



HAL
open science

Self super-resolution of anisotropic volumes in prostate MRI with normalized edge priors

Vincent Jaouen, Pierre-Henri Conze, Julien Bert, Dimitris Visvikis

► To cite this version:

Vincent Jaouen, Pierre-Henri Conze, Julien Bert, Dimitris Visvikis. Self super-resolution of anisotropic volumes in prostate MRI with normalized edge priors. ICCR 2024: XXth International Conference on the use of Computers in Radiation therapy, Jul 2024, Lyon, France. <hal-04714832>

HAL Id: hal-04714832

<https://imt-atlantique.hal.science/hal-04714832v1>

Submitted on 30 Sep 2024

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

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



HAL Authorization

Self super-resolution of anisotropic volumes in prostate MRI with normalized edge priors

Vincent Jaouen^{1,2}, Pierre-Henri Conze^{1,2}, Julien Bert^{2,3}, and Dimitris Visvikis²

¹IMT Atlantique, Brest, France

²UMR 1101 INSERM LaTIM, IMT Atlantique, UBO, Brest, France

³CHRU Brest, France

Abstract Due to involuntary organ motion, prostate magnetic resonance images are often acquired in 2D with thick through plane spacings and reconstructed into anisotropic voxelized volumes. In prostate cancer (PCa) radiotherapy, annotations such as zonal or lesion segmentation are then performed on the few axial slices in which the prostate appears, leading to coarse and often inconsistent delineations across slices, which are nevertheless used as ground truth for deep segmentation model training and performance reports.

Recently, deep self-supervised super resolution (SSR) was proposed in brain MRI to remove the need for calibrated training data, by learning the low-resolution to high-resolution mapping from the anisotropic image itself. In this work, we propose a new training objective (loss function) for SSR based on the alignment of normalized image gradient vectors, a technique normally used for image registration. Preliminary experiments on the public Prostate158 dataset suggest that automatic PCa detection could be improved using SSR with the proposed training objective. However, the low quality of lesion annotations obtained on a few slices is a major bottleneck to confirm these findings quantitatively.

1 Introduction

1.1 Context

Prostate cancer (PCa) is a major form of cancer that affects about one man in six during his lifetime [1]. Approximately 75% of PCa originate in the peripheral zone of the gland (PZ), as opposed to the more central transition zone (TZ) [2, 3]. Radiotherapy treatment planning of PCa relies on multiparametric magnetic resonance imaging (mpMRI) comprising T2-weighted and diffusion-weighted sequences for lesion detection and characterization [4].

1.2 Image anisotropy

Due to involuntary movements, recommendations are to perform fast 2D T2-weighted pulse sequences, preferably along the three anatomical planes but sometimes along the axial direction only [4]. However, such acquisitions lead to high resolution (HR) in-plane images (e.g. from $0.5 \times 0.5 \text{ mm}^2$ down to $0.25 \times 0.25 \text{ mm}^2$, Fig. 1a) but to low resolution (LR) through-plane images due to large spacings (typically of around 3 mm, Fig. 1b). Such anisotropy in the data representation may hamper image analysis downstream. In particular, the detection and characterization of the PCa lesions from such degraded acquisitions is difficult and subjected to high inter reader variability. In this context, the automatization of these tasks of high clinical relevance.

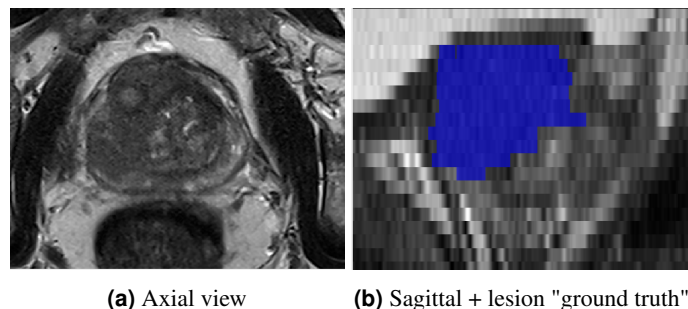


Figure 1: Example 2D axial prostate MRI acquisition with high in-plane resolution (0.27 mm^2) and thick through-plane spacing (3 mm). The lesion mask is also imprecise with strong "staircasing" effect.

1.3 Objective

In this work, we study whether enhancing image resolution through Single Image Super-Resolution (SISR) techniques can improve PCa lesion deep learning-based segmentation. More specifically, we leverage a self-supervised super resolution (SSR) method, SMORE, recently proposed for brain MR SISR [5]. SMORE alleviates the need for any training data outside the anisotropic image itself. In this context, we propose a new reconstruction objective (loss function) to improve SMORE results based on the alignment of normalized gradient fields (NGF) between the LR and HR planes, a technique normally used for image registration [6]. Preliminary results on the Prostate158 public dataset suggest that SSR is indeed beneficial to PCa detection, and that the addition of the proposed NGF loss term may further improve the quality of the detection.

2 Related works

The majority of modern computerized medical image analysis approaches rely on (deep) machine learning techniques. A canonical example is the U-Net convolutional neural network (CNN) model [7] for segmentation [8] and detection [9]. When dealing with highly anisotropic data, little to no gain is achieved using 3D models, and 2D models are generally preferred [8]. However, 2D models inherently discards the consistency of the signal along the (degraded) third dimension, which is suboptimal.

Among major research directions, careful curation and pre-processing of input data is generally considered a critical aspect for the improvement of learning-based model performances [8]. In this work, we hypothesize that improving the resolution of highly anisotropic images through self-supervised SSIR techniques can be beneficial to a variety of downstream image analysis tasks.

2.1 Single image super resolution

Current state of the art SSIR methods generally require curated external supervised training data with predetermined downsampling factors to learn a mapping between the LR and HR spaces using e.g. deep residual networks [10]. However, such training data may be difficult to obtain in practice, as data collections are typically gathered from multiple centers with heterogeneous practices. Moreover, these methods are subject to the domain shift between training and deployment conditions.

2.2 Self-supervised super-resolution with SMORE

Recently, Zhao et al. proposed Synthetic Multi-Orientation Resolution Enhancement (SMORE), a self-supervised super resolution (SSR) approach for anisotropic volumes [5]. SMORE removes the need for external training data as the LR to HR mapping is learned directly on the 3D volume itself (i.e. between the two LR planes and the HR plane). First, SMORE proposes a forward degradation model on the HR 2D slices to mimic the degradation in the LR planes due to the 2D acquisition protocol, taking into consideration both aliasing due to sampling in the slice direction and blur. In practice, the degradation is achieved through downsampling followed by upsampling and blurring. A supervised deep-learning super-resolution model, such as EDSR [11] is then used to learn the mapping between synthetically degraded LR patches and HR patches.

SMORE operates under a dual objective that combines a pixel-level reconstruction loss (using the ℓ_1 norm) with an edge loss ensuring similarity between image gradients using a Sobel filter.

Once trained, the model is applied to LR orientations (here sagittal and coronal, without loss of generality). Finally, a unique HR isotropic volume is reconstructed by merging volumes upsampled by SSR in the frequency domain through Fourier burst accumulation [12].

3 Method

3.1 Problem statement

Let \mathbf{x}_{LR} be a LR T2 image having thick through-plane spacing along the last dimension, i.e. a resolution of $r_x \times r_y \times r_z$ mm³, with voxel spacing $r_z > r_x$ and where $r_y = r_x$ for simplicity. We hypothesize that \mathbf{x}_{LR} is a downsampled version of an unobserved HR image \mathbf{x}_{HR} of superior resolution $r_x \times r_x \times r_x$

mm³. Our objective is to recover an estimate $\hat{\mathbf{x}}_{\text{HR}} = \mathcal{G}(\mathbf{x}_{\text{LR}})$ from \mathbf{x}_{LR} using a neural network \mathcal{G} , which is a typical single image super resolution (SiSR) problem [13].

3.2 Normalized Gradient Field loss

The original SMORE is trained under a joint ℓ_1 reconstruction and Sobel gradient objective, the latter promoting edge similarity. However, gradient similarity prioritizes high-contrast edges over low-contrast ones (i.e. of weaker gradient magnitude). As a result, some of the finer aspects of the image might be lost or blurred in the enhancement process. To counter this effect, we propose to give equal weight to weak and strong edges through gradient normalization. The idea of normalizing gradients is popular in image registration to obtain contrast-independent similarity metrics and known as the normalized gradient field (NGF) [6].

The NGF loss is defined as :

$$\mathcal{L}_{\text{NGF}} := 1 - \frac{1}{2} \langle \mathbf{n}_\varepsilon(\mathbf{x}_{\text{HR}}), \mathbf{n}_\varepsilon(\mathcal{G}(\mathbf{x}_{\text{LR}})) \rangle^2, \quad (1)$$

where $\langle \cdot, \cdot \rangle$ is the dot product and $\mathbf{n}_\varepsilon(\mathbf{x})$ is the NGF :

$$\mathbf{n}_\varepsilon(\mathbf{x}) = \frac{\nabla \mathbf{x}}{\|\nabla \mathbf{x}\| + \varepsilon}, \quad (2)$$

with ε a noise threshold scalar parameter controlling the influence of noise, with higher values corresponding to higher suppression. The reason for normalization is to encourage gradient vector alignment between generated and target images while being agnostic to contrast.

In this work, we thus propose to use the NGF in the context of image super resolution by training the SSR model under a joint ℓ_1 and a NGF objective (instead of the Sobel loss).

4 Experiments and Results

4.1 Experiments

The evaluation of super resolution can be done either using reconstruction metrics when ground truth HR images are available, or task-based considering a downstream objective (e.g. segmentation, detection, prediction, ...). In this work, we studied to what extent the proposed SSR approach can be beneficial to the training of deep learning-based PCa lesion detection models. In our experiments, we used the public Prostate158 dataset [14] constituted of 158 patient image data, of which we selected 63 images with annotated PCa lesions on axial T2-weighted MR images, showing voxel anisotropy ranging from $0.27 \times 0.27 \times 3$ mm³ to $0.47 \times 0.47 \times 3$ mm³. Images were clipped to the 1%-99% percentile and rescaled in the range [0-1] following common practice. The noise parameter ε was set to 0.1. Note that the annotations were performed on the LR axial images and are thus imperfect ground truth (Fig. 1b).

The SSR method was implemented using PyTorch and Kornia and applied individually to each image of the training and

testing set. The SSR model required approximately one hour per image on a NVIDIA RTX A6000 GPU. Once processed by the various upsampling methods, images were fed to a 3D full resolution nn-UNet model [8] for 1000 epochs using 5-fold cross validation and ensembling. Our baseline comparisons comprised results achieved the original SMORE model (using a joint ℓ_1 and Sobel loss), or by letting the nnU-Net chose the resampling space dynamically using its inner heuristics (resampling to the median resolution of the training set). We used the first 47 images in lexicographic order patients for training and the remaining 16 for testing. Detection evaluation was performed using volumetric overlap between predicted and real lesion masks. A detection was considered positive if the Dice score exceeded 0.2.

4.2 Results

Fig. 2 shows representative upsampling results obtained in one patient of the Prostate158 set, with much better recovery of high frequency details using SSR, and especially when considering the NGF loss with a clearly sharper output. Out of the 16 images tested, 8 were deemed unsuccessful as they recorded a Dice score of less than 0.2 across all methodologies. This can be attributed to the inherent challenges associated with prostate cancer (PCa) segmentation on complex images. Achieving accurate segmentation in these cases often requires the support of additional modalities, typically Apparent Diffusion Coefficient (ADC) MRI, which were not utilized in our study. However, using SMORE allowed for a better detection on average, which translated into better quantitative scores (Table 1). Note one failure case for nnU-Net interpolation and one failure case for baseline SMORE. Our proposed NGF training objective for SMORE allowed better results on average, with a Dice score of around 50. However, these quantitative results are to be taken with caution due to the very limited quality of the lesion segmentation in such anisotropic images. For instance, patient 118 has only two annotated axial slices.

Patient ID	nnU-Net interp.	SMORE	SMORE+NGF
112	18.1	29.0	13.7
118	59.2	71.2	62.2
125	64.5	62.8	34.1
131	72.7	69.2	70.6
133	79.4	80.3	79.9
138	31.7	46.9	59.8
140	30.8	0.0	53.6
148	0.0	9.3	60.0
Mean	44.55	46.09	54.24
Std Dev	26.59	28.25	19.69

Table 1: Dice scores for prostate cancer detection using three different upsampling methods.

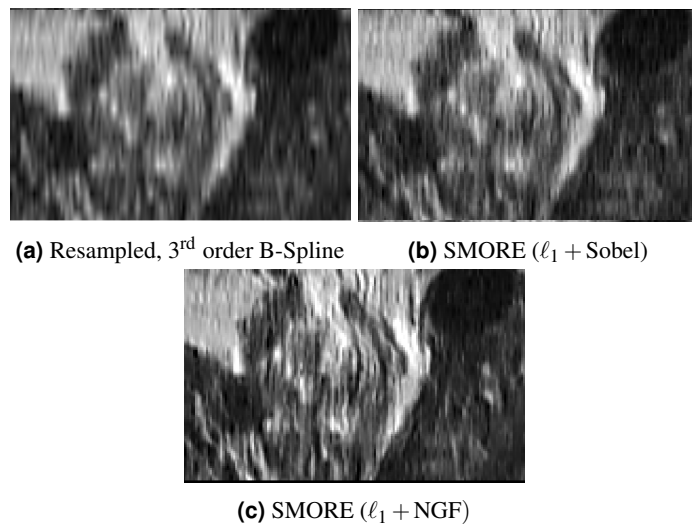


Figure 2: Upsampling result in the Prostate158 dataset. Self Super-Resolution with the proposed NGF objective lead to better detail recovery and sharper outputs.

5 Conclusion

In this preliminary work, we have investigated whether self supervised resolution could be used to improve the detection of PCa images from T2-weighted MR images. Leveraging the SMORE model, we have proposed a new loss function that aims at promoting the alignment of normalized gradient vectors between high resolution and low-resolution patches during training. Preliminary results tend to suggest a positive impact of such an approach, both visually and quantitatively. However, further experiments need to be conducted in a more controlled environment to ensure the interest of the approach, for instance using better annotations or synthetically degraded pseudo low resolution images. Other downstream tasks will also be investigated, such as e.g. segmentation of vascular structure in liver CT images for surgical planning [15].

References

- [1] O. W. Brawley, D. P. Ankerst, and I. M. Thompson. “Screening for prostate cancer”. *CA: a cancer journal for clinicians* 59.4 (2009), pp. 264–273.
- [2] R. J. Cohen, B. A. Shannon, M. Phillips, et al. “Central zone carcinoma of the prostate gland: a distinct tumor type with poor prognostic features”. *The Journal of urology* 179.5 (2008), pp. 1762–1767.
- [3] V. Jaouen, J. Bert, K. A. Mountris, et al. “Prostate volume segmentation in TRUS using hybrid edge-Bhattacharyya active surfaces”. *IEEE Transactions on Biomedical Engineering* 66.4 (2018), pp. 920–933.
- [4] T. Barrett, B. Turkbey, and P. L. Choyke. “PI-RADS version 2: what you need to know”. *Clinical radiology* 70.11 (2015), pp. 1165–1176.
- [5] C. Zhao, B. E. Dewey, D. L. Pham, et al. “SMORE: a self-supervised anti-aliasing and super-resolution algorithm for MRI using deep learning”. *IEEE transactions on medical imaging* 40.3 (2020), pp. 805–817.

- [6] E. Haber and J. Modersitzki. "Intensity gradient based registration and fusion of multi-modal images". *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2006: 9th International Conference, Copenhagen, Denmark, October 1-6, 2006. Proceedings, Part II* 9. Springer. 2006, pp. 726–733.
- [7] O. Ronneberger, P. Fischer, and T. Brox. "U-net: Convolutional networks for biomedical image segmentation". *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III* 18. Springer. 2015, pp. 234–241.
- [8] F. Isensee, P. F. Jaeger, S. A. Kohl, et al. "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation". *Nature methods* 18.2 (2021), pp. 203–211.
- [9] P. F. Jaeger, S. A. Kohl, S. Bickelhaupt, et al. "Retina U-Net: Embarrassingly simple exploitation of segmentation supervision for medical object detection". *Machine Learning for Health Workshop*. PMLR. 2020, pp. 171–183.
- [10] R. Sood and M. Rusu. "Anisotropic super resolution in prostate MRI using super resolution generative adversarial networks". *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. IEEE. 2019, pp. 1688–1691.
- [11] B. Lim, S. Son, H. Kim, et al. "Enhanced deep residual networks for single image super-resolution". *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2017, pp. 136–144.
- [12] M. Delbraccio and G. Sapiro. "Burst deblurring: Removing camera shake through fourier burst accumulation". *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015, pp. 2385–2393.
- [13] W. Yang, X. Zhang, Y. Tian, et al. "Deep learning for single image super-resolution: A brief review". *IEEE Transactions on Multimedia* 21.12 (2019), pp. 3106–3121.
- [14] L. C. Adams, M. R. Makowski, G. Engel, et al. "Prostate158: An expert-annotated 3T MRI dataset and algorithm for prostate cancer detection". *Computers in Biology and Medicine* 148 (2022), p. 105817.
- [15] V. Jaouen, Z. Wang, P.-H. Conze, et al. "Self super resolution for hepatic vessel CT segmentation". *2023 IEEE Nuclear Science Symposium, Medical Imaging Conference and International Symposium on Room-Temperature Semiconductor Detectors (NSS MIC RTSD)*. IEEE. 2023, pp. 1–1.