

Representative Patch-based Active Appearance Models Generated from Small Training Populations

Matthias Wilms, Heinz Handels, and Jan Ehrhardt

Institute of Medical Informatics, University of Lübeck, Lübeck, Germany,
wilms@imi.uni-luebeck.de

Abstract. Active Appearance Models (AAMs) and Constrained Local Models (CLMs) are classical approaches for model-based image segmentation in medical image analysis. AAMs consist of global statistical models of shape and appearance, are known to be hard to fit, and often suffer from insufficient generalization capabilities in case of limited training data. CLMs model appearance only for local patches and relax or completely remove the global appearance constraint. They are, therefore, much easier to optimize but in certain cases they lack the robustness of AAMs. In this paper, we present a framework for patch-based active appearance modeling, which elegantly combines strengths of AAMs and CLMs. Our models provide global shape and appearance constraints and we make use of recent methodological advances from computer vision for efficient joint optimization of shape and appearance parameters during model fitting. Furthermore, the insufficient generalization abilities of those global models are tackled by incorporating and extending a recent approach for learning representative statistical shape models from small training populations. We evaluate our approach on publicly available chest radiographs and cardiac MRI data. The results show that the proposed framework leads to competitive results in terms of segmentation accuracy for challenging multi-object segmentation problems even when only few training samples are available.

1 Introduction

Active Appearance Models (AAMs) [2] and Constrained Local Models (CLMs)¹ [8] are widely used in medical image analysis for robust model-based segmentation (see [1, 4, 6] for examples). Both approaches rely on the classical point- and PCA-based statistical shape modeling framework [3] and represent the shape space observed in a training population by a mean shape and a variability model. In addition to this global shape prior, AAMs also learn a global statistical model of

The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-66182-7_18

¹ Throughout this paper, we use the CLM definition of Saraghi *et al.* [8].

the typical (gray value) appearance/texture of those objects. During the segmentation process, an AAM, therefore, not only provides global a-priori information about plausible shape instances but also about global appearance variations. These properties are favourable for robust segmentation, but lead to complicated exact or heuristic fitting algorithms [7, 2] and to the limited generalization abilities of AAMs in case of small training populations [12]. CLMs, therefore, nowadays typically replace the global appearance model with (independent) local detectors for each landmark (i.e. random forests in [6]) based on small image patches. During model fitting, the shape model is used to globally regularize the local detector responses to guarantee plausibility of the resulting shape. This leads to efficient optimization strategies but ambiguities cannot always be resolved, which negatively impacts the segmentation performance [8].

Methodological advances and the availability of large annotated training data sets have led to a recent successful revival of AAM-based face tracking methods in computer vision (CV; see [12] for an overview). The main goal of this work is to show that this recent success in CV can be transferred to medical problems when those approaches are adequately adapted and the lack of training data is accounted for. Therefore, the contributions of this paper are threefold: (1) We adapt the recent patch-based facial shape and appearance modeling approach from [11] and the Fast-SIC fitting algorithm from [12] to multi-object segmentation in medical image data. This approach elegantly combines strengths of AAMs (global shape and appearance models) and CLMs (efficient optimization, use of small patches). (2) We incorporate our recent approach for learning of representative statistical shape models from few training data [13] into this framework and extend it to patch-based appearance modeling. (3) We show that this novel combination of methods leads to competitive results on publicly available chest radiographs and cardiac MRI data and outperforms traditional AAM methods.

2 Methods

Although being independent of the image/data dimensionality (2D or 3D), we will describe our methods in a 2D scenario for ease of understanding. We start with a description of our patch-based AAM framework in Sec. 2.1, which is followed by an explanation of the method we adapted and extended for learning representative AAMs from small training populations in Sec. 2.2. See Fig. 1 for a graphical overview of the proposed framework.

2.1 Patch-based AAM framework: Definition and Optimization

In the following, we assume a set $\{\mathbf{s}_i\}_{i=1}^{N_S}$ of N_S 2D training shapes/contours \mathbf{s}_i and a set of $\{I_i\}_{i=1}^{N_S}$ corresponding images $I_i : \mathbb{R}^2 \rightarrow \mathbb{R}$ to be given. Each contour is defined by N_M points $\mathbf{s}_i = [x_{i,1}, y_{i,1}, \dots, x_{i,N_M}, y_{i,N_M}]^T \in \mathbb{R}^{2N_M}$ distributed across its surface. We also assume that the landmarks are in correspondence across the training samples and that shape differences due to similarity transformations have been removed from the data (contours and images).

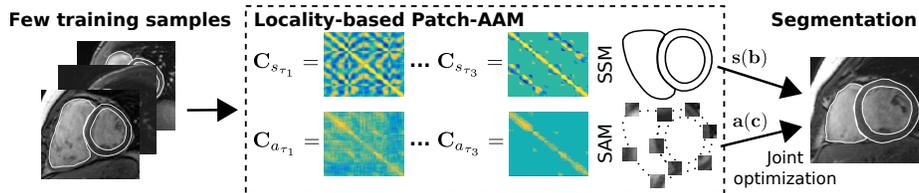


Fig. 1. Graphical overview of the proposed framework for patch-based active appearance modeling with few training samples. See text for details.

According to [12] an AAM consists of a statistical shape model (SSM), a statistical appearance model (SAM), and a suitable motion/transformation model used to map shape-free textures onto a new shape instance. For our patch-based AAM framework, which largely follows [11], we start by defining the SSM as a standard point distribution model [3]:

$$\mathbf{s}(\mathbf{b}) = \bar{\mathbf{s}} + \mathbf{P}\mathbf{b}, \quad \text{with } \bar{\mathbf{s}} = \frac{1}{N_S} \sum_{i=1}^{N_S} \mathbf{s}_i. \quad (1)$$

Here, $\mathbf{P} \in \mathbb{R}^{2N_M \times N_b}$ denotes an orthonormal matrix whose N_b columns are eigenvectors of the covariance matrix $\mathbf{C}_s = 1/N_S \sum_{i=1}^{N_S} (\mathbf{s}_i - \bar{\mathbf{s}})(\mathbf{s}_i - \bar{\mathbf{s}})^T$, and which compactly represents the subspace of plausible shapes. New shapes can be generated from Eq. (1) by varying parameter vector $\mathbf{b} \in \mathbb{R}^{N_b}$. As proposed in [7], \mathbf{P} also includes 4 orthonormal vectors to describe similarity transforms of the shapes generated (included in N_b).

To build a SAM, we first define that we are only interested in modeling the object appearance at small quadratic patches centered around each landmark. Furthermore, each patch defines a $N_e \times N_e$ regular grid with $N_p = N_e^2$ sampling locations. For model training, appearance information for each image I_i is obtained by simply sampling the image information at patches placed at the corresponding landmark locations. For statistical analysis, the appearance information of each image is concatenated to form an appearance vector $\mathbf{a}_i \in \mathbb{R}^{N_a}$, with $N_a = N_M N_p N_f$ where N_f denotes the number of features extracted (gray values, descriptors, ...) at each sampling location (e.g., for raw gray values $N_f = 1$). After applying an eigenvalue decomposition to the covariance matrix $\mathbf{C}_a = 1/N_S \sum_{i=1}^{N_S} (\mathbf{a}_i - \bar{\mathbf{a}})(\mathbf{a}_i - \bar{\mathbf{a}})^T$, we end up with a SAM similar to Eq. (1):

$$\mathbf{a}(\mathbf{c}) = \bar{\mathbf{a}} + \mathbf{Q}\mathbf{c}, \quad \text{with } \bar{\mathbf{a}} = \frac{1}{N_S} \sum_{i=1}^{N_S} \mathbf{a}_i, \quad \mathbf{Q} \in \mathbb{R}^{N_a \times N_c}, \quad \text{and } \mathbf{c} \in \mathbb{R}^{N_c}. \quad (2)$$

The SSM in Eq. (1) and the SAM in Eq. (2) define our patch-based AAM (Patch-AAM). Please note that for computational efficiency, we refrain from explicitly coupling both models as i.e. done in [2] and that our Patch-AAM implicitly defines a simple translational motion model instead of the complicated,

traditional piecewise-affine warp [7]. Instances of Eq. (1) and Eq. (2) are simply combined by translating the generated patches $\mathbf{a}(\mathbf{c})$ to the landmark locations given by $\mathbf{s}(\mathbf{b})$. Global scale changes and rotations can be handled by applying the associated similarity transform to the patches/image. Moreover, multiple objects can be easily handled by merging the landmarks/appearance information of all objects into one vector.

Fitting the Patch-AAM to an unseen image $I : \mathbb{R}^2 \rightarrow \mathbb{R}$ is now (with a slight abuse of notation and a patch-sampling/vectorization function $\phi(\cdot, \cdot)$) defined as a joint non-linear least-squares problem:

$$\arg \min_{\mathbf{b}, \mathbf{c}} \|\phi(I, \mathbf{s}(\mathbf{b})) - \mathbf{a}(\mathbf{c})\|^2. \quad (3)$$

Parameters \mathbf{b} and \mathbf{c} that optimally explain the image content in a least-squares sense are sought. Optimizing Eq. (3) is hard due to the non-linearity in \mathbf{b} . We follow the computationally efficient Gauss-Newton-like Fast-SIC optimization strategy presented in [12] to iteratively minimize Eq. (3). Fast-SIC was chosen due to its demonstrated ability to produce state-of-the-art results in CV [12].

After linearizing Eq. (3) with respect to the model and omission of second-order terms, we arrive at

$$\arg \min_{\Delta \mathbf{b}, \Delta \mathbf{c}} \|\phi(I, \mathbf{s}(\mathbf{b})) - \mathbf{a}(\mathbf{c}) - \mathbf{Q}\Delta \mathbf{c} - \mathbf{J}\Delta \mathbf{b}\|^2 \quad (4)$$

to compute updates $\Delta \mathbf{b}$ and $\Delta \mathbf{c}$, where $\mathbf{J} \in \mathbb{R}^{N_a \times N_b}$ is the Jacobian of \mathbf{Q} with respect to $\mathbf{s}(\mathbf{0})$ (see [11] for details). With $\mathbf{J}_F = \mathbf{J} - \mathbf{Q}\mathbf{Q}^T\mathbf{J}$, closed-form solutions $\Delta \mathbf{b} = (\mathbf{J}_F^T\mathbf{J}_F)^{-1}\mathbf{J}_F(\phi(I, \mathbf{s}(\mathbf{b})) - \bar{\mathbf{a}})$ and $\Delta \mathbf{c} = \mathbf{Q}^T(\phi(I, \mathbf{s}(\mathbf{b})) - \mathbf{a}(\mathbf{c}) - \mathbf{J}\Delta \mathbf{b})$ for both updates can be obtained in an alternating fashion. Because of the simple translational motion model, the shape and appearance parameters can finally be updated by $\mathbf{b} \leftarrow \mathbf{b} - \Delta \mathbf{b}$ and $\mathbf{c} \leftarrow \mathbf{c} + \Delta \mathbf{c}$.

2.2 Building representative Patch-AAMs from few training samples

The generalization capabilities and the segmentation performance of the Patch-AAMs presented in Sec. 2.1 will be mainly influenced by the quality and quantity of the training samples used to build the models. In medical image analysis, building AAMs often results in high-dimension-low-sample-size (HDLSS) problems, because the number of training samples is typically much smaller than the dimensionality of the data to be modeled (e.g., $N_S \ll 2N_M$ and $N_S \ll N_a$). In practice, this limits the dimension of the subspaces defined by \mathbf{P} and \mathbf{Q} to $N_S - 1$, the rank of $\mathbf{C}_s/\mathbf{C}_a$. It is unlikely that for very small sample sizes (e.g., $N_S = 10$), the space of plausible shapes/appearances is appropriately approximated by such low-dimensional subspaces. We, therefore, aim to extend the dimension of those subspaces to $N_b > N_S - 1$ and $N_c > N_S - 1$ in a plausible way to improve the generalization ability of the Patch-AAM in HDLSS scenarios.

Here, we utilize and extend the recent approach from [13] for building representative SSMs from few data, which is based on manipulations of the covariance

matrices and has several major advantages: It generates a single consistent shape model with global and local variability, seamlessly integrates with existing SSM-based frameworks, naturally handles multi-object scenarios, and was shown to outperform competing methods in [13].

In [13] it is assumed that in HDLSS scenarios some covariances (esp. those between distant points) in \mathbf{C}_s are overestimated. Therefore, a principle of locality (= interaction between distant landmarks is limited) is applied by defining a distance measure $d(s_i, s_j)$ on the set of landmarks and a cascade of thresholds $\tau_1 > \tau_l > \dots > \tau_{N_l}$. For each τ , a manipulated covariance matrix \mathbf{C}_{s_τ} is computed by enforcing the correlation between landmarks $d(s_i, s_j) > \tau$ to be 0. The eigenvectors of \mathbf{C}_{s_τ} define a subspace $\text{span}(\mathbf{P}_\tau) \subset \mathbb{R}^{2N_M}$ of dimension $N_{b_\tau} \geq N_S - 1$ because $\text{rank}(\mathbf{C}_{s_\tau}) \geq \text{rank}(\mathbf{C}_s)$, where the exact value of N_{b_τ} depends on τ . Now, starting with $\mathbf{P}_*^1 = \mathbf{P}_{\tau_1}$, the N_l subspaces are combined into a single multi-level shape model by successively searching for orthonormal bases \mathbf{P}_*^l of increasing locality in way that preserves global information [13]:

$$\mathbf{P}_*^l = \arg \min_{\mathbf{P}} d_G(\mathbf{P}, \mathbf{P}_{\tau_l}) \quad \text{s.t.} \quad \text{span}(\mathbf{P}_*^{l-1}) \subseteq \text{span}(\mathbf{P}) . \quad (5)$$

Here, $d_G(\cdot, \cdot)$ is a geodesic distance between subspaces and \mathbf{P}_*^l can be efficiently computed with an algorithm given in [13]. Finally, the orthonormal basis $\mathbf{P}_* = \mathbf{P}_*^{N_l}$, which represents global and local variability, can be plugged into Eq. (1).

In [13], this locality-based approach is only defined/used to enhance SSMs. We, however, strongly believe that in an AAM framework, the fitting algorithm can only make full use of the additional flexibility of \mathbf{P}_* when the SAM is enhanced in a similar way. We, therefore, propose to apply the same method to the covariance matrix \mathbf{C}_a of the SAM. Due to the patch-based definition of our AAM, we can simply use a distance d on the shape landmarks to manipulate \mathbf{C}_a , if we define that all points of a patch have the same distance to all points of another patch. Hence, sampling points of one patch are not separated by the manipulations. As for the SSM, the resulting orthonormal matrix \mathbf{Q}_* can be used to replace \mathbf{Q} in Sec. 2.1. This extension leads to a flexible Patch-AAM framework to build representative shape and appearance models from few training data.

3 Experiments and Results

The objectives of our evaluation are as follows: (1) Analysis of the segmentation performance of the proposed Patch-AAM approach on medical data. We focus on multi-object problems and HDLSS scenarios. (2) Analysis of the hypothesis that the fitting algorithm can only make full use of the flexibility of the locality-based SSM when the approach is also employed to build the SAM.

Data: Two publicly available 2D data bases are used for the experiments. (1) The JSRT/SCR data base [10, 4] that consists of 247 chest radiographs (2047×2047 pixels; 0.175 mm pixel spacing) and provides ground-truth segmentations for 5 structures (right/left lung, heart, right/left clavicle; represented by in total 166 corresponding landmarks; see [4]) for all cases. In [4], the data was divided

Table 1. Mean symmetric surface distances to the ground-truth segmentations obtained for the different patch-based AAM approaches on both data sets and for different numbers of training samples. See text for explanations. Results are given as mean \pm std. dev. in mm over all available test cases and repetitions. Italic font indicates a statistically significant difference to P-AAM. Bold font indicates a statistically significant difference to L-SSM. Significance is assessed by paired t-tests with $p < 0.05$.

# Train. samples	Gray values			SSC descriptor		
	P-AAM	L-SSM	L-AAM	P-AAM	L-SSM	L-AAM
<i>JSRT/SCR chest radiographs</i>						
15	4.34 \pm 1.98	3.93 \pm 2.03	3.66\pm2.03	3.62 \pm 1.40	2.84 \pm 1.37	2.83\pm1.35
30	3.28 \pm 1.20	3.09 \pm 1.21	2.89\pm1.73	2.77 \pm 1.08	2.41 \pm 1.03	2.37\pm1.02
40	3.06 \pm 1.14	2.91 \pm 1.16	2.77\pm1.12	2.54 \pm 0.94	2.28 \pm 0.89	2.26\pm0.89
70	2.73 \pm 0.99	2.63 \pm 0.98	2.54\pm0.97	2.22 \pm 0.78	2.12 \pm 0.75	2.11\pm0.75
all 123	2.65 \pm 1.10	2.48 \pm 1.02	2.44\pm0.95	2.05 \pm 0.71	2.02 \pm 0.73	2.00\pm0.73
<i>Cardiac MRI data</i>						
5	2.46 \pm 0.96	2.30 \pm 0.86	2.10\pm0.82	2.54 \pm 2.84	2.37 \pm 2.84	2.36 \pm 2.88
10	2.10 \pm 0.78	1.94 \pm 0.71	1.77\pm0.70	2.00 \pm 1.66	1.95 \pm 1.87	1.91\pm1.77
15	1.94 \pm 0.73	1.85 \pm 0.71	1.67\pm0.69	1.76 \pm 0.69	1.72 \pm 0.66	1.69\pm0.64
20	1.86 \pm 0.75	1.77 \pm 0.71	1.62\pm0.70	1.59 \pm 0.57	1.59 \pm 0.57	1.58 \pm 0.54

into two disjunct folds of 124 (*fold1*) and 123 cases (*fold2*), respectively. Here, *fold2* is employed to train the models, while *fold1* is used as test data. (2) 32 mid-ventricular slices (256×256 pixels; avg. pixel spacing: 1.40 mm) taken from different end-diastolic short axis cardiac MRI scans from [1]. For each case, ground-truth contours for the endo- and epicardium of the left ventricle (LV) are provided. We additionally segmented the right ventricle (RV). Each case consists of 104 landmarks (manually placed at corresponding locations). Random subsets of the data are employed for model training while the remaining cases are used for testing. Both data bases contain challenging cases due to i.e. fuzzy boundaries, the projective nature of the data (JSRT), and large anatomical variability.

Experimental design: We compare 3 different variants of the Patch-AAM approach: (1) Patch-AAMs directly learned on the training samples (P-AAM, see Sec. 2.1), (2) Patch-AAMs where the SSM is learned using the locality-based approach (L-SSM, see Sec. 2.2), and (3) Patch-AAMs where the SSM and the SAM are learned using the locality-based approach (L-AAM, see Sec. 2.2). For the locality-based variants, we use the multi-object distance defined in [13] and build 5 locality levels for the SSM and 3 for the SAM. Each variant is once learned on the raw gray values ($N_f=1$ and $N_e = 11$) and once by using the well-known SSC descriptor [5] with $N_f = 6$ and $N_e = 5$ to show the flexibility of the framework. To mimic HDLSS scenarios, models are generated for varying numbers of available training samples N_S randomly sampled from the training data (JSRT: 15, 30, 40, 70, all 123; Cardiac: 5, 10, 15, 20). Those models are then used to segment the objects in the test images. The experiments are repeated 10 (JSRT)/20 (Cardiac) times to reduce the bias introduced via random sampling. For P-AAM, the fitting algorithm is initialized with the mean shape and

a multi-resolution scheme with 3 levels is employed. The locality-based variants performed best with 2 multi-resolution levels. See our MATLAB code² for additional parameter settings. The segmentation accuracy is quantitatively assessed by computing mean symmetric contour distances to the ground-truth contours weighted by the number of landmarks of each object.

Results: Our results are summarized in Tab. 1 and exemplarily illustrated in Fig. 2. All 3 variants lead to competitive results for $N_S = 123$ (JSRT)/ $N_S = 20$ (Cardiac) when compared to the literature. For the JSRT data, all variants outperform the model-based approaches tested in [4, 13] (e.g., ASMs and AAMs; best in [4]: Hybrid ASM with 2.77 mm) and the SSC-based models achieve results comparable to [9] who obtain a mean distance of 2.10 mm on the same data. For the cardiac data, the AAM in [1] achieves a mean error of ≈ 1.5 mm for LV segmentation in 3D on the same data. We think our results obtained in 2D are at least comparable to theirs, given the fact that we also segment the RV.

Regarding the specific performances of the locality-based variants (L-SSM/L-AAM) for $N_S < 123$ (JSRT)/ $N_S < 20$ (Cardiac), we can see effects comparable to those observed in [13] for locality-based ASMs: In most cases, L-SSMs significantly outperform P-AAMs (see Tab. 1; paired t-tests with $p < 0.05$). The improvements tend to be larger for the JSRT data for which our results are also at least comparable to those achieved in [13] for the same sample sizes and data with a locality-based ASM. Most of ours are clearly better (e.g. $N_S = 40$; L-SSM w/ SSC: 2.28 mm; ASM in [13]: 2.82 mm). The results also confirm our initial hypothesis that the Patch-AAM fitting algorithm performs better with a locality-based SSM and SAM (L-AAM in Tab. 1). Nearly all results reported for L-AAM are significantly better (see Tab. 1; paired t-tests with $p < 0.05$) than those obtained by L-SSM. However, improvements when using the SSC descriptor are less prominent or even not existent. The exact reason for this behavior remains unclear and is subject to further research. Computationally, Patch-AAMs are efficient (0.5–6 s to process an image on a six-core Xeon CPU).

4 Conclusion

In this paper, a flexible framework for patch-based active appearance modeling that elegantly combines strengths of AAMs and CLMs is presented. Patch-AAMs consist of global shape and appearance models whose parameters are jointly optimized during the efficient segmentation procedure. The often insufficient generalization abilities of those global models are tackled by incorporating and extending a recent approach for learning representative SSMs from small training populations to patch-based appearance modeling. Our experiments on publicly available data show that our framework leads to competitive segmentation results for challenging multi-object problems even when only few training samples are available. Furthermore, the evaluation shows that our framework is able to make use of structural image representations like image descriptors

² www.imi.uni-luebeck.de/en/content/representative-patch-based-appearance-models

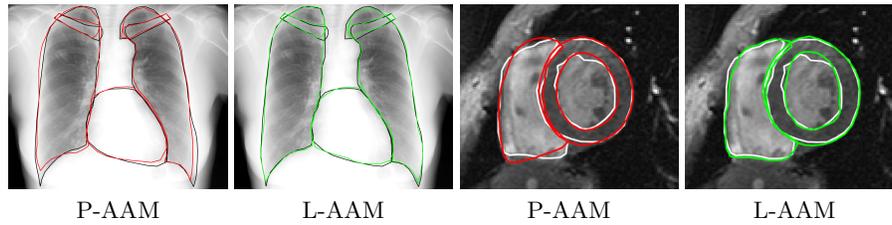


Fig. 2. Illustration of segmentation results (colored contours) for Patch-AAM variants on both data sets (left: JSRT; right: Cardiac). Black/White contours: ground-truth. Please note the improved coverage of local details by the L-AAM approach.

in addition to raw gray values. Although being only applied to 2D data in this work, the approach is not limited to 2D and readily generalizes to 3D cases.

Acknowledgments. This work is funded by the DFG (EH 224/6-1).

References

1. Andreopoulos, A., Tsotsos, J.K.: Efficient and generalizable statistical models of shape and appearance for analysis of cardiac MRI. *Med Image Anal* 12(3), 335 – 357 (2008)
2. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. *IEEE Trans PAMI* 23(6), 681–685 (2001)
3. Cootes, T., Taylor, C., Cooper, D., Graham, J.: Active shape models-their training and application. *CVIU* 61(1), 38 – 59 (1995)
4. van Ginneken, B., Stegmann, M.B., Loog, M.: Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database. *Med Image Anal* 10(1), 19 – 40 (2006)
5. Heinrich, M.P., Jenkinson, M., Papież, B.W., Brady, S.M., Schnabel, J.A.: Towards realtime multimodal fusion for image-guided interventions using self-similarities. In: Mori, K., Sakuma, I., Sato, Y. (eds.) *MICCAI 2013*. LNCS, vol. 8149, pp. 187–194. Springer (2013)
6. Lindner, C., Bromiley, P.A., Ionita, M.C., Cootes, T.F.: Robust and accurate shape model matching using random forest regression-voting. *IEEE Trans PAMI* 37(9), 1862–1874 (2015)
7. Matthews, I., Baker, S.: Active appearance models revisited. *IJCV* 60(2), 135–164 (2004)
8. Saragih, J.M., Lucey, S., Cohn, J.F.: Deformable model fitting by regularized landmark mean-shift. *IJCV* 91(2), 200–215 (2011)
9. Seghers, D., Loeckx, D., Maes, F., Vandermeulen, D., Suetens, P.: Minimal shape and intensity cost path segmentation. *IEEE Trans Medical Imaging* 26(8), 1115–1129 (Aug 2007)
10. Shiraiishi, J., Katsuragawa, S., Ikezoe, J., Matsumoto, T., Kobayashi, T., Komatsu, K.i., Matsui, M., Fujita, H., Kodera, Y., Doi, K.: Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists’ detection of pulmonary nodules. *American Journal of Roentgenology* 174(1), 71–74 (2000)

11. Tzimiropoulos, G., Pantic, M.: Gauss-newton deformable part models for face alignment in-the-wild. In: CVPR 2014. pp. 1851 – 1858 (2014)
12. Tzimiropoulos, G., Pantic, M.: Fast algorithms for fitting active appearance models to unconstrained images. IJCV 122(1), 17–33 (2017)
13. Wilms, M., Handels, H., Ehrhardt, J.: Multi-resolution multi-object statistical shape models based on the locality assumption. Med Image Anal 38, 17 – 29 (2017)