

## Research Article

# Automatic Spine Tissue Segmentation from MRI Data Based on Cascade of Boosted Classifiers and Active Appearance Model

Dominik Gawel <sup>1</sup>, Paweł Główka,<sup>2</sup> Tomasz Kotwicki <sup>2</sup>, and Michał Nowak<sup>1</sup>

<sup>1</sup>Chair of Virtual Engineering, Poznań University of Technology, 60-965 Poznań, Poland

<sup>2</sup>Department of Spine Disorders and Pediatric Orthopedics, University of Medical Sciences, 61-545 Poznań, Poland

Correspondence should be addressed to Dominik Gawel; dominik.gawel@gmail.com

Received 25 October 2017; Revised 14 March 2018; Accepted 19 March 2018; Published 29 April 2018

Academic Editor: Volker Rasche

Copyright © 2018 Dominik Gawel et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The study introduces a novel method for automatic segmentation of vertebral column tissue from MRI images. The paper describes a method that combines multiple stages of Machine Learning techniques to recognize and separate different tissues of the human spine. For the needs of this paper, 50 MRI examinations presenting lumbosacral spine of patients with low back pain were selected. After the initial filtration, automatic vertebrae recognition using Cascade Classifier takes place. Afterwards the main segmentation process using the patch based Active Appearance Model is performed. Obtained results are interpolated using centripetal Catmull–Rom splines. The method was tested on previously unseen vertebrae images segmented manually by 5 physicians. A test validating algorithm convergence per iteration was performed and the Intraclass Correlation Coefficient was calculated. Additionally, the 10-fold cross-validation analysis has been done. Presented method proved to be comparable to the physicians (FF =  $90.19 \pm 1.01\%$ ). Moreover results confirmed a proper algorithm convergence. Automatically segmented area correlated well with manual segmentation for single measurements ( $\bar{r} = 0.8336$ ) and for average measurements ( $\bar{r} = 0.9068$ ) with  $p = 0.05$ . The 10-fold cross-validation analysis (FF =  $91.37 \pm 1.13\%$ ) confirmed a good model generalization resulting in practical performance.

## 1. Introduction

Pathology of the intervertebral disk is one of the common causes of pain in the lumbar spine. In 40% of cases, pain of the lumbosacral spine is diagnosed as a discogenic [1]. What is more, 80% of the general population will have or already have had pain of the lumbosacral spine [2–4], in 5–10% of them a chronic pain develops [1, 5].

In contemporary diagnostics Magnetic Resonance Imaging (MRI) is the modality of choice for intervertebral disc visualization. Magnetic Resonance Imaging, for almost all spinal disorders, provides robust images of the spine [6] with high quality soft-tissue visualization, much more detailed than results obtained with other modalities [7]. The additional advantage of the MRI is the lack of radiation.

Automatic tissue segmentation from Magnetic Resonance Imaging data is a challenging task, because the quality

of the data affects the process; what is more, the differences between medical facilities, used protocols, and imaging machines force the necessity of universality.

Till now multiple approaches have been presented. Dong and Zheng in [8] divided the common solutions into methods that rely on graphical model [9], probabilistic model [10], watershed algorithm [11], atlas registration [12], graph cuts [13], Statistical Shape Model [14], anisotropic oriented flux [15], and random forest regression and classification [16].

The methods mentioned above are based on discrete classification returning a limited and inaccurate information about the tissue. The paper describes a method that combines multiple stages of Machine Learning (ML) [17] techniques to recognize and separate different tissues of the spine.

The objective of this study is to introduce a novel method for automatic segmentation of vertebral column tissue from MRI images.

## 2. Materials and Methods

**2.1. The Data.** For the needs of this paper 50 MRI examinations presenting lumbosacral (LS) spine of patients with low back pain were selected. The examinations were made with Siemens MAGNETOM Spectra 3T MR device. For vertebral body recognition T1 TSE (Turbo Spin Echo) Sagittal sequences, with Echo Time 9.3 ms and Repetition Time ranging from 550 ms to 700 ms, were chosen. The image sets consisted of 17 to 31 images with 4 mm slice thickness, 4.8 mm slice distance, and  $384 \times 384$  px resolution.

**2.2. General Procedure for Segmentation.** Presented solution is based on well-known Machine Learning (ML) [18] techniques combining Cascade of Boosted Classifiers [19–21] with patch based Active Appearance Model (AAM) [22, 23] algorithm and Principal Component Analysis (PCA) [24] (Figure 1).

At the beginning DICOM images are read. After that initial filtration is made to increase the quality of the data. Afterwards automatic vertebrae recognition using Cascade Classifier [21, 25] takes place. After the initial recognition the main tissue segmentation process is made using the patch based Active Appearance Model [22, 23]. Combined information about location, shape, and appearance provides a high quality model used for search and extraction of desired tissue. The results are afterwards interpolated using centripetal Catmull–Rom splines [26–28].

**2.3. Initial Filtration.** Due to low quality of the data (low resolution, intensity inhomogeneity, and high noise), initial filtration is needed. At the beginning the images are being resized to increase resolution. For the needs of presented method a high-resolution cubic spline has been chosen [29] (Figure 2).

After that the developed intensity inhomogeneity (IIH) correction method is performed. The method is based on recalculating local intensities in such a way to fit the global exponential function defined from the boundary fat-skin tissue intensity contrast. After calculations a nonlinear selective Gaussian Blur [30] using the same global exponential function for parameterization is performed to remove the noise amplified through the correction process. As a result of this method, an intensity inhomogeneity correction is achieved (Figure 3).

**2.4. Preliminary Vertebrae Recognition.** At the beginning, to achieve accurate segmentation results and reduce number of MRI examinations needed for training, vertebrae recognition is made. The goal of this action is to extract each vertebra from the whole image containing spine MRI examination. To achieve this the Machine Learning [18] training of Cascade of Boosted Classifiers [19–21] based on extended set Haar-like features [31] was made.

The vertebrae recognition consists of two major stages: training the classifier and vertebrae detection. Both were done using OpenCV library [25]. For the training two types of information are needed: positive examples presenting

desired object that one is looking for and negative examples presenting background. To prepare the data, special software allowing fast cutting, artificial data generation, and automatic background reconstruction was developed. For the training process 50 MRI examinations were used. From those examinations over 1000 vertebrae images were extracted manually and used for automatic creation of 10,000 artificial positive examples with Thin Plate Splines (TPS) transformations [32, 33]. Afterwards negative examples were reconstructed from the same examinations by covering previously cut out vertebrae using Image Inpainting method [34].

Both positive and negative examples are afterwards used for classifiers training based on the AdaBoost algorithm [35]. Multiple weak classifiers are then combined in a cascade resembling Decision Tree [36] creating a strong classifier [37]. To achieve best performance, after the recognition, additional size constraints were introduced, removing the false positive hits. Obtained model allows proper vertebrae recognition (Figure 4).

**2.5. Tissue Segmentation.** After the vertebrae recognition the main tissue segmentation is made. The solution is based on Active Appearance Model (AAM) [22, 23, 38] algorithm and combines a Statistical Shape Model based on Principal Component Analysis [39], with a gray-level Appearance Model. The method focuses on recognizing the predefined characteristic features from previously extracted vertebrae images by combining the information about each pretrained characteristic feature appearance with the information about features' mean position, their arrangement, and possible deviation. Similarly to preliminary vertebrae recognition, the tissue segmentation procedure consists of two stages: training and detection; however, contrary to previously trained classifiers, the built model is used for recognition of small patches instead of a whole vertebra. Each image used for the training originates from the prepared vertebrae database and was previously manually labeled by the group of five experts (physicians trained in MRI images assessment) with 16 characteristic points corresponding to vertebra features. Introduced information is used for building Point Distribution Model and creating training examples for the Appearance Model. The Point Distribution Model is used in a PCA [39] analysis to obtain the Shape Model containing information about the mean shape, eigenvectors, and eigenvalues. The positive and negative training examples are used for training to obtain the Appearance Model. Trained AAM model is afterwards used for spine tissue detection and classification. The detection procedure starts with an initial guess based on a perturbed ground truth shape. For this study a patch based AAM approach [22, 23, 38] has been chosen, representing the appearance of features as a rectangular patches distinguished around each landmark. Finally optimization of the cost function is solved by Lucas–Kanade Optimization [40, 41] method with Wiberg Inverse Compositional algorithm [42–44] (Figure 5).

**2.6. Shape Interpolation.** Automatically extracted 16 feature points for each vertebra image visible in the MRI examination

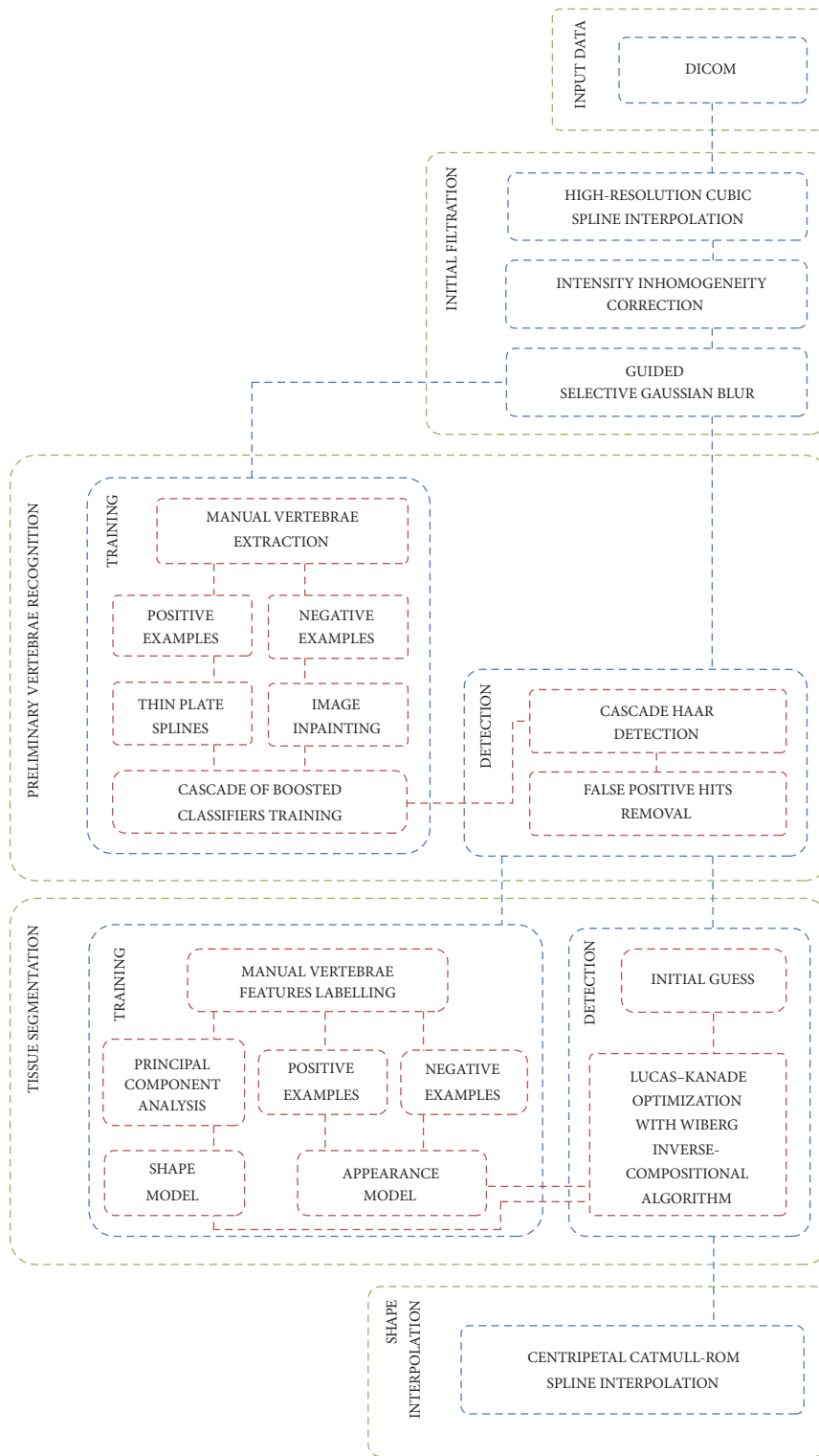


FIGURE 1: Flow chart of presented method. The solution is based on Machine Learning techniques combining Cascade of Boosted Classifiers with patch based Active Appearance Model and Principal Component Analysis.

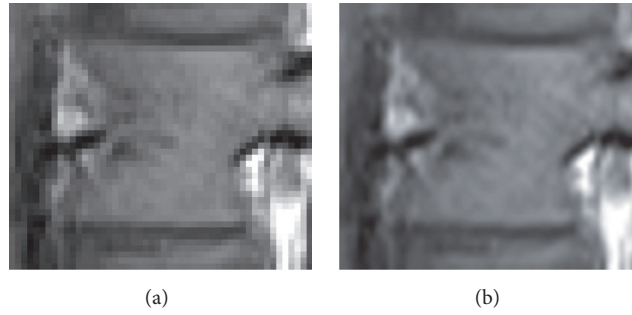


FIGURE 2: (a) Magnified image presenting sagittal slice of a single vertebra extracted from the input data. (b) The same image after initial resizing using a high-resolution cubic spline.

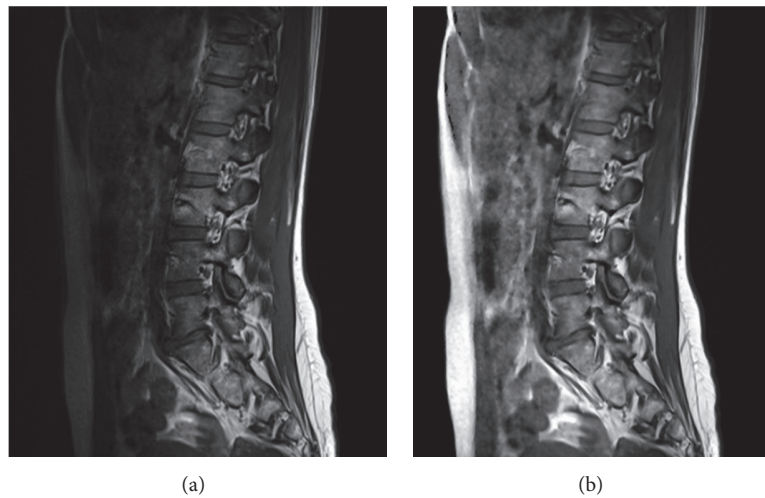


FIGURE 3: Initial filtration by recalculating local intensities to fit the global exponential function defined from boundary fat-skin tissue intensity contrast. (a) Image before IIH compensation. (b) Image after IIH compensation.

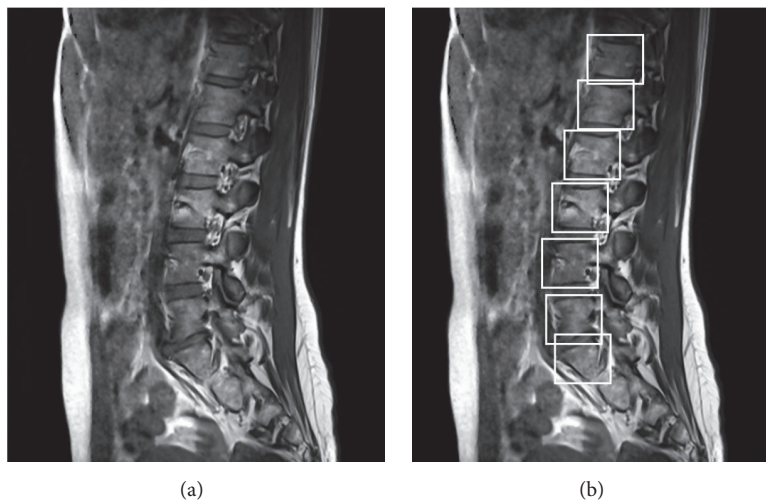


FIGURE 4: Vertebrae recognition using Cascade of Boosted Classifiers based on extended set of Haar-like features. Classifiers training based on the AdaBoost algorithm. (a) Initial image. (b) Positively detected vertebrae marked with bounding boxes for visualization.

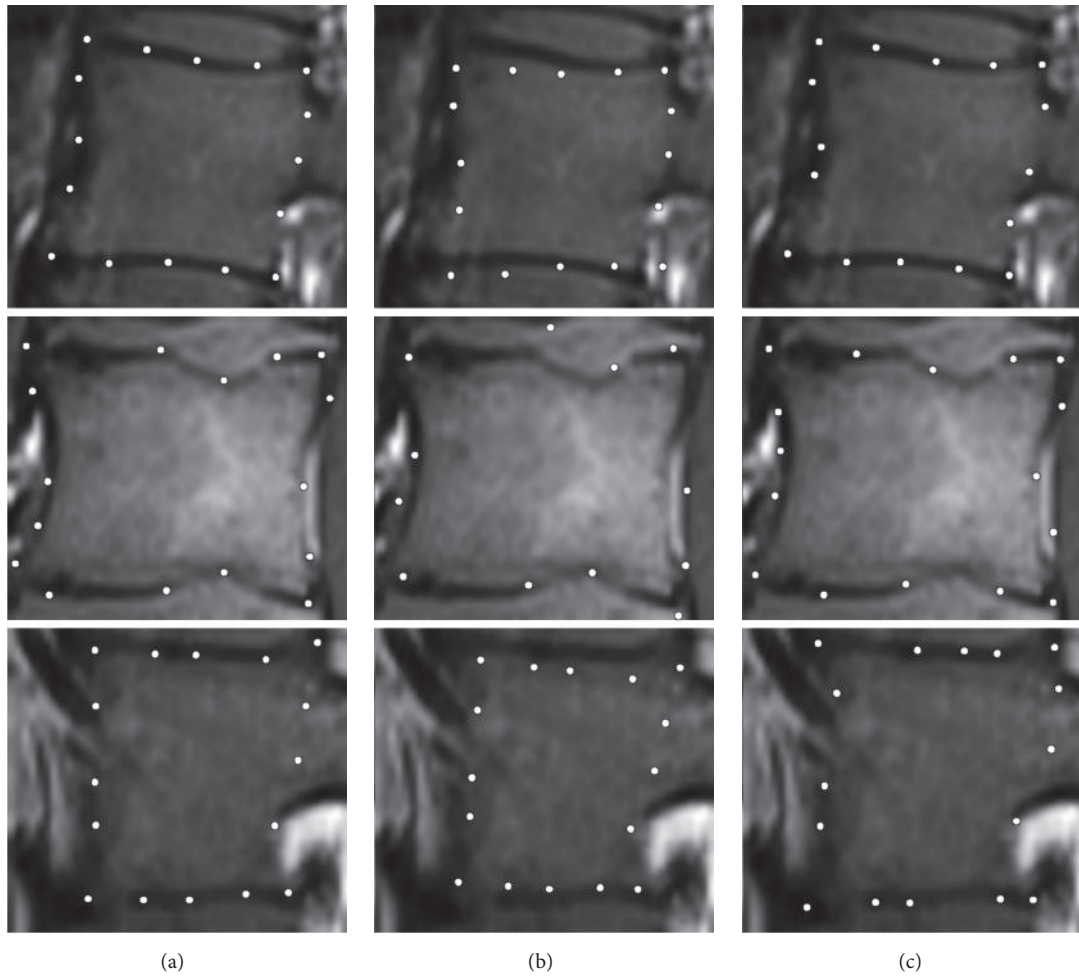


FIGURE 5: Example of landmark localization results for three different slices. (a–c) Ground truth shape, initial guess, and final result.

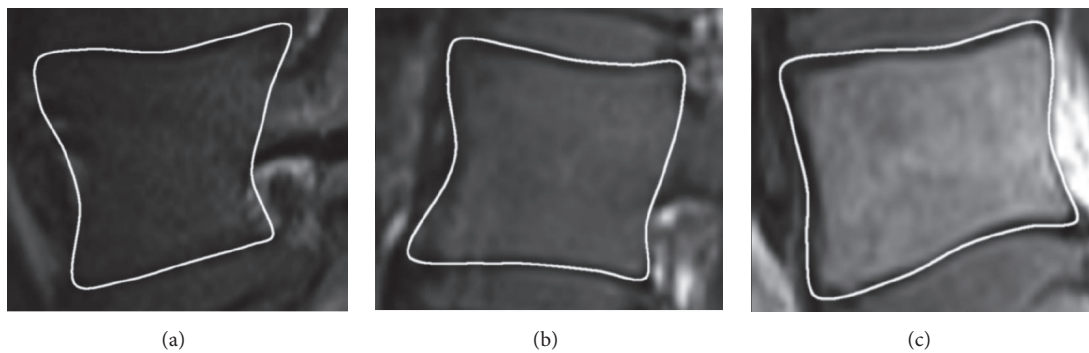


FIGURE 6: Information between the points interpolated with centripetal Catmull–Rom splines, ensuring C1 continuity, proper tightness with no self-intersections, and knot parameterization, leaving an area for further curve optimization. (a–c) Examples of a lateral, intermediate, and central part of the vertebral body.

are afterwards used for spine tissue segmentation. The information between the points is interpolated with centripetal Catmull–Rom Splines [26–28] (Figure 6), ensuring C1 continuity, proper tightness with no self-intersections, and knot parameterization, leaving an area for further curve optimization.

### 3. Results

The method was tested on a set of 50 previously unseen vertebrae images. The spine tissue was manually segmented by 5 physicians and compared with Machine Learning results. For the numerical evaluation three measures were used [45–47]:

TABLE 1: Comparison (percentage) of True Positive Fraction, False Negative Fraction, and False Fraction values and their standard deviations for automatically segmented data (significance level  $\alpha = 0.05$ ).  $\sigma_{\text{TPF}}$ : standard deviation for True Positive Fraction,  $\sigma_{\text{FNF}}$ : standard deviation for False Negative Fraction,  $\sigma_{\text{FF}}$ : standard deviation for False Fraction.

Algorithm	TPF	FNF	FF	$\sigma_{\text{TPF}}$	$\sigma_{\text{FNF}}$	$\sigma_{\text{FF}}$
WIC	92.46 $\pm$ 1.11	7.54 $\pm$ 1.11	90.04 $\pm$ 1.20	4.02	4.02	4.31
SIC	92.07 $\pm$ 1.23	7.93 $\pm$ 1.23	89.99 $\pm$ 1.30	4.45	4.45	4.69
POIC	89.51 $\pm$ 1.93	10.49 $\pm$ 1.93	85.89 $\pm$ 1.95	6.97	6.97	7.04
AIC	92.47 $\pm$ 1.10	7.53 $\pm$ 1.10	90.06 $\pm$ 1.21	3.95	3.95	4.37
MAIC	92.33 $\pm$ 1.10	7.67 $\pm$ 1.10	90.00 $\pm$ 1.20	3.97	3.97	4.33

True Positive Fraction (TPF) (1), False Negative Fraction (FNF) (2), and False Fraction (FF) (3):

$$\text{TPF (True Positive Fraction)} = \frac{A_{\text{TP}}}{A_T} \rightarrow \text{Sensitivity}, \quad (1)$$

where true positive area  $A_{\text{TP}} = A_S \cap A_T$ ,  $A_T$  is a manually segmented (by an expert) tissue area and  $A_S$  is an automatically segmented (by a computer) area.

$$\text{FNF (False Negative Fraction)} = \frac{A_{\text{FN}}}{A_T} \rightarrow \text{Specificity}, \quad (2)$$

where false negative area  $A_{\text{FN}} = A_T - A_S$ .

$$\text{FF (False Fraction)} = 1 - \frac{A_{\text{FP}} + A_{\text{FN}}}{A_T} \rightarrow \text{Accuracy}, \quad (3)$$

where false positive area  $A_{\text{FP}} = A_S - A_T$ .

To achieve better performance five different optimization algorithms, available in Menpo Framework [22, 38], were tested.

Only certain Inverse Compositional algorithms [38, 41–43, 48–52] were chosen to be tested: Wiberg Inverse Compositional (WIC) [38, 53–55] algorithm, Simultaneous Inverse Compositional (SIC) [49, 52] algorithm, Project-Out Inverse Compositional (POIC) [41] algorithm, Alternating Inverse Compositional (AIC) [42, 48] algorithm, and Modified Alternating Inverse Compositional (MAIC) [42, 48] algorithm. Three algorithms (WIC, AIC, and MAIC) achieved almost identical results (Table 1) (Figures 7, 8, and 9). Because of the best stability and lowest standard deviation (Table 1) Wiberg Inverse Compositional algorithm was chosen for further calculations.

What is more, to achieve reliable results, a mean value obtained from 100 procedure passes with 25 algorithm iterations each was computed and compared to results obtained manually by five experts (Table 2). The False Fraction is a general segmentation evaluation measure and is defined by the difference between manually segmented area and automatically segmented area, divided by the total area resulting from the manual segmentation. In this case the AAM algorithm (FF = 90.19  $\pm$  1.01%) proved to be almost identical to expert 5, has almost the lowest standard deviation ( $\sigma_{\text{FF}} = 3.64\%$ ), and is almost unnoticeably worse than other

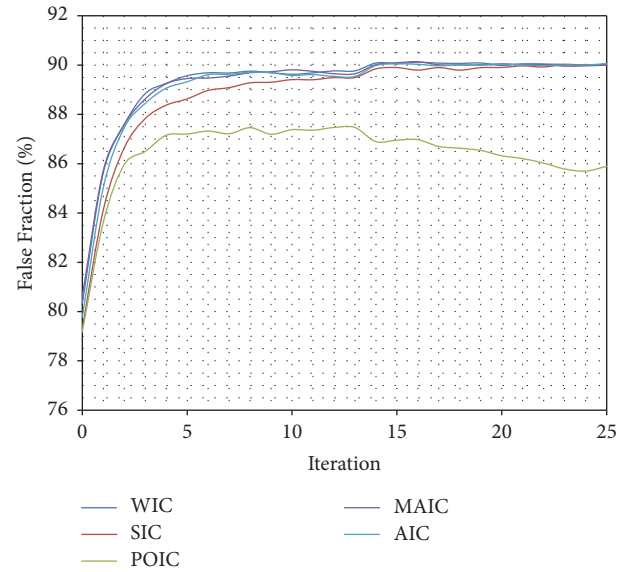


FIGURE 7: Comparison (percentage) of False Fraction mean values for subsequent interactions of automatic segmentation for different optimization algorithms.

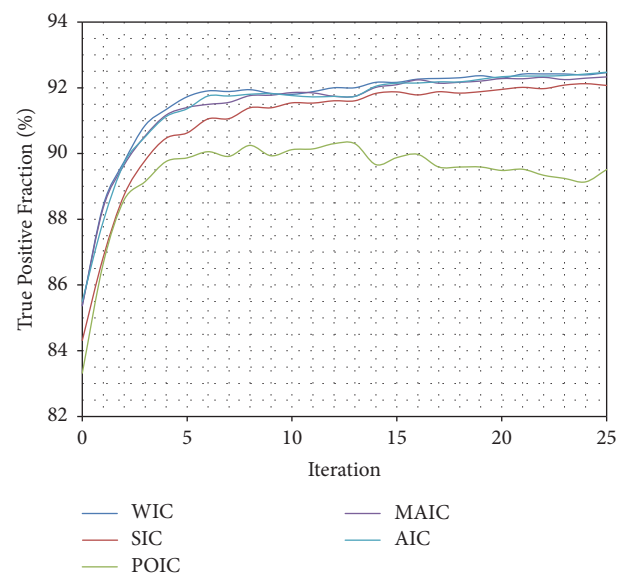


FIGURE 8: Comparison (percentage) of True Positive Fraction mean values for subsequent interactions of automatic segmentation for different optimization algorithms.

TABLE 2: Comparison (percentage) of True Positive Fraction, False Negative Fraction, and False Fraction for automatically segmented data using presented method and manually segmented data from each expert (significance level  $\alpha = 0.05$ ). To achieve reliable results a mean value obtained from 100 procedure passes with 25 algorithm iterations each is presented.  $\sigma_{\text{TPF}}$ : standard deviation for True Positive Fraction,  $\sigma_{\text{FNF}}$ : standard deviation for False Negative Fraction,  $\sigma_{\text{FF}}$ : standard deviation for False Fraction.

	TPF	FNF	FF	$\sigma_{\text{TPF}}$	$\sigma_{\text{FNF}}$	$\sigma_{\text{FF}}$
Computer	$92.28 \pm 0.95$	$7.72 \pm 0.95$	$90.19 \pm 1.01$	3.42	3.42	3.64
Expert 1	$96.17 \pm 0.92$	$3.83 \pm 0.92$	$91.67 \pm 1.27$	3.32	3.32	4.58
Expert 2	$95.13 \pm 0.83$	$4.87 \pm 0.83$	$92.11 \pm 1$	2.98	2.98	3.60
Expert 3	$97.74 \pm 0.43$	$2.26 \pm 0.43$	$91.09 \pm 1.33$	1.56	1.56	4.78
Expert 4	$97.56 \pm 0.37$	$2.44 \pm 0.37$	$91.49 \pm 1.47$	1.33	1.33	5.32
Expert 5	$92.46 \pm 1.33$	$7.54 \pm 1.33$	$90.22 \pm 1.31$	4.81	4.81	4.74

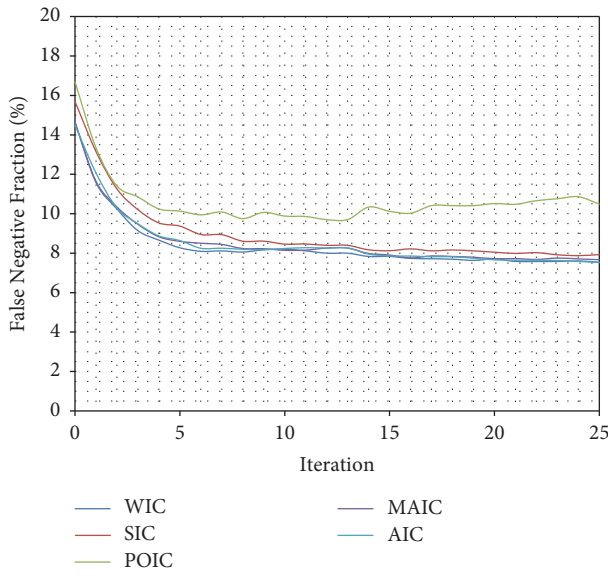


FIGURE 9: Comparison (percentage) of False Negative Fraction mean values for subsequent interactions of automatic segmentation for different optimization algorithms.

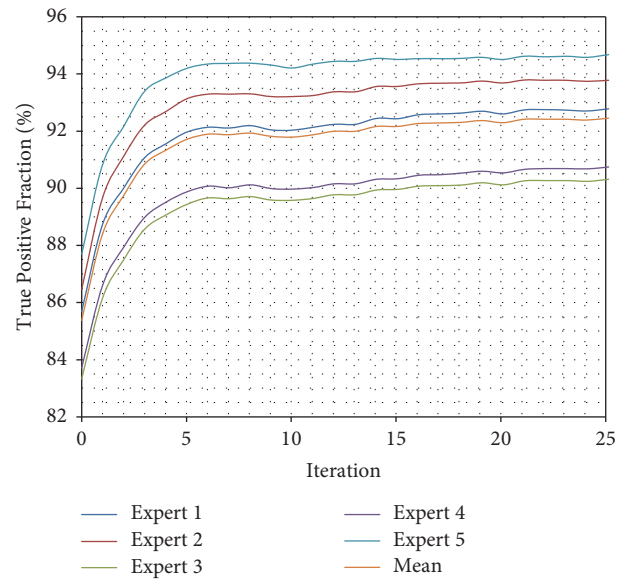


FIGURE 10: Comparison of True Positive Fraction for every iteration of automatic segmentation using presented method and manually segmented data from each expert. The True Positive Fraction provides information about percentage of properly segmented pixels. The TPF increases every iteration confirming proper functioning of presented algorithm—its convergence to manually segmented data.

experts. False Negative Fraction provides information about percentage of nonselected pixels classified by the investigators as a spine tissue and is an amount of manually segmented area not indicated by the automatic segmentation, divided by the total area resulting from the manual segmentation. The AAM has the highest False Negative Fraction ( $\text{FNF} = 7.72 \pm 0.95\%$ ,  $\sigma_{\text{FNF}} = 3.42\%$ ) of all investigators. The True Positive Fraction provides information about percentage of properly segmented pixels and is an amount of automatically segmented area consistent with manual segmentation, divided by the total area resulting from the manual segmentation. The AAM method has the True Positive Fraction value of  $\text{TPF} = 92.28 \pm 0.95\%$  and standard deviation of  $\sigma_{\text{TPF}} = 3.42\%$ .

What is more, a test validating algorithm convergence by comparing automatic segmentation per iteration results with manual segmentation results was performed. A single algorithm pass with 25 iterations for a set of 50 previously unseen vertebrae images was executed and a change of TPF, FNF, and FF values for each iteration was calculated (Table 3). The True Positive Fraction increases every iteration (Figure 10), while the False Negative Fraction decreases

(Figure 11) simultaneously leading the False Fraction to increase per iteration (Figure 12), confirming proper functioning of presented algorithm—its convergence to manually segmented data.

The Intraclass Correlation Coefficient (ICC) was calculated to evaluate the consistency of the vertebral bodies area determined by the experts and the computer (Tables 4 and 5). The high ICC results for single measurements ( $\bar{r} = 0.8336$ ) and for average measurements ( $\bar{r} = 0.9068$ ) with the  $p = 0.05$  confirmed that the automatically and manually obtained segmentation results are comparable.

Additionally, the 10-fold cross-validation [56–58] analysis has been done. The database of 1000 training images was divided into equal parts and tested iteratively 10 times by training the model from 90% of images and performing a test on the remaining 10% with ground truth annotations. The ground truth landmarks have been used for shape

TABLE 3: Comparison (percentage) of True Positive Fraction, False Negative Fraction, and False Fraction for automatically segmented data using presented method and manually segmented data from each expert.

Iteration	Expert 1			Expert 2			Expert 3			Expert 4			Expert 5			Mean		
	TPF	FNF	FF	TPF	FNF	FF	TPF	FNF	FF	TPF	FNF	FF	TPF	FNF	FF	TPF	FNF	FF
0	85.69	14.31	81.50	86.46	13.54	81.13	83.33	16.67	79.05	83.72	16.28	79.23	87.70	12.30	80.81	85.38	14.62	80.34
1	88.76	11.24	86.81	89.69	10.31	86.76	86.18	13.82	84.03	86.62	13.38	84.31	90.86	9.14	86.28	88.42	11.58	85.64
2	90.02	9.98	88.58	91.13	8.87	88.91	87.51	12.49	86.00	87.95	12.05	86.27	92.18	7.82	88.15	89.76	10.24	87.58
3	91.08	8.92	89.74	92.22	7.78	90.13	88.58	11.42	87.23	89.00	11.00	87.45	93.42	6.58	89.65	90.86	9.14	88.84
4	91.56	8.44	90.14	92.68	7.32	90.50	89.06	10.94	87.68	89.52	10.48	87.97	93.87	6.13	90.00	91.34	8.66	89.26
5	91.97	8.03	90.51	93.13	6.87	90.93	89.43	10.57	87.97	89.89	10.11	88.25	94.21	5.79	90.20	91.73	8.27	89.57
6	92.14	7.86	90.61	93.29	6.71	91.00	89.66	10.34	88.18	90.08	9.92	88.41	94.36	5.64	90.24	91.91	8.09	89.69
7	92.11	7.89	90.58	93.28	6.72	91.02	89.63	10.37	88.15	90.03	9.97	88.34	94.38	5.62	90.31	91.89	8.11	89.68
8	92.19	7.81	90.70	93.30	6.70	91.02	89.70	10.30	88.25	90.13	9.87	88.51	94.39	5.61	90.28	91.94	8.06	89.75
9	92.04	7.96	90.57	93.20	6.80	90.99	89.58	10.42	88.16	90.01	9.99	88.42	94.32	5.68	90.32	91.83	8.17	89.69
10	92.03	7.97	90.54	93.20	6.80	90.99	89.56	10.44	88.14	89.99	10.01	88.37	94.22	5.78	90.10	91.80	8.20	89.63
11	92.14	7.86	90.64	93.24	6.76	90.95	89.63	10.37	88.17	90.05	9.95	88.38	94.36	5.64	90.27	91.88	8.12	89.68
12	92.24	7.76	90.71	93.37	6.63	91.06	89.77	10.23	88.28	90.17	9.83	88.49	94.45	5.55	90.30	92.00	8.00	89.77
13	92.24	7.76	90.71	93.37	6.63	91.06	89.77	10.23	88.28	90.17	9.83	88.49	94.45	5.55	90.30	92.00	8.00	89.77
14	92.45	7.55	91.11	93.56	6.44	91.42	89.93	10.07	88.61	90.33	9.67	88.79	94.56	5.44	90.49	92.17	7.83	90.08
15	92.44	7.56	91.07	93.56	6.44	91.40	89.95	10.05	88.63	90.35	9.65	88.80	94.52	5.48	90.40	92.17	7.83	90.06
16	92.58	7.42	91.21	93.65	6.35	91.45	90.06	9.94	88.71	90.47	9.53	88.90	94.55	5.45	90.31	92.26	7.74	90.11
17	92.61	7.39	91.19	93.67	6.33	91.41	90.09	9.91	88.69	90.49	9.51	88.87	94.56	5.44	90.23	92.28	7.72	90.08
18	92.64	7.36	91.19	93.68	6.32	91.37	90.10	9.90	88.66	90.54	9.46	88.92	94.56	5.44	90.19	92.31	7.69	90.07
19	92.70	7.30	91.21	93.75	6.25	91.39	90.18	9.82	88.72	90.60	9.40	88.94	94.60	5.40	90.16	92.37	7.63	90.09
20	92.61	7.39	91.07	93.68	6.32	91.30	90.11	9.89	88.62	90.55	9.45	88.87	94.53	5.47	90.05	92.30	7.70	89.98
21	92.75	7.25	91.17	93.78	6.22	91.33	90.25	9.75	88.73	90.67	9.33	88.94	94.65	5.35	90.10	92.42	7.58	90.05
22	92.75	7.25	91.17	93.78	6.22	91.30	90.26	9.74	88.72	90.69	9.31	88.97	94.63	5.37	90.05	92.42	7.58	90.04
23	92.74	7.26	91.12	93.78	6.22	91.28	90.25	9.75	88.71	90.69	9.31	88.96	94.64	5.36	90.07	92.42	7.58	90.03
24	92.71	7.29	91.09	93.74	6.26	91.24	90.23	9.77	88.70	90.68	9.32	88.96	94.61	5.39	90.03	92.39	7.61	90.00
25	92.78	7.22	91.14	93.78	6.22	91.21	90.30	9.70	88.74	90.74	9.26	88.98	94.70	5.30	90.11	92.46	7.54	90.04



TABLE 4: The Intraclass Correlation Coefficient (ICC) of the vertebral bodies area determined by experts and computer, for single measurements with the  $p = 0.05$ .

	Computer	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5
Computer		0.8815	0.9153	0.7368	0.7210	0.9134
Expert 1	0.8815		0.9442	0.8761	0.8622	0.8685
Expert 2	0.9153	0.9442		0.8721	0.8695	0.9266
Expert 3	0.7368	0.8761	0.8721		0.9586	0.7830
Expert 4	0.7210	0.8622	0.8695	0.9586		0.7670
Expert 5	0.9134	0.8685	0.9266	0.7830	0.7670	

TABLE 5: The Intraclass Correlation Coefficient (ICC) of the vertebral bodies area determined by experts and computer, for average measurements with the  $p = 0.05$ .

	Computer	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5
Computer		0.9370	0.9558	0.8485	0.8379	0.9547
Expert 1	0.9370		0.9713	0.9339	0.9260	0.9296
Expert 2	0.9558	0.9713		0.9317	0.9302	0.9619
Expert 3	0.8485	0.9339	0.9317		0.9789	0.8783
Expert 4	0.8379	0.9260	0.9302	0.9789		0.8681
Expert 5	0.9547	0.9296	0.9619	0.8783	0.8681	

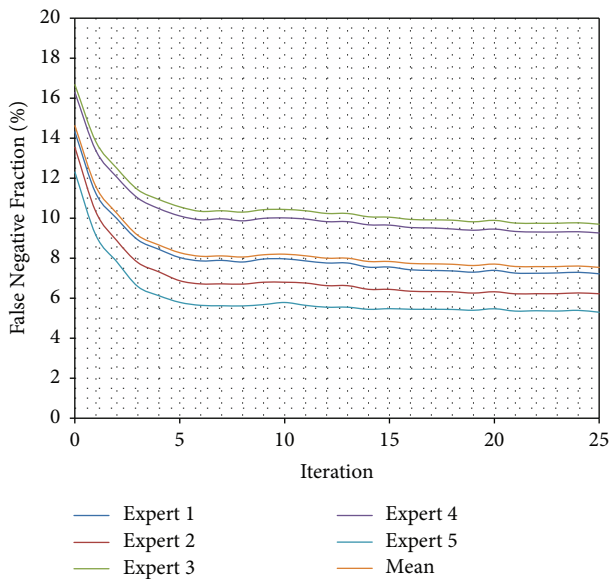


FIGURE 11: Comparison of False Negative Fraction for every iteration of automatic segmentation using presented method and manually segmented data from each expert. The FNF provides information about percentage of nonselected pixels classified by the investigators as a spine tissue. The FNF decreases every iteration confirming proper functioning of presented algorithm—its convergence to manually segmented data.

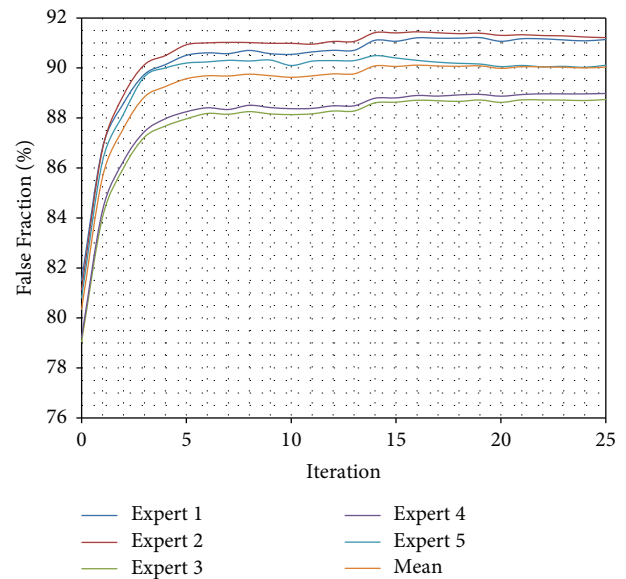


FIGURE 12: Comparison of False Fraction for every iteration of automatic segmentation using presented method and manually segmented data from each expert. The FF is a general segmentation evaluation measure. The FF increases every iteration confirming proper functioning of presented algorithm—its convergence to manually segmented data.

interpolation using Catmull–Rom splines and compared with automatic segmentation results using TPF, FNE, and FF measures (Table 6). The False Fraction mean value of  $91.37 \pm 1.13\%$  with standard deviation  $\sigma_{FF} = 5.76\%$  confirmed a good model generalization to an independent dataset and resulting practical performance.

#### 4. Discussion

Low resolution of presented data, high noise, and non-homogeneous information about the tissues enforced the increasing of the quality of input data by initial filtration. From multiple interpolation methods [29, 59] widely used for image resampling, a high-resolution cubic spline has been

TABLE 6: 10-fold cross-validation comparison (percentage) of True Positive Fraction (TPF), False Negative Fraction (FNF), and False Fraction (FF) for automatically segmented data using presented method and ground truth annotations (significance level  $\alpha = 0.05$ ).  $\sigma_{\text{TPF}}$ : standard deviation for True Positive Fraction,  $\sigma_{\text{FNF}}$ : standard deviation for False Negative Fraction,  $\sigma_{\text{FF}}$ : standard deviation for False Fraction.

$k$ -fold	TPF	FNF	FF	$\sigma_{\text{TPF}}$	$\sigma_{\text{FNF}}$	$\sigma_{\text{FF}}$
1	$95.72 \pm 1.27$	$4.28 \pm 1.27$	$90.12 \pm 1.38$	6.50	6.50	7.02
2	$97.39 \pm 0.77$	$2.61 \pm 0.77$	$91.29 \pm 1.07$	3.91	3.91	5.47
3	$95.73 \pm 0.86$	$4.27 \pm 0.86$	$91.18 \pm 1.72$	4.41	4.41	8.75
4	$95.13 \pm 0.85$	$4.87 \pm 0.85$	$91.98 \pm 0.98$	4.35	4.35	5.00
5	$95.00 \pm 0.85$	$5.00 \pm 0.85$	$91.32 \pm 1.01$	4.32	4.32	5.14
6	$96.18 \pm 0.76$	$3.82 \pm 0.76$	$90.87 \pm 0.94$	3.88	3.88	4.80
7	$96.35 \pm 0.82$	$3.65 \pm 0.82$	$92.13 \pm 1.06$	4.17	4.17	5.42
8	$96.14 \pm 0.69$	$3.86 \pm 0.69$	$92.20 \pm 0.95$	3.50	3.50	4.82
9	$96.03 \pm 0.71$	$3.97 \pm 0.71$	$92.09 \pm 0.98$	3.64	3.64	5.02
10	$93.56 \pm 1.10$	$6.44 \pm 1.10$	$90.53 \pm 1.21$	5.60	5.60	6.17
Mean	$95.72 \pm 0.87$	$4.28 \pm 0.87$	$91.37 \pm 1.13$	4.43	4.43	5.76

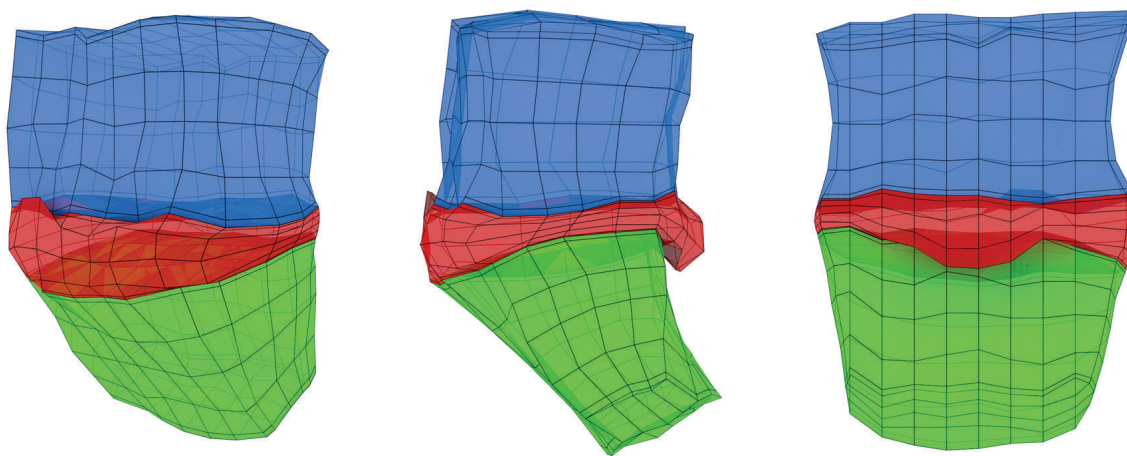


FIGURE 13: Discrete STL 3D model created manually from detected feature points (landmarks). The pathology of vertebrae and intervertebral disc is clearly visible.

chosen to increase the resolution of the input data, because of its good high-frequency response and high-frequency enhancement. What is more, a novel method of intensity inhomogeneity correction useful for sagittal MRI spine images has been presented. In last years, multiple methods used for intensity inhomogeneity correction emerged [60, 61]; however, they were mostly used for and tested with brain MRI scans. Because of a different application, segmentation of bone tissue instead of brain tissue, an additional method of initial filtration was developed.

For MRI images, a robust method for spine segmentation was prepared. The procedure combined well-known and widely tested Machine Learning methods [18]: Cascade of Boosted Classifiers [19–21] based on extended set Haar-like features [31] for preliminary vertebrae detection, with patch based Active Appearance Model [22, 23, 38] and Principal Component Analysis [39] for precise tissue segmentation. Usage of feature localization method and interpolation of the resulting information with centripetal Catmull–Rom splines [26, 27] omitted the problem of low quality. Due to the nature of Catmull–Rom splines [28] further optimization can

be done to achieve better interpolation results. The paper [62] presents multiple recent methods for intervertebral disc segmentation, which can be treated as a similar task, including Machine Learning and deep learning based approaches. The segmentation results presented in the paper [62] were measured with dice overlap coefficients and varied from 81.6% to 92% for different methods. Comparing those results with obtained segmentation and generalization results of 90.19% and 91.37%, one can conclude that presented AAM approach provides a good segmentation performance and moreover can be applied for intervertebral discs localization and segmentation.

In the future, automatically defined landmark localizations could be used for automatic creation of a discrete (Figure 13) and continuous (Figure 14) 3D spine models, which can be easily used in Finite Element Analysis [63–65], contrary to standard voxel representation.

Obtained three-dimensional model (Figures 13 and 14) contains information about the size and shape of the intervertebral disk and the adjacent vertebral bodies. Based on it, one can determine morphology of the intervertebral

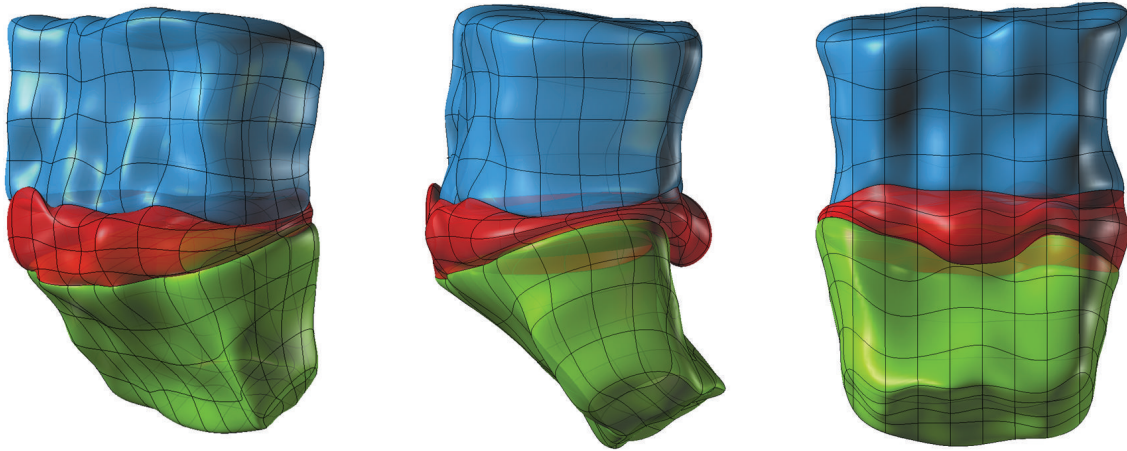


FIGURE 14: Continuous NURBS model created manually from detected feature points (landmarks), easily convertible to Finite Element mesh. The pathology of vertebrae and intervertebral disc is clearly visible.

disk including direction, dimensions, and volume of the herniation of intervertebral disc, giving the clinicians a tool for better understanding of the pathology.

### Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

### Acknowledgments

This work was supported by the National Centre for Research and Development under Grant no. PBS3/B9/34/2015.

### References

- [1] P. Finch, "Technology insight: imaging of low back pain," *Nature Clinical Practice Rheumatology*, vol. 2, no. 10, pp. 554–561, 2006.
- [2] O. Airaksinen, J. I. Brox, C. Cedraschi et al., "Chapter 4: european guidelines for the management of chronic nonspecific low back pain," *European Spine Journal*, vol. 15, no. 2, pp. S192–S300, 2006.
- [3] A. J. Schoenfeld and B. K. Weiner, "Treatment of lumbar disc herniation: evidence-based practice," *Journal of General Internal Medicine*, vol. 3, pp. 209–214, 2010.
- [4] H. Yang, H. Liu, L. Zemin et al., "Low back pain associated with lumbar disc herniation: role of moderately degenerative disc and annulus fibrous tears," *International Journal of Clinical and Experimental Medicine*, vol. 8, no. 2, p. 1634, 2015.
- [5] R. C. Lawrence, C. G. Helmick, F. C. Arnett et al., "Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States," *Arthritis & Rheumatology*, vol. 41, no. 5, pp. 778–799, 1998.
- [6] S. S. Eun, H.-Y. Lee, S.-H. Lee, K. H. Kim, and W. C. Liu, "MRI versus CT for the diagnosis of lumbar spinal stenosis," *Journal of Neuroradiology*, vol. 39, no. 2, pp. 104–109, 2012.
- [7] L. Ros, J. Mota, A. Guedea, and D. Bidgood, "Quantitative measurements of the spinal cord and canal by MR imaging and myelography," *European Radiology*, vol. 8, no. 6, pp. 966–970, 1998.
- [8] X. Dong and G. Zheng, "Automated 3D lumbar intervertebral disc segmentation from MRI data sets," in *Computational Radiology for Orthopaedic Interventions*, pp. 25–40, Springer, 2016.
- [9] S. Schmidt, J. Kappes, M. Bergtholdt et al., "Spine detection and labeling using a parts-based graphical model," in *Information Processing in Medical Imaging*, vol. 20, pp. 122–133, Springer, 2007.
- [10] J. J. Corso, R. S. Alomari, and V. Chaudhary, "Lumbar disc localization and labeling with a probabilistic model on both pixel and object features," in *Medical Image Computing and Computer-Assisted Intervention-MICCAI*, pp. 202–210, Springer, 2008.
- [11] C. Chevrefils, F. Chretien, C.-É. Aubin, and G. Grimard, "Texture analysis for automatic segmentation of intervertebral disks of scoliotic spines from MR images," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 4, pp. 608–620, 2009.
- [12] S. K. Michopoulou, L. Costaridou, E. Panagiotopoulos, R. Speller, G. Panayiotakis, and A. Todd-Pokropek, "Atlas-based segmentation of degenerated lumbar intervertebral discs from MR images of the spine," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 9, pp. 2225–2231, 2009.
- [13] B. A. Ismail, K. Punithakumar, G. Gregory, W. Romano, and L. Shuo, "Graph cuts with invariant object-interaction priors: application to intervertebral disc segmentation," in *Proceedings of the Biennial International Conference on Information Processing in Medical Imaging*, pp. 221–232, Springer, 2011.
- [14] A. Neubert, J. Fripp, C. Engstrom et al., "Automated detection, 3D segmentation and analysis of high resolution spine MR images using statistical shape models," *Physics in Medicine and Biology*, vol. 57, no. 24, pp. 8357–8376, 2012.
- [15] M. W. K. Law, K. Tay, A. Leung, G. J. Garvin, and S. Li, "Intervertebral disc segmentation in MR images using anisotropic oriented flux," *Medical Image Analysis*, vol. 17, no. 1, pp. 43–61, 2013.
- [16] B. Glocker, D. Zikic, E. Konukoglu, D. R. Haynor, and A. Criminisi, "Vertebrae localization in pathological spine CT via dense classification from sparse annotations," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, vol. 16, pp. 262–270, Springer, 2013.

- [17] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [18] S. Theodoridis, *Machine Learning: A Bayesian and Optimization Perspective*, Academic Press, 2015.
- [19] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, vol. 1, Springer, Berlin, Germany, 2001.
- [20] J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: a statistical view of boosting," *The Annals of Statistics*, vol. 28, no. 2, pp. 337–407, 2000.
- [21] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. I511–I518, December 2001.
- [22] J. Alabort-I-Medina, E. Antonakos, J. Booth, P. Snape, and S. Zafeiriou, "Menpo: a comprehensive platform for parametric image alignment and visual deformable models," in *Proceedings of the 2014 ACM Conference on Multimedia (MM '14)*, pp. 679–682, ACM, New York, NY, USA, November 2014.
- [23] G. J. Edwards, C. J. Taylor, and T. F. Cootes, "Interpreting face images using active appearance models," in *Proceedings of the 3rd IEEE International Conference on Automatic Face and Gesture Recognition (FG '98)*, pp. 300–305, April 1998.
- [24] I. Jolliffe, *Principal Component Analysis*, Wiley Online Library, 2002.
- [25] G. Bradski, "The opencv library," *Doctor Dobbs Journal*, vol. 25, no. 11, pp. 120–126, 2000.
- [26] P. J. Barry and R. N. Goldman, "A recursive evaluation algorithm for a class of Catmull-Rom splines," *Computer Graphics*, vol. 22, no. 4, pp. 199–204, 1988.
- [27] E. Catmull and R. Raphael, "A class of local interpolating splines," *Computer Aided Geometric Design*, vol. 74, pp. 317–326, 1974.
- [28] C. Yuksel, S. Schaefer, and J. Keyser, "Parameterization and applications of Catmull-Rom curves," *Computer-Aided Design*, vol. 43, no. 7, pp. 747–755, 2011.
- [29] J. A. Parker, R. V. Kenyon, and D. E. Troxel, "Comparison of Interpolating Methods for Image Resampling," *IEEE Transactions on Medical Imaging*, vol. 2, no. 1, pp. 31–39, 1983.
- [30] F. Catté, P.-L. Lions, J.-M. Morel, and T. Coll, "Image selective smoothing and edge detection by nonlinear diffusion," *SIAM Journal on Numerical Analysis*, vol. 29, no. 1, pp. 182–193, 1992.
- [31] R. Lienhart and J. Maydt, "An extended set of Haar-like features for rapid object detection," in *Proceedings of the International Conference on Image Processing (ICIP '02)*, vol. 1, September 2002.
- [32] F. L. Bookstein, "Principal warps: thin-plate splines and the decomposition of deformations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 6, pp. 567–585, 1989.
- [33] K. Rohr, H. S. Stiehl, R. Sprengel, T. M. Buzug, J. Weese, and M. H. Kuhn, "Landmark-based elastic registration using approximating thin-plate splines," *IEEE Transactions on Medical Imaging*, vol. 20, no. 6, pp. 526–534, 2001.
- [34] A. Telea, "An image inpainting technique based on the fast marching method," *Journal of Graphics Tools*, vol. 9, no. 1, pp. 23–34, 2004.
- [35] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," in *Proceedings of the European Conference on Computational Learning Theory*, pp. 23–37, Springer, 1995.
- [36] Y. Amit, D. Geman, and K. Wilder, "Joint induction of shape features and tree classifiers," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 11, pp. 1300–1305, 1997.
- [37] P. Viola and M. J. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.
- [38] J. Alabort-i-Medina and S. Zafeiriou, "A unified framework for compositional fitting of active appearance models," *International Journal of Computer Vision*, vol. 121, no. 1, pp. 26–64, 2017.
- [39] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 4, pp. 433–459, 2010.
- [40] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," 1981.
- [41] J. Matthews and S. Baker, "Active appearance models revisited," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 135–164, 2004.
- [42] G. Papandreou and P. Maragos, "Adaptive and constrained algorithms for inverse compositional active appearance model fitting," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08)*, pp. 1–8, IEEE, June 2008.
- [43] G. Tzimiropoulos and M. Pantic, "Optimization problems for fast AAM fitting in-the-wild," in *Proceedings of the 14th IEEE International Conference on Computer Vision (ICCV '13)*, pp. 593–600, IEEE, Sydney, Australia, December 2013.
- [44] G. Tzimiropoulos and M. Pantic, "Gauss-Newton deformable part models for face alignment in-the-wild," in *Proceedings of the 27th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '14)*, pp. 1851–1858, June 2014.
- [45] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.
- [46] A. Fenster and B. Chiu, "Evaluation of Segmentation algorithms for Medical Imaging," in *Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, pp. 7186–7189, Shanghai, China, January 2006.
- [47] C. E. Metz, "Special articles roc methodology in radiologic imaging," *Investigative Radiology*, vol. 21, no. 9, pp. 720–733, 1986.
- [48] J. Alabort-I-Medina and S. Zafeiriou, "Bayesian active appearance models," in *Proceedings of the 27th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '14)*, pp. 3438–3445, June 2014.
- [49] S. Baker, R. Gross, and I. Matthews, "Lucas-kanade 20 years on: a unifying framework," 2003.
- [50] S. Baker and I. Matthews, "Equivalence and efficiency of image alignment algorithms," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '01)*, vol. 1, December 2001.
- [51] S. Baker and I. Matthews, "Lucas-Kanade 20 years on: a unifying framework," *International Journal of Computer Vision*, vol. 56, no. 3, pp. 221–255, 2004.
- [52] R. Gross, I. Matthews, and S. Baker, "Generic vs. person specific active appearance models," *Image and Vision Computing*, vol. 23, no. 12, pp. 1080–1093, 2005.
- [53] T. Okatani and K. Deguchi, "On the wiberg algorithm for matrix factorization in the presence of missing components," *International Journal of Computer Vision*, vol. 72, no. 3, pp. 329–337, 2007.

- [54] D. Strelow, "General and nested wiberg minimization," in *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR '12)*, pp. 1584–1591, IEEE, June 2012.
- [55] T. Wiberg, "Computation of principal components when data are missing," in *Proceedings of the Second Symp. Computational Statistics*, pp. 229–236, 1976.
- [56] P. A. Devijver and J. Kittler, *Pattern Recognition: A Statistical Approach*, vol. 761, Prentice Hall, London, UK, 1982.
- [57] S. Geisser, *Predictive Inference: An Introduction*, Prentice Hall, New York, NY, USA, 1993.
- [58] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI '95)*, vol. 14, pp. 1137–1145, Stanford, Calif, USA, 1995.
- [59] T. M. Lehmann, C. Gönner, and K. Spitzer, "Survey: interpolation methods in medical image processing," *IEEE Transactions on Medical Imaging*, vol. 18, no. 11, pp. 1049–1075, 1999.
- [60] Z. Hou, "A review on MR image intensity inhomogeneity correction," *International Journal of Biomedical Imaging*, vol. 2006, Article ID 49515, 11 pages, 2006.
- [61] U. Vovk, F. Pernuš, and B. Likar, "A review of methods for correction of intensity inhomogeneity in MRI," *IEEE Transactions on Medical Imaging*, vol. 26, no. 3, pp. 405–421, 2007.
- [62] G. Zheng, C. Chu, D. L. Belavý et al., "Evaluation and comparison of 3D intervertebral disc localization and segmentation methods for 3D T2 MR data: A grand challenge," *Medical Image Analysis*, vol. 35, pp. 327–344, 2017.
- [63] T. J. R. Hughes, *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*, Courier Corporation, 2012.
- [64] B. A. Szabo and I. Babuška, *Finite Element Analysis*, John Wiley & Sons, 1991.
- [65] O. C. Zienkiewicz, R. L. Taylor, and R. L. Taylor, *The Finite Element Method*, vol. 3, McGraw-hill, London, UK, 1977.