Dense Iterative Contextual Pixel Classification using Kriging

Melanie Ganz
DIKU, University of Copenhagen
Universitetsparken 1, Cph, Denmark
ganz@diku.dk

Marco Loog
Delft University of Technology, Faculty of EEMCS
Mekelweg 4, 2628 CD Delft, The Netherlands
M.Loog@tudelft.nl

Sami Brandt
Nordic Bioscience A/S
Herlev Hovedgade 207, Herlev, Denmark
sbrandt@nordicbioscience.com

Mads Nielsen
DIKU, University of Copenhagen
Universitetsparken 1, Cph, Denmark
madsn@diku.dk

Abstract

In medical applications, segmentation has become an ever more important task. One of the competitive schemes to perform such segmentation is by means of pixel classification. Simple pixel-based classification schemes can be improved by incorporating contextual label information. Various methods have been proposed to this end, e.g., iterative contextual pixel classification, iterated conditional modes, and other approaches related to Markov random fields. A problem of these methods, however, is their computational complexity, especially when dealing with high-resolution images in which relatively long range interactions may play a role. We propose a new method based on Kriging that makes it possible to include such long range interactions, while keeping the computations manageable when dealing with large medical images.

1. Introduction

In medical applications, image segmentation tasks become ever more important to aid quantitative analysis. In this paper, we focus on the application of medical imaging to aid the diagnosis and prognosis of cardiovascular diseases (CVD) [7]. Images are segmented by e.g. active contour [18], active appearance [8] or level set models [20]. Competitors to these models, especially in the domain of medical imaging, are pixel-wise classifiers [23]. To achieve even better segmentations, in cooperation with pixel-wise classification, various methods have been developed that take contextual information in the images into account. Examples are iterated conditional modes (ICM) [6], iterative contextual pixel classification (ICPC) [17] or Markov Random Fields (MRF) [10].

A pure pixel-wise classification takes only the contributions of a neighborhood in the image into account when assigning the class label to a pixel. It does not consider any other class labels in its decision. The technique of ICM on the contrary employs contextual knowledge. It is divided in two steps: First, the pixels are classified by a pixel-wise classifier. Second, the neighboring class labels are included into a label decision. An advanced version of ICM was presented by Loog and van Ginneken. Their ICPC method went a step further than ICM by creating a simultaneous dependency of a class label on surrounding image values and class labels.

A problem of these methods, however, lies in their computational burden when dealing with high resolution images like medical data. Additionally, if also long range interactions are present in the image, the convergence of the methods often becomes an issue.

We propose a new model that can be applied after a pixel-wise classification, Dense Iterative Contextual Pixel Classification (DICPC). It employs the context of all class labels and can take long range interactions into account. We implement this by approximating the contextual interactions in label space with a linear model based on Kriging [15]. With this approximation it becomes feasible to converge to an optimal segmentation in manageable time, even for high-resolution images with a long interaction ranges.

The paper is organized as follows: Sect. 2 restates the problem and gives an introduction to a statistical interpretation of segmentation and Kriging. Sect. 3 introduces the DICPC algorithm. Sect. 4 introduces the problem on which we exemplify our method. It concerns the difficult task of quantifying aortic calcifications. Information on the study population and the exact classification settings are presented there as well as evaluation methods. The results can also be found in Sect.4 while Sect.5 comprises the discussion and

978-1-4244-3993-5/09/$25.00 ©2009 IEEE

conclusion.

2. Problem Description

Let an image I be described by its pixels $\vec{p} = (p_1, \ldots, p_i)$. In a pixel classification scheme, there exists a feature vector $\vec{f}_i$ for each pixel p_i that consists of one or multiple features, e.g. intensity values or filter responses, at the pixel location i. The matrix $F = (\vec{f}_1, \ldots, \vec{f}_i)$ is comprised of all the feature vectors. The labels for every pixel i also shape a vector, $\vec{c} = (c_1, \ldots, c_i)$, that consists of the class label at each pixel location. Class labels are part of the set $\Gamma = (1, 2, \ldots, \gamma)$, where γ is finite.

The problem lies now in finding the optimal segmentation $C^\star$ for the image I.

2.1. Statistical Interpretation

One can view the segmentation as a statistical process. To find the optimal segmentation $C^\star$ one might pursue a Bayesian minimal risk approach [4]. For the risk function of all wrong classifications being equally risky this becomes the maximum-a-posteriori (MAP) method, which we apply in the following. In a MAP estimation the optimal segmentation is given as

$$C^\star = \underset{C \in \mathcal{C}}{\operatorname{argmax}}\, P(C|I) = \underset{C \in \mathcal{C}}{\operatorname{argmax}}\, P(I|C)P(C), \quad (1)$$

where $\mathcal{C}$ is the set of all possible segmentations.

There exist different approaches to solve Eq. 1. Assuming the Hammersley-Clifford theorem [5] holds in the label space this can be transformed into

$$\begin{aligned} C^\star &= \underset{C \in \mathcal{C}}{\operatorname{argmax}} \prod_i P(C_i|C_{N_i}, I) \\ &= \underset{C \in \mathcal{C}}{\operatorname{argmax}} \prod_i P(I|C_i, C_{N_i})P(C_i|C_{N_i}), \end{aligned} \quad (2)$$

where C_i denotes the label for a pixel i and C_{N_i} the neighborhood labels of a pixel i. This equation can now be taken as the starting point to describe the different techniques mentioned before in a more formal framework.

The technique of ICM is, in addition to Eq. 2, based on assuming that the observation components I_i are conditionally independent given C and that each I_i has the same known conditional density function $p(I_i|C_i)$ dependent only on C_i. This leads to

$$C^\star = \underset{C \in \mathcal{C}}{\operatorname{argmax}} \prod_i P(C_i|C_{N_i})P(I|C_i), \quad (3)$$

where $P(I|C_i)$ is solved by a pixel-wise classification.

On the contrary, ICPC turns Eq. 2 into

$$C^\star = \underset{C \in \mathcal{C}}{\operatorname{argmax}} \prod_i P(C_i|C_{N_i}, I_{N_i}) \quad (4)$$

by assuming that the Hammersley-Clifford theorem holds in the image space as well. Then it combines image feature and label space and maximizes $P(C_i|C_{N_i}, I_{N_i})$ in an iterative manner. Here, I_{N_i} is the image neighborhood of pixel i.

But both ICM and ICPC are computationally unfeasible, when we deal with high-resolution images. Each iteration of ICM or ICPC is computationally intensive. Here, our proposed method, DICPC, can be seen as an alternative to ICM and ICPC. We also assume that the Hammersley-Clifford theorem is valid and that the contributions of the image and the labels can be considered independently leading to Eq. 3. First, we use a pixel-wise classifier to maximize $P(C_i, I)$. Then instead of using a complete neighborhood model in Eq. 3, we approximate the neigborhood interactions by a linear model given by Kriging and solve it. This speeds our calculations up tremendously, since most of the calculations can be formulated in a filtering formulation, which becomes apparent in the next subsection.

2.2. Kriging

Kriging [13, 11] is a geostatistical method that is used to spatially interpolate the value z_0 at any location $\vec{r}_0$ from irregularly sampled data $\vec{z}$ at N points $\vec{r}_N$. This is done by the local affine model $z_0 = z(\vec{r}_0)$ such that

$$z_0 = w_0 + \sum_{1=1}^{N} w_i z_i = w_0 + \vec{w}^t \vec{z}, \quad (5)$$

where w_0 is an offset and w_i is the weight applied to z_i. If we regard the z_i as realizations of random variables Z_i and request our measure to be unbiased, $E(Z_0 - \hat{Z}_0) = 0$, we can define the estimation variance

$$\sigma_{\mathrm{E}}^2 = \mathrm{Var}(Z_0 - \hat{Z}_0). \quad (6)$$

Using the linear model we can simplify this to

$$\begin{aligned} \sigma_{\mathrm{E}}^2 &= \mathrm{Var}(Z_0) + \mathrm{Var}(w_0 + \vec{w}^t \vec{Z}) - 2\mathrm{Cov}(Z_0, w_0 + \vec{w}^t \vec{Z}) \\ &= \sigma^2 + \vec{w}^t \mathbf{C} \vec{w} - 2\vec{w}^t \mathrm{Cov}(Z_0, \vec{Z}), \end{aligned} \quad (7)$$

where $\mathbf{C}$ is the variance/covariance matrix of $\vec{Z}$. $\mathrm{Cov}(Z_0, \vec{Z})$ is a column vector of covariances between data points a locations $\vec{r}_i$ and $\vec{r}_j$ that can be calculated based on the assumption of spatial stationarity from the entries in $\mathbf{C}$. We minimize the estimation variance after the weights w_i by solving

$$\frac{\partial \sigma_{\mathrm{E}}^2}{\partial \vec{w}} = 2\mathrm{Cov}(\vec{Z}, \vec{Z})\vec{w} - 2\mathrm{Cov}(Z_0, \vec{Z}) = 0. \quad (8)$$

This results in the simple Kriging system

$$\mathrm{Cov}(\vec{Z}, \vec{Z})\vec{w} = 2\mathrm{Cov}(Z_0, \vec{Z}), \quad (9)$$

which can be solved for the interpolation weights $\vec{w}$ and is in our case expanded to include the necessary condition $0 \leq w_i \leq 1$ in order to avoid negative weights.

2.3. Application of Kriging in DICPC

The same principle of Kriging can with regularly distributed samples be applied to an image and has been used for image restoration [21]. Kriging is as stated in Eq. 5 based on a linear estimation model. In the case of a segmentation task where manual segmentations are available, one can learn the weights that minimize the estimation variance, σ_E^2, from the manual segmentations S_{man} via the linear model

$$z_{0,\text{man}} = w_{0,\text{man}} + \vec{w}_{\text{man}}^t \vec{z}_{\text{man}}. \tag{10}$$

Then we use these weights to compose a linear model for the automated segmentations

$$z_{0,\text{aut}} = w_{0,\text{man}} + \vec{w}_{\text{man}}^t \vec{z}_{\text{aut}}. \tag{11}$$

This is possible because we may assume that the covariance structure of the manual segmentations can be transferred to the automated segmentation. The weights of the linear model can then be applied to the automated segmentation S_{aut} in a filtering manner to give a kriged estimate of the segmentation

$$K(S_{\text{aut}}) = k * S_{\text{aut}}, \tag{12}$$

where k is a 2D-filter built from the weights $\vec{w}_{\text{man}}$. Because our method is based on this type of filtering the computational cost stays low compared to ICM and ICPC. Using this formulation of simple Kriging, we now turn to our application in pixel-based segmentation.

3. DICPC - Dense Iterative Contextual Pixel Classification

The solution for an optimal contextual segmentation has the form of Eq. 2,

$$C^\star = \underset{C \in \mathcal{C}}{\text{argmax}} \prod_i P(C_i|C_{N_i}, I). \tag{13}$$

Using Bayes formula this can be transformed into

$$\begin{aligned} C^\star &= \underset{C \in \mathcal{C}}{\text{argmax}} \prod_i P(I|C_i, C_{N_i}) P(C_i|C_{N_i}) \\ &= \underset{C \in \mathcal{C}}{\text{argmin}} \sum_i \Big(-\log(P(I|C_i, C_{N_i})) - \log(P(C_i|C_{N_i})) \Big). \end{aligned} \tag{14}$$

If we, just as in ICM, assume independence of C_{N_i} and I, we can write this as

$$C^\star = \underset{C \in \mathcal{C}}{\text{argmin}} \sum_i \Big(-\log(P(I|C_i)) - \log(P(C_i|C_{N_i})) \Big). \tag{15}$$

3.1. Gaussian Distributions

To solve Eq. 15 we need to define $P(I|C_i)$ and $P(C_i|C_{N_i})$. For now we assume that both are Gaussian distributed, but we will relax this assumption later. A probability prior $P(C_i|C_{N_i})$ for the segmentation C can be formulated as follows

$$\begin{aligned} P(C_i|C_{N_i}) &= G_\sigma(C_i - K(C_i)) \\ &= \frac{1}{\sqrt{2\pi}\sigma_i} \exp\Big(-\frac{(C_i - K(C_i))^2}{2\sigma_i^2} \Big) \end{aligned} \tag{16}$$

Furthermore, we define $P(I_i|C_i)$ to be of the form

$$\begin{aligned} P(I_i|C_i) &= G_\sigma(I_i - C_i) \\ &= \frac{1}{\sqrt{2\pi}\sigma} \exp\Big(-\frac{(I_i - C_i)^2}{2\sigma^2} \Big). \end{aligned} \tag{17}$$

In the case of a Gaussian prior $P(C_i|C_{N_i})$ and a Gaussian likelihood $P(I_i|C_i)$, the posterior distribution is Gaussian again and in that case a closed form solution exists for Eq. 15. Plugging in $P(I_i|C_i)$ and $P(C_i|C_{N_i})$ into Eq. 17 leads to

$$E_{\text{G}} = \sum_i \left(a + \frac{(I_i - C_i)^2}{2\sigma^2} + \frac{(C_i - K(C_i))^2}{2\sigma_i^2} \right), \tag{18}$$

where $a = \log(2\pi\sigma\sigma_i)$. We replace $K(C_i)$ with $k * C_i$ according to the definition of Eq. 12 and get

$$E_{\text{G}} = \sum_i \left(a + \frac{(I_i - C_i)^2}{2\sigma^2} + \frac{(C_i - (k * C_i))^2}{2\sigma_i^2} \right). \tag{19}$$

Due to the Parseval theorem the energy is preserved in the Fourier transform, so the cost can be equivalently computed in the Fourier domain. The minimization in the Fourier domain is performed by differentiating the energy functional with respect to the real and imaginary parts of the Fourier coefficients and setting the result zero. This way we arrive at

$$C^\star = \mathcal{F}^{-1} \left(\frac{\tilde{I}_i}{1 + \varpi(\tilde{k'})^\star(\tilde{k'})} \right), \tag{20}$$

where ϖ is the ratio between the neighborhood and the global variance, $\mathcal{F}$ describes the Fourier transform and $\tilde{I} = \mathcal{F}(I)$ as well as $\tilde{k'} = \mathcal{F}(k') = \mathcal{F}(k-1)$.

3.2. Other Distributions

If $P(I_i|C_i)$ and $P(C_i|C_{N_i})$ are not Gaussian, but e.g. Laplace distributed, the energy function of Eq. 18 changes to

$$E_{\text{L}} = \sum_i \left(a + \frac{|I_i - C_i|}{2\sigma^2} + \frac{|C_i - K(C_i)|}{2\sigma_i^2} \right), \tag{21}$$

A solution to Eq. 21 can be found via variational methods by any approach for total variation minimization. In the case of distributions where the mode of the distribution is equal its mean this solution is equal to the MAP solution. Thus, for the Laplacian distribution the MAP solution is identical to the minimum variance solution. In the case of other distributions, it depends on the distribution if the total variation minimization equals the MAP solution.

3.3. DICPC algorithm

All in all, the contextual classification approach that is proposed looks as follows:

1. Learn the weights $\vec{w}_{man}$ from manual segmentations and construct a 2D-filter k.
2. Define a distribution to be used in the prior $P(C_i|C_{N_i})$.
3. Define a distribution to be used in the conditional probability $P(I_i|C_{N_i})$.
4. Use a direct or a variational approach to solve Eq. 15.

4. Experiments and Results

Cardiovascular diseases (CVD) are the most common cause of death in Europe [9] and the United States [1]. This is the case despite the general knowledge that a healthy lifestyle and the treatment of risk factors can prevent the development of CVD [19]. It is known that abdominal aortic calcifications (AAC) are strong predictors of cardiovascular morbidity and mortality [24]. They correlate strongly with coronary artery calcifications and can hence predict the risk of coronary artery problems [25]. The early detection of AAC helps to predict the risk of related coronary diseases and provides the possibility of intervention. Manual detection of AAC is tedious and costly, which gives rise to a great need for reliable automatic detection schemes.

4.1. Study population

The study population used in this paper is the EPI Followup population, which was part of the multi-centered PERF Study [3]. The EPI part of the PERF study was an epidemiological study addressing the role of a number of metabolic risk factors in the pathogenesis of CVD and osteoporosis [22] and carried out in Ballerup, Denmark, in 1992 and 2001. Three trained radiologists, unaware of the patients conditions, annotated the vertebrae, the aorta and the calcifications in the digitized images. They used Sectra radiological reading units and annotation software implemented in MatLab (Mathworks, MA, USA).

Five images where randomly selected and 18 patches with one or multiple calcifications as illustrated in Fig. 1 were acquired. The testing of the different methods was performed on these 18 calcification patches.

4.2. Evaluation Methods

To measure the compliance of an annotation A_1 and the outcome of a classification A_2, first the Jaccard index [12] is used. This is a quantification measure for segmented areas. The ratio of the number of pixels present in both segmentations to the total number of pixels in the segmentations is taken. The Jaccard index varies from 0 which equals no overlap to 1 corresponding to complete overlap.

$$r_{\text{Jacc}} = \frac{|A_1 \cap A_2|}{|A_1 \cup A_2|} \tag{22}$$

Furthermore we use sensitivity r_{Sens} and specificity r_{Spec} to evaluate the performance of our contextual classification methods.

These measurements require the images to have the same resolution. Furthermore, the errors of r_{Jacc}, r_{Sens} and r_{Spec} increase with smaller objects.

4.3. Classification Settings

We use a k-Nearest-Neighbor classifier [2] with $k = 25$. Training pixels are selected with a bias toward calcified pixels, meaning that 30% of the training pixels chosen are calcified pixels which are defined according to the manual annotations. The features used were the intensity, the gradient magnitude, the Hessian trace, Hessian determinant and Hessian eigenvalues, an adaption of Koenderink's shape classification measure [14] and the ratio of the difference and sum of the Hessian eigenvalues. All features were calculated at three different scales, corresponding to 2, 5 and 17 mm.

4.4. Kriging Settings

In our method, we learn the Kriging filter for calcifications from manual segmentations by calculating the weights for a seven by seven neighborhood in which we krig to the central pixel. This way we arrive at a seven by seven filter that is used as k in Eq. 12. In principle one can use any size neighborhood, since the filtering process is computationally very fast even for large filters.

4.5. Comparison to other methods

To investigate the performance of our method we compare it to other post-processing methods used for the segmentation of lumbar aortic calcifications.

4.5.1 Disk Morphology

The morphological operations used are an opening and a closing with a disk of the size of 1 mm. We chose 1 mm in order to remove pixel noise, but not parts of calcifications.

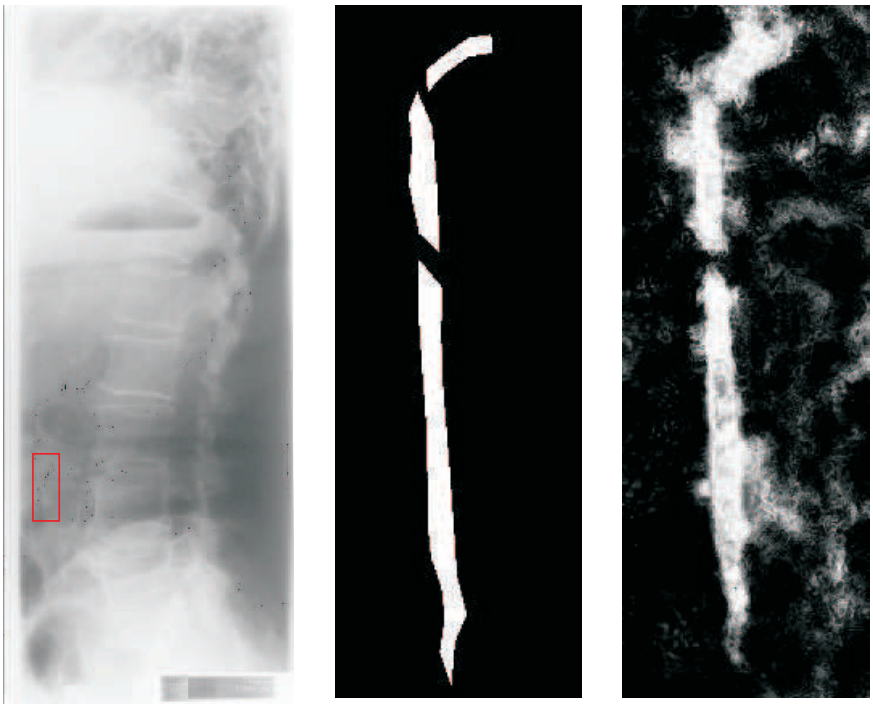

Figure 1. The graph displays one of the original images and a manual as well as an automatic segmentation of a calcification.

4.5.2 Biological Morphology

The morphological operations used are an opening and a closing with structuring elements derived from biological findings. Larsen et al. showed in [16] that the size of lumbar aortic calcifications is approx. 5 ± 3 mm in height and 2 ± 1 mm in width. According to these findings the first structuring element was designed to remove everything smaller than a standard deviation of a calcification in size. Therefore we used a disk of the size of 1 mm for the opening. For the closing, we made use of a rectangular structuring element of the size of one standard deviation of a calcification in width (1 mm) and one standard deviation of a calcification in height (3 mm).

4.6. Results

The pure pixels classification as well as the pixel classification in correspondence with the three different post-processing methods, disk morphology (Fig. 2), biological morphology (Fig. 3) and DICPC (Fig. 4), were evaluated for the 18 calcification patches. The average values for the Jaccard index, sensitivity and specificity for the 18 patches are given in Tab. 1. Furthermore the statistical significance of difference between the means of the different methods was tested via a paired one-sided Student's t-test. The results of the tests are shown in Fig. 5, 6, 7.

In general we can observe that the Kriging produces results that are much closer to the original pixel classification than the morphological operations. The morphology imprints the shapes of its structuring elements onto the pixel classification result and produces harsh boundaries. The Kriging, on the contrary, makes the pixel classification boundaries finer and even prescinds structures out of the background around the calcifications.

Figure 2. The result of the disk morphology.

5. Discussion and Conclusion

When dealing with high resolution medical images that present long range interactions one runs into computational problems when trying to use standard contextual classifica-

	Pure Pixel Classification	Disk Morphology	Biological Morphology	Kriging Prior
Jaccard index	40%	35%	34%	41%
Sensitivity	0.40	0.41	0.43	0.62
Specificity	0.96	0.99	0.98	0.95

Table 1. The peak area overlap results for all the populations.

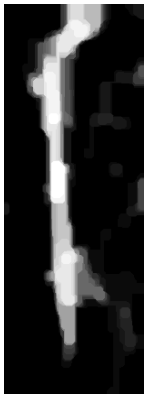

Figure 3. The result of the biological morphology.

Figure 4. The result of our method.

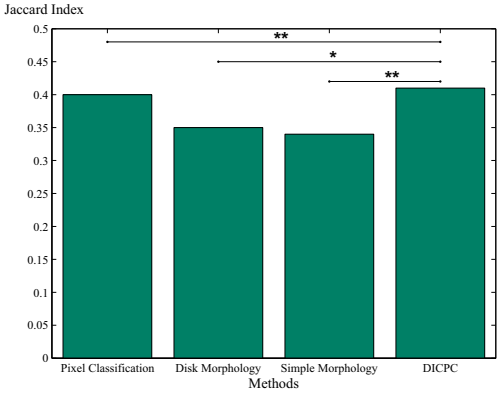

Figure 5. The graph displays the results for the paired student's t-tests between the different methods for the Jaccard index. The stars indicate the outcome of a paired one-tailed Students' t-test: $\star < 0.05$, $\star\star < 0.01$ and $\star\star\star < 0.001$.

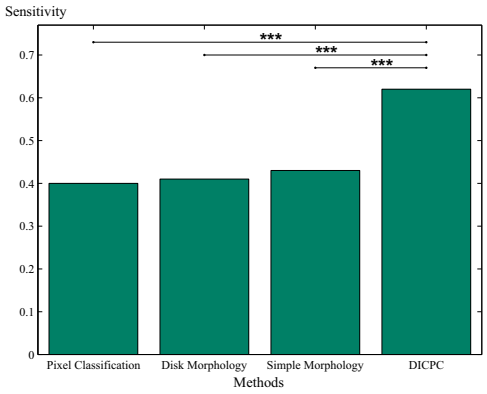

Figure 6. The graph displays the results for the paired student's t-tests between the different methods for the sensitivity. The stars indicate the outcome of a paired one-tailed Students' t-test: $\star < 0.05$, $\star\star < 0.01$ and $\star\star\star < 0.001$.

tion techniques like ICM or ICPC. This is why we compare our new method to other techniques, disk and biological morphology, which are common post-processing methods for this application.

We observe that our new method, DICPC, improves the sensitivity drastically and the Jaccard index slightly, while it leaves the specificity unchanged. The other post-processing methods, disk morphology and biological morphology, even lower the Jaccard index while improving the sensitivity and the specificity inconsiderably. The lowering of the Jaccard index is caused by the relatively harsh boundaries that the morphological operations produce in contrast to the Kriging.

Kriging and therefore DICPC is only the first step in the right direction. A weakness of DICPC is the linear model that underlies Kriging. It implies only pairwise interactions. The goal is to develop a contextual method that is as fast and computationally feasible as DICPC, but based on joint probabilities of the class labels.

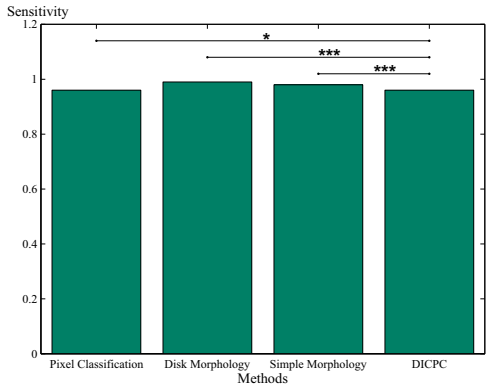

Figure 7. The graph displays the results for the paired student's t-tests between the different methods for the specificity. The stars indicate the outcome of a paired one-tailed Students' t-test: $\star < 0.05$, $\star\star < 0.01$ and $\star\star\star < 0.001$.

References

[1] AHA. American heart association. http://www.americanheart.org/, 2008. 4

[2] S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. Wu. An optimal algorithm for approximate nearest neighbor searching fixed dimensions. *J. ACM*, 45(6):891–923, 1998. 4

[3] Y. Bagger, L. Tanko, P. Alexandersen, H. Hansen, G. Qin, and C. Christiansen. The long-term predictive value of bone mineral density measurements for fracture risk is independent of the site of measurement and the age at diagnosis: results from the Prospective Epidemiological Risk Factors study. *Osteoporos Int*, 17(3):471–7, 2006. 4

[4] J. Berger. *Statistical decision theory: foundations, concepts, and methods*. Springer-Verlag, 1980. 2

[5] J. Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, 32(2):192–236, 1974. 2

[6] J. E. Besag. Discussion of a paper by p. switzer. *Bull. Int. Stat. Inst.*, 50(3):422–425, 1983. 1

[7] L. A. Conrad-Hansen, M. de Bruijne, F. Lauze, L. B. Tanko, and M. Nielsen. Quantizing calcification in the lumbar aorta on 2-d lateral x-ray images. *Lecture Notes in Computer Science, Computer Vision for Biomedical Image Applications*, 3765:409–418, 2005. 1

[8] T. Cootes, G. Edwards, and T. CJ. Active appearance models. *Lecture Notes in Computer Science, Computer Vision ECCV98*, 1407, 1998. 1

[9] EHN. European heart network. http://www.ehnheart.org/content/, 2008. 4

[10] S. Geman and D. Geman. *IEEE Trans. Pattern Anal. Mach. Intell.*, 6:721–741, 1984. 1

[11] E. Isaaks and R. Srivastava. *Applied Geostatistics*. Oxford University Press, 1989. 2

[12] P. Jaccard. tude comparative de la distribution florale dans une portion des alpes et des jura. *Bulletin del la Socit Vaudoise des Sciences Naturelles*, 37:547–579, 1901. 4

[13] A. Journel and C. Huijbregts. *Floating search methods in feature selection*. Academic Press, 1978. 2

[14] J. J. Koenderink and A. J. van Doorn. Surface shape and curvature scales. *Image Vision Comput.*, 10(8):557–565, 1992. 4

[15] D. G. Krige. A statistical approach to some basic mine valuation problems on the witwatersrand. *OR*, 4(1), 1953. 1

[16] L. Larsen, M. Ganz, E. Dam, and M. Nielsen. Growth pattern of atherosclerotic calcifications. 2008. 5

[17] M. Loog and B. van Ginneken. Supervised segmentation by iterated contextual pixel classification. *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, 2:925–928 vol.2, 2002. 1

[18] A. W. Michael Kass and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4), 1988. 1

[19] L. Mosca, L. Appel, and et al. Evidence-Based Guidelines for Cardiovascular Disease Prevention in Women. *Circulation*, 109(5):672–693, 2004. 4

[20] S. Osher and J. Sethian. Fronts propagating with curvature dependent speed: Algorithms based on hamilton-jacobi formulations. *Journal of Computational Physics*, 79:12–49, 1988. 1

[21] T. Pham. Image restoration by ordinary kriging with convexity. In *ICPR '00: Proceedings of the International Conference on Pattern Recognition*, page 3334, Washington, DC, USA, 2000. IEEE Computer Society. 3

[22] L. Tanko, Y. Bagger, G. Qin, P. Alexandersen, P. Larsen, and C. Christiansen. Enlarged Waist Combined With Elevated Triglycerides Is a Strong Predictor of Accelerated Atherogenesis and Related Cardiovascular Mortality in Postmenopausal Women. *Circulation*, 111(15):1883–1890, 2005. 4

[23] B. van Ginneken, M. Stegmann, and M. Loog. Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database. *Medical Image Analysis*, 10(1):19–40, 2006. 1

[24] P. W. F. Wilson, L. I. Kauppila, and et al. Abdominal Aortic Calcific Deposits Are an Important Predictor of Vascular Morbidity and Mortality. *Circulation*, 103(11):1529–1534, 2001. 4

[25] J. C. Witteman, F. J. Kok, and et al. Aortic calcification as a predictor of cardiovascular mortality. *Lancet*, 2:1120–1122, 1986. 4