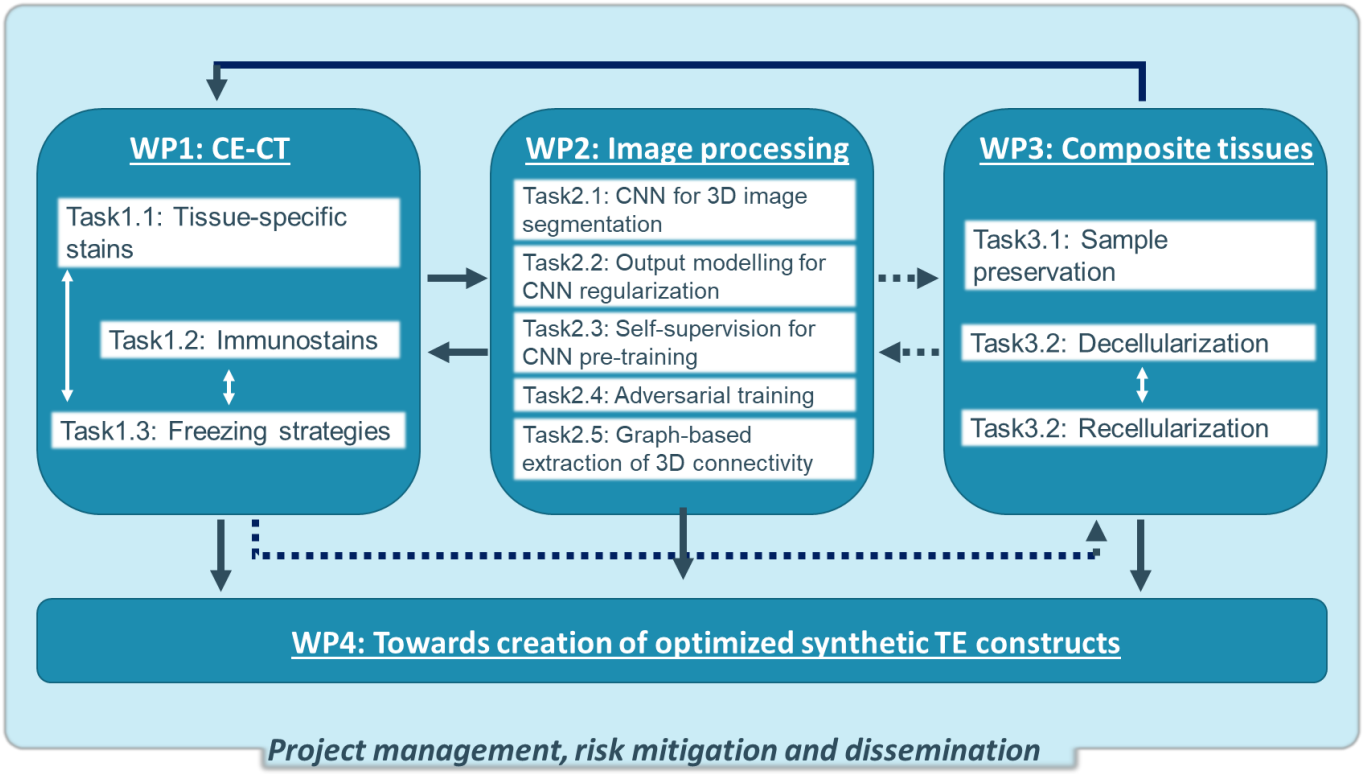


Figure 4: Schematic overview of the Bio-blueprints organization, as presented by its work packages and their tasks. The solid arrows indicate the input from one work package to another. The dashed arrows indicate feedback and potential reiteration based on the outcome of a WP or task.

5. DESCRIPTION OF THE WORK ENVIRONMENT

Please provide the information accounting for the adequacy of the environment (available intellectual and/or material means) to carry out the research as detailed in the submitted project. Please specify the assets of the research environment related to the project and the main publications of the laboratory/promoter (max. 1 page).

TODO

6. SUMMARY OF THE MASTER'S THESIS OR EQUIVALENT

Please provide a summary of your master's thesis or any equivalent, even if you have not been graduated yet (max. 1 page).

7. ADDITIONAL COMMENTS (OPTIONAL)

If you want to communicate elements that have not been mentioned elsewhere in the file, please provide this information below in max. 2 pages.

Please note that in case the presented project provides for the involvement of patients and/or human or animal subjects, it is important that the project includes justifications on the planned sample size (number of subjects included in the study/studies) and how the size is relevant (based on statistical power calculations, for instance). It

is also important to explain how the number of patients/subjects expected can be reached. In case the project provides for the involvement of patients and/or subjects, please provide those pieces of information under this section (if not already mentioned elsewhere in the project). Ultimately, this information (or the lack of information) may be taken into account by experts in the frame of the evaluation of your funding application.

[Enter text here. Format: Arial 12, single space]

8. PH.D. WORK CALENDAR PER MONTH

*Please provide a calendar on a monthly basis for your Ph.D. works planned for the next 3 years (*1st grant - 2nd year) *or 4 years (*1st grant - 1st year) *(max. 2 pages).*

[Enter text here. Format: Arial 12, single space]