Learning a sparse database for patch-based medical image segmentation

M. Freiman[1], H. Nickisch[2], H. Schmitt[2], P. Maurovich-Horvat[3], P. Donnelly[4], M. Vembar[5], and L. Goshen[1]

[1] GAT, CT, Philips Healthcare, Haifa, Israel.
[2] Philips Research, Hamburg, Germany
[3] Heart and Vascular Center, Semmelweis University, Budapest, Hungary
[4] South Eastern Health and Social Care Trust, Queen's University, Belfast, Ireland
[5] Clinical Science, CT, Philips Healthcare, Cleveland, Ohio, USA
moti.freiman@philips.com

Abstract. We introduce a functional for the learning of an optimal database for patch-based image segmentation with application to coronary lumen segmentation from coronary computed tomography angiography (CCTA) data. The proposed functional consists of fidelity, sparseness and robustness to small-variations terms and their associated weights. Existing work address database optimization by prototype selection aiming to optimize the database by either adding or removing prototypes according to a set of predefined rules. In contrast, we formulate the database optimization task as an energy minimization problem that can be solved using standard numerical tools. We apply the proposed database optimization functional to the task of optimizing a database for patch-base coronary lumen segmentation. Our experiments using the publicly available MICCAI 2012 coronary lumen segmentation challenge data show that optimizing the database using the proposed approach reduced database size by 96% while maintaining the same level of lumen segmentation accuracy. Moreover, we show that the optimized database yields an improved specificity of CCTA based fractional flow reserve (0.73 vs 0.7 for all lesions and 0.68 vs 0.65 for obstructive lesions) using a training set of 132 (76 obstructive) coronary lesions with invasively measured FFR as the reference.

Keywords: Energy minimization, Prototype sampling, K-nearest neighbor, Coronary Lumen Segmentation.

1 Introduction

Segmentation of anatomical structures from medical images plays an important role in many clinical applications. Automatic segmentation can be challenging due to the large variability in anatomical structures shape and appearance. Patch-based, non-parametric segmentation algorithms such as the K-nearest neighbor (KNN) algorithm [2] have demonstrated their potential in automatic segmentation of challenging anatomical structures. For example, Mechrez et al.

[7] use the KNN algorithm followed by a spatial consistency refinement step to segment Multiple-Sclerosis lesions from MRI data, and Wang et al. [11] demonstrate the potential of KNN algorithm in defining a search-space for improved patch-based segmentation of cardiac MR data and abdominal CT data. Specifically in the cardiovascular domain, Olabarriaga et al [9] used the KNN algorithm to steer a model-based segmentation of abdominal aortic aneurysms and more recently, Freiman et al. [3] used the KNN algorithm to estimate the likelihood component within a graph min-cut framework for coronary artery lumen segmentation.

While the KNN algorithm has several theoretical and practical advantages [2], two main limitations of this algorithm are: 1) large storage to retain the set of examples which defines the training set, and 2) low efficiency due to the re-calculation of the similarity between the test and training samples at each evaluation [4].

Among the approaches previously proposed to address these issues, database optimization by prototype selection is an attractive approach as it maintains originally an-notated data rather than generating artificial data. The optimal prototype selection is an NP-hard problem which can be mapped onto a set-cover problem and solved using an approximation algorithm [1]. Alternatively, the prototype selection problem can be relaxed by introducing some order on the prototypes. Then, the prototype selection approaches can be divided into three categories: 1) incremental search, in which the algorithm adds prototypes to the reduced data based on some rule, 2) decremental search in which the algorithm aims to remove prototypes from the database according to some rule, and; 3) hybrid approach which combines both incremental and decremental steps. For a comprehensive review of methods aimed to reduce KNN algorithm storage and computational demand we refer the reader to García et al [4].

In this work we formulate the database optimization problem as an energy minimization which enables the optimization using common numerical approaches. Our functional consists of fidelity, sparsity and robustness to small variations terms along with their associated weights. We applied the proposed functional to optimize a database used in Freiman's et al patch-based coronary artery lumen segmentation algorithm [3]. We evaluated the influence of database optimization on the segmentation performance by means of segmentation accuracy using the publicly available MICCAI 2012 Coronary Lumen Segmentation Challenge Database [5]. We also evaluated the impact of the database optimization on CCTA based fractional flow reserve (CT-FFR) estimates using a database of 132 coronary lesions with invasively measured FFR as the reference.

Our results show that the database size can be reduced by 96% while maintaining the same level of coronary lumen segmentation accuracy on the MICCAI 2012 Coronary lumen segmentation challenge database [5] and even improving the performance of CT-FFR estimates obtained with 3D models generated from the automatic segmentation results.

2   Method

Our goal is to select a subset of prototypes from a given database so that the classification performance for any new sample will be as accurate as possible. The sub-sampled database should represent the full structure of the population with as few prototypes as possible. First, we define a property describing the distribution of the prototypes in the original database. Next, we define a set of parameters used to generate a sub-sampled database, and finally, we formulate the optimal parameter finding as an energy minimization problem.

2.1   Prototype ranking

Inspired by the work of Bein and Tibshirani [1], we rank prototypes in the original database according to their location on the manifold. Specifically, we will consider a prototype as located in the center of its class when its neighboring most similar prototypes according to some pre-defined metric are from of the same class and as located on the boundary between classes if it has many neighbors similar samples which are belonging to other from different classes. Formally, for a prototype feature-vector $\boldsymbol{x}$, we define the sample ranking score $R(\boldsymbol{x})$ as follows:

$$R(\boldsymbol{x}) = \frac{\#\text{NN with other class}}{\#\text{NN with same class}} = \frac{\sum_{k=1}^{K} 1 - \delta(C[\boldsymbol{x}], C[\boldsymbol{x_k}])}{\max(\sum_{k=1}^{K} \delta(C[\boldsymbol{x}], C[\boldsymbol{x_k}]), 1)} \qquad (1)$$

where K is the number of the nearest prototype neighbors $\boldsymbol{x_i}$ that are similar according to the chosen distance metric, $C[\boldsymbol{x_i}]$ is the class of $\boldsymbol{x_i}$, $\boldsymbol{x}$ is a prototype similar to $\boldsymbol{x_i}$, and:

$$\delta(C[\boldsymbol{x_i}], C[\boldsymbol{x_k}]) = \begin{cases} 1, & C[\boldsymbol{x_i}] = C[\boldsymbol{x_k}] \\ 0, & C[\boldsymbol{x_i}] \neq C[\boldsymbol{x_k}] \end{cases} \qquad (2)$$

According to this definition $R(\boldsymbol{x})$ gets a high value when the prototype $\boldsymbol{x}$ has many similar prototypes from other classes, and a low value when the prototype $\boldsymbol{x}$ has many similar prototypes from its own class.

2.2   Database sparsification

We describe the distribution of the classification sample ranking scores of the prototypes in the original database for each class using a histogram with N bins. The bins boundaries are set (and kept fixed) to be percentiles of the overall samples per class to normalize against various sample ranks distributions. We define a sparsified database as a function of the percentile of prototypes to be selected from each bin of the histogram $DB(\boldsymbol{N})$, where $\boldsymbol{N} \in \mathbb{N}^N$ is the vector containing the number of prototypes to be put in each of the N bins. The subsampling is done deterministically by selecting the N_i prototypes with highest ranking score for bin i.

2.3 Database optimization

Given the function $DB(\boldsymbol{N})$ to sample the original database, we define a functional to estimate the sampling parameters $\boldsymbol{N}$ i.e. the number of prototypes per bin. Our goal is to find $\boldsymbol{N}$ that maximizes the capability of the sampled database to correctly classify each sample while minimizing the overall number of samples. We also would like the classification to be robust to small variations in the samples. Formally, our functional is defined as:

$$E_{\alpha,\beta}(\boldsymbol{N}) = \sum_{i=1}^{M} \overbrace{\rho(\boldsymbol{N},\boldsymbol{x_i})}^{\text{robustness}} + \alpha \sum_{i=1}^{M} \overbrace{\left(C(\boldsymbol{x_i}) - f\left(DB(\boldsymbol{N}),\boldsymbol{x_i}\right)\right)^2}^{\text{fidelity}} + \beta \overbrace{\|DB(\boldsymbol{N})\|}^{\text{sparsity}} \quad (3)$$

where: $DB(\boldsymbol{N})$ is the sampled database constructed from the original database by sampling the different bins according to the percentiles specified by $\boldsymbol{N}$, $f\left(DB(\boldsymbol{N}),\boldsymbol{x_i}\right)$ is the classification of the prototype $\boldsymbol{x_i}$ using the sampled database $DB(\boldsymbol{N})$, $\|DB(\boldsymbol{N})\|$ is the number of the prototype in the sampled database $DB(\boldsymbol{N})$, M is the number of prototypes in the original database, α,β are weighting meta-parameters controlling the contribution of each term, and $TV(\boldsymbol{N})$ measures the robustness of the the classification of each prototype using $DB(\boldsymbol{N})$ to a small variation in its appearance as follows:

$$\rho(\boldsymbol{N},\boldsymbol{x}) = \sum_{j=1}^{J} \|f\left(DB(\boldsymbol{N}),\boldsymbol{x}\right) - f\left(DB(\boldsymbol{N}),\boldsymbol{x}+\boldsymbol{e_j}\right)\|_1 \quad (4)$$

where J is the dimension of the prototype $\boldsymbol{x}$, the unit vector $\boldsymbol{e_j}$ is of the same dimension as with zeros at all entries except at entry j, and $\|f\left(DB(\boldsymbol{N}),\boldsymbol{x}\right) - f\left(DB(\boldsymbol{N}),\boldsymbol{x}+\boldsymbol{e_j}\right)\|_1$ is the absolute difference between the classification of $\boldsymbol{x}$ and $\boldsymbol{x}+\boldsymbol{e_j}$ given the database $DB(\boldsymbol{N})$.

We find the optimal database sampling parameters by minimizing the energy functional:

$$\hat{\boldsymbol{N}} = \operatorname*{argmin}_{\boldsymbol{N}} E_{\alpha,\beta}(\boldsymbol{N}) \quad (5)$$

2.4 Coronary Lumen Segmentation

We apply the proposed approach to reduce the database size required for the coronary lumen segmentation algorithm proposed by Freiman et al. [3]. For the sake of completeness, we briefly describe the relevant parts of the algorithm here. We refer the interested reader to a detailed and complete description provided in [3]. The algorithm formulates the segmentation task as an energy minimization problem over a cylindrical coordinate system [6] where the warped volume along the coronary artery centerline is expressed with the coordinate i representing the index of the cross-sectional plane, and θ, r represent the angle and the radial distance determining a point in the cross-sectional plane

$$E(X) = \sum_{p \in P} \psi_p(x_p) + \lambda \sum_{p,q \in E} \varphi_{p,q}(x_p, x_q) \quad (6)$$

where P is the set of sampled points, x_p is a vertex in the graph representing the point $(i^{x_p}, \theta^{x_p}, r^{x_p})$ sampled from the original CCTA volume, $\psi_p(x_p)$ represents the likelihood of the vertex to belong to the lumen or the background class, p, q are neighboring vertices according to the employed neighboring system E, and $\varphi_{p,q}(x_p, x_q)$ is a penalty for neighboring vertices belonging to different classes ensure the smoothness of the resulted surface. The algorithm uses the KNN algorithm [2] to calculate the likelihood of each vertex x_p belonging to the coronary lumen from a large training database with rays sampled from cardiac CTA data along with manually edited lumen boundary location represented as a binary rays serving as the database prototypes. The likelihood term is additionally adjusted to account for partial volume effects and a $L2$ norm used as the regularization term as described in [3].

2.5   Application of the database optimization

The original database used in the coronary lumen segmentation algorithm consists of ~2,130,000 prototypes obtained from 97 CCTA datasets segmented manually by a cardiac CT expert. We consider the lumen radii as the different classes of the rays. We use the L2 norm as the distance metric to define the sample rank (Eq. 1), and experimentally set the number of bins in the histogram N (Eq. 3) to 5. We optimize the functional hyperparameters to achieve the maximal area under the curve (AUC) for CT-FFR estimates with the segmentations obtained using the optimized database with invasive FFR measurements as the reference. Formally, we define a two-phase optimization task, where the outer loop optimized the model hyperparameters α and β:

$$\widehat{\alpha, \beta} = \underset{\alpha, \beta}{\arg\max} \, AUC \left(FFR_{CT} \left(DB \left(\boldsymbol{N} \right) \right) - FFR_{GT} \right) \tag{7}$$

and the inner loop find the optimal model parameters for given α, β using Eq. 5. We carried out the optimization using the derivative-free BOBYQA algorithm by Powel et al. [10].

3   Experimental results

We use two data sets as follows. The first dataset consists of CCTA data of 132 coronary lesions that were retrospectively collected from the medical records of 97 subjects who underwent a CCTA and invasive coronary angiography with invasive FFR measurements due to suspected CAD. CCTA data was acquired using either a Philips Brilliance iCT (gantry rotation time of 0.27 sec.) or Philips Brilliance 64 (gantry rotation time of 0.42 sec.). Acquisition mode was either helical retrospective ECG gating (N=54) or prospectively ECG triggered axial scan (N=43). The kVp range was 80 – 140 kVp and the tube output range was 600 – 1000 mAs for the helical retrospective scans and 200 – 300 mAs for the prospectively ECG triggered scans.

Table 1: Summary statistics of hemodynamic significance assessment of coronary lesions by means of CT-FRR based on automatic segmentation with the entire database and with the optimized database. Results presented for all (N=132) lesions and specifically for obstructive (Obst.) lesions (N=76) separately.

	Sensitivity		Specificity		Accuracy		AUC	
	All	Obst.	All	Obst.	All	Obst.	All	Obst.
Optimized database	0.85	0.86	0.73	0.68	0.77	0.76	0.84	0.83
Full database	0.85	0.86	0.70	0.65	0.76	0.75	0.84	0.82

Cross-sectional area (CSA) based stenosis quantification was performed by an expert reader on 132 lesions, out of which 56 were diagnosed as non-obstructive lesions (CSA stenosis less than 50%) and 76 diagnosed as obstructive lesions (CSA stenosis 50%-90%). According to the invasive FFR measurements, 48 lesions were hemodynamically significant (FFR$\leq$0.8) and 84 lesions were non-significant (FFR$>$0.8). The coronary artery centerlines and the aorta segmentation were computed automatically and adjusted manually by a cardiac CT expert (–) to account for algorithm inaccuracies using a commercially available software dedicated to cardiac image analysis (Comprehensive Cardiac Analysis, IntelliSpace Portal 6.0, Philips Healthcare).

We tuned the model hyper-parameters and optimized the database using Eq. 5 and 7. Next, we segmented the coronary tree with Freiman's et al algorithm [3] with the full and the optimized databases. Finally, we performed the flow simulations using the lumped parameter model (LM) proposed by Nickisch et al. [8].

We compared the flow simulation results from the 3D coronary tree models generated using the coronary lumen segmentation algorithm described in section 2.4, with the full and optimized training databases. The optimization process reduced the database size by 96% from $\sim$2,130,000 prototypes to $\sim$84,000 prototypes. Fig. 1 illustrates the reduction in the number of prototypes in the database at the different bins of the prototypes histogram. The Functional-based optimization prefers to keep prototypes with low sample ratio (i.e closer to the center of mass of each class) in the optimized database. Table 1 summarizes the performance metrics for assessing the hemodynamic significance of coronary lesions with automatic segmentation using the entire and optimized database for entire set of coronary lesions and specifically for obstructive lesions (Cross Sectional Area (CSA) stenosis$\geq$50%). The flow simulation results are slightly better using the optimized database compared to the results of the full database, although the optimized database has much less prototypes compared to the full database.

Fig. 2 depicts representative examples of straight multi-planar reconstructed images of coronary artery segmentation results with the optimized (green) and the full (red) database. Table 2 presents the segmentation accuracy results of
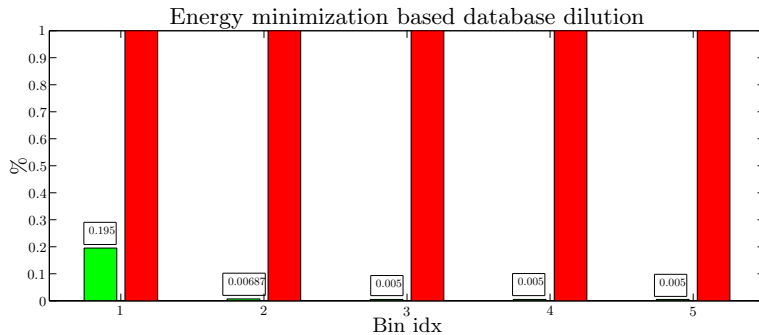
Fig. 1: The database reduction for each bin of the histogram of prototypes. The red bars indicate 100% of the original database and the green bars indicates the percentage of the remaining prototypes after the optimization. Optimized database sampling parameter for each bin is listed above the optimized histogram.
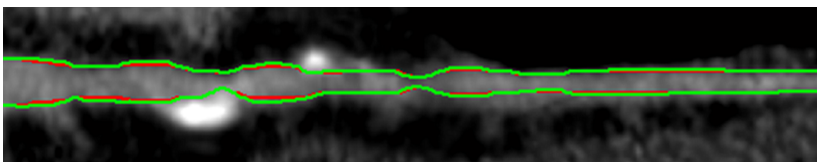

Fig. 2: Representative example of straight multi-planar reconstructed images of coronary artery segmentation results with the optimized (green) and full database.

Table 2: Summary statistics of coronary lumen segmentation accuracy using the MICCAI 2012 challenge evaluation framework [5] for the training datasets (18 cases, 78 coronary segments). Results presented for healthy and diseased segments separately and in the relevant metric units.

	Dice (%)		MSD (mm)		MAX SD (mm)	
	Healthy	Disease	Healthy	Disease	Healthy	Disease
Optimized database	0.69	0.74	0.49	0.27	1.69	1.24
Full database	0.69	0.74	0.49	0.27	1.69	1.22

our algorithm with the original and optimized database using the MICCAI 2012 challenge framework training data [5]. We refer the reader to the challenge website [5] for further comparison with the rest of the methods and with the observer performance.

4 Conclusion

We presented an energy functional for optimizing the training database in patch-based medical image segmentation algorithms. We define a 'sample rank' order on the training database prototypes and formulate the prototype sampling as an energy minimization task with hyper-parameters that can be adjusted to the specific task. We demonstrated the application of this approach to reducing database size and improving the performance of coronary lumen segmentation algorithm from CCTA data. Our experiments show that the optimized database can maintain overall segmentation results with added incremental improvements of CT-FFR estimates based on the 3D models generated from the segmentation results while substantially reducing the memory demand of the algorithm.

References

1. Bien, J., Tibshirani, R.: Prototype selection for interpretable classification. Annals of Applied Statistics 5(4), 2403–2424 (2011)
2. Cover, T., Hart, P.: Nearest neighbor pattern classification. IEEE Trans. Inform. Theory 13(1), 21–27 (1967)
3. Freiman, M., et al.: Improving CCTA-based lesions' hemodynamic significance assessment by accounting for partial volume modeling in automatic coronary lumen segmentation. Med Phys 44(3), 1040–1049 (2017)
4. García, S., et al.: Prototype Selection for Nearest Neighbor Classification: Taxonomy and Empirical Study. IEEE Trans Patt Anal Mach Intell 34(3), 417–435 (2012)
5. Kirişli, H., et al.: Standardized evaluation framework for evaluating coronary artery stenosis detection, stenosis quantification and lumen segmentation algorithms in computed tomography angiography. Med Image Anal 17(8), 859–876 (2013), `http://coronary.bigr.nl/stenoses/`
6. Lugauer, F., et al.: Precise Lumen Segmentation in Coronary Computed Tomography Angiography. In: Menze, B., et al. (eds.) Proc. of MICCAI 2014 Workshop on Medical Computer Vision: Algorithms for Big Data, Lecture Notes in Computer Science, vol. 8848, pp. 137–147. Springer (2014)
7. Mechrez, R., Goldberger, J., Greenspan, H.: Patch-Based Segmentation with Spatial Consistency: Application to MS Lesions in Brain MRI. Int J Biomed Imaging 2016, Article ID 7952541 (2016)
8. Nickisch, H., et al.: Learning Patient-Specific Lumped Models for Interactive Coronary Blood Flow Simulations. In: Navab, N., et al. (eds.) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Lecture Notes in Computer Science, vol. 9350, pp. 433–441. Springer International Publishing (2015)
9. Olabarriaga, S.D., et al.: Segmentation of thrombus in abdominal aortic aneurysms from CTA with nonparametric statistical grey level appearance modeling. IEEE Trans Med Imaging 24(4), 477–485 (2005)
10. Powell, M.: The BOBYQA algorithm for bound constrained optimization without derivatives. NA Report NA2009/06 p. 39 (2009), `http://www6.cityu.edu.hk/rcms/publications/preprint26.pdf`
11. Wang, Z., et al.: Geodesic patch-based segmentation. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). vol. 8673 LNCS, pp. 666–673 (2014)